Reports of the Institute of Biostatistics

No 06 / 2007

Leibniz University of Hannover

Natural Sciences Faculty

Titel: *Confidence intervals for the proof of safety for abundance data*

Authors: *Frank Schaarschmidt*

1 Introduction

1.1 Distributional assumptions for count data

If the effect of novel agricultural treatments on the fauna is to be assessed, usually abundance data are collected. Here, abundance is measured by counting the number of individuals in certain unit of time and space. Counts randomly occurring with a fixed expectation in a certain unit of time and space can be assumed to follow the Poisson distribution. For a random variable Y following Poisson distribution, the probability of observing y is

$$P\left(Y=y\right)=e^{-m}m^{y}/y! \tag{1}$$

A random variable Y following the Poisson distribution has the property $E\left(Y\right)=m=V\left(Y\right)$. However, if the expectation differs between subunit in time or space, the count follows a mixture of Poisson distributions, with the property $E\left(Y\right)\geq V\left(Y\right)$ This is usually the case for field data of insect abundance. A common distribution to model such data is the negative binomial distribution.

$$P\left(Y=y\right)=\left(\begin{array}{c}k+y-1\\ k-1\end{array}\right)\left(\frac{k}{m+k}\right)^{k}\left(\frac{m}{m+k}\right)^{y} \tag{2}$$

Here, $1/k$ is a dispersion parameter, and m is the expectation of Y. In the Negative Binomial distribution, the variance is a quadratic function of the expectation and the dispersion parameter $1/k$: $V\left(Y\right)=m+\frac{m^2}{k}$. As $k\rightarrow\infty$, the Poisson distribution results. In the following we will denote the Negative binomial distribution $NB\left(m,k\right)$.

1.2 Proof of Equivalence

We consider a single species only. We consider only two treatments, Novum and Standard, where $E\left(Y_1\right)=m_1$ denotes the mean abundance in Novum and $E\left(Y_0\right)=m_0$ denotes the mean abundance in Standard. The parameter of interest is the ratio of mean abundances $\rho=m_1/m_0$. To perform a proof of safety, we are interested in rejecting the null hypothesis

$$H_0: m_1/m_0\leq\theta^l\cup m_1/m_0\geq\theta^u \tag{3}$$

in favor of the alternative hypothesis:

$$H_0: m_1/m_0>\theta^l\cap m_1/m_0<\theta^u \tag{4}$$

where θ^l,θ^u are the lower and upper limits of change in mean abundance which is considered as non-relevant. Preserving a type I error of $\alpha=0.05$, a proof of safety can be performed using a lower 95% confidence limit and an upper 95% confidence limit for m_1/m_0. Lower and upper confidence limits will be denoted $\hat{\rho}^{l,0.95}$, $\hat{\rho}^{u,0.95}$. One can reject H_0, if

$$\hat{\rho}^{l,0.95}>\rho^l\cap\hat{\rho}^{u,0.95}<\rho^u \tag{5}$$

i.e., if and only if the lower and upper confidence limit estimated for the parameter ρ are both included in the range of $\left[\theta^l,\theta^u\right]$ of non-relevant change Wellek (2003).

1.3 Objectives

This report investigates different methods to construct confidence intervals for ρ. Special attention is paid to the two-sided coverage probability of such intervals (since they may serve as a tool for inference as well as estimation), on the minimum coverage probability of the lower and upper limit, since this will lead to maximal type-I-error for a two-sided proof of safety, and the coverage probability of lower limits, since they might serve to proof savety one-sided, if it is only of interest to show that the abundance of a species is not relevantly decreased.

2 Methods

We consider the setting of independent observations $Y_{ij} \sim NB\left(m_i, k\right)$. Such data might be derived from trials comprising two treatments $i = 0, 1$ and n_i, $j = 1, ..., n_i$ replications of each treatment, arranged in a completely randomized design.

The following Sections describe confidence limits $\left[\hat{\rho}^{l,0.95}, \hat{\rho}^{u,0.95}\right]$ for ρ and the available software implementations in R.

2.1 Asymptotic confidence interval from a glm with log-link, assuming Poisson distribution (*GLMP*)

In a generalized linear model with log-link, we assume that effects are additive on the scale of the log:

$$E\left(log\left(E\left(Y\right)\right)\right) = E\left(\eta\right) = X\beta \tag{6}$$

i.e.,

$$E\left(Y\right) = exp\left(X\beta\right) \tag{7}$$

where X denotes the $(N \times p)$ design matrix, $N = \sum_{i=1}^{2} n_i$, $p = 2$. We assume Poisson distribution for the fit. For details, see McCullagh and Nelder (1989). From the model fit we can obtain estimators $\hat{\eta}_1 - \hat{\eta}_0$ for $\eta_1 - \eta_0$ and estimates for the standard error of this estimator, $\hat{\sigma}_{\hat{\eta}_1 - \hat{\eta}_0}$. Asymptotic lower and upper confidence limits for ρ can be obtained using:

$$\left[\hat{\rho}^{l,0.95}; \hat{\rho}^{l,0.95}\right] = \left[exp\left(\hat{\eta}_1 - \hat{\eta}_0 \pm z_{0.95}\sqrt{\hat{\sigma}_{\hat{\eta}_1 - \hat{\eta}_0}}\right)\right] \tag{8}$$

with $z_{0.95}$ the 95% quantile of the standard normal distribution. In the simulation study, we used the function `glm` in the `stats` package of R to fit the model.

2.2 Asymptotic confidence interval from a glm with log-link, assuming using Quasipoisson method (*GLMQP*)

A model can be fitted according to Equation 7 using the quasipoisson method McCullagh and Nelder (1989), estimators are equivalently obtained, but additionally depend on the variance estimated from the data, assuming $V\left(Y\right) = \phi m_i$, where $\phi = 1/k$ is a common dispersion parameter. Analogously, Equation 8 is used to calculate confidence limits. In the simulation study, we used the function `glm` in the `stats` package of R to fit the model.

2.3 Asymptotic confidence interval from a glm with log-link, assuming Negative binomial distribution (*GLMNB*)

A model can be fitted according to Equations 7 using the likelihood of the negative binomial distribution McCullagh and Nelder (1989), estimators and confidence limits are analogously obtained. The problem of fitting glms with negative binomial assumptions, is that the dispersion parameter $1/k$ has to be estimated from the data. For this problem, different estimators exist. We used the implementation `glm.nb` by Venables and Ripley (2002) in the R package `MASS` to fit the negative binomial model. Additionally, the algorithm of Rigby and Stasinopoulos (2005) was used as provided in the R-package `gamlss` and is denoted *GLMNBI*.

2.4 Asymptotic confidence interval for the ratio of means assuming lognormal distribution (*LN*)

Often, strictly non-negative random variables with positively skewed distribution, as are count data, are approximated by assuming lognormal distribution. Chen and Zhou (2006) compare interval estimators for the difference and ratio of means of two independent lognormal samples.
Let Y_{ij}, $j = 1, ..., n_i$, $i = 1, 2$ denote two samples of log-normal distributions, and $Z_{ij} = \log(Y_{ij})$. The ratio of means $\rho = E(Y_1)/E(Y_0) = \mu_1/\mu_0$ is to be estimated, and $\Psi = \log(\mu_1/\mu_0)$. Let the $\hat{\zeta}_i$ denote the estimated mean and $\hat{\sigma}_i$ the estimated standard deviation of samples Z_{ij}. The MLE for Ψ is $\hat{\Psi} = \hat{\zeta}_1 - \hat{\zeta}_0 + \frac{1}{2}(\hat{\sigma}_1 - \hat{\sigma}_0)$, and the maximum likelihood estimator of the ratio of means μ_1/μ_0 of the Y_{ij} then is: $\exp\left(\hat{\Psi}\right)$. A large sample confidence interval can be constructed using $\hat{\tau} = \sqrt{\hat{V}\left(\hat{\Psi}\right)}$. An estimator for $\hat{\tau}$ is $\sqrt{\frac{\sigma_1}{n_1} + \frac{\sigma_1^2}{2n_1} - \frac{\sigma_0}{n_0} - \frac{\sigma_0^2}{2n_0}}$. The confidence interval then is:

$$\left[\hat{\rho}^{l,0.95}; \hat{\rho}^{l,0.95}\right] = \left[exp\left(\hat{\Psi} \pm z_{0.95}\hat{\tau}\right)\right] \tag{9}$$

However, the assumption of continuous data which is included in the assumption of lognormal distribution is violated for count data. Moreover, for low mean abundances the frequent occurrence of $Y = 0$ is a numerical problem for the above confidence interval. Here, Y is replaced by $Y + 0.1$ in case that $Y = 0$ occurs.

