Journal of Modern Applied Statistical Methods

Volume 5 | Issue 2 Article 14

11-1-2005

Interval Estimation of Risk Difference in Simple Compliance Randomized Trials

Kung-Jong Lui
San Diego State University, kjl@rohan.sdsu.edu

Journal of Modern Applied Statistical Methods
November, 2006, Vol. 5, No. 2, 395-407

1538 – 9472/06/$95.00

Interval Estimation of Risk Difference in Simple Compliance Randomized Trials

Kung-Jong Lui
San Diego State University

Consider the simple compliance randomized trial, in which patients randomly assigned to the experimental treatment may switch to receive the standard treatment, while patients randomly assigned to the standard treatment are all assumed to receive their assigned treatment. Six asymptotic interval estimators for the risk difference in probabilities of response among patients who would accept the experimental treatment were developed. Monte Carlo methods were employed to evaluate and compare the finite-sample performance of these estimators. An example studying the effect of vitamin A supplementation on reducing mortality in preschool children was included to illustrate their practical use.

Key words: interval estimation; coverage probability; average length; efficiency; simple compliance trial.

Introduction

In randomized clinical trials, because the characteristics of the experimental treatment effect are often not completely known, some patients randomly assigned to this treatment may not comply with their assigned treatment. For example, consider the study of vitamin A supplementation to reduce mortality among preschool children in rural Indonesia (Sommer & Zeger, 1991, Sommer, Tarwotjo, Djunaedi, et al., 1986). Children who resided in 225 randomly selected villages out of 450 villages were assigned to receive a large oral dose of vitamin A two to three months following baseline enumeration and again six months later. Children in the remaining 225 villages would receive no vitamin A supplementation. The number of deaths in both comparison groups was ascertained in a second population census 12 months following the baseline census. Nearly 20 percent of children assigned to the experimental group failed to receive vitamin A supplementation. The mortality rates between the groups were then compared. Because the data structure for this simple compliance study is essentially the same as that of the single consent randomized design (Zelen, 1979, 1990, Anbar, 1983, Brunner & Neumann, 1985, Ellenberg, 1984, Bernhard & Compagnone, 1989, McHugh, 1984, Matts & McHugh, 1987, 1993, Lui & Lin, 2003), all statistical methods developed here for the simple compliance trial are applicable to the latter (Sommer & Zeger, 1991, Zelen, 1986) as well.

In this article, six asymptotic interval estimators of the risk difference between two treatments among patients who would accept the experimental treatment were developed. These included the interval estimator using Wald's statistic (Casella & Berger, 1990, Sommer & Zeger, 1991), the interval estimator using $\tanh^{-1}(x)$ transformation (Edwardes, 1995, Lui, 2002) of the maximum likelihood estimator (MLE), the interval estimator derived from a quadratic equation based on the MLE and its asymptotic properties, the interval estimator using an idea similar to that of Fieller's Theorem (Casella & Berger, 1990), and the interval estimator using a randomization-based approach with and without the continuity correction (Mark & Robins, 1993, Sato, 2000). To evaluate and compare the finite sample performance of these

Kung-Jong Lui is Professor of Mathematics and Statistics at San Diego State University. His current research interests are in statistical methods in epidemiology and randomized clinical trials, categorical data analysis, and sample surveys. He is a fellow of the American Statistical Association, and is serving as an Associate Editor for *Biometrical Journal* as well as an Editorial Advisor for *Journal of Probability and Statistical Science*.

estimators, Monte Carlo simulation was employed. Finally, an example studying the effect of vitamin A supplementation on reducing mortality in preschool children (Sommer, Tarwotjo, Djunaedi, et al., 1986) was included to illustrate their practical use.

Methods and Notations

Consider comparing an experimental treatment with a standard treatment in a simple compliance randomized trial. Patients randomly assigned to the experimental treatment are allowed to switch to receive the standard treatment, while patients randomly assigned to the standard treatment are assumed to all receive their assigned treatment. For clarity, the probabilities p_{ij} of response for the experimental treatment are summarized in the following table. For example, the parameter p_{11} denotes probability of obtaining a patient who would accept the experimental treatment and have a positive response. Define $p_{i+} = p_{i1} + p_{i0}$ and $p_{+j} = p_{1j} + p_{0j}$ for $i = 1, 0$ and $j = 1, 0$ (see Table 1).

Similarly, the parameter p^*_{ij} ($i = 1, 0$ and $j = 1, 0$) denotes the corresponding cell probability of response in the following table for the standard treatment. Define $p^*_{i+} = p^*_{i1} + p^*_{i0}$ and $p^*_{+j} = p^*_{1j} + p^*_{0j}$ for $i = 1, 0$ and $j = 1, 0$ (see Table 2).

Because a patient assigned to the experimental treatment will receive the standard treatment if he/she declines to receive the experimental treatment and because patients are randomly assigned to either treatment, the equality $p_{i0} = p^*_{i0}$ (for $i = 1, 0$) can be reasonably assumed to hold (Sommer & Zeger, 1991). These imply that the proportions of patients who would consent to accept the experimental treatment between the two treatment groups are equal (i.e., $p_{+1} = p^*_{+1}$).

Suppose that there are n and m patients independently randomly assigned to receive the experimental and standard treatments, respectively. Let n_{ij} (and m_{ij}) denote the observed frequencies corresponding to the cell probabilities p_{ij} (and p^*_{ij}) in the experimental (and the standard) treatment. Then the random vector $n = (n_{11}, n_{10}, n_{01}, n_{00})'$ follows the multinomial distribution with parameters n and $(p_{11}, p_{10}, p_{01}, p_{00})'$. Because patients assigned to the standard treatment were not asked whether they would accept the experimental treatment, only the marginal total number of responses $m_{1+} (= m_{11} + m_{10})$, following the binomial distribution with parameters m and $p^*_{1+} (= p^*_{11} + p^*_{10})$, was observed.

