*Far East J. Theo. Stat.* **13**(1) (2004), 1-13
(A special volume on Biostatistics and related areas)

# INTERVAL BASED SAMPLE SIZE CALCULATION IN CASE-CONTROL STUDIES WITH MATCHED PAIRS UNDER INVERSE SAMPLING

**CHII-DEAN LIN and KUNG-JONG LUI**

Department of Mathematics and Statistics
San Diego State University, San Diego, CA 92182, U. S. A.

## Abstract

This paper provides a quantitative discussion on the effect resulting from the predetermined number of indexed matched pairs, defined as the pair in which the case is unexposed but the control is exposed, on the average length of the confidence interval. For a given ratio of the average length of the confidence interval relative to the underlying odds ratio, this paper provides a sample size calculation procedure to determine the required number of the indexed matched pairs under inverse sampling. To further facilitate use of the proposed sample size determination procedure and to help understanding of the effect due to the underlying mean and variation of the probability of exposure in the case group, this paper presents tables that summarize the required number of indexed matched pairs derived on the basis of the average length of the 95% confidence interval, the expected total number of matched pairs, and the minimum total number of matched pairs needed to obtain the required number of indexed matched pairs with a given probability in a variety of situations.



## 1. Introduction

Considering two independent samples under inverse sampling, in which we continue to sample subjects until we collect the desired number of index subjects (Haldane [7, 8], Finney [5, 6], Lui [15, 16, 17]). Barton

2000 Mathematics Subject Classification: 62D05.

Key words and phrases: sample size, case-control study, matched pairs, inverse sampling.

Received November 18, 2003

[1] notes use of large sample approximation to calculate the required number of index subjects for a desired power to detect a specific alternative hypothesis at a given type I error. When the underlying disease is rare, Bennett [2] calculates the power of an approximate F-test procedure as a function of the number of index subjects under inverse sampling. Lui [13] discusses estimation and testing hypothesis for case-control studies with matched pairs under inverse sampling. Considering the case of two independent samples, Lui [14] further derives a sample size calculation procedure for a given power based on hypothesis testing under inverse sampling. On the other hand, interval estimation can often provide more information on the contents of data than hypothesis testing. None of these papers focuses, however, discussion on sample size calculation on the basis of the average length of the confidence interval for case-control studies with matched pairs under inverse sampling.

The purpose of this paper is to provide a quantitative discussion on the effect resulting from the predetermined number of indexed matched pairs, in which the case is unexposed but the control is exposed, on the average length of the confidence interval. For a given ratio of the average length of the confidence interval relative to the underlying odds ratio, this paper provides a sample size calculation procedure of the minimum required number of the indexed matched pairs under inverse sampling. To facilitate the use of the results presented here and to study the effect due to the underlying mean and variation of the probability of exposure in the case group, this paper presents tables that summarize the minimum required number of indexed matched pairs derived on the basis of the average length of the 95% confidence interval, the expected total number of matched pairs, and the minimum total number of matched pairs needed to obtain the required number of indexed matched pairs with a given probability in a variety of situations. To simplify the calculation procedure, we also derive a sample size formula based on large sample theory. A quantitative evaluation of the accuracy of this approximate sample size is also included.



## 2. Notations and Model Assumptions

For a randomly selected subject from the case group, let $p$ denote the probability of exposure to a risk factor. Note that because each selected case has distinct individual characteristics, the probability of exposure $p$ may vary between cases. To account for this variation, we assume that $p$ follows a beta distribution with mean $\pi$ $(= \alpha/(\alpha + \beta))$ and variance $\pi(1 - \pi)/(T + 1)$, where $T = \alpha + \beta$, and $\alpha > 0$ and $\beta > 0$, because this family is rich in shapes and is commonly used to model the Bernoulli responses (Johnson and Kotz [10], Lui [11, 18]). Note that for a fixed $\pi$, the parameter $T$ may be regarded as a measure of the variation for $p$; the larger the value of $T$, the smaller is the variation of $p$. As $T$ goes to $\infty$, the probability $p$ converges to the constant $\pi$. For each randomly selected case, we form a matched pair by finding a control according to certain matching criteria. Let $p'$ denote the corresponding probability of exposure for the matched control. As commonly assumed in matched pair studies (Jewell [9], Ejigou and McHugh [4], Cox [3]), we assume further that the odds ratio of exposure for the selected case versus the matched control, $p(1 - p')/[(1 - p)p']$, is equal to a constant $Q$ across all matched pairs. Thus, for a given value of the common odds ratio $Q$, the corresponding probability $p'$ to the matched control of a selected case with the probability of exposure $p$ is then equal to $p/[p + Q(1 - p)]$. Let $\pi_{ij}$ $(i = 1, 2,\ j = 1, 2)$ denote the probabilities for the following $2 \times 2$ table.

