

Article

Importance and Uncertainty of λ -Estimation for Box–Cox Transformations to Compute and Verify Reference Intervals in Laboratory Medicine

Frank Klawonn ^{1,2,*} , Neele Riekeberg ¹ and Georg Hoffmann ³¹ Institute for Information Engineering, Ostfalia University, 38302 Braunschweig, Germany² Biostatistics Group, Helmholtz Centre for Infection Research, 38124 Braunschweig, Germany³ Medizinischer Fachverlag Trillium GmbH, 82284 Grafrath, Germany; georg.hoffmann@trillium.de

* Correspondence: f.klawonn@ostfalia.de or frank.klawonn@helmholtz-hzi.de

Abstract: Reference intervals play an important role in medicine, for instance, for the interpretation of blood test results. They are defined as the central 95% values of a healthy population and are often stratified by sex and age. In recent years, so-called indirect methods for the computation and validation of reference intervals have gained importance. Indirect methods use all values from a laboratory, including the pathological cases, and try to identify the healthy sub-population in the mixture of values. This is only possible under certain model assumptions, i.e., that the majority of the values represent non-pathological values and that the non-pathological values follow a normal distribution after a suitable transformation, commonly a Box–Cox transformation, rendering the parameter λ of the Box–Cox transformation as a nuisance parameter for the estimation of the reference interval. Although indirect methods put high effort on the estimation of λ , they come to very different estimates for λ , even though the estimated reference intervals are quite coherent. Our theoretical considerations and Monte–Carlo simulations show that overestimating λ can lead to intolerable deviations of the reference interval estimates, whereas $\lambda = 0$ produces usually acceptable estimates. For λ close to 1, its estimate has limited influence on the estimate for the reference interval, and with reasonable sample sizes, the uncertainty for the λ -estimate remains quite high.

Keywords: reference interval; Box–Cox transformation; Monte–Carlo simulation; nuisance parameter; confidence interval



Citation: Klawonn, F.; Riekeberg, N.; Hoffmann, G. Importance and Uncertainty of λ -Estimation for Box–Cox Transformations to Compute and Verify Reference Intervals in Laboratory Medicine. *Stats* **2024**, *7*, 172–184. <https://doi.org/10.3390/stats7010011>

Academic Editor: Hari Mohan Srivastava

Received: 31 December 2023

Revised: 5 February 2024

Accepted: 7 February 2024

Published: 9 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Reference intervals are among the most important interpretation aids in laboratory medicine. By definition, they comprise the central 95 percent of a presumably healthy reference population [1] and thus indicate the probability of whether an observed laboratory result is to be classified as “normal” or “pathological”.

In the so-called direct methods for determining a reference interval, one performs a sufficient number of measurements on a presumably healthy reference collective and determines the quantiles 0.025 and 0.975 either non-parametrically or parametrically under the assumption of a certain distribution (e.g., Gaussian). More significant in medical practice are indirect estimation methods based on routinely collected laboratory results from laboratory information systems [2]. Since these always represent mixtures of normal and pathological values, the estimation of the respective quantiles is only possible under the assumption of an underlying distribution model for the sub-collective of the healthy reference population contained therein.

For practical reasons, some authors rely on only one or two different distribution models, e.g., lognormal distribution for all kinds of quantitative results [3] or a choice between Gaussian for symmetrically distributed and lognormal for right-skewed data [4]. However, in the recent literature there is a trend to assume a continuum of distributions of

different skewnesses, which can be represented by a normal distribution after the Box–Cox transformation of the original data using a lambda exponent between 0 and 1 or slightly beyond these limits [5–9].

The motivation to perform this study results from the observation that the application of different published algorithms to identical laboratory data [10] provides a very weak correlation between λ estimates (see Figure 1). Hence, the question arises whether state-of-the-art procedures are suitable at all to optimise λ values under the specific conditions of indirect reference interval estimation. Our results seem to indicate that the reliability of λ estimates is poor even in ideal theoretical distributions and close to pure guesswork in mixed populations including pathological values as one would find them in real data sets from laboratory information systems.

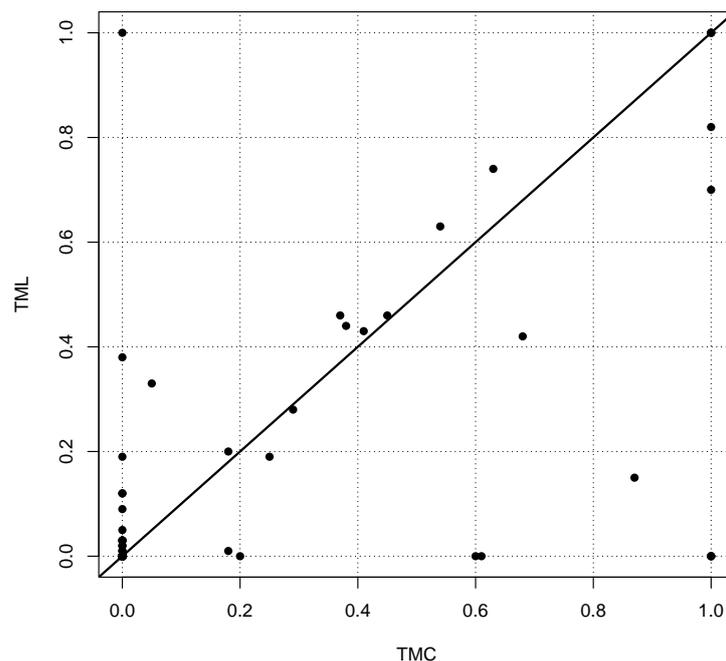


Figure 1. Estimates of the parameter λ for the Box–Cox transformation for 66 reference intervals for two different algorithms, called TMC (Truncated Minimum Chi-square, see [8]) and TML (Truncated Maximum Likelihood, see [5]). Each point represents the two λ values estimated by TMC and TML. In the ideal case, the estimates should be on the diagonal shown as a bold line. The estimated values have been taken from the supplementary Table 3, published in [10]. It should be noted that more than one-third of the values are concentrated at the origin of the coordinate system, i.e., for $\lambda = 0$ the methods show a certain coherence.

The Pearson, Spearman and Kendall correlations between the two methods compared in Figure 1 are 0.60, 0.51 and 0.46, respectively. Comparing other methods yields similar poor correlations, although the estimated reference intervals are in most cases very similar. A similar result was observed by Meyer et al. (see Figure 6 in [11]).

The aim of this paper is not to provide a new indirect method for the estimation of reference intervals but to investigate whether it is meaningful to incorporate an arbitrary Box–Cox transformation involving the estimation of its parameter λ . We do not question the assumption that the values from the healthy sub-population originate from a Box–Cox transformed (truncated) normal distribution. This assumption is common to all state-of-the-art indirect methods for the estimation of reference intervals. It might be advantageous to replace the (truncated) normal distribution by another family of distributions, e.g., the Weibull distribution, yielding only positive values, or to replace the Box–Cox transformation with a Yeo–Johnson transformation [12] that can also handle negative values. We leave such ideas to further research in the future. Our main aim is to demonstrate that the

standard indirect methods for reference interval estimation pay a high price in terms of computational costs to estimate the parameter λ of an arbitrary Box–Cox transformation without a significant benefit. The estimation of λ alone does not lead to high computational effort. But the combination of sorting out the pathological values and finding the optimal value for λ simultaneously requires iterative methods and high computational costs.