In this article, searching for a good interval estimator for the risk difference $\Delta (= (p_{11} / p_{+1}) - (p^*_{11} / p^*_{+1}) = (p_{1+} - p^*_{1+}) / p_{+1})$ between two treatments among patients who would accept the experimental treatment is the main focus of interest here. Note that the range for Δ is, by definition, $-1 < \Delta < 1$.

Table 1. Experimental Treatment: Consent to Receive the Experimental Treatment

Response	Yes	No	Total
Positive	p_{11}	p_{10}	p_{1+}
Negative	p_{01}	p_{00}	p_{0+}
Total	p_{+1}	p_{+0}	1

Table 2. Standard Treatment: Consent to Receive the Experimental Treatment

Response	Yes	No	Total
Positive	p^*_{11}	p^*_{10}	p^*_{1+}
Negative	p^*_{01}	p^*_{00}	p^*_{0+}
Total	p^*_{+1}	p^*_{+0}	1

Based on the intent-to-treat analysis (Zelen, 1979), patients are compared according to the treatments to which they are originally designated despite whether patients comply with their regimen. The maximum likelihood estimator (MLE) for Δ is given by

$$\hat{\Delta} = (\hat{p}_{1+} - \hat{p}^*_{1+}) / \hat{p}_{+1} \qquad (1)$$

where $\hat{p}_{1+} = \hat{p}_{11} + \hat{p}_{10}$, $\hat{p}_{+1} = \hat{p}_{11} + \hat{p}_{01}$, $\hat{p}_{ij} = n_{ij} / n$, and $\hat{p}^*_{1+} = m_{1+} / m$. Furthermore, on the basis of the delta method (Appendix), an estimated asymptotic variance of $\hat{\Delta}$ is obtained as

$$\hat{Var}(\hat{\Delta}) = [\hat{p}_{1+}(\hat{p}_{10} + \hat{p}_{01}) - \hat{p}^*_{1+}(2\hat{p}_{10} - \hat{p}^*_{1+}\hat{p}_{+0})]/(n\hat{p}^3_{+1}) + \hat{p}^*_{1+}(1 - \hat{p}^*_{1+}) / (m\hat{p}^2_{+1}) \qquad (2)$$

Based on (1) and (2), an asymptotic 100(1-α) percent confidence interval using Wald's statistic for Δ is given by

$$[\max\{\hat{\Delta} - Z_{\alpha/2}(\hat{Var}(\hat{\Delta}))^{1/2}, -1\}, \min\{\hat{\Delta} + Z_{\alpha/2}(\hat{Var}(\hat{\Delta}))^{1/2}, 1\}] \qquad (3)$$

where $Z_{\alpha/2}$ is the upper $100(\alpha/2)th$ percentile of the standard normal distribution.

Because the sampling distribution of $\hat{\Delta}$ can be skewed, interval estimator (3) may not perform well when the number of patients assigned to either treatment is not large. On the basis of some empirical results that the transformation $\tanh^{-1}(x)(=\frac{1}{2}\log((1+x)/(1-x)))$ has been successfully applied to improve the performance of statistics relevant to the difference in proportions under other situations (Edwardes, 1995, Lui, 2002), this transformation is considered in this article as well. Based on the delta method again, an estimated asymptotic variance can be shown to be given by $\hat{Var}(\tanh^{-1}(\hat{\Delta})) = \hat{Var}(\hat{\Delta}) / (1 - \hat{\Delta}^2)^2$. This leads to produce an asymptotic 100(1-α) percent confidence interval for Δ to be

$$[\tanh(\tanh^{-1}(\hat{\Delta}) - Z_{\alpha/2}(\hat{Var}(\tanh^{-1}(\hat{\Delta})))^{1/2}), \tanh(\tanh^{-1}(\hat{\Delta}) + Z_{\alpha/2}(\hat{Var}(\tanh^{-1}(\hat{\Delta})))^{1/2})] \qquad (4)$$

Note that when both *n* and *m* are large, the probability $P((\hat{\Delta} - \Delta)^2 / Var(\hat{\Delta}) \le Z^2_{\alpha/2}) \approx 1 - \alpha$, where $Var(\hat{\Delta})$ is, as shown in Appendix, given by

$$\Delta[(p_{1+} - p^*_{1+})(1 - p_{+1}) - 2(p_{11} - p_{1+}p_{+1})] / (np^2_{+1}) + p_{1+}(1 - p_{1+}) / (np^2_{+1}) + p^*_{1+}(1 - p^*_{1+}) / (mp^2_{+1}).$$

This suggests the following quadratic inequality:

$$\Delta^2 - 2B\Delta + C \le 0, \qquad (5)$$

where

$$A = 1$$

$$\mathrm{B} = \hat{\Delta} + Z_{\alpha/2}^2[(\hat{p}_{1+} - \hat{p}_{1+}^*)(1 - \hat{p}_{+1}) - 2(\hat{p}_{11} - \hat{p}_{1+}\hat{p}_{+1})]/(2n\hat{p}_{+1}^2)$$

$$\mathrm{C} = \hat{\Delta}^2 - Z_{\alpha/2}^2[\hat{p}_{1+}(1-\hat{p}_{1+})/(n\hat{p}_{+1}^2) + \hat{p}_{1+}^*(1-\hat{p}_{1+}^*)/(m\hat{p}_{+1}^2)].$$

Thus, if $B^2 - AC > 0$, then an asymptotic 100(1-α) percent confidence interval for Δ would be given by

$$[\max\{(B - \sqrt{B^2 - AC})/A, -1\}, \min\{(B + \sqrt{B^2 - AC})/A, 1\}]. \qquad (6)$$

