Confidence Interval Construction for Rate Ratio in Matched-pair Studies With Incomplete Data*

Hui-Qiong Li[1], **Ivan S. F. Chan**[2,*], **Man-Lai Tang**[3], **Guo-Liang Tian**[4] **and** **Nian-Sheng Tang**[1]

[1]Department of Statistics, Yunnan University, Kunming 650091, P. R. China

[2]Clinical Biostatistics, Merck Research Laboratories, North Wales, PA 19454, U. S. A.

[3]Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong, P. R. China

[4]Department of Statistics and Actuarial Science, The University of Hong Kong,

Pokfulam Road, Hong Kong, P. R. China

Summary

Matched-pair design is often used in clinical trials to increase the efficiency of establishing equivalence between two treatments with binary outcomes. In this article, we consider such a design based on rate ratio in the presence of incomplete data. The rate ratio is one of the most frequently used indices in comparing efficiency of two treatments in clinical trials. In this paper, we propose ten confidence interval estimators for the rate ratio in incomplete matched-pair designs. A hybrid method that recovers variance estimates required for the rate ratio from the confidence limits for single proportions is proposed. It is noteworthy that confidence intervals based on this hybrid method have closed-form solution. The performance of the proposed confidence intervals is evaluated with respect to their exact coverage probability, expected confidence interval widths, and distal and mesial non-coverage probability. The results show that the hybrid Agresti-Coull CI based on Fieller's theorem perform satisfactorily for small to moderate sample sizes. Two real examples from clinical trials will be used to illustrate the proposed confidence intervals.

Key words: Agresti–Coull interval; Correlated proportions; Jeffreys interval; Incomplete data; Method of variance estimations recovery; Wilson interval

---

*Corresponding author: Ivan S. F. Chan, Clinical Biostatistics, Merck Research Laboratories, North Wales, PA 19454, U. S. A. E-mail: ivan_chan@merck.com.

1 Introduction

Incomplete matched-pair data are often encountered in comparison studies of two treatments/diagnostics /reviewers of two different status for the same treatment. For instance, Osoba *et al.* (1986) considered a clinical trial with the objective of comparing methylprednisolone sodium succinate (MPRED) and metoclopramide (METCLO) with respect to efficacy in the prevention of vomiting induced by moderately emetogenic chemotherapy in patients with previously untreated cancer. 157 patients were randomized to receive either 250 mg MPRED or 10 mg METCLO for the first chemotherapy period, and then crossed over to the other study drug for the second chemotherapy cycle. After each chemotherapy cycle, patients are asked to complete a questionnaire measuring the number of episodes of vomiting. Define X (or Y) = 1 if a patient vomited at least once during the last six-hour period (i.e., hour 18 - 24) after receiving MPRED (or METCLO); = 0 otherwise. It was reported that amongst the 157 eligible patients, (i) 115 received **both** treatments in the two cycles (i.e., with **both** X and Y being observed); (ii) 16 received **only** MPRED for the first cycle **but not** METCLO for the second cycle (i.e., with **only** X being observed); and (iii) 26 received **only** METCLO for the first cycle **but not** MPRED for the second cycle (i.e., with **only** Y being observed). For scenario (i), it was reported that 106 patients experienced vomiting in **both** treatments, 6 had vomiting **only** in MPRED treatment, 23 had vomiting **only** in METCLO treatment, 9 had no vomiting experience. For scenarios (ii) and (iii), assume that 12 and 14 patients vomited at least once, respectively. Then, the final data consist of two parts: the complete observations which correspond to a 2×2 table from correlated series, and the incomplete observations which correspond to a 2×2 table from two independent binomial populations. We summarize the data in Table 1.

Table 1 here

Suppose that we would like to test the equivalence between MPRED and METCLO with respect to their rates of vomiting experience for the above crossover clinical trial. For this purpose, we can compute a (1 - α)100% confidence interval for the ratio of the two rates of vomiting experience. If the resultant confidence interval lies entirely in the interval (δ_0, $1/\delta_0$) with δ_0 (> 0) being some prestated clinically acceptable threshold, then we could conclude the equivalence between the two treatments at the α significance level. As a result, reliable confidence intervals for rate ratio in the presence of incomplete data are necessary.

Tests of non-inferiority/equivalence based on rate ratio of two independent treatments have been widely studied in the literature (see, for example, Miettinen and Nurminen, 1985; Farrington and Manning, 1990; Chan, 1998; Chen *et al.*, 2000). Parallel development for matched-pair designs have not been discussed until recently. Lachenbruch and Lynch (1998) proposed two statistics for establishing equivalence of a new HIV screening test to a current standard based on the rate ratio measure in a matched-pair design setting. Tang *et al.* (2003) empirically demonstrated that both statistics proposed by Lachenbruch and Lynch (1998) could produce empirical type I error rates that can be more than twice the pre-chosen nominal level in many cases and a score statistic perform satisfactorily in general situation. Tang *et al.* (2002) derived a score-test-based confidence interval for assessing equivalence based on the rate ratio. Zou and Donner (2008) proposed a so-called hybrid method to form approximate confidence intervals for rate ratio. However, all the above-mentioned work were confined to matched-pair data without missing data. Tang et al. (2009) proposed the exact and approximate unconditional confidence intervals for proportion difference in the presence of incomplete data. But these methods could be computationally intensive for moderate to large sample sizes and simple explicit formulas are impossible. Besides, the score test-based and likelihood-ratio-based confidence intervals have not yet been considered in incomplete matched-paired data. Hence, it is the aim of this article to consider the score test-based and likelihood-ratio-based confidence intervals and to generalize the aforementioned hybrid method to matched-pair studies based on rate ratio in the presence of incomplete data. These methods can be used for analysis of incomplete data as well as complete data.

In this article, we develop ten confidence interval estimators for correlated rate ratio with incomplete matched-pair data. A hybrid method that recovers variance estimates required for the rate ratio from the confidence limits for single proportions is also considered. The rest of this paper is organized as follows. In Section 2, we describe confidence intervals with incomplete matched-pair data based on the asymptotic method. The hybrid approach for confidence interval construction is presented in Section 3. In Section 4, numerical evaluations are conducted to investigate the performance of the proposed confidence intervals in terms of their exact coverage probability, expected confidence width, the ratio of mesial non-coverage probability and total non-coverage probability. We will illustrate our proposed methodologies with real examples from two clinical studies in Section 5. A brief discussion is given in Section 6.

2 Confidence intervals for rate ratio with incomplete data based on asymptotic approach

Let X and Y be the outcomes of two different treatments applied to each subject with joint probability $\Pr(X = i, Y = j) = \pi_{ij}$, $i, j = 0, 1$. Suppose that the dichotomous response is observed on n subjects for both treatments, and in addition m_1 subjects are observed only for the first treatment (i.e., X) and m_2 only for the second treatment (i.e., Y). The observed counts and the corresponding cell probabilities for the n complete data and $m_1 + m_2$ partially incomplete data can be summarized in Table 2.

Table 2 here

Here, n_{ij} is the number of subjects who go through both treatments with $X = i$, $Y = j$ for $i, j = 0, 1$, u is the number of subjects who go through ONLY treatment X with $X = 1$, and v is the number of subjects who go through ONLY treatment Y with $Y = 1$. It is assumed that the probabilities governing the complete and the incomplete data are the same, and that the mechanisms causing incomplete data are independent of the outcomes of the trials; all trials are assumed to be independent (see, Choi and Stablein, 1982; Tang and Tang, 2004; Lin et al.(2009)). With no missing data, i.e., $m_1 = m_2 = 0$, the random vector $\mathbf{n} = (n_{00}, n_{01}, n_{10}, n_{11})$ is then multinomially distributed with parameters n and $(\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11})$. The random variable u and the random varible v, respectively, follow $Binomial(m_1, \pi_{1+})$ and $Binomial(m_2, \pi_{+1})$, where $\pi_{1+} = \pi_{10} + \pi_{11}$ and $\pi_{+1} = \pi_{01} + \pi_{11}$. Under the random mechanism (i.e., independent of treatment and outcome), the observed data $Y_{obs} = \{n_{00}, n_{01}, n_{10}, n_{11}, u, m_1 - u, v, m_2 - v\}$ can be assumed to come from the following multinomial distribution:

$$\begin{aligned}\Pr(Y_{obs}|n, m_1, m_2, \boldsymbol{\pi}) &= c \cdot (\pi_{00})^{n_{00}}(\pi_{01})^{n_{01}}(\pi_{10})^{n_{10}}(\pi_{11})^{n_{11}} \\ &\quad \times(\pi_{1+})^u(1 - \pi_{1+})^{m_1-u}(\pi_{+1})^v(1 - \pi_{+1})^{m_2-v}\end{aligned} \tag{1}$$

where $c = \frac{n!}{n_{00}!n_{01}!n_{10}!n_{11}} \frac{m_1!}{u!(m_1-u)!} \frac{m_2!}{v!(m_2-v)!}$ and $\boldsymbol{\pi} = (\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11})$.

Let $\delta = \pi_{1+}/\pi_{+1}$. We have $\pi_{11} = \pi_{+1}$ - π_{01}, $\pi_{10} = (\delta$ - $1)\pi_{+1} + \pi_{01}$, and $\pi_{00} = 1$ - $\delta\pi_{+1}$ - π_{01}. Therefore, the log-likelihood function for the observed data is given by

$$l(\pi|Y_{obs}) = constant + \Sigma n_{ij} log \pi_{ij} + (m_1 - u) log(1 - \delta\pi_{+1}) + u log(\delta\pi_{+1}) + (m_2 - v) log(1 - \pi_{+1}) + v log \pi_{+1}.$$

Here, δ is the parameter of interest, π_{01} and π_{+1} become the nuisance parameters. In this paper, our main purpose is to construct confidence interval for the correlated rate ratio $\delta = \pi_{1+}/\pi_{+1}$. We describe various confidence interval estimators for the rate ratio as follows.

2.1 Likelihood-Ratio-Test-Based Confidence Interval(TlCI)

Let $\breve{\pi}_{+1}$, $\breve{\pi}_{01}$ and $\breve{\delta}$ be maximum likelihood estimator (MLE) of π_{+1}, π_{01} and δ. There is no analytical solution for the MLE of π_{+1}, π_{01} and δ, but the MLE of π_{+1}, π_{01} and δ satisfy the following equations:

$$\begin{cases} \frac{\partial l(\pi,\delta|Y_{obs})}{\partial \delta} = \frac{-n_{00}\pi_{+1}}{1-\delta\pi_{+1}-\pi_{01}} + \frac{n_{10}\pi_{+1}}{(\delta-1)\pi_{+1}+\pi_{01}} + \frac{u}{\delta} - \frac{(m_1-u)\pi_{+1}}{1-\delta\pi_{+1}}, \\ \frac{\partial l(\pi,\delta|Y_{obs})}{\partial \pi_{01}} = \frac{-n_{00}}{1-\delta\pi_{+1}-\pi_{01}} + \frac{n_{01}}{\pi_{01}} + \frac{n_{10}}{(\delta-1)\pi_{+1}+\pi_{01}} - \frac{n_{11}}{\pi_{+1}-\pi_{01}}, \\ \frac{\partial l(\pi,\delta|Y_{obs})}{\partial \pi_{+1}} = \frac{-\delta n_{00}}{1-\delta\pi_{+1}-\pi_{01}} + \frac{n_{10}(\delta-1)}{(\delta-1)\pi_{+1}+\pi_{01}} + \frac{n_{11}}{\pi_{+1}-\pi_{01}} + \frac{u+v}{\pi_{+1}} - \frac{\delta(m_1-u)}{1-\delta\pi_{+1}} - \frac{m_2-v}{1-\pi_{+1}}. \end{cases}$$

