STATISTICS IN MEDICINE, VOL. 7, 1279–1288 (1988)

ON SAMPLE SIZES TO ESTIMATE THE PROTECTIVE EFFICACY OF A VACCINE

ROBERT T. O'NEILL

*Food and Drug Administration, Center for Drug Evaluation and Research, Office of Epidemiology and Biostatistics, Division of Biometrics, Parklawn Building, Room 18B-45, 5600 Fishers Lane, Rockville, Maryland 20857, U.S.A.**

SUMMARY

To estimate vaccine protective efficacy, defined as VE = 1 − ARV/ARU where ARV is the disease attack rate in the vaccinated group and ARU is the disease attack rate in the controls, investigators have used both cohort and case-control designs. For each design, we present a method for calculation of the sample size required to provide an approximate confidence interval for VE of predetermined width and probability of coverage. The required sample size is a function of the desired width of the confidence interval, the probability of coverage, the assumed VE, and, for cohort designs, the assumed disease attack rate in the controls, and for case-control designs, the assumed vaccine exposure prevalence for the controls.

KEY WORDS Sample size Cohort Case-control Confidence intervals Relative risk Odds ratio Vaccine efficacy

INTRODUCTION

Evaluation of the protective efficacy of a vaccine under field conditions[1] has involved two study designs. The comparative cohort study, which involves one group of subjects allocated the vaccine and a second group allocated a control agent (placebo or another vaccine), is the design with most experience behind its use and evaluation. Randomized controlled trials are generally acknowledged as the best method to assess the protective efficacy of a vaccine under field conditions. The case-control study has also been used to evaluate a vaccine's efficacy particularly when prior evidence of vaccine efficacy exists or if the vaccine has had previous licence for use and interest now turns towards further evaluations of efficacy in various subgroups. Experience with case-control designs and a consensus on its utility for vaccine evaluation is evolving. This paper considers determination of the sample size required in comparative cohort studies to estimate a measure of vaccine efficacy VE with a confidence interval of specified width and probability of coverage. I also consider case-control designs so as to contrast their sample size considerations when one uses the odds ratio to approximate the relative risk. Schlesselman[2] has addressed sample size considerations for the hypothesis testing framework in which the goal is to demonstrate a difference in attack rates of the disease under study for cohort and case-control designs; one can adapt his results for this purpose.

The traditional measure of protective efficacy of a vaccine is the index

$$VE = (ARU - ARV)/ARU = 1 - ARV/ARU$$

where ARV and ARU, respectively, are the attack rates of the disease under study among the

* The views expressed in this paper are those of the author and not necessarily of the Food and Drug Administration

0277–6715/88/121279–10$05.00

Received December 1987
Revised June 1988

vaccinated and unvaccinated cohorts. It is assumed that an effective vaccine will produce disease attack rates that are lower than in the control (that is, ARV < ARU). We interpret the index VE as the proportion of cases of disease prevented by the vaccine. The index VE has been used to characterize the effectiveness of the Salk polio vaccine,[3] measles vaccine,[4,5] pertussis vaccines[6] and Hib capsular polysaccharide vaccine.[7]

In this paper I present a method for calculating the sample sizes required to estimate the index VE with a prespecified degree of precision. Although a field trial might provide sufficient information to reject the null hypothesis that the disease attack rate is reduced relative to that for the non-vaccinated cohort, the trial may not be large enough to afford a precise estimate of the true protective efficacy rate. The confidence interval reflects the magnitude of precision of the estimator of VE; the wider is this interval, the less assurance is there regarding true vaccine effectiveness. As an example, Daum[8] reported the efficacy of *Haemophilus influenzae* type b polysaccharide vaccine in Finnish children age 24–35 months as 79·9 per cent with 95 per cent confidence limits of (7 per cent, 95 per cent). He based this on 8453 vaccinated children and 8573 control children for whom the statistical comparison of differences in attack rates was significant at $p = 0{\cdot}04$. While there is strong statistical evidence that the vaccine reduces the attack rate, the estimate of vaccine efficacy VE is less informative because of the magnitude of the width of the confidence interval.

