

TUTORIAL

Variability in the Log Domain and Limitations to Its Approximation by the Normal Distribution

Jeroen Elassaiss-Schaap^{1,2,*} and Kevin Duisters³

Pharmacometric models using lognormal distributions have become commonplace in pharmacokinetic–pharmacodynamic investigations. The extent to which it can be interpreted by traditional description of variability through the normal distribution remains elusive. In this tutorial, the comparison is made using formal approximation methods. The quality of the resulting approximation was assessed by the similarity of prediction intervals (PIs) to true values, illustrated using 80% PIs. Approximated PIs were close to true values when lognormal standard deviation (ω) was smaller than about 0.25, depending mostly on the desired precision. With increasing ω values, the precision of approximation worsens and starts to deteriorate at ω values of about 1. With such high ω values, there is no resemblance between the lognormal and normal distribution anymore. To support dissemination and interpretation of these nonlinear properties, some additional statistics are discussed in the context of the three regions of behavior of the lognormal distribution.

The lognormal distribution has a widespread application in the pharmacometric community. The focus of nonlinear mixed effects approaches in pharmacometrics has been on the development of pharmacokinetic models initially, shifting toward pharmacokinetic–pharmacodynamic (PKPD) models later on.¹ The variability in pharmacokinetic models is typically of a limited magnitude, specifically when measured in smaller and well-controlled studies, for example, with coefficient of variation (CV) in a range of 5–30%. Variability in outpatient studies tends to be larger, e.g., 25–75 %CV, whereas biomarkers as assessed with PKPD models often present with much larger variability, e.g., 50–150 %CV.

Rationale for applying the lognormal distribution

Large variability, as typically encountered in pharmacological data sets with biomarkers, requires modeling assumptions different from those captured by the normal distribution. An early example in pharmacology dates back to 1972,² a study of effective doses in various tissues in which a clear case for the use of the lognormal distribution was made.

Most of the physiological processes modeled by pharmacometricians are strictly positive because drug exposure and biomarkers are often measured as concentrations in blood that cannot be negative. At the same time, parameters in pharmacokinetic and pharmacological models also need to remain positive to retain their meaning. Clearance, for example, describes the body's capacity to remove xenobiotics that it cannot produce on its own. Normal or Gaussian standard deviation (SD) cannot extend to more than half of the mean without the distribution also getting substantial coverage into negative values. The lognormal distribution is a natural alternative that can span well beyond this range while never producing negative values.

Moreover, a form of asymmetry is commonly required for models on physiological processes. This has been shown for observations on pharmacokinetic evaluations^{3,4} but also for general applications.⁵ The skewness has even been derived on the basis of theoretical grounds.^{6,7} Paraphrasing Gronholm and Annala,⁶ thermodynamic laws require that results from multiple reactions depend on intermediate states and therefore produce skewed distributions, independent of the distribution of individual reaction steps. The simplest distribution that complies with the criteria of positivity and asymmetry is the lognormal distribution. It is obtained by raising the normal distribution to the power of the base of choice, typically e (corresponding to the natural logarithm used as default). The result is strictly positive but has many more properties of interest. Although more distributions are available and used by the pharmacometric community through transformations,⁸ the lognormal distribution is by far the most commonly encountered.

Formal comparison of the approximated lognormal through PIs

Many pharmacometric applications have the objective to support decision making during pharmaceutical development or in the regulatory review process.¹ Whether the actual topic is trial outcome probability, population coverage, or the impact of special populations, the underlying property that drives the result is the model-derived PI, typically established at 80 or 90% of the population. Properties of the lognormal distribution will be evaluated against such PIs as the proverbial yardstick. The base distribution that we will use to compare the lognormal distribution against is the normal or Gaussian distribution. The normal distribution is obviously a popular distribution throughout scientific analysis. Also in pharmacometric textbooks the normal distribution is discussed

¹PD-Value B.V., Houten, The Netherlands; ²Leiden Academic Center for Drug Research (LACDR), Leiden University, The Netherlands; ³Mathematical Institute, Leiden University, Leiden, The Netherlands. *Correspondence: Jeroen Elassaiss-Schaap (jeroen@pd-value.com)

Received: November 20, 2019; accepted: March 3, 2020. doi:10.1002/psp4.12507

extensively, see, for example, ref. 9. Up to about 1990, the normal distribution was the only one available for the pharmacometric community.^{10–12} The normal distribution is therefore the reference distribution of choice. It will be investigated to which extent PIs generated using the lognormal distribution can be approximated by those of the normal distribution.

The standard approaches to describe variability in pharmacometric models are better suited for Gaussian distributions, with a focus on CV; see also the description in the previous few paragraphs. A lack of helpful statistical instruments and/or the application thereof may impede proper interpretation of results. Although the CV directly relates to the variability that is described, it does not reflect on the shape of the lognormal distribution. We therefore also revisit old metrics such as “skewness” and discuss the possibility of applying other metrics to capture the impact of variability magnitude on the shape of the distribution. The primary objective, however, is to precisely compare the lognormal distribution to the default Gaussian or normal distribution.

To be able to evaluate how well the normal distribution is capable of describing the lognormal distributions, a formal procedure is set up to calculate optimal normal parameters for given lognormal parameters using the Kullback–Leibler (KL) divergence of distributions and their probability density functions (PDFs). With these optimal parameters, the quality of approximated PIs is explored as a function of variability magnitude. To enable a wider audience to understand how these results are derived, the properties of the normal and lognormal distributions, as well as the KL divergence, will be derived in this article starting from first principles.

The derived optimal approach will then be used to determine the extent to which the lognormal distribution can be approximated by a normal distribution. The approximation was devised specifically for this purpose, and its use assures the limitations in approximating the lognormal distribution can be assessed objectively and robustly.

THEORETICAL

Formal properties of the lognormal distribution

This section presents theoretical properties of the lognormal distribution as can be found in textbooks using the normal distribution as starting point. Highlighting formal differences between these distributions will aid the more intuitive discussion in the remainder of this article. A glossary of terms is provided in **Supplemental Materials 13**. To preserve the flow, mathematical derivations are provided in the **Supplemental Materials 11**.

Motivation. One of the basic properties of the lognormal distribution is that its mean is not equal to its median, in contrast to the normal distribution. This has a large impact on interpretation. For example, pharmacometricians often need to explicitly explain to outsiders that predictions with mixed effects models are representative of the typical individual rather than the mean of the data. The impact of nonlinear mixed effects makes this necessary, e.g., in the classic example of averaging Emax curves. Notwithstanding, another driver of this representation is the difference between the mean and median of the lognormal distribution, e.g., in the discussion of linear

pharmacokinetics. It therefore is worthwhile to derive the basic properties of the lognormal distribution from first principles and establish them robustly.

Some basic definitions. Statistical models, such as those used in pharmacometrics, are typically calibrated on observations of random variables (i.e., collected data). In turn, these random variables are defined by the distribution they are assumed to follow. The first step in deriving the basic properties of a random variable is to define the distribution itself, which will be done by specification of its PDF. The collection of all values a random variable could take in theory is known as the “support.” For instance, a Gaussian (or normal) random variable can attain any value on the real line, i.e., its support is $(-\infty, \infty)$. The PDF defines for each value in the support what the probability density of (the occurrence of) that particular value is. This concept of density is a generalization of probability for discrete random variables, for instance, 1/6 being the probability a single roll of a fair die generates six eyes. Similarly, one reads off a density for a particular realization from a continuous random variable, such as the normal and lognormal. An important characteristic of any PDF is that its area under the curve (integrating over the entire support) equals one.

The normal distribution. Suppose random variable X is normally distributed with mean μ and SD $\sigma > 0$, denoted as $X \sim N(\mu, \sigma)$. Then, the PDF f_X of X for some realization x on $(-\infty, \infty)$ is given by:

$$f_X(x) := \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right). \quad (1)$$

Throughout, equalities that follow by definition are denoted as “:=”. The reader may verify that indeed $\int_{-\infty}^{\infty} f_X(x) dx = 1$, for which a standard Gaussian integration result is required. Using the PDF, basic characteristics such as the mean ($\mathbb{E}[X]$) and SD (the square root of the variance $\text{Var}[X]$) can be derived.

$$\mathbb{E}[X] := \int_{-\infty}^{\infty} x f_X(x) dx \quad (2)$$

In words, the expected value consists of the values that X can take multiplied by how often those values occur relative to each other; a density-weighted average. Again, intuition may be borrowed from discrete random variables. For instance, a random variable supported on $\{2, 4, 8\}$, realized with probabilities 1/4, 1/2, and 1/4, respectively, has an expected value of 4.5.