Let $\tilde{\pi}_{+1}$ and $\tilde{\pi}_{01}$ are respectively the constrained maximum likelihood estimates(CMLE) of π_{+1} and π_{01} under the null hypothesis $H_0 : \delta = \delta_0$. The CMLE of π_{+1} and π_{01} can not be expressed explicitly, hence we use the expectation-maximization (EM) algorithm to find CMLE. By the EM algorithm, the M-step finds the complete-data CMLE. Hence, we introduce latent vectors $u_i = (u_{i0}, u_{i1})^T, i = 0, 1$ and $v_j = (v_{0j}, v_{1j})^T, j = 0, 1$ such that $u_{00} + u_{01} = m_1 - u$, $u_{10} + u_{11} = u$, $v_{00} + v_{10} = m_2 - v$, and $v_{01} + v_{11} = v$. Denote these latent (or missing) data by $Y_{mis} = \{u_{00}, u_{01}, u_{10}, u_{11}, v_{00}, v_{01}, v_{10}, v_{11}\}$ and the complete data by $Y_{com} = \{Y_{obs}, Y_{mis}\}$. Consequently, the complete-data likelihood function is

$$L(\pi|Y_{com}) \propto \prod_{i=0}^{1} \prod_{j=0}^{1} \pi_{ij}^{n_{ij}+u_{ij}+v_{ij}}$$

which is a Dirichlet distribution up to a constant. Thus, the complete-data log-likelihood function under H_0 is given by

$$l(\pi_{01}, \pi_{+1}|Y_{com}) = (n_{11} + u_{11} + v_{11})log(\pi_{+1} - \pi_{01}) + (n_{10} + z_{10} + v_{10})log[(\delta - 1)\pi_{+1} + \pi_{01}]+$$
$$(n_{01} + u_{01} + v_{01})log\pi_{01} + (n_{00} + u_{00} + v_{00})log(1 - \delta\pi_{+1} - \pi_{01}) + constant. \quad (2)$$

The complete-data CMLE solve the following equations

$$\begin{cases} \frac{\partial l(\pi_{01},\pi_{+1}|Y_{com})}{\partial \pi_{01}} = -\frac{n_{11}+u_{11}+v_{11}}{\pi_{+1}-\pi_{01}} + \frac{n_{10}+u_{10}+v_{10}}{(\delta_0-1)\pi_{+1}+\pi_{01}} + \frac{n_{01}+u_{01}+v_{01}}{\pi_{01}} - \frac{n_{00}+u_{00}+v_{00}}{1-\delta_0\pi_{+1}-\pi_{01}} = 0, \\ \frac{\partial l(\pi_{01},\pi_{+1}|Y_{com})}{\partial \pi_{+1}} = \frac{n_{11}+u_{11}+v_{11}}{\pi_{+1}-\pi_{01}} + \frac{(n_{10}+u_{10}+v_{10})(\delta_0-1)}{(\delta_0-1)\pi_{+1}+\pi_{01}} - \frac{\delta_0(n_{00}+u_{00}+v_{00})}{1-\delta_0\pi_{+1}-\pi_{01}} = 0, \end{cases}$$

This is equivalence to the following equations

$$\begin{cases} \frac{n_{10}+u_{10}+v_{10}}{(\delta_0-1)\pi_{+1}+\pi_{01}} + \frac{n_{01}+u_{01}+v_{01}}{\pi_{01}} = \frac{n_{11}+u_{11}+v_{11}}{\pi_{+1}-\pi_{01}} + \frac{n_{00}+u_{00}+v_{00}}{1-\delta_0\pi_{+1}-\pi_{01}}, \\ \frac{(n_{10}+u_{10}+v_{10})\delta_0}{(\delta_0-1)\pi_{+1}+\pi_{01}} + \frac{n_{01}+u_{01}+v_{01}}{\pi_{01}} = \frac{(\delta_0+1)(n_{00}+u_{00}+v_{00})}{1-\delta_0\pi_{+1}-\pi_{01}}, \end{cases}$$

Following the arguments of Tang et al.(2003), the CMLE of π_{+1} and π_{01} under H_0 are respectively given by

$$\tilde{\pi}_{01} = \{-b + (b^2 - 4ac)^{\frac{1}{2}}\}/(2a), \tag{3}$$

$$\tilde{\pi}_{+1} = (\frac{(n + m_1 + m_2) - (n_{00} + u_{00} + v_{00})}{n + m_1 + m_2} - \tilde{\pi}_{01})/\delta_0, \tag{4}$$

where $a = (n+m_1+m_2)(1+\delta_0)$, $b = (n_{11}+u_{11}+v_{11}+n_{01}+u_{01}+v_{01})\delta_0^2-(n_{11}+u_{11}+v_{11}+n_{10}+u_{10}+v_{10}+2n_{01}+2u_{01}+2v_{01})$, and $c = (n_{01}+u_{01}+v_{01})(1-\delta_0)[(n+m_1+m_2)-(n_{00}+u_{00}+v_{00})]/(n+m_1+m_2)$. The E-step of the EM algorithm is to replace $u_i, i = 0, 1$ and $v_j, j = 0, 1$ by their conditional expectations which are given by

$$E(u_{ij}|Y_{obs}, \pi) = (m_1 - u)\tfrac{\pi_{ij}}{\pi_{i+}},\ i = 0, j = 0, 1,$$
$$E(u_{ij}|Y_{obs}, \pi) = u\tfrac{\pi_{ij}}{\pi_{i+}},\ i = 1, j = 0, 1,$$
$$E(v_{ij}|Y_{obs}, \pi) = (m_2 - v)\tfrac{\pi_{ij}}{\pi_{+j}},\ i = 0, 1, j = 0,$$
$$E(v_{ij}|Y_{obs}, \pi) = v\tfrac{\pi_{ij}}{\pi_{+j}},\ i = 0, 1, j = 1.$$

When there are no missing data, the maximum likelihood estimates $\check{\pi}_{+1}$, $\check{\pi}_{01}$ and $\check{\delta}$ of π_{+1}, π_{01} and δ can be obtained by solving the above same equation for incomplete data with $m_1 = 0$, $m_2 = 0$, $u = 0$ and $v = 0$. The CMLE $\tilde{\pi}_{+1}$ and $\tilde{\pi}_{01}$ of π_{+1} and π_{01} under H_0 have been given by Tang et al.(2003).

The likelihood-ratio statistic based on both complete and missing data for testing $H_0 : \delta = \delta_0$ is given by

$$T_l = 2\{l(\check{\pi}_{01}, \check{\pi}_{+1}, \check{\delta}|Y_{obs}) - l(\tilde{\pi}_{01}, \tilde{\pi}_{+1}, \delta_0|Y_{obs})\},$$

which is asymptotical distributed as the chi-squared distribution with one degree of freedom under H_0. Therefore, the approximate $(1 - \alpha)100\%$ likelihood-ratio-test-based confidence interval is given by $[\delta_L, \delta_U]$, where δ_L and δ_U are the solutions to the following equation:

$$T_l(\delta) = \chi^2_{1,\alpha},$$

where $\chi^2_{1,\alpha}$ is the upper α percentile point of central χ^2 distribution with one degree of freedom. There are no closed-forms for δ_L and δ_U. Hence, the bisection searching algorithm can be used to obtain δ_L and δ_U.

2.2 Score-Test-Based Confidence Interval (TsCI)

The score function with respect to δ and the Fisher information matrix with respect to δ, π_{01} and π_{+1} under $\delta = \delta_0$ are given by

$$\frac{\partial l(\pi, \delta|Y_{obs})}{\partial \delta|\delta = \delta_0} = \frac{-n_{00}\pi_{+1}}{1 - \delta_0\pi_{+1} - \pi_{01}} + \frac{n_{10}\pi_{+1}}{(\delta_0 - 1)\pi_{+1} + \pi_{01}} + \frac{u}{\delta_0} - \frac{(m_1 - u)\pi_{+1}}{1 - \delta_0\pi_{+1}},$$

$$I(\pi_{01}, \pi_{+1}) = \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{12} & I_{22} & I_{23} \\ I_{13} & I_{23} & I_{33} \end{pmatrix},$$

respectively, where

$$I_{11} = \frac{n\pi_{+1}^2}{1-\delta_0\pi_{+1}-\pi_{01}} + \frac{n\pi_{+1}^2}{(\delta_0-1)\pi_{+1}+\pi_{01}} + \frac{m_1\pi_{+1}^2}{\delta_0\pi_{+1}} + \frac{m_1\pi_{+1}^2}{1-\delta_0\pi_{+1}},$$

$$I_{12} = \frac{n\pi_{+1}}{1-\delta_0\pi_{+1}-\pi_{01}} + \frac{n\pi_{+1}}{(\delta_0-1)\pi_{+1}+\pi_{01}},$$

$$I_{13} = \frac{n(1-\pi_{01})}{1-\delta_0\pi_{+1})-\pi_{01}} - \frac{n\pi_{01}}{(\delta_0-1)\pi_{+1}+\pi_{01}} + \frac{m_1}{1-\delta_0\pi_{+1}},$$

$$I_{22} = \frac{n}{1-\delta_0\pi_{+1}-\pi_{01}} + \frac{n}{\pi_{01}} + \frac{n}{(\delta_0-1)\pi_{+1}+\pi_{01}} + \frac{n}{\pi_{+1}-\pi_{01}},$$

$$I_{23} = \frac{n\delta_0}{1-\delta_0\pi_{+1}-\pi_{01}} + \frac{n(\delta_0-1)}{(\delta_0-1)\pi_{+1}+\pi_{01}} - \frac{n}{\pi_{+1}-\pi_{01}},$$

$$I_{33} = \frac{\delta_0^2 n}{1-\delta_0\pi_{+1}-\pi_{01}} + \frac{n(\delta_0-1)^2}{(\delta_0-1)\pi_{+1}+\pi_{01}} + \frac{n}{\pi_{+1}-\pi_{01}} + \frac{m_1\delta_0+m_2}{\pi_{+1}} + \frac{m_1\delta_0^2}{1-\delta_0\pi_{+1}} + \frac{m_2}{1-\pi_{+1}}.$$

Thus, the left upper element I^{11} of I^{-1} can be expressed as

$$I^{11} = [I_{11} - \begin{pmatrix} I_{12} & I_{13} \end{pmatrix} \begin{pmatrix} I_{22} & I_{23} \\ I_{23} & I_{33} \end{pmatrix}^{-1} \begin{pmatrix} I_{12} \\ I_{13} \end{pmatrix}]^{-1}$$

Hence, the score statistic for testing H_0: $\delta = \delta_0$ is given by

$$T_S(\delta_0) = (\frac{\partial l(\pi,\delta|Y_{obs})}{\partial \delta}|_{\delta=\delta_0,\pi_{01}=\tilde{\pi}_{01},\pi_{+1}=\tilde{\pi}_{+1}})(I^{11}|_{\delta=\delta_0,\pi_{01}=\tilde{\pi}_{01},\pi_{+1}=\tilde{\pi}_{+1}})^{\frac{1}{2}}$$

which is asymptotically distributed as standard normal distribution under H_0. When data are complete, T_s is the same to that of Tang et al.(2003). Therefore, the approximate $(1 - \alpha)100\%$ score-test-based confidence interval for complete and incomplete data is given by $[\delta_L, \delta_U]$, where δ_L and δ_U are the solutions to the following equation:

$$T_S(\delta) = \pm z_{\alpha/2},$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ percentile point of the standard normal distribution, and the plus and the minus signs correspond to the lower limit δ_L and the upper limit δ_U, respectively. These two limits can be easily obtained by secant method (see, Tango, 1998).

2.3 Wald-Test-Based Confidence Interval (TwCI)

Noticing that n_{1+}/n and u/m_1 are two unbiased point estimates for π_{1+}, Choi and Stablein (1982) suggested the following unbiased estimator for π_{1+} and π_{+1}, which utilizes both the complete and incomplete data:

$$\hat{\pi}_{1+} = \frac{\psi_1 n_{1+}}{n} + \frac{(1-\psi_1)u}{m_1},$$

$$\hat{\pi}_{+1} = \frac{\psi_2 n_{+1}}{n} + \frac{(1-\psi_2)v}{m_2},$$

where $\psi_1 = n/(n + m_1)$ and $\psi_2 = n/(n + m_2)$. The asymptotic expectation of $\hat{\delta}$ is $E(\hat{\delta}) = \delta$, the asymptotic variance and covariance of π_{1+} and π_{+1} can be estimated by

$$\widehat{var}(\hat{\pi}_{1+}) = \frac{n_{1+}(n-n_{1+})\psi_1^2}{n^3} + \frac{u(m_1-u)(1-\psi_1)^2}{m_1^3},$$

$$\widehat{var}(\hat{\pi}_{+1}) = \frac{n_{+1}(n-n_{+1})\psi_2^2}{n^3} + \frac{v(m_2-v)(1-\psi_2)^2}{m_2^3},$$

$$\widehat{cov}(\hat{\pi}_{+1}, \hat{\pi}_{1+}) = \frac{(n_{00}n_{11}-n_{10}n_{01})\psi_1\psi_2}{n^3}.$$

The asymptotic variance of $\hat{\delta}$ can be given by

$$\widehat{var}(\hat{\delta}) = \frac{\widehat{var}(\hat{\pi}_{1+})}{\hat{\pi}_{1+}^2} + \frac{\hat{\pi}_{1+}^2 \widehat{var}(\hat{\pi}_{+1})}{\hat{\pi}_{+1}^4} - 2\frac{\hat{\pi}_{1+}}{\hat{\pi}_{+1}^3}\widehat{cov}(\hat{\pi}_{+1}, \hat{\pi}_{1+}),$$

Hence, an approximate $(1-\alpha)100\%$ confidence interval for δ on the basis of Wald-type statistic $T_w = (\hat{\delta} - \delta)/\sqrt{\widehat{var}(\hat{\delta})}$, which is asymptotically distributed as the standard normal distribution, is given by

$$[max\{0, \hat{\delta} - z_{\alpha/2}\sqrt{\widehat{var}(\hat{\delta})}\}, \hat{\delta} + z_{\alpha/2}\sqrt{\widehat{var}(\hat{\delta})}].$$

When data are complete, let ψ_1, ψ_2 be one in the above equations, we can obtain the Wald-type confidence interval.