In the following, we analyse the influence and uncertainty of the nuisance parameter λ in the context of reference interval estimation based on theoretical considerations and simulations. We deliberately restrict our investigations to the ideal context without pathological values. If the estimation of λ incorporates high uncertainty and the influence of a wrong estimation of λ is also quite limited in this ideal setting, then it is obvious that estimating λ in the more complex situation including pathological values is an academic task with little improvement for the results.

Section 2 provides a formal definition of the statistical problem that we consider here. Section 3 investigates the effect of a wrong estimate for λ on the reference interval estimate for infinite sample sizes-based Monte-Carlo simulations. Finite sample sizes are considered to quantify the uncertainty involved in the estimation of λ in Section 4. In the final discussion, we review our results and conclude that an exact estimation of the nuisance parameter λ is of limited benefit for reference intervals because it is afflicted with a high uncertainty, whereas a wrong estimation does not have a strong influence on the estimate for the reference interval.

2. Problem Formulation

Indirect methods for the estimation of reference intervals in laboratory medicine aim to estimate the 2.5% and the 97.5% quantiles of the healthy sub-population. It is not known which of the values belong to the healthy sub-population. Formally speaking, we have a sample of non-negative values from a mixture of distributions and we want to estimate the 2.5% and the 97.5% quantiles of one specific distribution in this mixture. It is obvious that—without further assumptions—this is an ill-defined task. Therefore, it is usually assumed that the specific distribution of interest represents the majority of values and follows (approximately) a normal distribution after a Box–Cox transformation. The challenge is then to estimate the parameter λ and simultaneously the expected value and the standard deviation of the normal distribution, which corresponds to the subset of values from the healthy population after the Box–Cox transformation. We argue that a reliable estimation of λ is more or less impossible in a realistic setting. In order to show this, we even assume that there are no pathological values, i.e., all values originate from the healthy population.

In mathematical terms, the problem is defined as follows. A sample is generated from a normal distribution $N(\mu, \sigma^2)$ truncated at 0, representing the healthy population after the Box–Cox transformation

$$t_\lambda(x) = \begin{cases} \frac{x^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0, \\ \log(x) & \text{if } \lambda = 0. \end{cases} \quad (1)$$

with parameter λ . The truncation at 0 is necessary because the Box–Cox transformation can only be applied to positive values. It is assumed that the truncation is more or less negligible, i.e., the parameters μ and σ of the normal distribution lead to a very small probability for negative values for the normal distribution. In any case, the 2.5% quantile must be above 0. In principle, one could use the Yeo–Johnson transformation [12]. But this does not seem to be a common practice in laboratory medicine.

Let $X \sim N(\mu, \sigma^2, 0, \infty)$, i.e., X is a random variable following a normal distribution with expected value μ and variance σ^2 , which is left-truncated at 0. We observe the random variable $t_\lambda^{-1}(X)$ where t_λ^{-1} is the inverse Box–Cox transformation

$$t_\lambda^{-1}(x) = \begin{cases} (\lambda x + 1)^{\frac{1}{\lambda}} & \text{if } \lambda \neq 0, \\ e^x & \text{if } \lambda = 0. \end{cases} \tag{2}$$

The parameters λ , μ and σ are not known. Our aim is to estimate the 2.5% and the 97.5% quantiles of $t_\lambda^{-1}(X)$.

It should be emphasised again that we consider the ideal case where we observe only values from the healthy population without any pathological values, so that we do not consider the more complex case of a mixture of distributions in which the healthy sub-population is only one part of the mixture model.

The aim of this paper is to show the following in a realistic setting:

- The estimation of λ will involve very high uncertainty;
- But this uncertainty has limited influence on the estimation of the 2.5% and the 97.5% quantiles.

3. Theoretical Analysis of Wrong Estimation of the Nuisance Parameter λ

We first analyse the influence of a wrong estimation of the parameter λ on the estimations of the 2.5% and the 97.5% quantiles. We assume that we observe the random variable $t_\lambda^{-1}(X)$ where $X \sim N(\mu, \sigma^2, 0, \infty)$ and we have a possibly wrong estimate $\hat{\lambda}$ for the parameter λ . From this estimate, we assume that the “true” underlying truncated normal distribution is $t_{\hat{\lambda}}(t_\lambda^{-1}(X))$ for which we use estimates of the expected value and the standard deviation to estimate its 2.5% and 97.5% quantiles. These quantiles are then transformed by the inverse Box–Cox transformation $t_{\hat{\lambda}}^{-1}$ to obtain estimates for the limits of the reference intervals. Figure 2 illustrates the situation we consider.

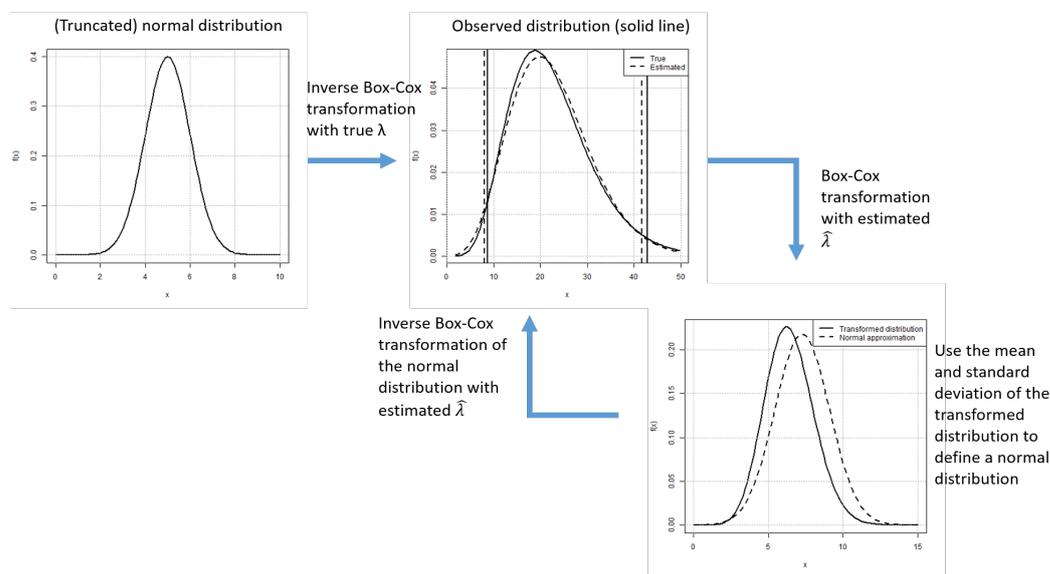


Figure 2. Scenario for the analysis of the influence of the wrong estimation of the parameter λ for infinite sample sizes.

