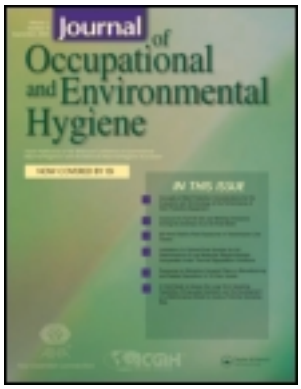

Journal of Occupational and Environmental Hygiene
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/uoeh20

Comparison of Means of Two Lognormal Distributions Based on Samples with Multiple Detection Limits

K. Krishnamoorthy[a], Thomas Mathew[b] & Zhao Xu[a]
[a] Department of Mathematics, University of Louisiana at Lafayette, Lafayette, Louisiana
[b] Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, Maryland
Accepted author version posted online: 27 Jan 2014.Published online: 27 Jan 2014.

To cite this article: K. Krishnamoorthy, Thomas Mathew & Zhao Xu (2014) Comparison of Means of Two Lognormal Distributions Based on Samples with Multiple Detection Limits, Journal of Occupational and Environmental Hygiene, 11:8, 538-546, DOI: 10.1080/15459624.2014.881487

To link to this article: http://dx.doi.org/10.1080/15459624.2014.881487

Journal of Occupational and Environmental Hygiene, 11: 538–546
ISSN: 1545-9624 print / 1545-9632 online

DOI: 10.1080/15459624.2014.881487

Comparison of Means of Two Lognormal Distributions Based on Samples with Multiple Detection Limits

K. Krishnamoorthy,[1] Thomas Mathew,[2] and Zhao Xu[1]

[1]Department of Mathematics, University of Louisiana at Lafayette, Lafayette, Louisiana
[2]Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, Maryland

The problem of comparing the means of two lognormal distributions based on samples with multiple detection limits is considered. Tests and confidence intervals for the ratio of the two means, based on pivotal quantities involving the maximum likelihood estimators, are proposed. The merits of the proposed approaches are evaluated by Monte Carlo simulation. Simulation study indicates that the procedures are satisfactory in terms of coverage probabilities of confidence intervals, and powers of tests. The proposed approach can also be applied to find confidence intervals for the difference between the means of the two lognormal distributions. Illustrative examples with a real data set and with a simulated data set are given.

Keywords confidence interval, coverage probability, difference of means, maximum likelihood estimator, power, ratio 'of means, Type I error

Address correspondence to: Thomas Mathew, Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, MD 21250; e-mail: mathew@umbc.edu

INTRODUCTION

Censored data resulting from the presence of one or more detection limits are quite common in industrial hygiene. Difficulties associated with the analysis of such data have been well documented in the literature. In particular, the commonly used substitution method (consisting of replacing the nondetects with a fraction of the detection limit) has now been recognized as an inappropriate strategy for data analysis, especially when the proportion of nondetects is not small. In fact, conclusions regarding exposures can be suspect if data analysis is performed based on the substitution method. While alternative approaches (for example, likelihood-based methods) can be used for estimating various parameters of interest (such as the arithmetic mean, geometric mean, percentiles, and so on), the computation of confidence limits for these parameters is a problem that has not been satisfactorily addressed, perhaps due to the unavailability of methodology that is satisfactory for data sets with small sample sizes and various proportions of sample values below the limits of detection. The problem is expected to be more complex when multiple detection limits are present.

The detection limit problem has been (and continues to be) an active topic of research in the industrial hygiene literature. The issues related to the analysis of such data have recently been highlighted by Helsel,[1,2] and earlier by this same author.[3] In the presence of a single detection limit, large sample procedures can be developed using likelihood approaches; see the report by Frome and Wambach,[4] and the articles by Ganser and Hewett,[5] Hewett,[6] and Hewett and Ganser[7] for a summary of the available procedures, especially in the context of lognormal exposure data.

However, *satisfactory* small sample procedures are currently unavailable. Also, it turns out that the usual likelihood-based methods can result in test procedures whose type I error rates can be significantly higher than or lower than a 5% nominal level; see Frome and Wambach,[4] Table I. In two recent articles, Krishnamoorthy, Mallick, and Mathew[8,9] have investigated an imputation approach and some large sample procedures to address the detection limit scenario. The authors have noted that the imputation approach can be calibrated to achieve a high degree of accuracy (in terms of providing coverage probability close to the assumed 95% confidence level) for the purpose of computing confidence intervals, tolerance intervals, prediction intervals, and so on, for the normal, lognormal, and gamma distributions. However, the above articles address only the situation of a single detection limit. For the multiple detection limit scenario, Krishnamoorthy and Xu[10] have recently developed a methodology for analyzing a lognormally distributed exposure sample. The authors exhibited pivotal quantities based on the maximum likelihood estimators, and derived confidence limits for the arithmetic mean and percentiles of the exposure distribution. Numerical results in their article have demonstrated *satisfactory* results even for small sample sizes, with a large proportion of nondetects.

TABLE I. Description and Notations for Samples with Multiple Detection Limits

	Sample 1	Sample 2
Devices or Laboratories	$(1, 1), \ldots, (1, k_1)$	$(2, 1), \ldots, (2, k_2)$
# items measured using *(i,j)* device	$n_{11}, \ldots, n_{1k_1}$	$n_{21}, \ldots, n_{2k_2}$
Detection limits	$DL_{11}, \ldots, DL_{1k_1}$	$DL_{21}, \ldots, DL_{2k_2}$
# nondetects	$m_{11}, \ldots, m_{1k_1}$	$m_{21}, \ldots, m_{2k_2}$
Sample sizes	$n_1 = \sum_{j=1}^{k_1} n_{1j}$	$n_2 = \sum_{j=1}^{k_2} n_{2j}$
Total # nondetects	$m_1 = \sum_{j=1}^{k_1} m_{1j}$	$m_2 = \sum_{j=1}^{k_2} m_{2j}$
Geometric means	$\exp(\mu_1)$ for Population 1	$\exp(\mu_2)$ for Population 2
Geometric SDs	$\exp(\sigma_1)$ for Population 1	$\exp(\sigma_2)$ for Population 2
MLEs	$(\hat{\mu}_1, \hat{\sigma}_1)$	$(\hat{\mu}_2, \hat{\sigma}_2)$
Arithmetic means	$\exp(\eta_1)$ for Population 1, where $\eta_1 = \mu_1 + .5\sigma_1^2$	$\exp(\eta_2)$ for Population 2, where $\eta_2 = \mu_2 + .5\sigma_2^2$

In this article, the multiple detection limit framework developed in Krishnamoorthy and Xu[10] is extended for the comparison of two lognormally distributed exposure populations. The problem of comparing two exposure populations can come up, for example, in the context of evaluating the effectiveness of engineering controls or interventions for reducing exposures; the two populations consist of exposure levels before and after the installation of engineering controls or interventions. In particular, an appropriate hypothesis testing problem can be formulated for verifying if the mean exposure level was improved by an intervention.

