ESTIMATING THE POWER OF A CASE-CONTROL STUDY CONDUCTED WITHIN A COHORT

Cologne, John
Department of Statistics, Radiation Effects Research Foundation

Izumi, Shizue
Department of Statistics, Radiation Effects Research Foundation

KYUSHU UNIVERSITY

ESTIMATING THE POWER OF A CASE-CONTROL STUDY CONDUCTED WITHIN A COHORT

by

John Cologne   and   **Shizue Izumi**

ESTIMATING THE POWER OF A CASE-CONTROL STUDY CONDUCTED WITHIN A COHORT

By

John Cologne[*] and **Shizue Izumi**[†]

Abstract

Power of epidemiologic studies can be studied via simulation, but simulations can require substantial computational effort and time. We describe a method for estimating power in matched and nested case-control studies using the non-central Chi-square approximation to the distribution of the log-likelihood ratio test statistic. The non-centrality parameter is estimated by computing the likelihood ratio statistic using the expected values of the parameters under the null and alternative hypotheses. The method is compared to simulation results from an actual study evaluating various numbers of matched controls. There was reasonably close agreement between simulated and calculated values of the mean likelihood ratio test statistics, but the power estimates differed, perhaps due to small sample failure of the asymptotic distribution assumption.

1. Introduction

A goal of epidemiologic research is to investigate so-called *risk factors*, the causes of disease or death (the *outcome*). A major part of this is the evaluation of confounding and/or effect modification by factors other than the risk factor of interest, for which Professor Yanagawa provided useful mathematical descriptions (Yanagawa, 1979). Due to rarity of disease occurrence and/or difficulty obtaining covariate information on all members of a large follow-up (*cohort*) study, it is often more practical to study risk factors using a sample of subjects selected from the cohort (Breslow, 1996; Sato, 1992). Although cases can be sampled, we focus on the situation where all cases are used and a subset of non-cases is selected.

Several study designs are available for sampling from a cohort. These include the retrospective matched case-control (Breslow and Day, 1980), nested case-control (Breslow et al, 1983; Langholz and Thomas, 1990), case-cohort (Prentice, 1986; Barlow et al, 1999), and two-stage (Zhao and Lipsitz, 1992; Breslow and Chatterjee, 1999) designs. The matched case-control design utilizes the conditional likelihood in its analysis. The nested case-control design utilizes a partial likelihood, analogous to that used for Cox regression. The latter two designs utilize pseudo likelihoods or missing data likelihood methods. In this paper we focus on the case-control designs because their corresponding likelihoods have tractible asymptotic properties—particularly, asymptotic Chi-square

[*] Department of Statistics, Radiation Effects Research Foundation, 5-2 Hijiyama Park, Minami-ku, Hiroshima 732-0815, Japan. tel 81-82-261-1957 jcologne@rerf.jp

[†] Department of Statistics, Radiation Effects Research Foundation, 5-2 Hijiyama Park, Minami-ku, Hiroshima 732-0815, Japan.

distributions of the likelihood ratio test statistic under both null and alternative hypotheses, which is useful for estimating power.

The classical retrospective case-control design uses cases identified up to the end of follow-up and selects as controls a subset of persons who are free of outcome (*cumulative incidence proportion sampling*). The controls may be matched to the cases on important risk factors, such as age and gender, which leads to use of the conditional likelihood. The nested case-control design involves sampling controls from the risk sets, the persons who are at risk—free of outcome but still under observation—at each time a case occurs (*incidence density sampling*). Controls are selected using random or stratified sampling. In the nested case-control design, sampling is performed within risk sets, so a subject can be selected in more than one risk set and a case can serve as a control at times prior to becoming a case, analogous to the cohort.