In this section, it is assumed that the estimations for μ and σ are based on a (infinite) sample from the distribution $t_{\hat{\lambda}}(t_\lambda^{-1}(X))$ where $X \sim N(\mu, \sigma^2, 0, \infty)$. Because the corresponding integrals have no obvious solutions, we use a sample size of $n = 100,000$ to estimate μ and σ by the corresponding mean and sample standard deviation of the sample, i.e., we generated 100,000 values from a truncated normal distribution with expected value

μ and standard deviation σ and applied the transformation $t_\lambda(t_\lambda^{-1})$ to obtain the theoretical “wrong” estimates $\hat{\mu}$ and $\hat{\sigma}$.

We vary the true value for λ between 0 and 1. Without a loss of generality, we assume that the lower limit of the reference interval is at 1. This can always be achieved by changing or scaling the measurement units. The upper limit of the reference interval is varied between slightly above 1 and up to 50. In order to quantify the error of the estimated reference interval caused by the wrong estimate $\hat{\lambda}$ for λ , we measure the difference in terms of the zlog value that was introduced in [13]. The zlog value is a transformation based on a logarithmic transformation and a normalisation such that logarithms of the reference interval limits are mapped to the 2.5% and the 97.5% quantiles of a standard normal distribution. Thus, the reference interval is transformed to the range $[-1.96, 1.96]$:

$$\text{zlog}_{l,u}(x) = \frac{\log(x) - \frac{\log(l)+\log(u)}{2}}{\log(u) - \log(l)} \left(\Phi^{-1}(0.975) - \Phi^{-1}(0.025) \right) \tag{3}$$

where Φ is the cumulative distribution function of the standard normal distribution.

Figure 3 shows the absolute zlog deviation for the lower (a) and upper (b) limit of the reference interval depending on λ and the upper limit of the reference interval when the lower limit is fixed at 1.

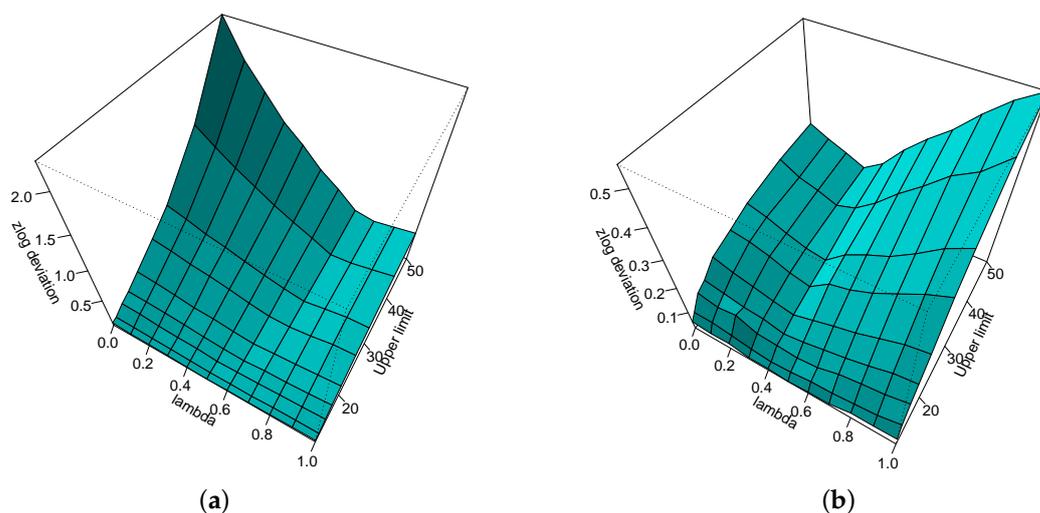


Figure 3. Absolute zlog deviation for the lower (a) and upper (b) limit of the reference interval depending on λ and the upper limit of the reference interval when the lower limit is fixed at 1.

It can be seen that the effect of the wrong λ estimation to the estimate of the upper limit of the reference interval is quite limited, never exceeding the value of 0.5. The effect on the lower limit of the reference interval is high for λ values close to zero and a large upper limit of the reference interval.

Figure 4 zooms into Figure 3, showing the results only for a maximum upper limit of the reference interval of 30. It can be seen that for $\lambda > 0.2$, a wrong estimate of λ has very little influence on the estimate of the upper limit of the reference interval and only when $\lambda < 0.2$ there is a non-negligible effect on the estimate for the lower limit of the reference interval.

In order to further analyse the influence of a wrong estimate for λ on the estimate of the reference interval, we consider the observed distribution $t_\lambda^{-1}(X)$ where $X \sim N(5, 1, 0, \infty)$. Figure 5a shows the estimate of the upper limit for $\lambda = 0$ and wrong estimate $\hat{\lambda}$ between 0 and 1. The dotted line is the true upper limit of the reference interval. The curve indicates the estimated upper limit of the reference interval depending on $\hat{\lambda}$. Figure 5b shows the same for $\lambda = 0.1$.

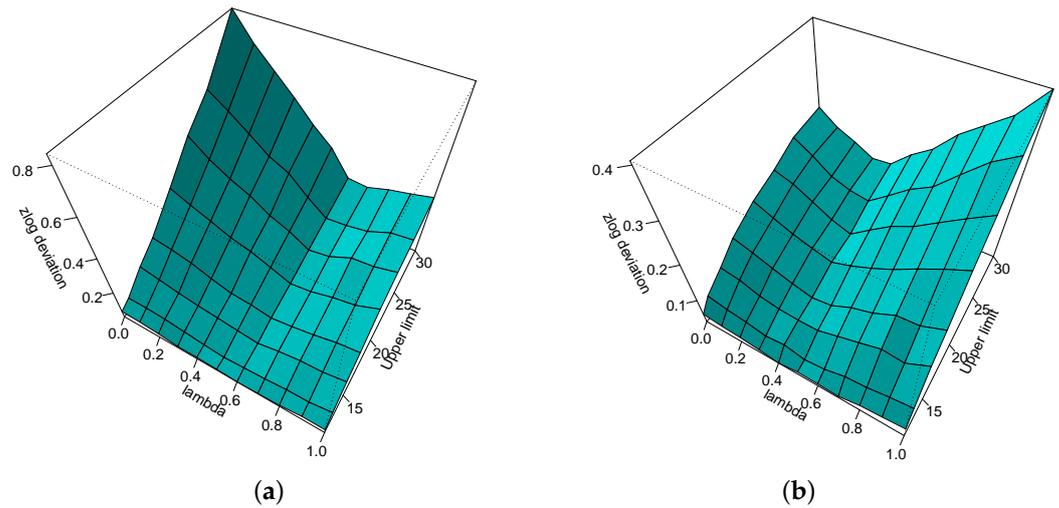


Figure 4. Absolute zlog deviation for the lower (a) and upper (b) limit of the reference interval depending on λ and the upper limit of the reference interval—here bounded to 30—when the lower limit is fixed at 1.