2.5 Fieller type confidence interval for ratio of normal means (*N*)

Tamhane and Logan (2004) proposed a test for the ratio of means of two normal populations in the presence of heteroscedasticity, using the Satterthwaite approximation for the degrees of freedom. The test can be inverted approximately using the method applied by Fieller (1954).

$$\left[\hat{\rho}^{l,0.95}; \hat{\rho}^{l,0.95}\right] = \left[\frac{-B}{2A} \pm \frac{\sqrt{(B/2)^2 - AC}}{A}\right] \tag{10}$$

, with

$$A = s_0 \frac{t^2_{dfS,1-\alpha}}{n_0} - \hat{\mu}_0^2$$
$$B = 2\hat{\mu}_1\hat{\mu}_0$$
$$C = s_1 \frac{t^2_{dfS,1-\alpha}}{n_1} - \hat{\mu}_1^2$$

where $s_i = \sqrt{\frac{\sum_{j=1}^{n_i}(y_{ij}-\hat{\mu}_i)^2}{n_i-1}}$, $\hat{\mu}_i = \frac{\sum_{j=1}^{n_i} y_i}{n_i}$, and $t_{dfS,1-\alpha}$ being a quantile from the t-distribution with $df = dfS$, and $dfS = \frac{\left(\hat{\rho}^2 s_0^2/n_0 + s_0^2/n_0\right)^2}{\hat{\rho}^4 s_0^4/n_0^2(n_0-1) + s_0^4/n_0^2(n_1-1)}$.

This interval can be contructed for $A > 0$, not otherwise. In case that $A < 0$, we define

2.6 Non-parametric confidence interval for the ratio of locations (*HL*)

A non-parametric confidence interval for the difference of location parameters has been described by Hodges and Lehmann (1963) and Bauer (1972). This interval is conditional to the sample. It is based on the assumption that two independent samples y_{1j}, $y_{0j'}$, with $j = 1, ..., n_1$, $j' = 1, ..., n_0$ are distributed $F(Y)$ and $G(Y) = F(Y/\rho)$, where $F(Y)$ is a continuous distribution, and ρ is the unknown parameter of interest.

Hothorn and Munzel (2002) show that an exact confidence interval for ρ can be constructed by first calculating all pairwise ratios $r_{jj'} = \frac{y_{1j}}{y_{0j'}}$, $j = 1, ..., n_1, j' = 1, ..., n_0$, second order $r_{jj'}$ according to their magnitude, resulting in ordered sample of pairwise ratios $r_{(1)} < r_{(2)} < ... < r_{(n_1 n_0)}$. A 0.95 lower and upper confidence limits can now be constructed by taking the 0.05- and 0.95-quantile of the Wilcoxon distribution with parameters denoted $w_{0.05,n_1,n_0}$ and $w_{0.95,m_1,n_0}$. The confidence interval is:

$$[\hat{\rho}_L, \hat{\rho}_L] = \left[r_{(w_{0.05,n_1,n_0})}, r_{(w_{0.95,n_1,n_0})} \right].$$

If the distribution is not continuous and ties occur in the data, the authors recommend the use of a conditional approach based on the Streitberg-Röhmel algorithm.

An implementation of the Hodges-Lehmann interval Hodges and Lehmann (1963) for the difference in locations based on the ordered sample of pairwise differences $d_{(1)} < d_{(2)} < ... < d_{(n_1 n_0)}$, with $d_{jj'} = y_{1j} - y_{0j'}$, $j = 1, ..., n_1, j' = 1, ..., n_0$ is available in the R-package `exactRankTests`. Since ordering $ld_{jj'} = \log(y_{1j}) - \log\left(y_{0j'}\right)$ results in exactly the same order as ordering the $r_{jj'}$. Then, applying the exponential function to the interval constructed based on the ordered sample of $ld_{(1)} < ld_{(2)} < ... < ld_{(n_1 n_0)}$, results in the interval described by Hothorn and Munzel. However, if the above interval shall be applied to count data, one has to deal with the occurrence of zeros in the samples y_1 and y_0. In case that zeros occur in only one of the samples, setting

$$ld = \log(0) - \log(\epsilon) = -\infty \Leftrightarrow r = 0/\epsilon = 0,$$

and

$$ld = \log(\epsilon) - \log(0) = \infty \Leftrightarrow r = \epsilon/0 = \infty,$$

for $\epsilon > 0$ we can numerically deal with the problem of ordering $ld_{jj'}$ and finally construct a confidence interval. There is no reasonable numerical replacement for $r = 0/0$ or $ld = \log(0) - \log(0)$, hence such values are omitted from the ordered sample of lds. If the samples n_1, n_0 and the mean abundances are small, frequently all observations of at least one sample $x_{ij} = 0 \forall j$, for $i = 1$ or $i = 2$. Then, the all $r_{jj'}$ or $ld_{jj'}$ have the same value and the confidence interval degenerates to a point. In this case, we formally define the interval to be $[\hat{\rho}_L, \hat{\rho}_L] = [0, \infty]$.

In the context of log-transformation in parametric analyses, the problem of occurring zeros is often solved by using the transformation $\log(y + \epsilon)$. However, the intervals discussed here are based on the stochastical order between the samples y_1 and y_0. By simple numerical examples it can be shown that result of ordering $ld^*_{jj'} = \log(y_{1j} + \epsilon) - \log\left(y_{0j'} + \epsilon\right)$ may markedly differ from ordering $r_{jj'}$. Moreover, the resulting confidence interval for ld^* can not be back transformed into a confidence interval for ρ. If the exponential function is applied to the confidence bounds for ld^*, the resulting values of confidence bounds $\hat{\rho}_*$ differ markedly depending on the choice of ϵ. For these reasons, the nonparametric confidence interval is not recommended for use, if abundance is low and samples are small.

Nevertheless, this interval was included in the simulation study under the acronym *HL*, based on $ld^*_{jj'} = \log(y_{1j} + \epsilon) - \log\left(y_{0j'} + \epsilon\right)$, with $\epsilon = 0.1$. Since it is based on Permutation, the exact solution using the Streitberg-Röhmel algorithm is only available for limited sample size. Here, for more than $n_0 + n_1 > 50$ observations an asymptotic version is computed.

Note, that the parameter ρ as defined for the *HL* interval differs from the definition of $\rho = E(Y_1)/E(Y_0)$ for the methods explicitly assuming a certain distribution. It can not be interpreted as the ratio of location parameters, but as a parameter of changing scale.

3 Conclusions

If the negative binomial distribution (of which the Poisson distribution is a special case) can be reasonably assumed for the data, constructing confidence intervals based on the fit of a GLM with family negative binomial or quasipoisson is recommended. Fits based on the Poisson assumption lead to unacceptably liberal confidence intervals also for modest overdispersion $k = 10$. If the ratio of mean abundance is extreme, i.e. $\rho = 0.1, 10$ the quasipoisson method might be conservative, but if the mean abundances are about equal, ρ is close to 1, the quasipoisson approach performs acceptable, also for data following the negative binomial distribution. If the model fit assumes the negative binomial distribution, the algorithm of Rigby and Stasinopoulos (2005) appears preferable in the considered settings, being slightly less liberal the implementation of Venables and Ripley (2002). However, it should be noted, that none of the methods is close to the nominal level for all considered situations. In a proof of safety with nominal 5% error probability, actual error probabilities of up to 8% may occur.
If the data indeed follow a negative binomial distribution, basic assumptions of the *LN*, *HL* and *N* method are violated. The *LN* method performs mostly acceptable if situations with low mean abundance (i.e. high discreteness of the distribution and frequent occurrence of zeros) are avoided. Adding a small number to the observed counts avoids numerical problems but the confidence interval described in this report should not be interpreted as intervals for $\rho = E(Y_1)/E(Y_0)$. The same comes true for the

HL method. In many situations, it does not cover the parameter $\rho = E(Y_1)/E(Y_0)$ with acceptable probability. It should be noted that the comparison of the HL method with the GLM based methods might be rather unfair, since we assume a certain distribution and compare methods that are not based on these assumptions with others which are based on that assumptions. It is then very likely that the latter perform better. Based on this report, the HL method can not be recommended because the underlying distributional assumptions differ substantially from the properties of count data (discreteness, frequent occurrence of zeros), and because the parameter ρ as defined by Hothorn and Munzel (2002) can not interpreted as the ratio of expectations or ratio of medians, but rather is a parameter of change in scale between two distributions. This is a problem the easy interpretation of the parameters is important to define the safety margins when objective is to perform a proof of safety. However, it should be investigated whether the *HL* method outperforms the *GLMQP*, *GLMNB*, or *GLMNBI* method for distributions which show the typical properties of count data, but are not following the negative binomial distribution.

