ORIGINAL RESEARCH ARTICLE

Drug Safety 2002; 25 (9): 677-687
0114-5916/02/0009-0677/$25.00/0

Adverse Effects of Observational Studies When Examining Adverse Outcomes of Drugs

Case-Control Studies with Low Prevalence of Exposure

Peter C. Austin,[1,2] *Muhammad Mamdani*[1,3] and *Ivan J. Williams*[1,2]

1 Institute for Clinical Evaluative Sciences, North York, Ontario, Canada
2 Department of Public Health Sciences, University of Toronto, Toronto, Ontario, Canada
3 Faculty of Pharmacy, University of Toronto, Toronto, Ontario, Canada

Abstract

Objectives: The case-control study is commonly used to examine adverse drug events, in which prevalence of exposure in the source population is frequently very low. The objective of the current study was to examine the bias inherent in the odds ratio assessing the association between exposure and an adverse outcome when prevalence of exposure in the source population is extremely low.
Design: Monte Carlo simulations examined the effect of sample size, exposure prevalence, and magnitude of the underlying odds ratio on the bias of the estimated risk ratio, and the power to detect a non-zero risk ratio.
Results: Once the underlying odds ratio was at least four, the adverse effects of low prevalence of exposure was minimal. Studies with small sample sizes and low prevalence of exposure, coupled with small to moderate effect sizes, can result in biased estimates of association between exposure and disease status. With a sample size of 200 and an exposure prevalence of 0.5% in the control population, the bias in the estimated odds ratio can be as large as 115%. However, bias becomes negligible as sample size becomes large ($n \geq 2000$), even when prevalence of exposure is very low. Once the expected number of exposed controls is at least eight, the bias in the estimated odds ratio was no more than 5%.
Conclusions: Studies with small sample sizes and low prevalence of exposure, coupled with small to moderate effect sizes can result in biased estimates of association between exposure status and adverse drug effects. However, bias becomes negligible as sample size becomes large.

Case-control studies[1,2] are frequently used to assess adverse effects of drugs infrequently prescribed, and therefore having limited exposure.[3-12] For example, a recent case-control study suggested that the use of products containing phenylpropanolamine (PPA) increased the risk of haemorrhagic stroke in women.[4] Of 750 female controls, 20 (2.7%) were exposed to any agent containing PPA. A secondary analysis examined the effect of exposure to an appetite suppressant containing PPA. Only one woman (0.1%) was exposed to an appetite suppressant containing PPA. The study suggested that among females, the risk of haemorrhagic stroke increased over 16-fold when exposed

to an appetite suppressant containing PPA. These findings received widespread media coverage, and immediate reaction from the US Food and Drug Administration (FDA). The FDA subsequently issued a public health advisory, took steps to remove PPA from all drug products, and requested all drug companies to discontinue marketing products containing PPA.

Peduzzi et al.[13,14] showed that the results of logistic regression and survival models can be biased when the outcome of interest occurs for only a few subjects. This highlights the possibility that case-control studies may potentially result in biased estimates of the underlying risk ratio when the prevalence of exposure is very low in the control population.

The maximum likelihood estimate of the odds ratio is biased.[15,16] However, the bias decreases with increasing sample size. Modified estimators have been suggested that reduce the bias in the estimated odds ratio. Haldane[17] suggested adding 0.5 to each of the four cells in a 2×2 table, eliminating the first-order bias. Jewell[18] suggested adding one to each of the off-diagonal cells in the 2×2 table. Becker[15] compared the performance of the maximum likelihood estimate of the odds ratio to Jewell's estimate. He showed that for 2×2 tables where at least one of the off-diagonal elements takes on the values one, two or three, then Jewell's estimates are far below the midpoint of the confidence interval. Based on this and other results, he recommended continued use of the maximum likelihood estimator. Gart and Zweifel[19] describe the bias inherent in several estimators for the logarithm of the odds of an event, assuming Binomial sampling, and estimators for the estimate of its variance. They showed that the bias of Haldane's estimate of the logarithm of the odds of an event is <10% when the expected number of events is at least one. They also showed that the different estimates of variance all perform poorly when the number of expected number of events (or non-events) is <1.5. Greenland et al.[20] describe how bias due to sparse matched sets can bias the estimate of odds ratio obtained from conditional logistic regression.

Furthermore, it has been shown that for a given odds ratio, and desired statistical power, there is an optimum prevalence of exposure amongst the controls in order to minimise sample size requirements.[21] Similarly, using the asymptotic estimate (Woolf's estimate[2]) of the variance of the logarithm of the odds ratio, one can show that for a given odds ratio and sample size, there is an optimum prevalence of exposure amongst the controls in order to minimise the width of the resultant 95% confidence interval for the exposure odds ratio. Figure 1 describes the effect of prevalence of exposure (in percent) on precision for different scenarios, based upon a sample size of 100 000 cases and 400 000 controls. As exposure prevalence increases and decreases from the optimum level, the study becomes less efficient, resulting in wider confidence intervals, and requiring a larger sample size for the same statistical power. For odds ratios between 0.1 and 10, optimum prevalence of exposure amongst the controls varies between 16 and 84%. Case-control studies where the prevalence of exposure is very low are never optimal with respect to minimising the width of the 95% confidence interval, unless the true odds ratio is extremely large (>500). In other words, unless the true odds ratio is very large, case-control studies with very low prevalence of exposure will result in wider confidence intervals than a case-control study with the same sample size, but with a higher prevalence of exposure. Figure 1 illustrates that those odds ratios that are reciprocals of each other result in curves that are symmetrical (i.e. the figures are mirror images of each other).

