On Confidence Intervals for a Single Proportion in the Analysis of CTC Images Data

Krishna Saha* Daniel Miller†

Abstract
In the estimation of sensitivity and specificity in diagnostic medicine when a patient or the diagnostic unit is considered as a cluster, the confidence interval of a single proportion is frequently used. A number of confidence intervals for a single proportion in the case of non-clustered binary data are proposed in the literature and implemented in standard software packages. However, little attention has been paid to extending such inferences using clustered binary data. In this study, we consider several asymptotic procedures, based on parametric and semiparametric models, to construct the confidence interval for a single proportion based on binary outcome data arising in cluster studies. We compare the performance of the five proposed methods, in terms of coverage and expected lengths, with a Monte Carlo simulation study and we illustrate the methodology with an example from the CTC images study.

Key Words: beta-binomial, clustered binary data, confidence interval, CTC images data, proportion

1. Introduction

In the diagnostic accuracy of CTC angiography (Zhou, 2002), interval estimation of a single proportion is frequently used to estimate sensitivity (true positive rate) and specificity (true negative rate) at the patient level, at the coronary artery level, and at the coronary artery segment level. For the data of CTC angiography given in Table 1, the observed variance in the estimated proportion of the colon polyps per patient is 0.1156, while the predicted variance obtained using a binomial model is 0.0834. This concludes that the observed variance is 1.39 times larger than the predicted variance. The responses within the same cluster are correlated and due to this intraclass correlation, the observed variance of these data exhibit greater or smaller binomial variability. The standard approaches of analyzing such data that ignore the cluster structure may result in under-estimation of the true standard error of the estimated risk rate. Furthermore, confidence intervals of a single proportion based on the binomial model in such data may show lack of coverage. In the simple binomial model case, several alternative methods for the interval estimation of a single proportion have been developed and compared. Newcombe (1998) compared seven interval methods for the binomial proportion and concluded that the Wilson score method produced reasonable intervals for all parameter values. Brown et al. (2001) compared a number of methods and concluded that the Wilson method performs quite well in terms of coverage for π away from 0 or 1 but the interval was unnecessarily long and exceeded that of the Clopper-Pearson interval when π was close to 0 or 1. They also summarized when to use which method as follows: (i) do not use the text-book Wald interval for any π; (ii) confidence intervals for transformations are generally too wide; (iii) the Clopper-Pearson interval is too conservative; (iv) the likelihood ratio interval was too hard to compute for the average user; (v) for not so rare events, use the Wilson score interval or Agresti-Coull; and

*Department of Mathematical Sciences, Central Connecticut State University, 1615 Stanley Street, New Britain, CT 06050, USA

†Department of Mathematical Sciences, Central Connecticut State University, 1615 Stanley Street, New Britain, CT 06050, USA

Table 1: CAD-enhanced CTC results for detection of colon polyps for reader 1

Patient's ID	x_i=number of polyps detected by radiologist	n_i=total number of polyps
1	1	1
2	2	2
3	2	2
4	1	1
5	2	2
6	2	2
7	1	1
8	1	1
9	1	1
10	1	1
11	2	2
12	0	1
13	2	3
14	2	2
15	1	1
16	1	1
17	1	1
18	2	2
19	1	2
20	0	2
21	1	1
22	2	2
23	2	2
24	2	2
25	0	1

(vi) for rare events, use the Wilson score interval. Although there has been much literature pertaining to the binomial model, little attention has been paid to developing confidence interval procedures for a single proportion based on clustered binary data.

For complex survey data, Korn and Graubard (1998) developed four alternative methods by estimating the variance based on the design effect as well as the effective sample size. They concluded that the modified Wald interval did not work well for rare events and the modified logit-transform was too wide, and they recommended the modified Clopper-Pearson interval using the effective sample size adjusted for the survey's nominal degrees of freedom. However, these proposed methods show lack of coverage for small sample sizes or an extremely small/large estimated proportion (see, Sukasih and Jang, 2005; Chen and Tipping, 2002). Lui (2004) discussed two versions of Wald-type intervals and the Wilson score interval based on the same variance in a beta-binomial model and using the ANOVA estimate of the intraclass correlation. However, the performance coverage properties of these methods were not investigated nor were the methods compared with other recommended approaches in the literature. Although several works in the literature investigated the joint estimation of the proportion and intraclass correlation based on parametric and semiparametric model assumptions for clustered binary data, these approaches were not extended. In this paper, we focus on extending these approaches to obtain confidence intervals for a single proportion and investigate the performances through simulation studies for small to moderate sample sizes in terms of coverage and expected lengths.

2. Parametric and Semiparametric Models

It is well known that there is no unique probability function for clustered binary data. Some authors (Altham, 1978; Kupper and Haseman, 1978; Efron 1986) proposed some probability distributions, such as the beta-binomial model, the correlated binomial model, the multiplicative binomial model, and the double binomial model. However, the beta-binomial model is the superior model compared to its competitive models (Paul, 1982; Saha, 2011). Let $x_1, x_2, \ldots, x_k$ be a sample from the beta-binomial distribution with probability mass function

$$P(x_i|n_i, \pi, \phi) \quad = \quad \begin{pmatrix} n_i \\ x_i \end{pmatrix} \frac{\prod_{r=0}^{x_i-1}[(1-\phi)\pi + r\phi] \prod_{r=0}^{n_i-x_i-1}[(1-\pi)(1-\phi) + r\phi]}{\prod_{r=0}^{n_i-1}[(1-\phi) + r\phi]}$$

for $x_i = 0, 1, \ldots, n_i$; $0 \le \pi \le 1$ and $0 < \phi < 1$. The mean and variance of X_i are E$(X_i) = n_i\pi$ and var$(X_i) = n_i\pi(1-\pi)\xi$, where $\xi = 1 + (n_i - 1)\phi$. When $\phi \to 0$, the limiting distribution of $BB(n_i, \pi, \phi)$ becomes Binomial(n_i, π).

An inadequate model assumption for the underlying data distribution may lead to making falsely significant inferences and one must be careful when applying the parametric model assumptions (Nikoloulopoulos and Karlis, 2008). In such cases, a more flexible model is available that only specifies the mean and variance of the data distribution. This model assumption has been widely used in many biomedical applications. See, for example, Breslow (1984), Moore (1986) and Paul and Saha (2007). Let $x_1, x_2, \ldots, x_k$ be a random sample drawn from a population with the same mean and variance as the beta-binomial model, that is, E$(X_i) = n_i\pi$ and Var$(X_i) = n_i\pi(1-\pi)\{1 + (n_i - 1)\phi\}$.