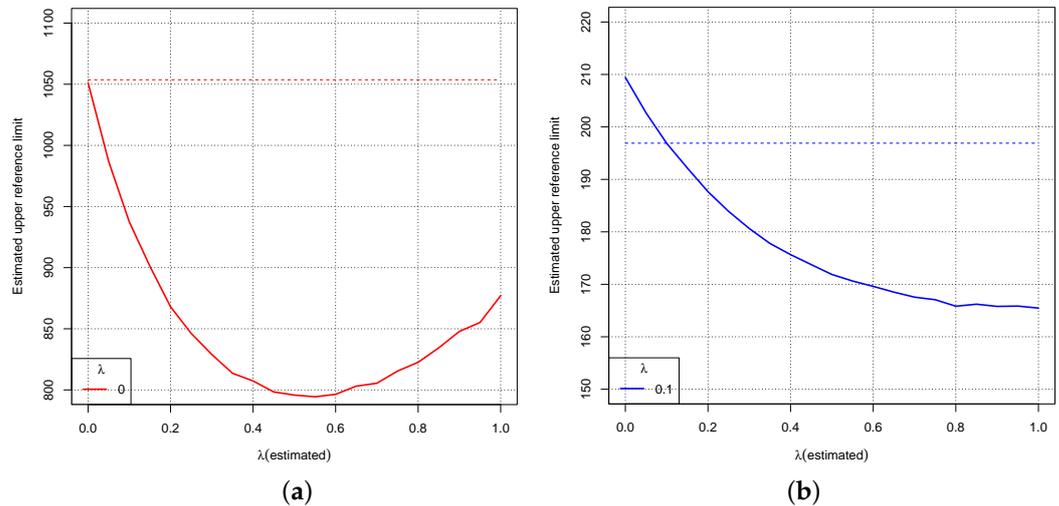


Figure 5. Dotted lines: True upper limit of the reference interval for the distribution $t_{\lambda}^{-1}(X)$ where $X \sim N(5, 1, 0, \infty)$ for $\lambda = 0$ (a) and $\lambda = 0.1$ (b). The curves show the estimated upper limit of the reference interval depending on the wrong estimate $\hat{\lambda}$ for λ .

Figure 6b shows the same as Figure 5 for $\lambda \in \{0.2, 0.3, 0.4, 0.5, 0.7, 1\}$. Figure 6a shows the same for the lower limit of the reference interval including the values 0 and 0.1 for λ .

Figures 5 and 6 indicate again that a wrong estimate for the parameter λ will not affect the estimate for the reference interval drastically unless λ is close to 0, i.e., the Box–Cox transformation should be close to the logarithmic transformation. For λ close to 1, even a completely wrong estimate shows little effect on the estimates for the reference intervals. For smaller values of λ , an overestimation of λ implies high errors for the reference interval estimate. Figures 5 and 6 might even suggest that always choosing $\lambda \approx 0$ leads to small errors with respect to the reference interval.

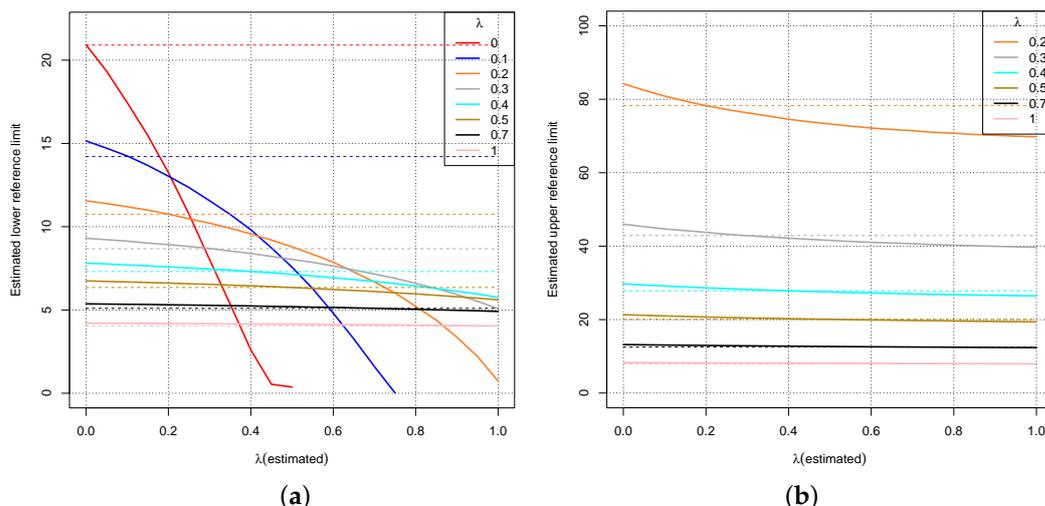


Figure 6. Dotted lines: True lower (a) and upper (b) limit of the reference interval for the distribution $t_{\lambda}^{-1}(X)$ where $X \sim N(5, 1, 0, \infty)$ for different values of λ . The curves show the estimated lower (a) and upper (b) limit of the reference interval depending on the wrong estimate $\hat{\lambda}$ for λ .

The observation is supported by investigating critical values of λ more closely where the logarithmic transformation could lead to unacceptable results for the reference interval. In [14], an equivalence test was proposed to check whether the deviation of an estimated reference interval from a given reference interval can be considered to be relevant in medical terms.

Figure 7 zooms into Figure 6 for $\lambda = 0.4$ and shows the ranges for which the equivalence test according to [14] would not reject the corresponding estimated reference limits as black lines. For both limits of the reference interval, the logarithmic transformation yields acceptable deviations in terms of the equivalence test, although the upper reference limit is almost at the boundary of the acceptance range. However, if λ is overestimated, i.e., for $\lambda > 0.7$, the lower limit of the reference interval can no longer be accepted.

Not shown here are the similar results that $\lambda = 0.5$ yields.

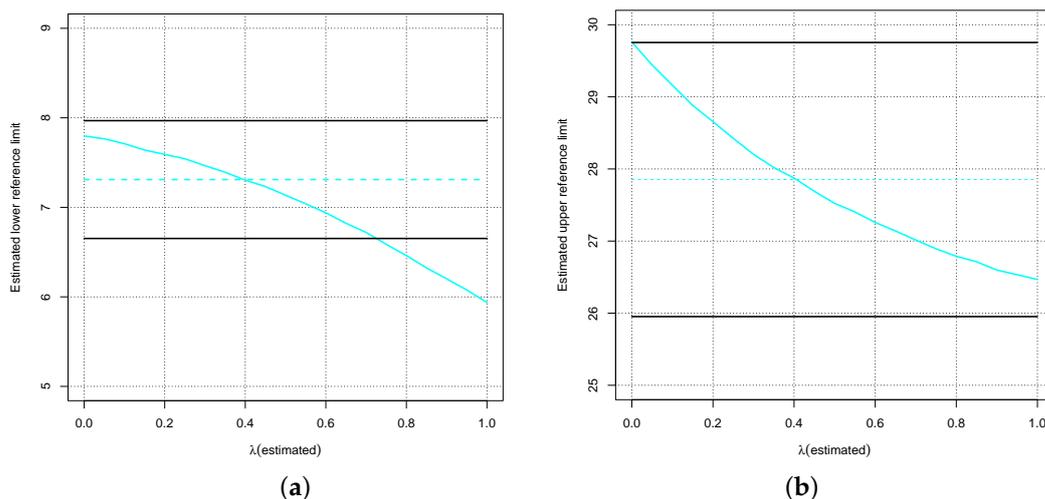


Figure 7. Blue dotted lines: True lower (a) and upper (b) limit of the reference interval for the distribution $t_{0.4}^{-1}(X)$ where $X \sim N(5, 1, 0, \infty)$. The curves show the estimated lower (a) and upper (b) limit of the reference interval depending on the wrong estimate $\hat{\lambda}$ for $\lambda = 0.4$. The black lines indicate the ranges for which the tolerance test according to [14] would not reject the corresponding estimated reference limits.