| | | Control | |
|---|---|---|---|
| | | Exposed | Non-exposed |
| Case | Exposed | $\pi_{11}$ | $\pi_{12}$ |
| | Non-exposed | $\pi_{21}$ | $\pi_{22}$ |

With the above assumption, these $\pi_{ij}$ are then given as follows:

$$\pi_{11} = \Gamma(\alpha + \beta)\left[\int_0^1 p^{\alpha+1}(1 - p)^{\beta-1}/[p + Q(1 - p)]dp\right]\Big/\Gamma(\alpha)\Gamma(\beta)$$

$$\pi_{10} = \Gamma(\alpha + \beta)\left[\int_0^1 p^{\alpha}(1 - p)^{\beta} Q/[p + Q(1 - p)]dp\right]\Big/\Gamma(\alpha)\Gamma(\beta)$$

$$\pi_{01} = \Gamma(\alpha+\beta)\left[\int_0^1 p^{\alpha}(1-p)^{\beta}/[p+Q(1-p)]dp\right]\Big/\Gamma(\alpha)\Gamma(\beta)$$

$$\pi_{00} = \Gamma(\alpha+\beta)\left[\int_0^1 p^{\alpha-1}(1-p)^{\beta+1}Q/[p+Q(1-p)]dp\right]\Big/\Gamma(\alpha)\Gamma(\beta). \quad (1)$$

Note that the odds ratio $Q$ is equal to $\pi_{10}/\pi_{01}$.

Consider the inverse sampling, in which we continue sampling subjects from the case group to form matched pairs until we obtain the predetermined number $c\,(\geq 1)$ of index matched pairs. For clarity, we summarize the resulting data by use of the following $2 \times 2$ table:

| | | Control: Exposed | Control: Non-exposed |
|---|---|---|---|
| Case | Exposed | $A$ | $B$ |
| | Non-exposed | $c$ | $D$ |

where $c$ is a positive integer and is predetermined, and random variables $A$, $B$ and $D$ denote the number of matched pairs falling into the corresponding categories before obtaining exactly the desired number $c$ of the indexed matched pairs. As noted elsewhere (Lui [14]), the marginal distribution $f_B(b \mid c,\ p)$ of $B$ is the negative binomial distribution with parameters $c$ and $p = \pi_{01}/(\pi_{10} + \pi_{01}) = 1/(1+Q)$:

$$f_B(b \mid c,\ p) = \frac{(b+c-1)!}{b!\,(c-1)!}\,p^c(1-p)^b, \quad \text{where} \quad b = 0, 1, 2, \ldots. \quad (2)$$

Therefore, with use of the confidence interval (Lui [12]) for $p$ and the monotonically decreasing transformation of $(1-x)/x$, the exact $100(1-\alpha)\%$ confidence interval for $Q$ is then given by $((1-u(b, c, \alpha/2))/u(b, c, \alpha/2), (1-\ell(b, c, \alpha/2))/\ell(b, c, \alpha/2))$, where $u(b, c, \alpha/2)$ $= cf_{\alpha/2}(2c,\ 2b)/[cf_{\alpha/2}(2c,\ 2b) + b]$, $\ell(b, c, \alpha/2) = c/[c + (b+1)f_{\alpha/2}(2(b+1), 2c)]$, and $f_{\alpha/2}(df_1,\ df_2)$ denotes the upper $100(\alpha/2)$th percentile of the central F-distribution with degrees of freedom equal to $df_1$ and $df_2$, respectively.