2.4 Log-test-Based Confidence Interval (TlogCI)

The log statistic for testing H_0: $\delta = \delta_0$ is given by

$$Tlog(\delta_0) = \frac{log\hat{\delta} - log\delta_0}{\sqrt{\widehat{var}(log\hat{\delta})}},$$

which is asymptotically distributed as standard normal distribution under H_0, where

$$\widehat{var}(log\hat{\delta}) = \frac{\widehat{var}(\hat{\pi}_{1+})}{\hat{\pi}_{1+}^2} + \frac{\widehat{var}(\hat{\pi}_{+1})}{\hat{\pi}_{+1}^2} - \frac{2\widehat{cov}(\hat{\pi}_{+1}, \hat{\pi}_{1+})}{\hat{\pi}_{1+}\hat{\pi}_{+1}},$$

Hence, an approximate $(1-\alpha)100\%$ confidence interval for $log\delta$ on the basis of log statistic is given by

$$[max\{0, log\hat{\delta} - z_{\alpha/2}\sqrt{\widehat{var}(log\hat{\delta})}\}, log\hat{\delta} + z_{\alpha/2}\sqrt{\widehat{var}(log\hat{\delta})}].$$

A $100(1-\alpha)\%$ confidence interval for δ can then be obtained as $[exp(L_{log}), exp(U_{log})]$. When there are no missing data, let ψ_1, ψ_2 be one in the above equations, we can obtain the log-test-based confidence interval.

3 Confidence intervals for rate ratio with incomplete data based on hybrid approach

Zou and Donner (2008) considered a so-called method of variance estimates recovery (MOVER) to construct confidence intervals for correlated rate ratio. The basic idea of their method is to construct hybrid confidence limits for single proportions based on Wilson score intervals by recovering variance estimates required for the rate ratio. In this article, we generalize their hybrid (i.e., MOVER) approach to situations in which incomplete data are present. For this purpose, we first briefly describe their MOVER for difference between two parameters. Let θ_1 and θ_2 denote any two parameters of interest. Let $\hat{\theta}_1$ and $\hat{\theta}_2$ be two estimates of θ_1 and θ_2, respectively. By the Central Limit Theorem and under the assumption of independent between $\hat{\theta}_1$ and $\hat{\theta}_2$, an approximate two-sided $(1-\alpha)100\%$ confidence interval for θ_1 - θ_2 can be constructed as (L, U), where

$$L = \hat{\theta}_1 - \hat{\theta}_2 - z_{\alpha/2}\sqrt{Var(\hat{\theta}_1) + Var(\hat{\theta}_2)}, \text{ and}$$

$$U = \hat{\theta}_1 - \hat{\theta}_2 + z_{\alpha/2}\sqrt{Var(\hat{\theta}_1) + Var(\hat{\theta}_2)},$$

where $Var(\hat{\theta}_i)$ is the variance estimate for $\hat{\theta}_i$ $(i = 1, 2)$. Unfortunately, this procedure performs well only if sample sizes are large or the sampling distributions of $\hat{\theta}_i (i = 1, 2)$ are close to normal. To improve the performance, we can obtain better estimates of $Var(\hat{\theta}_i)$ $(i = 1, 2)$ at the neighborhood of the confidence limits L and U separately. Let (l_1, u_1) and (l_2, u_2) be the two-sided $(1-\alpha)100\%$ confidence intervals for θ_1 and θ_2, respectively. We know that (l_i, u_i) contains plausible parameter values of θ_i $(i = 1, 2)$. Among these plausible values for θ_1 and θ_2, the values closest to the minimum L and maximum U are respectively $l_1 - u_2$ and $u_1 - l_2$ in spirit of the score-type CI (see, Bartlett, 1953). According to the Central Limit Theorem, the variance estimates can now be recovered from

$\theta_1 = l_1$ as $\hat{V}ar(\hat{\theta}_1) = (\hat{\theta}_1 - l_1)^2/z^2_{\alpha/2}$ and from $\theta_2 = u_2$ as $\hat{V}ar(\hat{\theta}_2) = (u_2 - \hat{\theta}_2)^2/z^2_{\alpha/2}$ for setting L. As a result,

$$L = \hat{\theta}_1 - \hat{\theta}_2 - \sqrt{(\hat{\theta}_1 - l_1)^2 + (u_2 - \hat{\theta}_2)^2}.$$

Similarly, we have

$$U = \hat{\theta}_1 - \hat{\theta}_2 + \sqrt{(u_1 - \hat{\theta}_1)^2 + (\hat{\theta}_2 - l_2)^2}.$$

For $\hat{\theta}_1$ and $\hat{\theta}_2$ being correlated, we can extend the above results in a straightforward fashion and the confidence limits for $\theta_1 - \theta_2$ are then

$$L = \hat{\theta}_1 - \hat{\theta}_2 - \sqrt{(\hat{\theta}_1 - l_1)^2 + (u_2 - \hat{\theta}_2)^2 - 2\widehat{corr}(\hat{\theta}_1, \hat{\theta}_2)(\hat{\theta}_1 - l_1)(u_2 - \hat{\theta}_2)}, \text{ and} \tag{5}$$

$$U = \hat{\theta}_1 - \hat{\theta}_2 + \sqrt{(u_1 - \hat{\theta}_1)^2 + (\hat{\theta}_2 - l_2)^2 - 2\widehat{corr}(\hat{\theta}_1, \hat{\theta}_2)(u_1 - \hat{\theta}_1)(\hat{\theta}_2 - l_2)}, \tag{6}$$

where $\widehat{corr}(\hat{\theta}_1, \hat{\theta}_2)$ is any sensible correlation estimate between $\hat{\theta}_1$ and $\hat{\theta}_2$.

3.1 Hybrid Fieller-type confidence interval

To construct confidence interval for $\delta = \pi_{1+}/\pi_{+1}$, we may let $\theta_1 = \pi_{1+}$ and $\theta_2 = \pi_{+1}$. We may first consider confidence interval for π_{1+} - $\delta\pi_{+1}$. Let L' denote the lower confidence limit for δ. The objective is to find L' such that $Pr(\pi_{1+}/\pi_{+1} \leq L') = \alpha/2$, that is,

$$\Pr(\pi_{1+} - L'\pi_{+1} \leq 0) = \alpha/2.$$

For any fixed L', applying (5) to $\pi_{1+} - L'\pi_{+1}$ gives

$$LL = \hat{\pi}_{1+} - L'\hat{\pi}_{+1} - \sqrt{(\hat{\pi}_{1+} - l_1)^2 + L'^2(u_2 - \hat{\pi}_{+1})^2 - 2L'\widehat{corr}(\hat{\pi}_{1+}, \hat{\pi}_{+1})(\hat{\pi}_{1+} - l_1)(u_2 - \hat{\pi}_{+1})}.$$

Setting $LL = 0$, we obtain

$$L' = \frac{[A - \hat{\pi}_{1+}\hat{\pi}_{+1}] + \sqrt{[A - \hat{\pi}_{1+}\hat{\pi}_{+1}]^2 - l_1(2\hat{\pi}_{1+} - l_1)u_2(2\hat{\pi}_{+1} - u_2)}}{u_2(u_2 - 2\hat{\pi}_{+1})}. \tag{7}$$

where $A = \widehat{corr}(\hat{\pi}_{1+}, \hat{\pi}_{+1})(\hat{\pi}_{1+} - l_1)(u_2 - \hat{\pi}_{+1})$.

Similarly, according to (6), the (1 - α)100% upper limit for π_{1+}/π_{+1} is given as

$$U' = \frac{[B - \hat{\pi}_{1+}\hat{\pi}_{+1}] - \sqrt{[B - \hat{\pi}_{1+}\hat{\pi}_{+1}]^2 - u_1(2\hat{\pi}_{1+} - u_1)l_2(2\hat{\pi}_{+1} - l_2)}}{l_2(l_2 - 2\hat{\pi}_{+1})}. \tag{8}$$

where $B = \widehat{corr}(\hat{\pi}_{1+}, \hat{\pi}_{+1})(u_1 - \hat{\pi}_{1+})(\hat{\pi}_{+1} - l_2)$. $\hat{\pi}_{1+}$ and $\hat{\pi}_{+1}$ are correlated in matched-pair design and the correlation can be estimated by

$$\widehat{corr}(\hat{\pi}_{1+}, \hat{\pi}_{+1}) = \frac{\frac{\psi_1\psi_2(\hat{\pi}_{11}\hat{\pi}_{00} - \hat{\pi}_{10}\hat{\pi}_{01})}{n}}{\sqrt{\frac{\hat{\pi}_{1+}(1-\hat{\pi}_{1+})\hat{\pi}_{+1}(1-\hat{\pi}_{+1})}{(n+m_1)(n+m_2)}}},$$

where $\hat{\pi}_{11} = n_{11}/n$, $\hat{\pi}_{00} = n_{00}/n$, $\hat{\pi}_{10} = n_{10}/n$, $\hat{\pi}_{01} = n_{01}/n$, $\psi_1 = n/(n+m_1)$, $\psi_2 = n/(n+m_2)$, $\hat{\pi}_{1+} = \frac{\psi_1 n_{1+}}{n} + \frac{(1-\psi_1)u}{m_1} = \frac{n_{10}+n_{11}+u}{n+m_1}$, and $\hat{\pi}_{+1} = \frac{\psi_2 n_{+1}}{n} + \frac{(1-\psi_2)u}{m_2} = \frac{n_{01}+n_{11}+v}{n+m_2}$.

To obtain confidence interval for π_{1+}/π_{+1} using (7) and (8), one needs two separate confidence intervals, denoted as $(l_1,\, u_1)$ and $(l_2,\, u_2)$, for $\theta_1 = \pi_{1+}$ and $\theta_2 = \pi_{+1}$, respectively. Here, we notice that $n_{1+} + u \sim Binomial(n + m_1, \pi_{1+})$ and $n_{+1} + v \sim Binomial(n + m_2, \pi_{+1})$. According to Brown, Cai and Dasgupta (2001), we consider only the Wilson, Jeffreys and Agresti-Coull intervals for $\theta_1 = \pi_{1+}$ and $\theta_2 = \pi_{+1}$. In general, let $Y_i \sim Binomial(n_i,\, \theta_i)$ and $\hat{\theta}_i = Y_i/n_i$ $(i = 1,\, 2)$. To save space, we simply report their formulae as follows:

(1) The Wilson score interval (WCI)

$$l_i = \tilde{\theta}_i - \frac{z_{\alpha/2}}{\tilde{n}_i}\sqrt{n_i\hat{\theta}_i(1-\hat{\theta}_i) + \frac{z^2_{\alpha/2}}{4}}, \text{ and } u_i = \tilde{\theta}_i + \frac{z_{\alpha/2}}{\tilde{n}_i}\sqrt{n_i\hat{\theta}_i(1-\hat{\theta}_i) + \frac{z^2_{\alpha/2}}{4}}, i = 1, 2,$$

where $\tilde{\theta}_i = (Y_i + 0.5z^2_{\alpha/2})/(n_i + z^2_{\alpha/2})$, $\tilde{n}_i = n_i + z^2_{\alpha/2}$. For $\theta_1 = \pi_{1+}$, $Y_1 = n_{1+} + u$, $n_1 = n + m_1$, and $Y_2 = n_{+1} + v$, $n_2 = n + m_2$ for $\theta_2 = \pi_{+1}$.