4. Influence of the Sample Size

After having shown in the previous section that a correct estimation of λ is only important for the estimate of the reference interval when λ is close to 0, this section investigates the uncertainty of the λ estimate depending on the sample size.

We use the same observed distribution as in the previous section, i.e., $t_{\lambda}^{-1}(X)$ where $X \sim N(5, 1, 0, \infty)$ for $\lambda \in \{0, 0.25, 0.5, 0.75, 1\}$. The sample sizes are varied between 100 and 1000. For each λ value and each sample size, we executed 10,000 Monte-Carlo simulations to obtain estimates for λ and the reference interval. For estimating λ , three methods were used: the probability plot correlation coefficient (PPCC), the Shapiro–Wilk test and a maximum likelihood estimator (log-likelihood). Based on the estimated λ , the data are Box–Cox transformed, the mean $\hat{\mu}$ and standard deviation $\hat{\sigma}$ of the transformed data are computed and the 2.5% and the 97.5% quantiles of the normal distribution $N(\hat{\mu}, \hat{\sigma}^2)$ are determined. The estimates for the reference interval are obtained by applying the corresponding inverse Box–Cox transformation to the quantiles. In addition, the reference interval is also estimated directly based on the sample’s corresponding quantiles.

Figure 8 shows the results for $\lambda = 0$ for the lower (a) and upper (b) limit of the reference interval. The sample size is indicated on the x-axis, and the values for the limit of the reference interval are on the y-axis. The grey line indicates the true limit of the reference interval. The coloured curves show the 95% confidence interval for the estimate of the limit of the reference interval based on 10,000 Monte-Carlo simulations for the above-mentioned estimation methods. The black curve represents the 95% confidence interval when the limit of the reference interval is estimated directly by the corresponding sample quantile. The three different methods for λ estimation—the coloured lines—lead to almost exactly the same confidence intervals for the reference interval. The confidence intervals based on the sample quantiles are broader, which is no surprise because they do not make use of the fact that the data originate from a normal distribution to which an inverse Box–Cox transformation was applied.

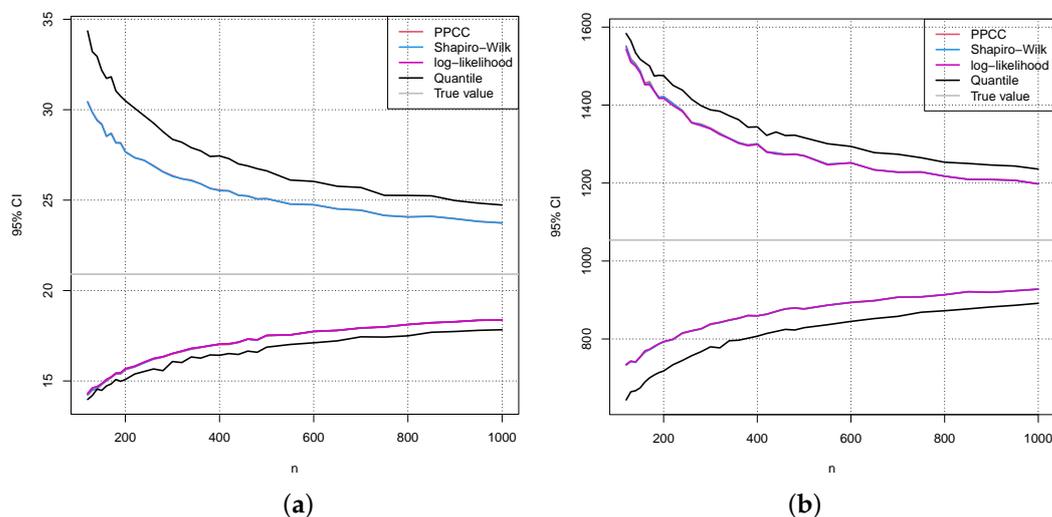


Figure 8. For each sample size, 10,000 samples from the distribution $t_0^{-1}(X)$ where $X \sim N(5, 1, 0, \infty)$ were taken to compute the 95% Monte-Carlo confidence intervals for the lower (a) and upper (b) limit of the reference interval depending on the sample size. The grey line indicates the true limit of the reference interval. The coloured lines use estimations for the reference intervals based on an estimate of λ . The black line is the 95% confidence interval when the 2.5% and the 97.5% sample quantiles are directly used for the reference interval.

Figures 9 and 10 show the same results as in Figure 8 for the values $\lambda = 0.5$ and $\lambda = 1$, respectively. Comparing the confidence intervals in Figures 8–10 to the errors caused by a wrong estimation of λ in Figures 5 and 6, one can see that they are in a similar range.

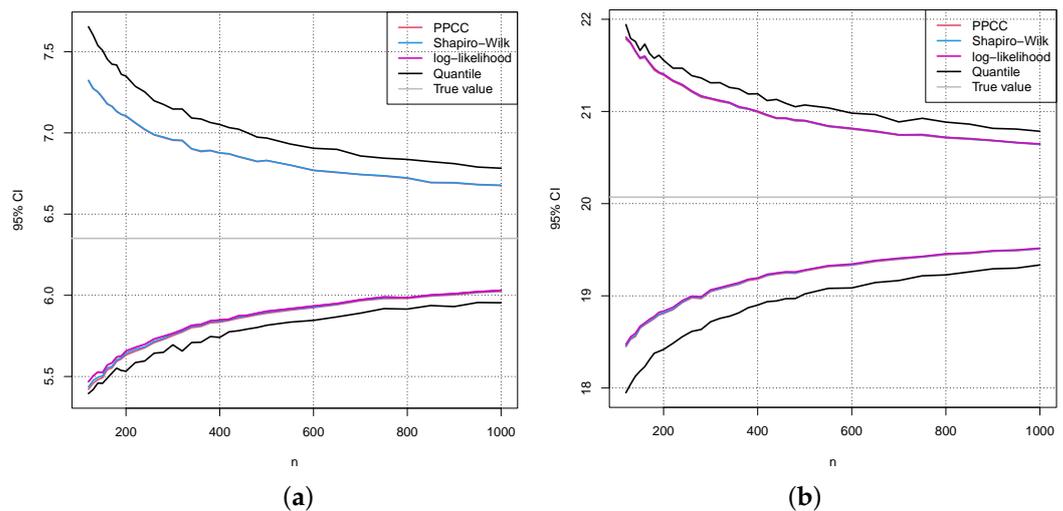


Figure 9. The same as in Figure 8 (a) and (b), except that $\lambda = 0.5$ is chosen.