The planning of a study—i.e., determining the required sample size—should involve an assessment of statistical power. Simulation studies of power can be time consuming, especially if there is time-dependent weighting as arises with stratified sampling in the nested case-control design (Langholz and Borgan, 1995). A more direct method of estimating power is therefore desirable. Recently Professor Yanagawa and his colleagues studied the asymptotic power of the Kudô-Tarumi test for the risk ratio in 2×2 tables and compared it with the Fisher exact test for the odds ratio (Makinosumi et al, 2003). Methods for the analysis of matched case-control studies include, but are not limited to, stratified 2×2 data (Breslow and Day, 1980). We describe and implement a method for estimating power in case-control studies using the non-central Chi-square distribution for the log-likelihood ratio test based on the likelihood (matched case-control studies) or partial likelihood (nested case-control studies), following methods proposed by Brown et al (1999) and Holford (2002).

2.   Likelihood ratio test

Let $l(\theta; y)$ be the log likelihood under parameter $\theta \in \Omega$ for observed outcome variable y. The likelihood ratio statistic for testing the null hypothesis $H_0 : \theta \in \Omega_0 \equiv \theta_0$ (i.e., $\theta = \theta_0$) is

$$V = 2 \left[l(\hat{\theta}; y) - l(\theta_0; y) \right] \quad , \tag{1}$$

where $\hat{\theta}$ is the maximum likelihood estimate of θ [i.e., $l(\hat{\theta}; y) = sup_{\theta \in \Omega} l(\theta; y)$]. For $y = i.i.d.$ $\{y_k\}$, (1) has asymptotically a Chi-square distribution with ν degrees of freedom under H_0, where ν is the difference in dimension between Ω and Ω_0, and is asymptotically equivalent to the score and Wald test statistics (Rao, 1973, pages 417-8).

Let x be a factor taking on either continuous or discrete values. Our interest is in ascertaining whether x increases the risk of suffering the outcome ($y = 1$) — i.e., testing whether x is a *risk factor*. The quantity of interest in epidemiologic studies is often the relative risk (risk ratio) $RR(x)$, or its analog in the retrospective case-control design, the odds ratio $OR(x)$. In relative-risk analyses based on cohort follow-up or sampling from the cohort risk sets, the relative risk is frequently modelled as a log-linear function,

$$RR(x) = \frac{\lambda(t|x)}{\lambda(t|0)} = e^{\beta_{\mathrm{RR}} x} \quad , \tag{2}$$

where $\lambda(t|\cdot)$ is the instantaneous hazard (rate of disease or death), so $\beta_{\mathrm{RR}} = \log\{RR(x = 1)\}$. In the retrospective case-control study, the log odds are often modelled with a linear function, so the odds ratio is

$$OR(x) = \frac{\Pr(y = 1|x)/[1 - \Pr(y = 1|x)]}{\Pr(y = 1|0)/[1 - \Pr(y = 1|0)]} = \frac{e^{\eta + \beta_{\mathrm{OR}}x}}{e^{\eta}} = e^{\beta_{\mathrm{OR}}x} \tag{3}$$

and $\beta_{\mathrm{OR}} = \log\{OR(x = 1)\}$. Other models may be used, such as the *excess relative risk model* $ERR(x) = RR(x) - 1 = \beta_{\mathrm{ERR}}x$.

The quantities $\lambda(t|0)$ and η generally involve further parameters ψ reflecting effects of age, birth cohort, gender, etc., so $\theta = (\psi, \beta)$. We consider all elements of ψ to be nuisance parameters subsumed in the nonparametrimetric estimation of $\lambda_\psi(t|0)$ or that covariates corresponding to η_ψ are matched. Then $\Omega = \{\psi, \beta\}$ with $\beta \in (-\infty, \infty)$. We typically test the null hypothesis of no effect of x, $\mathrm{H}_0 : \beta = \beta_0 = 0$ ($\Omega_0 = \{\psi, 0\}$) versus the general alternative $\mathrm{H}_A : \beta \neq 0$ ($\Omega_A = \Omega \setminus \beta_0$). To evaluate power, one specifies a specific alternative hypothesis, $\beta = \beta_A$ for some $\beta_A \neq 0$.