(2) The Agresti-Coull interval (ACI)

$$l_i = \tilde{\theta}_i - z_{\alpha/2}\sqrt{\tilde{\theta}_i(1-\tilde{\theta}_i)/\tilde{n}_i}, \text{ and } u_i = \tilde{\theta}_i + z_{\alpha/2}\sqrt{\tilde{\theta}_i(1-\tilde{\theta}_i)/\tilde{n}_i}, i = 1, 2.$$

(3) The Jeffreys interval(JCI)

$$l_i = \frac{2Y_i+1}{2Y_i+1+(2[n_i-Y_i]+1)F_{\alpha/2}(2[n_i-Y_i]+1, 2Y_i+1)}, \text{ and}$$

$$u_i = \frac{2Y_i+1}{2Y_i+1+(2[n_i-Y_i]+1)F_{1-\alpha/2}(2[n_i-Y_i]+1, 2Y_i+1)}, i = 1, 2,$$

where $F_r(\nu_1, \nu_2)$ is the upper r quantile from the F-distribution with (ν_1, ν_2) degrees of freedom.

3.2 Hybrid logarithmic transformation confidence interval

To construct confidence interval for the correlated proportion ratio π_{1+}/π_{+1}, one can also first construct a confidence interval for $log(\pi_{1+}/\pi_{+1})$ (i.e., $log\pi_{1+}$ - $log\pi_{+1}$), say $[L_{ln},\, U_{ln}]$. For this purpose, we can simply set $\theta_1 = log(\pi_{1+})$ and $\theta_2 = log(\pi_{+1})$. From (5) and (6), we can readily obtain a $100(1-\alpha)\%$ confidence interval for the log rate difference. Then, a $100(1-\alpha)\%$ confidence interval for π_{1+}/π_{+1} can be obtained as $[exp(L_{log}),\, exp(U_{log})]$.

To obtain confidence interval for $log\pi_{1+}$ - $log\pi_{+1}$, we need two separate confidence intervals for $log\pi_{1+}$ and $log\pi_{+1}$. Suppose $[l_\theta, u_\theta]$ is a $100(1-\alpha)\%$ confidence interval for θ. A $100(1-\alpha)\%$ CI for $log\theta$ can be obtained by substitution method as

$$[log(l_\theta), log(u_\theta)].$$

Hence, the three confidence intervals (i.e., the Wilson score, Agresti-Coull, and Jeffreys confidence intervals) for $\theta_1 = \pi_{1+}$ and $\theta_2 = \pi_{+1}$ described in Section 3.1 can be adopted here. Besides, the correlation coefficient estimate between $log(\hat{\pi}_{1+})$ and $log(\hat{\pi}_{+1})$, denoted as $\widehat{corr}(log\hat{\pi}_{1+}, log\hat{\pi}_{+1})$, can be obtained by the delta method. It can be easily shown that $\widehat{corr}(log\hat{\pi}_{1+}, log\hat{\pi}_{+1}) = \widehat{corr}(\hat{\pi}_{1+}, \hat{\pi}_{+1})$.

When data are complete, let ψ_1, ψ_2 be one and m_1, m_2, u and v be zero in the above equations, we can obtain the responding confidence intervals. This has been given by Tang et al. (2010).

4 Performance Evaluation Using Exact Approach

In this section, we investigate the performance of various confidence intervals in small to moderate sample sizes with respect to their exact coverage probabilities, expected confidence interval widths, and distal and mesial non-coverage probabilities. A summary of abbreviation for various confidence intervals are presented in Table 3.

Let $N = n + m_1 + m_2$ represent the total sample size. All these measures are then examined for small (e.g., N=20) and moderate (e.g., N=50) sample sizes via numerical evaluation. For each given total sample size N, we consider (i) (10% + 10% balanced missing data): $n = 0.8\times N$, and $m_1 = m_2 = 0.1\times N$; (ii) (20% + 0% imbalanced missing data): $n = 0.8\times N$, $m_1 = 0.2\times N$ and $m_2 = 0$; (iii) (20% + 20% balanced missing data): $n = 0.6\times N$, $m_1 = 0.2\times N$ and $m_2 = 0.2\times N$; (iv)(0% + 0% complete data): $n = N$, $m_1 = 0$ and $m_2 = 0$. Since the computing time for exact coverage probabilities is very tedious, we only consider $\delta_0 = 0.91$ and 1.1, $\pi_{+1} = 0.5$, and $N =$ 20 and 50. Here, $\pi_{1+} = \delta_0\pi_{+1}$. To introduce dependence/correlation between the paired binary outcomes, we assume the bivariate binary observations are coming from a bivariate distribution with the correlation coefficient defined by

$$\rho = (\pi_{11} - \pi_{1+}\pi_{+1})/[\pi_{1+}(1-\pi_{1+})\pi_{+1}(1-\pi_{+1})]^{1/2}.$$

Hence, given π_{1+}, π_{+1} and ρ, we have $\pi_{11} = \pi_{1+}\pi_{+1} + \rho[\pi_{1+}(1-\pi_{1+})\pi_{+1}(1-\pi_{+1})]^{1/2}$, $\pi_{01} = \pi_{+1}$ - π_{11}, $\pi_{10} = \delta\pi_{+1}$ - π_{11}, and $\pi_{00} = 1$ - $\delta\pi_{+1}$ - π_{01}. For the correlation coefficient, i.e., ρ, we consider $\rho = -0.9, -0.5, -0.1, 0.5, 0.9$.

For each configuration of (π_{+1}, δ_0, ρ), we can calculate $\boldsymbol{\pi}$=(π_{11}, π_{10}, π_{01}). For each given setting (n, m_1, m_2, $\boldsymbol{\pi}$), we can compute the corresponding exact coverage probabilities(ECP) by

$$ECP = \sum_{n_{10}=0}^{n} \sum_{n_{01}=0}^{n-n_{10}} \sum_{n_{11}=0}^{n-n_{10}-n_{01}} \sum_{u=0}^{m_1} \sum_{v=0}^{m2} I(L \leq \delta_0 \leq U) f(n_{11}, n_{10}, n_{01}, u, v | n, m_1, m_2, \boldsymbol{\pi}),$$

where $[L, U]$ is any of the ten confidence intervals under investigation, and

$$f(n_{11}, n_{10}, n_{01}, u, v | n, m_1, m_2, \boldsymbol{\pi}) = \frac{n!}{n_{11}! n_{10}! n_{01}! (n - n_{11} - n_{10} - n_{01})!} \frac{m_1!}{u!(m_1 - u)!} \frac{m_2!}{v!(m_2 - v)!} \pi_{11}^{n_{11}} \pi_{10}^{n_{10}} \pi_{01}^{n_{01}} (1 - \pi_{11} - \pi_{10} - \pi_{01})^{n - n_{11} - n_{10} - n_{01}} (\pi_{11} + \pi_{10})^u (1 - \pi_{11} - \pi_{10})^{m_1 - u} (\pi_{11} + \pi_{01})^v (1 - \pi_{11} - \pi_{01})^{m_2 - v}.$$

The corresponding expected confidence widths(ECW)is given by

$$ECW = \sum_{n_{10}=0}^{n} \sum_{n_{01}=0}^{n-n_{10}} \sum_{n_{11}=0}^{n-n_{10}-n_{01}} \sum_{u=0}^{m_1} \sum_{v=0}^{m2} (U - L) f(n_{11}, n_{10}, n_{01}, u, v | n, m_1, m_2, \boldsymbol{\pi}).$$

In addition, we will characterize the interval location by evaluating the mesial and distal non-coverage probabilities. The definitions of the Mesial Non-Coverage Probabilities (MNCP) and Distal Non-Coverage Probabilities (DNCP) are derived from the left non-coverage probability (LNCP) and the right non-coverage probability (RNCP) commonly used in the literature (see, Newcombe, 2011) Recall that LNCP and RNCP are defined by

$$LNCP = \sum_{n_{10}=0}^{n} \sum_{n_{01}=0}^{n-n_{10}} \sum_{n_{11}=0}^{n-n_{10}-n_{01}} \sum_{u=0}^{m_1} \sum_{v=0}^{m2} I(\delta_0 < L) f(n_{11}, n_{10}, n_{01}, u, v | n, m_1, m_2, \boldsymbol{\pi}),$$

and

$$RNCP = \sum_{n_{10}=0}^{n} \sum_{n_{01}=0}^{n-n_{10}} \sum_{n_{11}=0}^{n-n_{10}-n_{01}} \sum_{u=0}^{m_1} \sum_{v=0}^{m2} I(\delta_0 > U) f(n_{11}, n_{10}, n_{01}, u, v | n, m_1, m_2, \boldsymbol{\pi}).$$

The terms mesial and distal are defined relative to the true value of δ_0. For $\delta_0 > 1$, when the interval is too far to the right to include δ_0, this is sometimes referred to as non-coverage at the left or mesial end of the interval. Conversely, when the interval is too far to the left to include δ_0, this is sometimes referred to as non-coverage at the right or distal end of the interval. Consequently, the definitions of MNCP and DNCP are identical to LNCP and RNCP provided $\delta_0 > 1$. However, if $\delta_0 < 1$, when the interval is too far to the right to include δ_0, this is sometimes referred as non-coverage at the left or distal end of the interval. Conversely, when the interval is too far to the left to include δ_0, this is sometimes referred as non-coverage at the right or mesial end of the interval. Consequently, the definitions of MNCP and DNCP need to be interchanged here. When $\delta_0 = 1$, left and right non-coverage should be balanced.

The non-coverage probability (NCP) is the sum of these measures and the ratio $MNCP/(MNCP+DNCP) = MNCP/NCP$ can effectively separate the function of assessing location from assessment of overall coverage. For a balanced confidence interval, the ratio should be close to 0.5. We classify this ratio measure as satisfactory if it is between 0.4 and 0.6, the interval is too mesially located if it is below 0.4, and too distally located if it is above 0.6.

We expect good methods for constructing confidence intervals have their ECPs close to the pre-specified $1-\alpha$ level. When the $ECPs$ are well controlled, one then prefers confidence intervals with shorter widths: i.e., smaller ECW values. When the $ECWs$ are smaller, one would also prefer $MNCP/NCP$ to be between 0.4 and 0.6.

Results of numerical evaluations of these confidence intervals are presented in Tables 4, 5 and 6 for exact coverage probabilities, in Tables 7, 8 and 9 for expected confidence widths, and in Tables 10-12 for symmetry of non-coverage probabilities. From these results, we observe the following:

(1) In general, when the total sample size is larger, the expected confidence widths is narrower. Also, the confidence widths increase with proportion of missing observations.

(2) All confidence widths decrease as the correlation coefficient (i.e. ρ) increase and there is not any significant effect of ρ on $ECPs$.

(3)There is no any significant effect of δ on exact coverage probabilities and confidence widths.

(4)The asymptotic Wald, score, and log-test-based intervals can have substantial under-coverage probabilities when the correlations are extreme (i.e.$\rho = -0.9$ or 0.9). The likelihood ratio intervals can be overly conservative with high correlations (with $> 99\%$ coverage), which results in longer interval widths.

(5)The hybrid Wilson score confidence intervals (WCI and WCIlog)tend to have under-coverage probabilities. On the other hand, the hybrid Jeffrey confidence intervals (WCI and WCIlog) tend to be overly conservative ($> 99\%$ coverage in many cases) and have asymmetric non-coverage probabilities (MNCP/NCP > 0.6 in many cases). The hybrid Agresti-Coull confidence intervals (i.e., ACI and $ACIlog$) behave satisfactorily in the sense that they (i) generally well control their coverage probabilities around the pre-chosen confidence level; (ii) consistently yield shorter confidence widths; and (iii) usually guarantee their ratios of the $MNCP/NCP$ lying in the interval $[0.4, 0.6]$, indicating symmetry of the CI. In particular, if one would like a CI that yield the shortest confidence width, then the hybrid Agresti-Coull confidence intervals based on the Fieller's theorem is the optimal choice.

5 Two real examples

5.1 Osoba's example

We re-visit the study considered by Osoba et al. (1986). From Table 1, we have $n_{00} = 9$, $n_{01} = 23$, $n_{10} = 6$, $n_{11} = 77$, $n = 115$, $u = 14$, $m_1 = 16$, $v = 12$ and $m_2 = 26$. The ratio between the rates of vomiting experience after using MPRED and METCLO is estimated to be 0.9322. The 95% CIs for π_{1+}/π_{+1} based on various methods are summarized in Table 13. All CIs include 1, suggesting that the rates of vomiting experience between MPRED and METCLO are not significantly different.