The expected value (Eq. 2) is a general definition for any continuous random variable X . In particular, for X Gaussian, we have:

$$\mathbb{E}[X] := \int_{-\infty}^{\infty} x f_X(x) dx = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = \mu \quad (3)$$

Here, equality * is worked out in **Supplemental Materials 11**. For the variance, it always follows that

$$\text{Var}[X] := \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \quad (4)$$

For X Gaussian the first term on the right-hand side equals

$$\mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = \sigma^2 + \mu^2, \quad (5)$$

and from Eq. 3 we know

$$(\mathbb{E}[X])^2 = \mu^2. \quad (6)$$

In summary, $\text{Var}[X] = \sigma^2$ for $X \sim N(\mu, \sigma)$. Again, mathematical details (*) are discussed in **Supplemental Materials 11**. The CV can now be plugged in easily:

$$\text{CV}[X] := \frac{\sqrt{\text{Var}[X]}}{\mathbb{E}[X]} = \frac{\sigma}{\mu} \quad (7)$$

Finally, we mention that symmetry and unimodality of the Gaussian PDF around μ lead to some well-known properties; the median equals μ , the skewness is 0, and the mode of the distribution is located at its mean. This is in sharp contrast to the lognormal distribution discussed next.

The lognormal distribution. Suppose random variable U is lognormally distributed with statistical parameters θ and ω , denoted as $U \sim \text{logN}(\theta, \omega)$. Starting from a normal random variable X with mean θ and SD ω , this U can be defined as follows:

$$U := \exp(X) \text{ with } X \sim N(\theta, \omega). \quad (8)$$

Remark that θ now takes the role of μ and ω that of σ in the notation of the previous subsection. Again, let us define the values that the distribution of U can take as u . From the definition in Eq. 8, it follows that the support of U is $(0, \infty)$. For any realization u on this support, the PDF of the lognormal is given by:

$$f_U(u) := \frac{1}{u\sqrt{2\pi}\omega} \exp\left(-\frac{(\log(u) - \theta)^2}{2\omega^2}\right). \quad (9)$$

Throughout this article, $\log()$ is understood as the natural logarithm with respect to base e unless mentioned otherwise. This PDF expression can be taken at face value or derived through the normal PDF using Eq. 8 as is done in **Supplemental Materials 11**. To derive the expected value, variance, and CV for $U \sim \text{logN}(\theta, \omega)$, the following proposition is used, for which a proof is provided in **Supplemental Materials 11**:

$$\mathbb{E}[U^k] = \exp\left(k\theta + \frac{1}{2}k^2\omega^2\right) \text{ for any } k = 1, 2, \dots \quad (10)$$

Hence, the expected value is obtained with $k = 1$

$$\mathbb{E}[U] := \int_0^{\infty} u f_U(u) du = \exp\left(\theta + \frac{1}{2}\omega^2\right), \quad (11)$$

and the variance using $k = 2$ and $k = 1$ (squared).

$$\begin{aligned} \text{Var}[U] &:= \mathbb{E}[U^2] - (\mathbb{E}[U])^2 \\ &= \exp\left(2\theta + \frac{1}{2}2^2\omega^2\right) - \left(\exp\left(\theta + \frac{1}{2}\omega^2\right)\right)^2, \end{aligned} \quad (12)$$

which can be rewritten as $\exp(2\theta + \omega^2)(\exp(\omega^2) - 1)$. The CV is trivially implied:

$$\text{CV}[U] = \frac{\sqrt{\text{Var}[U]}}{\mathbb{E}[U]} = \sqrt{\exp(\omega^2) - 1}. \quad (13)$$

Finally, as established in **Supplemental Materials 11**, the median of $U \sim \text{logN}(\theta, \omega)$ is located at $\exp(\theta)$ and the mode at $\exp(\theta - \omega^2)$. Observe that the mean of U is given by $\exp(\theta + \frac{1}{2}\omega^2)$; a property clearly different from the Gaussian “mean = median = mode” characteristic. This is also reflected in the skewness of U given by $(\exp(\omega^2) + 2)\sqrt{\exp(\omega^2) - 1}$.

Overview. Table 1 summarizes the differences between the normal and lognormal distributions established so far. The variables X or U can be replaced by the name of the parameter of interest to the pharmacometrician, for example, clearance (CL), i.e., $CL \sim \text{logN}(\theta, \omega)$.

In conclusion, several properties of the lognormal distribution can be straightforwardly derived from first statistical principles. The take-home messages are that the mean of the lognormal distribution, in contrast to the normal, is different from the median and that it is related to both θ and ω . The CV is only a function of ω . Lastly, the parameter value with largest density, i.e., the mode, is different from the mean and the median.

Graphical exploration of the lognormal distribution

To augment the formal definition of normal and lognormal parameters, a graphical exploration is presented. A normally distributed parameter with $N(\mu, \sigma)$ is shown in **Figure S1** with $\mu = 8$ and $\sigma = 4$, as its probability density and its cumulative density, overlaid with a histogram of 1000 draws from the normal distribution. Remark that for any mean and (positive) SD, the normal distribution has at least some mass below zero, in other words, the parameter can attain negative values. The figure also exemplifies that the mean lies at the peak of the distribution and that the cumulative distribution is exactly 50% at that point, in other words: The mean, mode, and median are equal.

Table 1 Main properties of the normal and lognormal distribution

	Normal $X \sim N(\mu, \sigma)$	Lognormal $U \sim \text{logN}(\theta, \omega)$
Support	$(-\infty, \infty)$	$(0, \infty)$
Mean	μ	$\exp\left(\theta + \frac{1}{2}\omega^2\right)$
SD	σ	$\exp\left(\theta + \frac{1}{2}\omega^2\right) \sqrt{\exp(\omega^2) - 1}$
CV	σ/μ	$\sqrt{\exp(\omega^2) - 1}$
Median	μ	$\exp(\theta)$
Mode	μ	$\exp(\theta - \omega^2)$
Skewness	0	$(\exp(\omega^2) + 2) \sqrt{\exp(\omega^2) - 1}$

CV, coefficient of variation; SD, standard deviation.

An example of the lognormal distribution is shown in **Figure S2** with $\theta = \log(8)$ and $\omega = 0.5$ together with the normal distribution from **Figure S1**. The lognormal ω was set equal to the CV of the normal distribution ($4/8$). The example visualizes that the lognormal distribution cannot assume negative values in contrast to the normal distribution. Other differences are that the point of maximum density not coincides with its mean and that there is a heavier tail in its density toward higher values compared with the normal distribution. In other words, the cumulative densities of the lognormal distribution occur at higher values. The difference between the normal and the lognormal distribution increases with increasing variability; see **Table 1**. It is notable that the mean of the lognormal distribution now lies higher than 50% cumulative probability; in other words, the mean has become larger than the median (the mean of log-transformed values is equal to the log of the median); see Eq. 11.

The extent of the difference between the normal and lognormal distribution is further evaluated in **Figure 1**. The 10th and 90th percentiles are plotted as a function of σ for the normal and as a function of ω for the lognormal distribution, where the same values for σ and ω are used. The plot therefore reflects what would happen if one would mistakenly

assume the normal and the lognormal distribution behave similar. The percentiles of the normal distribution relate linear to σ . Although at low values the 10th and 90th percentiles of lognormal are close to those of the normal distribution, at mildly higher values they start to deviate. The increasing skewness of the lognormal distribution with increasing ω is clearly visible, as the 10th percentile shrinks toward zero and the 90th percentile increases exponentially.

These graphs demonstrate that the behavior of the lognormal distribution is not easy to capture in a single number. At the backtransformed scale, the median and the mean become further separated as a function of ω ; see **Figure 1**. This can have counterintuitive effects, as the basic expectation is the mean to be located at the center of the distribution. The mean of the lognormal distribution also is not the value with the highest probability density. In addition, the distribution becomes more and more asymmetric, as shown in the left panel of **Figure 1**. These properties all are governed by one parameter, ω , but are not easily and transparently derived from it. For example, in the right panel of the same figure it can be observed how different the mean and the mode become as a function of ω . The increasing skewness of the lognormal distribution is not always presented and