COMPARATIVE COHORT TRIALS

Confidence intervals for VE

With vaccine trials, several methods for computing the confidence intervals for a ratio of two binomial outcomes have appeared. Ederer and Mantel[9] proposed a method based on the assumption that counts follow a Poisson distribution; Santosham *et al.*[7] used this method to evaluate bacterial polysaccharide immune globulin. Another method based upon the log of the ratio of two binomial random variables (Katz *et al.*[10]) appears more frequently; we have chosen it here because of its simplicity of interpretation and of the confidence interval's symmetry on the log scale.

Let ARV(T) and ARU(T), respectively, denote the probabilities of experiencing the disease under study in a defined interval of time T for the vaccinated and unvaccinated cohorts. Let $\psi(T) = \text{ARV}(T)/\text{ARU}(T)$ so that the efficacy index is defined as $\text{VE}(T) = 1 - \psi(T)$ and $\beta(T) = \ln\psi(T)$. For the remainder of this section we drop the notation with respect to T and assume a fixed period of time T.

In sample cohorts of N_1 vaccinated and N_2 unvaccinated subjects, we can display the cases of disease according to their presence or absence as in Table I. The estimate of ψ is $\hat{\psi} = (x/N_1)/(y/N_2)$ and $\hat{\beta} = \ln\hat{\psi}$. For large samples, β is asymptotically normally distributed with variance

$$\sigma^2 = (1 - \text{ARV})/N_1(\text{ARV}) + (1 - \text{ARU})/N_2(\text{ARU}) \tag{1}$$

(see Koopman,[11] and Katz *et al.*[10]). Substitution of the sample estimates $\hat{\text{ARU}} = x/N_1$ and $\hat{\text{ARV}} = y/N_2$ yields an estimate of the variance of β,

$$\sigma^2 = (N_1 - x)/N_1 x + (N_2 - y)/N_2 y = 1/x - 1/N_1 + 1/y - 1/N_2. \tag{2}$$

Usually, because the incidence rate of disease in a vaccine trial is small, the sample sizes N_1 and N_2 required to observe disease cases will be very large so that the variance of β is largely influenced by and approximated by $1/x + 1/y$, which is a function of the numbers of cases x in the vaccinated and unvaccinated y groups.

Table I

	Cases	Non-cases	Total
Vaccinated	x	N_1-x	N_1
Unvaccinated	y	N_2-y	N_2
Total	$x+y$	$N_1+N_2-(x+y)$	N_1+N_2

The derivation of a confidence interval for VE is equivalent to derivation of a confidence interval for ψ with subtraction of 1 from each of the limits. To derive a confidence interval for ψ, one method (Katz[10]) recommended for large sample sizes (as is the situation in vaccine trials) is first to derive a confidence interval for $\beta=\ln\psi$. That is, the 100 $(1-\alpha)$ per cent confidence interval for β is

$$\hat{\beta}\pm z\hat{\sigma} \tag{3}$$

where z denotes the $(1-\alpha)$ percentage point of the standardized normal distribution.

The 100 $(1-\alpha)$ per cent confidence interval for ψ is

$$\exp(\hat{\beta}\pm z\hat{\sigma}). \tag{4}$$

The 100 $(1-\alpha)$ per cent confidence interval for $\mathrm{VE}=1-\psi$ is

$$[1-\exp(\hat{\beta}+z\hat{\sigma}),\ 1-\exp(\hat{\beta}-z\hat{\sigma})] \tag{5}$$

whose width $W(\hat{\beta},d)$ is the difference between the upper and lower limits, namely:

$$(1-\exp(\hat{\beta}-z\hat{\sigma}))-(1-\exp(\hat{\beta}+z\hat{\sigma}))=W(\hat{\beta},\hat{d}). \tag{6}$$

Then

$$W(\hat{\beta},\hat{d})=\exp(\hat{\beta})\ (\exp(\hat{d})-\exp(-\hat{d}))\ \text{where}\ \hat{d}=z\hat{\sigma}. \tag{7}$$