5.2 A neurological study

A neurological study of meningitis patients reported in Choi and Stablein (1982) are revisited here to illustrated our proposed methodologies. According to our setting, we have $n_{00} = 6$, $n_{01} = 3$, $n_{10} = n_{11} = 8$, $n = 25$, $u = 4$, $m_1 = 6$, $v = 2$ and $m_2 = 2$. The ratio for incidence rates of neurological complication before and after the standard treatment $\delta = \pi_{1+}/\pi_{+1} = (\frac{n_{10}+n_{11}+u}{n+m_1})/(\frac{n_{01}+n_{11}+v}{n+m_2}) = 1.34$. The 95% CIs for π_{1+}/π_{+1} based on various methods are summarized in Table 14. Since all resulting CIs include the value 1, applying these confidence interval estimators leads to the conclusion that the incidence rates of neurological complication before and after the standard treatment are essentially the same. This result is consistent to Tang et al. (2009).

6 Discussion

In this article, we propose the use of a hybrid method for combining two individual confidence intervals for a single proportion to form a confidence interval for the ratio of the two proportions in the presence of incomplete data. We incorporate the hybrid method with (i) Fieller's theorem; (ii) logarithmic transformation. According to our numerical evaluation, the hybrid Agresti-Coull confidence intervals (i.e., ACI and $ACIlog$)behave satisfactorily. In particular, ACI generally yields the shortest confidence widths and the exact coverage probabilities are usually close to the pre-specified coverage level. Unlike the asymptotic score confidence interval, all hybrid confidence intervals described in this paper possess closed form solution. In terms of computational simplicity, they are more preferable than the asymptotic score confidence interval. It is also noteworthy that the asymptotic score confidence interval could produce overly inflated exact coverage probabilities, which may lead to reasonably wide expected widths. In view of the above observations, we highly recommend the hybrid Agresti-Coull confidence interval based on Fieller's theorem (i.e., ACI) in practice.

In order to compare with the method of Choi and Stablein(1982), we assume data are missing completely at random (MCAR) in this paper. This assumption is reasonable for some studies where missing data are mostly caused by loss-to-followup or invalid test results. For example, in a vaccine study comparing a new vaccine versus a placebo (or an active control vaccine), most missing data are generally related to out-of-day-range visits or loss-to-followup, and MCAR assumption is reasonable. In other studies, such as a crossover trial of drugs, patients missing treatment might be outcome related, and the missing at random assumption (or non-ignorable missing) may be more plausible. We are currently conducting further investigations on corresponding methods based on the missing at random and non-ignorable missing assumption.

ACKNOWLEDGMENTS

This research is fully supported by the Natural Science Foundation of China (11201412), and grants from the Science Foundation of Yunnan Province(2011FB016). The authors wish to thank associate editor and the two anonymous referees for the constructive suggestions and comments that substantially improved the original version of this article.

REFERENCES

1. Bartlett, M.S. (1953). Approximate confidence interval more than one unknown parameters. *Biometrika* **40**: 306–317.

2. Brown, L.D., Cai, T., and Dasgupta, A. (2001). Interval estimation for a binomial proportion (with discussion). *Statistical Science* **16**: 101–133.

3. Chan, I. S. F.(1998). Exact tests of equivalence and efficacy with a non-zero lower bound for comparative trials. *Statistics in Medicine* **17**: 1403–1413.

4. Chen, J., Tsong, Y, and Kang, S.H.(2000). Tests for equivalence or noninferiority between two proportions. *Drug information Journal* **34**: 569–578.

5. Choi, S.C. and Stablein, D.M.(1982). Practical tests for comparing two proportions with incomplete data. *Applied Statistics* **31**: 256–262.

6. Farrington, C.P., Manning, G.(1990). Test statistics and sample size formulae for comparative binomial trials with null hypothesis of non-zero risk difference or non-unity relative risk. *Statistics in Medicine* **9**: 1447–1454.

7. Lachenbruch, P.A. and Lynch, C. J.(1998). Assessing screening tests: extensions of McNemar's test. *Statistics in Medicine* **17**: 2207–2217.

8. Lin, Y., Lipsitz, S., Sinha, D., Gawande, A.A., Regenbogen, S.E., Greenberg, C.C. (2009). Using Bayesian p-values in a 2×2 table of matched pairs with incompletely classified data. *Journal of the Royal Statistics Society Series C* **58**: 237–246.

9. Miettinen, O., Nurminen, M.(1985). Comparative analysis of two rates. *Statistics in Medicine* **4**: 213–226.

10. Newcombe,R.G. (2011). Measures of location for confidence intervals for proportions. *Communications in Statistics-Theory and Methods* **40**: 1743–1767.

11. Tang, M.L., Tang, N.S.(2004). Exact tests for comparing two paired proportions with incomplete data. *Biometrical Journal* **46**: 72–82.

12. Tang, M.L., Li, H.Q., Tang, N.S. (2010). Confidence interval construction for proportion ratio in paired studies based on hybrid method. *Statistical Methods in Medical Research* **21**: 361–378.

13. Tang, M.L., Ling, M.H., Tian, G.L.(2009). Exact and approximate unconditional confidence intervals for proportion difference in the presence of incomplete data. *Statistics in Medicine* **28**: 625–641.

14. Tang, M.L., Tang, N.S., Chan, I.S.F., and Chan, B.P.S.(2002). Sample size determination for establishing equivalence/noninferiority via ratio of two proportions in matched-pair design. *Biometrics* **58**: 957–963.

15. Tang, N.S., Tang, M.L., Chan, I.S.F. (2003). On tests of equivalence via non-unity relative risk for matched-pair design. *Statistics in Medicine* **22**: 1217–1233.

16. Tango, T.(1998). Equivalence test and confidence interval for the difference in proportions for the paired-sample design. *Statistics in Medicine* **17**: 891–908.

17. Zou, G.Y., Donner, A.(2008). Construction of confidence limits about effect measures: A general approach. *Statistics in Medicine* **27**: 1693–1702.

Table 1: Clinical data for comparing MPRED and METCLO

	$Y=0$	$Y=1$	Subtotal	Supplement on X	Total
$X=0$	9	23	32	2	34
$X=1$	6	77	83	14	97
Subtotal	15	100	115	16	131
Supplement on Y	14	12	26		
Total	29	112	141		157

Table 2: Observed counts and cell probabilities for matched-pair design with incomplete data

	$Y=0$	$Y=1$	Subtotal	Supplement on X	Total
$X=0$	$n_{00}(\pi_{00})$	$n_{01}(\pi_{01})$	$n_{0+}(\pi_{0+})$	m_1-u	$n_{0+}+m_1-u$
$X=1$	$n_{10}(\pi_{10})$	$n_{11}(\pi_{11})$	$n_{1+}(\pi_{1+})$	u	$n_{1+}+u$
Subtotal	$n_{+0}(\pi_{+0})$	$n_{+1}(\pi_{+1})$	$n(1.0)$	m_1	$n+m_1$
Supplement on Y	m_2-v	v	m_2		
Total	$n_{+0}+m_2-v$	$n_{+1}+v$	$n+m_2$		$n+m_1+m_2$

Table 3 Summary of abbreviations for various confidence interval estimators

Abbreviation	Confidence interval CI
TlCI	Asymptotic Likelihood-ratio-test-based CI
TsCI	Asymptotic Score-test-based CI
TwCI	Asymptotic Wald-test-based CI
TlogCI	Asymptotic Log-test-based CI
ACI	Hybrid Agresti-Coull CI based on Fieller's theorem
WCI	Hybrid Wilson score CI based on Fieller's theorem
JCI	Hybrid Jeffrey CI based on Fieller's theorem
$ACIlog$	Hybrid Agresti-Coull CI based on logarithmic transformed method
$WCIlog$	Hybrid Wilson score CI based on logarithmic transformed method
$JCIlog$	Hybrid Jeffrey CI based on logarithmic transformed method

Table 4. Exact coverage probabilities(percent) of various 95 percent confidence intervals for δ with $N = 20$ and $\pi_{+1} = 0.5$ under incomplete data

n	m_1	m_2	δ	ρ	TlCI	TsCI	TwCI	TlogCI	ACI	WCI	JCI	$ACIlog$	$WCIlog$	$JCIlog$
12	4	4	0.91	-0.9	96.72	86.41	99.90	95.14	94.84	93.50	97.29	94.83	93.44	97.28
				-0.5	95.52	89.86	99.88	95.14	94.47	93.18	97.56	94.39	92.99	97.47
				-0.1	96.96	93.53	99.74	95.06	94.40	93.02	98.35	94.20	92.87	98.20
				0.1	95.53	95.32	99.48	94.97	94.50	93.12	98.76	94.25	93.02	98.59
				0.5	95.25	98.38	97.55	94.38	95.13	94.11	99.46	94.86	94.10	99.30
				0.9	97.24	99.92	90.92	89.75	95.18	97.20	99.95	94.91	96.64	99.85
			1.1	-0.9	94.92	86.70	99.43	94.96	94.90	93.41	96.75	94.89	93.39	96.74
				-0.5	95.60	90.25	99.40	94.97	94.49	92.93	96.82	94.45	92.82	96.77
				-0.1	96.90	93.85	99.40	94.92	94.47	92.75	97.76	94.34	92.57	97.63
				0.1	95.84	95.58	99.34	94.83	94.61	92.85	98.26	94.43	92.69	98.10
				0.5	95.64	98.49	98.24	94.19	95.35	93.71	99.19	95.11	93.63	99.00
				0.9	97.84	99.93	91.58	89.59	95.51	96.92	99.94	95.30	96.91	99.86
16	2	2	0.91	-0.9	94.83	85.36	99.96	95.58	95.23	94.29	96.29	95.22	94.18	96.29
				-0.5	95.70	89.28	99.94	95.19	94.68	93.92	96.33	94.49	93.83	96.17
				-0.1	96.68	93.56	99.68	95.01	94.19	93.42	96.66	94.07	93.32	96.34
				0.1	96.99	95.63	99.17	94.87	94.10	93.35	97.04	93.99	93.24	96.65
				0.5	95.32	98.92	95.55	93.69	94.98	94.35	98.33	94.86	94.40	97.91
				0.9	99.00	99.98	99.84	98.11	97.51	97.29	99.91	97.80	97.86	99.86
			1.1	-0.9	95.93	83.63	99.29	95.49	95.20	94.32	95.15	95.19	94.32	95.15
				-0.5	95.61	88.39	99.20	95.10	94.66	93.77	95.84	94.56	93.73	95.69
				-0.1	96.37	93.10	99.14	94.91	94.20	93.39	96.21	94.07	93.28	95.89
				0.1	96.70	95.33	99.05	94.78	94.13	93.30	96.62	94.04	93.22	96.23
				0.5	96.35	98.85	97.22	93.57	94.90	94.09	98.03	95.09	94.26	97.62
				0.9	99.23	99.98	91.53	98.28	94.78	97.56	99.96	95.33	97.82	99.94
16	4	0	0.91	-0.9	95.16	83.97	99.97	96.57	95.20	95.26	93.59	95.18	95.25	93.58
				-0.5	96.23	88.60	99.99	96.75	94.58	94.80	93.69	94.51	94.71	93.68
				-0.1	96.89	93.24	99.97	97.03	94.08	94.38	94.43	93.95	94.32	94.31
				0.1	94.07	95.44	99.93	97.31	93.98	94.28	95.01	93.81	94.24	94.70
				0.5	96.76	98.90	99.68	98.26	94.83	94.83	96.35	94.75	94.88	95.69
				0.9	98.20	99.98	98.14	99.84	94.82	97.73	98.45	95.08	98.09	98.61
			1.1	-0.9	95.54	84.75	97.90	95.95	95.15	95.79	94.50	95.14	95.77	94.50
				-0.5	96.22	89.46	98.44	96.46	94.60	94.89	94.37	94.55	94.87	94.28
				-0.1	94.07	93.74	98.82	96.74	94.21	94.31	94.83	94.03	94.29	94.57
				0.1	96.12	95.76	99.01	96.94	94.17	94.20	95.25	94.01	94.18	94.90
				0.5	95.43	98.92	99.37	97.70	95.00	94.98	97.08	95.27	95.14	96.64
				0.9	99.06	99.98	99.56	98.23	95.35	98.15	99.69	95.37	99.11	99.52