Suppose we have a cohort of individuals indexed by $k \in \{1, \ldots, N\}$, each observed until time T_k (for the kth individual, the time when outcome occurs or the end of follow-up, whichever is first), with time-dependent outcome variable $y_k(t)$ ($y_k(t) = 1$ if the outcome has occurred by time t, 0 otherwise). Note that $y_k(T_k) = 1$ for the cases and 0 for the non-cases. Denote the times that the cases occur (the failure times) by τ_i ($i = 1, \ldots C$). Let $\mathcal{R}(i)$ represent the *risk set*, the indices of all subjects who are at risk just prior to time τ_i ($\mathcal{R}(i) = \{k : T_k > \tau_i\}$), including the person who suffers the outcome at time τ_i (the *case*). The retrospective matched design matches to each case i m_i controls from among those who did not suffer the outcome during the entire follow-up period $[y_k(T_k) = 0]$. The nested case-control design selects for each case i m_i individuals from among those who are in $\mathcal{R}(i)$. Typically, $m_i \equiv m$.

The conditional likelihood for the matched case-control design has an exponential family form in the conditionally sufficient statistic (McCullagh and Nelder, 1989, p. 253). Under an exponential family,

$$l(\theta; y) = [y\theta - b(\theta)]/a(\sigma) + c(y, \sigma) \tag{4}$$

for known scale parameter σ and known functions $a(\cdot)$, $b(\cdot)$, and $c(\cdot)$ (McCullagh and Nelder, 1989). Define $\mathcal{S}(i)$ to be the set of indices of individuals in matched set i. The y_k are conditionally sufficient given $\sum y_k$ (the total number of cases), the canonical parameter is $\theta = \log\{\mu/(1 - \mu)\} = \beta x$, where $\mu = \mathrm{E}(y)$, and the log likelihood is

$$l_c(\beta; y) = \sum_i \left[\sum_{j \in \mathcal{S}(i)} y_j \beta x_j - \log \sum_{j \in \mathcal{S}(i)} e^{\beta x_j} \right] . \tag{5}$$

The likelihood ratio test statistic for the conditional likelihood is

$$V_c = 2 \left[l_c(\hat{\beta}; y) - l_c(0; y) \right] . \tag{6}$$

The partial likelihood for the nested case-control study is similar (Breslow and Day, 1987; Goldstein and Langholz, 1992), except that $\mathcal{S}(i)$ is replaced by the set of indices of subjects sampled from the risk sets, $\mathcal{R}^*(i) \subset \mathcal{R}(i)$, and y_j is replaced by $y_j(\tau_i)$, the

failure status of individual j at time τ_i [$y_j(\tau_i) = 0$ for all members of the risk set except the case]:

$$l_p(\beta; y) = \sum_i \left[\sum_{j \in \mathcal{R}^*(i)} y_j(\tau_i)\beta x_j - \log \sum_{j \in \mathcal{R}^*(i)} e^{\beta x_j} \right] \quad . \tag{7}$$

Unlike the conditional likelihood (5), the terms of the partial likelihood (7) are not independent, because subjects can appear in more than one sampled risk set. However, the partial likelihood has asymptotic properties similar to those of the conditional likelihood; in particular, its distribution is Chi-square under both H_0 and H_A (Anderson et al, 1993). The same for nested case-control sampling can be inferred from consistency and asymptotic normality of maximum partial likelihood estimates in the nested case-control study (Goldstein and Langholz, 1992). The likelihood ratio test statistic for the partial likelihood is

$$V_p = 2 \left[l_p(\hat{\beta}; y) - l_p(0; y) \right] \quad . \tag{8}$$

3. Estimating Power

Suppose that $\beta_A \neq 0$ is presumed to be the true effect of x. In practice we might estimate power by simulation: 1) generate many Monte Carlo data sets assuming β_A, 2) maximize the likelihood under the null and unconstrained models to compute the likelihood ratio statistic, and 3) examine the simulation distribution of the test statistic. In our experience, obtaining a reasonable estimate of this distribution generally requires many hundred simulations. In the case of the partial likelihood when there are risk-set dependent variables, such as stratified sampling weights, these simulations can take a great deal of time. We therefore describe an alternative approach based on the noncentral Chi-square approximation to the sampling distribution of the conditional or partial likelihood ratio test statistic under H_A.