The problem can also come up if we want to compare two methods in terms of their efficacy in reducing exposures; in fact, one of our examples is of this type. For computing confidence limits for the ratio or difference of the arithmetic means, we have obtained pivotal quantities based on the MLEs. We have used the concept of a *generalized pivotal quantity* (GPQ), as done in Krishnamoorthy and Xu,[10] and the concept is briefly explained in the next section. Following this, the confidence interval calculation is described and numerical results are reported to show the utility of the proposed algorithm. Hypothesis testing is then taken up for comparing two exposure populations, and we have also reported numerical results on the type I error probability and power of the proposed test. Two examples are provided to illustrate the proposed methodology. The first example is based on synthetic data sets, and the second example involves real industrial data sets. Our overall conclusion is that the proposed methodology for comparing the arithmetic means of two lognormally distributed exposure populations is *satisfactory* even for small samples.

We conclude our introductory remarks by noting that in practice, exact lognormality may be unlikely to hold for occupational exposure data. However, the lognormal distribution has been found to adequately describe the distribution of exposure data in many occupational settings. Some of the earlier work on the detection limit problem (Frome and Wambach,[4] Ganser and Hewett,[5] Hewett,[6] and Hewett and Ganser,[7] for example) do deal with lognormally distributed exposure data. In the present work also, the methodology developed is limited to the case of lognormally distributed exposure samples.

METHODS

The method of maximum likelihood estimators and their sampling distributions, coupled with the ideas of generalized confidence intervals and generalized p-values, are used to obtain the results in this article. We now provide the necessary background information on all of these.

Maximum Likelihood Estimation

Consider exposure measurements whose distribution can be satisfactorily modeled using the lognormal distribution, so that the log-transformed measurements follow a normal distribution. We shall denote this normal distribution by $N(\mu, \sigma^2)$, where μ and σ^2 are, respectively, the unknown mean and unknown variance of the log-transformed data. Consider a simple random sample of n observations subject to k detection limits, say, $DL_1, \ldots, DL_k$, from the normal distribution $N(\mu, \sigma^2)$. With respect to the exposure population, the DL_is are thus the log-transformed detection limits associated with the log-transformed data. We assume that n_i measurements were obtained by the ith laboratory procedure or device with detection limit DL_i, and $n = \sum_{i=1}^{k} n_i$. Suppose m_i nondetects are below DL_i, with $m = \sum_{i=1}^{k} m_i$. Thus the number of detected observations is n-m, and let the corresponding log-transformed observations be denoted by $x_1, \ldots, x_{n-m}$.

In the presence of k detection limits as mentioned above, with m_i observations below the detection limit DL_i, the likelihood function is given by

$$\prod_{i=1}^{k}\left[\Phi\left(\frac{DL_i - \mu}{\sigma}\right)\right]^{m_i}\prod_{i=1}^{n-m}\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{\frac{-(x_i-\mu)^2}{2\sigma^2}\right\}, \quad (1)$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. When there is only one detection limit (i.e., $k = 1$), the above likelihood function is given, for example, in Cohen.[11] The first part of the likelihood in Eq. (1) is a product of k terms since there are k detection limits; see also Eqs. (2.3)–2.6) in Helsel.[2] Note that the above expression for the likelihood assumes independence of the observations. The log-likelihood function, after omitting a constant term, can be written as

$$l(\mu, \sigma) = \sum_{i=1}^{k} m_i \ln \Phi\left(\frac{DL_i - \mu}{\sigma}\right) - (n-m)\ln\sigma - \sum_{i=1}^{n-m} \frac{(x_i - \mu)^2}{2\sigma^2}. \quad \textbf{(2)}$$

Now let $\hat{\mu}$ and $\hat{\sigma}$ denote the maximum likelihood estimates of μ and σ, respectively, obtained by maximizing the above log-likelihood function. These can be numerically obtained using the bivariate Newton-Raphson iterative method; for computational details we refer to the Appendix in Krishnamoorthy and Xu.[10]

In their work, Krishnamoorthy and Xu[10] have obtained the following approximate distributional result:

$$\left(\frac{\hat{\mu} - \mu}{\sigma}, \frac{\hat{\sigma}}{\sigma}\right) \textit{ is distributed as } \left(\hat{\mu}^*, \hat{\sigma}^*\right), \quad \textbf{(3)}$$

where $\hat{\mu}^*$ and $\hat{\sigma}^*$ are the MLEs based on a sample of size n from a standard normal distribution with detection limits $DL_i^* = (DL_i - \hat{\mu})/\hat{\sigma}$, $i = 1,\ldots,k$. That is, $\hat{\mu}^*$ and $\hat{\sigma}^*$ are the values of μ and σ, respectively, that maximize the log-likelihood function in Eq. (2) with $x_1, \ldots, x_{n-m}$ being a sample from a standard normal distribution with detection limits $DL_1^*, \ldots, DL_k^*$. Note that Eq. (3) implies that

$$\frac{\hat{\mu} - \mu}{\hat{\sigma}} = \frac{(\hat{\mu} - \mu)/\sigma}{\hat{\sigma}/\sigma} \text{ is distributed as } \frac{\hat{\mu}^*}{\hat{\sigma}^*}, \text{ approximately.} \quad \textbf{(4)}$$

Using the above approximate distributional results, Krishnamoorthy and Xu[10] have developed confidence intervals for the mean and the percentiles of a lognormal exposure distribution. They also used the generalized confidence interval idea. We shall also use the concept of a generalized confidence interval, as well as the concept of a generalized p-value. These are briefly explained below.