To alleviate the concern that the sampling distribution of the MLE $\hat{\Delta}(= (\hat{p}_{1+} - \hat{p}_{1+}^*)/\hat{p}_{+1})$ can be, as noted previously, skewed when n or m is small, the idea of Fieller's Theorem (Casella & Berger, 1990) may be considered. Define $Z^* = (\hat{p}_{1+} - \hat{p}_{1+}^*) - \Delta\hat{p}_{+1}$. First, note that the expectation $E(Z^*)$ is equal to 0. The variance of Z^* is

$$Var(Z^*) = p_{1+}(1-p_{1+})/n + p_{1+}^*(1-p_{1+}^*)/m + \Delta^2 p_{+1}(1-p_{+1})/n - 2\Delta(p_{11} - p_{1+}p_{+1})/n \qquad (7)$$

Thus, when both n and m are large, the probability $\mathrm{P}((Z^*)^2/Var(Z^*) \le Z_{\alpha/2}^2) \approx 1-\alpha$. This leads to considering the following quadratic equation:

$$A^*\Delta^2 - 2B^*\Delta + C^* \le 0 \qquad (8)$$

where

$$A^* = \hat{p}_{+1}^2 - Z_{\alpha/2}^2\hat{p}_{+1}(1-\hat{p}_{+1})/n$$

$$B^* = (\hat{p}_{1+} - \hat{p}_{1+}^*)\hat{p}_{+1} - Z_{\alpha/2}^2(\hat{p}_{11} - \hat{p}_{1+}\hat{p}_{+1})/n$$

$$C^* = (\hat{p}_{1+} - \hat{p}_{1+}^*)^2 - Z_{\alpha/2}^2\left(\frac{\hat{p}_{1+}(1-\hat{p}_{1+})}{n} + \frac{\hat{p}_{1+}^*(1-\hat{p}_{1+}^*)}{m}\right).$$

If $(B^*)^2 - A^*C^* > 0$, then an asymptotic 100(1-α) percent confidence interval for Δ would be given by

$$[\max\{(B^* - \sqrt{(B^*)^2 - A^*C^*})/A^*, -1\}, \min\{(B^* + \sqrt{(B^*)^2 - A^*C^*})/A^*, 1\}]. \qquad (9)$$

Following Mark and Robins (1993), Sato (1995, 2000) discussed sample size calculation using a randomization-based approach in binary outcome data. Based on the same arguments as those given by (Sato, 2000, pp. 2691-2692), an asymptotic 100(1-α) percent confidence interval for Δ is obtained as

$$[\max\{(B_-^{**} - \sqrt{(B_-^{**})^2 - A^{**}C_-^{**}})/A^{**}, -1\}, \min\{(B_+^{**} + \sqrt{(B_+^{**})^2 - A^{**}C_+^{**}})/A^{**}, 1\}], \qquad (10)$$

where

$$A^{**} = n_{+1}^2[m^2 + Z_{\alpha/2}^2(nm)/N]$$

$$B_\pm^{**} = mn_{+1}(mn_{1+} - nm_{1+} \pm c) - Z_{\alpha/2}^2 nmn_{+1}(N - 2(n_{1+} + m_{1+}))/(2N)$$

$$C_\pm^{**} = (mn_{1+} - nm_{1+} \pm c)^2 - Z_{\alpha/2}^2 nm(n_{1+} + m_{1+})(N - (n_{1+} + m_{1+}))/\mathrm{N}$$

$N = n + m$, $n_{1+} = n_{11} + n_{10}$, $n_{+1} = n_{11} + n_{01}$, and $c = N/2$ when one wishes to employ the continuity correction; and c = 0, otherwise.

Monte Carlo Simulation

To evaluate and compare the finite sample performance of interval estimators (3), (4), (6), (9), and (10), Monte Carlo simulation is employed. Given the population proportion of a randomly selected patient who would consent to accept the experimental treatment $p_{+1} (= p^*_{+1})$ and the probability of positive response $p^*_{r|c} (= p^*_{11} / p^*_{+1})$ among patients who would consent to accept the experimental treatment (if he/she had been assigned to the experimental treatment) in the standard treatment, the cell probability $p^*_{11} = p^*_{r|c} p^*_{+1} = p^*_{r|c} p_{+1}$ can then be uniquely determined. For given a value Δ, the cell probability: $p_{11} = p^*_{11} + \Delta p_{+1}$.

Similarly, the cell probability $p^*_{10} (= p_{10})$ can be determined by $p^*_{\bar{r}|c} (1 - p^*_{+1})$ when the probability $p^*_{r|\bar{c}}$ of positive response among patients who would decline to receive the experimental treatment in the standard treatment is given. In the simulation, $p^*_{r|\bar{c}}$ is arbitrarily set equal to $p^*_{r|c}/3$. For simplicity, the case of equal sample allocation (i.e., $n = m$) is focused here.

This article considers the situations, in which the probability of a randomly selected patient who would consent to accept the experimental treatment p_{+1} = 0.30, 0.50, 0.80; the underlying difference between two treatments among patients who would consent to accept the experimental treatment Δ = 0.0, 0.10, 0.20; the conditional probability of positive responses among patients who would consent to accept the experimental treatment in the standard treatment group $p^*_{r|c}$=0.20, 0.50; and the number of patients assigned to either treatment $n(= m)$ = 30, 50, 100.

For each configuration determined by a combination of these parameters, SAS (1990) is applied to generate 10000 repeated samples of observations following the desired multinomial distributions to estimate the coverage probability and the average length. Note that if a sample led to $\hat{p}_{+1}$ = 0 or an estimate $\hat{\Delta}$ was out of the range $-1 < \Delta < 1$, any interval estimator discussed here would be inapplicable.

Furthermore, if the two distinct real roots of a quadratic equation did not exist, the corresponding interval estimator could not be employed either. The estimated coverage probability and average length are calculated over those samples for which the corresponding interval estimator exists. For completeness, the probability of failing to produce a confidence interval for each interval estimator in all the above situations is also calculated.

