11-1-2009

Closed Form Confidence Intervals for Small Sample Matched Proportions

James F. Reed III

Christiana Care Hospital System, Newark, Delaware, JaReed@ChristianaCare.org

Part of the Applied Statistics Commons, Social and Behavioral Sciences Commons, and the Statistical Theory Commons

Closed Form Confidence Intervals for Small Sample Matched Proportions

James F. Reed III
Christiana Care Hospital System,
Newark, Delaware

The behavior of the Wald-z, Wald-c, Quesenberry-Hurst, Wald-m and Agresti-Min methods was investigated for matched proportions confidence intervals. It was concluded that given the widespread use of the repeated-measure design, pretest-posttest design, matched-pairs design, and cross-over design, the textbook Wald-z method should be abandoned in favor of the Agresti-Min alternative.

Key words: Matched proportions, Wald-z, Wald-c, Quesenberry-Hurst, Wald-m, Agresti-Min.

Introduction

Matched-pairs data are common in clinical trials. Study designs that use paired data include the repeated-measure design, pretest-posttest, the matched-pairs design, and the cross-over design. When the response variable is dichotomous and when two dichotomous measurements are available, the data may be summarized as shown in Table 1.

Table 1: Paired Data Study Design Responses

	Test II		
Test I	Success	Failure	Total
Success	a	b	a+b
Failure	c	d	c+d
Total	a+c	b+d	n

For binary responses, McNemar's test is the most commonly applied significance test for comparing the two response distributions. For interval estimation of the difference of proportions, textbooks present the Wald large sample interval (Wald-z). Like the one proportion (Figure 1) and the difference between

James F. Reed III, Ph.D., is a Senior Biostatistician. Email him at: JaReed@ChristianaCare.org.

two independent binomial confidence intervals (Figure 2), the Wald-z interval for matched-pair proportions behaves rather poorly (Figure 3a). Two problems are generally encountered. First, the coverage probability cannot be achieved exactly and secondly, in many instances the Wald-z method does not yield sensible intervals.

The purpose of this study was to investigate the coverage probability of alternative methods for computing confidence intervals to the typical textbook Wald-z or Wald-c (continuity correction). Those alternatives include a simple add four method proposed by Agresti and Min (AM) (2005), a method by Quesenberry and Hurst (QH) (1964), and a modified Wald (Wald-m) suggested by May and Johnson (1998).

Methodology
Notation and Computational Formula

Let $\mathbf{y} = (a, b, c, d)^T$ represent the observed frequencies for a sample from a multinomial distribution with underlying probabilities $\boldsymbol{\pi} = (\pi_a, \pi_b, \pi_c, \pi_d)^T$. Let b be the number of subjects who respond favorably on the first occasion but unfavorably on the second and let c be the number who responds unfavorably on the first occasion but favorably on the second. Let a be the number of subjects who respond favorably on both occasions and let d be the number who respond unfavorably on both occasions; then a + d represents the number of concordant pairs and b + d represents the number of discordant pairs.