Asymptotically the test of H_0 under H_A is trivial (has power 1) because the likelihood ratio test is consistent. However, for alternatives $\beta_A^{(N)}$ such that

$$(\beta_A^{(N)} - \beta_0) = O\left(\frac{1}{\sqrt{N}}\right), \tag{9}$$

the asymptotic distribution of the likelihood ratio test of H_0 is non-central Chi-square (Cox and Hinkley, 1974, p. 318). Thus, we can approximate the power of the likelihood ratio test of H_0 under a specific alternative hypothesis H_A using the non-central Chi-square distribution, as long as H_A is close to H_0.

We estimate the noncentrality parameter by using the expected value of (6) or (8) given the y_k. Because $\hat{\beta}$ consistently estimates β_A, we substitute β_A for $\hat{\beta}$ (note that the conditional and partial likelihoods do not depend on ψ).

We must also specify the distribution of the exposure variable x. For categorical x, Holford proposed specifying the risk parameter and the distribution of cases and controls among the categories of x (Holford, 2002). This approach does not work for continuous x; we therefore input the mean value of x given case/control status—i.e., in place of x we substitute $E(x|y)$. The case-control ratio in the sample is fixed by design, and we

do not need to further specify η. Then, since $E\left[\chi^2(\delta;\nu)\right] = \nu + \delta$ for a non-central Chi-square variable with ν degrees of freedom and noncentrality parameter δ,

$$\delta = 2\left[l(\beta_A;\tilde{x},y) - l(0;\tilde{x},y)\right] - \nu$$

is the noncentrality parameter. Brown et al noted that it made little difference whether or not ν was subtracted (Brown et al, 1999).

We can estimate the power to reject $H_0 : \beta = 0$, based on a level α test when $\beta = \beta_A$, by calculating the probability that a non-central Chi-square distributed variable V with non-centrality parameter δ, $V \sim \chi^2(\delta;1)$, is greater than the $(1-\alpha)$ quantile of the central Chi-square distribution with one degree of freedom:

$$\text{power} = \Pr\left[V > \chi^2_{1-\alpha}(0;1)\right] \quad . \tag{10}$$

For example, if we are testing at level $\alpha = 0.05$, the estimated power is the probability that the non-central Chi-square is greater than 3.8415. Thus, the problem of estimating power based on the likelihood ratio test statistic is reduced to calculating the noncentrality parameter for a large-sample approximating Chi-square distribution. This non-centrality parameter is estimated by calculating the likelihood ratio statistic based on the hypothetical parameter values and the distribution of the risk factor conditional on the outcome status.

Suppose the model is multidimensional with parameters (γ,β) and only β is to be tested. The method of substituting expected values of the parameters under H_A to derive the Chi-square noncentrality parameter is not strictly correct because the maximum likelihood estimates of the other parameters $(\hat{\gamma})$ do not converge to γ when the reduced model is fit, as described by Self et al (1992). They decomposed the likelihood ratio statistic as:

$$\begin{aligned}
2\left[l(\hat{\gamma},\hat{\beta}) - l(\hat{\gamma}_0,\beta_0)\right] &= 2[l(\hat{\gamma},\hat{\beta}) - l(\gamma,\beta)] \\
&\quad - 2[l(\hat{\gamma}_0,\beta_0) - l(\gamma_0^*,\beta_0)] \\
&\quad + 2[l(\gamma,\beta) - l(\gamma_0^*,\beta_0)]
\end{aligned} \tag{11}$$