The primary objective of the current study is to illustrate the effect of low prevalence of exposure on bias in the estimated odds ratios and on power to detect non-zero odds ratios. These two issues are critical in allowing one to assess the results of case-control studies that assess the association between adverse outcomes and drug exposure, for drugs with limited use in the general population. There are two secondary objectives of the current study.

First, to examine the effect of sparse data on the coverage of confidence intervals computed using the asymptotic, large sample properties of the odds ratio. Second, to examine the effect of sparse data on the asymptotic estimate of the variance of the logarithm of the odds ratio.

Methods

A factorial experimental design using Monte Carlo simulations was employed for the study. We designed the simulation so that disease (case/control) status was related to exposure status. For each

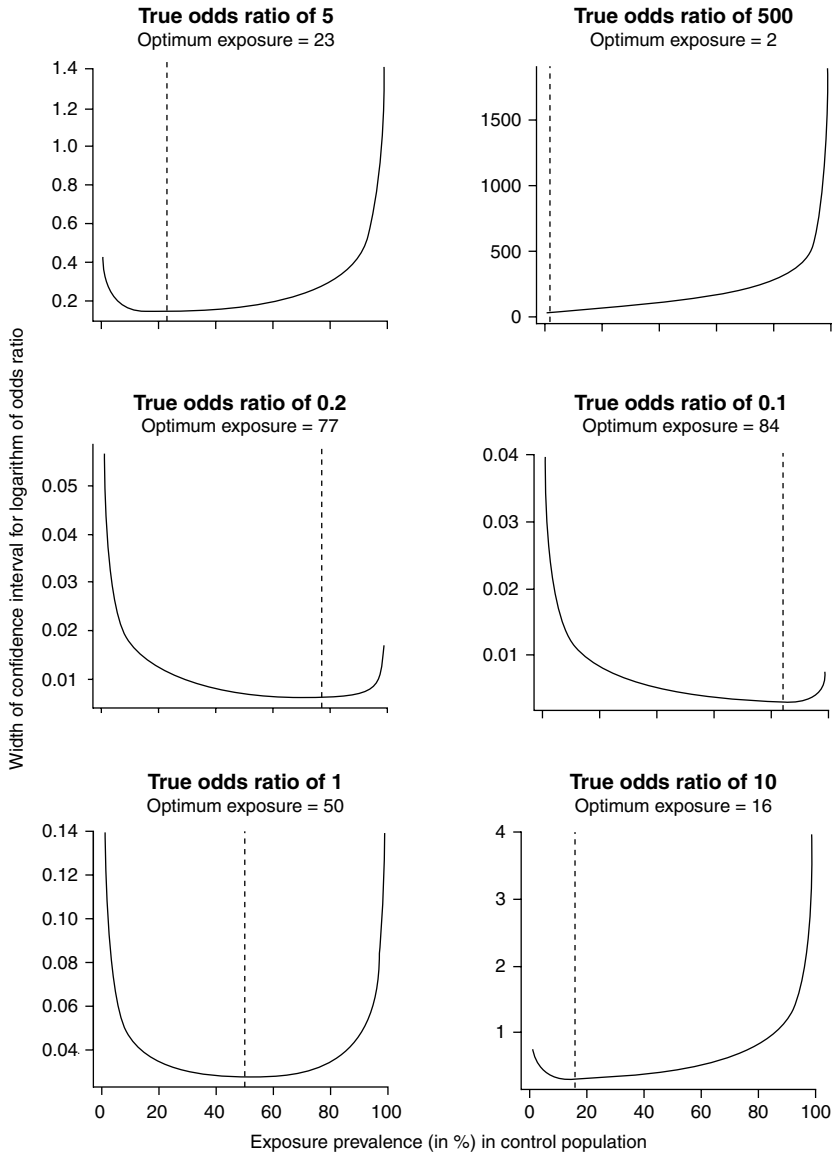

Fig. 1. The effect of prevalence of exposure on precision for different scenarios, based on a sample size of 100 000 cases and 400 000 controls.

scenario, the levels of three different factors were fixed. First, we fixed the number of cases. The number of controls was set to be four times the number of cases, to simulate four to one matching. Second, we fixed the population prevalence of exposure for the controls in the source population. Third, we fixed the odds ratio relating exposure status to disease status. We chose total sample sizes (i.e. number of cases plus number of controls) of 200, 500, 600, 800, 1000, 2000, 5000 and 10 000. The prevalence of exposure amongst the controls in the source population was chosen to be equal to 0.5, 1, 2 and 5%. The odds ratio relating exposure to disease status was chosen to be 1.25, 1.5, 2, and 4.