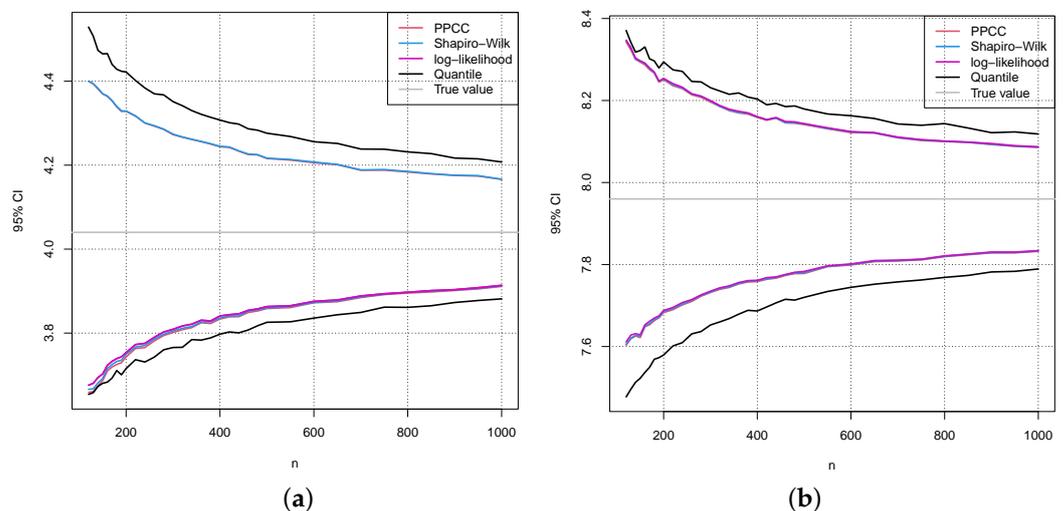


Figure 10. The same as in Figure 8 (a) and (b), except that $\lambda = 1$ is chosen.

Figure 11 shows the 95% confidence intervals for the parameter $\lambda \in \{0, 0.25, 0.5, 0.75, 1\}$ depending on the sample size. For $\lambda \in \{0.5, 0.75, 1\}$, the confidence intervals cover almost the full interval $[0, 1]$ for the small samples sizes and make it impossible to obtain a reliable estimate for λ in these cases. The maximum likelihood method was used here for estimating λ . But the other mentioned methods yield very similar results.

Finally, we also apply one of the established indirect methods for reference interval estimation to the ideal simulated data to demonstrate that the estimation of λ is a daunting task when λ is close to 1. For each of the sample sizes $n \in \{100, 200, \dots, 900, 1000\}$, we have carried out 100 simulations, i.e., generated data 100 times from a truncated normal distribution $X \sim N(5, 1, 0, \infty)$ with a Box-Cox transformation with $\lambda = 1$. Figure 12 shows the box plots for the 100 estimates for each of the simulated sample sizes. The estimates were computed using the indirect method RefineR [9].

It is obvious that the methods are close to random guessing for λ estimation in the case of smaller sample sizes, and even for larger sample sizes, the estimates for λ carry a high uncertainty. It also obvious that the methods seems to have a bias, systematically underestimating λ . Although this is not a desirable property in terms of statistics, it seems to be beneficial for the estimation of the reference intervals because we have seen that overestimating λ can lead to quite wrong reference interval estimates, whereas underestimating λ still produces reasonable reference interval estimates.

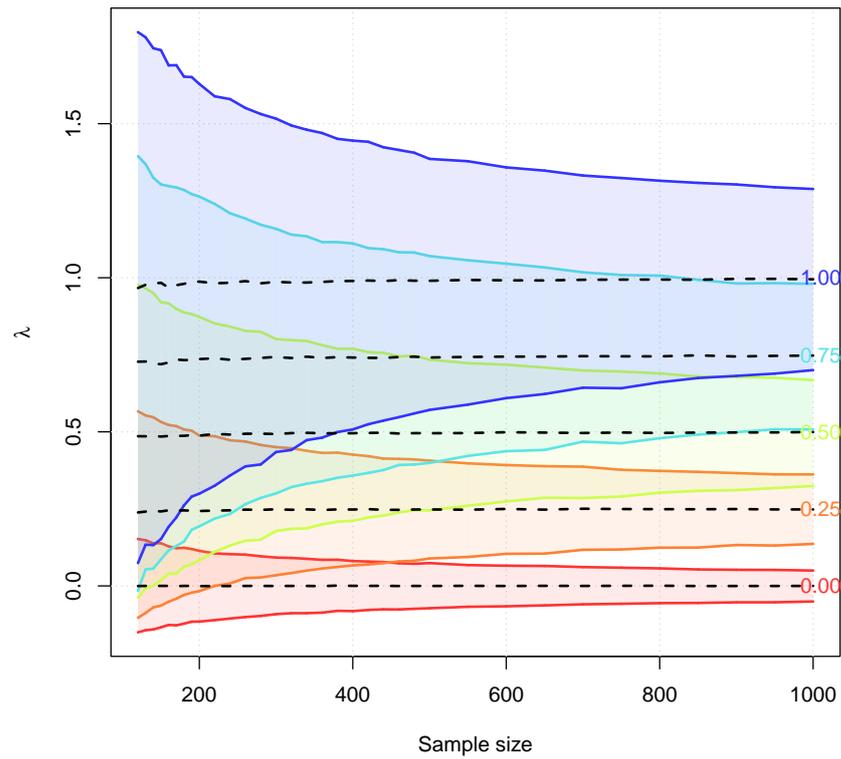


Figure 11. Dotted lines correspond to the mean of the estimates for $\lambda \in \{0, 0.25, 0.5, 0.75, 1\}$ based on 10,000 Monte-Carlo simulations. The shaded areas and the lines are based on the 95% confidence intervals for λ .