where γ_0^* is the limiting value of $\hat{\gamma}_0$. In other words, if H_A is true, $\hat{\gamma}_0$ is generally not consistent for γ (Self and Mauritsen, 1988). However, Self et al (1992) showed that in exponential families the third term of (11) is

$$2\sum_{i=1}^n \{b'(\theta_i)[\theta_i - \theta_i^*]/a_i(\sigma) - [b(\theta_i) - b(\theta_i^*)]\}$$

which is the computed estimate of the likelihood ratio test statistic using expected values, and the other terms cancel when $(\gamma_0^*,\beta_0) = (\gamma,\beta)$, the true parameter value. Thus, under assumption (9), the calculated Chi-square noncentrality parameter computed using expected values of y under H_A should be close to the noncentrality parameter of the true asymptotic Chi-square distribution.

Interest may be in estimating the effect of x only, or in estimating the effect of x with adjustment for confounding by another risk factor, z, or in testing whether there is effect modification by a factor z of the risk due to x (biological interaction). In the first

case, we need only specify β_0 and β_A. In the latter two instances, there are parameters other than β that must be accounted for. We additionally specify what their values are under the null and alternative models. For example, if we are interested in the effect of x with adjustment for possible confounding by another risk factor z, a possible model is

$$RR(x, z) = \exp\{\gamma z + \beta x\} \quad . \tag{12}$$

Then we must specify γ_0^*, the value of γ estimated under the constrained model with $\beta = 0$ when H_A is true, and γ_A, the value of γ for the full model when $\beta = \beta_A$. Or, we may be interested in possible effect modification between two factors z_1 and z_2, and consider the multiplicative interaction model

$$RR(z_1, z_2) = \exp\{\gamma z_1 + \phi z_2 + \beta x\} \quad ,$$

where $x = z_1 z_2$. Now we must specify two sets of values: $\{\gamma_0^*, \phi_0^*\}$, the maximum likelihood estimates of γ and ϕ under the constrained model fit with $\beta = 0$ (when H_A is true) and $\{\gamma_A, \phi_A\}$ when $\beta = \beta_A$.

If we are calculating power in a situation where one covariate (z_1, say) is already measured in the cohort, it is convenient to treat z_2 by centering it at its population mean ($z_2^* = z_2 - \bar{z}_2$) so that we can use the information already available to specify γ_A [$\gamma_A z_1 = RR(z_1 | z_2 = \bar{z}_2)$], which can be used in the case-control analysis if z_1 and z_2 are independent (Land et al, 1994).

4.   Example

This example is based on studying the risk of an additional factor measured in a case-control study. Power was previously assessed by simulation (Cologne et al, 2004). We illustrate the method for the test of a main effect using the model (12).

A nested case-control study was planned to investigate the joint risk of radiation and the endogenous hormone estradiol on pre-menopausal breast cancer. Radiation exposure (z) was known for all individuals in the cohort, but measuring estradiol (x) is costly and requires the use of precious, limited stored serum specimens. Eighty cases were diagnosed during the follow-up period among women for whom stored serum specimens were available; there were 5,644 cancer-free subjects with stored serum. Controls selected within risk sets were to be matched to cases on age and time of serum collection. The investigators wanted to know whether a small number of matched controls would be sufficient, to avoid consuming many serum specimens. Because the probability of repeated selection is small, we estimated power based on the conditional likelihood for a retrospective matched case-control design.

Information about γ was available from the cohort (i.e., based implicitly on unmeasured z, so that the cohort estimate of γ may be assumed to represent the effect of z for the average value of x, $\bar{x}$). Some information about β was available from a small previous study where controls were matched to cases on radiation dose (Kabuto et al, 2000). Given the known cases and the distribution of z among cases and controls, as well as the distribution of x from the previous study, we simulated case-control samples with values of x randomly generated according to a log-normal distribution with mean 2.19 among cases and 2.04 among controls, and $\beta = 2.393$. Five hundred simulations were run with 1, 2, 4, or 8 controls matched to each case. We fit the model (12) to the simulated data using the conditional likelihood (PECAN; Epicure software, HiroSoft

Table 1: Estimated power for the matched case-control study of interaction.