For a given scenario, the prevalence of exposure in the population of controls was defined to be p_2. We used the fact that the exposure odds ratio for a case-control study is equal to the disease odds ratio for a cohort study. Given the exposure odds ratio, $OR_{exposure}$, the prevalence of exposure amongst the cases was defined to be (equation 1):

$$p_1 = OR_{exposure} \times p_2 / (1 - p_2 + OR_{exposure}\, p_2)$$

Let N_{cases} denote the number of cases. The number of cases with positive exposure was randomly generated from a binomial distribution with parameters p_1 and N_{cases}. The number of exposed controls was generated from a binomial distribution with parameters p_2 and $4\,N_{cases}$. In this sample of controls, from the population of controls, the number of exposed and non-exposed controls was noted. We now have data drawn from the population of case-control studies with desired sample size, exposure prevalence amongst the controls, and exposure odds ratio. Note that in a simulated 2×2 table, the observed odds ratio may not be equal to the underlying exposure odds ratio, since cases and controls were randomly sampled from a population of cases and controls respectively. We have thus introduced an element of stochastic variability. For each scenario, 1000 simulated 2×2 tables were generated.

For each scenario, Haldane's estimate[17] of the odds ratio was computed for each of the 1000 simulated case-control studies. The mean estimated odds ratio was defined to be the exponential of the mean of the logarithms of the estimated odds ratios. The bias in the estimated odds ratio for exposure was defined to be the difference between the mean estimated exposure odds ratio estimated from the simulated datasets and the true odds ratio for exposure from the underlying data generating procedure. The relative bias was defined to be the percentage by which the mean estimated odds ratio deviated from the true odds ratio in the underlying model.

For each simulated dataset, a 95% confidence interval for the estimated odds ratio was calculated. Woolf's estimate of the standard error[2] of the logarithm of the odds ratio, and the asymptotic normality of the logarithm of the odds ratio were used to calculate 95% confidence intervals for the odds ratio. The percentage of confidence intervals that contained the true parameter value was then computed. If the intervals had the advertised level of coverage, then approximately 95% of the intervals should contain the true parameter. For each model fit to simulated data, the significance of the estimated odds ratio (at the $\alpha = 0.05$ level) was determined using Fisher's Exact Test. The percentage of times that the null hypothesis (i.e. the underlying odds ratio is equal to unity) was rejected was determined over the 1000 simulations.

For each scenario, Woolf's estimate of the standard error of the logarithm of the odds ratio was computed. For each scenario, the sample and asymptotic variances for the logarithm of the odds ratio for exposure were compared. The sample variance of the estimated logarithm of the odds ratio was defined to be the variance of the logarithm of the estimated odds ratio over the 1000 simulations. The asymptotic variance of the logarithm of the odds ratio was defined as the mean of Woolf's estimated variance of the logarithm of the odds ratio over the 1000 simulated 2×2 tables. If the ratio of these two quantities was different from one, then the asymptotic estimate of variance may be incorrectly estimating the variability of the odds ratio.

Table I. Mean estimated odds ratio (OR) for exposure and percent bias in estimated OR[a]

Sample size	Prevalence of exposure (in controls)															
	OR for exposure = 1.25				OR for exposure = 1.50				OR for exposure = 2.0				OR for exposure = 4.0			
	0.5%	1.0%	2.0%	5.0%	0.5%	1.0%	2.0%	5.0%	0.5%	1.0%	2.0%	5.0%	0.5%	1.0%	2.0%	5.0%
200	2.69 (115.2)	1.87 (49.6)	1.50 (20.0)	1.29 (3.2)	2.61 (74.0)	2.07 (38.0)	1.60 (6.7)	1.51 (0.7)	2.98 (49.0)	2.41 (20.5)	2.14 (7.0)	1.97 (–1.5)	4.13 (3.3)	4.07 (1.8)	3.95 (–1.3)	4.12 (3.0)
500	1.77 (41.6)	1.38 (10.4)	1.29 (3.2)	1.28 (2.4)	1.99 (32.7)	1.57 (4.7)	1.47 (–2.0)	1.48 (–1.3)	2.34 (17.0)	2.05 (2.5)	2.03 (1.5)	1.98 (–1.0)	3.97 (–0.8)	4.03 (0.8)	3.86 (–3.5)	3.98 (–0.5)
600	1.65 (32.0)	1.31 (4.8)	1.24 (–0.8)	1.24 (–0.8)	1.90 (26.7)	1.51 (0.7)	1.48 (–1.3)	1.49 (–0.7)	2.17 (8.5)	1.96 (–2.0)	1.97 (–1.5)	2.01 (0.5)	4.05 (1.3)	4.01 (0.3)	4.01 (0.3)	4.00 (0.0)
800	1.54 (23.2)	1.27 (1.6)	1.24 (–0.8)	1.23 (–1.6)	1.65 (10.0)	1.47 (–2.0)	1.55 (3.3)	1.51 (0.7)	2.15 (7.5)	1.94 (–3.0)	2.03 (1.5)	2.00 (0.0)	3.84 (–4.0)	3.95 (–1.3)	3.98 (–0.5)	3.97 (–0.8)
1000	1.45 (16.0)	1.27 (1.6)	1.28 (2.4)	1.25 (0.0)	1.62 (8.0)	1.54 (2.7)	1.52 (1.3)	1.49 (–0.7)	2.01 (0.5)	2.02 (1.0)	2.02 (1.0)	2.01 (0.5)	3.87 (–3.3)	4.00 (0.0)	4.09 (2.3)	4.03 (0.8)
2000	1.27 (1.6)	1.23 (–1.6)	1.24 (–0.8)	1.25 (0.0)	1.49 (–0.7)	1.48 (–1.3)	1.49 (–0.7)	1.50 (0.0)	1.97 (–1.5)	1.98 (–1.0)	2.00 (0.0)	2.01 (0.5)	3.99 (–0.3)	4.01 (0.3)	4.06 (1.5)	4.03 (0.8)
5000	1.31 (4.8)	1.26 (0.8)	1.25 (0.0)	1.25 (0.0)	1.47 (–2.0)	1.48 (–1.3)	1.49 (–0.7)	1.49 (–0.7)	2.01 (0.5)	1.98 (–1.0)	2.00 (0.0)	2.00 (0.0)	4.00 (0.0)	4.00 (0.0)	4.00 (0.0)	4.01 (0.3)
10 000	1.23 (–1.6)	1.25 (0.0)	1.26 (0.8)	1.25 (0.0)	1.51 (0.7)	1.49 (–0.7)	1.51 (0.7)	1.51 (0.7)	1.98 (–1.0)	2.00 (0.0)	1.99 (–0.5)	2.00 (0.0)	4.04 (1.0)	4.00 (0.0)	3.98 (–0.5)	4.02 (0.5)