Figure 1: Coverage Probabilities (n=50) for A Single Proportion Wald Confidence Interval Method
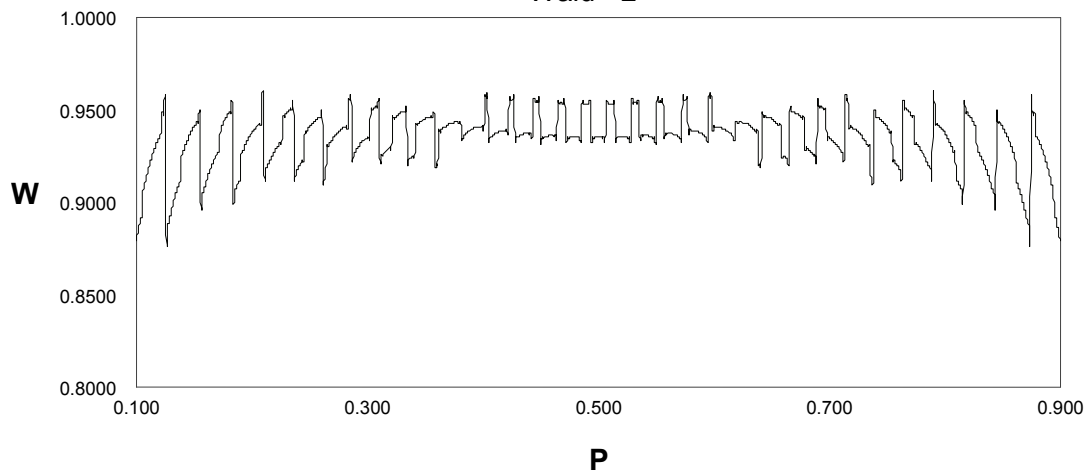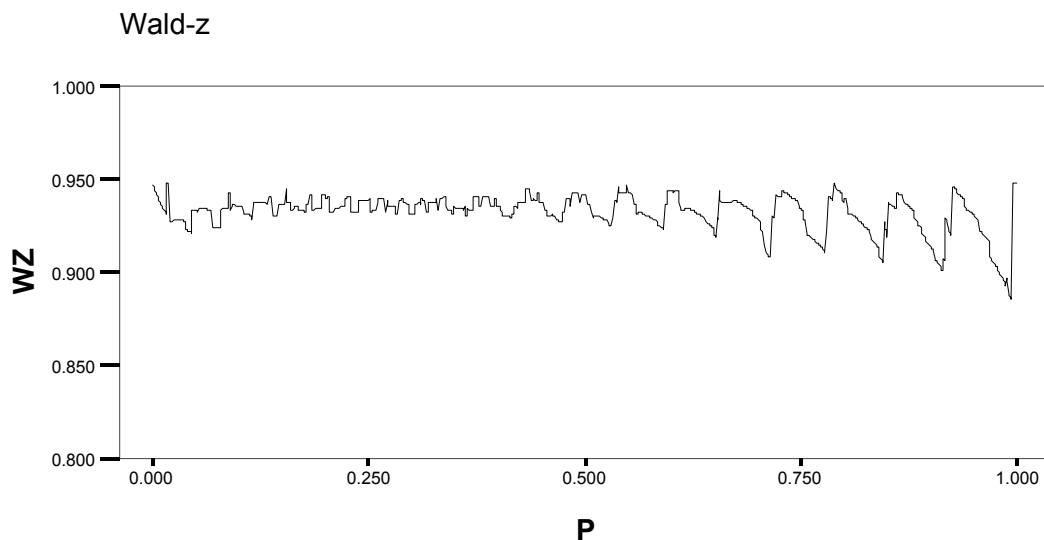
Wald - z

Figure 2: Coverage probabilities for the difference in nominal 95% Wald-z as a function of p_1 when p_2=0.3 with n_1=n_2=20 (Two Independent Proportions)

Wald-z

The confidence interval computational methods based on the data structure given in Table 1 are as follows:

Wald-z:

$$LB = (b-c)/n - z_{\alpha 2}\sqrt{[((b + c)/n - (b-c)^2)/n^2)/n]}$$

and

$$UB = (b-c)/n + z_{\alpha 2}\sqrt{[((b + c)/n - (b-c)^2)/n^2)/n]}.$$

Wald-c:

$$LB = (b-c)/n$$
$$- \left\{ \frac{z_{\alpha 2}\sqrt{[((b + c)/n - (b-c)^2)/n^2)/n]}+1}{n} \right\}$$

and

$$UB = (b-c)/n$$
$$+ \left\{ \frac{z_{\alpha 2}\sqrt{[((b + c)/n - (b-c)^2)/n^2)/n]}+1}{n} \right\}.$$

Wald-m:

$$LB = |p_b - p_c| - \{\chi^2 \left[(p_b + p_c + 1/n) - (p_b - p_c)^2 \right]/n)^{1/2}$$

and

$$UB = |p_b - p_c| + \{\chi^2 \left[(p_b + p_c + 1/n) - (p_b - p_c)^2 \right]/n)^{1/2}$$

with

$$\chi^2 = \chi^2 (\alpha, \ 1).$$

Quesenberry-Hurst (QH):

$$LB = \frac{(n|p_b - p_c|)}{\left(\frac{(\chi^2 + n) - \{\chi^2 \left[(\chi^2 + n)(p_b + p_c) - n(p_b - p_c)^2 \right]\}^{1/2}}{(\chi^2 + n)} \right)}$$