Results

Table 3 summarizes the estimated coverage probability and average length of 95% confidence interval using interval estimators (3), (4), (6), (9), as well as the interval estimator (10) with and without the continuity correction. Note that the coverage probability of interval estimator (10) using the randomization approach without the continuity correction can be frequently less than 95% by more than 1% when the underlying difference Δ is not equal to 0. By contrast, the interval estimator (10) with the continuity correction tends to be conservative and hence lose efficiency with respect to the average length as compared with the other estimators considered here (Table 3). Note further that interval estimators (3) and (6) are essentially equivalent and both are preferable to interval estimator (9) with respect to efficiency.

Finally, note that interval estimator (4) is consistently the most efficient with respect to the average length among all interval estimators with the coverage probability larger than or equal to the desired 95% confidence level considered here. In fact, interval estimator (4) is the only one estimator that has the estimated coverage probability larger than the 95% confidence level in all the situations discussed in Table 3.

Table 3. The estimated coverage probability and average length (in parenthesis) of 95% confidence interval using interval estimators (3), (4), (6), (9), as well as the interval estimator (10) with and without continuity correction. Each entry is calculated on the basis of 10000 repeated samples.

p_{+1}	Δ	$p^*_{r\|c}$	n	(3)	(4)	(6)	(9)	(10)	
0.30	0.00	0.20	30	0.965	0.992✓	0.949	0.946	0.995[a]	0.950[b]
				(1.060)	(1.004)	(1.044)	(1.166)	(1.204)	(1.016)
			50	0.956	0.984✓	0.937	0.937	0.980	0.944
				(0.831)	(0.792)	(0.829)	(0.915)	(0.941)	(0.815)
			100	0.956	0.970✓	0.950	0.950	0.976	0.951
				(0.582)	(0.565)	(0.583)	(0.613)	(0.646)	(0.580)
		0.50	30	0.972	0.998✓	0.964	0.962	0.988	0.972
				(1.395)	(1.328)	(1.352)	(1.453)	(1.483)	(1.341)
			50	0.965	0.995✓	0.950	0.950	0.975	0.956
				(1.161)	(1.096)	(1.145)	(1.231)	(1.237)	(1.125)
			100	0.954	0.983✓	0.946	0.946	0.966	0.949
				(0.835)	(0.795)	(0.836)	(0.877)	(0.886)	(0.821)
	0.10	0.20	30	0.965	0.994✓	0.949	0.948	0.987	0.944
				(1.093)	(1.034)	(1.084)	(1.210)	(1.229)	(1.036)
			50	0.958	0.987✓	0.945	0.946	0.976	0.938✗
				(0.868)	(0.825)	(0.868)	(0.959)	(0.956)	(0.828)
			100	0.953	0.968✓	0.947	0.946	0.964	0.939✗
				(0.605)	(0.587)	(0.606)	(0.638)	(0.647)	(0.581)
		0.50	30	0.969	0.999✓	0.960	0.958	0.987	0.967
				(1.391)	(1.325)	(1.362)	(1.461)	(1.499)	(1.354)
			50	0.961	0.998✓	0.949	0.950	0.975	0.955
				(1.158)	(1.091)	(1.148)	(1.236)	(1.244)	(1.130)
			100	0.951	0.986✓	0.943	0.943	0.964	0.945
				(0.837)	(0.795)	(0.838)	(0.879)	(0.891)	(0.825)
	0.20	0.20	30	0.960	0.994✓	0.944	0.944	0.980	0.932✗
				(1.107)	(1.061)	(1.106)	(1.227)	(1.247)	(1.047)
			50	0.956	0.989✓	0.944	0.946	0.970	0.931✗
				(0.882)	(0.841)	(0.886)	(0.975)	(0.961)	(0.831)
			100	0.952	0.972✓	0.946	0.946	0.960	0.933✗
				(0.617)	(0.599)	(0.619)	(0.650)	(0.647)	(0.580)
		0.50	30	0.973	1.000✓	0.964	0.962	0.991	0.975
				(1.372)	(1.328)	(1.362)	(1.440)	(1.504)	(1.361)
			50	0.959	0.998✓	0.946	0.946	0.978	0.955
				(1.142)	(1.088)	(1.143)	(1.229)	(1.248)	(1.134)
			100	0.958	0.991✓	0.952	0.952	0.972	0.957
				(0.824)	(0.785)	(0.828)	(0.868)	(0.889)	(0.823)

[a] :This column is calculated with the continuity correction.

[b]: This column is calculated without the continuity correction.

✗: means that the estimated coverage probability is less than the desired 95% confidence level by more than 1%.

✓: indicates that the interval estimator has the shortest average length among interval estimators subject to its estimated coverage probability less than the desired 95% confidence level by no more than 1%.

Table 3 (Continued)