Controls per case	Power		
	Simulated	Calculated	Percent over-estimated
cohort	1.0	0.995	-
1	0.738	0.839	13.7
2	0.880	0.939	6.7
4	0.934	0.976	4.5
8	0.984	0.988	0.4

International Corporation, Seattle, WA) and calculated the likelihood ratio statistic for testing $H_0 : \beta = 0$. The distributions of simulated test values are shown in Figure 1.

We used the existing information about radiation to compute the values of γ_A and γ_0^* as follows. The cohort estimate of γ (not adjusted for estradiol) was 0.638 (call this γ^c). Because z and x are independent, adjusting for x should not affect γ, so $\gamma_A = .638$ also (the simulated value was .631).

Using these values we calculated expected values of the likelihood ratio test statistic as described in section 3. (solid line in Figure 1). The calculated values are close to the means of the simulated values. We estimated power as the proportion of simulated likelihood ratio statistics that exceeded the 95% level of the central Chi-square distribution with 1 degree of freedom (Table 1). Power estimated according to the expected value of the likelihood ratio test statistic overestimated power compared with the simulation results (Table 1), with overestimates ranging from more than 13% with one control to less than 1% with 8 controls. One reason for this overestimation may be poor approximation of the asymptotic distribution. Another reason may be the magnitude of distance between H_A and H_0 in this example.

5.   Discussion and conclusions

The idea of using the asymptotic non-central Chi-square distribution to estimate power with likelihood ratio tests is not new (see, for example, Holford (2002) and Brown et al (1999)). We examined application of the approach to matched case-control studies using the conditional likelihood. Because the form of the partial likelihood for a nested case-control is identical, implementation of the method in nested case-control studies would be the same. However, the small-sample distribution properties of the partial likelihood under null and alternative hypotheses may differ from that of the conditional likelihood. Further work is required to better understand what sample sizes are required for good approximation of power using the proposed method. Further work is also needed to examine sensitivity of the method to the magnitude of difference between H_A and H_0. An alternative to power is the relative efficiency (ratio of cohort and case-control variances). Wacholder et al (1991) described a method for estimating relative efficiency based on cohort size and sample sizes at the risk sets.

The interpretation of power calculations can be misleading without considering the power of a full cohort analysis, because there is little point in conducting a relatively efficient study based on sampling if the full cohort does not have adequate power. It

is therefore important to evaluate not only the power of the sample, but the power of the full cohort as well. We endorse a two-step approach to planning a study involving sampling from a cohort:

1. Determine the power of the full cohort analysis. If the cohort power is inadequate, no further study should be contemplated. If the cohort power is adequate, then,

2. Assess the power of the sampling design relative to that of the full cohort.

The estimated relative power will aid in deciding what proportion of the cohort to sample.

If an exposure variable is already measured in the cohort, sampling based on that variable can improve efficiency and power. One effective method is to sample controls so as to provide balance across exposure strata in each individual risk set (counter matching; see Langholz and Goldstein (1996) and Cologne and Langholz (2003)). Another—albeit somewhat limited in generality—approach is to match controls to cases on exposure (Land et al, 1994; Thomas and Greenland, 1985). *Randomized recruitment* is a design that offers a compromise between matching and random sampling that can achieve probability (frequency) matching on exposure without the analysis limitations caused by matching on exposure (Weinberg and Wacholder, 1990; Weinberg and Sandler, 1991). Power computations as described here should be feasible for any of these exposure-based sampling designs.

Acknowledgement

References

Andersen PK., Borgan O., Gill RD., and Keiding N. (1993). *Statistical Models Based on Counting Processes*, Springer-Verlag, New York.