and

$$UB = \frac{(n|p_b - p_c|)}{\left(\frac{(\chi^2 + n) + \{\chi^2 \left[(\chi^2 + n)(p_b + p_c) - n(p_b - p_c)^2 \right]\}^{1/2}}{(\chi^2 + n)} \right)}$$

where

$$\chi^2 = \chi^2 (\alpha, \ 1).$$

Agresti-Min (AM):

$$LB = \left(c^* - b^* \right)/n^*$$

$$- z_{\alpha 2} \sqrt{\left[\left(b^* + c^* \right) - \left(c^* - b^* \right)^2 / n^* \right]/n^*}$$

and

$$UB = \left(c^* - b^* \right)/n^*$$

$$- z_{\alpha 2} \sqrt{\left[\left(b^* + c^* \right) - \left(c^* - b^* \right)^2 / n^* \right]/n^*}$$

with
$$b^* = b+1/2, \ c^* = c+1/2, \ n^* = n + 2.$$

The joint probability mass function of (Y_b, Y_c) is expressed as a function of Δ [$\Delta = (\pi_c - \pi_b) - (\pi_a - \pi_c) = \pi_b - \pi_c$] and is given by: $f(b,c | \Delta, \pi_c) = Pr(Y_b = b, Y_c = c | \Delta, \pi_c) = n!/[b!c!(n-b-c)!] (\pi_c + \Delta)^b \pi_c^c (1 - 2\pi_c - \Delta)^{n-b-c}$. Where, Δ and π_c satisfy the following inequality:

$$\pi_c \in [0, (1-\Delta)/2] \text{ if } 0 < \pi_c < 1,$$

and

$$\pi_c \in [-\Delta, (1-\Delta)/2] \text{ if } -\Delta < \pi_c < 0.$$

Coverage probability (CP) is generally used to evaluate $(1 - \alpha)$ confidence intervals. The coverage probability function CP (Δ) for matched proportions for any Δ is defined as:

$$CP(\Delta) = [\Sigma_k [\Sigma_b \Sigma_c I_T(b,c | \Delta, \pi_c) f(b,c | \Delta, \pi_c)]\},$$

where:

$$I_T(b, c | \Delta, \pi_c) = 1 \text{ if } \Delta \in [\Delta_l, \Delta_u]; 0 \text{ otherwise.}$$

Results

The 95% CP (Δ) for $p_c = 0.1$, n=20 and $p_b = 0.001, ..., 0.999$ for the Wald-z, Wald-c, AM, Wald-m and Quesenberry-Hurst methods are shown in Figure 3.

CP (Δ) probabilities are 0.9125, 0.9545, 0.9401, 0.9435 and 0.0541 respectively. The 95% CP (Δ) for $p_c = 0.25$, n=30 and $p_b = 0.001, ..., 0.999$ for the Wald-z, Wald-c, AM Wald-m and Quesenberry-Hurst methods are shown in Figure 4.

CP (Δ) probabilities are 0.9334, 0.9611, 0.9425, 0.9484 and 0.9448 respectively. And, the CP (Δ) for $p_c = 0.40$, n=40 and $p_b = 0.001, ..., 0.999$ for the Wald-z, Wald-c, AM Wald-m and Quesenberry-Hurst methods are shown in Figure 5. CP (Δ) probabilities are 0.9390, 0.9607, 0.9444, 0.9485 and 0.9451 respectively.

The CP (Δ) plots in figures 3-5 demonstrate that the Wald-z method is suboptimal over the range of p, the Wald-c and Wald-m methods are conservative and the Quesenberry-Hurst and Agresti-Min methods are slightly less than nominal.

Conclusion

A number of closed form methods for constructing confidence intervals for paired binary data were proposed. Newcombe (1998) conducted an empirical study to compare the CP (Δ) of ten confidence interval estimators for the difference between binomial proportions based on paired data. He concluded that the profile likelihood estimator and the score test based confidence interval proposed by Tango (1998) performed well in large-sample situations. May