p_{+1}	Δ	$p^*_{r\|c}$	n	(3)	(4)	(6)	(9)	(10)	
0.50	0.00	0.20	30	0.949	0.971✓	0.941	0.941	0.987	0.944
				(0.698)	(0.670)	(0.699)	(0.755)	(0.819)	(0.688)
			50	0.955	0.967✓	0.947	0.947	0.979	0.949
				(0.538)	(0.524)	(0.538)	(0.563)	(0.612)	(0.533)
			100	0.952	0.957✓	0.950	0.950	0.973	0.950
				(0.379)	(0.374)	(0.379)	(0.387)	(0.417)	(0.377)
		0.50	30	0.950	0.985✓	0.945	0.945	0.976	0.953
				(0.973)	(0.919)	(0.968)	(1.033)	(1.057)	(0.935)
			50	0.951	0.976✓	0.947	0.947	0.971	0.952
				(0.755)	(0.723)	(0.756)	(0.790)	(0.814)	(0.736)
			100	0.945	0.957✓	0.940	0.940	0.961	0.942
				(0.528)	(0.516)	(0.529)	(0.540)	(0.561)	(0.521)
	0.10	0.20	30	0.948	0.976✓	0.943	0.943	0.979	0.934✗
				(0.734)	(0.703)	(0.735)	(0.793)	(0.821)	(0.689)
			50	0.947	0.965✓	0.944	0.944	0.969	0.933✗
				(0.564)	(0.549)	(0.565)	(0.590)	(0.610)	(0.532)
			100	0.951	0.957✓	0.948	0.948	0.961	0.937✗
				(0.399)	(0.393)	(0.399)	(0.407)	(0.417)	(0.378)
		0.50	30	0.951	0.990✓	0.944	0.944	0.978	0.953
				(0.968)	(0.912)	(0.967)	(1.033)	(1.060)	(0.937)
			50	0.949	0.977✓	0.944	0.944	0.971	0.952
				(0.746)	(0.714)	(0.748)	(0.781)	(0.810)	(0.732)
			100	0.951	0.963✓	0.948	0.948	0.967	0.952
				(0.525)	(0.513)	(0.526)	(0.536)	(0.561)	(0.521)
	0.20	0.20	30	0.943	0.975✓	0.938	0.939	0.969	0.919✗
				(0.753)	(0.723)	(0.757)	(0.815)	(0.824)	(0.692)
			50	0.951	0.966✓	0.946	0.947	0.963	0.928✗
				(0.581)	(0.565)	(0.582)	(0.608)	(0.612)	(0.534)
			100	0.948	0.956✓	0.947	0.946	0.953	0.927✗
				(0.411)	(0.405)	(0.411)	(0.419)	(0.418)	(0.378)
		0.50	30	0.950	0.992✓	0.942	0.942	0.978	0.955
				(0.953)	(0.901)	(0.957)	(1.022)	(1.062)	(0.938)
			50	0.949	0.977✓	0.943	0.942	0.971	0.952
				(0.735)	(0.706)	(0.738)	(0.770)	(0.810)	(0.732)
			100	0.951	0.962✓	0.947	0.948	0.968	0.955
				(0.516)	(0.505)	(0.517)	(0.527)	(0.560)	(0.521)

Table 3 (continued)

p_{+1}	Δ	$p^*_{r\|c}$	n	(3)	(4)	(6)	(9)	(10)	
0.80	0.00	0.20	30	0.945	0.953✓	0.941	0.941	0.982	0.951
				(0.473)	(0.464)	(0.473)	(0.482)	(0.546)	(0.467)
			50	0.945	0.950✓	0.942	0.942	0.974	0.945
				(0.368)	(0.364)	(0.368)	(0.372)	(0.414)	(0.365)
			100	0.949	0.952✓	0.950	0.950	0.968	0.950
				(0.261)	(0.259)	(0.261)	(0.262)	(0.284)	(0.260)
		0.50	30	0.948	0.964	0.951	0.951	0.977	0.952✓
				(0.624)	(0.604)	(0.624)	(0.636)	(0.682)	(0.603)
			50	0.946	0.953	0.944	0.944	0.966	0.944✓
				(0.484)	(0.475)	(0.484)	(0.489)	(0.522)	(0.474)
			100	0.947	0.952	0.945	0.945	0.959	0.945✓
				(0.343)	(0.340)	(0.343)	(0.345)	(0.364)	(0.339)
	0.10	0.20	30	0.944	0.953✓	0.941	0.942	0.969	0.930✗
				(0.505)	(0.494)	(0.505)	(0.514)	(0.545)	(0.466)
			50	0.950	0.956✓	0.950	0.950	0.965	0.934✗
				(0.393)	(0.388)	(0.393)	(0.397)	(0.414)	(0.365)
			100	0.947	0.950✓	0.946	0.946	0.956	0.932✗
				(0.279)	(0.277)	(0.279)	(0.280)	(0.284)	(0.260)
		0.50	30	0.943	0.960✓	0.937	0.939	0.973	0.950
				(0.619)	(0.600)	(0.620)	(0.631)	(0.682)	(0.603)
			50	0.949	0.958✓	0.947	0.947	0.972	0.952
				(0.481)	(0.472)	(0.481)	(0.486)	(0.523)	(0.475)
			100	0.945	0.950✓	0.944	0.944	0.963	0.948
				(0.340)	(0.337)	(0.340)	(0.342)	(0.364)	(0.339)
	0.20	0.20	30	0.944	0.956✓	0.943	0.943	0.961	0.918✗
				(0.524)	(0.513)	(0.524)	(0.534)	(0.546)	(0.467)
			50	0.943	0.951✓	0.943	0.943	0.951	0.917✗
				(0.408)	(0.402)	(0.408)	(0.412)	(0.413)	(0.365)
			100	0.946	0.952	0.947	0.947	0.946✓	0.918✗
				(0.289)	(0.287)	(0.289)	(0.291)	(0.285)	(0.260)
		0.50	30	0.949	0.964✓	0.947	0.947	0.979	0.960
				(0.605)	(0.588)	(0.606)	(0.617)	(0.683)	(0.605)
			50	0.946	0.957✓	0.945	0.945	0.974	0.957
				(0.468)	(0.460)	(0.469)	(0.473)	(0.523)	(0.474)
			100	0.950	0.953✓	0.948	0.949	0.969	0.956
				(0.332)	(0.329)	(0.332)	(0.334)	(0.364)	(0.340)

Finally, Table 4 shows that except for the extreme cases where both the probability of consent $\hat{p}_{+1}$ and the number of patients assigned to each treatment n are small (i.e., $\hat{p}_{+1}$ = 0.30 and n = 30), the probability of failing to produce a 95% confidence interval by using interval estimators discussed here is generally negligible (≤ 0.002) in all the situations considered in Table 3.

