Empirical evaluation of the influence of control selection schemes on relative risk estimation: the Welsh nickel workers study

Alfredo Morabia, Tom Ten Have, J Richard Landis

Clinical Epidemiology
Division, University
Hospital of Canton
Geneva, Switzerland
A Morabia

Center for
Biostatistics and
Epidemiology, Penn
State University
College of Medicine,
Hershey,
Pennsylvania, USA
T Ten Have
J R Landis

Correspondence to:
Dr Alfredo Morabia,
Clinical Epidemiology
Division, University
Hospital of Canton Geneva,
Rue Micheli-du-Crest 25,
1211 Geneve 14,
Switzerland.

Accepted 20 March 1995

Abstract

Objective—To perform an empirical evaluation of the theory that relative incidence rate (RIR) and relative risk (RR) can be directly estimated from case-control studies that have different sampling schemes of controls.

Methods—With data from the South Wales nickel refinery workers (SWNRW) study, a nested case-control study of the relation of nickel exposure to respiratory cancers, was conducted within each of four fixed subcohorts that differed for stability of exposure, incidence rates and RIR. Odds ratios (ORs) and confidence intervals (CIs) were estimated either with all available controls or with randomly sampled subsets of controls.

Results—Respiratory cancers were not rare as risk of nasal and lung cancer in workers unexposed to nickel varied from 15% to 26% over the full risk period. The RIR was adequately estimated by the OR when controls were identified concurrently to case occurrence throughout the risk period. The RR was well approximated with the OR when controls were a sample of the study base.

Conclusions—These results add empirical support to the theory that the RIR or the RR can be validly estimated in case-control studies. Overall, this theory is relatively tolerant of large departures from the stability assumptions of exposure and of incidence.

(*Occup Environ Med* 1995;52:489–493)

A wealth of theoretical work shows that different measures of relative effect are obtained according to the way controls are selected in case-control studies.[1-8] It is usually accepted that "if a case-control study is based on incident cases occurring over a specified interval of time, with controls for each case being chosen from among those who are disease free at the time the case is diagnosed, then the odds ratio determined from a case-control study can be interpreted as the ratio of the instantaneous rates of disease in exposed *v* unexposed groups of individuals".[1] It is also accepted that the odds ratio (OR) can be interpreted as being the ratio of the risk of disease (defined in table 1) in exposed *v* unexposed people when controls are selected from all disease free subjects at the beginning of the risk period—that is, before the period during which new cases will be found.

The relations between the OR and relative risk (RR) or relative incidence rate (RIR) can be understood by viewing the case-control study as a variant of a cohort study[1 9] in which cases are all or a fraction of all the cases that will occur within the baseline population (population at risk at the start of the risk period, typically called the base) and controls are all or a fraction of all subjects within the base who did not develop the disease during the risk period.[8 9]

In this context, we can identify three types of case-control studies according to whether controls are sampled: (*a*) at the end of the risk period (traditional case-control study); (*b*) from the population at risk during the risk period (concurrent case-control study); or (*c*) from the base (case-base study). Table 1 gives the notation and definitions used to describe these three approaches.

(*a*) In the *traditional* case-control study, controls are sampled from subjects that remain at risk at the end of the risk period. This approach is the case-non-case design described in classic textbooks.[10 11] According to Miettinen, it corresponds to "cumulative incidence" sampling of controls.[8] The traditional odds ratio (OR$_T$) is computed as:

$$OR_T = C_E/C_U \times (N_U - C_U)/N_E - C_E)$$

where C_E and C_U are the number of exposed and unexposed cases; and N_E and N_U are the number of exposed and unexposed subjects in the base.