4 Simulation study

Objective of the simulation study is to investigate the coverage probabilities of the above confidence interval methods, with special respect to the

1. two-sided coverage probability $C_{ts} = P\left(\rho \in \left[\hat{\rho}^l, \hat{\rho}^u\right]\right)$,
2. the minimum coverage probability of the upper and the lower bound $C_{min} = min\left(P\left(\rho \geq \hat{\rho}^l\right), P\left(\rho \leq \hat{\rho}^u\right)\right)$,
3. and the coverage probability of the lower bound, $C_l = P\left(\rho \geq \hat{\rho}^l\right)$.

4.1 Assumption of the Negativ Binomial distribution

Samples were drawn from the family of negative binomial distribution $Y_i \sim NB(\mu_i, k)$, i.e. assuming equal overdispersion parameters in both samples.

The following parameter settings were considered:

$$m_0 = 1, 5, 10, 50,$$
$$\rho = 0.1, 0.5, 0.8, 1, 1.25, 2, 10,$$
$$n_0 = n_1 = 5, 10, 20, 50.$$

The dispersion parameter was set to:

$$k = 1000, 10, 1$$

where $k = 1000$ practically corresponds to a Poisson distribution.

4.2 Results

4.2.1 Overview

From Figures 1, 3, and 2 it is obvious, that if data in fact follow a Poisson distribution, the *GLMP*, *GLMQP*, and *GLMNB* methods are acceptable options. If data follow a negative binomial distribution with marked overdispersion $k = 10$, or $k = 1$, only the *GLMQP* and *GLMNB* methods perform acceptable. All other methods severely fail to cover the true parameter in at least a few cases.

Summarizing over all considered parameter settings, severe violations of the nominal confidence level of the lower bound, here defined as $C_l < 0.93$, occurred in 52% of the cases for *GLMP*, in 4% of the cases for *GLMNB*, never for *GLMNBI* and *GLMQP*, and in 14%, 12%, and 16% of the cases for the *HL*, *LN*, and *N* method, respectively. Severely lowered minimal coverage probabilities, here defined as $C_{min} < 0.93$, occured in 52% of the cases for *GLMP*, in 7% of the case for *GLMNB*, never for *GLMNBI* and *GLMQP*, in 29% of the cases for the *HL* method, and in 21% and 22% for the *LN* and *N* method, respectively.

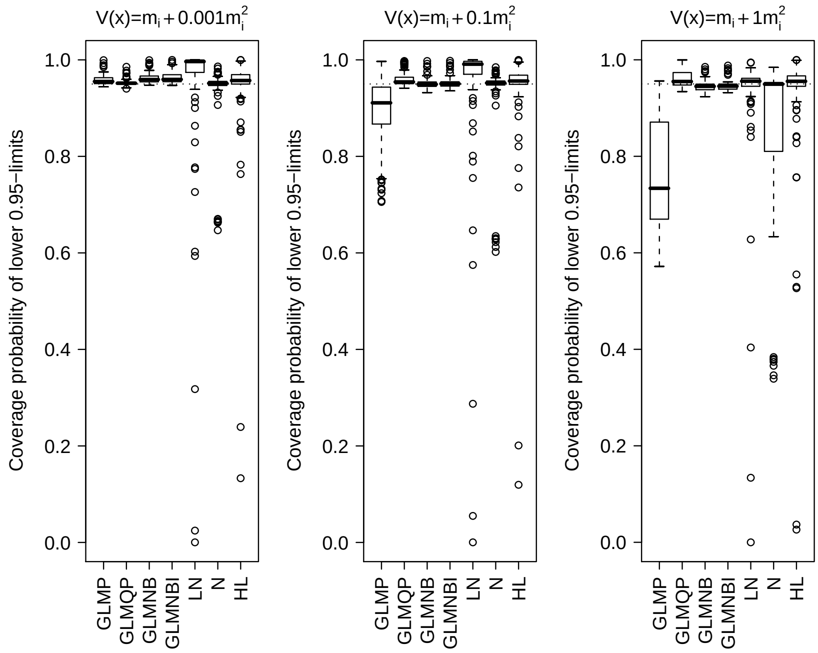

Figure 1: Coverage probabilities of nominal 95% lower confidence limits

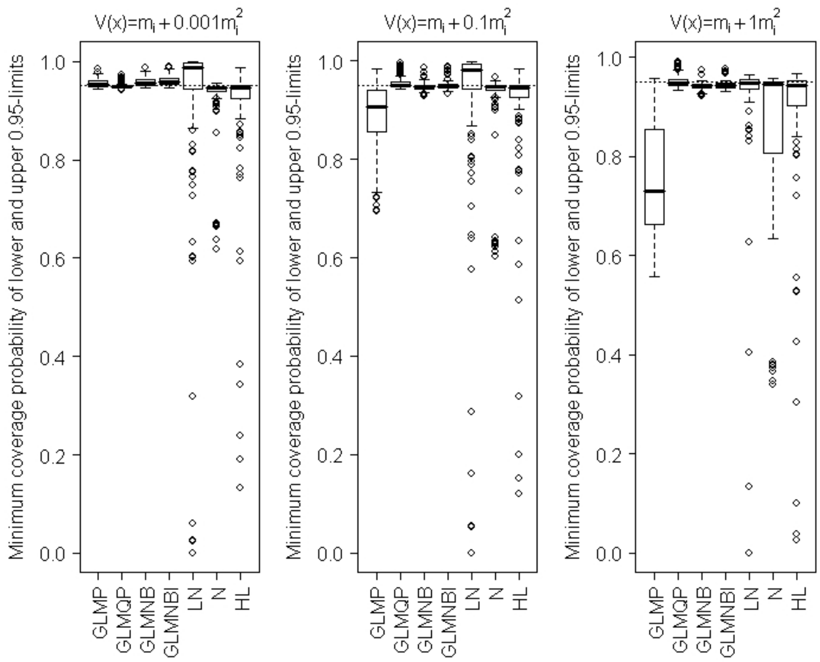

Figure 2: Minimal coverage probabilities of nominal 95% lower and upper confidence limits

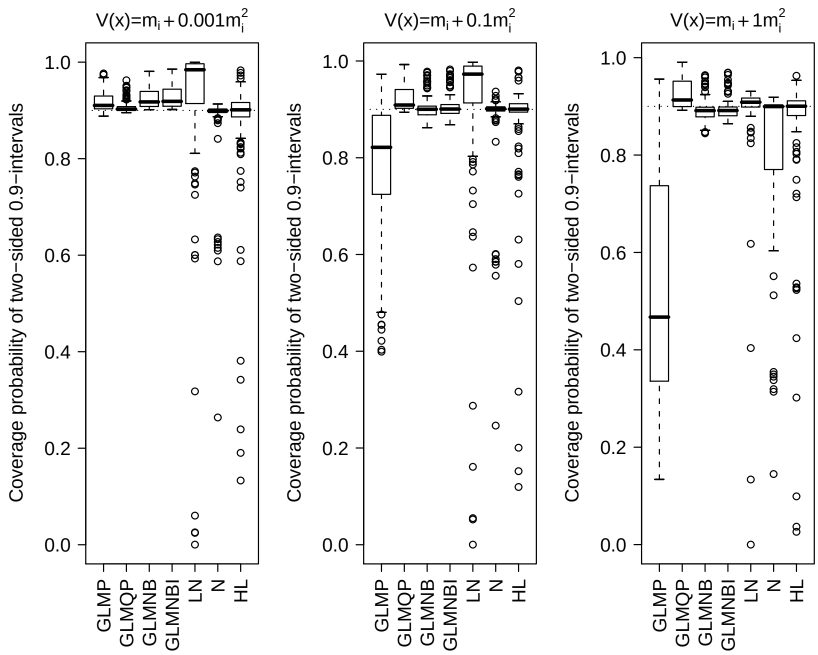

Figure 3: Coverage probabilities of nominal 90% two-sided confidence intervals

4.2.2 Detailed results for *GLMQP* and *GLMNB*

In Figures 4, 5, 6, 7, the coverage probabilities of lower 95% bounds of the *GLMNB*, and *GLMQP* are summarized. The *GLMNB* method becomes liberal for marked overdispersion and small sample size. The GlmQP method becomes conservative for large overdispersion and extreme values of ρ, due to mis-specification of the variance-mean-relation. In other cases, using the GLMQP method for negative binomial data has no severe effect on the coverage probability.

If the intervals are used to estimate lower bounds for ρ one has to expect coverage probabilities as low as 92% for nominal 95% limits if large overdispersion is present ($k = 1$) and sample size is small ($n_i = 5$). For sample sizes of $n_i = 10$, one has to expect coverage probabilities of about 94% for nominal 95% confidence intervals.

Using the *GLMNB* for a proof of non-inferiority with margins $r_L = 0.8, 0.5$, will have type-I-error rates up to $0.08, 0.065, 0.06$ if overdispersion is large ($k = 1$) and sample sizes are $n = 5, 10, 20$, respectively. Using the GLMQP method in the same situations, the type-I-error rate will rarely be higher than 0.06. Hence, if interest is in assessing non-inferiority one should be aware of the slightly liberal performance.