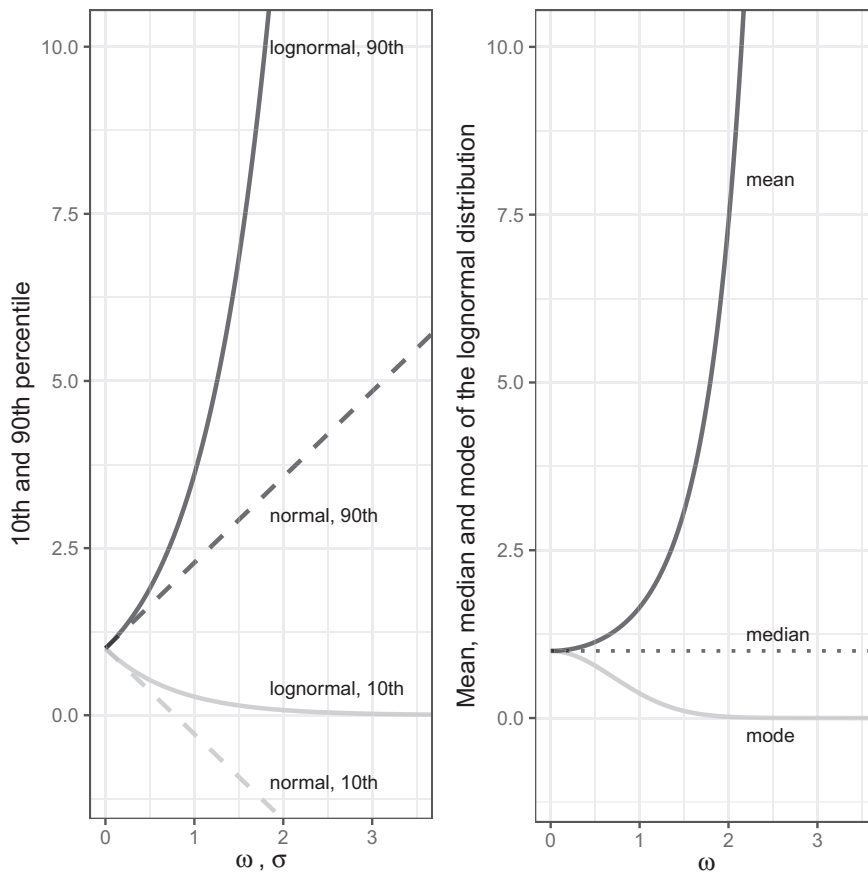


Figure 1 Differences between the lognormal distribution relative to the normal. The 10th and 90th percentiles of the normal (dashed, $\mu = 1, \sigma$ varies) and the lognormal (straight, $\theta = 0$, i.e., the same median value, $\omega = \sigma$), as a function of σ (left panel). The mean (dark straight), median (dotted), and mode (light straight) plotted against ω where θ is set to zero; at zero the curve therefore starts with one (right panel). Note how quickly the mean runs off the scale and overtakes the position associated with the upper probability, as plotted in the left panel, at an ω of about two. The mode similarly decreases to a position lower than associated with the lower probability, at an ω of one.

appreciated as such. A perhaps more intuitive plot demonstrating the increase in skewness with increasing ω can be found in **Figure 2**. The skewness clearly increases more than linearly with increasing ω , and it becomes more and more difficult to summarize the distribution.

In the following part of this article, the consequences of interpretation of lognormal distributions as if they were normal will be discussed. The conclusion for now is that the shape of the lognormal distribution does clearly depend on the value of its variability.

RESULTS

Interpretation of the lognormal distribution as normal Proxy interpretation as normal and its limitations. One aspect of potential misrepresentation is the difference between mean and median. As discussed previously, the median and mean are equal for the normal but not for the lognormal distribution. The median of the lognormal distribution is $\exp(\theta)$ and therefore is independent of ω . All other characteristics of the lognormal distribution, however, are dependent on ω , including the mean and also the mode. A first deviation when interpreting the lognormal distribution as normal therefore relates to the difference between the median, mean, and mode, as is discussed further in section “Statistics to Describe the Lognormal Distribution”.

A second risk that is inherent in the interpretation of a lognormal distribution as if it was normal is that the tails of the distribution are misjudged. Frequently, the end result of a pharmacometric evaluation regards the tails of the distribution, e.g., to determine whether exposure in a subpopulation corresponds with that of the general population for at least 90% of the subpopulation. Therefore, it is important to establish how closely the lognormal distribution corresponds with a normal distribution in its tails.

Suppose one has estimated θ and ω^2 (estimates denoted as $\hat{\theta}$ respectively $\hat{\omega}^2$ here) based on gathered data about the random variable $CL \sim \text{logN}(\theta, \omega)$, for instance, in NONMEM with declaration $CL = \text{EXP}(\text{THETA}(1) + \text{ETA}(1))$. Model results for CL could be: $\hat{\theta} = 0.45$; $\hat{\omega}^2 = 0.45$, and $\hat{CV} = 75\%$. A pharmacokinetic parameter was chosen as a relevant example, but the same principles hold for any parameter or response variable such as concentration.

Table 2 lists the implied mean, SD, mode, median, and two percentiles for CL in the population for several model interpretations treated next, the top line (A) being the correct one. In scenario B, the parameters are mistaken for those of the normal distribution. Scenario C is a rule of thumb that interprets $\hat{\theta}$ as (Gaussian) mean. Scenarios D to F apply a formal approximation of $\hat{\theta}$ assuming, respectively, the mean, mode, and median are equal to $\hat{\theta}$.

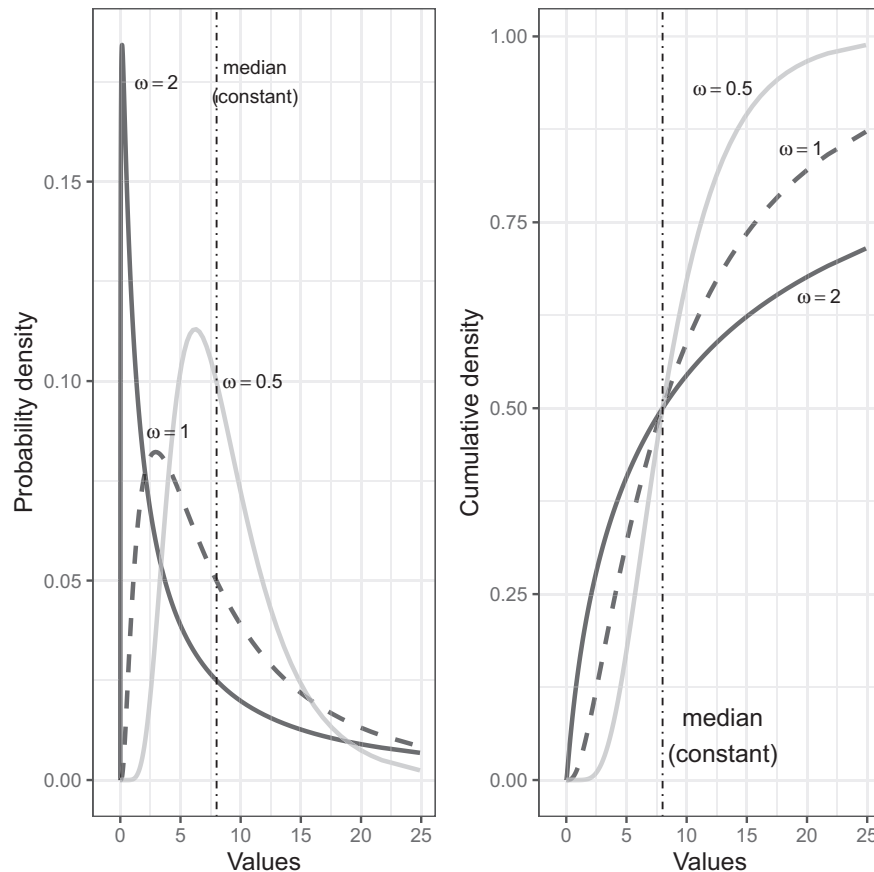


Figure 2 Probability density (left) and cumulative density (right) of the lognormal distribution at different values of ω and fixed median. The probabilities according to the lognormal distribution at ω values of 0.5 (gray), 1 (black stripes), and 2 (black) is plotted against parameter values between 0 and 25; θ was set to $\log(8)$.

Table 2 Mean; SD; mode; 10%, 50% (median), 90% percentiles; and associated RE% for several candidate model interpretations for $\hat{\theta} = \log(8.0)$ and $\hat{\omega} = 0.67$

Interpretation	Comment	Mean	SD	Mode	Percentile			RE(%) percentile			
					10	50	90	10	50	90	
A	$\log N(\hat{\theta}, \hat{\omega})$	Best case	10.0	7.5	5.1	3.4	8.0	18.9			
B	$N(\hat{\theta}, \hat{\omega})$	Worst case	2.1	0.67	2.1	1.2	2.1	2.9	-65	-74	-84
C	$N(\exp(\hat{\theta}), \exp(\hat{\theta}) \cdot \hat{CV})$	Rule of thumb	8.0	6.0	8.0	-4.0	8.0	20.0	-220	0	6
D	$N(\hat{\mu}, \hat{\sigma})$	$\hat{\mu}$ = actual mean	10.0	7.5	10.0	0.35	10.0	19.7	-90	25	4
E	$N(\hat{\mu}, \hat{\sigma})$	$\hat{\mu}$ = actual mode	5.1	9.0	5.1	-6.4	5.1	16.6	-290	-36	-12
F	$N(\hat{\mu}, \hat{\sigma})$	$\hat{\mu}$ = actual median	8.0	7.8	8.0	-2.0	8.0	18.0	-160	0	5

The normal distribution parameter $\hat{\sigma}$ is approximated by Kullback–Leibler divergence minimization of $N(\hat{\mu}, \hat{\sigma})$ with regard to $\log N(\hat{\theta}, \hat{\omega})$. RE%, relative errors; SD, standard deviation.