The Generalized Confidence Interval and Generalized *P*-Value

The generalized confidence interval idea was introduced by Weerahandi,[12] and is useful to obtain confidence intervals in many problems where conventional methods are difficult to apply. The computation of a generalized confidence interval requires a generalized pivotal quantity (GPQ). Here we shall give a general definition of the GPQ. For this, let X be a random variable whose distribution depends on a scalar parameter of interest θ and a nuisance parameterδ, where δcould be a vector. Let x denote the observed value of X. Here X and x could be vectors. A generalized pivotal quantity (GPQ) for θ is a function of X, x, θ, and δ, denoted by G(X, x; θ, δ), and is required to satisfy the following conditions:

(i) given the observed value x, the distribution of G(X, x; θ, δ) is free of the nuisance parameter δ,
(ii) the observed value of G(X, x; θ, δ), namely G(x, x; θ, δ), is free of the nuisance parameter δ.

Here we would like to point out that it is the distribution of G(X, x; θ, δ) and its observed value that are required to be free of δ, even though G(X, x; θ, δ) itself could depend on the nuisance parameter δ. This aspect will become clearer from the GPQ given later in Eq. (6).

Under the above two conditions, the percentiles of G(X, x; θ, δ) can be computed, and can be used to obtain confidence limits for θ. The resulting confidence interval is referred to as a generalized confidence interval. We would like to emphasize that the GPQ G(X, x; θ, δ) involves the random variable X as well as the observed data x. While computing the percentiles of the GPQ, x is held fixed, and the GPQ is to be treated as a random variable since X is random.

The generalized p-value concept was introduced by Tsui and Weerahandi[13] for the purpose of hypothesis testing. In the above set-up, consider the problem of testing

$$H_0 : \theta \geq \theta_0 \; vs. \; H_1 : \theta < \theta_0,$$

for a specified θ_0. A *generalized test variable* T(X, x; θ, δ) satisfies the following conditions:

i. given the observed value x, the distribution of T(X, x; θ, δ) is free of the nuisance parameter δ,
ii. the observed value of T(X, x; θ, δ), namely T(x, x; θ, δ), is free of the nuisance parameter δ,
iii. given the observed data and the nuisance parameter δ, the distribution of T(X, x; θ, δ) is stochastically monotone in θ (i.e., stochastically increasing or decreasing in θ).

If the distribution of T(X, x; θ, δ) is stochastically increasing in θ, the generalized p-value for testing the above hypotheses is given by $P\left[T(X, x; \theta_0, \delta) \leq T(x, x; \theta_0, \delta)\right]$. If the distribution of T(X, x; θ, δ) is stochastically decreasing in θ, the generalized p-value is given by $P\left[T(X, x; \theta_0, \delta) \geq T(x, x; \theta_0, \delta)\right]$.

The generalized confidence interval method and generalized p-value approach are together referred to as the generalized variables approach. Numerous applications of generalized variables methodology have appeared in the literature, especially for the interval estimation and hypothesis testing for problems for which traditional solutions are difficult to obtain. Several such applications are given in the books by Weerahandi.[14,15] In the industrial hygiene context, such applications are also discussed in Krishnamoorthy and Mathew[16] and Krishnamoorthy, Mathew, and Ramachandran.[17,18] In particular, the generalized variables methodology has resulted in useful procedures for comparing the arithmetic means of two lognormal distributions based on complete

(i.e., uncensored) samples, even when the sample sizes are small; see Krishnamoorthy, Mathew, and Ramachandran.[17]

RESULTS

In this section, we first provide results on confidence intervals for the problem of comparing two lognormally distributed exposure populations, followed by results on hypothesis testing. Numerical results will also be reported to assess the performance of the proposed confidence intervals and test procedures.

Consider a simple random sample of n_i observations subject to k_i detection limits (on the log-transformed data set), say, $DL_{i1}, \ldots, DL_{ik_i}$, from a lognormal distribution with parameters μ_i and σ_i^2, which are the mean and variance, respectively, of the log-transformed exposure population, $i = 1,2$. We assume that for the ith exposure population, n_{ij} measurements were obtained by the jth laboratory procedure or device, with a detection limit that can be expressed in the same measurement unit as that of the detected observations; $j = 1, \ldots, k_i$ and $i = 1,2$. The quantity DL_{ij} denotes the log-transformed detection limits. Furthermore, let m_{ij} denote the number of nondetects below DL_{ij}, $j = 1, \ldots, k_i$, $i = 1,2$, and let $m_i = \sum_{j=1}^{k_i} m_{ij}$ denote the total number of nondetects from the ith sample, $i = 1,2$. Note also that the ith sample size is $n_i = \sum_{j=1}^{k_i} n_{ij}$, $i = 1,2$. We use the notation *(i, 1), (i, 2),. . . , (i,k$_i$)* to denote the k_i laboratory procedures or devices that made measurements from the ith exposure population. For easy reference, we describe the samples and notation in Table I.

Recall that the arithmetic mean of a lognormal distribution with parameters μ and σ^2 is given by $\exp(\eta)$, where $\eta = \mu + .5\sigma^2$. Let $\eta_i = \mu_i + .5\sigma_i^2$, $i = 1,2$. Note that μ is the log of the true geometric mean, and σ is the log of true geometric standard deviation.

Comparison of the Means: Confidence Intervals

Following Krishnamoorthy, Mathew, and Ramachandran,[17] for comparing the arithmetic means of two lognormally distributed exposure populations, we shall consider the ratio of the arithmetic means, which can be written as

$$R = \frac{\exp(\eta_1)}{\exp(\eta_2)} = \exp(\eta_1 - \eta_2). \quad \textbf{(5)}$$

Thus, a comparison between the two means can be made by comparing η_1 and η_2. In particular, a CI and test for R can be obtained from those for $\eta_1 - \eta_2$. For this we shall use the generalized confidence interval idea mentioned in the previous section. In fact a generalized pivotal quantity (GPQ) for $\eta_1 - \eta_2$ can be readily obtained from the one for the one-sample case given in Krishnamoorthy and Xu[10] as follows:

Let $(\hat{\mu}_{i0}, \hat{\sigma}_{i0})$ be an observed value of $(\hat{\mu}_i, \hat{\sigma}_i)$, $i = 1,2$. That is, $(\hat{\mu}_i, \hat{\sigma}_i)$ is a random quantity representing the MLEs, and $(\hat{\mu}_{i0}, \hat{\sigma}_{i0})$ denote the numerical values of the MLEs that are computed from the given exposure data, $i = 1,2$. An approximate GPQ for η_i is given by

$$G_{\eta_i} = \hat{\mu}_{i0} + \frac{\mu_i - \hat{\mu}_i}{\hat{\sigma}_i}\hat{\sigma}_{i0} + .5\frac{\hat{\sigma}_{i0}^2}{\hat{\sigma}_i^2}\sigma^2 \sim \hat{\mu}_{i0} - \frac{\hat{\mu}_i^*}{\hat{\sigma}_i^*}\hat{\sigma}_{i0} + .5\frac{\hat{\sigma}_{i0}^2}{\hat{\sigma}_i * 2} \text{ approximately,} \quad \textbf{(6)}$$

where the notation " $\sim$" means "distributed as," and $\hat{\mu}_i^*$ and $\hat{\sigma}_i^*$ are the MLEs based on a sample of size n_i from a standard normal distribution with detection limits $DL_{i1}^* < \cdots < DL_{ik_1}^*$, with $DL_{ij}^* = (DL_{ij} - \hat{\mu}_{i0})/\hat{\sigma}_{i0}$, $j = 1, \ldots, k_i$, $i = 1,2$; see the discussion on maximum likelihood estimation outlined in the previous section. The above GPQ is only an approximate GPQ since the distribution of $\frac{\mu_i - \hat{\mu}_i}{\hat{\sigma}_i}$ is the same as that of $\frac{\hat{\mu}_i^*}{\hat{\sigma}_i^*}$, only approximately. An approximate GPQ for $\eta_1 - \eta_2$ is then given by $G_{\eta_1} - G_{\eta_2}$. Confidence intervals as well as one-sided confidence limits for the difference $\eta_1 - \eta_2$ can be obtained using the percentiles of the GPQ $G_{\eta_1} - G_{\eta_2}$, and the required percentiles can be evaluated using the Monte Carlo algorithm given below.

We emphasize that the GPQ in (6) involves the MLEs $(\hat{\mu}_{i0}, \hat{\sigma}_{i0})$ obtained from a given set of data, as well as the MLEs $(\hat{\mu}_i^*, \hat{\sigma}_i^*)$ obtained from a standard normal distribution. The latter quantities are random variables, and consequently, the GPQ itself is a random variable, even though $(\hat{\mu}_{i0}, \hat{\sigma}_{i0})$ are held fixed while computing the percentiles of the GPQ.

Monte Carlo Algorithm

For given log-transformed samples with multiple detection limits, compute $\hat{\mu}_{10}$, $\hat{\mu}_{20}$, $\hat{\sigma}_{10}$ and $\hat{\sigma}_{20}$, where these quantities are, respectively, the MLEs of μ_1, μ_2, σ_1 and σ_2, computed from the given exposure samples. Let k_i be the number of DLs in the ith sample, and set

$$DL_{ij}^* = \frac{DL_{ij} - \hat{\mu}_{i0}}{\hat{\sigma}_{i0}}, \; j = 1, \ldots, k, \; i = 1, 2.$$

1. Generate k_1 independent samples of sizes $n_{11}, \ldots, n_{1k_1}$ from $N(0,1)$ distribution. Find m_{1j} the number of observations below DL_{1j}^* in the sample of size n_{1j}, $j = 1, \ldots, k$. Pool the samples so that the pooled sample is of size $n_1 = \sum_{j=1}^{k_1} n_{1j}$ with a total of m_{1j} nondetects below DL_{1j}^*, $j = 1, \ldots, k_1$
2. Generate k_2 independent samples of sizes $n_{21}, \ldots, n_{2k_2}$ from $N(0,1)$ distribution. Find m_{2j} the number of observations below DL_{2j}^* in the sample of size n_{1j}, $j = 1, \ldots, k_2$. Pool the samples so that the pooled sample is of size $n_2 = \sum_{j=1}^{k_2} n_{2j}$ with a total of m_{2j} nondetects below DL_{2j}^*, $j = 1, \ldots k_2$.
3. Compute the MLEs $\hat{\mu}_1^*$ and $\hat{\sigma}_1^*$ based on the pooled sample generated in step 1, and compute the MLEs $\hat{\mu}_2^*$ and $\hat{\sigma}_2^*$ based on the pooled sample generated in step 2.
4. Set $Q = (\hat{\mu}_{10} - \frac{\hat{\mu}_1^*}{\hat{\sigma}_1^*}\hat{\sigma}_{10} + .5\frac{\hat{\sigma}_{10}^2}{\hat{\sigma}_1^{*2}}) - (\hat{\mu}_{20} - \frac{\hat{\mu}_2^*}{\hat{\sigma}_2^*}\hat{\sigma}_{20} + .5\frac{\hat{\sigma}_{20}^2}{\hat{\sigma}_2^{*2}})$

Repeat steps 1 – 4 a large number of times, say, 10,000.

The $100(1-\alpha)$ percentile of the 10,000 Q's (denoted by $Q_{1-\alpha}$) is a $100\,(1-\alpha)\,\%$upper confidence limit for $\eta_1 - \eta_2$. The interval $(Q_{\alpha/2}, Q_{1-\alpha/2})$ is a $100\,(1-\alpha)\,\%$confidence interval of $\eta_1 - \eta_2$, and $\big(\exp(Q_{\alpha/2}), \exp(Q_{1-\alpha/2})\big)$ is a $100\,(1-\alpha)\,\%$CI for the ratio of the means of two lognormal distributions defined in Eq. (5). An R program to compute confidence limits for the mean of a single lognormal distribution, using the above algorithm, is posted at http://www.ucs.louisiana.edu/∼kxk4695.

Note that, to find a CI for the difference between the means of two lognormal distributions, that is, for $\exp(\mu_1 + .5\sigma_1^2) - \exp(\mu_2 + .5\sigma_2^2)$, using the algorithm, we simply need to modify Q in Step 4 as

$$Q = \exp\left(\hat{\mu}_{10} - \frac{\hat{\mu}_1^*}{\hat{\sigma}_1^*}\hat{\sigma}_{10} + .5\frac{\hat{\sigma}_{10}^2}{\hat{\sigma}_1^{*2}}\right) - \exp\left(\hat{\mu}_{20} - \frac{\hat{\mu}_2^*}{\hat{\sigma}_2^*}\hat{\sigma}_{20} + .5\frac{\hat{\sigma}_{20}^2}{\hat{\sigma}_2^{*2}}\right).$$

Similarly, to find a CI for the difference between the geometric means (i.e., medians) of two lognormal distributions, we simply need to set $Q = \exp(\hat{\mu}_{10} - \frac{\hat{\mu}_1^*}{\hat{\sigma}_1^*}\hat{\sigma}_{10}) - \exp(\hat{\mu}_{20} - \frac{\hat{\mu}_2^*}{\hat{\sigma}_2^*}\hat{\sigma}_{20})$. We also note that the choices of the n_{ij}'s do not affect the MLEs; see the log-likelihood function Eq. (2). So, if n_{ij}'s are not recorded or not reported for some data sets, then we can choose n_{ij}'s in the Monte Carlo algorithm arbitrarily with the constraints that $n_{ij} \geq m_{ij}$and $\sum_{j=1}^{k_i} n_{ij} = n_i$, $i = 1,2$.