The width of the confidence interval relative to $\hat{\mathrm{VE}}=1-\hat{\psi}=1-\exp(\hat{\beta})$ is

$$\begin{aligned} R\hat{W}(d,\hat{\mathrm{VE}}) &= (\exp(\hat{\beta})/[(1-\exp(\hat{\beta}))\ (\exp(\hat{d})-\hat{\exp}(-\hat{d}))] \\ &= [(1-\hat{\mathrm{VE}})/\hat{\mathrm{VE}}]\ [(\exp(\hat{d})-\exp(-\hat{d})]. \end{aligned} \tag{8}$$

RELATIVE WIDTH

The index VE is bounded above since vaccine efficacy is maximum at 100 per cent and is unbounded below in that VE could be negative and substantially so depending on the relationship between ARV and ARU. For planning purposes, the situation of most interest is when VE is between 0 and 100 per cent, so it is within this range of VE that sample sizes will be focused. The confidence interval for VE is asymmetric about the point estimator $\hat{\mathrm{VE}}$, the lower and upper limit being, respectively, $1-\exp(\hat{\beta}+\hat{d})$ and $1-\exp(\hat{\beta}-\hat{d})$ where $\hat{d}=z\hat{\sigma}$. Thus, the distance between the upper limit and the point estimate is $\exp(\hat{\beta})(\exp(\hat{d})-1)$ and the distance between the point estimate and the lower limit is $\exp(\hat{\beta})(1-\exp(-\hat{d}))$. Thus, to specify the desired width of a confidence interval for VE, an investigator must be aware of this asymmetry.

If an investigator chooses to specify the half width d of the interval on the log scale, then it is useful to interpret the interval width on the original scale in terms of relative width. For example, a relative width equal to 0·40 means that the total width of the confidence interval for VE is 40 per cent of the magnitude of the point estimate of VE. If VE = 0·80 or 80 per cent efficacy, the

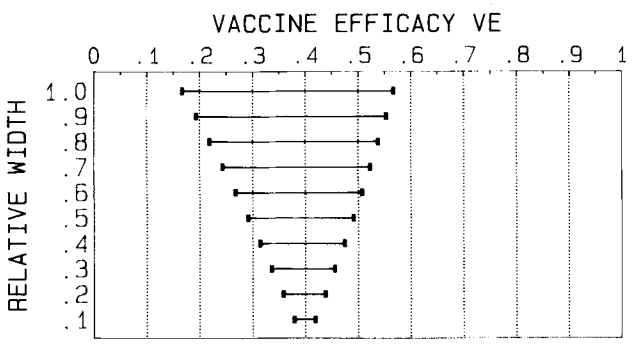

Figure 1. 95 per cent confidence intervals for VE, VE = 40 per cent

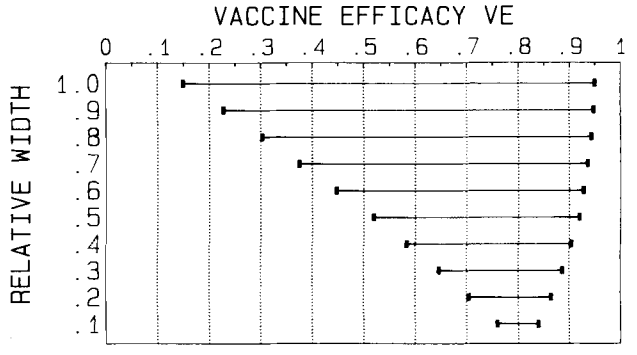

Figure 2. 95 per cent confidence intervals for VE, VE = 80 per cent

confidence interval for VE has width (0·80)(0·40) = 0·322, with the total width proportioned in a manner such that the half length for the lower bound is (0·80)(exp(0·2) − 1) = 0·177 and the half length for the upper bound is (0·80)(1 − exp(−0·2)) = 0·145. The confidence interval is then (0·80 − 0·177, 0·80 + 0·145) = (0·623, 0·945) whose absolute width is 0·945 − 0·623 = 0·322 and whose width relative to the point estimate 0·8 is 0·322/0·80 = 0·40 or 40 per cent.