In the best case (A), the model estimates $\hat{\theta}$ and $\hat{\omega}$ are used to infer the mean, SD, mode, and percentile statistics from the corresponding lognormal distribution. This may sound trivial, but it is not uncommon to encounter one of the following five alternative interpretations (B–F) based on the Gaussian distribution in practice instead. A worst case scenario (B) would result from interpreting the estimators directly as if they specify a normal distribution $CL \sim N(\hat{\theta}, \hat{\omega})$, leading to a substantial underrepresentation of effect size and variability (80% PI [1.2–2.9]). Furthermore, the authors have seen a form of the following rule of thumb (C) being informally applied. Using $\hat{\theta}$ and \hat{CV} , approximate the mean by $\exp(\hat{\theta})$ and consequently the SD by $\exp(\hat{\theta}) \cdot \hat{CV}$. Then, based on characteristics of the Gaussian, suppose the mode and median are similar to the mean and construct 10% and 90% percentiles symmetrically based on $\exp(\hat{\theta}) \pm 2\hat{SD}$, which is a strong inflation with respect to the normal 80% PI that uses factor 1.28 instead of 2. In case this inflation causes the 10% percentile to become negative (which is impossible for a lognormal random variable), it is replaced by 0 (or considered “small”).

Realistically, interpretations D, E, and F try to mimic the lognormal distribution with a Gaussian by matching the mean, mode, or median of the underlying lognormal, respectively. By fixing the normal mean $\hat{\mu}$ in this way, focus is placed on what value of $\hat{\sigma}$ does the best job in approximating the targeted $\log N(\hat{\theta}, \hat{\omega})$ distribution. Despite $\hat{\mu}$ having improved, it may be clear that $\hat{\sigma} = \hat{\omega}$ is completely on the wrong scale as before. With the goal of approximating the $\log N(\hat{\theta}, \hat{\omega})$ density with the $N(\hat{\mu}, \hat{\sigma})$ in mind, we turn to the widely adopted KL divergence¹² as pseudometric for approximation error, see the Proofs section, which defines different $\hat{\sigma}$ s in Eq. 20. Examining **Table 2**, the mean-match Gaussian (D) seems to be closest to the benchmark (A) in terms of the 10% and 90% percentiles, although the 80% PI [0.35, 19.7] is too wide (compare [3.4, 18.9]) and the median too high (10.0 vs. 8.0). Despite the unfeasible estimate for the lower 10% percentile (–4.0), the rule of thumb (C) does not do much worse than the KL mean-match (D) in terms of the 90% percentile.

How large the deviations at the tails of the approximation are also depends on the magnitude of variability. At very small values of ω , the difference will be negligible because the shape of the distributions is similar, but at larger values

the locations will diverge exponentially; see **Figure 3** (In the Proofs section it is shown that the ratio of any percentile in KL optimized normal approximation against the true log-normal percentiles is only dependent on $\hat{\omega}$, i.e., not on $\hat{\theta}$.) The tails of the distribution are specifically important when simulations are used to generate PIs and the tolerance for deviating locations is not very high. The relative error in the lower 10% tail of the approximation gets to 10% at an ω of about 0.25. The upper tail hits the 10% error level at higher values because the approximated SDs are large compared with the lognormal SDs, analogs to the difference between the correct equation for CV compared with the first-order Taylor approximation as also described in the next paragraph, to be able to cover the bulk of the density. The rule of thumb approximates the 10% tail worse, but the 90% tail the better among all approximations, and performs reasonably well up to an ω of about 1.1. The results at the 10% tail nevertheless lead to the conclusion that above an ω of ≈ 0.25 , the lognormal distribution cannot be interpreted as a normal distribution even when using KL-optimal estimates. The results remain similar when the 95% PI is chosen instead of the 80% PI; see **Figure S3**.

STATISTICS TO DESCRIBE THE LOGNORMAL DISTRIBUTION

CV

The usual way of representing a distribution in pharmacometric reports and papers is the mean and CV of the distribution. The CV is an adequate measure to summarize a normal distribution as it is clear how far values extend below and above the mean and what the probability is of finding negative values. The CV is, however, more difficult to interpret for the lognormal distribution with larger values of ω . Optimally, one would convert the CV back into the lognormal SD ω and use that for further interpretation. In the pharmacometric literature, some scrutiny needs to be applied before doing that because often the CV is approximated by its first-order Taylor expansion around zero, $\sqrt{\omega^2}$ instead of the correct $\sqrt{\exp(\omega^2) - 1}$ (see also ref. 3). At higher values, for example, 2 for ω , the traditionally reported CV would be 200% instead of 732%. Regardless of this confusion, the CV remains difficult to interpret beyond ω values that allow for sufficient precision in the interpretation of the lognormal as a normal distribution as explored previously.

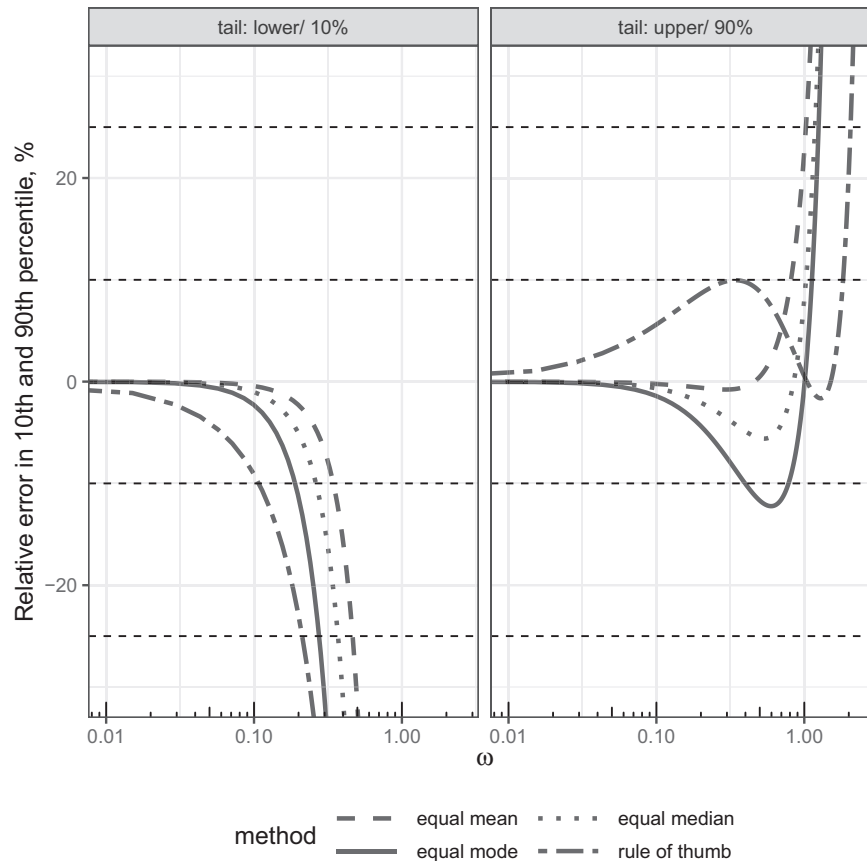


Figure 3 The percentage error of the location of the lower (left panel) and upper (right panel) 10% tail of the Kullback–Leibler optimally approximated normal distribution representing a given lognormal distribution or as assessed by rule of thumb, see text for an explanation. The relative error in the location of the approximate normal tail compared with the given true, i.e., the tail of the lognormal distribution, is plotted as percentage against the lognormal standard deviation ω . The Kullback–Leibler optimal approximation was calculated for an assumed equal mean, median, or mode. The horizontal lines indicate the 10% and 25% error levels in either direction.