(*b*) In the *concurrent* case-control study,

Table 1 Notations and definitions

Label	Definition
C_E	Number of exposed incident cases
C_U	Number of unexposed incident cases
N_E	Number of exposed subjects in the base
N_U	Number of unexposed subjects in the base
T_E	Average duration of follow up of exposed subjects
T_U	Average duration of follow up of unexposed subjects
PY_E	Person-years from exposed subjects ($PY_E = N_E \times T_E$)
PY_U	Person-years from unexposed subjects ($PY_U = N_U \times T_U$)
$Risk_E$	Proportion of the exposed population in the base who become a case during the risk period
$Risk_U$	Proportion of the unexposed population in the base who become a case during the risk period
$Rate_E$	Number of new cases per person-years during the risk period among the exposed population in the base
$Rate_U$	Number of new cases per person-years during the risk period among the unexposed population in the base
RR	Relative risk = $Risk_E/Risk_U$
RIR	Relative incidence rate = $Rate_E/Rate_U$
Dt	Duration of risk period
RR_{CIR}	Relative risk = cumulative incidence rate (CIR) ratio = 1-exponential ($e^{-(Rate_E \times Dt)}$)/1-exponential ($e^{-(Rate_U \times Dt)}$)
RIR_{PH}	Relative incidence rate based upon the proportional hazard model

subjects in the population at risk are eligible as controls at many points in time within the risk period, given that they are disease free at the time of selection. They may also be sampled later as cases if they develop the disease. As controls are selected from all subjects in the base still free of disease at the time of occurrence of the case in question, the number of available controls for each incident case is a function of the duration of follow up. It is obtained by multiplying the number of subjects at risk in the base (N) times the average duration of follow up (T), which is equivalent to computing person-years. These can be expressed, for the exposed subjects, as $N_E \times T_E = PY_E$, and for the unexposed subjects as $N_U \times T_U = PY_U$. Thus, it is as if controls were counted as person-years rather than people.[78] Miettinen refers to this approach as "incidence density" sampling.[7] The OR_C is defined as

$$OR_C = C_E/C_U \times PY_U/PY_E$$

where PY_U and PY_E are sums of person-years at risk for the unexposed and exposed groups, respectively. The OR_C is strictly speaking a rate ratio rather than an OR.

(c) In the *case-base* (or case-cohort) study, controls are sampled from the base, regardless of their disease state at the time of occurrence of a new case.[12] The case-base odds ratio (OR_{CB}) is computed as follows:

$$OR_{CB} = C_E/C_U \times N_U/N_E$$

The OR_{CB} is strictly speaking an RR.

The scientific literature is very specific about which ORs obtained under these three control selection schemes should be used to estimate the RIR and RR. When the disease is rare—that is, when the risk of disease is lower than about 10% over the risk period[257]—the values of all risk ratios and ORs will be similar,[13] and all control sampling schemes will yield similar estimates of effect. When the disease is not rare but both exposure and incidence are stable, the OR_C estimates the relative incidence rate (RIR), OR_{CB} estimates the RR, and the OR_T overestimates both the RR and RIR. Unstable incidence or exposure during the risk period theoretically act on all these relations but unstable exposure is not a practical concern unless, as shown in several examples by Greenland and Thomas,[2] the prevalence of exposure decreases substantially, probably more than 50%, over the risk period. Instability of RIR during the risk period creates the problem of the interpretation of any of the summary measures of effect.

In this paper, we used data from the South Wales nickel refinery workers (SWNRW)[14 15] to evaluate the biases associated with departures from the assumptions of rare disease, stability of exposure, incidence, and RIR over time. These effects are investigated within the context of each type of case-control design.

Materials and methods

The principal cohort study of the SWNRW was first reported in 1970[16] and extended in 1977.[17] Further analyses have been published by Peto *et al*,[18] Kaldor *et al*,[15] and Breslow and Day.[14] In the present study we use the data set published in appendix VII of Breslow and Day.[14] A detailed description of the study can be found elsewhere.[14 18] Briefly, the SWNRW study was conducted to determine whether nickel production by decomposition of gaseous nickel compounds is associated with increased risk of cancers of the nasal sinus and lung. The study population was selected from one nickel factory in South Wales. Information on employment was obtained from inspecting factory pay sheets for the first week in April of the years 1929, 1934, 1939, 1944, and 1949.

Figure 1 shows the study design. Our analysis is based on 679 men who had been continuously employed for at least five years, who were first employed before 1925, and who were still employed in 1934 or later. The risk period extended from 1 April 1934 to 31 December 1981. Among the 679 study subjects, 193 developed lung or nasal cancer, and 42 of the remaining subjects were lost to follow up before the end of the risk period.