Note that if $b$ were 0, then we would define the upper limit $u(0, c, \alpha/2) = 1$ by convention and this leads the lower limit for the common odds ratio $Q$ to be 0. On the basis of the above interval estimation, the average length of the $100(1-\alpha)\%$ confidence interval is then given by

$$L(Q, c, \alpha) = \sum_{b=0}^{\infty} [(1 - \ell(b, c, \alpha/2))/\ell(b, c, \alpha/2)$$

$$- (1 - u(b, c, \alpha/2))/u(b, c, \alpha/2)] f_B(b \mid c, p). \qquad (3)$$

Note that the average length in (3) depends only on the underlying common odds ratio $Q$, the number $c$ of indexed matched pairs, and the desired confidence level $100(1-\alpha)\%$. Therefore, for a given $Q$ and a desired precision $\delta$, we can apply equation (3) to find the minimum number of indexed matched pairs $c$ such that the ratio $L(Q, c, \alpha)/Q \leq \delta$ by use of the trial-and-error method. To study the effect resulting from the underlying mean and variation of $p$ in the case group on the total number of matched pairs that we need for collecting the desired number $c$ of indexed matched pairs, we consider the expected total number $n_1$ of matched pairs by calculation of $n_1 = E(A + B + D) + c = c/\pi_{01}$. We further consider the minimum total number of matched pairs $n_2$ required for obtaining the $c$ indexed matched pairs with a given probability $p_r$ by use of the negative binomial cumulative distribution such that

$$n_2 = \min_n \sum_{i=c}^{n} \binom{i-1}{c-1} \pi_{01}^c (1 - \pi_{01})^{i-c} \geq p_r.$$

Note that since $n_1$ and $n_2$ are functions of probability $\pi_{01}$, which depends on $\pi$, $T$ and $Q$ and cannot be expressed explicitly in a closed form, we apply numerical procedures, such as trapezoidal rule or Monte Carlo integration to calculate $\pi_{01}$ given in (1) (Thisted [19]).

For comparison purposes, we calculate the required number of indexed matched pairs using asymptotical approximation. As shown in Lui [13], the expectation and variance of $B/c$ are $E(B/c) = Q$ and $Var(B/c) = Q(Q+1)/c$. An asymptotic $100(1-\alpha)\%$ confidence interval for the odds ratio $Q$ can be constructed by

$$(B/c - Z_{\alpha/2}\sqrt{Q(Q+1)/c},\ B/c + Z_{\alpha/2}\sqrt{Q(Q+1)/c}), \tag{4}$$

where $Z_{\alpha/2}$ denotes the upper $100(\alpha/2)$th percentile of the standard normal distribution. With a predetermined confidence interval length $\ell$, odds ratio $Q$, and nominal level $\alpha$, the required $c$ indexed matched pairs can be approximately calculated by

$$c^* = 4Z_{\alpha/2}^2 Q(Q+1)/\ell^2. \tag{5}$$

To examine the accuracy of the sample size calculation procedure (5) based on large sample approximation, we also calculate the relative error defined as