3. The Confidence Interval Methods

3.1 The ML based interval

The kernel of the beta-binomial log-likelihood is given by

$$l(\pi, \phi) \quad = \quad \sum_{i=1}^{k} \left[\sum_{r=0}^{x_i-1} \ln\{(1-\phi)\pi + r\phi\} + \sum_{r=0}^{n_i-x_i-1} \ln\{(1-\pi)(1-\phi) + r\phi\} - \sum_{r=0}^{n_i-1} \ln\{(1-\phi) + r\phi\} \right] \quad (1)$$

Then the ML estimates $\hat{\pi}_{ml}$ and $\hat{\phi}_{ml}$ of π and ϕ are obtained by solving the estimating equations

$$\frac{\partial l}{\partial \pi} = \sum_{i=1}^{k} \left\{ \sum_{r=0}^{x_i-1} \frac{1-\phi}{(1-\phi)\pi + r\phi} - \sum_{r=0}^{n_i-x_i-1} \frac{1-\phi}{(1-\phi)(1-\pi) + r\phi} \right\} = 0$$

and

$$\frac{\partial l}{\partial \phi} = \sum_{i=1}^{k} \left\{ \sum_{r=1}^{x_i-1} \frac{-\pi + r}{(1-\phi)\pi + r\phi} + \sum_{r=0}^{n_i-x_i-1} \frac{-(1-\pi) + r}{(1-\phi)(1-\pi) + r\phi} - \sum_{r=0}^{n_i-1} \frac{r-1}{(1-\phi) + r\phi} \right\} = 0$$

simultaneously. Based on the expected Fisher information matrix, one can easily obtain an asymptotic variance of $\hat{\pi}_{ml}$ as

$$\text{Var}(\hat{\pi}_{ml}) = \frac{\Lambda_{22}}{\Lambda_{11}\Lambda_{22} - \Lambda_{12}^2}, \quad (2)$$

where the expressions of the quantities, Λ_{11}, Λ_{12}, and Λ_{22} are given by

$$\begin{aligned}
\Lambda_{11} &= (1-\phi)^2 \sum_{i=1}^{m} \left(A_{1i}^{(2,0)} + A_{2i}^{(2,0)} \right), \\
\Lambda_{12} &= \sum_{i=1}^{m} \left(A_{1i}^{(2,2)} + A_{2i}^{(2,2)} - A_{3i}^{(2,2)} \right), \\
\Lambda_{22} &= (1-\phi) \sum_{i=1}^{m} \left(A_{1i}^{(2,1)} - A_{2i}^{(2,1)} \right) + \sum_{i=1}^{m} \left(A_{1i}^{(1,0)} - A_{2i}^{(1,0)} \right),
\end{aligned}$$

with, for $i = 1, ..., m$; $p = q = 0, 1, 2$,

$$\begin{aligned}
A_{1i}^{(p,q)} &= \sum_{j=1}^{n_i} \frac{(j-\pi-1)^q}{E_{j-1}^p} Pr(X_i \geq j), \\
A_{2i}^{(p,q)} &= \sum_{j=1}^{n_i} \frac{(j+\pi-2)^q}{F_{j-1}^p} Pr(X_i \leq n_i - j), \\
A_{3i}^{(p,q)} &= \sum_{j=1}^{n_i} \frac{(j-2)^q}{G_{j-1}^p},
\end{aligned}$$

where $E_j = (1-\phi)\pi + j\phi$, $F_j = (1-\phi)(1-\pi) + j\phi$, and $G_j = 1 + j\phi$. An approximate 100(1 - α)% confidence interval for π, based on ML, is then given by

$$\hat{\pi}_{ml} - z_{\alpha/2}\sqrt{\widehat{\text{Var}}(\hat{\pi}_{ml})} \leq \pi \leq \hat{\pi}_{ml} + z_{\alpha/2}\sqrt{\widehat{\text{Var}}(\hat{\pi}_{ml})},$$

where $z_{\alpha/2}$ is the100(1- $\alpha/2$)th percentile of the standard normal distribution and $\widehat{\text{Var}}(\hat{\pi}_{ml})$ is the estimated variance of $\hat{\pi}_{ml}$ obtained from (2) by replacing the parameters π and ϕ by $\hat{\pi}_{ml}$ and $\hat{\phi}_{ml}$, respectively.

3.2 The EQL based interval

Based on the above semiparametric model, the kernel of the extended quasi-likelihood can be obtained following Nelder and Pregibon (1987) as

$$Q^+ = \sum_{i=1}^{k} \left[-\frac{1}{2} \ln \{1 + (n_i - 1)\phi\} + \frac{x_i \ln\left(\frac{n_i \pi}{x_i}\right) + (n_i - x_i) \ln\left(\frac{n_i\{1-\pi\}}{n_i - x_i}\right)}{1 + (n_i - 1)\phi} \right].$$

Then the EQL estimates $\hat{\pi}_{eql}$ and $\hat{\phi}_{eql}$ can be obtained by solving the estimating equations

$$\sum_{i=1}^{m} (x_i - n_i\pi)[1 + (n_i - 1)\phi]^{-1} = 0$$

and

$$\sum_{i=1}^{m} \frac{(n_i - 1)}{[1 + (n_i - 1)\phi]^2} \left[x_i \ln\left(\frac{x_i}{n_i\pi}\right) + (n_i - x_i) \ln\left(\frac{n_i - x_i}{n_i\{1-\pi\}}\right) - \frac{1}{2}\{1 + (n_i - 1)\phi\} \right] = 0,$$

simultaneously. Using the observed Fisher information matrix, an asymptotic variance of $\hat{\pi}_{eql}$ is obtained by

$$\text{Var}(\hat{\pi}_{eql}) = \frac{\Upsilon_{22}}{\Upsilon_{11}\Upsilon_{22} - \Upsilon_{12}^2}, \tag{3}$$

Table 2: Coverage probability estimates (expected interval lengths) based on confidence intervals by the methods with nominal level, $1 - \alpha$ = 95% and fixed litter sizes for sample size $k = 19$.