Skewness

The traditional statistic to describe asymmetry of unimodal distributions, such as the normal and lognormal, is the third standardized moment known as skewness. It was previously noted that the symmetry of any normal distribution implies zero skew, and any lognormal distribution has positive skew, which increases exponentially in ω . Although caution in interpreting the skewness statistic is generally warranted,¹³ the positive skew indicates a long tail to the right of the lognormal density here. In other words, more than half of the mass of the distribution is located to the left of the mean, i.e., the mean is larger than the median, and the 90th percentile will be further removed from the median than the 10th percentile, again in this specific case. This effect becomes stronger for larger values of skewness (i.e., ω), such that nonzero skew is a clear warning signal that the modeled lognormal distribution is not approximated well by any normal distribution. However, it is argued here that it is easier to directly assess relative difference against the percentiles of a matched normal distribution (see **Figure 3**) or the relative difference between the mean and median in the lognormal distribution described next because these indicators are more closely related to typically reported results.

Percentiles of the distribution

A straightforward and intuitive way to characterize the distribution is by percentiles of interest, for example, the 10th and the 90th percentiles. These numbers directly provide information about properties of interest, the coverage of the population. One can normalize the percentiles to the median because percentiles of the lognormal distribution, relative to the median, only depend on ω ; see the previous section and the Proofs section. The normalized percentiles, and especially the 90th percentile, form a useful indicator of how the lognormal distribution extends into extreme values; relative comparisons to the normal distribution with the 10th percentile are somewhat more difficult to interpret because of the negative values realized with the normal distribution at higher variances. How much different the 90th percentile is from the normal distribution is still complicated to assess because the relative percentiles are a function of variability in both distributions. The difference of the KL-optimally mean-match percentiles, see **Figure 3**, would provide this information in a scale-invariant way. These values are, however, rather complex to calculate and therefore less practical for normal use.

Mean-to-median ratio

Let us therefore find additional statistics to characterize the lognormal distribution. We have clearly established that a key property of the lognormal distribution is the difference between the mean and median. Depending on the application, the tolerance for interpretation as a normal distribution with the mean and median being equal will vary. Rearrangement of Eq. 11 results in a concise equation to calculate the lognormal ω associated with a chosen tolerance for the mean and median being equal:

$$\omega_{tol} = \sqrt{2 \log(1 + tol/100)}, \quad (14)$$

with *tol* expressed in %. Using this expression, a somewhat liberal tolerance of 25%, a number that is often used as acceptable standard error for estimated PKPD parameters, results in a cutoff ω of 0.67. In other words, above this cutoff, the mean is more than 25% higher than the median. The next step is to restate the ratio of mean over median as a statistic to characterize the distribution:¹⁴

$$\text{MMR} := \frac{\text{mean}}{\text{median}} = \frac{\exp\left(\theta + \frac{1}{2}\omega^2\right)}{\exp(\theta)} = \exp\left(\frac{1}{2}\omega^2\right). \quad (15)$$

This parameter is easy to understand, as the general idea of what mean and median are is common knowledge. It is straightforward to interpret this parameter as an indication of the asymmetry of the distribution and as guidance for how accurate a proxy interpretation as normal could be for assessing the mean. The area between the median and the mean lends itself to further interpretation, at moderate ω values, as an indication for how far comparable probabilities extend into the distribution. A higher ratio of mean over median therefore can be quite intuitively used as an indicator for how far the distribution is extended into high numbers. Within a report, space is always limited and a suggested abbreviation of this ratio is the mean-to-median ratio (MMR). We conclude here that MMR can be used to indicate (i) asymmetry of the lognormal distribution and (ii) its spread into higher sample values.

Mode density inflation

The point where the distribution has its highest probability density, the mode, is equal to the mean for the normal Gaussian distribution. The mode of the lognormal distribution however is lower than the mean and the median, see the Theoretical section. The consequence is that at higher ω , the most likely discrete value the lognormal distribution can take becomes lower and lower. At an ω of ≈ 0.83 the mode lies at half of the median, meaning the lognormal is not only skewed, but also has its peak far away from the median. This is further illustrated in **Figure 2** (left) by a sharpening of the peak at high lognormal variability. At $\omega=2$, the peak in density occurs at values close to zero while the median is at 8. Interpretation of such distributions includes acknowledging that the probability of values between zero and its maximum changes with a very steep slope at the normal scale. Likewise, a disproportional fraction of values will be realized at low values. For $\omega=2$, a quarter of all

probability mass is located below 25% of the median, as can be read out from **Figure 2** (right).

The impact of ω on how normal the distribution appears can be read out from **Figure 4**. In the left panel the peak density of the distribution, i.e. the density at the mode or mode density, is depicted as a function of ω . Initially, the peak density declines similarly to the normal distribution, but starts to flatten out around ω values of 0.4–0.67 and reaches its minimum at 1. Beyond this value, peak densities start to increase slowly again and become 10% larger than the minimum at about 1.3. The peak density increases more than linearly with higher values and the lognormal distribution unequivocally no more resembles a normal-like distribution. Values of ω that are beyond the point where the lognormal density starts to sharpen, i.e. 1, are therefore unequivocally misrepresented by any normal distribution.

Therefore, another property of high lognormal variability that we need to cover is the high density peak close to zero. A second additional statistic should therefore speak to the increased sharpness of the peak. To emphasize its dependence on statistical parameters θ and ω , we denote the $\log N(\theta, \omega)$ density at value u by $f_{\theta, \omega}(u)$ and its mode density as $\max_{u>0} f_{\theta, \omega}(u)$. Then, the Mode Density Inflation, which is a function of ω , can be defined as follows.

$$\text{MDI} := \frac{\max_{u>0} f_{\theta, \omega}(u)}{\min_{\omega>0} \left(\max_{u>0} f_{\theta, \omega}(u) \right)} = \frac{1}{\omega} \exp\left(\frac{1}{2}(\omega^2 - 1)\right) \quad (16)$$

The final equation follows since one can show that

$$\min_{\omega>0} \left(\max_{u>0} f_{\theta, \omega}(u) \right) = \max_{u>0} f_{\theta, 1}(u). \quad (17)$$

The mode Density Inflation clarifies that the lognormal distribution gets a sharp peak at higher variabilities. The material in for example **Figure 4** demonstrates that the sharpened peak always occurs close to zero and it is therefore not necessary to provide an additional number to indicate the position of the mode. The proposed abbreviation of the ratio would be mode density inflation (MDI). The MDI is large at lower values of ω , where the lognormal distribution is very similar to normal and the peak is sharp. It is almost similar to the density of the normal distribution to an ω of about 0.5, after which it slowly approaches a nadir of 1 at an ω of 1. The point where the the peak density gets 10% above the minimum is the end of the gray zone. Beyond this point the lognormal distribution is unequivocally different from a normal distribution, with a mode close to zero, a mean that is more than twice the median, and a sharpness in probability at the mode that is indicated by the MDI. An illustration of the lognormal distribution at a number of ω values with respective CV, MMR, MDI, relative 90th percentile, and skewness values can be found in **Figure 5**.

Representing the center of the lognormal distribution

The median is the preferable statistic to reflect the center of the lognormal distribution, in contrast to the normal distribution, for which the mean is equal to the median and the mode, such that the center of the normal distribution can also be expressed by the mean or mode. With the