If interest is in a proof of equivalence the minimal coverage of the lower and upper 95% bounds is the criterion for recommendation. Figures 8, 9, 10, and 11 summarize this criterion for the *GLMNB* and *GLMQP* method. As can be expected, the situation is slightly worse than for the lower bounds alone.

Using the *GLMNB* method, one has to be aware that the actual type-I-errors can be as low as 0.08 for $n_i = 5$ and large overdispersion ($k = 1$). Also for sample sizes $n_1 = 10, 20$, the actual type-I- error can be markedly lower than 0.05 but in the here considered situations was not lower than 0.65. For $n_i = 20$, actual type-I-errors below 0.06 did not occur in the present simulation study.

Using the GLMQP method with margins $r_L = 0.8, 0.5$ and $r_U = 1.25, 2.0$, also results in slightly liberal performance. For $n_i = 5$, the actual type-I-errors as low as 6.5% were occasionally observed.

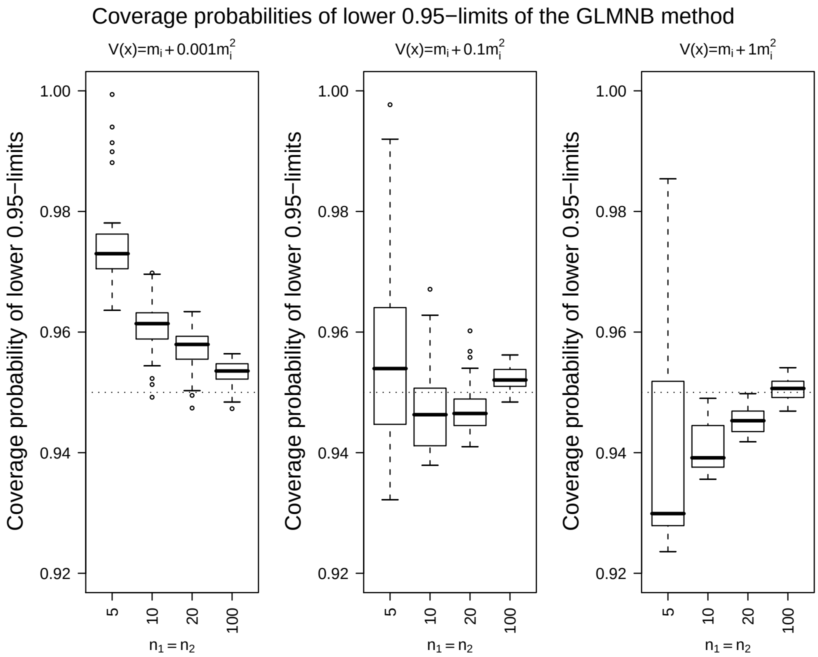

Figure 4: Coverage probabilities of nominal 95% lower confidence limits of the *GLMNB* method, in dependence of the sample size $n_0 = n_1$

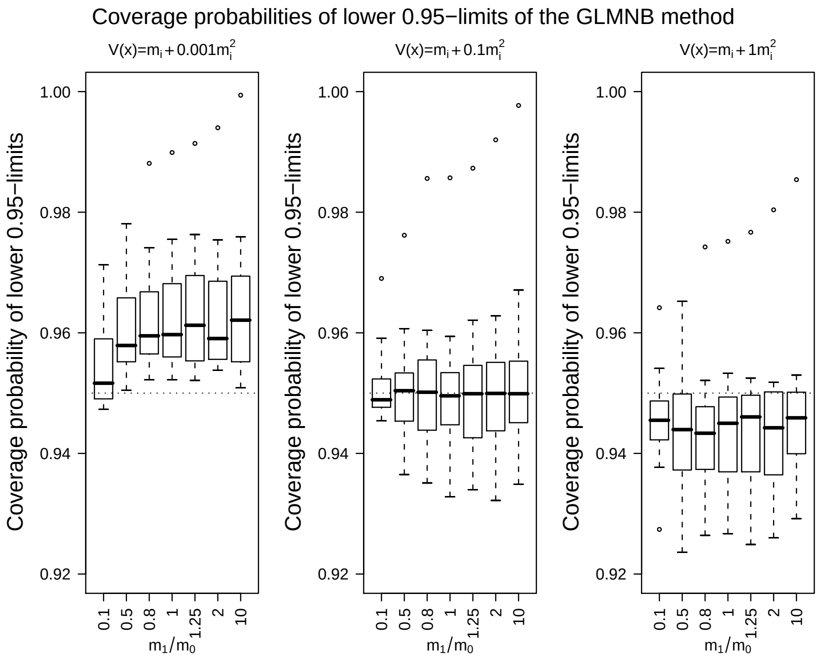

Figure 5: Coverage probabilities of nominal 95% lower confidence limits of the *GLMNB* method, in dependence of the true ratio $\rho = m_1/m_0$

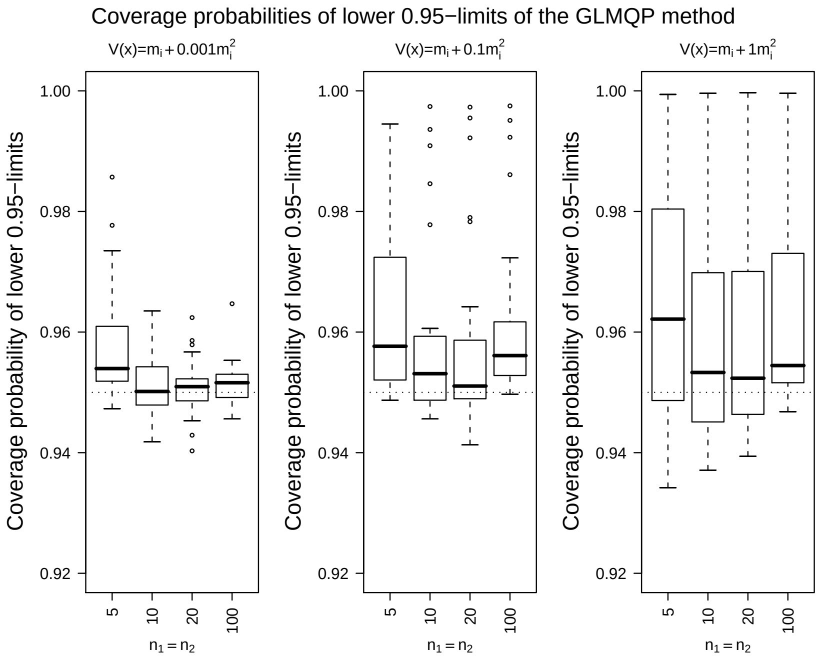

Figure 6: Coverage probabilities of nominal 95% lower confidence limits of the *GLMQP* method, in dependence of the sample size $n_0 = n_1$

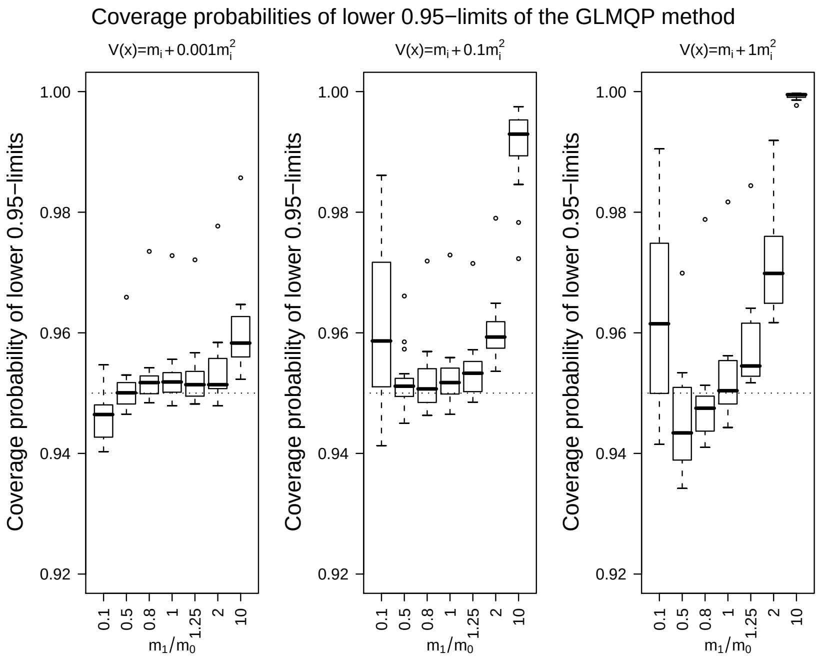

Figure 7: Coverage probabilities of nominal 95% lower confidence limits of the *GLMQP* method, in dependence of the true ratio $\rho = m_1/m_0$