An Example

Consider the randomized trial of studying vitamin A supplementation to reduce the mortality among preschool children in rural Indonesia (Sommer, Tarwotjo, Djunaedi, et al., 1986). As described previously, children were randomly assigned to either the treatment group of receiving a large oral dose of vitamin A for two to three months following baseline evaluation and again six months later, or to the control group without receiving any vitamin A supplementation. In the control group, children were precluded from receiving a placebo for ethical reasons and therefore, only the total number of survival children without the information on children who would fail to take vitamin A if they were assigned to receive this vitamin was observed. The results on mortality from month 4 (following completion of the first distribution cycle) to month 12 were compared.

As shown elsewhere (Sommer & Zeger, 1991), the number of survival children in the control group, $(m_{1+} =)$ was 11,514 out of $(m =)$ 11,588 children. Furthermore, the frequencies: $n_{11} = 9,663$, $n_{10} = 2,385$, $n_{01} = 12$, $n_{00} = 34$, were obtained for the total number 12,094 of children assigned to the treatment of receiving vitamin A. Suppose that one is interested in estimation of vitamin A effect on the non-death rate between the group receiving vitamin A and the group without receiving vitamin A supplement among children who would consent to accept vitamin A.

Given the above data, the MLE $\hat{\Delta}$ is 0.0032. Applying interval estimators (3), (4), (6), (9), as well as (10) with and without the continuity correction leads to produce 95% confidence intervals to be: [0.0010, 0.0055], [0.0010, 0.0055], [0.0010, 0.0055], [0.0010, 0.0055], [0.0008, 0.0061], and [0.0009, 0.0060], respectively. Because the total number of children in this randomized trial is quite large, the resulting 95% confidence interval using interval estimators (3), (4), (6), and (9) are essentially identical. Note that the 95% confidence interval using (10) with the continuity correction tends to have the length larger than the others. This is actually consistent with the previous findings obtained in simulations. Because all the above lower limits fall above 0, there is a significant evidence to support that taking vitamin A can increase the survival rate for preschool children at 5%-level.

Conclusion

It was found that the interval estimator (10) using the randomization-based approach with the continuity correction can lose efficiency, while this estimator without the continuity correction can be slightly liberal when the underlying difference Δ is not equal to 0. This article also finds that the interval estimator (4) using the $\tanh^{-1}(\hat{\Delta})$ transformation is probably preferable to all the other estimators discussed here with respect to the coverage probability and the average length. Thus, interval estimator (4) is recommended for general use.

Following Brunner and Neumann (1985) as well as Bernhard and Compagnone (1989), one can also discuss interval estimation of the selection effect, defined as $p^*_{11} / p^*_{+1} - p^*_{10} / p^*_{+0}$ which is the difference in probabilities of positive response for the standard treatment between patients who would agree and patients who would decline to receive the experimental treatment. By the functional invariance property, the MLE for this selection effect is simply equal to $(\hat{p}^*_{1+} - \hat{p}_{10} / \hat{p}_{+0}) / \hat{p}_{+1}$ (Appendix). Thus, it is straightforward to extend the above discussion to account for interval estimation of this selection effect. However, the detailed derivation and discussion on the selection effect are beyond the scope of this article and can be a future possible research topic.

Table 4. The estimated probability of failing to apply interval estimators (3), (4), (6), (9), as well as the interval estimator (10) with and without the continuity correction.

p_{+1}	Δ	$p^*_{r\|c}$	n	(3)	(4)	(6)	(9)	(10)	
0.30	0.00	0.20	30	0.012	0.012	0.013	0.017	0.047[a]	0.012[b]
			50	0.001	0.001	0.001	0.001	0.006	0.001
			100	0.000	0.000	0.000	0.000	0.000	0.000
		0.50	30	0.041	0.041	0.041	0.043	0.041	0.041
			50	0.010	0.010	0.010	0.010	0.010	0.010
			100	0.000	0.000	0.000	0.000	0.000	0.000
	0.10	0.20	30	0.010	0.010	0.010	0.014	0.041	0.010
			50	0.002	0.002	0.002	0.002	0.004	0.002
			100	0.000	0.000	0.000	0.000	0.000	0.000
		0.50	30	0.038	0.038	0.038	0.041	0.038	0.038
			50	0.008	0.008	0.008	0.008	0.008	0.008
			100	0.000	0.000	0.000	0.000	0.000	0.000
	0.20	0.20	30	0.016	0.016	0.016	0.020	0.048	0.016
			50	0.002	0.002	0.002	0.002	0.005	0.002
			100	0.000	0.000	0.000	0.000	0.000	0.000
		0.50	30	0.040	0.040	0.040	0.043	0.040	0.040
			50	0.009	0.009	0.009	0.009	0.009	0.009
			100	0.000	0.000	0.000	0.000	0.000	0.000
0.50	0.00	0.20	30	0.000	0.000	0.000	0.000	0.015	0.000
			50	0.000	0.000	0.000	0.000	0.001	0.000
			100	0.000	0.000	0.000	0.000	0.000	0.000
		0.50	30	0.002	0.002	0.002	0.002	0.002	0.002
			50	0.000	0.000	0.000	0.000	0.000	0.000
			100	0.000	0.000	0.000	0.000	0.000	0.000
	0.10	0.20	30	0.000	0.000	0.000	0.000	0.015	0.000
			50	0.000	0.000	0.000	0.000	0.000	0.000
			100	0.000	0.000	0.000	0.000	0.000	0.000
		0.50	30	0.001	0.001	0.001	0.001	0.001	0.001
			50	0.000	0.000	0.000	0.000	0.000	0.000
			100	0.000	0.000	0.000	0.000	0.000	0.000
	0.20	0.20	30	0.000	0.000	0.000	0.000	0.014	0.000
			50	0.000	0.000	0.000	0.000	0.001	0.000
			100	0.000	0.000	0.000	0.000	0.000	0.000
		0.50	30	0.002	0.002	0.002	0.002	0.002	0.002
			50	0.000	0.000	0.000	0.000	0.000	0.000
			100	0.000	0.000	0.000	0.000	0.000	0.000

[a] : This column is calculated with the continuity correction.