Barlow WE., Ichikawa L., Rosner D., and Izumi S. (1999). Analysis of case-cohort designs, *J Clin Epidemiol*, **52**, 1165–72.

Breslow NE. (1996). Statistics in epidemiology: The case-control study, *J Amer Statist Assoc* **91**, 14–28.

Breslow NE. and Chatterjee N. (1999). Design and analysis of two-phase studies with binary outcome applied to Wilms tumour prognosis, *Appl Statist*, **48**, 457–68.

Breslow NE. and Day NE. (1980). *Statistical Methods in Cancer Research: Volume I–The Analysis of Case-Control Studies*, IARC Scientific Publications, Lyon.

Breslow NE. and Day NE. (1987). *Statistical Methods in Cancer Research: Volume II–The Design and Analysis of Cohort Studies*, IARC Scientific Publications, Lyon.

Breslow NE., Lubin JH., Marek P., and Langholz B. (1983). Multiplicative models and cohort analysis, *J Amer Statist Assoc*, **78**, 1–12.

Brown BW., Lovato J., and Russell K. (1999). Asymptotic power calculations: Description, examples, computer code, *Statistics in Medicine*, **18**, 3137–51.

Cologne J. and Langholz B. (2003). Selecting controls for assessing interacton in nested case-control studies, *J Epidemiol*, **13**, 184–93.

Cologne JB., Sharp GB., Neriishi K., Verkasalo PK., Land CE., and Nakachi K. (2004). Improving the efficiency of nested case-control studies of interaction by selecting controls using counter matching on exposure, *Int J Epidemiol*, **33**, 1–8.

Cox DR. and Hinkley DV. (1974). *Theoretical Statistics*, Chapman and Hall, London.

Goldstein L. and Langholz B. (1992). Asymptotic theory for nested case-control sampling in the Cox regression model, *Ann Statist*, **20**, 1903–28.

Holford TR. (2002). *Multivariate Methods in Epidemiology*, Oxford University Press, Oxford (pages 295–310).

Kabuto M., Akiba S., Stevens RG., Neriishi K., and Land CE. (2000). A prospective study of estradiol and breast cancer in Japanese women, *Cancer Epidemiol Biomark Prev*, **9**, 575–9.

Land CE., Hayakawa N., Machado SG., Yamada Y., Pike MC.,, Akiba S., and Tokunaga M. (1994). A case-control interview study of breast cancer among Japanese A-bomb survivors. II. Interactions with radiation dose, *Cancer Causes Control*, **5**, 167–76.

Langholz B. and Borgan Ø. (1995). Counter-matching: A stratified nested case-control sampling method, *Biometrika*, **82**, 69–79.

Langholz B. and Goldstein L. (1996). Risk set sampling in epidemiologic cohort studies, *Stat Sci*, **11**, 35–53.

Langholz B. and Thomas DC. (1990). Nested case-control and case-cohort methods of sampling from a cohort: A critical comparison, *Am J Epidemiol*, **131**, 169–76.

Makinosumi Y., Siraha A., and Yanagawa T. (2003). A conditional test for 2×2 tables: Asymptotic power of the Kudô-Tarumi test, *Japan J Biomet*, **24**, 5–16. (in Japanese)

McCullagh P., and Nelder JA. (1989). *Generalized Linear Models, Second Edition*, Chapman & Hall, London.

Prentice RL. (1986). A case-cohort design for epidemiologic cohort studies and disease prevention trials, *Biometrika*, **73**, 1–11.

Rao, CR. (1973). *Linear Statistical Inference and its Applications, Second Edition*, John Wiley & Sons, New York.

Sato T. (1992). Epidemiologic study designs and the estimation of relative risk, *Proceedings of the Institute of Statistical Mathematics*, **40**, 173–84. (in Japanese)

Self SG. and Mauritsen RH. (1988). Power/sample size calculations for generalized lin-

ear models, *Biometrics*, **44**, 79–86.