The exposure period lasted from 1901 to 1934. Men first employed after 1925 were not retained because of evidence that exposure was considerably reduced after 1925 due to changes in the production of nickel. Exposure to nickel was assessed in terms of the number of years worked in jobs at high risk of exposure to nickel before the start of the risk period. In earlier reports there was a category with 0 years of exposure to nickel,[14] but we
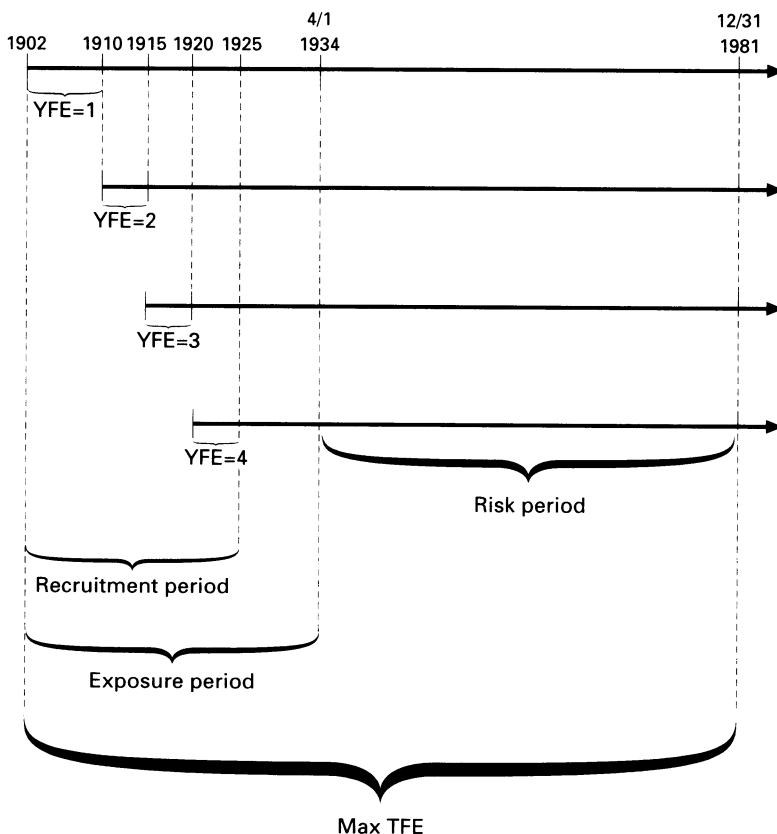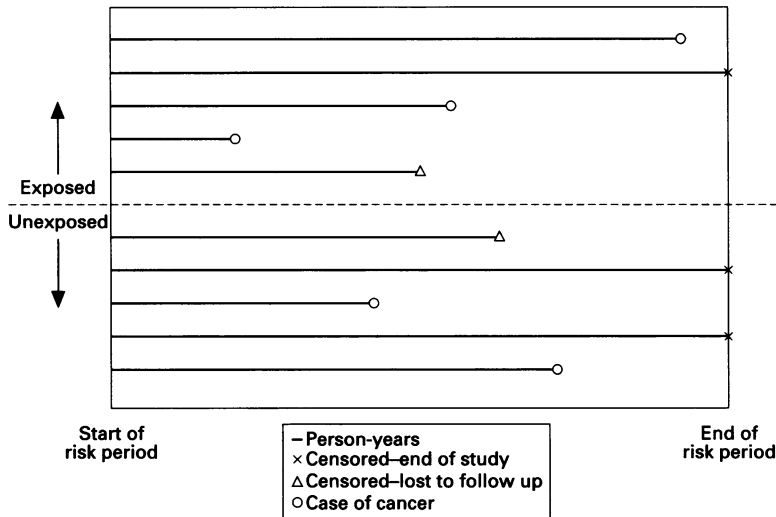

Figure 1   Recruitment, exposure, and risk periods of four subcohorts by year of first employment (YFE) within the south Wales nickel refinery worker study. TFE = time since first employment.

$$OR_{Traditional} = \frac{3_o}{2_o} \cdot \frac{2_x}{1_x}$$

$$OR_{Concurrent} = \frac{3_o}{2_o} \cdot \frac{\text{Sum of person-years in unexposed}}{\text{Sum of person-years in exposed}}$$

$$OR_{Case-base} = \frac{3_o}{2_o} \cdot \frac{(2_x + 1_\Delta + 2_o)}{(1_x + 1_\Delta + 3_o)}$$