ϕ	π	ML	EQL	DEQL	QEE	WA
0.1	0.05	0.986 (0.122)	0.987 (0.132)	0.977 (0.157)	0.988 (0.191)	0.987 (0.104)
	0.10	0.984 (0.139)	0.994 (0.143)	0.991 (0.166)	0.987 (0.211)	0.975 (0.136)
	0.15	0.959 (0.161)	0.973 (0.171)	0.978 (0.188)	0.967 (0.204)	0.957 (0.160)
	0.20	0.949 (0.177)	0.966 (0.198)	0.972 (0.193)	0.953 (0.188)	0.948 (0.175)
	0.25	0.947 (0.192)	0.975 (0.238)	0.967 (0.208)	0.945 (0.195)	0.936 (0.190)
	0.30	0.956 (0.201)	0.980 (0.251)	0.972 (0.221)	0.960 (0.205)	0.944 (0.200)
	0.35	0.957 (0.209)	0.984 (0.273)	0.973 (0.237)	0.954 (0.210)	0.930 (0.207)
	0.40	0.961 (0.218)	0.989 (0.295)	0.971 (0.251)	0.963 (0.219)	0.929 (0.214)
	0.45	0.939 (0.223)	0.992 (0.317)	0.955 (0.276)	0.934 (0.224)	0.939 (0.216)
	0.50	0.933 (0.230)	0.991 (0.332)	0.954 (0.290)	0.929 (0.231)	0.946 (0.216)
0.3	0.05	0.992 (0.147)	0.991 (0.153)	0.972 (0.199)	0.990 (0.165)	0.971 (0.130)
	0.10	0.991 (0.167)	0.998 (0.170)	0.998 (0.209)	0.993 (0.245)	0.959 (0.174)
	0.15	0.936 (0.192)	0.942 (0.202)	0.971 (0.238)	0.951 (0.295)	0.948 (0.206)
	0.20	0.919 (0.215)	0.931 (0.238)	0.949 (0.264)	0.933 (0.371)	0.939 (0.229)
	0.25	0.925 (0.235)	0.943 (0.275)	0.954 (0.285)	0.939 (0.261)	0.944 (0.248)
	0.30	0.921 (0.250)	0.948 (0.314)	0.954 (0.301)	0.933 (0.272)	0.928 (0.261)
	0.35	0.933 (0.261)	0.958 (0.347)	0.959 (0.313)	0.935 (0.278)	0.924 (0.273)
	0.40	0.948 (0.270)	0.964 (0.376)	0.959 (0.322)	0.947 (0.273)	0.933 (0.281)
	0.45	0.950 (0.275)	0.974 (0.409)	0.975 (0.327)	0.953 (0.279)	0.921 (0.285)
	0.50	0.926 (0.279)	0.982 (0.437)	0.961 (0.331)	0.931 (0.282)	0.932 (0.288)
0.5	0.05	0.995 (0.172)	0.995 (0.176)	0.986 (0.225)	0.996 (0.195)	0.962 (0.156)
	0.10	0.986 (0.195)	0.994 (0.198)	0.994 (0.245)	0.990 (0.266)	0.940 (0.210)
	0.15	0.919 (0.222)	0.923 (0.235)	0.964 (0.275)	0.940 (0.326)	0.937 (0.247)
	0.20	0.904 (0.249)	0.911 (0.275)	0.941 (0.306)	0.919 (0.335)	0.927 (0.275)
	0.25	0.906 (0.273)	0.919 (0.315)	0.950 (0.335)	0.922 (0.322)	0.932 (0.297)
	0.30	0.917 (0.293)	0.936 (0.355)	0.962 (0.357)	0.931 (0.315)	0.928 (0.316)
	0.35	0.941 (0.306)	0.952 (0.390)	0.973 (0.373)	0.949 (0.322)	0.934 (0.330)
	0.40	0.945 (0.315)	0.966 (0.416)	0.971 (0.383)	0.956 (0.322)	0.936 (0.338)
	0.45	0.924 (0.321)	0.957 (0.437)	0.965 (0.389)	0.933 (0.328)	0.927 (0.344)
	0.50	0.922 (0.325)	0.962 (0.453)	0.962 (0.394)	0.923 (0.341)	0.941 (0.346)

where the elements of the observed Fisher information matrix, Υ_{11}, Υ_{12}, and Υ_{22}, are given by

$$\begin{aligned}
\Upsilon_{11} &= -\frac{\partial^2 Q^+}{\partial \pi^2} = \frac{1}{\pi^2(1-\pi)^2}\sum_{i=1}^{m}\frac{2\pi x_i - x_i - n_i\pi^2}{1+(n_i-1)\phi},\\
\Upsilon_{12} &= H_{21}^E = -\frac{\partial^2 Q^+}{\partial \pi \partial \phi} = \frac{1}{\pi(1-\pi)}\sum_{i=1}^{m}\frac{(\pi n_i - x_i)(n_i-1)}{[1+(n_i-1)\phi]^2}, \text{ and}\\
\Upsilon_{22} &= -\frac{\partial^2 Q^+}{\partial \phi^2} = \sum_{i=1}^{m}\left[\frac{2(n_i-1)^2}{[1+(n_i-1)\phi]^3}\left\{x_i \ln\left(\frac{n_i x_i}{\pi}\right) + (n_i - x_i)\ln\left(\frac{n_i(1-\pi)}{n_i - x_i}\right)\right\} - \frac{(n_i-1)^2}{2[1+(n_i-1)\phi]^2}\right].
\end{aligned}$$

The approximate 100(1 - α)% confidence interval of π based on EQL is given by

$$\hat{\pi}_{eql} - z_{\alpha/2}\sqrt{\widehat{\text{Var}}(\hat{\pi}_{eql})} \le \pi \le \hat{\pi}_{eql} + z_{\alpha/2}\sqrt{\widehat{\text{Var}}(\hat{\pi}_{eql})},$$

where $\widehat{\text{Var}}(\hat{\pi}_{eql})$ is the estimated variance of $\hat{\pi}_{eql}$ obtained from (3) after replacing the parameters π and ϕ by $\hat{\pi}_{eql}$ and $\hat{\phi}_{eql}$, respectively.

3.3 The DEQL based interval

From Paul and Saha (2007), we obtain the kernel of the profile double extended quasi-likelihood which is given by

$$\begin{aligned}
p_v(Q) = \sum_{i=1}^{k}\Bigg[&\left(x_i + \frac{\pi}{\delta} - \frac{1}{2}\right)\ln\left(x_i + \frac{\pi}{\delta}\right) + \left(n_i - x_i + \frac{1-\pi}{\delta} - \frac{1}{2}\right)\ln\left(n_i - x_i + \frac{1-\pi}{\delta}\right)\\
&- \left(n_i + \frac{1}{\delta} - \frac{1}{2}\right)\ln\left(n_i + \frac{1}{\delta}\right) + \frac{\delta}{12(\pi + \delta x_i)} + \frac{\delta}{12\{1-\pi+\delta(n_i - x_i)\}} - \frac{\delta}{12(1+\delta n_i)}\\
&- \left(\frac{\pi}{\delta} + \frac{1}{2}\right)\ln\left(\frac{\pi}{\delta}\right) - \left(\frac{1-\pi}{\delta} - \frac{1}{2}\right)\ln[(1-\pi)/\delta] + \left(\frac{1}{\delta} + \frac{1}{2}\right)\ln(1/\delta) - \frac{\delta}{12\pi}\\
&- \frac{\delta}{12(1-\pi)} + \frac{\delta}{12}\Bigg],
\end{aligned}$$

with $\delta = \phi(1-\phi)^{-1}$. Then the DEQL estimates $\hat{\pi}_{de}$ and $\hat{\delta}_{de}$ are obtained by solving

$$\sum_{i=1}^{k}\left[\frac{1}{\delta}\ln\left(\frac{P_{1i}(1-\pi)}{\pi P_{2i}}\right) + \frac{P_{1i} - P_{2i}}{2P_{1i}P_{2i}}\left(1 + \frac{\delta P_{3i}}{6P_{1i}P_{2i}}\right) + \frac{2\pi - 1}{2\pi(1-\pi)}\left(1 + \frac{\delta}{6\pi(1-\pi)}\right)\right] = 0$$

and

$$\begin{aligned}
\sum_{i=1}^{k}\Bigg[&\frac{\pi}{\delta^2}\ln\left(\frac{\pi P_{2i}}{P_{1i}(1-\pi)}\right) + \frac{1}{\delta}\ln\left(\frac{(1-\pi)P_{3i}}{P_{2i}}\right) + \frac{\pi}{2P_{1i}}\left(\frac{1}{\delta} + \frac{1}{6P_{1i}}\right)\\
&+ \frac{1-\pi}{2P_{2i}}\left(\frac{1}{\delta} + \frac{1}{6P_{2i}}\right) - \frac{1}{2P_{3i}}\left(\frac{1}{\delta} + \frac{1}{6P_{3i}}\right) - \frac{1}{2\delta} - \frac{1-\pi(1-\pi)}{12\pi(1-\pi)}\Bigg] = 0,
\end{aligned}$$

where $P_{1i} = \pi + \delta x_i$, $P_{2i} = 1 - \pi + \delta(n_i - x_i)$ and $P_{3i} = 1 + \delta n_i$. The DEQL estimate of ϕ is $\hat{\phi}_{de} = \hat{\delta}_{de}(1 + \hat{\delta}_{de})^{-1}$. Using the observed Fisher information, we obtain the asymptotic variance of $\hat{\pi}_{de}$, which is given by

$$\text{Var}(\hat{\pi}_{de}) = \frac{\Psi_{22}}{\Psi_{11}\Psi_{22} - \Psi_{12}^2}, \tag{4}$$

where Ψ_{11}, Ψ_{12}, and Ψ_{22} are the elements of the observed Fisher information matrix given

Table 3: Coverage probability estimates (expected interval lengths) based on confidence intervals by the methods with nominal level, $1-\alpha$ = 95% and fixed litter sizes for sample size $k = 73$.