Estimated Coverage Probabilities

As the procedures in the preceding sections are approximate, we shall numerically investigate their statistical properties. In the context of a confidence interval, a quantity of interest is its coverage probability, which is the proportion of confidence intervals that will include the true parameter value if such intervals are repeatedly computed based on independent samples. For a 95% confidence interval, a satisfactory interval will have coverage probability close to 0.95. In view of this, we estimated the coverage probabilities of the confidence interval for the ratio of the arithmetic means as follows:

Recall that the ratio of the arithmetic means is the parameter R given in Eq. (5), and a confidence interval for the ratio can be obtained from a confidence interval for the difference $\eta_1 - \eta_2$; see Eq. (5). We first generated 2500 pairs of samples from the two lognormal exposure populations for specified values of the parameters and specified values for the proportions of nondetects. For each generated pair of samples, we used the algorithm to compute a 95% confidence interval for the ratio of the arithmetic means using 5000 simulation runs (we used 5000 simulation runs for computing the coverage probability instead of the 10,000 simulation runs recommended in the Monte Carlo algorithm given earlier, since the coverage probability calculation is time-consuming, as opposed to analyzing a given pair of samples). The percentage of these 2500 confidence intervals that include the true ratio of the arithmetic means is a Monte Carlo estimate of the coverage probability. For convenience, we assumed that the number of DLs is the same in both samples. In the simulation, instead of specifying the DLs, we have specified the proportions of nondetects. For instance, with two DLs, if the proportions of nondetects are $(p_1, p_2) = (.2, .3)$, the DLs are$DL_{11} = \mu_1 + z_{.2}\sigma_1$, $DL_{12} = \mu_1 + z_3\sigma_1$,$DL_{21} = \mu_2 + z_{.2}\sigma_2$, and $DL_{22} = \mu_2 + z_{.3}\sigma_2$, where z_p denotes the *100p*th percentile of the standard normal distribution. The parameter choices that we have made for the simulations are all on the log scale. For example, our choice $\mu_1 = \mu_2 = 1$ on the log scale corresponds to lognormal distributions with geometric means each equal to 2.71. The estimated coverage probabilities are given in Table II for some selected values of the parameters, proportions of nondetects, and sample sizes. In the table the notations p_1, (p_1, p_2), and (p_1, p_2, p_3) are used to denote the proportion of nondetects when the number of detection limits for each sample is one, two, and three, respectively.

We see from Table II that the estimated coverage probabilities are close to the nominal confidence level .95 for most cases considered, and they are slightly larger than the nominal level when the proportions of nondetects are larger. Overall, the proposed interval estimation procedure is *satisfactory* even for small samples when the proportions of nondetects are not large. For some parameter configurations, the confidence intervals could be conservative for small samples with large proportions of nondetects. However, none of the coverage probabilities reported in Table II is unacceptably large.

Comparison of the Means: Hypothesis Testing

A generalized test variable for testing hypotheses on the difference $\eta_1 - \eta_2$ is given by

$$T_{\eta_1-\eta_2} = G_{\eta_1} - G_{\eta_2} - (\eta_1 - \eta_2), \tag{7}$$

where $G_{\eta_i}$is given in Eq. (6). The generalized p-value for testing the hypotheses

$$H_0 : \eta_1 \geq \eta_2 \; vs. \; H_\alpha : \eta_1 < \eta_2 \tag{8}$$

is given by

$$\begin{aligned} &\underset{H_0}{Sup}\, P(T_{\eta_1-\eta_2} \geq 0) \\ &= \underset{H_0}{Sup}\, P(G_{\eta_1} - G_{\eta_2} \geq \eta_1 - \eta_2) = P(G_{\eta_1} \geq G_{\eta_2}). \end{aligned} \tag{9}$$

The generalized p-value for testing hypotheses $H_0 : \eta_1 = \eta_2 vs. H_\alpha : \eta_1 \neq \eta_2$ is

$$2\min\{P(G_{\eta_1} \geq G_{\eta_2}),\, P(G_{\eta_1} < G_{\eta_2})\} \tag{10}$$

We refer to Krishnamoorthy, Mathew, and Ramachandran[17] for further details on such a test procedure in the absence of nondetects. The above generalized p-values can be estimated using the Monte Carlo algorithm. Specifically, the proportion of Q's (generated using the Monte Carlo algorithm) that are greater than or equal to zero is a Monte Carlo estimate of the generalized p-value for testing the hypotheses in Eq. (8).

Note that for comparing the arithmetic means using confidence intervals, we can consider the ratio parameter as

TABLE II. Coverage Probabilities of 95% CIs for the Ratio of Two Lognormal Means $\mu_1 = \mu_2 = 1$