Figure 2   Methods for calculating traditional, concurrent, and case-base ORs.

tions entailed the formulas 1–3 already mentioned for these ORs. All three ORs were based upon all cases, whereas different groups of controls were used for each OR: (a) all non-cases for OR_T; (b) all person-years for all subjects for OR_C; and (c) all subjects for OR_{CB}. Figure 2 shows a line plot of person-years and corresponding calculations for the three ORs under a hypothetical example to illustrate their computations. Asymptotic 95% CIs for the SWNRW data were based on previously reported methods.[9 19 20]

We also estimated the RIR (CIs) with the proportional hazards model (RIR_{PH})[21] and the RR with formulas in table 1 for RR_{CIR}. The estimates of the RIR_{PH} were provided by PROC PHREG in SAS.[22] These estimates of the RIR_{PH} and RR_{CIR} were used as standards of comparison for the OR estimates.

To assess the influence of taking a random sample of controls (as is usually done in case-control studies) v taking all available controls, each type of OR and its approximate 95% CI were computed twice: with all available controls that belonged to one of the three groups of controls; and with a random sampling approach with estimates computed at each of 1000 iterations. For each iteration, 50% of the controls were sampled randomly and the median of the 1000 ORs or CIs were presented. A sampling fraction of 50% was chosen because lower sampling fractions yielded cell counts that were insufficient to justify the asymptotic CI approximations based upon the delta method.[19]

used broader categories (less than three years or three years or more) to obtain reliable estimates of cancer rates for each level of exposure.

The cohorts of interest were defined for the analysis on the basis of year of first employment (YFE): 1902–9, 1910–14, 1915–19, and 1920–24. Rates of respiratory cancer varied substantially among these cohorts.[14] The table in the appendix shows the frequencies of subjects and person-years for each combination of cancer outcome, exposure, and year of first employment.

Figure 1 shows that person-years at risk were calculated for the period 1934–81. For purposes of assessing the stability assumptions, time since first employyment (TFE) was grouped into <20, 20–29, and ≥30 years. Finally, the small number of lung and nasal cancer cases led us to pool the two outcomes in a single category of respiratory cancers.

Estimates of the following ORs specific for YFE were obtained for the SWNRW data: (a) traditional OR_T, (b) unmatched concurrent OR_C, and (c) case-base OR_{CB}. The computa-

Results

Table 2 shows that risks of respiratory cancers were greater than 10%—that is, they were not rare—over the risk period in the unexposed groups of the four YFE cohorts. Table 3 shows the variation in the incidence and RIR across the risk period for each of the YFE cohorts. Table 4 shows the percentages exposed to nickel at the beginning and end of each risk period defined by YFE. Tables 3 and 4 are summarised in table 5. They show that the four YFE cohorts have various combinations of departures from the stability assumptions of exposure, incidence, or RIR involved in the estimation of the RR and RIR by the OR. In the 1902–9 YFE cohort, incidence, RIRs, and exposures are stable across the two last TFE (no workers in this cohort had <20 years of employment). The 1910–14 YFE cohort is stable for incidence, but RIR

Table 2   Risk of respiratory cancers over the full risk period by year of first employment in unexposed workers of the South Wales refinery

Year of first employment	Risk$_U$ (%)
1902–9	18·2
1910–14	25·7
1915–19	14·7
1920–24	22·5

Risk$_U$ = risk of respiratory cancers in unexposed workers, computed with the denominator of the RR_{CIR} formula given in table 1.

Table 3   Incidence of respiratory cancers/1000 person-years in unexposed workers (IR_U) of the South Wales refinery and corresponding relative incidence rate (RIR_{PH})

| Time since first employment | Year of first employment | | | | | | | |
| | 1902–9 | | 1910–14 | | 1915–19 | | 1920–24 | |
	IR_U	RIR_{PH}	IR_U	RIR_{PH}	IR_U	RIR_{PH}	IR_U	RIR_{PH}
0–19·9	NA	NA	NA	NA	0·6	2·8	0·6	1·8
20–29·9	9·9	2·2	6·0	3·1	3·5	4·7	7·5	4·2
≥30	10·3	2·0	15·0	3·2	7·2	3·6	11·9	1·4

NA = no workers fall into these categories.