Figure 3: 95% Coverage Probability for Matched Proportions
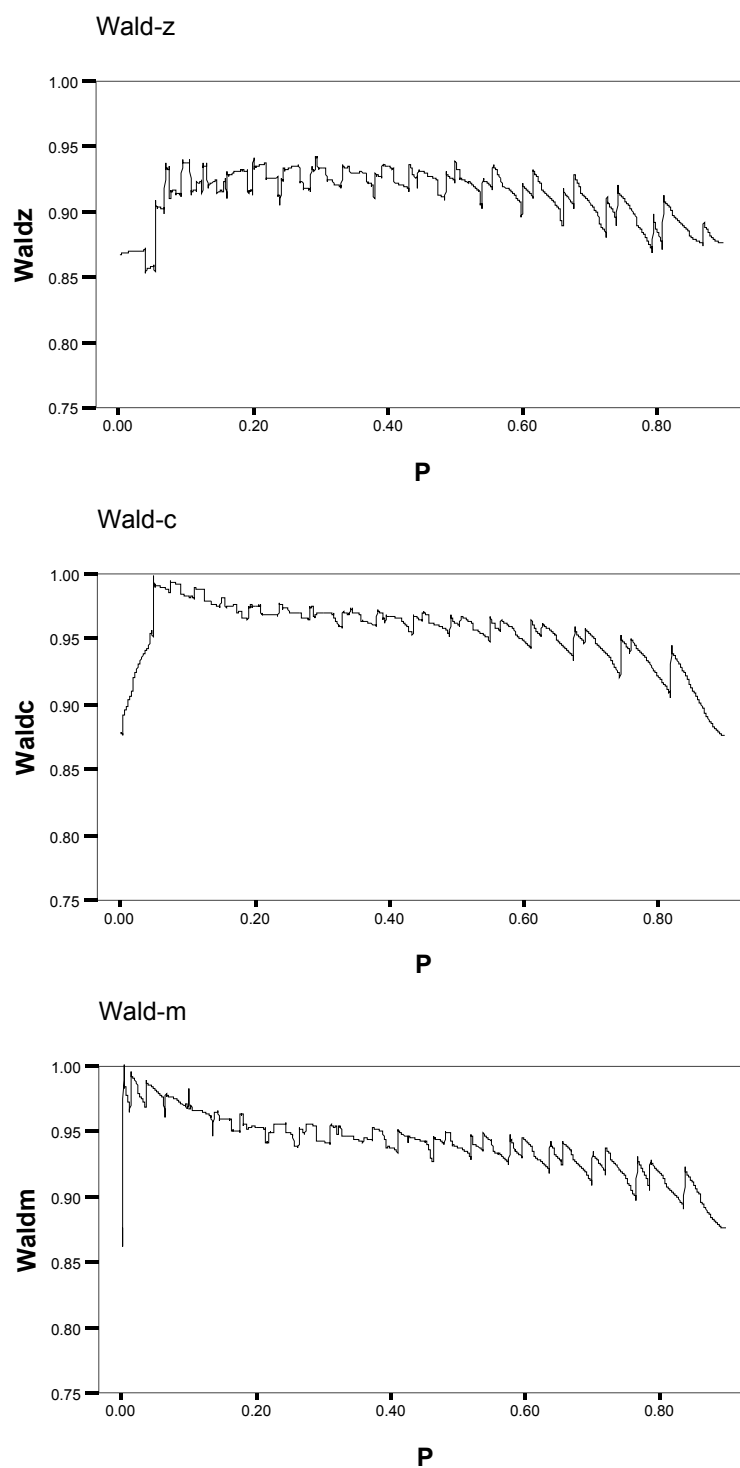$p_c = 0.1$, n=20 $p_b = 0.001, ..., 1-p_c$

Wald-z

Wald-c

Wald-m

Figure 3 (continued): 95% Coverage Probability for Matched Proportions
$p_c=0.1$, n=20 $p_b= 0.001, ..., 1-p_c$