[b] : This column is calculated without the continuity correction

Table 4 (continued)

0.80	0.00	0.20	30	0.000	0.000	0.000	0.000	0.004	0.000
			50	0.000	0.000	0.000	0.000	0.000	0.000
			100	0.000	0.000	0.000	0.000	0.000	0.000
		0.50	30	0.000	0.000	0.000	0.000	0.000	0.000
			50	0.000	0.000	0.000	0.000	0.000	0.000
			100	0.000	0.000	0.000	0.000	0.000	0.000
	0.10	0.20	30	0.000	0.000	0.000	0.000	0.003	0.000
			50	0.000	0.000	0.000	0.000	0.000	0.000
			100	0.000	0.000	0.000	0.000	0.000	0.000
		0.50	30	0.000	0.000	0.000	0.000	0.000	0.000
			50	0.000	0.000	0.000	0.000	0.000	0.000
			100	0.000	0.000	0.000	0.000	0.000	0.000
	0.20	0.20	30	0.000	0.000	0.000	0.000	0.003	0.000
			50	0.000	0.000	0.000	0.000	0.000	0.000
			100	0.000	0.000	0.000	0.000	0.000	0.000
		0.50	30	0.000	0.000	0.000	0.000	0.000	0.000
			50	0.000	0.000	0.000	0.000	0.000	0.000
			100	0.000	0.000	0.000	0.000	0.000	0.000

In summary, six interval estimators for estimating the risk difference in a simple compliance randomized trial have been developed. The evaluation and comparison of the finite-sample performance of these interval estimators have been carried out in a variety of situations. The interval estimator using the $\tanh^{-1}(\hat{\Delta})$ transformation of the MLE has been shown to be the best among all interval estimators for the risk difference considered here. The results and the discussion presented in this article should have use for biostatisticians and clinicians when they encounter data under a simple compliance randomized trial.

References

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, *91*, 444-455.

Anbar, D. (1983). The relative efficiency of Zelen's prerandomization design for clinical trials. *Biometrics*, *39*, 711-718.

Bernhard, G. & Compagnone, D. (1989). Binary data in prerandomized designs. *Biometrical Journal*, *31*, 29-33.

Brunner, E. & Neumann, N. (1985). On the mathematical basis of Zelen's prerandomized designs. *Methods of Information in Medicine*, *24*, 120-130.

Casella, G. & Berger, R. L. (1990). *Statistical Inference*. Belmont, CA: Duxbury.

Edwardes, M. D. (1995). A confidence interval for Pr(X < Y) – P(X > Y) estimated from simple cluster samples. *Biometrics*, *51*, 571-578.

Ellenberg, S. S. (1984). Randomization designs in comparative clinical trials. *The New England Journal of Medicine*, *310*, 1404-1408.

Lui, K.-J. (2002). Notes on estimation of the general odds ratio and the general risk difference for paired-sample data. *Biometrical Journal*, *44*, 957-968.

Lui, K.-J. & Lin, C.-D. (2003). Interval estimation of treatment effects in double consent randomized design. *Statistica Sinica*, *13*, 179-187.

Mark, S. D. & Robins, J. M. (1993). A method for the analysis of randomized trials with compliance information: an application to the multiple risk factor intervention trial. *Controlled Clinical Trials*, *14*, 79-97.

Matts, J. & McHugh, R. (1987). Randomization and efficiency in Zelen's single consent design. *Biometrics*, *43*, 885-894.

Matts, J. P. and McHugh, R. B. (1993). Precision estimation in Zelen's single-consent design. *Biometrical Journal*, *35*, 65-72.

McHugh, R. (1984). Validity and treatment dilution in Zelen's single consent design. *Statistics in Medicine*, *3*, 215-218.

SAS Institute Inc. (1990). *SAS Language, Reference Version 6*, 1st edition. Cary, North Carolina: SAS Institute.

Sato, T. (2000). Sample size calculations with compliance information. *Statistics in Medicine*, *19*, 2689-2697.

Sommer, A., Tarwotjo, I., Djunaedi, E., West, K. P., Loedin, A. A., Tilden, R., & Mele, L. (1986). Impact of vitamin A Supplementation on childhood mortality: A randomized controlled community trial. *Lancet i,* 1169-1173.

Sommer, A. & Zeger, S. L. (1991). On estimating efficacy from clinical trials. *Statistics in Medicine*, *10*, 45-52.

Zelen, M. (1979). A new design for randomized clinical trials. *The New England Journal of Medicine*, *300*, 1242-1245.

Zelen, M. (1986). Response. *Journal of Chronic Disease*, *39*, 247-249.

Zelen, M. (1990). Randomized consent designs for clinical trials: an update. *Statistics in Medicine*, *9*, 645-656.