To illustrate the concept of relative width of the confidence interval for VE, Figures 1 and 2 provide graphical representations of the widths of 95 per cent confidence intervals for VE for assumed values of VE = 40 per cent and 80 per cent, respectively. Each figure displays the 95 per cent confidence interval for VE with a relative width ranging from 0·1 to 1·0. That is, the figures display the width of the confidence interval for intervals that range from 10 per cent to 100 per cent of the magnitude of the point estimate of VE, for each of VE = 40 per cent and VE = 80 per cent. Note the asymmetry of all intervals about $\hat{\text{VE}}$ but the intervals become more so as the magnitude of the relative widths increases towards 1·0.

Sample size determination

Assume equal numbers N of vaccinated and unvaccinated subjects. Recalling that $d = z\sigma$ and that σ^2 is given by (1) then one can solve this equation for N and obtain the relationship

$$N = (z/d)^2((1 - \text{ARV})/\text{ARV} + (1 - \text{ARU})/\text{ARU})). \tag{9}$$

Table II. Cohort sample sizes for a 95 per cent confidence interval for VE = 40 per cent, for selected relative widths and attack rates, ARU, in controls

	ARU			
Relative width	0·01	0·005	0·001	0·0005
1·0	9,482	19,037	95,469	191,011
0·9	11,630	23,348	117,093	234,274
0·8	14, 632	29,375	147,317	294,745
0·7	19,010	38,164	191,395	382,933
0·6	25,755	51,704	259,300	518,795
0·5	36,940	74,159	371,910	744,100
0·4	57,530	115,494	579,208	1,158,851
0·3	102,013	204,796	1,027,064	2,054,898
0·2	229,106	459,943	2,306,604	4,615,012
0·1	915,408	1,837,773	9,216,337	18,439,591

Table III. Cohort sample sizes for a 95 per cent confidence interval for VE = 80 per cent, for selected relative widths and attack rates, ARU, in controls

	ARU			
Relative width	0·01	0·005	0·001	0·0005
1·0	1,102	2,208	11,056	22,116
0·9	1,260	2,524	12,653	25,274
0·8	1,473	2,950	14,771	29,547
0·7	1,774	3,554	17,793	35,592
0·6	2,226	4,456	22,323	44,654
0·5	2,957	5,924	29,662	59,334
0·4	4,280	8,573	42,924	85,863
0·3	7,100	14,224	71,213	142,450
0·2	15,101	30,252	151,464	302,979
0·1	22,184	44,409	222,205	444,451

Noting that ARV $=\psi$ ARU, and substitution of this into the above equation, yields

$$N=(z/d)^2(1/\mathrm{ARU})((1-\mathrm{ARU})/\mathrm{ARU})+(1-\mathrm{ARU})/\mathrm{ARU}).$$
$$=(z/d)^2((1+1/\psi)/\mathrm{ARU}-2). \quad (10)$$

Thus, to determine the required sample sizes N, an investigator must specify the probability of coverage $(1-\alpha)$ for the confidence interval, the assumed attack rate in the unvaccinated group ARU, the anticipated vaccine efficacy VE, and the desired width W (or desired relative width RW) of the confidence interval. The desired value of d is determined by solving either equation (7) or (8) where the expected value is substituted for the observed value. It is a matter of choice as to whether an investigator chooses d based upon a desired width or relative width of the confidence interval. To illustrate the calculations, Tables II and III, respectively, provide the sample sizes required for each of the vaccinated and unvaccinated cohorts for 95 per cent confidence intervals for VE, for an

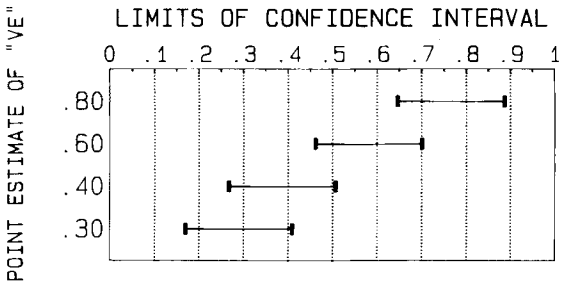

Figure 3. 95 per cent confidence intervals of width 0·24 for different estimates of 'VE', 'ARU' = 0·01 The sample sizes required for each interval differ

Table IV. Sample sizes required for a constant width of 0·24 for specified 'VE' and associated d and RW, ARU = 0·01

VE	ARV/ARU	RW	d	Confidence limits	Subject size
0·8	0·2	0·3	0·569	(0·65, 0·89)	7,100
0·6	0·4	0·4	0·296	(0·46, 0·70)	15,292
0·4	0·6	0·6	0·199	(0·27, 0·51)	25,755
0·3	0·7	0·8	0·171	(0·17, 0·41)	31,792

assumed $\hat{V}E = 40$ per cent and $\hat{V}E = 80$ per cent; and for selected assumed attack rates in the unvaccinated cohort of 0·01, 0·005, 0·001 and 0·0005; and for relative interval widths from 0·1 to 1·0 (0·1).