Table 5. Exact coverage probabilities(percent) of various 95 percent confidence intervals for δ with $N = 50$ and $\pi_{+1} = 0.5$ under incomplete data

n	m_1	m_2	δ	ρ	TlCI	TsCI	TwCI	TlogCI	ACI	WCI	JCI	$ACIlog$	$WCIlog$	$JCIlog$
30	10	10	0.91	-0.9	95.66	86.24	99.52	95.11	94.93	93.32	97.57	94.93	93.30	97.56
				-0.5	96.42	90.02	99.33	95.12	94.74	93.36	98.57	94.71	93.33	98.55
				-0.1	96.22	93.75	98.93	95.07	94.59	93.26	99.29	94.55	93.23	99.27
				0.1	95.41	95.53	98.58	95.04	94.52	93.25	99.54	94.48	93.21	99.52
				0.5	96.51	98.51	97.36	94.91	94.51	93.41	99.86	94.49	93.39	99.83
				0.9	99.50	99.93	94.52	93.89	95.43	94.42	99.98	95.59	94.85	99.96
			1.1	-0.9	96.09	85.89	96.82	95.11	95.06	93.13	96.81	95.06	93.12	96.79
				-0.5	95.01	89.67	96.91	95.05	94.75	93.05	98.01	94.71	93.02	97.99
				-0.1	95.53	93.51	96.87	95.02	94.62	92.88	98.90	94.56	92.86	98.87
				0.1	95.62	95.35	96.77	95.00	94.55	92.79	99.24	94.50	92.77	99.21
				0.5	96.20	98.45	96.06	94.88	94.54	92.74	99.72	94.53	92.73	99.67
				0.9	99.47	99.92	92.70	93.94	95.43	94.10	99.95	95.67	94.15	99.91
40	5	5	0.91	-0.9	95.93	86.48	99.23	95.12	95.06	94.50	96.43	95.05	94.50	96.43
				-0.5	96.56	90.24	99.09	95.12	94.81	94.21	96.97	94.77	94.17	96.93
				-0.1	96.45	94.21	98.70	95.03	94.60	94.05	97.52	94.57	93.99	97.40
				0.1	96.13	96.14	98.33	94.97	94.50	93.98	97.87	94.47	93.90	97.71
				0.5	95.49	99.14	96.82	94.74	94.36	93.95	98.72	94.37	93.89	98.46
				0.9	99.22	99.99	91.96	92.28	95.45	95.40	99.68	95.79	95.68	99.32
			1.1	-0.9	96.05	85.74	96.55	95.02	94.95	94.36	96.19	94.95	94.35	96.17
				-0.5	95.89	89.73	96.65	95.06	94.79	94.09	96.57	94.78	94.08	96.55
				-0.1	95.96	93.94	96.55	95.00	94.62	93.88	97.10	94.57	93.85	96.99
				0.1	95.66	95.95	96.36	94.94	94.52	93.78	97.43	94.47	93.74	97.29
				0.5	96.69	99.09	95.26	94.72	94.38	93.62	98.30	94.42	93.64	98.03
				0.9	99.62	99.99	85.15	91.99	95.98	95.33	99.42	95.11	95.65	99.02
40	10	0	0.91	-0.9	96.73	85.55	98.99	95.70	94.93	95.48	90.88	94.93	95.40	90.88
				-0.5	95.09	89.62	99.03	95.72	94.82	95.14	91.94	94.77	95.08	91.93
				-0.1	94.76	93.87	98.97	95.84	94.60	94.87	93.31	94.55	94.81	93.26
				0.1	94.57	95.91	98.86	95.96	94.48	94.74	94.14	94.44	94.67	94.04
				0.5	96.31	99.09	98.33	96.43	94.36	94.51	96.04	94.33	94.48	95.78
				0.9	97.40	99.99	97.12	97.29	95.75	95.77	97.87	95.97	96.02	97.32
			1.1	-0.9	96.21	84.99	96.39	95.47	94.97	95.25	92.61	94.97	95.24	92.61
				-0.5	95.84	89.41	96.61	95.66	94.80	94.96	93.52	94.79	94.95	93.51
				-0.1	96.20	93.72	96.81	95.76	94.62	94.67	94.66	94.58	94.69	94.56
				0.1	95.98	95.77	96.88	95.86	94.54	94.52	95.30	94.50	94.53	95.15
				0.5	95.18	99.00	97.00	96.21	94.51	94.26	96.80	94.53	94.28	96.54
				0.9	96.79	99.99	96.68	96.66	95.63	95.66	99.11	94.93	96.17	98.63

Table 6. Exact coverage probabilities(percent) of various 95 percent confidence intervals for δ with $N = 20, 50$ and $\pi_{+1} = 0.5$ under complete data

n	m_1	m_2	δ	ρ	TlCI	TsCI	TwCI	TlogCI	ACI	WCI	JCI	$ACIlog$	$WCIlog$	$JCIlog$
20	0	0	0.91	-0.9	96.09	85.72	99.67	96.53	94.69	94.69	95.57	94.68	94.69	95.57
				-0.5	98.44	89.88	99.98	96.61	94.75	94.36	95.34	94.61	94.29	95.34
				-0.1	95.89	94.27	99.82	96.82	94.10	94.36	94.94	94.06	93.69	94.91
				0.1	94.50	96.36	99.59	97.03	94.88	94.29	94.78	93.82	93.45	94.74
				0.5	94.47	99.39	98.35	98.19	94.60	94.92	95.75	95.03	94.91	95.57
				0.9	95.53	99.99	99.99	99.99	95.87	99.91	99.95	99.90	99.90	99.93
			1.1	-0.9	94.42	86.09	98.62	96.52	94.69	94.66	95.69	94.69	94.66	95.69
				-0.5	94.45	90.32	99.04	96.40	95.05	94.54	95.38	95.04	94.52	95.38
				-0.1	95.52	94.47	99.31	96.76	94.50	94.30	94.83	94.39	94.03	94.81
				0.1	96.03	96.49	99.45	97.02	94.28	94.16	94.77	94.19	93.82	94.70
				0.5	96.45	99.41	99.78	98.11	94.86	94.92	96.02	95.46	95.31	95.84
				0.9	98.04	99.99	93.00	99.99	94.93	99.89	99.95	99.90	99.89	99.91
50	0	0	0.91	-0.9	95.76	84.77	98.87	95.64	95.02	95.02	95.75	95.02	95.02	95.71
				-0.5	95.82	89.06	98.92	95.70	94.85	94.83	95.77	94.82	94.78	95.75
				-0.1	94.05	93.81	98.68	95.78	94.63	94.57	95.60	94.55	94.56	95.58
				0.1	94.72	96.08	98.43	95.88	94.50	94.47	95.53	94.42	94.44	95.51
				0.5	94.42	99.39	97.48	96.38	94.28	94.32	95.42	94.22	94.22	95.46
				0.9	98.37	99.99	96.39	95.08	94.49	94.55	94.84	98.27	98.28	98.57
			1.1	-0.9	94.36	85.39	96.71	95.74	95.06	95.06	95.95	95.06	95.06	95.95
				-0.5	94.28	89.63	96.72	95.62	94.82	94.81	95.72	94.82	94.78	95.72
				-0.1	94.16	94.10	96.84	95.75	94.65	94.62	95.60	94.61	94.55	95.60
				0.1	94.35	96.26	96.80	95.86	94.53	94.48	95.52	94.46	94.48	95.51
				0.5	94.78	99.42	96.57	96.39	94.26	94.33	95.40	94.37	94.44	95.40
				0.9	98.62	99.99	93.02	94.16	94.47	92.47	92.65	99.26	99.26	99.38

Table 7. Expected confidence widths of various 95 percent confidence intervals for δ with $N = 20$ and $\pi_{+1} = 0.5$ under incomplete data

n	m_1	m_2	δ	ρ	TlCI	TsCI	TwCI	TlogCI	ACI	WCI	JCI	$ACIlog$	$WCIlog$	$JCIlog$
12	4	4	0.91	-0.9	1.9278	1.8287	2.1543	2.3409	2.3126	2.1582	1.8584	2.3113	2.1568	1.8562
				-0.5	1.8706	1.7733	1.9186	2.0296	2.0233	2.1276	1.6634	2.0262	2.1293	1.6591
				-0.1	1.8109	1.7179	1.6654	1.7063	1.7258	1.9040	1.4655	1.7378	1.9141	1.4576
				0.1	1.6005	1.6902	1.5296	1.5379	1.5722	1.7121	1.3648	1.5908	1.7277	1.3539
				0.5	1.5323	1.6348	1.2296	1.1789	1.2589	1.2593	1.1645	1.2947	1.2858	1.1438
				0.9	1.0755	1.0763	0.8592	0.7647	0.9648	0.8923	0.9861	1.0214	0.9254	0.9492
			1.1	-0.9	2.4320	2.0385	2.2293	2.5228	2.5291	2.1497	1.9182	2.5285	2.1491	1.9177
				-0.5	1.9943	1.9834	1.9731	2.2033	2.2372	2.0990	1.7215	2.2410	2.1016	1.7194
				-0.1	1.9248	1.9284	1.6939	1.8697	1.9346	2.0200	1.5184	1.9485	2.0317	1.5142
				0.1	1.8994	1.9009	1.5421	1.6952	1.7777	1.9171	1.4141	1.7992	1.9360	1.4082
				0.5	1.6772	1.8460	1.2001	1.3213	1.4564	1.5548	1.2046	1.4987	1.5914	1.1921
				0.9	1.2911	1.3009	0.7617	0.8878	1.1506	1.0982	1.0131	1.2196	1.1517	0.9885
16	2	2	0.91	-0.9	2.0621	1.6658	2.0674	2.2313	2.2079	2.0915	2.1214	2.2066	2.0903	2.1196
				-0.5	1.8115	1.6151	1.8196	1.9085	1.9077	1.8104	1.8340	1.9087	1.8109	1.8320
				-0.1	1.5643	1.5644	1.5497	1.5702	1.5902	1.5224	1.5315	1.5997	1.5300	1.5313
				0.1	1.4309	1.5390	1.4026	1.3917	1.4217	1.3625	1.3722	1.4383	1.3760	1.3730
				0.5	1.0832	1.4883	1.0655	0.9990	1.0624	1.0180	1.0383	1.1007	1.0489	1.0378
				0.9	0.8234	0.8377	0.5949	0.4765	0.6990	0.6794	0.7153	0.7675	0.7334	0.6991
			1.1	-0.9	1.9675	1.8536	2.1436	2.4024	2.4056	2.2297	2.2495	2.4048	2.2289	2.2491
				-0.5	1.9023	1.8032	1.8735	2.0701	2.1010	1.9496	1.9616	2.1028	1.9510	1.9613
				-0.1	1.7582	1.7528	1.5747	1.7196	1.7764	1.6669	1.6541	1.7876	1.6765	1.6575
				0.1	1.6207	1.7277	1.4090	1.5337	1.6036	1.5145	1.4907	1.6227	1.5311	1.4969
				0.5	1.5704	1.6774	1.0179	1.1222	1.2337	1.1706	1.1442	1.2779	1.2094	1.1558
				0.9	0.9107	0.9271	0.4257	0.5753	0.8534	0.8078	0.7983	0.9367	0.8820	0.8080
16	4	0	0.91	-0.9	1.9649	1.7694	1.7615	1.9088	2.3332	2.2390	2.8839	2.3322	2.2401	2.8803
				-0.5	1.9019	1.7091	1.5808	1.6719	2.0128	2.1238	2.4425	2.0147	2.1253	2.4418
				-0.1	1.6458	1.6489	1.3834	1.4234	1.6743	1.7268	1.9829	1.6863	1.7378	1.9942
				0.1	1.6076	1.6188	1.2762	1.2927	1.4946	1.5127	1.7437	1.5150	1.5324	1.7638
				0.5	1.5409	1.5587	1.0373	1.0111	1.1108	1.1267	1.2517	1.1572	1.1787	1.2899
				0.9	0.8654	0.8586	0.7451	0.6804	0.7191	0.7604	0.7923	0.8033	0.8788	0.8301
			1.1	-0.9	2.0987	1.9839	1.8591	2.0558	2.5556	2.0999	3.0548	2.5551	2.1087	3.0531
				-0.5	1.9293	1.9239	1.6576	1.8120	2.2319	2.0644	2.6288	2.2344	2.0680	2.6306
				-0.1	1.8907	1.8639	1.4361	1.5553	1.8877	1.8737	2.1794	1.9007	1.8846	2.1951
				0.1	1.8430	1.8340	1.3151	1.4199	1.7046	1.6851	1.9434	1.7266	1.7038	1.9702
				0.5	1.7014	1.7741	1.0420	1.1275	1.3132	1.2568	1.4505	1.3630	1.3028	1.5067
				0.9	1.1534	1.1143	0.6982	0.7843	0.9093	0.8858	0.9678	1.0023	0.9899	1.0547