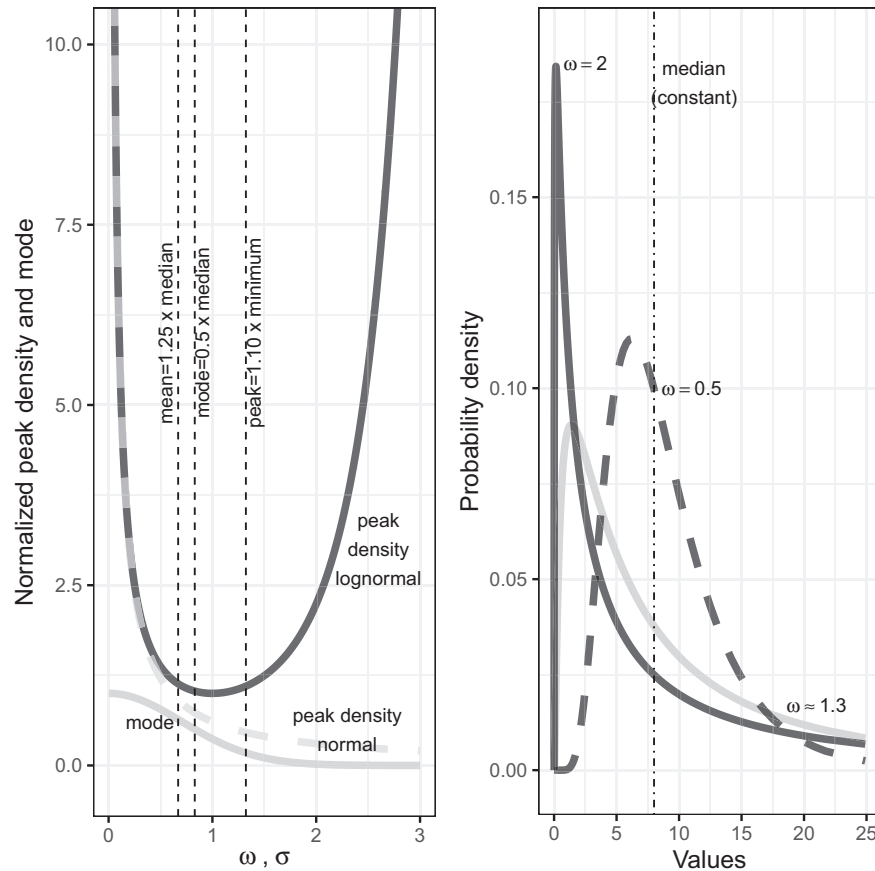


Figure 4 The lognormal distribution increases in sharpness after reaching a minimum at an ω of about one. Combined plot of the peak density value (solid) of the lognormal distribution, normalized to its lowest value, and of the mode of the lognormal distribution (dark gray) as a function of the lognormal standard deviation ω (left panel). The peak density of the normal distribution as a function of σ (dashed gray) has been added as reference. Vertical lines highlight some ω values of special interest; from left to right, the values where the mean is 25% higher than the median, where the mode is half of the median, and where the peak density becomes 10% larger than the minimum density. Selected examples of the lognormal probability density at several values of ω ; θ was set to $\log(8)$ (right panel).

lognormal distribution these three statistics are different and their difference furthermore increases with increasing variability, see also **Figure 1**. The mode shrinks toward zero and the mean increases more than exponentially (exponentially with ω^2). Given large enough ω values, the mean will be even higher than most percentiles because percentiles above 50% increase exponentially with ω ; for example, the mean of the lognormal distribution is above its 90th percentile for ω larger than ≈ 2.56 . At such values, the mode and the mean therefore represent extreme values of the lognormal distribution. The median on the other hand is at the center of the lognormal distribution regardless of ω value.

DISCUSSION

We have explored the properties of the lognormal distribution in mathematical and graphical detail, and the limitations one encounters when interpreting it as if it were a normal distribution. These limitations have been established using formal and therefore objective approximation methods.

The lognormal distribution can be considered to have normal-like properties up to a value of $\omega \approx 0.25$, the point at

which a normal approximation would lead to more than 10% error in the location of its lower tail. A gray area follows, with a distribution peak that is still somewhat normal-like, while increasingly eccentric mode and mean values are realized. At an ω of about ≈ 1.1 , the gray area ends where the best normal approximation would lead to more than 10% error in the location of its upper tail.

The previous interpretation can be considered reasonable but somewhat strict. Depending on the context and application, e.g., in the context of small efficacy–safety windows, a more stringent set of cutoff values could be selected, 0.12 and 0.83 as the border where the error in the tail would be 5% and the mode is found at 50% of the median, respectively. Or in case modeled distributions are used in a descriptive fashion, a more liberal lower cutoff value could be selected with ≈ 0.67 , the point at which the mean gets 25% larger than the median. A more liberal alternative for the higher cutoff could be an ω of ≈ 1.3 , where the distribution sharpens by more than 10% and the mode and mean become too eccentric. The interpretation of a distribution below the lower cutoff is interpreted as almost identical to that of a normal distribution, whereas above the higher cutoff the distribution is considered as a completely different

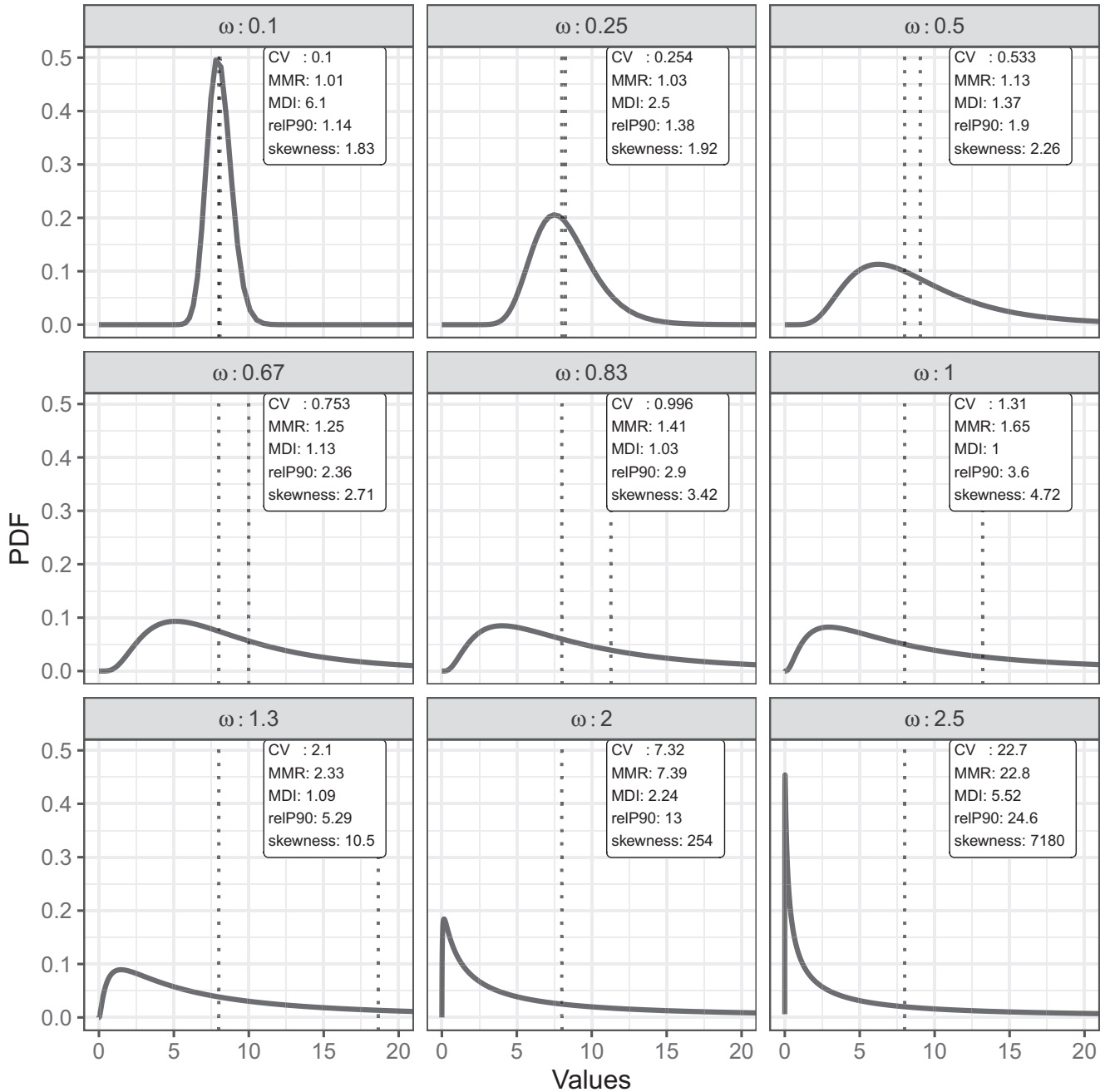


Figure 5 Probability densities at ω values of interest computed at $\theta = \log(8)$. The inserts tabulate values of CV, MMR, MDI, the relP90, and skewness. Vertical dotted lines indicate the position of the median (left) and the mean (right); from an ω value of two, onward the position of the mean falls outside the horizontal axis range. CV, coefficient of variation; MDI, mode density inflation; MMR, mean-to-median ratio; relP90, 90th percentile relative to the median.

distribution, with diverging mean and medians, and a peak in density that is close to zero.

Cutoff values and regions of ω values to guide interpretation

Three scenarios can be discerned when reporting on lognormal distributions; see **Table 3** and **Supplemental Material 12**. In case the ω is small, the results can be interpreted comparable to normal distributions, focusing on ω and,

certainly if normal distributions are presented alongside, %CV. Furthermore, instead of the mean, the median would be reported as appropriate. From ω values of about 0.25–0.67, the lognormal starts to deviate from the normal distribution, and the mean-to-mode ratio or MMR starts to deviate from one. The MMR can function as a warning of asymmetry and as an indication that the mean explicitly cannot be assumed to be similar to the median. It also starts to indicate the extent of spread of the distribution into higher values. With these ω

Table 3 Suggestions for reporting lognormal distributions at different levels of variability

Range ω	ω^2	MMR	MDI	Shape of distribution	Suggestion
$0 < \omega < 0.25$	0.0625	1.00 to 1.03	∞ 1 to 2.5	"Close" to normal	-
$0.25 < \omega < 1.10$	1.2	1.03 to 1.83	2.5 to ≈ 1	Heavy tail and shallow peak	Consider to discuss
$1.10 < \omega < \infty$	-	1.83 to ∞	1.01 to ∞	Elongated heavy tail, sharp peak near 0	Discuss shape

MDI, mode density inflation; MMR, mean-to-median ratio.

values, the tails of the distribution can no longer be recovered using an interpretation as normal. Beyond an ω of 1.1, the distributional shape gets very different, and it is suggested to include the third parameter, the mode-density ratio or MDI that serves to illustrate the increased peak sharpness at values close to zero. At this scenario, the text of the report would include some discussion of additional support for the distributional shape. The tails of the lognormal have lost any similarity to those of the normal distribution at these high ω values.