a Each cell contains the mean estimated odds ratio for exposure over the 1000 simulations (percent bias). Numbers are rounded to two decimal places.

The simulations were carried using the S-Plus version 5.1[22] and SAS version 6.12[23] software packages.

Results

The results for bias are presented in table I. The power to detect a non-zero odds ratio are presented in table II. The results comparing the ratio between the mean asymptotic estimate of the variance of the logarithm of the odds ratio to its sampling-based estimate of variability are presented in table III.

Table I outlines three major trends. First, for a given odds ratio and prevalence of exposure, bias tended to increase with decreasing sample size. Second, for a given sample size and odds ratio, bias tended to increase as the prevalence of exposure decreased. Third, for a given sample size and prevalence of exposure, bias in the estimated odds ratio tended to increase as the true underlying odds ratio decreased. If the true odds ratio was four, then the bias was no larger than 4%, regardless of the sample size and prevalence of exposure. When the sample size was 500 and prevalence of exposure was 0.5% (two controls exposed on average), the estimated odds ratio was subject to substantial bias if the true odds ratio was no greater than two. When the sample size was 1000 and the prevalence of exposure was 0.5% (four controls exposed on average) there was moderate bias in the odds ratio.

The expected number of exposed controls was determined for each scenario. For each odds ratio, the percent bias was plotted against the expected number of exposed controls. These results are presented in figure 2 (in order to increase resolution, two additional sample sizes of 700 and 900 were added to the simulation). Once the expected number of exposed controls was at least eight, bias in the estimated odds ratio was at most 5%, regardless of the underlying odds ratio. With an underlying odds ratio of two, once the expected number of exposed controls was at least four, the bias in the estimated odds ratio was minimal.

Table II outlines the effect of low prevalence of exposure and sample size on the power to detect a non-zero odds ratio. Three trends are apparent in table II. First, for a given prevalence of exposure

and odds ratio, power increased with increasing sample size. Second, for a given sample size and prevalence of exposure, power increased as the underlying odds ratio increased. Third, for a given odds ratio and sample size, power increased as the prevalence of exposure increased. When the prevalence of exposure was low (≤5%) and the sample size was 1000 or less, the power to detect a non-zero odds ratio of no more than two was low. When the sample size was large (n = 10 000), the effect of very low prevalence of exposure was to substantially reduce power if the true odds ratio was no more than two. When the true odds ratio was equal to four, the effect of low prevalence of exposure on power was greatly diminished.

For each scenario, the expected number of exposed controls was determined. For each odds ratio, the power to detect a non-zero odds ratio was plotted against the expected number of exposed controls. The results are displayed in figure 3 (in order to increase resolution, two additional sample sizes of 700 and 900 were added). With a true odds ratio of four, one requires only that the expected number of exposed controls be at least 10, to have a power of 80%.

In most of the scenarios examined, at least 95% of the 95% confidence intervals covered the true parameter value. The confidence intervals tended to be conservative, with coverage greater than the advertised nominal level of 95%.

Table III outlines the effect of prevalence of exposure and sample size on the quality of the asymptotic estimate of the variance of the logarithm of the odds ratio. Two trends are apparent. First, for a given sample size and exposure odds ratio, the ratio between the two variance estimates diverged from unity as the exposure prevalence decreased. Second, for a given exposure prevalence and odds ratio, the ratio diverged from unity as the sample size decreased. In general, if the expected number of exposed controls was at least 20 then the two estimates of variance were within 10% of each other. If the true odds ratio is four, then the asymptotic variance estimate closely approximated the sampling based estimate of variance once the prevalence of exposure was at least two percent, regardless of sample size.