Table 8. Expected confidence widths of various 95 percent confidence intervals for δ with $N = 50$ and $\pi_{+1} = 0.5$ under incomplete data

n	m_1	m_2	δ	ρ	TlCI	TsCI	TwCI	TlogCI	ACI	WCI	JCI	*ACIlog*	*WCIlog*	*JCIlog*
30	10	10	0.91	-0.9	1.0326	0.9362	1.2006	1.2078	1.1996	1.1018	1.0279	1.1992	1.1014	1.0274
				-0.5	0.9965	0.9265	1.0851	1.0748	1.0719	0.9871	0.9851	1.0722	0.9872	0.9838
				-0.1	0.9186	0.9168	0.9589	0.9320	0.9335	0.8627	0.9406	0.9358	0.8640	0.9378
				0.1	0.9002	0.9120	0.8904	0.8554	0.8590	0.7956	0.9176	0.8630	0.7980	0.9139
				0.5	0.8932	0.9023	0.7381	0.6863	0.6950	0.6483	0.8704	0.7041	0.6538	0.8639
				0.9	0.8926	0.9062	0.5513	0.4791	0.5156	0.4881	0.8250	0.5331	0.4985	0.8146
			1.1	-0.9	1.0380	1.0380	1.2436	1.3157	1.3142	1.1914	1.0455	1.3140	1.1911	1.0451
				-0.5	1.0283	1.0283	1.1132	1.1747	1.1797	1.0712	0.9964	1.1804	1.0717	0.9952
				-0.1	1.0187	1.0187	0.9686	1.0225	1.0336	0.9398	0.9448	1.0368	0.9420	0.9421
				0.1	1.0138	1.0138	0.8889	0.9406	0.9547	0.8685	0.9178	0.9598	0.8721	0.9142
				0.5	1.0042	1.0042	0.7068	0.7591	0.7804	0.7102	0.8616	0.7918	0.7184	0.8551
				0.9	0.9946	0.9946	0.4685	0.5362	0.5906	0.5370	0.8069	0.6123	0.5531	0.7959
40	5	5	0.91	-0.9	1.1253	0.9743	1.1658	1.1741	1.1651	1.1248	1.1153	1.1648	1.1245	1.1149
				-0.5	0.9982	0.9649	1.0417	1.0319	1.0283	0.9938	0.9952	1.0284	0.9938	0.9938
				-0.1	0.9757	0.8555	0.9043	0.8773	0.8781	0.8498	0.8657	0.8800	0.8513	0.8625
				0.1	0.9065	0.8508	0.8286	0.7932	0.7960	0.7710	0.7965	0.7996	0.7738	0.7916
				0.5	0.8655	0.6414	0.6552	0.6016	0.6093	0.5919	0.6455	0.6186	0.5993	0.6341
				0.9	0.4809	0.4920	0.4216	0.3390	0.3783	0.3713	0.4790	0.4001	0.3886	0.4506
			1.1	-0.9	0.9949	0.9684	1.2116	1.2786	1.2745	1.2194	1.1794	1.2743	1.2192	1.1791
				-0.5	0.9921	0.9590	1.0715	1.1276	1.1302	1.0829	1.0529	1.1306	1.0832	1.0519
				-0.1	0.9136	0.9496	0.9139	0.9626	0.9713	0.9315	0.9154	0.9739	0.9337	0.9132
				0.1	0.8994	0.9449	0.8254	0.8723	0.8840	0.8480	0.8411	0.8886	0.8519	0.8377
				0.5	0.8484	0.9355	0.6154	0.6661	0.6845	0.6562	0.6763	0.6961	0.6664	0.6679
				0.9	0.5383	0.5261	0.2942	0.3832	0.4392	0.4191	0.4911	0.4665	0.4437	0.4679
40	10	0	0.91	-0.9	0.9943	0.8934	1.1145	1.1270	1.1915	1.1649	1.4044	1.1912	1.1646	1.4034
				-0.5	0.9882	0.8828	1.0003	0.9972	1.0506	1.0324	1.2249	1.0508	1.0326	1.2232
				-0.1	0.8932	0.8723	0.8742	0.8563	0.8958	0.8842	1.0356	0.8982	0.8864	1.0336
				0.1	0.8745	0.8670	0.8050	0.7799	0.8112	0.8022	0.9364	0.8154	0.8063	0.9339
				0.5	0.7546	0.8564	0.6480	0.6079	0.6183	0.6135	0.7249	0.6293	0.6243	0.7185
				0.9	0.6009	0.6459	0.4459	0.3840	0.3781	0.3759	0.4996	0.4040	0.4020	0.4759
			1.1	-0.9	1.0048	0.9984	1.1684	1.2300	1.3109	1.2160	1.4779	1.3107	1.2158	1.4774
				-0.5	1.0008	0.9878	1.0413	1.0930	1.1630	1.0879	1.2973	1.1634	1.0881	1.2968
				-0.1	0.9745	0.9773	0.8992	0.9439	1.0003	0.9427	1.1045	1.0031	0.9449	1.1055
				0.1	0.9202	0.9720	0.8203	0.8629	0.9110	0.8614	1.0018	0.9160	0.8655	1.0041
				0.5	0.8641	0.9114	0.6371	0.6804	0.7075	0.6723	0.7779	0.7200	0.6828	0.7827
				0.9	0.6509	0.6690	0.3849	0.4449	0.4586	0.4354	0.5284	0.4866	0.4599	0.5326

Table 9. Expected confidence widths of various 95 percent confidence intervals for δ with $N = 20, 50$ and $\pi_{+1} = 0.5$ under complete data

n	m_1	m_2	δ	ρ	TlCI	TsCI	TwCI	TlogCI	ACI	WCI	JCI	$ACIlog$	$WCIlog$	$JCIlog$
20	0	0	0.91	-0.9	1.5405	1.5370	1.7070	1.8316	2.1005	2.1046	2.3419	2.0998	2.1039	2.3404
				-0.5	1.4976	1.4912	1.5217	1.5913	1.8011	1.8585	1.9999	1.8013	1.8586	2.0003
				-0.1	1.4502	1.4454	1.3176	1.3375	1.4787	1.5483	1.6339	1.4862	1.5562	1.6443
				0.1	1.4198	1.4226	1.2056	1.2028	1.3047	1.3593	1.4373	1.3192	1.3744	1.4572
				0.5	1.1734	1.1768	0.9507	0.9067	0.9194	0.9334	1.0046	0.9575	0.9711	1.0568
				0.9	0.7296	0.7311	0.6237	0.5342	0.4847	0.4852	0.5195	0.5664	0.5610	0.6342
			1.1	-0.9	1.7109	1.7082	1.7856	1.9592	2.2835	2.2232	2.5465	2.2831	2.2228	2.5457
				-0.5	1.6701	1.6627	1.5769	1.7109	1.9785	1.9671	2.1999	1.9794	1.9678	2.2011
				-0.1	1.6189	1.6172	1.3447	1.4471	1.6476	1.6945	1.8251	1.6567	1.7039	1.8371
				0.1	1.5894	1.5945	1.2158	1.3065	1.4681	1.5245	1.6225	1.4851	1.5423	1.6447
				0.5	1.2478	1.2491	0.9165	0.9959	1.0687	1.1034	1.1737	1.1135	1.1501	1.2322
				0.9	0.8235	0.8337	0.5067	0.6041	0.6124	0.6129	0.6631	0.7136	0.7160	0.7963
50	0	0	0.91	-0.9	1.0231	0.8227	1.0816	1.0878	1.1288	1.1297	1.2079	1.1286	1.1296	1.2077
				-0.5	0.9264	0.8138	0.9645	0.9553	0.9879	0.9888	1.0560	0.9878	0.9888	1.0559
				-0.1	0.8765	0.8049	0.8340	0.8102	0.8316	0.8323	0.8879	0.8332	0.8338	0.8898
				0.1	0.7992	0.8005	0.7616	0.7306	0.7451	0.7457	0.7950	0.7483	0.7487	0.7989
				0.5	0.7688	0.7916	0.5945	0.5472	0.5429	0.5431	0.5784	0.5520	0.5520	0.5899
				0.9	0.4567	0.4827	0.3657	0.2843	0.2533	0.2539	0.2687	0.2829	0.2821	0.3064
			1.1	-0.9	1.0976	0.9105	1.1233	1.1799	1.2338	1.2271	1.3213	1.2336	1.2270	1.3211
				-0.5	1.0481	0.9016	0.9915	1.0391	1.0849	1.0805	1.1605	1.0850	1.0806	1.1607
				-0.1	0.9432	0.8927	0.8423	0.8843	0.9191	0.9164	0.9820	0.9212	0.9184	0.9846
				0.1	0.8850	0.8883	0.7581	0.7991	0.8269	0.8247	0.8829	0.8309	0.8287	0.8877
				0.5	0.7608	0.7794	0.5563	0.6020	0.6100	0.6085	0.6502	0.6215	0.6200	0.6640
				0.9	0.4235	0.4705	0.2352	0.3196	0.3029	0.3013	0.3214	0.3394	0.3380	0.3650

Table 10. $MNCP/NCP$ of various 95 percent confidence intervals for δ with $N = 20$ and $\pi_{+1} = 0.5$ under incomplete data

n	m_1	m_2	δ	ρ	TlCI	TsCI	TwCI	TlogCI	ACI	WCI	JCI	$ACIlog$	$WCIlog$	$JCIlog$
12	4	4	0.91	-0.9	0.4906	0.5701	0.3852	0.5413	0.5146	0.5138	0.5807	0.5143	0.5094	0.5805
				-0.5	0.4819	0.5693	0.6447	0.5302	0.5229	0.5304	0.6968	0.5167	0.5181	0.6987
				-0.1	0.4631	0.5723	0.6062	0.5503	0.5409	0.5481	0.7820	0.5260	0.5369	0.7840
				0.1	0.4508	0.5740	0.6676	0.5648	0.5547	0.5616	0.6320	0.5352	0.5492	0.6334
				0.5	0.4353	0.5649	0.6961	0.6244	0.5937	0.5988	0.6316	0.5612	0.5647	0.6357
				0.9	0.4799	0.1507	0.6987	0.7174	0.5889	0.6230	0.6080	0.5431	0.5201	0.6473
			1.1	-0.9	0.4892	0.5591	0.6230	0.5270	0.5174	0.5118	0.5814	0.5172	0.5108	0.5774
				-0.5	0.4863	0.5645	0.6017	0.5363	0.5233	0.5201	0.6958	0.5211	0.5137	0.6933
				-0.1	0.4778	0.5705	0.6739	0.5570	0.5426	0.5307	0.7798	0.5317	0.5222	0.7766
				0.1	0.4674	0.5734	0.7751	0.5714	0.5586	0.5401	0.6343	0.5411	0.5326	0.6271
				0.5	0.4400	0.5661	0.3524	0.6326	0.6049	0.5662	0.6488	0.5676	0.5600	0.6410
				0.9	0.5131	0.1298	0.2295	0.7356	0.5794	0.4839	0.6960	0.5369	0.4869	0.6962
16	2	2	0.91	-0.9	0.4934	0.4654	0.2032	0.5057	0.5314	0.5164	0.5251	0.5309	0.5066	0.5245
				-0.5	0.4935	0.4759	0.7238	0.5227	0.5196	0.5141	0.5750	0.5215	0.5065	0.5592
				-0.1	0.4799	0.4816	0.6740	0.5493	0.5261	0.5263	0.6266	0.5217	0.5201	0.6045
				0.1	0.4671	0.4833	0.6932	0.5754	0.5423	0.5438	0.6763	0.5302	0.5329	0.6519
				0.5	0.4260	0.4679	0.6993	0.6935	0.6134	0.6282	0.6436	0.5794	0.5740	0.6078
				0.9	0.2429	0.2376	0.5999	0.8202	0.5721	0.8748	0.6526	0.5249	0.7571	0.6332
			1.1	-0.9	0.4951	0.5023	0.2835	0.5155	0.5201	0.5049	0.5818	0.5194	0.5049	0.5814
				-0.5	0.4834	0.4926	0.6680	0.5260	0.5221	0.5192	0.5629	0.5289	0.5166	0.5586
				-0.1	0.4813	0.4892	0.6667	0.5520	0.5289	0.5247	0.6171	0.5341	0.5187	0.5984
				0.1	0.4699	0.4876	0.7358	0.5807	0.5452	0.5351	0.6707	0.5429	0.5301	0.6480
				0.5	0.4155	0.4648	0.3203	0.7051	0.6249	0.6008	0.6425	0.5832	0.5845	0.6072
				0.9	0.2797	0.2416	0.3286	0.8189	0.5457	0.7541	0.6302	0.5813	0.7203	0.6182
16	4	0	0.91	-0.9	0.4193	0.4907	0.4000	0.5098	0.5178	0.5797	0.6341	0.5199	0.5807	0.6340
				-0.5	0.4373	0.4818	0.6203	0.4971	0.5316	0.5557	0.6207	0.5368	0.5627	0.6198
				-0.1	0.4398	0.4744	0.6368	0.5104	0.5379	0.5594	0.6253	0.5368	0.5602	0.6218
				0.1	0.4336	0.4691	0.6921	0.5246	0.5487	0.5709	0.6420	0.5424	0.5664	0.6366
				0.5	0.3745	0.4351	0.6000	0.6138	0.6253	0.6428	0.6869	0.6185	0.6255	0.6837
				0.9	0.2030	0.2694	0.5007	0.6997	0.5369	0.8429	0.6977	0.6016	0.7970	0.6984
			1.1	-0.9	0.5312	0.4984	0.5312	0.6512	0.5442	0.4710	0.1731	0.5452	0.4688	0.1732
				-0.5	0.5406	0.5043	0.6406	0.6602	0.5191	0.4577	0.1752	0.5177	0.4594	0.1747
				-0.1	0.5236	0.5077	0.6236	0.7043	0.5279	0.4797	0.2117	0.5052	0.4764	0.2115
				0.1	0.5122	0.5093	0.6122	0.7487	0.5456	0.5015	0.2289	0.5086	0.4878	0.2295
				0.5	0.4778	0.4980	0.6778	0.6776	0.6040	0.5930	0.2455	0.5314	0.5222	0.2416
				0.9	0.7079	0.5997	0.6079	0.6951	0.5648	0.7784	0.6255	0.5170	0.5921	0.6204