Self SG., Mauritsen RH., and Ohara J. (1992). Power calculations for likelihood ratio tests in generalized linear models, *Biometrics*, **48**, 31–9.

Thomas DC. and Greenland S. (1985). The efficiency of matching in case-control studies of risk-factor interactions, *J Chron Dis*, **38**, 569–74.

Wacholder S., Gail M., and Pee D. (1991). Selecting an efficient design for assessing exposure-disease relationships in an assembled cohort, *Biometrics*, **47**, 63–76.

Weinberg CR. and Wacholder S. (1990). The design and analysis of case-control studies with biased sampling, *Biometrics*, **46**, 963–75.

Weinberg CR. and Sandler DP. (1991). Randomized recruitment in case-control studies, *Am J Epidemiol*, **134**, 421–32.

Yanagawa T. (1979). Designing case-control studies, *Environ Health Perspect*, **32**, 143–56.

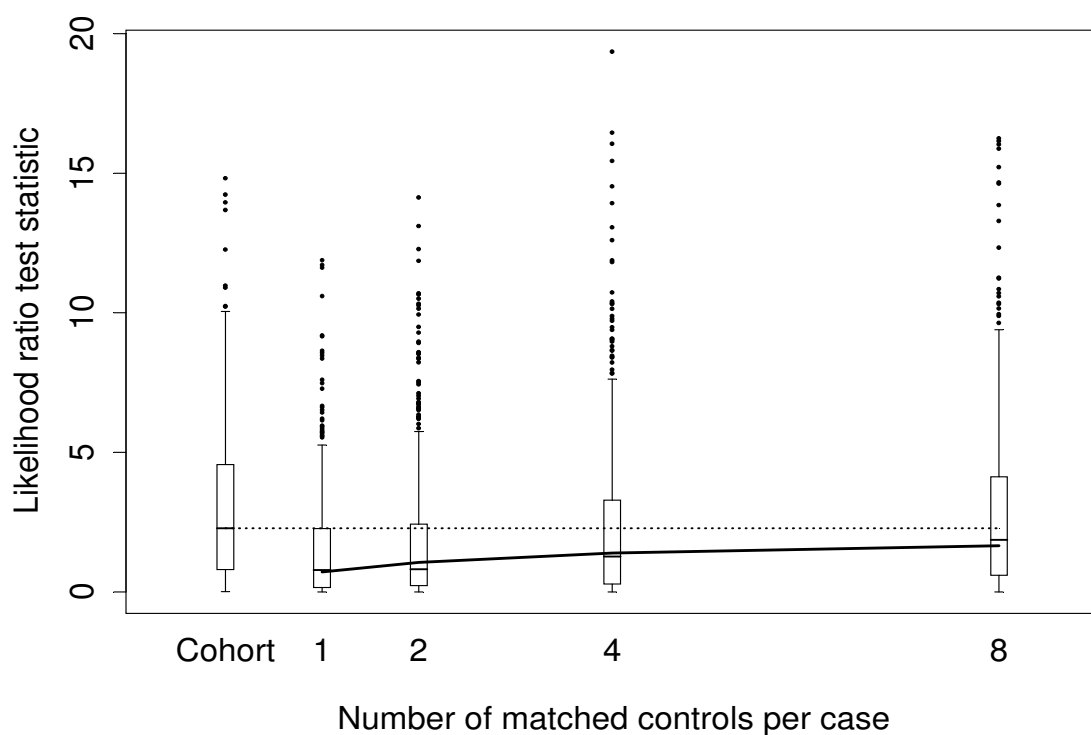Zhao LP. and Lipsitz S. (1992). Designs and analysis of two-stage studies, *Stat Medicine*, **11**, 769–82.

Figure 1: Likelihood ratio statistics for testing the effect of risk factor x in the matched case-control study example. Boxplots summarize the simulated values (500 values per sample size). The solid line connects the calculated expected values of the likelihood ratio statistics based on the proposed method. The horizontal dotted line is the median likelihood ratio test value for the cohort.