ϕ	π	ML	EQL	DEQL	QEE	WA
0.1	0.05	0.951 (0.043)	0.961 (0.053)	0.963 (0.057)	0.096 (0.117)	0.941 (0.043)
	0.10	0.942 (0.058)	0.946 (0.060)	0.957 (0.069)	0.955 (0.134)	0.932 (0.059)
	0.15	0.937 (0.069)	0.957 (0.077)	0.955 (0.080)	0.950 (0.076)	0.943 (0.071)
	0.20	0.945 (0.078)	0.976 (0.096)	0.969 (0.087)	0.942 (0.081)	0.948 (0.079)
	0.25	0.932 (0.085)	0.982 (0.126)	0.953 (0.094)	0.935 (0.085)	0.935 (0.086)
	0.30	0.945 (0.090)	0.996 (0.147)	0.965 (0.098)	0.948 (0.090)	0.950 (0.090)
	0.35	0.939 (0.094)	0.998 (0.162)	0.957 (0.102)	0.938 (0.094)	0.937 (0.094)
	0.40	0.943 (0.096)	1.000 (0.187)	0.955 (0.104)	0.943 (0.096)	0.938 (0.097)
	0.45	0.970 (0.099)	1.000 (0.201)	0.979 (0.107)	0.975 (0.099)	0.951 (0.098)
	0.50	0.938 (0.101)	1.000 (0.227)	0.955 (0.110)	0.939 (0.101)	0.943 (0.098)
0.3	0.05	0.968 (0.061)	0.950 (0.058)	0.852 (0.108)	0.979 (0.070)	0.930 (0.060)
	0.10	0.917 (0.074)	0.915 (0.075)	0.975 (0.096)	0.948 (0.092)	0.939 (0.083)
	0.15	0.925 (0.092)	0.940 (0.100)	0.971 (0.117)	0.942 (0.129)	0.943 (0.100)
	0.20	0.934 (0.105)	0.958 (0.122)	0.974 (0.131)	0.947 (0.122)	0.947 (0.112)
	0.25	0.941 (0.115)	0.970 (0.143)	0.982 (0.142)	0.951 (0.125)	0.953 (0.121)
	0.30	0.936 (0.123)	0.978 (0.167)	0.975 (0.149)	0.942 (0.126)	0.949 (0.128)
	0.35	0.936 (0.129)	0.987 (0.199)	0.969 (0.155)	0.938 (0.131)	0.948 (0.133)
	0.40	0.943 (0.133)	0.989 (0.221)	0.977 (0.159)	0.939 (0.135)	0.949 (0.137)
	0.45	0.967 (0.136)	1.000 (0.257)	0.985 (0.161)	0.965 (0.137)	0.953 (0.140)
	0.50	0.943 (0.137)	0.998 (0.275)	0.977 (0.162)	0.944 (0.138)	0.953 (0.141)
0.5	0.05	0.973 (0.072)	0.957 (0.065)	0.967 (0.112)	0.988 (0.085)	0.935 (0.075)
	0.10	0.909 (0.089)	0.908 (0.088)	0.968 (0.114)	0.944 (0.102)	0.945 (0.103)
	0.15	0.906 (0.109)	0.926 (0.117)	0.966 (0.138)	0.934 (0.131)	0.943 (0.123)
	0.20	0.922 (0.125)	0.954 (0.143)	0.976 (0.157)	0.942 (0.154)	0.954 (0.138)
	0.25	0.942 (0.137)	0.965 (0.167)	0.979 (0.171)	0.954 (0.148)	0.952 (0.149)
	0.30	0.930 (0.147)	0.970 (0.189)	0.979 (0.182)	0.934 (0.155)	0.950 (0.158))
	0.35	0.938 (0.155)	0.984 (0.212)	0.982 (0.190)	0.932 (0.161)	0.956 (0.164)
	0.40	0.956 (0.160)	0.990 (0.237)	0.989 (0.195)	0.955 (0.165)	0.955 (0.169)
	0.45	0.962 (0.163)	0.991 (0.262)	0.990 (0.198)	0.965 (0.168)	0.954 (0.172)
	0.50	0.939 (0.164)	0.988 (0.283)	0.981 (0.199)	0.947 (0.169)	0.952 (0.173)

by

$$\Psi_{11} = -\frac{\partial^2 p_v(Q)}{\partial \pi^2} = \sum_{i=1}^{m}\left[\frac{2P_{1i}-\delta}{2\delta P_{1i}^2} - \frac{2}{\delta P_{1i}} - \frac{2}{\delta P_{2i}} + \frac{2P_{2i}-\delta}{2\delta P_{2i}^2} + \frac{\delta}{6P_{1i}^3} - \frac{\delta}{6P_{2i}^3}\right]$$
$$+m\left[\frac{2}{\pi\delta} - \frac{2\pi-\delta}{2\pi^2\delta} + \frac{2}{(1-\pi)\delta} - \frac{2(1-\pi)-\delta}{2(1-\pi)^2\delta} + \frac{\delta}{6\pi^3} + \frac{1}{6(1-\pi)^3}\right],$$

$$\Psi_{12} = H_{21}^D = -\frac{\partial^2 p_v(Q)}{\partial\pi\partial\delta} = \sum_{i=1}^{m}\left[\frac{1}{\delta^2}\ln\left(\frac{P_{1i}}{\delta}\right) + \frac{2\pi}{\delta^2 P_{1i}} + \frac{2P_{1i}-2\delta}{\delta P_{1i}} - \frac{\pi(2P_{1i}-\delta)}{2\delta^2 P_{1i}^2} - \frac{1}{\delta^2}\ln\left(\frac{P_{2i}}{\delta}\right)\right.$$
$$\left. -\frac{2(1-\pi)}{2\delta^2 P_{1i}} - \frac{2P_{2i}-\delta}{2\delta^2 P_{2i}} + \frac{(1-\pi)(2P_{2i}-\delta)}{2\delta^2 P_{2i}^2} + \frac{\pi}{6P_1 i^3} - \frac{1}{12P_{1i}^2} + \frac{1-\pi}{P_{2i}^3} + \frac{1}{12P_{2i}^2}\right]$$
$$+m\left[\frac{1}{\delta^2}\ln\left(1-\frac{\pi}{\delta}\right) - \frac{1}{\delta}\ln\left(\frac{\pi}{\delta}\right) - \frac{1}{12\pi^2} + \frac{1}{12(1-\pi)^2}\right], \text{ and}$$