$$RE = (L(Q, c^*, \alpha) - \ell)/\ell. \tag{6}$$

## 3. Results

To investigate the effect resulting from an increase in the number $c$ of indexed matched pairs on the average length of a $100(1-\alpha)\%$ confidence interval, we calculate the average length defined in (3) of the 95% confidence interval in the situation where the odds ratio ranges from .25 to 4.0 and the number $c$ of indexed matched pairs ranges from 1 to 200 in Table 1. As one would expect, Table 1 shows that the average length of the confidence interval decreases as the number $c$ of indexed matched pairs increases from 1 to 200. For a given precision of $\delta$, Table 2 presents the minimum number $c$ of indexed matched pairs such that the average length of the 95% confidence interval relative to the underlying odds ratio $L(Q, c, \alpha)/Q \leq \delta$ in the situation where $\delta$ equals 1.0 and .5, and the odds ratio $Q$ ranges from .25 to 4. Furthermore, to assess the expected total number $n_1\ (= c/\pi_{01})$ of matched pairs for collecting the required number $c$ of indexed matched pairs and the corresponding minimum total number $n_2$ of matched pairs with a probability $\sum_{i=c}^{n}\binom{i-1}{c-1}\pi_{01}^{c}(1-\pi_{01})^{i-c} \geq .95$, Table 3 summaries the estimates of $n_1$ and $n_2$ in the situations, where the underlying mean probability of exposure $\pi$ in the case group ranges from .1 to .5, the measure of probability variation $T$ in the case group ranges from 1 to 9, the precision $\delta\,(= L(Q,\, c,\, \alpha)/Q)$ equals 1, .75, and .5, and the odds ratio $Q$ ranges from .25 to 4. To examine the accuracy of approximated sample size calculation procedure (5), Table 4 shows the required number of indexed matched pairs calculated such that the 95% interval length relative to the underlying odds ratio $\ell/Q \leq \delta$ in the situation where $\delta$ equals 1.0, .75, and .5, and the odds ratio $Q$ ranges from .25 to 4. Additionally, the relative error of the average length (6) is presented. Note that a positive value of $RE$ represents an underestimation of the number of indexed matched pairs. That is, we have a wider confidence interval than we expected. Similarly, a negative value of $RE$ means an overestimation of the required number of indexed matched pairs and we have a narrower confidence interval than we expected.

## 4. An Example

Suppose that the underlying odds ratio of exposure to a risk factor for the case versus the control is equal to 2. Say, we want to find out what the minimum number $c$ of indexed matched pairs is required to assure that the ratio $\delta$ of the average length of the 95% confidence interval relative to the underlying odds ratio of exposure is $\leq 1$. From Table 2, we can see that we need to take 28 indexed matched pairs. In this case, if the underlying mean probability of exposure $\pi$ is 30% and the coefficient of variation for $p$ were known to be 48% $(= \sqrt{(1-\pi)/[(T+1)\pi]}$, which leads the value of $T$ to be approximately equal to 9), respectively, the expected total number of matched pairs $n_1$ for collecting the required 28 indexed matched pairs would be 245 (Table 3). Furthermore, with a budget that can cover the exposure of collecting $n_2 = 321$ matched pairs, the probability that we will be able to collect the required 28 matched pairs within the budget limit is $\geq 95\%$. On the other hand, if we should have only the resource to collect 20 rather than 28 indexed matched pairs, as seen from Table 1, the average length of the 95% confidence interval would increase from approximate 2 to approximate 2.4 (or equivalently, the precision $\delta = L(Q,\, c,\, \alpha)/Q$ increases from $\approx 1$ to $\approx 1.2$) as the underlying odds ratio is 2 (Table 1).

## 5. Discussion

When the number $c$ of indexed matched pairs is small or moderate, note that increasing the number of indexed matched pairs is quite effective to reduce the average length of the 95% confidence interval (Table 1). For example, when the underlying odds ratio is 2, the average length of the 95% confidence interval reduces from 117.69 for $c = 1$ to 21.61 for $c = 2$, an increase by a single pair can reduce the average length of the 95% confidence interval for $Q$ by more than 80%. The effect on the reduction of the average length due to the number of $c$ gradually diminish, however, as $c$ is large. Therefore, in a study with a small number of matched pairs, taking a few additional matched pairs (if it is possible) is certainly desirable. This is similar to the finding appeared elsewhere (Lui [12]), but which focuses discussion on the interval estimation of the population prevalence rate under inverse sampling.