Discussion

The case-control design is an important and popular epidemiological tool. Case-control studies allow one to efficiently assess the relationship between disease status and exposure to agents hypothesised to cause the disease. Case-control studies have been responsible for many important findings such as the positive association between smoking and lung cancer[24] and the primary prevention of coronary heart disease in patients with hypertension with the use of β-blocker drug therapy.[25] However, the ease and relatively low cost

Table II. Power to detect a non-zero odds ratio (OR)[a]

Sample size	Prevalence of exposure (in controls)															
	OR for exposure = 1.25				OR for exposure = 1.50				OR for exposure = 2.0				OR for exposure = 4.0			
	0.5%	1.0%	2.0%	5.0%	0.5%	1.0%	2.0%	5.0%	0.5%	1.0%	2.0%	5.0%	0.5%	1.0%	2%	5%
200	1.7	2.5	2.6	4.4	2.7	3.9	4.2	9.1	5.6	6.9	8.1	19.2	10.8	21.1	32.7	65.2
500	3.5	3.3	6.0	9.5	3.8	5.2	9.6	12.8	8.1	12.4	19.7	37.0	24.6	44.9	69.3	95.6
600	2.9	3.6	6.3	8.5	4.0	5.1	9.9	17.0	7.9	13.4	23.6	43.3	25.9	49.7	77.2	97.2
800	2.9	4.5	7.3	9.2	4.3	7.3	10.7	21.5	7.1	18.9	28.5	53.6	37.6	63.0	87.7	99.4
1000	3.6	4.6	8.4	10.3	3.8	6.5	11.2	20.8	11.6	21.4	32.7	62.2	46.0	71.9	90.6	99.8
2000	5.5	8.9	8.1	16.9	7.4	12.8	20.1	41.5	21.3	33.0	56.8	89.3	71.1	94.2	99.4	100.0
5000	8.3	10.2	15.1	30.3	16.3	27.7	43.5	79.1	40.6	67.8	90.2	100.0	97.2	99.8	100.0	100.0
10 000	10.3	17.5	26.1	53.4	26.4	44.2	71.4	97.7	68.2	91.2	99.5	100.0	100.0	100.0	100.0	100.0

a Each cell is the percentage of times that the null hypothesis of a zero OR for exposure was rejected.

Table III. Ratio of variance of estimated log odds for exposure and mean asymptotic estimate of the variance of log odds for exposure

Sample size	Prevalence of exposure (in controls)															
	OR for exposure = 1.25				OR for exposure = 1.5				OR for exposure = 2.0				OR for exposure = 4.0			
	0.5%	1.0%	2.0%	5.0%	0.5%	1.0%	2.0%	5.0%	0.5%	1.0%	2.0%	5.0%	0.5%	1.0%	2.0%	5.0%
200	3.85	2.54	1.65	1.16	3.73	2.16	1.49	1.07	3.21	1.99	1.29	1.06	2.47	1.49	1.09	1.06
500	2.02	1.43	1.21	1.01	2.11	1.41	1.04	1.02	1.73	1.28	1.05	1.02	1.42	1.04	1.03	1.04
600	2.05	1.42	1.13	0.93	1.87	1.32	1.08	1.01	1.70	1.12	0.93	0.99	1.16	1.07	0.99	1.04
800	1.59	1.24	1.03	1.07	1.52	1.25	1.07	0.96	1.35	1.09	1.04	0.97	1.16	0.98	1.06	1.04
1000	1.49	1.17	0.95	1.10	1.39	1.06	1.02	0.98	1.22	1.03	1.04	1.06	1.07	1.04	0.93	1.00
2000	1.11	0.95	1.04	1.01	1.08	0.96	0.97	1.02	1.00	1.04	0.99	1.00	1.02	0.97	0.95	0.98
5000	1.01	1.05	0.99	0.98	0.99	0.94	0.96	1.04	1.04	0.88	0.97	0.99	0.98	0.93	1.04	1.00
10 000	0.99	0.95	1.12	1.05	1.05	0.94	1.00	0.96	1.10	1.08	1.04	0.94	0.87	1.09	0.96	1.04

OR = odds ratio.

with which case-control studies can be carried out can result in studies with poor designs.[1] The association between medication use and adverse events are commonly examined using case-control studies. Frequently, such studies have very low prevalence of exposure in the control population. This low prevalence of exposure in the control population can lead one to question how to interpret the study results.

We showed that in the presence of very low prevalence of exposure, coupled with a small effect size and low sample size, the bias in the estimated odds ratio was approximately 100%. In such a setting, the estimated odds ratio systematically overstates the true effect size.

There are many examples in the pharmacoepidemiological literature of case-control studies with small sample sizes and low prevalence of exposure.[3-12] A recent case-control study by Hu et al.[26] with moderately low exposure rates and sample sizes highlights the need for recognising these limitations, as they found no association between arsenic exposure and lung cancer, a relationship which has been reasonably well established in numerous previous studies with greater exposure rates.[27,28] Hu et al.[26] suggest cautious interpretation of negative findings when exposure rates are particularly low. Our current research illustrated that the study may have had insufficient power to detect an odds ratio different from one, due to the small sample size and low prevalence of exposure.