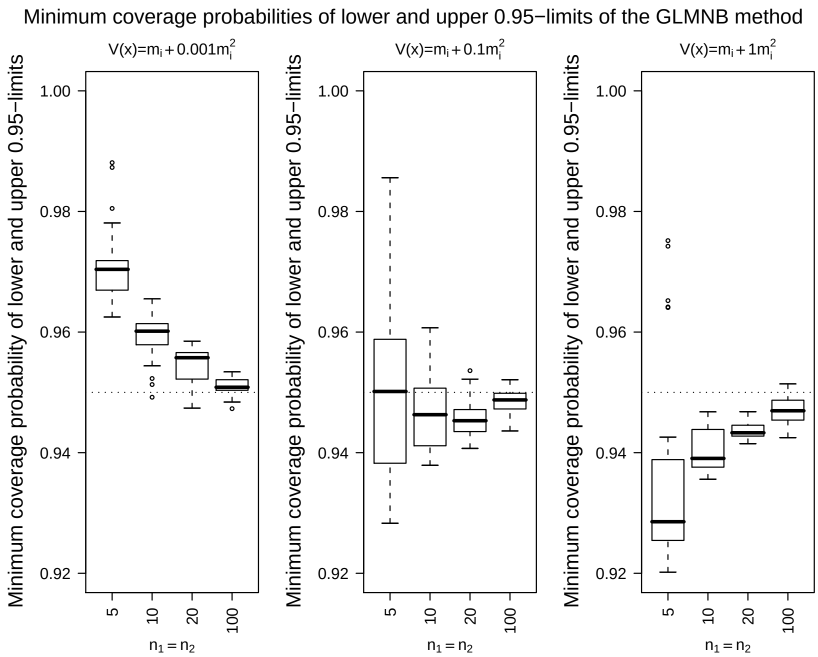

Figure 8: Minimal coverage probabilities of nominal 95% lower and upper confidence limits of the *GLMNB* method, in dependence of $n_0 = n_1$

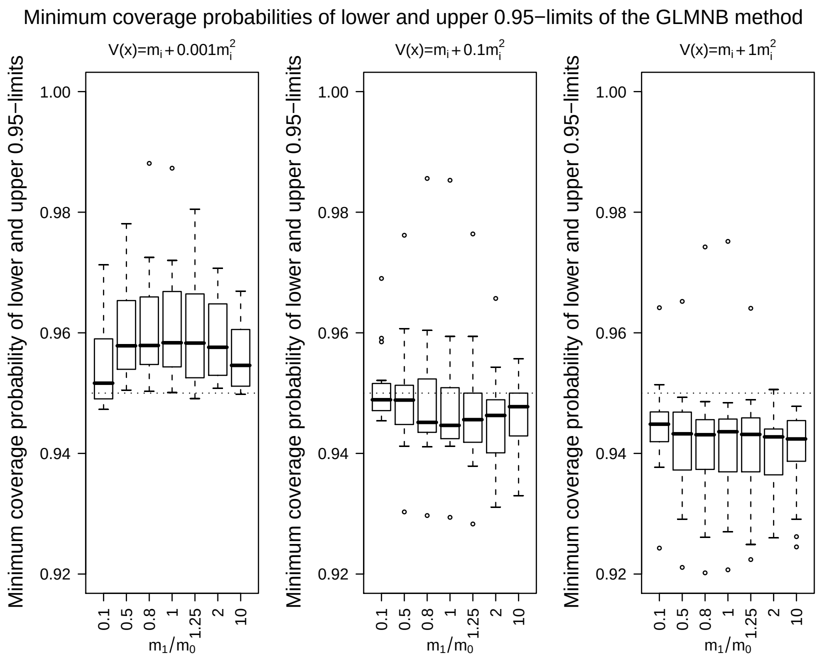

Figure 9: Minimal coverage probabilities of nominal 95% lower and upper confidence limits of the *GLMNB* method, in dependence of the true ratio $\rho = m_1/m_0$

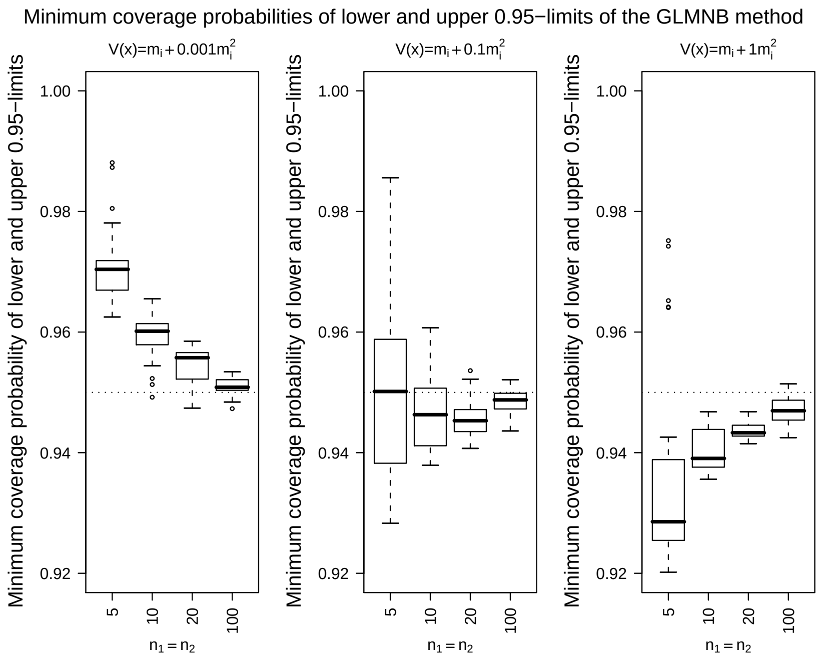

Figure 10: Minimal coverage probabilities of nominal 95% lower and upper confidence limits of the *GLMNB* method, in dependence of $n_0 = n_1$

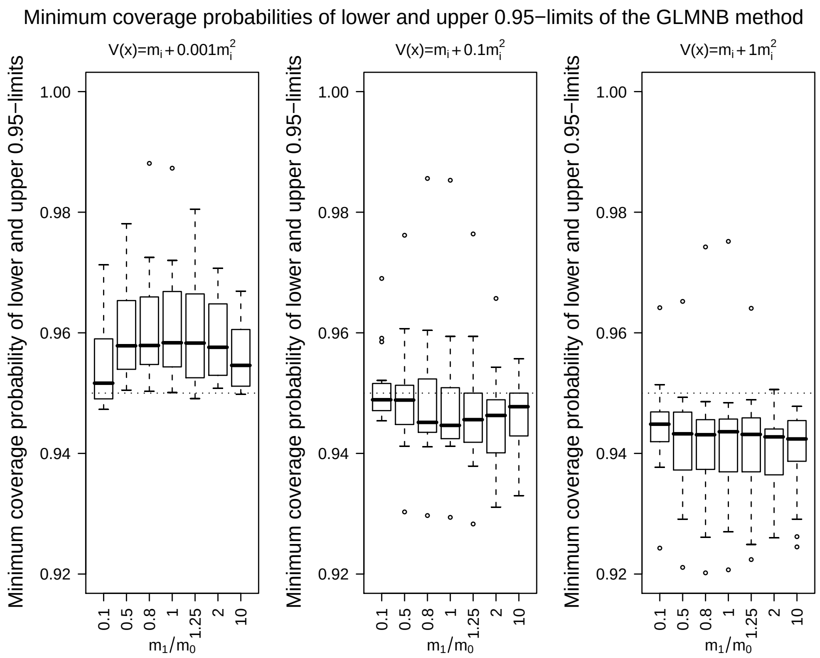

Figure 11: Minimal coverage probabilities of nominal 95% lower and upper confidence limits of the *GLMNB* method, in dependence of the true ratio $\rho = m_1/m_0$

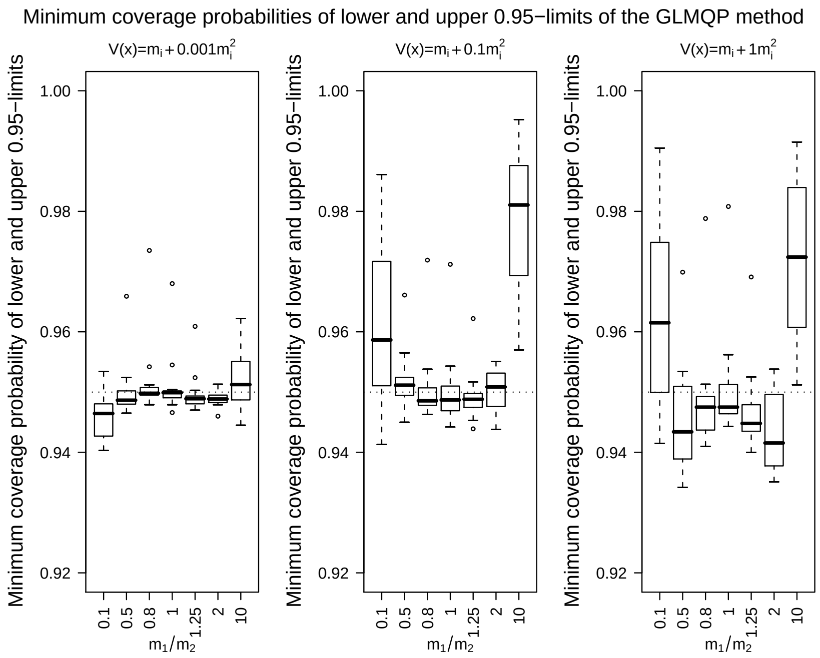

Figure 12: Coverage probabilities of nominal 95% lower confidence limits of the *GLMQP* method, in dependence of the true ratio $\rho = m_1/m_0$