	$(\sigma_1^2, \sigma_2^2) = (1, 3)$ (n_1, n_2)				$(\sigma_1^2, \sigma_2^2) = (4, 5)$ (n_1, n_2)			
p_1	(7,11)	(11,11)	(11,18)	(20,20)	(7,11)	(11,11)	(11,18)	(20,20)
.2	.95	.95	.95	.95	.95	.95	.95	.95
.3	.95	.96	.95	.95	.95	.95	.94	.95
.5	.96	.96	.94	.95	.95	.96	.95	.95
.6	.96	.95	.95	.95	.95	.95	.95	.96
.7	.95	.96	.96	.95	.95	.96	.95	.95
.8	.95	.96	.96	.95	.96	.96	.96	.96
	$(\sigma_1^2, \sigma_2^2) = (1, 1)$ (n_1, n_2)				$(\sigma_1^2, \sigma_2^2) = (1, 3)$ (n_1, n_2)			
(p_1, p_2)	(10,10)	(10,15)	(15,15)	(20,20)	(10,10)	(10,15)	(15,15)	(20,20)
(.2,.3)	.95	.95	.94	.95	.94	.95	.95	.95
(.4,.5)	.95	.95	.95	.95	.95	.95	.96	.95
(.5,.6)	.95	.95	.94	.95	.94	.95	.96	.95
(.6,.7)	.97	.95	.96	.95	.96	.96	.96	.95
(.7,.8)	.97	.97	.97	.96	.96	.97	.96	.95
	$(\sigma_1^2, \sigma_2^2) = (2, 5)$ (n_1, n_2)				$(\sigma_1^2, \sigma_2^2) = (1, 4)$ (n_1, n_2)			
(p_1, p_2, p_3)	(12,12)	(15,18)	(15,24)	(24,24)	(12,12)	(15,18)	(15,24)	(24,24)
(.2,.4,.5)	.95	.95	.95	.95	.95	.95	.95	.95
(.3,.4,.5)	.95	.95	.95	.95	.95	.95	.95	.95
(.4,.5,.6)	.95	.95	.95	.95	.95	.95	.95	.95
(.3,.6,.7)	.96	.96	.95	.95	.96	.95	.95	.95
(.5,.6,.7)	.97	.96	.95	.95	.96	.96	.96	.95

mentioned in Eq. (5), and we can also consider the difference between the arithmetic means, as pointed out earlier. However, for testing the hypothesis that the true ratio is one (equivalently, the true difference is zero), formulating the problem in terms of the ratio of the arithmetic means, or their difference, does not matter, since the p-values in Eqs. (9) and (10) do not depend on which formulation is used.

The coverage properties of the confidence intervals along with the duality between the interval estimation procedure and hypothesis testing indicate that the type I error rates of a two-sided test should be close to the nominal level .05, and they could be smaller than the nominal level .05 for small samples with large proportions of nondetects. To judge the type II error rates, and the efficiency of the test for one-sided hypotheses, we evaluated the powers of the test for testing the hypotheses in Eq. (8) for a few sample size and parameter configurations, and when the samples include two or three DLs. The estimated powers are reported in Table III for testing $H_0 : \eta_1 \geq \eta_2$ *vs.* $H_\alpha : \eta_1 < \eta_2$. The powers in Table III are reported for the parameter values $\sigma_1 = \sigma_2 = 1$, and for different choices of (μ_1, μ_2). Recall that $\eta_i = \mu_i + .5\sigma_i^2, i = 1,2$.

The reported powers in Table III clearly indicate that, for fixed sample sizes and assumed values of the percentage of nondetects, the power is increasing with increasing difference between μ_1 and μ_2. Note that an increasing difference between μ_1 and μ_2 implies an increasing difference between η_1 and η_2 (since $\eta_i = \mu_i + .5\sigma_i^2, i = 1,2,$ and we have assumed $\sigma_1 = \sigma_2 = 1$). In other words, the test should detect the difference between η_1 and η_2 with a higher probability if the difference $\eta_2 - \eta_1$ is large. For fixed parameter values, the power is also an increasing function of the sample sizes. Thus, the test possesses some natural properties of an efficient test. Furthermore, by comparing the powers given in Table III, we notice that the power is decreasing with increasing proportions of nondetects. Overall, we see that the test is efficient and possesses some natural properties.

Finally, we note that censoring intensity certainly affects the properties of a test or a confidence interval. Also, we note that our methods require at least two detected values to compute the MLEs of the parameters. Our extensive simulation studies (not all reported here) indicated that the properties of proposed tests are affected by the overall percentage of nondetects, not by the individual proportions of nondetects below each of the DLs. If the sample size is 6 to 9, then our methods are satisfactory as long as the percentage of nondetects is around 50% or less. If the sample size is 10 or more, then the proposed methods are expected to be satisfactory provided the percentage of nondetects is no more than 80%. If the percentage of nondetects exceeds the aforementioned values, then our results could be conservative yielding confidence

TABLE III. Powers of the Test for $H_0 : \eta_1 \geq \eta_2$ vs. $H_\alpha : \eta_1 < \eta_2$; Level $\alpha = 0.05$, and $\sigma_1 = \sigma_2 = 1$

	$n_1 = 10, n_2 = 10$ (μ_1, μ_2)				$n_1 = 10, n_2 = 16$ (μ_1, μ_2)				$n_1 = 16, n_2 = 16$ (μ_1, μ_2)			
(p_1, p_2)	(2,2)	(2,2.5)	(2,3)	(2,3.5)	(2,2)	(2,2.5)	(2,3)	(2,3.5)	(2,2)	(2,2.5)	(2,3)	(2,3.5)
(.2,.3)	.05	.16	.43	.63	.05	.41	.79	.92	.05	.54	.89	.97
(.3,.5)	.05	.15	.33	.55	.05	.39	.75	.89	.05	.52	.86	.96
(.4,.5)	.04	.13	.33	.47	.05	.37	.75	.89	.05	.51	.87	.95
(.5,.6)	.04	.12	.28	.46	.04	.25	.57	.75	.03	.43	.79	.91
(.6,.7)	.03	.10	.26	.35	.04	.20	.48	.64	.03	.38	.70	.85
	$n_1 = 9, n_2 = 15$ (μ_1, μ_2)				$n_1 = 15, n_2 = 21$ (μ_1, μ_2)				$n_1 = 24, n_2 = 24$ (μ_1, μ_2)			
(p_1, p_2, p_3)	(2,2)	(2,2.5)	(2,3)	(2,3.5)	(2,2)	(2,2.5)	(2,3)	(2,3.5)	(2,2)	(2,2.5)	(2,3)	(2,3.5)
(.2,.4,.5)	.05	.18	.36	.59	.05	.27	.36	.78	.05	.31	.77	.94
(.3,.4,.5)	.05	.18	.35	.56	.05	.26	.35	.77	.05	.30	.76	.94
(.4,.5,.6)	.05	.18	.35	.54	.05	.26	.35	.74	.05	.28	.76	.92
(.3,.6,.7)	.04	.15	.33	.54	.05	.22	.33	.74	.05	.28	.73	.92
(.5,.6,.7)	.04	.12	.27	.52	.04	.19	.28	.71	.05	.24	.63	.86

intervals that are too wide and tests that may not be very powerful.

Illustrative Examples

We shall now use two examples to illustrate the generalized variable procedures for finding the confidence intervals for comparing the means of the two lognormal distributions. The first example involves simulated data sets from two assumed lognormal models, and the second example involves real data representing concentrations of an analyte in the urine of exposed workers. The first example is included so that we can check if the true difference or the ratio of the means of the two lognormal distributions is contained in the corresponding confidence interval.