Several drug studies, however, may not have recognised such biases, leading to potentially inaccurate conclusions. Table IV summarises several case-control studies that examine the association between drug use and adverse outcomes. In each of these studies, the prevalence of exposure amongst the controls was very low for at least one of the exposures studied. In table IV we have chosen to focus on only one exposure per study, although a given study may have reported results for multiple exposures. For example, one study[3] examined the association between medication use and the risk of Stevens-Johnson syndrome or toxic epidermal necrolysis. This study consisted of 245 cases and 1147 controls. Among the medications studied were 23 different medications or classes of medication for which the prevalence of exposure amongst the controls was ≤1%. There were several drugs to which less than five controls were exposed. Hence, the estimate of the risk ratio in the source population may be biased upwards due to the low number of exposed controls. A second study found that for women, the odds ratio was 16.58 for the association between appetite suppressant containing PPA and the risk of a haemorrhagic stroke.[4] However, out of 750 female controls, only one (0.1%) was exposed to an appetite suppressant containing PPA. A third study found

that the odds ratio was 23.1 for the association between the use of appetite-suppressant drugs for a total of more than 3 months, and the risk of primary pulmonary hypertension.[5] In this study, out of 355 controls, only five (1.4%) used appetite-suppressant drugs for a total of more than three months. A fourth study found that the odds ratio was 24.1 for the association between low potency antipsychotic drugs and venous thromboembolism.[6] In this study, out of 168 controls, only two (1%) were exposed to low potency antipsychotic drugs. Those studies described in table IV that described their methods, used either maximum likelihood estimation or exact methods, whereas our study used Haldane's estimator of the odds ratio. Haldane's estimator is subject to less bias than the maximum likelihood estimator is. Hence, the above studies are potentially subject to greater bias than we demonstrated in our simulations. The studies summarised in table IV were all able to demonstrate a statistically significant odds ratio, despite low prevalence of exposure in the control population.

Of the studies described in table IV, only one described the sample size calculations upon which the study was based. The study examining the association between PPA and haemorrhagic stroke expected a true odds ratio of five for haemorrhagic stroke, and a prevalence of exposure of 0.5% in the source population. Based upon our findings, one would expect the estimated odds ratio to be subject to only minimal bias due to the large expected underlying odds ratio. Those studies that did not report the hypothesised underlying odds ratio do not allow us to accurately estimate the degree of bias inherent in the estimated odds ratio. However, the largest bias that we describe in table I is of the order of 100%. The risk ratios presented in table IV are all at least 9.2 (some as large as 172). Hence, even in the presence of bias, it is very likely that there is a large risk ratio in the source population.

A recent study identified an increased risk of gastrointestinal bleeding with the use of certain types of antidepressants but not others.[12] However, the prevalence of use of the group of antidepressants that failed to demonstrate an association with gastrointestinal bleeding was <0.5% in both the case and control groups. Coupled with a moderate sample size of nearly 12 000 patients, the logistic regression analysis used may have had insufficient power, due to the very low prevalence of

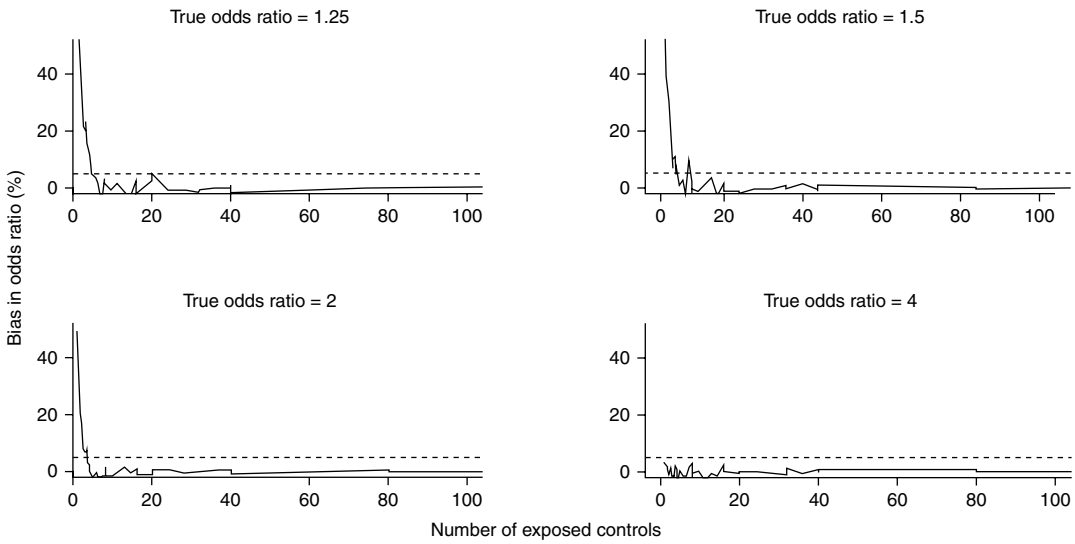

Fig. 2. Bias in estimated odds ratio.