Table 4   *Workers exposed to nickel at the beginning and end of the risk period, by year of first employment*

Year of first employment	Risk period	Exposed workers (%)
1902–9:	Start	63·6
	End	61·6
1910–14:	Start	57·1
	End	54·5
1915–19:	Start	30·6
	End	26·9
1920–24:	Start	14·8
	End	9·4

Table 5   *Incidence and stability of exposure of respiratory cancers during the risk period in the four cohorts defined by year of first employment*

Year of first employment	Incidence*	RIR*	Exposure†
1902–9	Stable	Stable	Stable
1910–14	Unstable	Stable	Stable
1915–19	Unstable	Unstable	Stable
1920–24	Unstable	Unstable	Unstable

*From table 3; †from table 4.

and exposure vary. In the 1915–19 YFE cohort, incidences, and RIRs show considerable variation by TFE, whereas exposure is moderately stable. The 1920–24 YFE cohort is unstable with respect to all three variables.

In table 6, the full sampling (denoted as "All") estimates of the ORs (95% CIs) can be computed from formulas 1–3 from the introduction and the data given in the appendix. The 50% and full sampling point estimates of the ORs do not differ by more than 2%, regardless of the case-control design. The subsampling approximate CIs are wider than the full sampling CIs for both the traditional and case-based estimates, whereas the subsampling and full sampling CIs are similar for the concurrent estimates.

Table 6 also shows that: (*a*) in the 1902–9 YFE cohort, the OR_C (2·1) equals the RIR_{PH} (2·1), whereas the OR_{CB} (1·9) slightly overestimates the RR_{CIR} (1·8); (*b*) in the 1910–14 cohort, the OR_C (2·8) underestimates the RIR_{PH} (3·2) by about 15% and the OR_{CB} (2·0) underestimates the RR_{CIR} (2·3) by 20%; (*c*) in the 1915–19 cohort, the OR_C (3·6) equals the RR_{PH} (3·6) and the OR_{CB} (2·7) underestimates RR_{CIR} (3·1) by 15%; and (*d*) in the 1920–4 cohort, OR_C (2·2) underestimates RIR_{PH} (2·2) and OR_{CB} (1·5) underestimates RR_{CIR} (1·8) by about 20%. Finally, the relation $OR_T > OR_C > OR_{CB}$ holds for each YFE cohort.

Discussion

The influence of control sampling schemes on effect estimation has been previously discussed on a theoretical basis, with true measures of relative risk and relative incidence rate compared with various OR estimators— for example, the true RIR *v* its estimation by the concurrent OR_C.[35] In contrast, we have compared empirical estimates of both sets of measures of effect with an example in which the rarity of disease assumption could not be invoked, as risks in the unexposed population over the full risk period varied from 15% to 26% for the four cohorts defined by year of first employment. Hence, conditions other than the rare disease assumption had to be satisfied for an unbiased estimation of RR with OR.

The results are compatible with the theory when stability assumptions of exposure, incidence, and RIR are met. Our empirical evaluation showed, as expected from the theory, that, when exposure, incidence, and RIR are stable during the risk period, OR_C equals RIR_{PH}, and OR_T overestimates RR_{CIR} and RIR_{PH}. On the other hand, our analysis did not show that OR_{CB} estimated exactly RR_{CIR} in the stable cohort. This could be due possibly to differential losses to follow up between exposed and unexposed workers in this cohort (6% *v* 0%, respectively).

The present results suggest that the theory is relatively tolerant of large departures from the stability assumptions of exposure, of incidence, and of RIR. There are no wide differences (>25%) from the predicted relations between the RR_{PH} and RIR_{CIR} and their respective OR estimates even in the cohorts for which incidence rates were clearly unstable. The largest differences between OR_C and RIR_{PH} and between OR_{CB} and RR_{CIR} were found in the cohort of subjects first employed between 1920 and 1924 for whom none of the stability assumptions were met.