Example 1

To illustrate the methods of earlier sections, we shall use simulated samples from two lognormal distributions. Sample 1 is generated as follows:

We generated a sample of $n_{11} = 16$ observations, assuming a log-transformed detection limit of $DL_{11} = -.049$, from a lognormal $(\mu = 1, \sigma = 2)$ distribution, and another independent sample of $n_{12} = 16$ observations with the log-transformed detection limit $DL_{12} = .493$ from the same lognormal $(\mu = 1, \sigma = 2)$ distribution. These two samples were pooled so that the pooled sample consists of $n = 32$ observations with two detection limits. Similarly, sample 2 consists of $n_2 = 39$ observations with $(n_{21}, n_{22}, n_{23}) = 13$, generated form a lognormal $(\mu = 0.5, \sigma = 1.5)$ distribution with three log-transformed detection limits: $-.287, .120, .500$. Log-transformed samples with the detection limits are given in Table IV. The statistics required to find confidence intervals using the Monte Carlo algorithm are provided in Table V.

Using the MLEs given in Table V, we computed an estimate for the ratio of means as $\hat{R} = [\exp(\hat{\mu}_{10} + .5\hat{\sigma}_{10}^2)]/[\exp(\hat{\mu}_{20} + .5\hat{\sigma}_{20}^2)] = 4.094$. Note that the true ratio defined in Eq. (5) is 3.955. Furthermore, using the quantities given in Table V in the Monte Carlo algorithm with 10,000 simulation runs, we computed the 95% generalized confidence interval for the difference $(\mu_1 + .5\sigma_1^2) - (\mu_2 + .5\sigma_2^2)$ as (.171, 3.983). The 95% CI for the ratio of the means of the two lognormal distributions is thus (exp(.171),exp(3.983)) = (1.2, 53.7). Note that the true ratio 3.955, and the MLE value of 4.094, are both included in the 95% CI.

TABLE IV. Simulated Log-transformed Samples from Two Lognormal Distributions

Sample 1			Sample 2		
1.719	$< -.049$	$<.493$	0.509	$<.500$	0.746
$< -.049$	2.013	0.638	-0.054	1.664	$<.500$
$< -.049$	3.907		$<.120$	3.336	1.799
2.116	$< -.049$		$<.120$	$< -.287$	$<.500$
4.140	1.638		1.571	1.617	1.866
$< -.049$	$<.493$		0.206	$<.500$	$< -.287$
2.371	0.494		$< -.287$	0.353	1.909
3.201	$<.493$		$<.500$	$<.120$	$<.120$
$<.493$	1.634		-0.059	2.678	$<.500$
2.203	$< -.049$		1.174	1.864	
$< -.049$	1.553		$<.120$	$<.120$	
1.335	3.260		0.249	1.662	
$<.493$	2.663		0.387	2.334	
4.138	$< -.049$		$< -.287$	$< -.287$	
1.651	2.518		1.693	1.080	

Note: Sample 1 is generated with parameters $(\mu_1, \sigma_1) = (1,2)$, and sample 2 is generated with parameters $(\mu_2, \sigma_2) = (0.5,1.5)$. The log-transformed detection limits for the first sample are $-.049$ and .493; for the second sample, they are $-.287$, .120, and .500.

TABLE V. Summary Statistics for Example 1

	Sample Sizes	DLs	#non-detects	MLEs	$DL^* = \frac{DL-\hat{\mu}_0}{\sigma_0}$
1	$n_1 = 32$	$DL_{11} = -.049$	$m_{11} = 8$	$\hat{\mu}_{10} = .846$	$DL^*_{11} = -.440$
	$n_{11} = n_{12} = 16$	$DL_{12} = 0.493$	$m_{12} = 5$	$\hat{\sigma}_{10} = 2.032$	$DL^*_{12} = -.174$
2	$n_2 = 39$	$DL_{21} = -.287$	$m_{21} = 5$	$\hat{\mu}_{20} = .328$	$DL^*_{21} = -.436$
	$n_{21} = n_{22} = n_{23} = 13$	$DL_{22} = .120$	$m_{22} = 6$	$\hat{\sigma}_{20} = 1.409$	$DL^*_{22} = -.148$
		$DL_{23} = .500$	$m_{23} = 6$		$DL^*_{23} = .122$

An estimate of the mean difference based on the MLEs is given by $\exp(\hat{\mu}_{10} + .5\hat{\sigma}^2_{10}) - \exp(\hat{\mu}_{20} + .5\hat{\sigma}^2_{20}) = 14.62$, and the true difference is 15.01. We also computed the 95% generalized confidence interval for the difference of the means, i.e., the differenceexp$(\mu_1 + .5\sigma_1^2) - \exp(\mu_2 + .5\sigma_2^2)$. The interval came out to be (1.3, 190), which includes the true difference 15.01 and the estimated value 14.6.

We note that with a standard group-comparison t-test there is an assumption of equality of variances. However, this assumption is not needed here; the assumption has not been used while developing the GPQ. Also, such an assumption is unlikely to hold in the above example; this should be clear from Tables IV and V.

Example 2

This example is on the effectiveness of implementing a work practice in reducing the concentration of an analyte in the urine of exposed workers. Two control methods (A and B, say) were used, and six workers were sampled before and after the work practice was implemented. Urine samples were collected and analyzed for the analyte. The analyte concentrations (micromoles/L), obtained before implementing the work practice, are given below; the data are subject to a single detection limit of 1 1.4.

Method A: 22.2, <11.4, 16.8, <11.4, 22.2, <11.4
Method B: 14.7, <11.4, 34.1, <11.4, 18.2, <11.4

Here we shall consider the problem of comparing the means of the two control methods A and B.. Summary statistics for these data are given in Table VI.