Quesenberry-Hurst

Agresti-Min

Wald-z
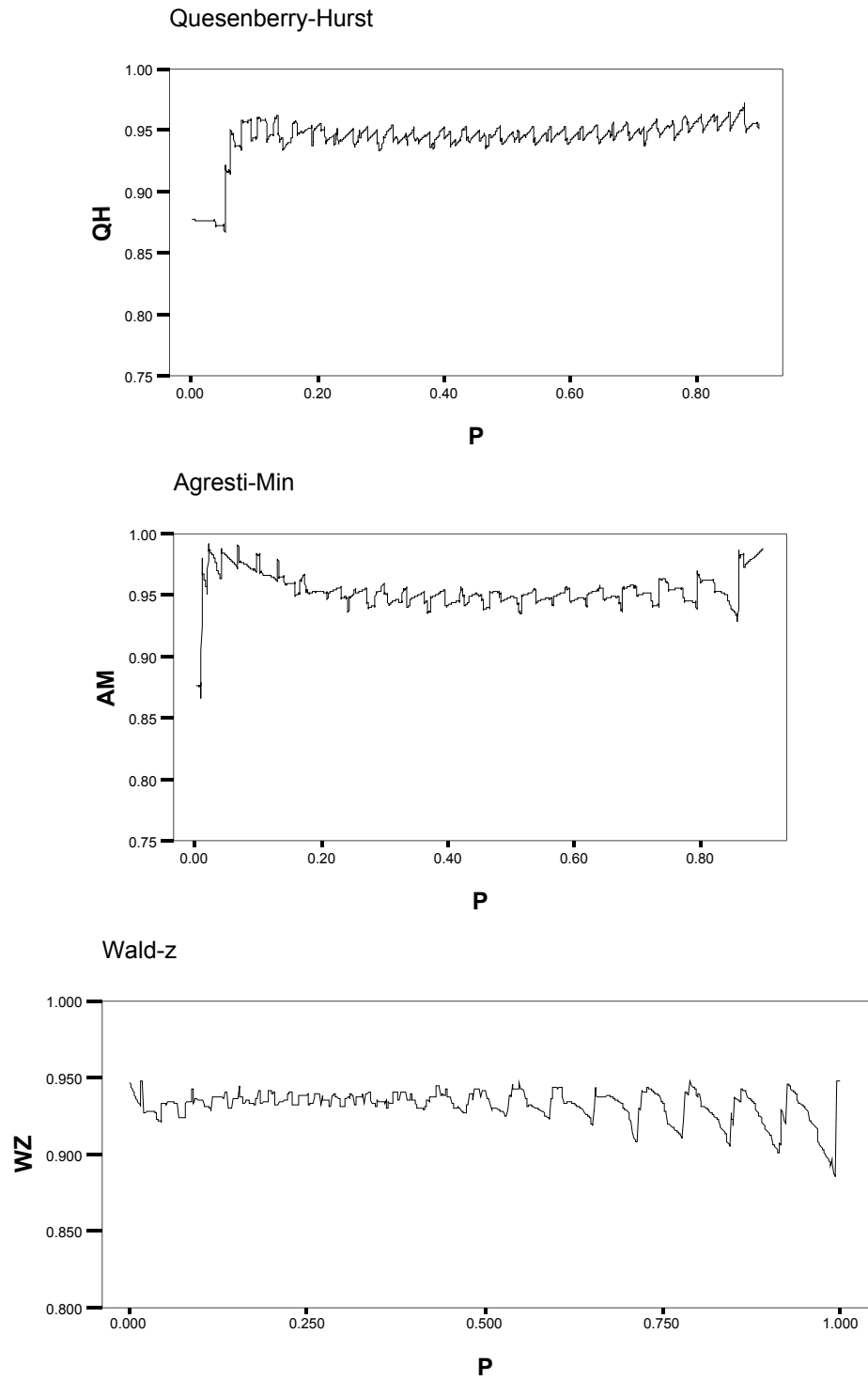
Figure 4: 95% Coverage Probability for Matched Proportions
$p_c=0.25$, $n=30$, $p_b= 0.001, \ldots, 1-p_c$