The sample size required to achieve a confidence interval of a fixed prespecified width can change dramatically as a function of the attack rate in the control cohort and the magnitude of the index VE. Figure 3 and Table IV illustrate these points. Figure 3 shows for an attack rate in the control cohort of 0·01, the location of 95 per cent confidence intervals of the same constant width for assumed VE of 0·80, 0·60, 0·40 and 0·30. In all four situations the confidence intervals have the same width, namely $W = 0·2$ to 4. The width relative to the point estimate VE, however, differs for each interval. Figure 3 displays the relative width as well as the limits of the confidence intervals and the sample sizes in the vaccinated and unvaccinated cohorts required to achieve a confidence interval of the specified width of $W = 0·24$. Clearly, as the efficacy of the vaccine increases towards 100 per cent the required sample size decreases.

An example

An investigator desires a 95 per cent confidence interval for VE where he/she anticipates the attack rate in the unvaccinated cohort is 0·005, the vaccine efficacy is 80 per cent and the desired relative width of the interval is 0·30. That is, the absolute width of the interval is (0·7)(0·3) = 0·21. The tables indicate choice of $N = 14{,}224$ in each of the vaccinated and unvaccinated cohorts to meet this objective.

CASE-CONTROL STUDIES

Although case-control studies to assess efficacy of some vaccines have appeared, such as the pertussis vaccine, thorough evaluation of their utility to assess efficacy accurately continues.[1,12]

Table V. Data layout for case-control design

Vaccine exposure	Cases	Controls	Total
Yes	a	b	M_1
No	c	d	M_2
Total	N_1	N_2	N

These observational studies are prone to considerable bias and may best suit situations where reliable comparative cohort trials have already demonstrated vaccine efficacy. With case-control studies one cannot assess disease attack rates directly but can approximate the ratio of attack rates (relative risk) with the odds ratio and estimate vaccine efficacy in this manner. This situation requires knowledge of the vaccination histories of (disease) cases and controls (no disease). Under certain circumstances, the odds ratio may not approximate the relative risk such as when attack rates are high. With attack rates in the vaccinated cohorts greater than 10 per cent, one will estimate vaccine efficacy as erroneously high with use of a case-control study.[1]

We approach the sample sizes required to estimate vaccine efficacy VE (where VE = 1 − OR) with confidence interval of prescribed width W similarly to our approach with cohort trials except that with the case-control design the parameter we must specify is the prevalence of vaccine exposure in the controls and we define VE = 1 − OR. The expression for the standard error of the log odds ratio differs from that for the log relative risk with the cohort approach. O'Neill[13] has examined sample size considerations to estimate the odds ratio with specified precision for case-control designs. We use the symbol Φ to denote the odds ratio $P_1(1-P_2)/P_2(1-P_1)$ where P_1 and P_2 are, respectively, the prevalences of vaccine exposure in the cases and controls. Table V provides the data layout for a case-control design.

The variance of ln Φ is

$$\sigma^2 = 1/N_1P_1(1-P_1) + 1/N_2P_2(1-P_2), \tag{11}$$

estimated by

$$\hat{\sigma}^2 = \frac{1}{N_1(a/N_1)(c/N_1)} + \frac{1}{N_2(b/N_2)(d/N_2)}$$
$$= \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}. \tag{12}$$

Use of the same reasoning as with the cohort design, but now with $N_2 = CN_1$, we can set $\sigma = dz$ in (11) and solve for N_1, to obtain the relationship

$$N_1 = (z/d)^2[1/P_1(1-P_1) + (1/CP_2(1-P_2)]$$

which we can rewrite as

$$N_1 = (z/d)^2[1/A(1-A) + 1/CP_2(1-P_2)]. \tag{13}$$

where $A = P_2(1-\text{VE})/[1-P_2(\text{VE})]$.