In reports that discuss population coverages, for example, 80% or 90% PIs, under such high ω it is especially advisable to carefully convey to how different the tails of the lognormal distribution are. For example, the sensitivity of coverages to parameter uncertainty or misspecification could become quite different for lower and upper intervals. Explicit discussion of such properties where relevant may help prevent unrealistic expectations among the audience.

Notwithstanding the preceding discussion, the mathematically best presentation of a lognormal distribution is indeed strictly as a lognormal distribution, with θ and ω presented in their untransformed values. Such a presentation, however, might receive a less-than-favorable reception by a wider target audience and therefore cannot be deemed optimal. We hope that our findings and discussion of possible statistics can help in interpretation.

Confirmation of lognormal distributions with high variability

Fleming *et al.*² investigated the distribution of potency values of acetylcholine. Their paper clearly showed highly skewed distributions of observed potencies with high counts close to zero, consistent with the skewed and asymmetric properties of high-variability lognormal distribution, i.e., a high MDI. It is, however, not always clear that such a curvature is indeed needed. Therefore it is recommended to perform additional investigations to confirm the modeled skewness if a lognormal ω falls beyond the gray zone. The actual indicators could vary dependent on the amount of data available, the background knowledge, and the available software. Three types of confirmation could be sought: (i) typical run completion checks, foremost whether the standard error of ω is not overly large; (ii) *post hoc* checks, such as whether the *post hoc*s show indeed—at the normal scale—a high probability at low values consistent with the MDI and whether the mode of the distribution does indeed occur frequently; and (iii) simulation-based diagnostics such as mirror plots and visual predictive checks. When the number of observations/individuals is small, it is easy to see that these diagnostics perform poorly and research of additional (external) support for large ω values would be advisable. In case the results of the investigations

decrease the support for the large ω lognormal distribution, alternative parameterizations such as mixture models or semiparametric conversions could be explored.

PROOFS

Analytical expression for KL divergence of the normal distribution against the lognormal distribution

The KL divergence¹² between two continuous distributions P and Q with densities $p(u)$ respectively $q(u)$ is defined as

$$KL(P||Q) := \int_{-\infty}^{\infty} \log \left(\frac{p(u)}{q(u)} \right) p(u) du. \quad (18)$$

Because it is our goal to evaluate how different a given lognormal distribution is from (the best-fitting) normal distribution, we will denote $logN(\theta, \omega)$ by P and the $N(\mu, \sigma)$ distribution by Q . The following may be considered a technicality. Because $p(u)$ is only supported on $(0, \infty)$ we have to assume that $p(u) \log(p(u)) = 0$ for $p(u) = 0$, i.e., for $u \leq 0$. Hence, the difference between the lognormal and the normal is merely considered on $(0, \infty)$ in this setting. Strictly speaking, that is not how the KL divergence is intended because Q is formally no longer a probability measure on this restricted space $(0, \infty)$, but it serves its purpose here. Finally, one may observe that the reverse definition $KL(Q||P)$ is ill defined ($+\infty$), which is why we speak of "divergence" rather than "distance."

As a first step toward optimizing $\hat{\sigma}$ by KL minimization (on $(0, \infty)$) for fixed θ , ω , and μ , we derive an analytical expression for the divergence.

$$KL(P||Q) = -\frac{1}{2} - \theta + \log \left(\frac{\sigma}{\omega} \right) + \frac{1}{2\sigma^2} \left(\exp \left(2\theta + 2\omega^2 \right) - 2\mu \exp \left(\theta + \frac{1}{2}\omega^2 \right) + \mu^2 \right) \quad (19)$$

Proof.

$$KL(P||Q) = \int_0^{\infty} p(u) \log \frac{p(u)}{q(u)} du = I + II + III + IV$$

Substituting p and q gives the following four expressions after rewriting.

Part I.

$$I = -\frac{1}{2} \int_0^{\infty} \frac{1}{u\omega\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{\log(u) - \theta}{\omega} \right)^2 \right) \left(\frac{\log(u) - \theta}{\omega} \right)^2 du$$

Substitute $(\log(u) - \theta)/\omega = s$, $u = \exp(\theta + \omega s)$, $du = \omega \exp(\theta + \omega s) ds$.

$$I = -\frac{1}{2} \int_{-\infty}^{\infty} s^2 \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2}s^2 \right) ds = -\frac{1}{2}$$

Here, the last equality follows from a standard Gaussian integral (which one may recognize as the variance of a standard normal random variable).

Part II.

$$II = - \int_0^{\infty} \frac{1}{u\omega\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\log(u)-\theta}{\omega}\right)^2\right) \log(u) du$$

Substitute $\log(u) = s$, $s = \exp(u)$, $du = \exp(u)ds$.

$$II = - \int_{-\infty}^{\infty} s \frac{1}{\sqrt{2\pi}\omega} \exp\left(-\frac{1}{2}\left(\frac{s-\theta}{\omega}\right)^2\right) ds = -\theta$$

Here, the last equality follows by recognizing the expression as expected value of a Gaussian random variable with mean θ (and variance ω^2).

Part III.

$$III = \int_0^{\infty} \frac{1}{u\omega\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\log(u)-\theta}{\omega}\right)^2\right) \log\left(\frac{\sigma}{\omega}\right) du = \log\left(\frac{\sigma}{\omega}\right)$$

The result follows immediately since the integral of the lognormal PDF $\int_0^{\infty} p(u)du$ over its entire support equals 1.

Part IV.

$$IV = \frac{1}{2\sigma^2} \int_0^{\infty} \frac{1}{u\omega\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\log(u)-\theta}{\omega}\right)^2\right) (u-\mu)^2 du$$

Writing $(u-\mu)^2 = u^2 - 2\mu u + \mu^2$, we again recognize three terms. First, for u^2 we plug in $k=2$ in Eq. 10.

$$\int_0^{\infty} u^2 \frac{1}{u\omega\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\log(u)-\theta}{\omega}\right)^2\right) du = \exp(2\theta + 2\omega^2)$$

Second, for $-2\mu u$, we use $k=1$.

$$\begin{aligned} & -2\mu \int_0^{\infty} u \frac{1}{u\omega\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\log(u)-\theta}{\omega}\right)^2\right) du \\ & = -2\mu \exp\left(\theta + \frac{1}{2}\omega^2\right) \end{aligned}$$

Third, μ^2 is again simply a multiplication factor to the total area under the curve for the lognormal PDF (which equals one) as in part III.

$$\mu^2 \int_0^{\infty} \frac{1}{u\omega\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\log(u)-\theta}{\omega}\right)^2\right) du = \mu^2$$

Collecting all terms for part IV gives

$$IV = \frac{1}{2\sigma^2} \left(\exp(2\theta + 2\omega^2) - 2\mu \exp\left(\theta + \frac{1}{2}\omega^2\right) + \mu^2 \right)$$

Finally, adding parts I + II + III + IV completes the proof.

Optimizing $\hat{\sigma}$ of the distribution to match a given lognormal distribution

Show that $\hat{\sigma}$ is equal to one of the following three cases:

$$\begin{cases} \exp(\hat{\theta}) \sqrt{\exp(2\hat{\omega}^2) - \exp(\hat{\omega}^2)} & \text{(D)} \\ \exp(\hat{\theta}) \sqrt{\exp(2\hat{\omega}^2) - 2\exp\left(-\frac{1}{2}\hat{\omega}^2\right) + \exp(-2\hat{\omega}^2)} & \text{(E)} \\ \exp(\hat{\theta}) \sqrt{\exp(2\hat{\omega}^2) - 2\exp\left(\frac{1}{2}\hat{\omega}^2\right) + 1} & \text{(F)} \end{cases} \quad (20)$$

Proof.

Fixing θ , ω , and μ , we set the partial derivative of Eq. 19 to zero.

$$\frac{\delta KL}{\delta \sigma} = \frac{1}{\sigma} \left(1 - \frac{1}{\sigma^2} \left(\exp(2\theta + 2\omega^2) - 2\mu \exp\left(\theta + \frac{1}{2}\omega^2\right) + \mu^2 \right) \right) = 0$$

We work out the details of this equation case by case, substituting μ .