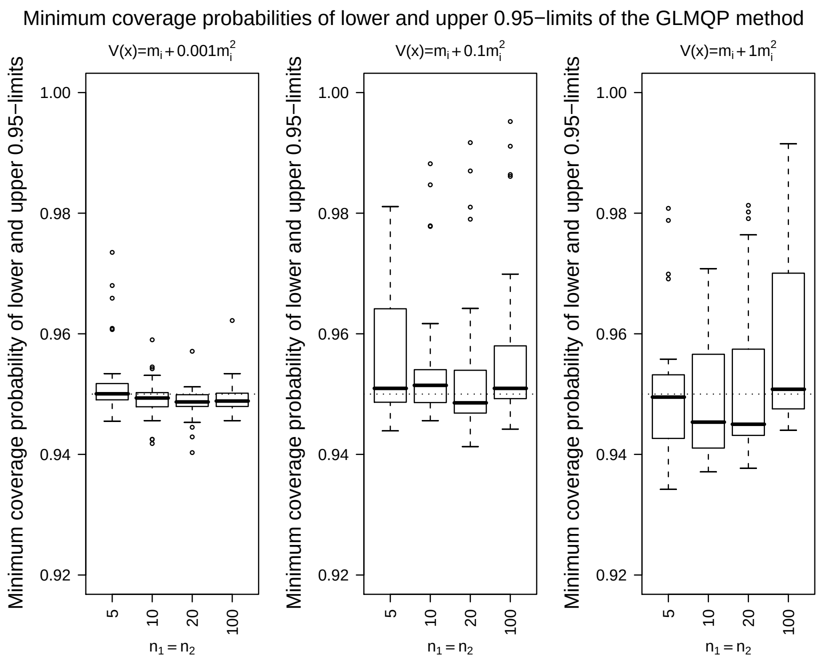

Figure 13: Minimal coverage probabilities of nominal 95% lower and upper confidence limits of the *GLMQP* method, in dependence of $n_0 = n_1$

4.3 Comparison of two algorithms to fit the negative binomial distribution

Two algorithms of Venables and Ripley (2002) (*GLMNB*) and Rigby and Stasinopoulos (2005) (*GLMNBI*) were available for estimation of means and overdispersion parameters of the negative binomial distribution. Figures 14, 15, and 16 show, that using the algorithm of Rigby and Stasinopoulos (2005) (*GLMNBI*) leads to about equal or slightly higher coverage probability than using the implementation of Venables and Ripley (2002) (*GLMNB*) in all considered cases. Especially, for the situations with high overdispersion, where the confidence intervals generally are too liberal, *GLMNBI* is markedly less liberal than *GLMNB*.

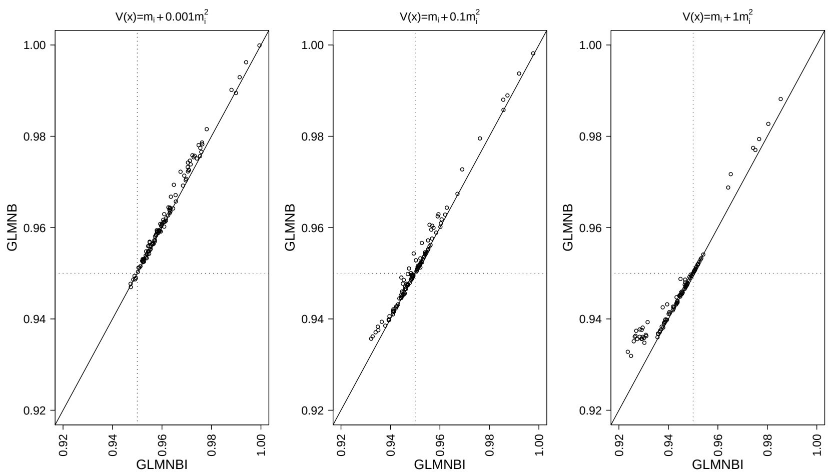

Figure 14: Comparison of coverage probabilities of nominal 95% lower confidence limits of the *GLMNB* and the *GLMNBI* method

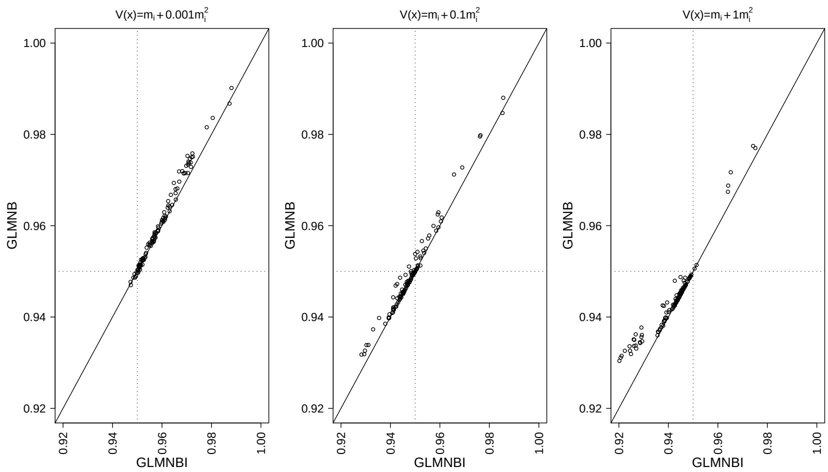

Figure 15: Minimal coverage probabilities of nominal 95% lower and upper confidence limits of the *GLMNB* and the *GLMNBI* method

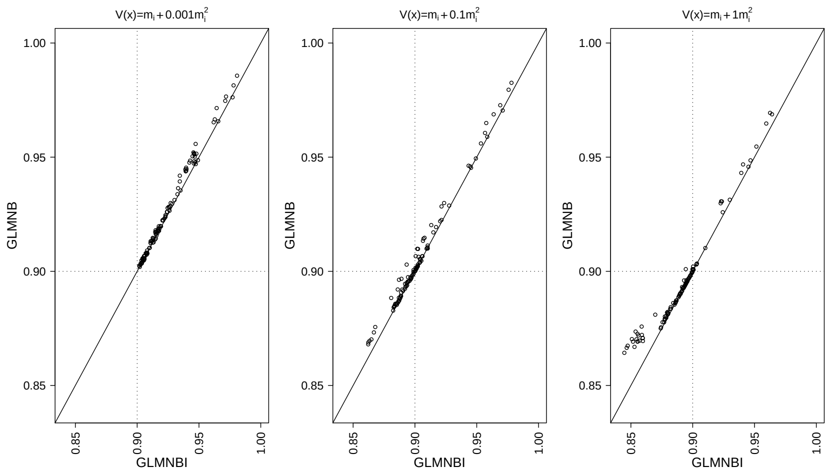

Figure 16: Coverage probabilities of nominal 90% two-sided confidence limits of the *GLMNB* and the *GLMNBI* method

4.4 Detailed results for the *HL* method

The performance of the *HL* method is complicated to describe. For lower 95% limits, severe violations do not occur if $m_0 = 50$, $\rho = 1$, OR! $n = 5$. Severe violations of the nominal levels may occur when $m_0 = 1, 5$, $\rho = 0.5, 0.8, 2, 10$, $n = 20, 100$, but they do not occur in the majority of these situations.