Table 11. $MNCP/NCP$ of various 95 percent confidence intervals for δ with $N = 50$ and $\pi_{+1} = 0.5$ under incomplete data

n	m_1	m_2	δ	ρ	TlCI	TsCI	TwCI	TlogCI	ACI	WCI	JCI	$ACIlog$	$WCIlog$	$JCIlog$
30	10	10	0.91	-0.9	0.5345	0.5221	0.3816	0.5263	0.5069	0.5233	0.7314	0.5070	0.5236	0.7285
				-0.5	0.5292	0.5188	0.7009	0.5187	0.5109	0.5149	0.7732	0.5111	0.5128	0.7732
				-0.1	0.5137	0.5176	0.6038	0.5282	0.5155	0.5213	0.6261	0.5120	0.5158	0.6240
				0.1	0.5104	0.5173	0.6527	0.5368	0.5194	0.5267	0.6571	0.5144	0.5179	0.6534
				0.5	0.5281	0.5112	0.6920	0.5695	0.5441	0.5533	0.6491	0.5314	0.5280	0.6445
				0.9	0.6093	0.3290	0.6990	0.6479	0.6101	0.6433	0.6991	0.5633	0.5554	0.6989
			1.1	-0.9	0.4917	0.5025	0.9408	0.5140	0.5137	0.5079	0.7251	0.5135	0.5083	0.7218
				-0.5	0.4897	0.5011	0.8721	0.5193	0.5110	0.5105	0.7627	0.5103	0.5103	0.7624
				-0.1	0.4872	0.4995	0.7614	0.5298	0.5150	0.5132	0.6122	0.5122	0.5128	0.6103
				0.1	0.4834	0.4981	0.6882	0.5397	0.5194	0.5165	0.6445	0.5146	0.5166	0.6411
				0.5	0.4603	0.4865	0.4998	0.5765	0.5436	0.5396	0.6439	0.5324	0.5397	0.6383
				0.9	0.4111	0.2778	0.3198	0.6556	0.6199	0.5838	0.6988	0.5842	0.5889	0.6985
40	5	5	0.91	-0.9	0.4731	0.5016	0.7506	0.4916	0.4989	0.4925	0.3713	0.4988	0.4920	0.3711
				-0.5	0.4867	0.5004	0.4280	0.4859	0.4897	0.4871	0.3807	0.4915	0.4889	0.3821
				-0.1	0.4476	0.4997	0.4562	0.4761	0.4867	0.4817	0.3466	0.4883	0.4866	0.3559
				0.1	0.5270	0.5003	0.7480	0.4656	0.4830	0.4761	0.3207	0.4856	0.4835	0.3309
				0.5	0.5057	0.5186	0.4073	0.4198	0.4495	0.6405	0.6075	0.4645	0.4660	0.6261
				0.9	0.5049	0.5896	0.6000	0.6379	0.5545	0.6404	0.6094	0.4137	0.6271	0.6128
			1.1	-0.9	0.5081	0.4932	0.9399	0.5000	0.4919	0.5261	0.6040	0.4919	0.5250	0.6018
				-0.5	0.4899	0.4990	0.8601	0.5158	0.5111	0.5076	0.6170	0.5101	0.5077	0.6159
				-0.1	0.4894	0.5014	0.7377	0.5276	0.5143	0.5108	0.6504	0.5106	0.5090	0.6436
				0.1	0.4875	0.5018	0.6558	0.5377	0.5181	0.5150	0.6754	0.5125	0.5118	0.6684
				0.5	0.4647	0.4885	0.4473	0.5896	0.5521	0.5467	0.7895	0.5353	0.5359	0.7729
				0.9	0.3119	0.3152	0.3390	0.7853	0.5206	0.6976	0.6901	0.5881	0.6743	0.6859
40	10	0	0.91	-0.9	0.5072	0.4989	0.5981	0.5031	0.4931	0.5548	0.5823	0.4928	0.5600	0.5822
				-0.5	0.4904	0.4999	0.3125	0.4944	0.5088	0.5490	0.5808	0.5096	0.5482	0.5800
				-0.1	0.4868	0.4970	0.6271	0.5015	0.5120	0.5446	0.5808	0.5132	0.5432	0.5789
				0.1	0.4832	0.4943	0.7791	0.5089	0.5161	0.5447	0.5828	0.5168	0.5419	0.5797
				0.5	0.4497	0.4710	0.5655	0.5495	0.5481	0.5692	0.5942	0.5503	0.5688	0.5915
				0.9	0.2624	0.2697	0.2996	0.7493	0.5367	0.7501	0.7000	0.5376	0.7488	0.7000
			1.1	-0.9	0.4921	0.5023	0.5744	0.5824	0.5079	0.4781	0.5551	0.5079	0.4765	0.5551
				-0.5	0.4918	0.5031	0.5370	0.5932	0.5084	0.4870	0.5675	0.5082	0.4859	0.5679
				-0.1	0.4928	0.5041	0.5746	0.6172	0.5122	0.4993	0.5849	0.5072	0.4950	0.5864
				0.1	0.4931	0.5037	0.5332	0.6373	0.5164	0.5090	0.5968	0.5069	0.5005	0.6006
				0.5	0.4825	0.4865	0.5308	0.7122	0.5399	0.5572	0.5362	0.5175	0.5306	0.5516
				0.9	0.4365	0.4521	0.6044	0.6160	0.5792	0.7333	0.6862	0.5635	0.6421	0.6617

Table 12. $MNCP/NCP$ of various 95 percent confidence intervals for δ with $N = 20, 50$ and $\pi_{+1} = 0.5$ under complete data

n	m_1	m_2	δ	ρ	TlCI	TsCI	TwCI	TlogCI	ACI	WCI	JCI	$ACIlog$	$WCIlog$	$JCIlog$
20	0	0	0.91	-0.9	0.5045	0.4871	0.4666	0.4505	0.4793	0.4802	0.5677	0.4795	0.4802	0.5677
				-0.5	0.4411	0.5104	0.6913	0.4137	0.5075	0.5193	0.4629	0.5167	0.5255	0.4636
				-0.1	0.4894	0.5167	0.7042	0.3945	0.4884	0.4560	0.4569	0.4866	0.5137	0.4602
				0.1	0.4823	0.5208	0.7391	0.3743	0.4780	0.4296	0.4467	0.4794	0.5051	0.4512
				0.5	0.4790	0.5721	0.7796	0.2323	0.4437	0.2803	0.3508	0.4247	0.4369	0.3764
				0.9	0.3265	0.6451	0.6758	0.6245	0.5203	0.3038	0.6577	0.6812	0.6807	0.6716
			1.1	-0.9	0.4444	0.4955	0.5370	0.4771	0.4698	0.4699	0.5788	0.4697	0.4699	0.5790
				-0.5	0.4478	0.4931	0.5576	0.4149	0.4918	0.5098	0.4766	0.4927	0.5116	0.4768
				-0.1	0.4728	0.4920	0.5128	0.4066	0.4754	0.4931	0.4649	0.4850	0.5145	0.4673
				0.1	0.4049	0.4934	0.4444	0.3883	0.4685	0.4806	0.4498	0.4793	0.5088	0.4568
				0.5	0.3644	0.5379	0.7032	0.5350	0.5404	0.5873	0.3571	0.4281	0.4466	0.3865
				0.9	0.2477	0.6362	0.7176	0.5881	0.5830	0.6526	0.6807	0.6780	0.6861	0.6973
50	0	0	0.91	-0.9	0.2385	0.4832	0.7599	0.4721	0.5101	0.5098	0.4843	0.5102	0.5098	0.4898
				-0.5	0.4754	0.4914	0.4604	0.4611	0.4899	0.4883	0.5015	0.4900	0.4908	0.5004
				-0.1	0.4260	0.4905	0.6844	0.4458	0.4888	0.4881	0.4945	0.4899	0.4890	0.4965
				0.1	0.3761	0.4902	0.6888	0.4299	0.4852	0.4801	0.4925	0.4848	0.4825	0.4946
				0.5	0.2621	0.5160	0.6055	0.3558	0.4411	0.4374	0.4454	0.4633	0.4578	0.4613
				0.9	0.2621	0.3000	0.6000	0.6120	0.5323	0.6253	0.6235	0.6933	0.6861	0.6492
			1.1	-0.9	0.5370	0.4971	0.6797	0.4748	0.4687	0.4688	0.5004	0.4687	0.4687	0.5004
				-0.5	0.4831	0.4960	0.6378	0.4552	0.4916	0.4941	0.5009	0.4915	0.4921	0.5000
				-0.1	0.3521	0.4944	0.5456	0.4397	0.4869	0.4865	0.4952	0.4903	0.4946	0.4961
				0.1	0.1964	0.4942	0.3101	0.4239	0.4813	0.4860	0.4890	0.4893	0.4929	0.4951
				0.5	0.4242	0.5198	0.4950	0.3436	0.4422	0.4498	0.4358	0.4589	0.4644	0.4605
				0.9	0.3294	0.3403	0.6094	0.6096	0.4383	0.4837	0.5231	0.4261	0.4270	0.5707

Table 13 Various 95% CIs for π_{1+}/π_{+1} based on the Osoba study

Method	Lower limit	Upper limit
TlCI	0.8129	1.0624
TsCI	0.8126	1.0631
TwCI	0.8210	1.0434
TlogCI	0.8328	1.0441
ACI	0.8238	1.0488
WCI	0.8265	1.0410
JCI	0.7904	1.0694
$ACIlog$	0.8235	1.0489
$WCIlog$	0.8262	1.0413
$JCIlog$	0.7909	1.0665

Table 14 Various 95% CIs for π_{1+}/π_{+1} based on the neurological data set

Method	Lower limit	Upper limit
TlCI	0.8621	2.2049
TsCI	0.8504	2.2143
TwCI	0.8217	1.8582
TlogCI	0.8769	2.0475
ACI	0.8835	2.1248
WCI	0.8896	1.9948
JCI	0.9318	2.2490
$ACIlog$	0.8833	2.1348
$WCIlog$	0.8893	2.0018
$JCIlog$	0.9342	2.2625