TABLE VI. Summary Statistics for Example 2

	Sample size	DLs	#non-detects	MLEs	$DL^* = \frac{DL-\hat{\mu}_0}{\sigma_0}$
Method A	$n_1 = 6$	$DL_{11} = 11.4$	$m_1 = 3$	$\hat{\mu}_{10} = 2.517$ $\hat{\sigma}_{10} = 0.546$	$DL^*_{11} = -.152$
Method B	$n_2 = 6$	$DL_{21} = 11.4$	$m_2 = 3$	$\hat{\mu}_{20} = 2.468$ $\hat{\sigma}_{20} = 0.688$	$DL^*_{21} = -.050$

Based on the summary statistics, and 10,000 simulation runs, we computed 95% confidence interval for the difference $(\mu_1 + .5\sigma_1^2) - (\mu_2 + .5\sigma_2^2)$ as (−63.5, 18.3). The 95% confidence interval for the ratio of means is (exp(−63.5), exp(18.3)) = (0, 8.86×10^7). This confidence interval indicates that there is no significant difference between the group means. A point estimate of the ratio of means is $\hat{R} = [\exp(\hat{\mu}_{10} + .5\hat{\sigma}^2_{10})]/[\exp(\hat{\mu}_{20} + .5\hat{\sigma}^2_{20})] = .962$, and an estimate of the difference of means is $\exp(\hat{\mu}_{10} + .5\hat{\sigma}^2_{10}) - \exp(\hat{\mu}_{20} + .5\hat{\sigma}^2_{20}) = -.570$.

To compare the medians of these two, we found 95% confidence interval for$\exp(\mu_1) - \exp(\mu_2)$ as (−5.11, 9.03). As this confidence interval includes zero, we can conclude that the group medians are not significantly different. A point estimate for the difference between the two group medians is$\exp(\hat{\mu}_{10}) - \exp(\hat{\mu}_{20}) = .0001$.

DISCUSSION

This article addresses a commonly occurring problem in the analysis of exposure data in the presence of nondetects. As is well known, lognormality can be assumed for many exposure data sets, and the estimation of the lognormal parameters can be carried out using standard statistical methodologies, for example, using maximum likelihood even if multiple detection limits are present. However, it appears that for computing confidence limits no satisfactory procedure was available that is satisfactory regardless of the sample size, and regardless of the proportion of nondetects, until the work of Krishnamoorthy and Xu.[10] The work by these authors dealt with a single exposure sample. The present work extends their ideas to the problem of comparing two lognormally distributed exposure samples.

CONCLUSION

The work reported in this article shows that the maximum likelihood estimates can be used to compute satisfactory confidence limits and tests for the problem of comparing two exposure populations that follow lognormal distributions, in the presence of multiple detection limits. We would like to emphasize that the likelihood only uses information on the number of values below each detection limit; no substitution values are used to replace them. Furthermore, based on simulation results, we have demonstrated the utility of the

proposed solution developed in Eqs. (6)–(10), for comparing the arithmetic means of two lognormally distributed exposure populations. We believe that we have developed a much–needed framework for analyzing lognormally distributed industrial hygiene data subject to multiple detection limits.

ACKNOWLEDGMENTS

The authors thank Professor Gurumurthy Ramachandran for reviewing the manuscript and for several helpful comments. They are also grateful to Dr. Paul Hewett for suggesting Example 2, and to two reviewers for providing detailed comments and suggestions. The work was supported by grant R01OH003628 from the National Institute of Occupational Safety and Health (NIOSH).

REFERENCES

1. **Helsel D. R.:** Much ado about next to nothing: incorporating nondetects in science. *Ann. Occup. Hyg. 54*: 257–262 (2010).
2. **Helsel, D. R.:** *Statistics for Censored Environmental Data Using Minitab and R*. New York: Wiley, 2012.
3. **Helsel, D. R.:** Fabricating data: How substituting values for nondetects can ruin results, and what can be done about it. *Chemosphere 65*:2434–2439 (2006).
4. **Frome, E.L., and P.F. Wambach:** "Statistical Methods and Software for the Analysis of Occupational Exposure Data with Non-detectable Values." Report by Oak Ridge National Laboratory for U.S. Department of Energy, 2005. Available at www.hss.energy.gov/healthsafety/iipp/hservices/statmethods.pdf (accessed 2005).
5. **Ganser, G.H., and P. Hewett:** An accurate substitution method for analyzing censored data. *J. Occup. Environ. Hyg. 7*:233–244 (2010).
6. **Hewett, P.:** Analysis of censored data. In *A Strategy for Assessing and Managing Occupational Exposures*, J. S. Ignacio and W. H. Bullock, eds. Fairfax, Va.: AIHA Press, 2006. pp. 415–422.
7. **Hewett, P., and G.H. Ganser:** A comparison of several methods for analyzing censored data. *Ann. Occup. Hyg. 51*:611–632 (2007).
8. **Krishnamoorthy, K., A. Mallick, and T. Mathew:** Model based imputation approach for data analysis in the presence of nondetectable values: Normal and related distributions. *Ann. Occup. Hyg. 59*:249–261 (2009).
9. **Krishnamoorthy, K., A. Mallick, and T. Mathew:** Inference for the lognormal mean and quantiles based on samples with left and right type I censoring. *Technometrics 53*:72–83 (2011).
10. **Krishnamoorthy, K., and Z. Xu:** Confidence limits for lognormal percentiles and for lognormal means based on samples with multiple detection limits. *Ann. Occup. Hyg. 55*:495–509 (2011).
11. **Cohen, A.C.:** On the solution of estimating equations for truncated and censored samples from normal populations. *Biometrika 44*: 225–236 (1957).
12. **Weerahandi, S.:** Generalized confidence intervals. *J. Amer. Statist. Assoc. 88*:899–905 (1993).
13. **Tsui, K.W., and S. Weerahandi:** Generalized p-values in significance testing of hypotheses in the presence of nuisance parameters. *J. Amer. Statist. Assoc. 84*:602–607 (1989).
14. **Weerahandi, S.:** *Exact Statistical Methods for Data Analysis*. New York: Springer-Verlag, 1995.
15. **Weerahandi, S.:** *Generalized Inference in Repeated Measures: Exact Methods in MANOVA and Mixed Models*. New York: Wiley, 2004.
16. **Krishnamoorthy, K., and T. Mathew:** Assessing occupational exposure via the one-way random model with balanced data. *J. Agric., Biol. Environ. Statist. 7*:440–457 (2002).
17. **Krishnamoorthy, K., T. Mathew, and G. Ramachandran:** Generalized p-values and confidence intervals: A novel approach for analyzing lognormally distributed exposure data. *J. Occup. Environ. Hyg. 3*:642–650 (2006).
18. **Krishnamoorthy, K., T. Mathew, and G. Ramachandran:** Upper limits for the exceedance probabilities in the one-way random effects model. *Ann. Occup. Hyg. 51*:397–406 (2007).