$$\Psi_{22} = -\frac{\partial^2 p_v(Q)}{\partial\delta^2} = \sum_{i=1}^{m}\left[-\frac{2\pi^3}{\delta^3 P_{1i}} - \frac{2\pi}{\delta^3}\ln\left(\frac{p_{1i}}{\delta}\right) - \frac{2(1-\pi)}{\delta^3}\ln\left(\frac{p_{2i}}{\delta}\right) + \frac{\pi}{6\delta P_{1i}^2} + \frac{1-\pi}{6\delta P_{2i}^2}\right.$$
$$+\frac{2P_{3i}-\delta}{\delta^3 P_{3i}} + \frac{2}{\delta^3 P_{3i}} + \frac{1}{6\delta P_{3i}^3} + \frac{2}{\delta^3}\ln\left(\frac{p_{3i}}{\delta}\right) - \frac{2P_{3i}-\delta}{2\delta^3 P_{3i}^2} - \frac{\pi^2}{6\delta P_{1i}^3} - \frac{(1-\pi)^2}{6\delta P_{2i}^3} - \frac{2(1-\pi)^2}{\delta^3 P_{2i}}$$
$$\left.-\frac{1}{6\delta P_{3i}^2} - \frac{2\pi(2P_{1i}-\delta)}{2\delta^3 P_{1i}} - \frac{2(1-\pi)(2P_{2i}-\delta)}{2\delta^3 P_{2i}} + \frac{\pi^2(2P_{1i}-\delta)}{2\delta^3 P_{1i}^2} + \frac{(1-\pi)^2(2P_{2i}-\delta)}{2\delta^3 P_{2i}^2}\right]$$
$$+\frac{2m\pi}{\delta^3}\ln\left(\frac{\pi}{\delta}\right) - \frac{2m}{\delta^3}\ln\left(\frac{1}{\delta}\right) - \frac{2m}{\delta^3} + \frac{m(2\pi-\delta)}{2\delta^3} + \frac{2m(1-\pi)-m\delta}{2\delta^3} - \frac{m(2-\delta)}{2\delta^3}$$
$$+\frac{2m\pi}{\delta^3} + \frac{2m(1-\pi)}{\delta^3} + \frac{2m(1-\pi)}{\delta^3}\ln\left(\frac{1-\pi}{\delta}\right),$$

where P_{1i}, P_{2i}, and P_{3i} are defined above. Then, the 100(1 - α)% confidence interval of π, based on the DEQL, is

$$\hat{\pi}_{de} - z_{\alpha/2}\sqrt{\widehat{\text{Var}}(\hat{\pi}_{de})} \leq \pi \leq \hat{\pi}_{de} + z_{\alpha/2}\sqrt{\widehat{\text{Var}}(\hat{\pi}_{de})},$$

where $\widehat{\text{Var}}(\hat{\pi}_{de})$ is the estimated variance of $\hat{\pi}_{de}$ obtained from (4) after replacing the parameters π and ϕ by $\hat{\pi}_{de}$ and $\hat{\delta}_{de}$, respectively.

3.4 The QEE based interval

Using the above semi-parametric model, the optimal quadratic estimating equations (QEE) of Crowder (1987) for the parameters π and ϕ are given by

$$U_\pi = \textstyle\sum_{i=1}^{m}[\Phi_{i\pi}(z_i-\pi) + \Delta_{i\pi}\{(z_i-\pi)^2 - \sigma_{i\lambda}^2\}] = 0 \text{ and}$$
$$U_\phi = \textstyle\sum_{i=1}^{m}[\Phi_{i\phi}(z_i-\pi) + \Delta_{i\phi}\{(z_i-\pi)^2 - \sigma_{i\lambda}^2\}] = 0,$$

where $z_i = x_i/n_i$, $\lambda = (\pi,\phi)$, $\sigma_{i\lambda}^2 = \pi(1-\pi)\{1+(n_i-1)\phi\}/n_i = \mu_{2i}$, $\Phi_{i\pi} = [-(\gamma_{2i\lambda}+2) + \gamma_{1i\lambda}(1-2\pi)\sigma_\lambda/\pi(1-\pi)]/\sigma_{i\lambda}^2\gamma_{i\lambda}$, $\Delta_{i\pi} = [\gamma_{1i\lambda} - (1-2\pi)\sigma_\lambda/\pi(1-\pi)]/\sigma_{i\lambda}^3\gamma_{i\lambda}$, $\Phi_{i\phi} = \gamma_{1i\lambda}\pi(1-\pi)(n_i-1)/n_i\sigma_{i\lambda}^3\gamma_{i\lambda}$, $\Delta_{i\phi} = -\pi(1-\pi)(n_i-1)/n_i\sigma_{i\lambda}^4\gamma_{i\lambda}$ and $\gamma_{i\lambda} = \gamma_{2i\lambda} + 2 - \gamma_{1i\lambda}^2$, with $\gamma_{1i\lambda}$ and $\gamma_{2i\lambda}$ being the skewness and kurtosis measures of x_i, respectively. In practice, $\gamma_{1i\lambda}$ and $\gamma_{2i\lambda}$ are unknown. However, one can estimate these parameters using the 3rd and 4th order moments of the beta-binomial model, which are $\mu_{3i} = \mu_{2i}(1-2\pi)\{1+(2n_i-1)\phi\}/n_i(1+\phi)$ and $\mu_{4i} = \mu_{2i}[\{1+(2n_i-1)\phi\}\{1+(3n_i-1)\phi\}\{1-3\pi(1-\pi)\} + (n_i-1)(1-\phi)\{\phi+3n_i\mu_{2i}\}]/[(1+\phi)(1+2\phi)n_i^2]$, respectively. The QEE estimates $\hat{\pi}_{qee}$ and $\hat{\phi}_{qee}$ are the solutions to the above equations. Following the results of Inagaki (1973), the sandwich variance, as $k \to \infty$, of $\hat{\pi}_{qee}$ can be obtained as Var$(\hat{\pi}_{qee}) = \Gamma_{11}$, where Γ_{11} is the 1st diagonal element of the variance-covariance matrix

Table 4: Coverage probability estimates (expected interval lengths) based on confidence intervals by the methods with nominal level, $1 - \alpha$ = 95% and ED litter sizes for sample size $k = 20$.