Note that as the odds ratio $Q$ increases, the probability of exposure $p/[p + Q(1 - p)]$ for the matched control decreases. This suggests that for given all the other conditions fixed the probability $\pi_{01}$ of obtaining a matched pair, in which the case is unexposed but the control is exposed, is small when $Q$ is large. Therefore, although the required number of $c$ indexed matched pairs is smaller for large $Q$ than that for small $Q$ (Table 2), both estimates of the total number of matched pairs $n_1$ and $n_2$ can still be larger in the former than in the later (Table 3). Furthermore, note that both $n_1$ and $n_2$ increase as the coefficient of variation $(= \sqrt{(1 - \pi)/[(T + 1)]}$ (or equivalently, as $T$ decreases) increases for all situations considered in Table 3). This suggests that for a given precision and an odds ratio, on average we may need to take the total number of matched pairs when the underlying probability of exposure $p$ in the case group has a large variation more than that when $p$ has a small variation.

When using an asymptotic confidence interval (4) to calculate the required number of the indexed matched pairs, the numbers are all underestimated (Table 4). For example, when $Q = .25$ and $\delta = 1$, the required number of indexed pairs calculated using (5) is equal to 77 whereas the required number of indexed pairs is 89 (Table 2) using (3).

The relative error is large when $\delta$ is large (Table 4). For example, when $\delta = 1$, the relative error ranges from 8% to 12% for various $Q$ (Table 4). When $\delta$ is small, however, the relative error (6) using the approximate sample size formula (5) is reasonably small (i.e., $\leq 3\%$).

In summary, this paper proposes a sample size calculation procedure on the basis of the average length of the confidence interval for studies with matched pairs under inverse sampling. This paper provides tables that summarize the required number of indexed matched pairs in a variety of situations. To simplify the sample size calculation procedure, this paper also provides a formula based on large sample theory. The approximate sample size is easy to calculate and performs well when $\delta$ is small. The discussion and the results presented here should be useful for applied statisticians and epidemiologists to determine the required sample size when employing a retrospective study design with matched pairs under inverse sampling.

## References

[1] D. Barton, Comparison of sample sizes in inverse binomial sampling, Technometrics 9 (1967), 337-339.

[2] B. Bennett, On the use of the negative binomial in epidemiology, Biometrical J. 23 (1981), 69-72.

[3] D. Cox, Two further applications of a model for binary regression, Biometrika 45 (1958), 552-556.

[4] A. Ejigou and R. McHugh, Estimation of relative risk from matched pairs in epidemiologic research, Biometrics 33 (1977), 552-556.

[5] D. Finney, Errors of estimation in inverse sampling, Nature 160 (1947), 195-196.

[6] D. Finney, On a method of estimating frequencies, Biometrika 36 (1949), 233-234.

[7] J. Haldane, A labour-saving method of sampling, Nature 155 (1945a), 49-50.

[8] J. Haldane, On a method of estimating frequencies, Biometrika 33 (1945b), 222-225.

[9] N. Jewell, Small-sample bias of point estimators of the odds ratio from matched sets, Biometrics 40 (1984), 421-435.

[10] N. Johnson and S. Kotz, Continuous Univariate Distributions, 2nd ed., Wiley, New York, 1994.

[11] K. J. Lui, Sample sizes for repeated measurements in dichotomous data, Statistics in Medicine 10 (1991), 463-472.

[12] K. J. Lui, Confidence limits for the population prevalence rate based on the negative binomial distribution, Statistics in Medicine 14 (1995), 1471-1477.

[13] K. J. Lui, Notes in case-control studies with matched pairs under inverse sampling, Biometrical J. 38 (1996a), 681-693.

[14] K. J. Lui, Sample size for the exact conditional test under inverse sampling, Statistics in Medicine 15 (1996b), 671-678.

[15] K. J. Lui, Conditional maximum likelihood estimate and exact test of the common relative difference in combination of $1 \times 2$ tables under inverse sampling, Biometrical J. 39 (1997a), 215-225.

[16] K. J. Lui, Exact equivalence test for risk ratio and its sample size determination under inverse sampling, Statistics in Medicine 16 (1997b), 1777-1786.

[17] K. J. Lui, Estimation of rate ratio and relative difference in matched-pairs under inverse sampling, EnvironMetrics 12 (2001a), 539-546.

[18] K. J. Lui, Interval estimation for simple difference in dichotomous data with repeated measurements, Biometrical J. 43 (2001b), 845-861.