Wald-z

Wald-c

Wald-m
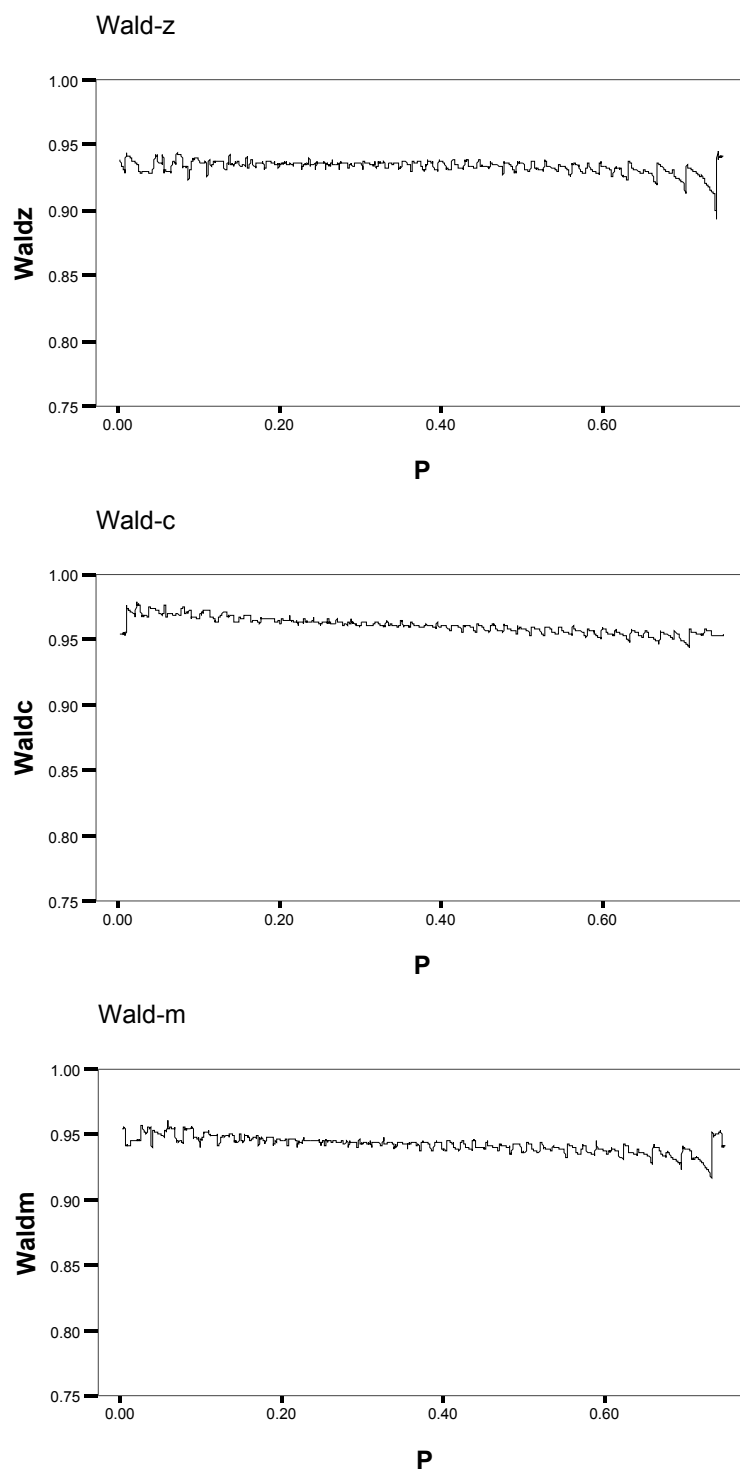
Figure 4 (continued): 95% Coverage Probability for Matched Proportions
p_c=0.25, n=30, p_b= 0.001, …, 1−p_c