D. Set $\mu = \exp(\theta + \frac{1}{2}\omega^2)$.

$$\begin{aligned} 0 &= 1 - \frac{1}{\sigma^2} (\exp(2\theta + 2\omega^2) - \exp(2\theta + \omega^2)) \Rightarrow \\ \sigma &= \exp(\theta) \sqrt{\exp(2\omega^2) - \exp(\omega^2)} \end{aligned}$$

E. Set $\mu = \exp(\theta - \omega^2)$

$$\begin{aligned} 0 &= 1 - \frac{1}{\sigma^2} (\exp(2\theta + 2\omega^2) - 2\exp\left(2\theta - \frac{1}{2}\omega^2\right) \\ & \quad + \exp(2\theta - 2\omega^2)) \Rightarrow \\ \sigma &= \exp(\theta) \sqrt{\exp(2\omega^2) - 2\exp\left(-\frac{1}{2}\omega^2\right) + \exp(-2\omega^2)} \end{aligned}$$

F. Set $\mu = \exp(\theta)$

$$\begin{aligned} 0 &= 1 - \frac{1}{\sigma^2} (\exp(2\theta + 2\omega^2) - 2\exp\left(2\theta + \frac{1}{2}\omega^2\right) \\ & \quad + \exp(2\theta)) \Rightarrow \\ \sigma &= \exp(\theta) \sqrt{\exp(2\omega^2) - 2\exp\left(\frac{1}{2}\omega^2\right) + 1} \end{aligned}$$

Now, substituting θ and ω by their estimators $\hat{\theta}$ and $\hat{\omega}$ provides the resulting expression for $\hat{\sigma}$.

Ratio of percentiles of lognormal and KL-optimized, matched normal do not depend on $\hat{\theta}$

In **Supplemental Materials 11**, it is explained that the median of $U \sim \log N(\theta, \omega)$, i.e., $U = \exp(X)$ with $X \sim N(\theta, \omega)$, can be found as $\exp(\theta)$ (note: θ being the median of X) because the exponential is a strictly increasing function. The same holds for other percentiles than the median. Denote the $p \times 100\%$ -percentile of the standard normal distribution $Z \sim N(0, 1)$ by Z_p . Then, the $p \times 100\%$ -percentile of U is given by

$$U_p = \exp(\theta + Z_p \omega), \quad (21)$$

where we have used the well-known fact that for normal distributions, $X \sim N(\mu, \sigma)$, percentiles X_p can be stated as

$$X_p = \mu + Z_p \sigma. \quad (22)$$

Now, in **Table 2**, the best-case (A) percentiles of the lognormal distribution are given by Eq. 21, plugging in $\hat{\theta}$ and $\hat{\omega}$. One may note that $-Z_{0.10} = Z_{0.90} \approx 1.28$. On the other hand, for any of the KL-matched scenarios (D,E,F) percentiles follow Eq. 22 because of the interpretation as Gaussian plugging in $\hat{\mu}$ according to **Table 2** and corresponding $\hat{\sigma}$ from Eq. 20. The relative difference X_p/U_p is given for the respective cases D, E, and F, by:

$$\left(\frac{\exp(\hat{\theta} + \frac{1}{2}\hat{\omega}^2) + Z_p \exp(\hat{\theta}) \sqrt{\exp(2\hat{\omega}^2) - \exp(\hat{\omega}^2)}}{\exp(\hat{\theta} + Z_p \hat{\omega})} \right) \frac{\exp(\hat{\theta} - \hat{\omega}^2) + Z_p \exp(\hat{\theta}) \sqrt{\exp(2\hat{\omega}^2) - 2\exp(-\frac{1}{2}\hat{\omega}^2)} + \exp(-2\hat{\omega}^2)}{\exp(\hat{\theta} + Z_p \hat{\omega})} \\ \frac{\exp(\hat{\theta} + Z_p \exp(\hat{\theta}) \sqrt{\exp(2\hat{\omega}^2) - 2\exp(\frac{1}{2}\hat{\omega}^2)} + 1}{\exp(\hat{\theta} + Z_p \hat{\omega})}$$

and all $\exp(\hat{\theta})$ terms cancel in the denominator and nominator.

Derivation of the properties of the normal and lognormal distributions

Basic textbook results with dense proofs are included in the **Supplemental Material** to provide all that is needed to derive properties of the lognormal distribution algebraically; see **Supplemental Materials 11**.

CONCLUSIONS

The application of the lognormal distribution in pharmacometrics, its evolution, and strong rationale have been discussed. The field of pharmacometrics has widened, and with that also larger variability in the lognormal domain are encountered. The characteristics of the lognormal distribution change dramatically with increasing variability values (ω). The distribution initially becomes skewed. With further increase of variability the center mass of the distribution shifts, with increasing mean and decreasing mode relative to the median. At even higher values the lognormal distribution gets extremely sharp with its mode close to zero and its mean positioned at multitudes of the median. Some additional statistics to evaluate these characteristics are described, the MMR to describe asymmetry and spread, and the MDI to describe peak density sharpening. We investigated the exact consequences of interpreting the lognormal distribution as normal by defining optimal approximations of the lognormal distributions and determining the highest variabilities at which the approximation was still valid in describing the 10th and 90th percentiles. The 10th percentile could be approximated with a normal distribution up to ω values of up to about 0.25, whereas the 90th percentile

approximation remained valid up to about an ω of 1.1, with an error of 25%. Above this level of variability, the lognormal distribution does not resemble a normal distribution anymore, and other statistics may be helpful in the reporting and discussion of lognormal distributions at high variability values.

Supporting Information. Supplementary information accompanies this paper on the *CPT: Pharmacometrics & Systems Pharmacology* website (www.psp-journal.com).

Acknowledgments. The authors thank Coen van Hasselt and Erik Olofsen for their review of an earlier version of the article.

Funding. No funding was received for this work.

Conflict of Interest. The authors declared no competing interests for this work.

Author Contributions. J.E.-S. and K.D. wrote the manuscript. J.E.-S. and K.D. designed the research. J.E.-S. and K.D. performed the research.

- Williams, P.J. & Ette, E.I. Pharmacometrics: Impacting drug development and pharmacotherapy. *Pharmacomet. Sci. Quantit. Pharmacol.* 1–21 (2007).
- Fleming, W.W., Westfall, D.P., De La Lande, I.S. & Jellett, L.B. Log-normal distribution of equieffective doses of norepinephrine and acetylcholine in several tissues. *J. Pharmacol. Exp. Therap.* **181**, 339–345 (1972).
- Julious, S.A. & DeBarnot, C.A.M. Why are pharmacokinetic data summarized by arithmetic means? *J. Biopharm. Stat.* **10**, 55–71 (2000).
- Lacey, L.F., Keene, O.N., Pritchard, J.F. & Bye, A. Common noncompartmental pharmacokinetic variables: are they normally or log-normally distributed? *J. Biopharm. Stat.* **7**, 171–178 (1997).
- Gaddum, J.H. Lognormal distributions. *Nature* **156**, 463 (1945).
- Grönholm, T. & Annala, A. Natural distribution. *Math. Biosci.* **210**, 659–667 (2007).
- Koch, A.L. The logarithm in biology 1. Mechanisms generating the lognormal distribution exactly. *J. Theoret. Biol.* **12**, 276–290 (1966).
- Petersson, K.J.F., Hanze, E., Savic, R.M. & Karlsson, M.O. Semiparametric distributions with estimated shape parameters. *Pharmaceut. Res.* **26**, 2174–2185 (2009).
- Bonate, P.L. Pharmacokinetic-pharmacodynamic modeling and simulation. Springer, New York (2011).
- Lindstrom, M.J. & Bates, D.M. Nonlinear mixed effects models for repeated measures data. *Biometrics* **46**, 673–687 (1990).
- Boeckmann, A.J., Beal, S.L. & Sheiner, L.B. *NONMEM Users Guide Part VI PREDPP Guide* (NONMEM Project Group, UCSF, San Francisco, 1992).
- Kullback, S. & Leibler, R.A. On information and sufficiency. *Ann. Math. Statist.* **22**, 79–86 (1951).
- Von Hippel, P.T. Mean, median, and skew: correcting a textbook rule. *J. Stat. Educat.* **13**, (2005).
- Johnson, N.L., Kotz, S. & Balakrishnan, N. *Lognormal Distributions* 2nd edn, Vol. 1 (John Wiley & Sons, Hoboken, NJ, 1994).

© 2020 The Authors *CPT: Pharmacometrics & Systems Pharmacology* published by Wiley Periodicals, Inc. on behalf of the American Society for Clinical Pharmacology and Therapeutics. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.