As well as the comparisons of the OR and RR estimates, random sampling of 50% of controls was compared with the use of all available controls. Absence of differences suggest that comparisons of the various ORs should not be affected by whether all or a fraction of the controls are used, and that the issue of sampling bias in the interpretation of the findings related to RR_{CIR} and RIR_{PH} estimation is not relevant in our study.

Table 6   *Odds ratios, RR_{CIR}, and RIR_{PH} of respiratory cancers and exposure to nickel in the four fixed cohorts defined by year of first employment: the OR (95% CIs) were computed under 50% random sampling of controls and under sampling of all available controls for three case-control designs*

Measure of effect	Control sampling	Year of first employment				
		1902–9	1910–14	1915–19	1920–24	Total
OR_T:	50%	2·4 (0·9–6·6)	3·1 (1·3–7·0)	4·3 (1·3–14·0)	2·6 (1·0–6·4)	2·7 (1·8–4·0)
	All*	2·4 (1·0–6·0)	3·1 (1·6–6·2)	4·3 (1·5–12·0)	2·6 (1·2–5·5)	2·7 (1·9–3·8)
OR_C:	50%	2·1 (0·9–4·6)	2·8 (1·6–4·8)	3·6 (1·6–8·4)	2·2 (1·3–3·8)	3·0 (2·2–3·9)
	All*	2·1 (0·9–4·6)	2·8 (1·6–4·8)	3·6 (1·6–8·3)	2·2 (1·2–3·8)	3·0 (2·2–3·9)
RIR_{PH}		2·1 (1·0–4·7)	3·2 (1·9–5·6)	3·6 (1·6–8·5)	2·7 (1·5–4·7)	3·3 (2·5–4·4)
OR_{CB}:	50%	1·9 (0·9–4·2)	1·9 (1·2–3·4)	2·7 (1·2–6·3)	1·5 (0·9–2·7)	1·9 (1·4–2·5)
	All*	1·9 (0·9–3·7)	2·0 (1·3–3·1)	2·7 (1·3–5·6)	1·5 (1·0–2·4)	1·9 (1·5–2·4)
RR_{CIR}		1·8 (0·8–4·1)	2·3 (1·5–3·6)	3·1 (1·5–6·2)	1·8 (1·2–2·9)	2·1 (1·7–2·7)

*Can be computed with formulas 1–3 of the introduction and the data given in the appendix. T = Traditional; CB = case–base; C = concurrent; CIR = cumulative incidence rate; PH = proportional hazards.

Appendix　*Frequencies of subjects and person-years by year of first employment, cancer, and exposure*

Measure (label)	Year of first employment				
	1902–9	1910–14	1915–19	1920–24	Total
Exposed cases (C_E)	26	47	12	15	100
Unexpected cases (C_U)	8	18	10	57	93
Exposed base (N_E)	77	93	30	44	244
Unexposed base (N_U)	44	70	68	253	435
Exposed controls* (N'_E-C_E)	48	38	14	16	116
Unexposed controls* (N'_U-C_U)	36	45	50	155	286
Exposed person-years (PY_E)	1230	1413	577	877	4097
Unexposed person-years (PY_U)	782	1505	1723	7241	11251

*Subjects free of disease at the end of the risk period; N' = number of subjects in the base minus subjects lost to follow up.

We did not present age matched analyses for each type of control, because the consequences of matching vary substantially according to the control selection schemes. This tends to blur rather than to illuminate the relation between control sampling schemes and RR estimates. Thus, although it would have been important to match for age for the understanding of the relation between nickel exposure and cancer risk, this was not crucial for the main objective of this paper.

It may be argued that the results of this single study may not be extrapolated to other studies and that simulations may yield better control of factors that influence effect estimation. It is unlikely, however, that our results are related to specific characteristics of the cohort. It is remarkable that exposure had been relatively stable among workers first employed before 1920, which reflects stability of the production process but also of the employees who had been working continuously for many years in the plant. Overall, however, the four subcohorts offered a relatively wide range of variation of stability of exposure, incidence rates, or RIRs to assess the tolerance of the theory to departure from the stability assumptions. As theoretical analyses cannot completely replace empirical evaluation because of the complexity of the studied phenomena, both empirical and theoretical approaches are needed to fully understand the range of validity of a given theory.