ϕ	π	ML	EQL	DEQL	QEE	WA
0.1	0.05	0.985 (0.100)	0.988 (0.099)	0.961 (0.130)	0.987 (0.153)	0.975 (0.085)
	0.10	0.976 (0.114)	0.991 (0.119)	0.988 (0.136)	0.980 (0.170)	0.939 (0.113)
	0.15	0.947 (0.134)	0.973 (0.153)	0.967 (0.155)	0.950 (0.148)	0.946 (0.134)
	0.20	0.952 (0.148)	0.970 (0.187)	0.966 (0.168)	0.951 (0.155)	0.950 (0.148)
	0.25	0.954 (0.162)	0.983 (0.201)	0.965 (0.181)	0.955 (0.164)	0.945 (0.161)
	0.30	0.955 (0.172)	0.992 (0.237)	0.967 (0.189)	0.954 (0.173)	0.937 (0.170)
	0.35	0.951 (0.180)	0.991 (0.251)	0.966 (0.217))	0.954 (0.181)	0.941 (0.178)
	0.40	0.950 (0.186)	0.997 (0.267)	0.961 (0.231)	0.951 (0.187)	0.923 (0.180)
	0.45	0.951 (0.193)	1.000 (0.293)	0.967 (0.258)	0.955 (0.194)	0.929 (0.184)
	0.50	0.930 (0.197)	1.000 (0.327)	0.957 (0.283)	0.933 (0.198)	0.923 (0.185)
0.3	0.05	0.988 (0.128)	0.994 (0.126)	0.943 (0.204)	0.995 (0.147)	0.962 (0.118)
	0.10	0.972 (0.147)	0.982 (0.152)	0.984 (0.191)	0.976 (0.178)	0.942 (0.158)
	0.15	0.904 (0.173)	0.916 (0.191)	0.954 (0.219)	0.919 (0.270)	0.932 (0.189)
	0.20	0.906 (0.196)	0.935 (0.232)	0.951 (0.243)	0.925 (0.231)	0.930 (0.211)
	0.25	0.918 (0.215)	0.948 (0.274)	0.956 (0.263)	0.927 (0.225)	0.938 (0.228)
	0.30	0.927 (0.232)	0.966 (0.292)	0.964 (0.280)	0.930 (0.238)	0.936 (0.243)
	0.35	0.940 (0.241)	0.971 (0.323)	0.966 (0.289)	0.941 (0.246)	0.934 (0.251)
	0.40	0.951 (0.247)	0.984 (0.342)	0.969 (0.294)	0.950 (0.251)	0.930 (0.257)
	0.45	0.950 (0.252)	0.987 (0.374)	0.971 (0.299)	0.955 (0.256)	0.942 (0.262)
	0.50	0.927 (0.255)	0.994 (0.397)	0.964 (0.302)	0.925 (0.260)	0.929 (0.263)
0.5	0.05	0.995 (0.154)	0.999 (0.149)	0.979 (0.232)	0.997 (0.181)	0.955 (0.149)
	0.10	0.972 (0.174)	0.978 (0.179)	0.986 (0.226)	0.989 (0.217)	0.944 (0.194)
	0.15	0.891 (0.205)	0.905 (0.225)	0.959 (0.257)	0.921 (0.324)	0.933 (0.231)
	0.20	0.886 (0.232)	0.908 (0.271)	0.946 (0.289)	0.908 (0.267)	0.917 (0.258)
	0.25	0.894 (0.256)	0.920 (0.315)	0.951 (0.315)	0.907 (0.322)	0.919 (0.279)
	0.30	0.925 (0.275)	0.953 (0.361)	0.965 (0.337)	0.933 (0.293)	0.936 (0.297)
	0.35	0.935 (0.288)	0.955 (0.389)	0.970 (0.352)	0.945 (0.302)	0.929 (0.309)
	0.40	0.939 (0.296)	0.965 (0.413)	0.975 (0.360)	0.943 (0.307)	0.928 (0.317)
	0.45	0.932 (0.303)	0.967 (0.443)	0.971 (0.367)	0.940 (0.314)	0.930 (0.322)
	0.50	0.918 (0.307)	0.970 (0.477)	0.959 (0.371)	0.919 (0.316)	0.932 (0.323)

Γ given by

$$\Gamma = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix} = [A(\hat{\pi}, \hat{\phi})]^{-1} B(\hat{\pi}, \hat{\phi}) \{[A(\hat{\pi}, \hat{\phi})]^{-1}\}'.$$

The expressions for the elements of the 2×2 matrix $A(\hat{\pi}, \hat{\phi})$ in the variance-covariance matrix Γ are obtained by

$$\begin{aligned}
A_{\pi\pi}(\hat{\pi}, \hat{\phi}) &= \mathrm{E}\left(-\partial U_\pi/\partial \pi\right) = \sum_{i=1}^{m} \left[\Phi_{i\pi} + \Psi_{i\pi}(1-2\pi)\{1+(n_i-1)\phi\}/n_i\right] \\
A_{\pi\phi}(\hat{\pi}, \hat{\phi}) &= \mathrm{E}\left(-\partial U_\pi/\partial \phi\right) = \sum_{i=1}^{m} \left[\Psi_{i\pi}\pi(1-\pi)(n_i-1)/n_i\right] \\
A_{\phi\pi}(\hat{\pi}, \hat{\phi}) &= \mathrm{E}\left(-\partial U_\phi/\partial \pi\right) = \sum_{i=1}^{m} \left[\Phi_{i\phi} + \Psi_{i\phi}(1-2\pi)\{1+(n_i-1)\phi\}/n_i\right] \quad \text{and} \\
A_{\phi\phi}(\hat{\pi}, \hat{\phi}) &= \mathrm{E}\left(-\partial U_\phi/\partial \phi\right) = \sum_{i=1}^{m} \left[\Psi_{i\phi}\pi(1-\pi)(n_i-1)/n_i\right].
\end{aligned}$$

Similarly, we obtain the expressions for the elements of the 2×2 matrix $B(\hat{\pi}, \hat{\phi})$ in the variance-covariance matrix Γ by

$$\begin{aligned}
B_{\pi\pi}(\hat{\pi}, \hat{\phi}) &= \mathrm{E}(U_\pi U_\pi) = \sum_{i=1}^{m} \left[\Phi_{i\pi}^2 \mu_{i2} + 2\Phi_{i\pi}\Psi_{i\pi}\mu_{3i} + \Psi_{i\pi}^2(\mu_{4i} - \mu_{2i}^2)\right] \\
B_{\pi\phi}(\hat{\pi}, \hat{\phi}) &= \mathrm{E}(U_\pi U_\phi) = \sum_{i=1}^{m} \left[\Phi_{i\pi}\Phi_{i\phi}\mu_{i2} + \Phi_{i\pi}\Psi_{i\phi}\mu_{3i} + \Phi_{i\phi}\Psi_{i\pi}\mu_{3i} + \Psi_{i\pi}\Psi_{i\phi}(\mu_{4i} - \mu_{2i}^2)\right] \\
&= B_{\phi\pi}(\hat{\phi}) \quad \text{and} \\
B_{\phi\phi}(\hat{\pi}, \hat{\phi}) &= \mathrm{E}(U_\phi U_\phi) = \sum_{i=1}^{m} \left[\Phi_{i\phi}^2 \mu_{i2} + 2\Phi_{i\phi}\Psi_{i\phi}\mu_{3i} + \Psi_{i\phi}^2(\mu_{4i} - \mu_{2i}^2)\right],
\end{aligned}$$

where $\Phi_{i\pi}$, $\Psi_{i\pi}$, $\Phi_{i\pi}$, $\Psi_{i\pi}$, μ_{2i}, μ_{3i}, and μ_{4i} are defined above. Then, the approximate 100(1 - α)% confidence interval for π based on the QEE method is given by

$$\hat{\pi}_{qee} - z_{\alpha/2}\sqrt{\widehat{\mathrm{Var}}(\hat{\pi}_{qee})} \leq \pi \leq \hat{\pi}_{qee} + z_{\alpha/2}\sqrt{\widehat{\mathrm{Var}}(\hat{\pi}_{qee})},$$

where $\widehat{\mathrm{Var}}(\hat{\pi}_{qee})$ is the estimated variance of $\hat{\pi}_{qee}$ obtained from $\mathrm{Var}(\hat{\pi}_{qee})$ after replacing the parameters π and ϕ by $\hat{\pi}_{qee}$ and $\hat{\phi}_{qee}$, respectively.