[19] R. Thisted, Elements of Statistical Computing, Chapman & Hall, New York, 1988.



**Table 1.** The average length (rounded to 3 decimals) of the 95% confidence interval in the situations where the odds ratio $(Q)$ ranges from .25 to 4 and the number of indexed matched pairs $c$ ranges from 1 to 200

| | | | $Q$ | | |
|---|---|---|---|---|---|
| $c$ | .25 | .5 | 1 | 2 | 4 |
| 1 | 48.86 | 58.709 | 78.384 | 117.687 | 196.217 |
| 2 | 7.432 | 9.501 | 13.575 | 21.607 | 37.517 |
| 3 | 3.676 | 4.882 | 7.217 | 11.764 | 20.716 |
| 4 | 2.474 | 3.374 | 5.090 | 8.400 | 14.894 |
| 5 | 1.900 | 2.641 | 4.037 | 6.712 | 11.950 |
| 10 | .989 | 1.439 | 2.258 | 3.810 | 6.843 |
| 20 | .599 | .894 | 1.421 | 2.417 | 4.365 |
| 30 | .464 | .698 | 1.115 | 1.903 | 3.444 |
| 40 | .391 | .590 | .946 | 1.618 | 2.932 |
| 50 | .344 | .520 | .835 | 1.431 | 2.595 |
| 100 | .234 | .356 | .575 | .989 | 1.796 |
| 200 | .162 | .247 | .401 | .690 | 1.256 |

**Table 2.** The required number of indexed matched pairs, in which the case is unexposed but the control is exposed for the ratio δ of the average length of the 95% confidence interval to the underlying odds ratio ($Q$) to equal 1, .75, and .5 in the situations where the underlying odds ratio ranges from .25 to 4

| | $Q$ | | | | |
|---|---|---|---|---|---|
| δ | .25 | .5 | 1 | 2 | 4 |
| 1 | 89 | 54 | 37 | 28 | 24 |
| .75 | 151 | 91 | 61 | 46 | 39 |
| .5 | 325 | 197 | 131 | 98 | 82 |

**Table 3.** The expected total number of matched pairs $(n_1)$ and the minimum number $(n_2)$ of matched pairs needed to obtain the required number of indexed matched pairs with a probability ≥ 95% for the ratio δ of the average length of the 95% confidence interval to the underlying odds ratio $(Q)$ to equal 1 and .5 in the situations where the underlying odds ratio ranges from .25 to 4, the mean π and the variation $T$ of the probability of exposure in the case group ranges from .1 to .5, and 9 to 1, respectively