Quesenberry-Hurst

Agresti-Min
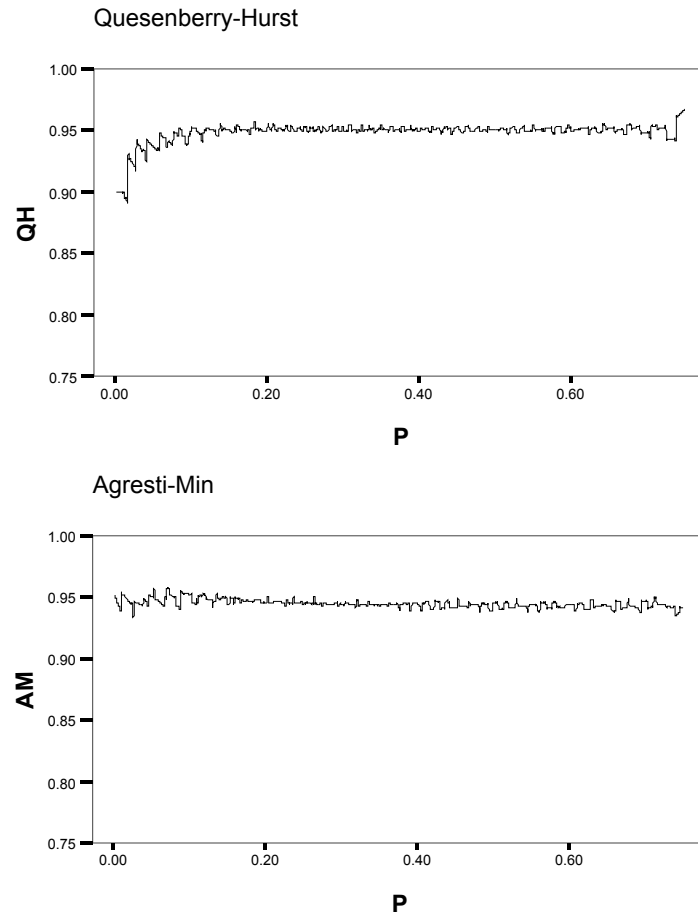
Figure 5: 95% Coverage Probability for Matched Proportions
p_c= 0.4, n=40 and p_b= 0.001, ..., 1 − p_c
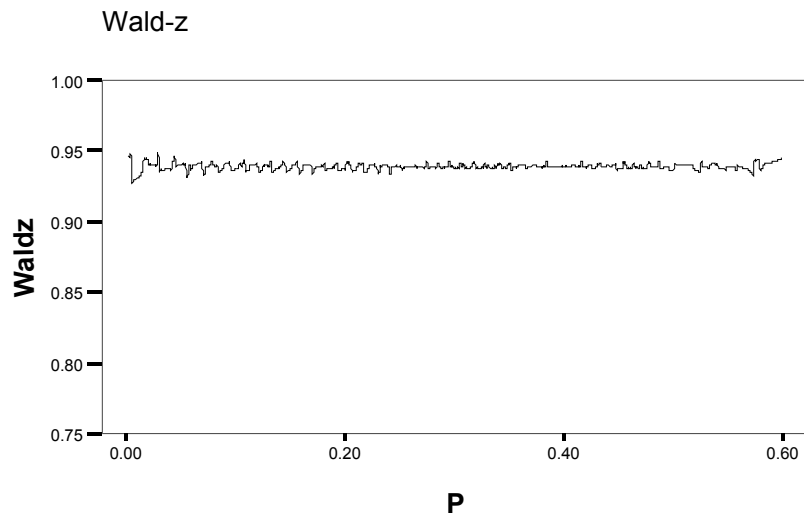
Wald-z