We choose d in a manner similar to that for the cohort design. That is, for a given desired relative width RW and VE, we can express d as the solution to the equation

$$\exp(d) - \exp(-d) = RW((1-\text{VE})/\text{VE}). \tag{14}$$

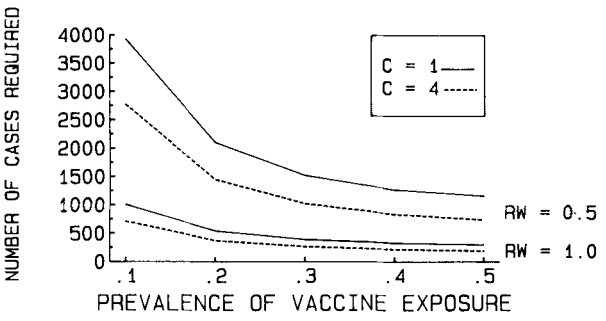

Figure 4. Cases required to estimate VE = 40 per cent for case-control ratios of 1 and 4

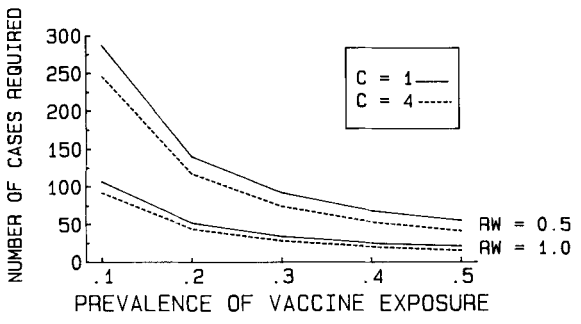

Figure 5. Cases required to estimate VE = 80 per cent for case-control ratios of 1 and 4

O'Neill[13] provides a discussion for case-control designs of the rationale for a selection of d to obtain a desired relative width W.

Figures 4 and 5 illustrate for VE = 40 per cent and 80 per cent respectively, the number of cases required to estimate VE with a 95 per cent confidence interval for two values of the relative width $RW = 0{\cdot}5$ and $1{\cdot}0$, and for a range of exposure prevalence rates from $0{\cdot}1$ to $0{\cdot}5$. The dotted line figures illustrate the sample sizes required for a matching ratio between cases and controls of 1 and 4.

An example

An investigator desires a 95 per cent confidence interval for VE where he anticipates that the prevalence rate of vaccine exposure in the control group is 20 per cent, and the vaccine efficacy VE is 80 per cent and the desired relative width of the interval is $0{\cdot}30$. That is, the absolute width of the confidence interval will be $(0{\cdot}8)(0{\cdot}3) = 0{\cdot}24$. Use of formula (13) for $C = 1$ where the numbers of cases and controls are equal yields a case sample size of 336. For $C = 4$, a 4 to 1 control to case size ratio, the case sample size is 280.

DISCUSSION

The formulae for cohort sample size calculations rely on the assumption that the denominators are large and the attack rates are small and that the confidence intervals are approximate in their coverage. The validity of the sample size formulae relies on asymptotic normality theory for large

sample sizes, which should apply in most vaccine trials. For the cohort design, the number of cases has the most influence in determination of the precision of the estimator. One should be cautious that the validity of these sample size calculations is questionable when the expected number of cases of disease becomes small.

In a cohort trial, the sample size required to estimate a specific vaccine efficacy VE with a specified magnitude of precision relates to the magnitude of the attack rates in the control group. In general, the smaller the attack rates the larger are the required sample sizes. Smith, Rodrigues and Fine[12] point out that several measures of vaccine efficacy might apply depending upon whether one measures risk of disease development in a follow-up period in terms of numbers of subjects at risk or follow-up time at risk. These authors propose several models for how the vaccine might protect the vaccinated population that can dramatically influence the estimate of VE. The design considered here assumes that the accumulated number of cases over a specified interval of time T results from two independent binomial populations. Different considerations will apply when subjects have differential follow-up or exposure time or when the comparative effectiveness of the vaccine follows a non-constant hazard rate.