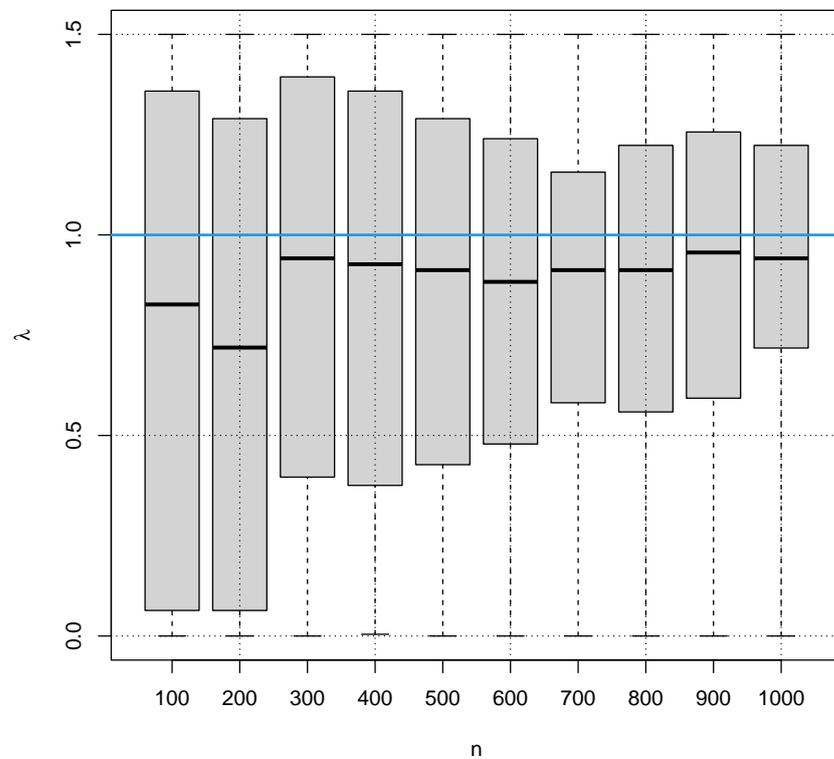


Figure 12. Box plots for the estimates $\hat{\lambda}$ of the indirect method RefineR for 100 simulations for different samples n from the distribution $X \sim N(5, 1, 0, \infty)$ with a Box-Cox transformation with $\lambda = 1$ (blue line).

Figure 13a,b show the corresponding box plots for the estimates of the lower and upper limit of the reference interval, respectively. The method seems to have a bias tending to too narrow reference intervals. The bias decreases with increasing sample size.

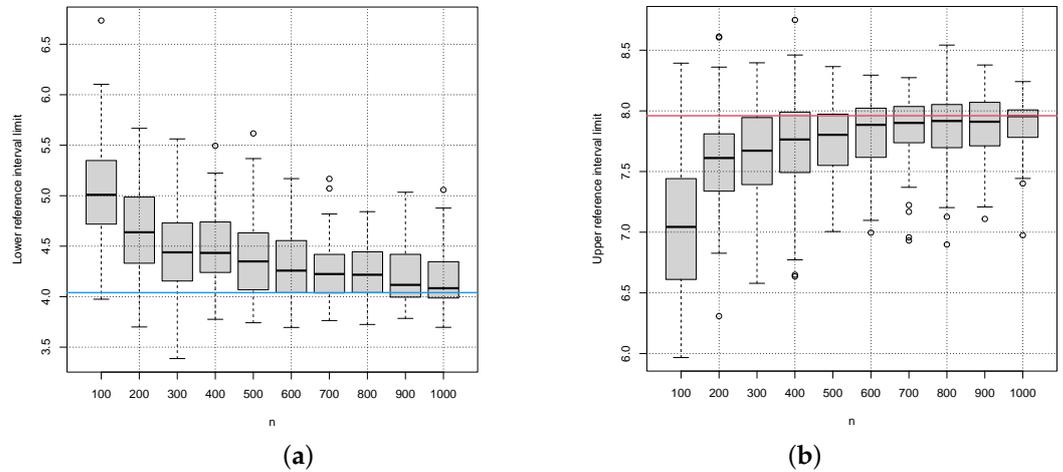


Figure 13. Box plots for the estimates of lower (a) and upper (b) reference interval limit of the indirect method RefineR for 100 simulations for different samples n from the distribution $X \sim N(5, 1, 0, \infty)$ with a Box–Cox transformation with $\lambda = 1$. The red and blue lines indicate the true lower (blue line) and true upper limit (red line) of the reference interval.

Figure 10 shows much better results than RefineR in Figure 13. This is, however, an unfair comparison because RefineR was designed to deal with mixed populations including pathological values, whereas the methods in Figure 10 only work under the ideal assumption of an exclusively healthy population.

It is notable that a wrong estimate for λ does not show a high correlation with wrong estimates for the reference limits as Figure 14 indicates, where the absolute errors for the λ estimates are shown on the x -axis and the absolute errors for the estimates of the reference interval limits are shown on the y -axes.

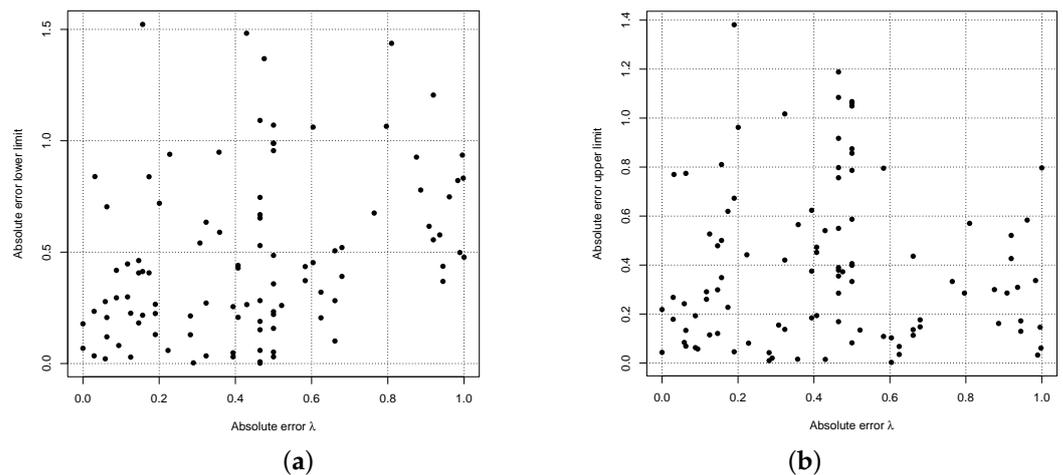


Figure 14. RefineR’s absolute errors of the λ estimate and the lower (a) and upper (b) reference interval limit for sample size $n = 300$.

Other indirect methods yield similar results as RefineR for the simulated ideal data. The results are therefore not shown here.

5. Discussion and Conclusions

The advantage of indirect methods for reference interval estimation and verification is that they do not require a careful selection and examination of a healthy sub-population and they avoid ethical considerations, i.e., they rely on measurements of blood samples that were needed and taken for other reasons [2]. This is especially important when reference intervals for young children are considered where ethical reasons make it difficult to collect larger amounts of blood samples just for reference interval validation. The price that needs to be paid is that additional distributional assumptions cannot be avoided in order to at least roughly identify the healthy sub-population among a mixed sample from healthy and pathological cases. The usual assumptions are that the healthy sub-population contributes the majority of values and follows a normal distribution after a “suitable” transformation. The overlap of the healthy sub-population with “pathological distributions” should also be limited [15]. It is generally agreed that some kind of Box–Cox transformation should be applied to the data. The crucial question that was discussed in this paper is whether the full range of the parameter λ for the Box–Cox transformation should be considered or whether it might be sufficient to restrict the choice between $\lambda = 0$ and $\lambda = 1$, meaning that the values of the healthy sub-population follow a lognormal or a normal distribution. The advantage of this restriction would also be the avoidance of the otherwise required truncation for arbitrary Box–Cox transformations.