Figure 5 (continued): 95% Coverage Probability for Matched Proportions
$p_c= 0.4$, n=40 and $p_b= 0.001, ..., 1 - p_c$
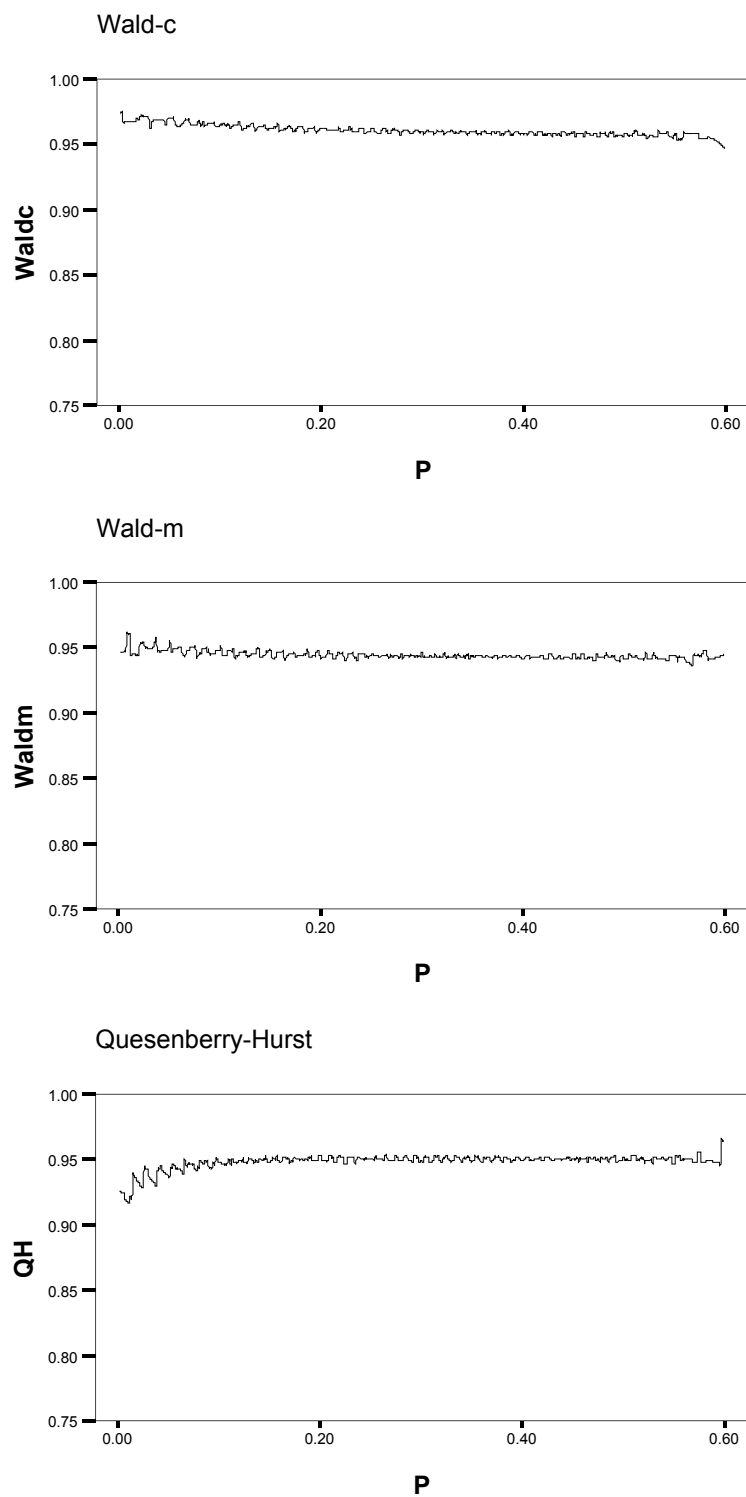
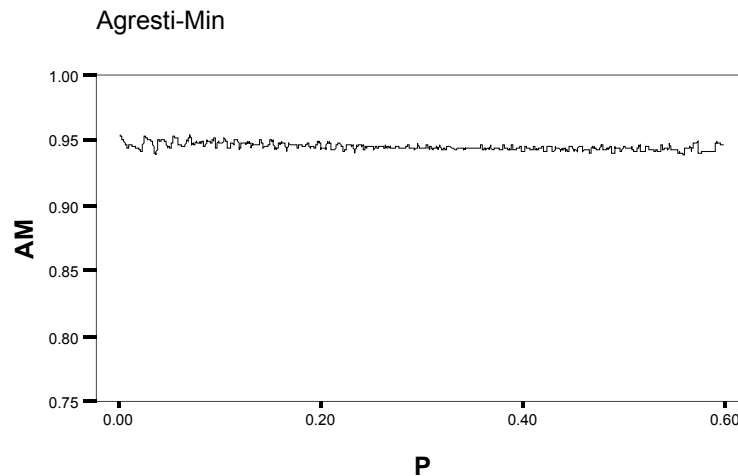Wald-c

Wald-m

Quesenberry-Hurst

Figure 5 (continued): 95% Coverage Probability for Matched Proportions
$p_c= 0.4$, n=40 and $p_b= 0.001, ..., 1 - p_c$

Agresti-Min

and Johnson (1998), Quesenberry and Hurst (1964) and Agresti and Min (2005) proposed closed form computationally friendly alternatives.

This article focused on constructing confidence intervals using closed form methods for paired data under small-sample designs. In this setting, based on the results, either the Quesenberry-Hurst or Agresti-Min methods are recommended. Given the widespread use of the repeated-measure, pretest-posttest, the matched-pairs, and the cross-over designs, the textbook Wald-z method should be abandoned in favor of either the closed form of Quesenberry-Hurst or Agresti-Min.

References

Agresti, A. & Min, Y. (2005). Simple improved confidence intervals for comparing matched proportions. *Statistics in Medicine, 24,* 729-740.

Newcombe, R. G. (1998). Improved confidence intervals for the difference between binomial proportions based on paired data. *Statistics in Medicine, 17,* 2635-2650.

May, W. L., & Johnson, W. D. (1998). Confidence intervals for differences in correlated binary proportions. *Statistics in Medicine, 16,* 2127-2136.

Quesenberry, C. P., & Hurst, D. C. (1964). Large sample simultaneous confidence intervals for multinomial proportions. *Technometrics, 6,* 191-195.

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