For the case control design, an analogous situation occurs. For a specified VE, the prevalence of vaccine exposure in the control group directly influences the required sample sizes for the cases and controls. The lower the prevalence of vaccine exposure in the controls, the larger is the required sample size. With case-control designs, however, there is more flexibility in sample size planning because a feasible option exists to select many more controls than cases (usually maximum efficiency accrues with a 4 to 1 matching ratio) to increase the efficiency of the estimator. One attains this advantage in efficiency to a greater extent and with less resource outlay for the case-control design relative to the cohort design. The method of sample size calculation proposed here pertains to unmatched designs. The analysis of matched case-control designs requires methodology that accounts for the matching factors and might take into account adjustments for other multivariate risk factors.[14] The precision of the estimators are affected in a more complex manner in these latter situations.

ACKNOWLEDGEMENTS

I would like to express my appreciation to Jane Tata for her assistance in preparing this manuscript and to the referees and editor for their suggestions which improved the manuscript.

REFERENCES

1. Orenstein, W. A., Bernier, R. H., Dondero, T. J., Hinman, A. R., Marks, J. S., Bart, K. J. and Sirotkin, B. 'Field evaluation of vaccine efficacy', *Bulletin of the World Health Organization*, **63** (6), 1055–1068 (1985).
2. Schlesselman, J. J. 'Sample size requirements in cohort and case-control studies of disease', *American Journal of Epidemiology*, **99**, 381–384 (1974).
3. Francis, T., Korns, R. F., Voight, R. B., Boisen, M., Hemphill, F. M., Napier, J. A. and Tolchinsky, E. 'An evaluation of the 1954 poliomyelitis vaccine trials-summary report', *American Journal of Public Health*, **45**, 5, 1–63 (1955).
4. Linnemann, C. C., Jr. 'Measles vaccine: immunity, reinfection and revaccination', *American Journal of Epidemiology*, **97**, 365–371 (1973).
5. Marks, J. S. 'Methodologic issues in the evaluation of vaccine effectiveness. Measles vaccine at 12 versus 15 months', *American Journal of Epidemiology*, **166**, 510–523 (1982).
6. Fine, P. E. M. and Clarkson, J. A. 'Reflections on the efficacy of pertussis vaccines', *Reviews of Infectious Diseases*, **9**, No. 5, 866–883 (1987).
7. Santosham, M., Reid, R., Ambrosino, D. M., Wolff, M. C., Almeido-Hill, J., Prieks, C., Aspery, K., Garrett, S., Croll, L., Foster, S., Burge, G., Page, P., Zacher, B., Moxon, R. and Siber, G. 'Prevention of Haemophilus Influenzae Type B infections in high-risk infants treated with bacterial polysaccharide immune globulin', *New England Journal of Medicine*, **317**, 923–929 (1987).

8. Daum, R. *Immunization Monitor*, **1**, No. 1, 3–7 (1987).
9. Ederer, F. and Mantel, N. 'Confidence limits on the ratio of two Poisson variables', *American Journal of Epidemiology*, **100**, 165–167 (1974).
10. Katz, D., Baptista, J., Azen, S. P. and Pike, M. C. 'Obtaining confidence intervals for the risk ratio in cohort studies', *Biometrics*, **34**, 469–474 (1978).
11. Koopman, P. A. R. 'Confidence intervals for the ratio of two binomial proportions', *Biometrics*, **40**, 513–517 (1984).
12. Smith, P. G., Rodrigues, L. C. and Fine, P. E. M. 'Assessment of the protective efficacy of vaccines against common disease using case-control and cohort studies', *International Journal of Epidemiology*, **13**, No. 1, 87–93 (1984).
13. O'Neill, R. T. 'Sample sizes for estimation of the odds ratio in unmatched case-control studies', *American Journal of Epidemiology*, **120**, 145–153 (1984).
14. Breslow, N. and Day, N. *Statistical Methods in Cancer Research Volume 1-The Analyses of Case-Control Studies*, IARC Publications No. 32, Lyon, France, 1980.