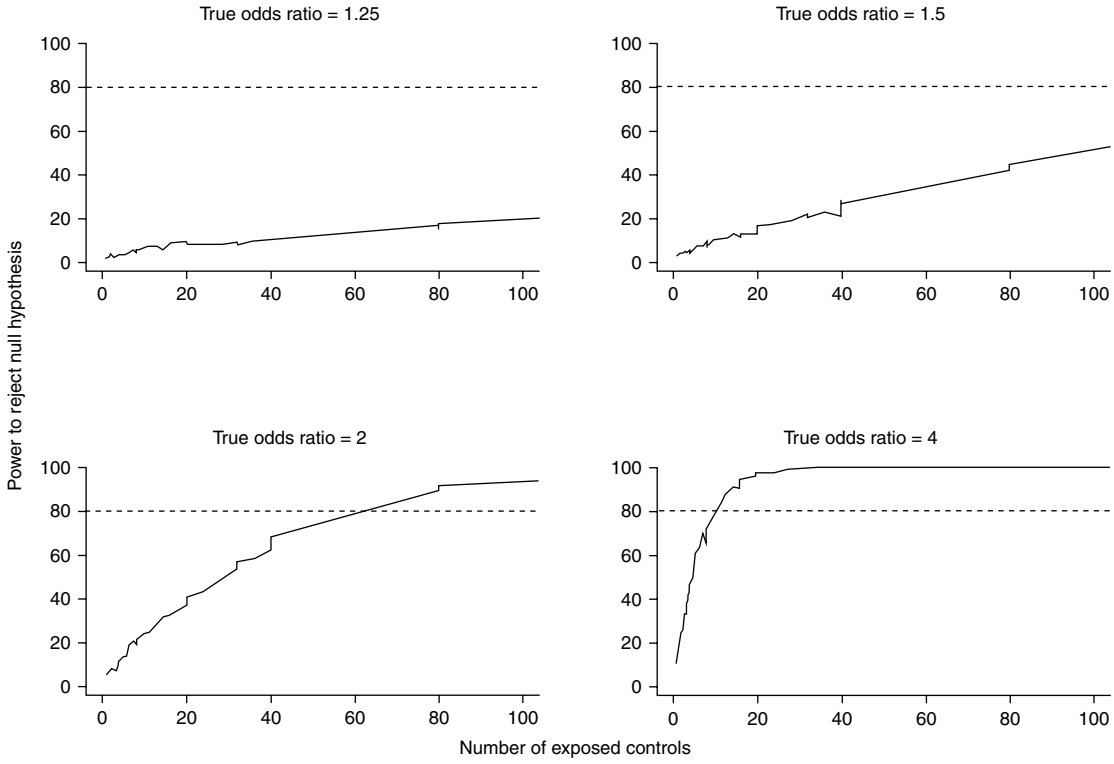

Fig. 3. Power to reject null hypothesis.

exposure. The authors of the above study did not present any formal sample size calculations, nor did they indicate the strength of the hypothesised association. Our findings indicate, however, that this may not have been a significant problem for other studies that also had extremely low prevalence of exposure rates but compensated for this limitation with much larger sample sizes.[29]

There are certain limitations to the current study. The Monte Carlo simulations were designed to assess bias in the simplest case-control design: the unmatched case-control study. There are many variations and refinements of the case-control study.[1] Many case-control studies (including those described in table IV) present adjusted odds ratio, whereas our study focuses on the unadjusted, or crude, odds ratio. It is likely that similar biases would be observed in the estimated adjusted odds ratio. After stratification by other covariates (age, gender, etc.), the expected cell sizes would be much smaller. For instance, Greenland et al.[20] describe the bias that can arise when conditional logistic regression models are applied to sparse matched data. We have examined only scenarios with odds ratios greater than one. However, the bias away from the null would still be observed if one examined odds ratios less than one, since one could examine the reciprocal of the odds ratio (i.e. results for an odds ratio of 0.25 would be directly related to the results for an odds ratio of 4.0). Two different types of bias are possible in epidemiological studies. The first is statistical bias in the performance of the statistical tests and measures of association that are used to report results. The second is bias in the study-design. This comprises sources of bias such as recall error in assessing

Table IV. Case-control studies examining adverse drug reactions

Citation	Exposure	Outcome	Cases exposed	Controls exposed	Odds ratio	Significant ($p < 0.05$)
Kernan et al.[4]	Appetite suppressant containing phenylpropanolamine	Hemorrhagic stroke in women	6/383	1/750	16.58[a]	Yes
IAAA Study[7]	Antithyroid drugs	Aplastic anaemia	4/135	5/2145	9.2[a]	Yes
Rawson et al.[8]	Tocainide (between 35 and 210 days before hospitalisation)	Aplastic anaemia	1/23	1/2726	121.1	Yes
Kelly et al.[9]	Procainamide	Agranulocytosis and aplastic anaemia	7/270	1/1870	50	Yes
Roujeau et al.[3]	Sulfonamides	Stevens-Johnson syndrome	32/245	1/1147	172	Yes
Abenhaim et al.[5]	Appetite suppressant with duration of use >3 months	Pulmonary hypertension	18/95	5/355	23.1[a]	Yes
Zornberg and Jick[6]	Low potency antipsychotic drugs	Idiopathic venous thromboembolism	7/42	2/168	24.1[a]	Yes
Pastuszak et al.[10]	Misoprostol during pregnancy	Mobius' syndrome in infants	47/96	3/96	29.7	Yes

a Adjusted odds ratio.