3.5 The Wald interval

From the above semiparametric model, one can obtain an unbiased estimate of π as the sample proportion $\hat{\pi} = \sum_i^k x_i / \sum_i^k n_i = x_. / n_.$. The variance of $\hat{\pi}$ is given by $\mathrm{Var}(\hat{\pi}) = \pi(1-\pi)\xi/n_.$. Using the central limit theorem, we can show that $\hat{\pi}$ follows the normal distribution with mean π and variance $\pi(1-\pi)\xi/n_.$, as $k \to \infty$. Then, an approximate $100(1-\alpha)\%$ Wald confidence interval for π is given by

$$\hat{\pi} - z_{\alpha/2}\sqrt{\hat{\pi}(1-\hat{\pi})\hat{\xi}/n_.} \leq \pi \leq \hat{\pi} + z_{\alpha/2}\sqrt{\hat{\pi}(1-\hat{\pi})\hat{\xi}/n_.}.$$

Note that we use the ML estimate of ϕ in the equation for $\hat{\xi}$ above and denote this interval by WA.

Table 5: Coverage probability estimates (expected interval lengths) based on confidence intervals by the methods with nominal level, $1-\alpha$ = 95% and ED litter sizes for sample size $k = 50$.

ϕ	π	ML	EQL	DEQL	QEE	WA
0.1	0.05	0.957 (0.060)	0.956 (0.059)	0.943 (0.078)	0.975 (0.074)	0.942 (0.052)
	0.10	0.958 (0.070)	0.969 (0.072)	0.971 (0.083)	0.970 (0.132)	0.943 (0.071)
	0.15	0.948 (0.084)	0.967 (0.093)	0.967 (0.097)	0.955 (0.090)	0.949 (0.085)
	0.20	0.960 (0.094)	0.990 (0.118)	0.982 (0.106)	0.966 (0.097)	0.954 (0.095)
	0.25	0.942 (0.103)	0.986 (0.127)	0.959 (0.114)	0.938 (0.104)	0.940 (0.104)
	0.30	0.945 (0.108)	0.992 (0.145)	0.954 (0.119)	0.944 (0.109)	0.940 (0.109)
	0.35	0.936 (0.113)	0.999 (0.172)	0.952 (0.122)	0.936 (0.113)	0.931 (0.113)
	0.40	0.934 (0.117)	1.000 (0.187)	0.949 (0.126)	0.931 (0.117)	0.932 (0.116)
	0.45	0.954 (0.120)	1.000 (0.206)	0.965 (0.130)	0.957 (0.121)	0.946 (0.118)
	0.50	0.934 (0.123)	1.000 (0.237)	0.951 (0.134)	0.931 (0.123)	0.949 (0.119)
0.3	0.05	0.983 (0.075)	0.967 (0.072)	0.884 (0.127)	0.989 (0.086)	0.942 (0.073)
	0.10	0.934 (0.091)	0.932 (0.092)	0.982 (0.117)	0.952 (0.1020	0.944 (0.101)
	0.15	0.912 (0.110)	0.922 (0.119)	0.965 (0.140)	0.922 (0.203)	0.930 (0.120)
	0.20	0.935 (0.126)	0.957 (0.147)	0.969 (0.158)	0.953 (0.308)	0.951 (0.136)
	0.25	0.925 (0.138)	0.971 (0.173)	0.975 (0.170)	0.937 (0.144)	0.943 (0.146)
	0.30	0.924 (0.148)	0.973 (0.202)	0.972 (0.179)	0.925 (0.152)	0.934 (0.155)
	0.35	0.939 (0.155)	0.980 (0.231)	0.970 (0.187)	0.937 (0.159)	0.944 (0.162)
	0.40	0.952 (0.160)	0.988 (0.257)	0.977 (0.191)	0.948 (0.163)	0.951 (0.166)
	0.45	0.962 (0.164)	0.992 (0.283)	0.984 (0.195)	0.963 (0.166)	0.952 (0.169)
	0.50	0.953 (0.165)	0.998 (0.317)	0.978 (0.195)	0.947 (0.167)	0.948 (0.170)
0.5	0.05	0.988 (0.089)	0.985 (0.082)	0.983 (0.142)	0.997 (0.105)	0.932 (0.091)
	0.10	0.928 (0.107)	0.917 (0.107)	0.987 (0.139)	0.964 (0.123)	0.955 (0.124)
	0.15	0.916 (0.131)	0.929 (0.141)	0.972 (0.166)	0.947 (0.174)	0.954 (0.148)
	0.20	0.917 (0.150)	0.938 (0.173)	0.973 (0.189)	0.932 (0.178)	0.938 (0.167)
	0.25	0.929 (0.165)	0.956 (0.201)	0.975 (0.206)	0.949 (0.183)	0.952 (0.180)
	0.30	0.940 (0.177)	0.973 (0.228)	0.980 (0.219)	0.944 (0.187)	0.951 (0.191)
	0.35	0.930 (0.186)	0.972 (0.256)	0.972 (0.228)	0.931 (0.194)	0.941 (0.199)
	0.40	0.957 (0.193)	0.984 (0.285)	0.988 (0.235)	0.964 (0.199)	0.957 (0.205)
	0.45	0.953 (0.196)	0.988 (0.313)	0.988 (0.239)	0.962 (0.202)	0.947 (0.208)
	0.50	0.943 (0.197)	0.987 (0.334)	0.984 (0.240)	0.942 (0.203)	0.965 (0.209)

4. A Simulation Study

We investigated the performance of the five interval procedures discussed in Section 3 (ML=maximum likelihood, EQL=extended quasi-likelihood, DEQL=double extended quasi-likelihood, QEE=quadratic estimating equations, and WA=Wald) by way of simulation. Using a pre-assigned confidence level of 95% we examined their coverage probabilities (CPs) and expected lengths (ELs) using ten values of the population proportion π = 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5 and three values measuring the lack of independence among observations within the same cluster ϕ = 0.1, 0.3, 0.5 to simulate data from the beta-binomial distribution. We used four different configurations of cluster sizes: (i) the low dose treatment group (k = 19) of the data in Table 1 of Paul (1982); (ii) the control group (k=73) for the data in Table 3 of Paul and Saha (2007); and random sizes from the empirical distribution (ED) of 523 litter sizes quoted by Kupper et al. (1986) for k=20 and 50. Based on 1000 samples, we computed the empirical coverage probabilities as the number of times the confidence interval contained the true value divided by 1000 and the expected coverage length as the mean of the 1000 lengths. The results are reported in Tables 2-5 from which we make the following observations:

- The CP and EL results between fixed and ED litter sizes for all five methods are in remarkable agreement for similar sample sizes. Specifically, the CPs and ELs for all five methods are virtually the same across all combinations of π and ϕ in the cases of fixed litter size with k=19 and ED litter size with k=20. The same is true in the cases of fixed litter size k=73 and ED litter size k=50.