| | | | $Q$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| π | δ | $T$ | .25 | | .5 | | 1 | | 2 | | 4 | |
| | | | $n_1$ | $n_2$ | $n_1$ | $n_2$ | $n_1$ | $n_2$ | $n_1$ | $n_2$ | $n_1$ | $n_2$ |
| .1 | 1 | 9 | 400 | 463 | 388 | 472 | 457 | 582 | 629 | 832 | 1021 | 1383 |
| | | 4 | 488 | 568 | 455 | 555 | 514 | 655 | 681 | 901 | 1072 | 1452 |
| | | 7/3 | 596 | 695 | 539 | 658 | 587 | 749 | 751 | 995 | 1147 | 1554 |
| | | 1 | 920 | 1077 | 796 | 975 | 822 | 1051 | 992 | 1315 | 1428 | 1935 |
| | .5 | 9 | 1474 | 1594 | 1415 | 1572 | 1617 | 1847 | 2202 | 2572 | 3490 | 4140 |
| | | 4 | 1800 | 1951 | 1661 | 1848 | 1820 | 2079 | 2383 | 2784 | 3664 | 4346 |
| | | 7/3 | 2195 | 2382 | 1966 | 2189 | 2079 | 2377 | 2630 | 3073 | 3920 | 4651 |
| | | 1 | 3389 | 3687 | 2903 | 3239 | 2911 | 3332 | 3472 | 4060 | 4879 | 5790 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| .2 | 1 | 9 | 261 | 299 | 233 | 281 | 257 | 325 | 338 | 444 | 531 | 717 |
| | | 4 | 307 | 354 | 270 | 326 | 289 | 366 | 368 | 484 | 562 | 759 |
| | | 7/3 | 365 | 422 | 315 | 383 | 330 | 419 | 409 | 539 | 607 | 820 |
| | | 1 | 543 | 633 | 458 | 559 | 463 | 589 | 546 | 721 | 770 | 1041 |
| | .5 | 9 | 961 | 1033 | 850 | 939 | 910 | 1034 | 1181 | 1376 | 1815 | 2149 |
| | | 4 | 1131 | 1220 | 984 | 1089 | 1023 | 1165 | 1287 | 1499 | 1921 | 2275 |
| | | 7/3 | 1344 | 1452 | 1151 | 1277 | 1170 | 1333 | 1430 | 1667 | 2074 | 2456 |
| | | 1 | 2002 | 2172 | 1672 | 1861 | 1638 | 1870 | 1910 | 2230 | 2631 | 3119 |
| .3 | 1 | 9 | 227 | 259 | 189 | 226 | 196 | 246 | 245 | 321 | 372 | 501 |
| | | 4 | 259 | 297 | 216 | 260 | 220 | 277 | 269 | 352 | 397 | 535 |
| | | 7/3 | 301 | 347 | 250 | 302 | 252 | 318 | 301 | 395 | 434 | 584 |
| | | 1 | 434 | 504 | 358 | 435 | 352 | 447 | 407 | 536 | 562 | 759 |
| | .5 | 9 | 836 | 896 | 689 | 759 | 693 | 785 | 857 | 995 | 1271 | 1503 |
| | | 4 | 956 | 1028 | 787 | 869 | 780 | 885 | 940 | 1093 | 1358 | 1606 |
| | | 7/3 | 1110 | 1197 | 912 | 1009 | 891 | 1013 | 1052 | 1224 | 1482 | 1753 |
| | | 1 | 1601 | 1733 | 1305 | 1449 | 1248 | 1422 | 1423 | 1659 | 1921 | 2274 |
| .5 | 1 | 9 | 239 | 274 | 178 | 213 | 164 | 206 | 185 | 241 | 258 | 346 |
| | | 4 | 263 | 301 | 199 | 239 | 185 | 232 | 207 | 270 | 283 | 380 |
| | | 7/3 | 294 | 339 | 226 | 273 | 211 | 266 | 235 | 308 | 317 | 427 |
| | | 1 | 401 | 464 | 315 | 382 | 296 | 375 | 326 | 429 | 432 | 582 |
| | .5 | 9 | 882 | 947 | 650 | 716 | 582 | 658 | 647 | 750 | 882 | 1041 |
| | | 4 | 968 | 1041 | 727 | 801 | 655 | 742 | 723 | 839 | 968 | 1142 |
| | | 7/3 | 1085 | 1169 | 826 | 913 | 749 | 849 | 822 | 955 | 1085 | 1281 |
| | | 1 | 1476 | 1596 | 1148 | 1274 | 1048 | 1193 | 1142 | 1330 | 1476 | 1746 |

**Table 4.** The required number of indexed matched pairs, in which the case is unexposed but the control is exposed for the ratio δ of the average length of the 95% confidence interval to the underlying odds ratio ($Q$) to equal 1, .75 and .5 in the situations where the underlying odds ratio ranges from .25 to 4 using normal approximation. The relative error of the length (defined in (6)) is presented in parenthesis

| | $Q$ | | | | |
|---|---|---|---|---|---|
| δ | .25 | .5 | 1 | 2 | 4 |
| 1 | 77 (0.08) | 46 (0.09) | 31 (0.10) | 23 (0.11) | 19 (0.12) |
| .75 | 137 (0.05) | 82 (0.06) | 55 (0.06) | 41 (0.06) | 34 (0.07) |
| .5 | 307 (0.03) | 184 (0.03) | 123 (0.03) | 92 (0.03) | 77 (0.03) |

■