Appendix

Let R denote the response random variable: R = 1 if the underlying response is positive and = 0, otherwise. Furthermore, let T and C denote the random variables for the treatment assignment and the acceptance of the experimental treatment, respectively: T = 1 for the experimental treatment, and T = 0 for the standard treatment; C = 1 for accepting the experimental treatment and C = 0, otherwise. Let $P(R=1|T=i)$ denote the probability of response for treatment i (i=1 and 0). Because patients were randomly assigned to the treatment, it may reasonably be assumed that the proportion of patients who would consent to accept the experimental between the experimental and standard treatment groups are equal

$$(\text{i.e., } P(C=1|T=1) = P(C=1|T=0) = P(C=1))$$

Note that

$$\begin{aligned} &P(R=1|T=1) - P(R=1|T=0) = \\ &[P(R=1|T=1, C=1) - \\ &P(R=1|T=0, C=1)]\ P(C=1|T=1) + \\ &[P(R=1|T=1, C=0) - \\ &P(R=1|T=0, C=0)]\ P(C=0|T=1)\,. \end{aligned} \tag{A.1}$$

Because a patient assigned to the experimental treatment who does not consent to accept his/her assigned treatment will receive the standard treatment (i.e., T = 0), by the exclusion restriction assumption (Angrist, Imbens, & Rubin, 1996), the equality $P(R=1|T=1, C=0) = P(R=1|T=0, C=0)$ holds. Thus, the second component of (A.1) is zero. These suggests that equation (A.1) can be expressed in terms of notations p_{ij} and p_{ij}^* as

$$p_{1+} - p_{1+}^* = \Delta\ p_{+1} \tag{A.2}$$

where $\Delta = (p_{11} / p_{+1}) - (p_{11}^* / p_{+1}^*)$ is the difference in probabilities of positive response among patients who would consent to accept the experimental treatment. Say, there are n and m patients independently randomly assigned to receive the experimental and standard treatments, respectively. Let n_{ij} (and m_{ij}) denote the observed frequency corresponding to the cell probability p_{ij} (and p_{ij}^*) in the experimental (and the standard) treatment. Then, the random vector $n = (n_{11}, n_{10}, n_{01}, n_{00})'$ follows the multinomial distribution with parameters n and $(p_{11}, p_{10}, p_{01}, p_{00})'$. Furthermore, the marginal number $m_{1+} (= m_{11} + m_{10})$ of patients with positive response in the standard treatment

follows the binomial distribution with parameters m and $p^*_{1+}(=p^*_{11}+p^*_{10})$. Therefore, by the functional invariance property (Casella & Berger, 1990) of the MLE, the MLE for Δ is

$$\hat{\Delta}=(\hat{p}_{1+}-\hat{p}^*_{1+})/\hat{p}_{+1} \qquad \text{(A.3)}$$

where $\hat{p}_{1+}=\hat{p}_{11}+\hat{p}_{10}$, $\hat{p}_{+1}=\hat{p}_{11}+\hat{p}_{01}$, and $\hat{p}^*_{1+}=\hat{p}^*_{11}+\hat{p}^*_{10}$, where $\hat{p}_{ij}=n_{ij}/n$ and $\hat{p}^*_{ij}=m_{ij}/m$. By the Central Limit Theorem (Casella & Berger, 1990), as both n and m are large, the random vector $(\hat{p}_{1+},\hat{p}_{+1},\hat{p}^*_{1+})'$ asymptotically follows the multivariate normal distribution with mean vector $(p_{1+},p_{+1},p^*_{1+})'$ and covariance matrix with diagonal terms given by

$$Var(\hat{p}_{1+})=p_{1+}(1-p_{1+})/n,$$
$$Var(\hat{p}_{+1})=p_{+1}(1-p_{+1})/n,$$
$$Var(\hat{p}^*_{1+})=p^*_{1+}(1-p^*_{1+})/m,$$

and off-diagonal terms given by

$$Cov(\hat{p}_{1+},\hat{p}_{+1})=(p_{11}-p_{1+}p_{+1})/n,$$

and

$$Cov(\hat{p}_{1+},\hat{p}^*_{1+})=Cov(\hat{p}_{+1},\hat{p}^*_{1+})=0.$$

On the basis of the delta method, an asymptotic variance of $\hat{\Delta}$ is given by

$$Var(\hat{\Delta})=\Delta[(p_{1+}-p^*_{1+})(1-p_{+1})$$
$$-2(p_{11}-p_{1+}p_{+1})]/(np^2_{+1})+$$
$$p_{1+}(1-p_{1+})/(np^2_{+1})+p^*_{1+}(1-p^*_{1+})/(mp^2_{+1}). \qquad \text{(A.4)}$$

From (A.4), the following variance estimator

$$\hat{Var}(\hat{\Delta})=[\hat{p}_{1+}(\hat{p}_{10}+\hat{p}_{01})$$
$$-\hat{p}^*_{1+}(2\hat{p}_{10}-\hat{p}^*_{1+}(1-\hat{p}_{+1}))]/(n\hat{p}^3_{+1})$$
$$+\hat{p}^*_{1+}(1-\hat{p}^*_{1+})/(m\hat{p}^2_{+1}) \qquad \text{(A.5)}$$

is obtained by simply substituting $\hat{\Delta}$ for Δ, $\hat{p}_{1+}$ for p_{1+}, $\hat{p}_{+1}$ for p_{+1}, $\hat{p}^*_{1+}$ for p^*_{1+}, and $\hat{p}_{11}$ for p_{11}. For assessing the selection effect (Brunner & Neumann, 1985, Bernhard & Compagnone, 1989), consider the following difference:

$$P(R=1|T=0)-P(R=1|T=1,C=0)$$
$$=[P(R=1|T=0,C=1)$$
$$-P(R=1|T=0,C=0)]P(C=1|T=1) \qquad \text{(A.6)}$$

In terms of notations p_{ij} and p^*_{ij}, formula (A.6) may be rewritten as

$$p^*_{1+}-p_{10}/p_{+0}=(p^*_{11}/p^*_{+1}-p^*_{10}/p^*_{+0})\,p_{+1}. \qquad \text{(A.7)}$$

Thus, the selection effect, defined as $p^*_{11}/p^*_{+1}-p^*_{10}/p^*_{+0}$ which is simply the difference in probabilities of positive response for the standard treatment between patients who would accept and patients who would decline to the experimental treatment, can be estimated by using the MLE:

$$(\hat{p}^*_{1+}-\hat{p}_{10}/\hat{p}_{+0})/\hat{p}_{+1}. \qquad \text{(A.8)}$$