In order to keep our analysis simple, we have restricted our investigations to the idealised case where we only deal with the healthy sub-population. If arbitrary Box–Cox transformations turn out to be of limited benefit even in this ideal case, the advantage of arbitrary Box–Cox transformation in the much more complicated setting of mixed populations might be questionable. Similar weak correlations between *lambda* values estimated with different methods on identical data sets have also been observed by [11]. More general transformations, e.g., Box–Cox transformations with an additional shift parameter as mentioned in [10], make the situation even more difficult and lead to less interpretable results.

Our analyses have shown that a wrong estimate for λ has very limited influence on the resulting estimate for the reference interval when λ is greater than 0.6. For smaller λ values, it is important to not overestimate λ . Otherwise, the upper reference limit will be too low. An underestimation of λ by $\hat{\lambda} = 0$ leads to tolerable errors for the reference interval in terms of an equivalence test described in [14].

Another result of our analyses is that at least for sample sizes clearly below 1000, the uncertainty for the estimation of λ remains relatively high so that a reliable estimate cannot be expected. A minimal sample size of 1000 has also been mentioned in [16].

Taking these results together, it might be sufficient to just consider the two options $\lambda = 0$ and $\lambda = 1$. This would also lead to a significant speed-up of the often very slow algorithms for indirect reference interval estimation because these algorithms try to estimate λ , the healthy sub-population and its parameters at the same time, involving long iterative procedures.

Author Contributions: Conceptualisation, F.K. and G.H.; methodology, F.K. and G.H.; software, F.K. and N.R.; validation, G.H. and N.R.; formal analysis, F.K.; investigation, G.H., F.K. and N.R.; resources, F.K.; writing—original draft preparation, F.K.; writing—review and editing, G.H. and N.R.; visualisation, F.K. and N.R.; supervision, F.K. and G.H.; project administration, F.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: Author Georg Hoffmann was employed by the company Medizinischer Fachverlag Trillium GmbH. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Horowitz, G.; Altaie, S.; Boyd, J.C.; Ceriotti, F.; Garg, U.; Horn, P.; Pesce, A.; Sine, H.E.; Zakowski, J. *Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory*; Tech Rep Document EP28-A3C; Clinical & Laboratory Standards Institute: Wayne, PA, USA, 2010.
2. Jones, G.; Haeckel, R.; Loh, T.; Sikaris, K.; Streichert, T.; Katayev, A.; Barth, J.; Ozarda, Y. Indirect methods for reference interval determination: Review and recommendations. *Clin. Chem. Lab. Med.* **2019**, *57*, 20–29. [[CrossRef](#)] [[PubMed](#)]
3. Haeckel, R.; Wosniok, W. Observed, unknown distributions of clinical chemical quantities should be considered to be log-normal: A proposal. *Clin. Chem. Lab. Med.* **2010**, *48*, 1393–1396. [[CrossRef](#)] [[PubMed](#)]
4. Klawonn, F.; Hoffmann, G.; Orth, M. Quantitative laboratory results: Normal or lognormal distribution. *J. Lab. Med.* **2020**, *44*, 143–150. [[CrossRef](#)]
5. Arzideh, F.; Wosniok, W.; Gurr, E.; Hinsch, W.; Schumann, G.; Weinstock, N.; Haeckel, R. A plea for intra-laboratory reference limits. Part 2. A bimodal retrospective concept for determining reference limits from intra-laboratory databases demonstrated by catalytic activity concentrations of enzymes. *Clin. Chem. Lab. Med.* **2007**, *45*, 1043–1057. [[CrossRef](#)] [[PubMed](#)]
6. Concordet, D.; Geffré, A.; Braun, J.P.; Trumel, C. A new approach for the determination of reference intervals from hospital-based data. *Clin. Chim. Acta* **2009**, *405*, 43–48. [[CrossRef](#)] [[PubMed](#)]
7. Ichihara, K.; Boyd, J.C.; IFCC Committee on Reference Intervals and Decision Limits (C-RIDL). An appraisal of statistical procedures used in derivation of reference intervals. *Clin. Chem. Lab. Med.* **2010**, *48*, 1537–1551. [[CrossRef](#)] [[PubMed](#)]
8. Wosniok, W.; Haeckel, R. A new indirect estimation of reference intervals: Truncated minimum chi-square (TMC) approach. *Clin. Chem. Lab. Med.* **2019**, *57*, 1933–1947. [[CrossRef](#)] [[PubMed](#)]
9. Ammer, T.; Schützenmeister, A.; Prokosch, H.U.; Rauh, M.; Rank, C.M.; Zierk, J. refineR: A Novel Algorithm for Reference Interval Estimation from Real-World Data. *Sci. Rep.* **2021**, *11*, 16023. [[CrossRef](#)] [[PubMed](#)]
10. Ozarda, Y.; Ichihara, K.; Jones, J.; Streichert, T.; Ahmadian, R. Comparison of reference intervals derived by direct and indirect methods based on compatible datasets obtained in Turkey. *Clin. Chim. Acta* **2021**, *520*, 186–195. [[CrossRef](#)] [[PubMed](#)]
11. Meyer, A.; Müller, R.; Hoffmann, M.; Skadberg, Ø.; Ladang, A.; Dieplinger, B.; Huf, W.; Stankovic, S.; Kapoula, G.; Orth, M. Comparison of three indirect methods for verification and validation of reference intervals at eight medical laboratories: A European multicenter study. *J. Lab. Med.* **2023**, *47*, 155–163. [[CrossRef](#)]
12. Yeo, I.-K.; Johnson, R.A. A new family of power transformations to improve normality or symmetry. *Biometrika* **2000**, *87*, 954–959. [[CrossRef](#)]
13. Hoffmann, G.; Klawonn, F.; Lichtinghagen, R.; Orth, M. The zlog value as a basis for the standardization of laboratory results. *J. Lab. Med.* **2017**, *41*, 20170135. [[CrossRef](#)]
14. Haeckel, R.; Wosniok, W.; Arzideh, F. Equivalence limits of reference intervals for partitioning of population data. Relevant differences of reference limits. *LaboratoriumsMedizin* **2016**, *40*, 199–205. [[CrossRef](#)]
15. Haeckel, R.; Wosniok, W.; Arzideh, F. A plea for intra-laboratory reference limits. Part 1. General considerations and concepts for determination. *Clin. Chem. Lab. Med.* **2007**, *45*, 1033–1042. [[CrossRef](#)] [[PubMed](#)]
16. Anker, S.C.; Morgenstern, J.; Adler, J.; Brune, M.; Brings, S.; Fleming, T.; Kliemank, E.; Zorn, M.; Fischer, A.; Szendroedi, J.; et al. Verification of sex- and age-specific reference intervals for 13 serum steroids determined by mass spectrometry: Evaluation of an indirect statistical approach. *Clin. Chem. Lab. Med.* **2023**, *61*, 452–463. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.