The carcinogenic effect of nickel on the respiratory tract has been shown previously.[14-18] It was not the aim of our study to contribute to the understanding of this association. Our results are not completely comparable with those reported by others because we used different categories of exposure and outcome. The combination of nasal sinus and lung cancers resulted in an underestimation of the association between nasal sinus cancer and nickel compounds as this association was much stronger than that for lung cancer. Further underestimation resulted from considering as unexposed workers those who had been exposed for one or two years to nickel compounds. Nevertheless, the results presented in tables 4 and 5 are consistent with the SMRs presented by Kaldor et al[15] for similar categories of exposure.

In conclusion, the four nested case-control

studies performed within the historical South Wales nickel factory cohort yield an empirical illustration that the RIR_{PH} is adequately estimated by the OR_C when controls are identified concurrently with case occurrence throughout the risk period and that the RR_{CIR} is well approximated by the OR_{CB} when controls are a sample of the study base.

1 Schlesselman JJ. *Case-control studies.* New York: Oxford, 1982:52.
2 Greenland S, Thomas DC. On the need for the rare disease assumption in case-control studies. *Am J Epidemiol* 1982;116:547–53.
3 Hogue CJR, Gaylor DW, Schulz KF. Estimators of relative risk for case-control studies. *Am J Epidemiol* 1983; 118:396–407.
4 Smith PG, Rodrigues LC, Fine PEM. Assessment of the protective efficacy of vaccines against common diseases using case-control and cohort studies. *Int J Epidemiol* 1984;13:87–93.
5 Greenland S, Thomas DT, Morgenstern H. The rare-disease assumption revisited: a critique of "estimators of relative risk for case-control studies". *Am J Epidemiol* 1986;124:869–76.
6 Hogue CJR, Gaylor DW, Schulz KF. The case exposure study a further explication and response to a critique. *Am J Epidemiol* 1986;124:877–83.
7 Rodrigues L, Kirkwood BR. Case-control designs in the study of common diseases: updates on the demise of the rare disease assumption and the choice of sampling scheme for controls. *Int J Epidemiol* 1990;19:205–13.
8 Miettinen OS. Estimability and estimation in case-referent studies. *Am J Epidemiol* 1976;103:226–35.
9 Rothman KJ. *Modern epidemiology.* Boston: Little Brown, 1986.
10 Lilienfeld AM, Lilienfeld DE. *Foundations of epidemiology.* New York: Oxford, 1980.
11 MacMahon B, Pugh TF. *Epidemiology: principles and methods.* Boston: Little Brown, 1970.
12 Prentice RL. A case-cohort design for epidemiologic cohort studies and disease prevention trials. *Biometrika* 1986;73:1–11.
13 Cornfield J. A method of estimating comparative rates from clinical data. *J Natl Cancer Inst* 1951;11:1269–75.
14 Breslow NE, Day NE. *Statistical methods in cancer research: the design and analysis of cohort studies. Vol II.* Lyon: IARC, 1987:174–220. (IARC Sci Publ No 53.)
15 Kaldor JM, Peto J, Easton D, Doll R, Hermon C, Morgan L. Models for respiratory cancer in nickel refinery workers. *J Natl Cancer Inst* 1986;77:841–8.
16 Doll R, Morgan LG, Speizer FE. Cancers of the lung and nasal sinuses in nickel workers. *Br J Cancer* 1970;24: 623–32.
17 Doll R, Matthews JD, Morgan LG. Cancers of the lung and nasal sinuses in nickel workers: a reassessment of the period of risk. *Br J Ind Med* 1977;34:102–5.
18 Peto J, Cuckle H, Doll R, Hermon R, Morgan LG. Respiratory cancer mortality of Welsh nickel refinery workers. In: Sunderman FW, ed. *Nickel in the human environment.* Lyon:IARC, 1984. (IARC Sci Publ No 53.)
19 Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic research: principles and quantitative methods.* Belmont, CA: Lifetime Learning, 1982.
20 Lawless JF. *Statistical models and methods for lifetime data.* New York: Wiley, 1982.
21 Prentice RL, Breslow NE. Retrospective studies and failure time models. *Biometrika* 1978;65:153–8.
22 SAS/STAT Software. *The PHREG procedure, preliminary documentation.* Carry NC: SAS Institute, 1991.