- For all five methods, the ELs increase as the true proportion π increases; the ELs decrease as sample size k increases; and the ELs increase as the deviation from independence ϕ increases.

- The ML, QEE, and WA methods tend to have similar ELs which are smaller than the ELs of the EQL and DEQL methods.

- The ML method has among the lowest ELs which in many situations is at the expense of under-coverage whereas good coverage properties of the QEE method tend to be at the expense of larger ELs. Both methods show inconsistent coverage properties for small values of π.

- The ML method shows inconsistent coverage with over-coverage for small values of π and under-coverage for larger values of π, except when ϕ = .1 where the CPs tend to be near goal.

- The EQL and DEQL methods show severe over-coverage across the board whereas the CPs for the QEE and WA methods are either at goal or show slight under-coverage.

- The WA method produced inconsistent coverage, namely, under-coverage for small sample sizes, especially for large values of π, but good coverage for large sample sizes.

5. CTC Images Data Example

Computed tomography colonography (CTC) is an imaging test that can detect polyps before they develop into cancer. Investigators have developed a computer algorithm, called computer aided detection (CAD), to help radiologists detect polyps on the CTC. Two hundred seventy patients from six institutions were compiled in the retrospective design. These patients had undergone CTC for several medical reasons. In order to assess the reader performance, 30 patients were randomly selected from the 119 test cases. In this study, there were actually multiple polyps in some patients and 25 of 30 patients had from 1 to 3 polyps. Hence the detection capabilities for each patient may be correlated. The data for reader 1 are displayed in Table 1. The purpose of this study was to assess the reader accuracy of CAD-enhanced CTC for detecting polyps. In this case, we considered the above confidence interval procedures to estimate the sensitivity of CAD-enhanced CTC for detecting polyps. The point estimates (standard errors) $\hat{\pi}_{ml}$, $\hat{\pi}_{eql}$, $\hat{\pi}_{de}$, and $\hat{\pi}_{qee}$ are given by 0.8464 (0.0633), 0.8473 (0.1458), 0.8456 (0.0777), and 0.8433 (0.0685), respectively. Also, the point estimates $\hat{\phi}_{ml}$, $\hat{\phi}_{eql}$, $\hat{\phi}_{de}$, and $\hat{\phi}_{qee}$ are given by 0.3426, 0.1919, 0.2371, and 0.2873, respectively. Then, the 95% confidence intervals for π using the ML, EQL, DEQL, QEE, and WA methods are given by (0.7223, 0.9705), (0.5614, 1.0), (0.6933, 0.9978), (0.7090, 0.9776), and (0.7192, 0.9736), respectively.

6. Discussion

We proposed five confidence intervals for estimating a population proportion based on binary outcome data taken from clusters assuming a beta-binomial distribution and compared their confidence probabilities and expected lengths using a simulation study and real-life application. Generally speaking, the results of the simulation study in Section 4 suggest the maximum likelihood method produced the shortest intervals followed by the quadratic estimating equations and Wald methods. However, all three suffer to some extent from under/over coverage, especially when the true proportion is small, and Wald's method requires the estimation of a nuisance parameter using iterative methods. The extended quasi-likelihood and double extended quasi-likelihood procedures produced much longer intervals indicating their severely conservative nature. When applied to CTC images data, the results were consistent with the simulation study in that the maximum likelihood method produced the shortest interval followed closely by Wald and then the quadratic estimating equations method. As expected, the extended quasi-likelihood and double extended quasi-likelihood methods gave longer intervals with the former producing a profoundly longer interval.

All five methods rely on the asymptotic normality distribution assumption which may not be valid, especially for small sample sizes or small parameter values. This, and the estimation of the asymptotic standard error, may explain the lack of coverage problems of these asymptotic confidence intervals. Further study using alternative distributional approximations may alleviate these issues.

Acknowledgements

This paper was supported in part by a CSU-AAUP University research grant.

REFERENCES

Altham, P.M.E. (1978), "Two Generalizations of the Binomial Distribution," *Applied Statistics*, **27**, 162-167.

Breslow, N.E. (1984), "Extra-Poisson Variation in Log-Linear Models," *Appl. Statist.*, 33, 38-44.

Brown, L.D., Cai, T.T., and DasGupta, A. (2001), "Confidence Intervals for A Binomial Proportion (with Discussion)," *Statist Sci*, 16, 101133.

Chen, C., and Tipping, R.W. (2002), "Confidence Interval of A Proportion with Over-Dispearson," *Biometrical Journal*, **44**, 877-886.

Crowder, M.J. (1987), "On Linear and Quadratic Estimating Functions," *Biometrika*, 74, 591597.

Efron, B. (1986), ”Double Exponential Families and Their Use in Generalized Linear Regression”. *Journal of the American Statistical Association*, 81, 709-721.

Inagaki, N. (1973), ”Asymptotic Relations Between the Likelihood Estimating Function and the Maximum Likelihood Estimator,” *Annals of the Institute of Statistical Mathematics*, **25**, 1-26.

Korn, E.L., and Graubard, B.I. (1998), ”Confidence Intervals for Proportions with Small Expected Number of Positive Counts Estimated from Survey Data,” *Survey Methodology*, **24**, 193-201.

Kupper, L.L., and Haseman, J.K. (1978), ”The Use of A Correlated Binomial Model for the Analysis of Certain Toxicological Experiments,” *Biometrics*, **34**, 69-76.

Kupper, L.L., Portier, C., Hogan, M.D., and Yamamoto, E. (1986), ”The Impact of Litter Effects on DoseResponse Modeling in Teratology,” *Biometrics*, **42**, 85-98.

Lui, K.J. (2004), *Statistical Estimation of Epidemiological Risk*, Wiley. New York.

Moore, D.F. (1986), ”Asymptotic Properties of Moment Estimators for Overdispersed Counts and Proportions,” *Biometrika*, 73, 583-588.

Nelder, J.A., and Pregibon, D. (1987), ”An Extended Quasi-Likelihood Function,” *Biometrika*, **74**, 221-232.

Newcombe, R.G. (1998), ”Two-Sided Confidence Intervals for the Single Proportion: Comparison of Seven Methods,” *Stat Med*, 17, 857872.

Nikoloulopoulos, A., and Karlis, D. (2008), ”On Modeling Count Data: A Comparison of Some Well-Known Discrete Distributions,” *J. Statist. Comput. Simul.*, 78, 437-457.

Paul, S.R. (1982), ”Analysis of Proportions of Affected Foetuses in Teratological Experiments,” *Biometrics*, **38**, 361-370.

Paul, S.R., and Saha, K.K. (2007), ”The Generalized Linear Model and Extensions: A Review and Some Biological and Environmental Applications”, *Environmetrics*, 18, 421-443.

Saha, K.K. (2011), ”A Comparison Study of Some Competing Discrete Models for Proportions or Counts, with Applications to Biological Data,” *Journal of the Indian Society of Agricultural Statistics*, 65, 143-153.

Sukasih, A., and Jang, D. (2005), ”An Application of Confidence Interval Methods for Small Proportions in the Health Care Survey of DoD Beneficiaries,” *Proceedings of the Survey Research Methods Section, ASA*, 3608-3612.

Zhou, X.H., Obuchowski, N.A., and McClish, D.K. (2002), *Statistical Methods in Diagnostic Medicine*, Wiley, New York.