	C_l	m_1	m_0	$n_0 = n_1$	k	ρ
1	0.239	0.50	1	100	1000	0.50
2	0.133	0.80	1	100	1000	0.80
3	0.851	10.00	1	100	1000	10.00
4	0.783	2.50	5	100	1000	0.50
5	0.924	4.00	5	100	1000	0.80
6	0.856	50.00	5	100	1000	10.00
7	0.922	5.00	10	100	1000	0.50
8	0.924	20.00	10	100	1000	2.00
9	0.871	100.00	10	100	1000	10.00
10	0.914	500.00	50	100	1000	10.00
11	0.201	0.50	1	100	10	0.50
12	0.119	0.80	1	100	10	0.80
13	0.776	10.00	1	100	10	10.00
14	0.821	2.50	5	100	10	0.50
15	0.883	50.00	5	100	10	10.00
16	0.912	5.00	10	100	10	0.50
17	0.903	100.00	10	100	10	10.00
18	0.555	0.10	1	100	1	0.10
19	0.037	0.50	1	100	1	0.50
20	0.026	0.80	1	100	1	0.80
21	0.529	10.00	1	100	1	10.00
22	0.827	2.50	5	100	1	0.50
23	0.878	4.00	5	100	1	0.80
24	0.896	50.00	5	100	1	10.00
25	0.906	5.00	10	100	1	0.50
26	0.925	100.00	10	100	1	10.00

Table 1: Situations for which the coverage probability of lower 95% confidence limits of the *HL* methods was lower than 0.93 and $n_0 = n_1 = 5, 10, 20$

	C_l	m_1	m_0	$n_0 = n_1$	k	ρ
1	0.239	0.50	1	100	1000	0.50
2	0.133	0.80	1	100	1000	0.80
3	0.851	10.00	1	100	1000	10.00
4	0.783	2.50	5	100	1000	0.50
5	0.924	4.00	5	100	1000	0.80
6	0.856	50.00	5	100	1000	10.00
7	0.922	5.00	10	100	1000	0.50
8	0.924	20.00	10	100	1000	2.00
9	0.871	100.00	10	100	1000	10.00
10	0.914	500.00	50	100	1000	10.00
11	0.201	0.50	1	100	10	0.50
12	0.119	0.80	1	100	10	0.80
13	0.776	10.00	1	100	10	10.00
14	0.821	2.50	5	100	10	0.50
15	0.883	50.00	5	100	10	10.00
16	0.912	5.00	10	100	10	0.50
17	0.903	100.00	10	100	10	10.00
18	0.555	0.10	1	100	1	0.10
19	0.037	0.50	1	100	1	0.50
20	0.026	0.80	1	100	1	0.80
21	0.529	10.00	1	100	1	10.00
22	0.827	2.50	5	100	1	0.50
23	0.878	4.00	5	100	1	0.80
24	0.896	50.00	5	100	1	10.00
25	0.906	5.00	10	100	1	0.50
26	0.925	100.00	10	100	1	10.00

Table 2: Situations for which the coverage probability of lower 95% confidence limits of the *HL* methods was lower than 0.93 and $n_0 = n_1 = 100$

4.5 Detailed results for the *LN* method

In the overview, the *LN* method showed a conservative performance for a large proportion of the situations, but severely liberal performance for a small proportion of situations. The Figures 17, 18, and 19 display, that severe violations of the nominal confidence level occur mainly in situations, where the mean abundance in the control group m_0 is low, hence the overall mean abundance is low. In cases with high mean abundance in both treatments, the *LN* method does not show coverage probabilities which are markedly lower than the nominal confidence level. The Tables 4.5, 4.5, and 4.5, show the parameter settings, for which the coverage probability of lower 95% limits (Table 4.5) and the minimum coverage probability of lower and upper 95% limits (Tables 4.5 and 4.5) was lower than 0.93. Violations occur for $m_0 = 1$, and mainly for large or small values of $rho = 0.1, 10$, but never for intermediate values of $\rho = 1$. If $m_0 > 1$ (Table 4.5), violations mainly occur for low mean abundances in the sample 1, i.e. in situtions where $m_1 = 0.1, 0.5, 1$, or in situations where ρ has extreme values, i.e. $\rho = 0.1, 10$.

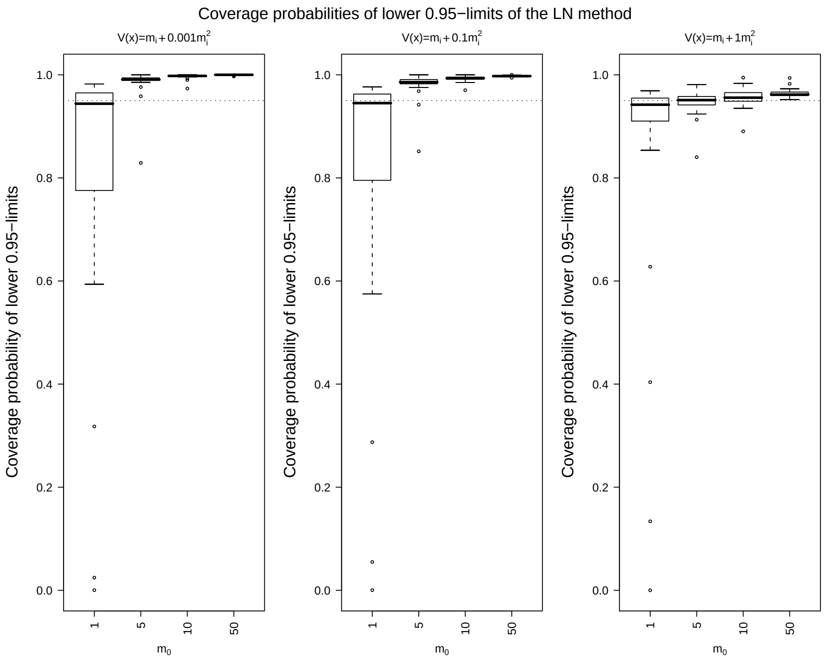

Figure 17: Coverage probabilities of nominal 95% lower confidence limits of the *LN* method, in dependence of m_1

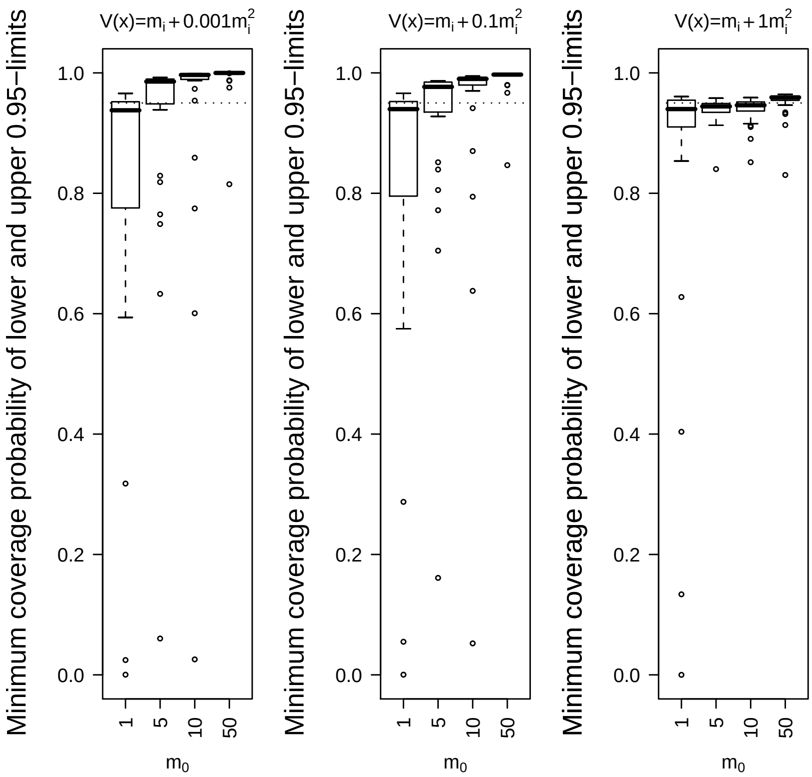

Figure 18: Minimum coverage probabilities of nominal 95% lower and upper confidence limits of the *LN* method, in dependence of m_1

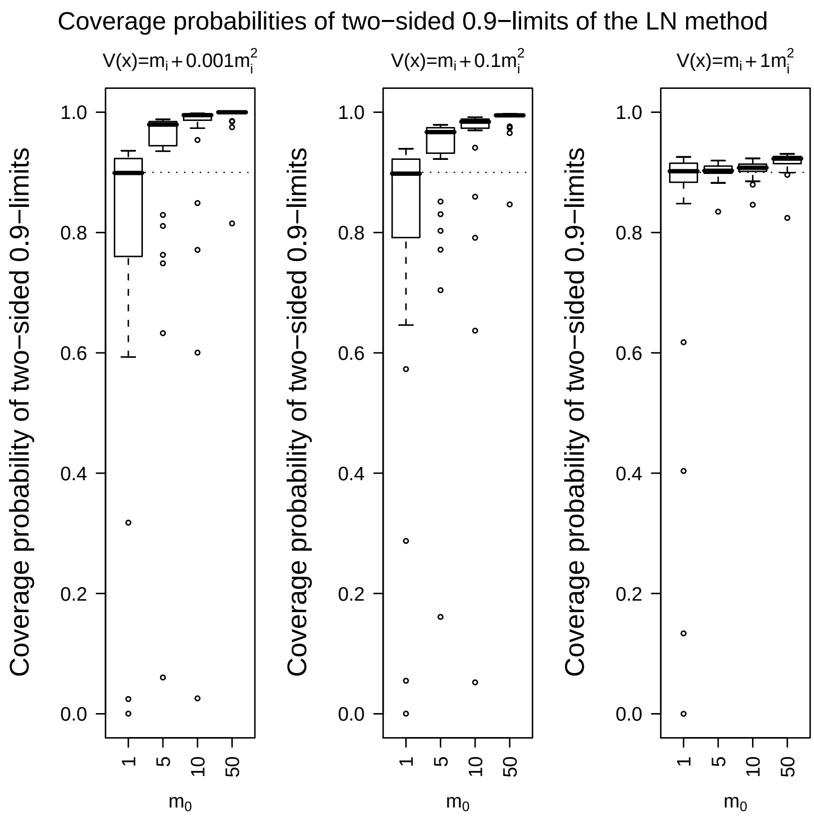

Figure 19: Minimum coverage probabilities of nominal 95% lower and upper confidence limits of the *LN* method, in dependence of m_1

	C_l	m_1	m_0	$n_0 = n_1$	k	ρ
1	0.774	0.10	1	5	1000	0.10
2	0.603	0.10	1	10	1000	0.10
3	0.318	0.10	1	20	1000	0.10
4	0.000	0.10	1	100	1000	0.10
5	0.922	2.00	1	5	1000	2.00
6	0.912	2.00	1	10	1000	2.00
7	0.900	2.00	1	20	1000	2.00
8	0.726	2.00	1	100	1000	2.00
9	0.863	10.00	1	5	1000	10.00
10	0.777	10.00	1	10	1000	10.00
11	0.594	10.00	1	20	1000	10.00
12	0.025	10.00	1	100	1000	10.00
13	0.829	50.00	5	100	1000	10.00
14	0.755	0.10	1	5	10	0.10
15	0.575	0.10	1	10	10	0.10
16	0.287	0.10	1	20	10	0.10
17	0.000	0.10	1	100	10	0.10
18	0.921	2.00	1	5	10	2.00
19	0.915	2.00	1	10	10	2.00
20	0.906	2.00	1	20	10	2.00
21	0.789	2.00	1	100	10	2.00
22	0.869	10.00	1	5	10	10.00
23	0.801	10.00	1	10	10	10.00
24	0.647	10.00	1	20	10	10.00
25	0.055	10.00	1	100	10	10.00
26	0.852	50.00	5	100	10	10.00
27	0.628	0.10	1	5	1	0.10
28	0.404	0.10	1	10	1	0.10
29	0.134	0.10	1	20	1	0.10
30	0.000	0.10	1	100	1	0.10
31	0.929	0.50	1	20	1	0.50
32	0.861	0.50	1	100	1	0.50
33	0.912	10.00	1	5	1	10.00
34	0.914	10.00	1	10	1	10.00
35	0.908	10.00	1	20	1	10.00
36	0.854	10.00	1	100	1	10.00
37	0.924	10.00	5	100	1	2.00
38	0.930	50.00	5	10	1	10.00
39	0.913	50.00	5	20	1	10.00
40	0.840	50.00	5	100	1	10.00
41	0.890	100.00	10	100	1	10.00

Table 3: Situations for which the coverage probability of lower 95% confidence limits of the *LN* methods was lower than 0.93

	C_{min}	m_1	m_0	$n_0 = n_1$	k	ρ
1	0.774	0.10	1	5	1000	0.10
2	0.603	0.10	1	10	1000	0.10
3	0.318	0.10	1	20	1000	0.10
4	0.000	0.10	1	100	1000	0.10
5	0.922	2.00	1	5	1000	2.00
6	0.912	2.00	1	10	1000	2.00
7	0.900	2.00	1	20	1000	2.00
8	0.726	2.00	1	100	1000	2.00
9	0.863	10.00	1	5	1000	10.00
10	0.777	10.00	1	10	1000	10.00
11	0.594	10.00	1	20	1000	10.00
12	0.025	10.00	1	100	1000	10.00
13	0.755	0.10	1	5	10	0.10
14	0.575	0.10	1	10	10	0.10
15	0.287	0.10	1	20	10	0.10
16	0.000	0.10	1	100	10	0.10
17	0.921	2.00	1	5	10	2.00
18	0.915	2.00	1	10	10	2.00
19	0.906	2.00	1	20	10	2.00
20	0.789	2.00	1	100	10	2.00
21	0.869	10.00	1	5	10	10.00
22	0.801	10.00	1	10	10	10.00
23	0.647	10.00	1	20	10	10.00
24	0.055	10.00	1	100	10	10.00
25	0.628	0.10	1	5	1	0.10
26	0.404	0.10	1	10	1	0.10
27	0.134	0.10	1	20	1	0.10
28	0.000	0.10	1	100	1	0.10
29	0.929	0.50	1	20	1	0.50
30	0.861	0.50	1	100	1	0.50
31	0.912	10.00	1	5	1	10.00
32	0.914	10.00	1	10	1	10.00
33	0.908	10.00	1	20	1	10.00
34	0.854	10.00	1	100	1	10.00

Table 4: Situations with $m_0 = 1$ for which the minimum coverage probability of lower and upper 95% confidence limits of the *LN* methods was lower than 0.93

	C_{min}	m_1	m_0	$n_0 = n_1$	k	ρ
1	0.819	0.50	5	5	1000	0.10
2	0.765	0.50	5	10	1000	0.10
3	0.633	0.50	5	20	1000	0.10
4	0.060	0.50	5	100	1000	0.10
5	0.749	2.50	5	100	1000	0.50
6	0.829	50.00	5	100	1000	10.00
7	0.859	1.00	10	5	1000	0.10
8	0.775	1.00	10	10	1000	0.10
9	0.601	1.00	10	20	1000	0.10
10	0.026	1.00	10	100	1000	0.10
11	0.815	5.00	50	100	1000	0.10
12	0.840	0.50	5	5	10	0.10
13	0.805	0.50	5	10	10	0.10
14	0.705	0.50	5	20	10	0.10
15	0.161	0.50	5	100	10	0.10
16	0.928	2.50	5	20	10	0.50
17	0.772	2.50	5	100	10	0.50
18	0.852	50.00	5	100	10	10.00
19	0.870	1.00	10	5	10	0.10
20	0.794	1.00	10	10	10	0.10
21	0.638	1.00	10	20	10	0.10
22	0.052	1.00	10	100	10	0.10
23	0.847	5.00	50	100	10	0.10
24	0.921	0.50	5	5	1	0.10
25	0.921	2.50	5	100	1	0.50
26	0.924	10.00	5	100	1	2.00
27	0.930	50.00	5	10	1	10.00
28	0.913	50.00	5	20	1	10.00
29	0.840	50.00	5	100	1	10.00
30	0.912	1.00	10	5	1	0.10
31	0.916	1.00	10	10	1	0.10
32	0.910	1.00	10	20	1	0.10
33	0.852	1.00	10	100	1	0.10
34	0.920	5.00	10	100	1	0.50
35	0.890	100.00	10	100	1	10.00
36	0.913	5.00	50	20	1	0.10
37	0.830	5.00	50	100	1	0.10

Table 5: Situations with $m_0 > 1$ for which the minimum coverage probability of lower and upper 95% confidence limits of the *LN* methods was lower than 0.93

5 Discussion

6 References

References

Fieller, EC (1954): Some problems in interval estimation. Journal of the Royal Statistical Society, Series B, 16: 175-185.

Tamhane, AC, and Logan, BR (2004): Finding the maximum save dose level for heteroscedastic data. Journal of Biopharmaceutical Statistics 14, 843-856.

Hothorn, T., and Munzel, U. (2002): Exact Nonparametric Confidence Intervals for the Ratio of Medians. Technical Report, Institut für Medizininformatik, Biometrie und Epidemiologie, Friedrich Alexander Universität Erlangen-Nürnberg.

Bauer, D.F. (1972): Constructing confidence sets using rank statistics. Journal of the American Statistical Association 67, 687-690.

Hodges, J.L., and Lehmann, E.L. (1963): Estimates of Location Based on Rank Tests. Annals of Mathematical Statistics 34, 598-611.

Chen, Y-H, Zhou, X-H (2006): Interval estimates for the ratio and the difference of two lognormal means. Statistics in Medicine 25, 4099-4113.

McCullagh, and Nelder (1989): Generalized Linear Models. Chapman & Hall / CRC, Boca Raton.

Venables, W.N. and Ripley, B.D. (2002): Modern Applied Statistics with S. Springer New York.

Rigby, R.A. and Stasinopoulos D.M. (2005). Generalized additive models for location, scale and shape. Applied Statististics 54, 507-554.

Wellek, S. (2003): Testing statistical hypotheses of equivalence. Chapman & Hall / CRC, Boca Raton.