IAAA = International Agranulocytosis and Aplastic Anaemia.

exposure status, poorly chosen controls, etc. The current study has focussed on the bias inherent in the odds ratio as a measure of association is case-control studies with low prevalence of exposure.

As a secondary objective, we chose to examine the properties of confidence intervals calculated using asymptotic, large sample properties of the odds ratio. Methods have been developed for computing exact confidence intervals that are appropriate for sparse data. The exact confidence intervals described by Thomas[30] are designed to be conservative, with at least 95% coverage. However, the exact method described by Thomas, as implemented in SAS,[23] is only applicable when no zero cells are observed, and all cells contain an integer number of observations. Given that many of our simulated 2×2 tables would have empty cells, Thomas' method (as implemented in SAS) would have been of limited utility. Exact confidence intervals can be computed for 2×2 tables with empty cells in Stata, version 7.[31] We have shown that Normal-theory methods tend to behave similarly to exact methods, in that Normal-theory methods, even when used inappropriately, tend to produce conservative confidence intervals.

Low prevalence of exposure is optimal in certain occasions. However, this occurs in situations in which the true odds ratio is much larger than odds ratios that are typically encountered. Hence, in general, low prevalence of exposure will be non-optimal. We have chosen to examine odds ratios representing small to moderate effect sizes.

Conclusion

The current study has illustrated that bias in the estimated odds ratio arising from a case-control study increases as prevalence of exposure decreases. Similarly, bias increases with decreasing sample size. In particular, case-control studies with a small sample size, and very low prevalence of exposure can be subject to moderate bias in the estimated exposure effect. Once the true underlying odds ratio is at least four, the effect of low prevalence of exposure is minimal. Confidence intervals constructed using the large sample Normal approximation to the distribution of the logarithm of the odds ratio, tended to have conservative cov-

erage. Furthermore, studies that have a small sample size and low prevalence of exposure may have insufficient power to detect non-zero odds ratios. In light of this research, the epidemiological community needs to carefully assess the results from small studies with a low number of exposed subjects that demonstrate a small to moderate risk ratio. Furthermore, investigators are encouraged to include the assumptions that informed the sample size calculations in the study methods.

Acknowledgements

The study was funded by the Institute for Clinical Evaluative Sciences. The funding arrangement allowed the authors freedom to design, analyse, and publish the results.

References

1. Rothman KJ, Greenland S. Modern epidemiology. New York: Lippincott Williams & Wilkins, 1998
2. Breslow NE, Day NE. Statistical methods in cancer research. vol 1: the analysis of case-control studies. Lyon: IARC, 1980
3. Roujeau JC, Kelly JP, Naldi L, et al. Medication use and the risk of Stevens-Johnson syndrome or toxic epidermal necrolysis. N Engl J Med 1995; 333: 1600-7
4. Kernan WN, Viscoli CM, Brass LM, et al. Phenylpropanolamine and the risk of hemorrhagic stroke. N Engl J Med 2000; 343: 1826-32
5. Abenhaim L, Moride Y, Brenot F, et al. Appetite-suppressant drugs and the risk of primary pulmonary hypertension. N Engl J Med 1996; 335: 609-16
6. Zornberg GL, Jick H. Antipsychotic drug use and the risk of first-time idiopathic venous thromboembolism: a case-control study. Lancet 2000; 356: 1219-23
7. International Agranulocytosis and Aplastic Anaemia Study. Risk of agranulocytosis and aplastic anaemia in relation to use of antithyroid drugs. BMJ 1988; 297: 262-5
8. Rawson NSB, Rutledge Harding S, Malcolm E, et al. Hospitalizations for aplastic anemia and agranulocytosis in Saskatchewan: incidence and associations with antecedent prescription drug use. J Clin Epidemiol 1998; 51: 1343-55
9. Kelly JP, Kaufman DW, Shapiro S. Risks of agranulocytosis and aplastic anemia in relation to the use of cardiovascular drugs: the International Agranulocytosis and Aplastic Anemia Study. Clin Pharmacol Ther 1991; 49: 330-41
10. Pastuszak AL, Schuler L, Speck-Martins CE, et al. Use of misoprostol during pregnancy and mobius' syndrome in infants. N Engl J Med 1998; 338: 1881-5
11. Jick H. The discovery of drug-induced illness. N Engl J Med 1977; 296: 481-5
12. de Abajo FJ, Rodriguez LAG, Montero D. Association between selective serotonin reuptake inhibitors and upper gastrointestinal bleeding: population-based case-control study. BMJ 1999; 319: 1106-9
13. Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996; 49: 1373-9
14. Peduzzi P, Concato J, Feinstein AR, et al. Importance of events per independent variable in proportional hazards regression analysis II: accuracy and precision of regression estimates. J Clin Epidemiol 1995; 48: 1503-10
15. Becker S. A comparison of maximum likelihood and jewell's estimators of the odds ratio and relative risk in single 2x2 tables. Stat Med 1989; 8: 987-96
16. Walter SD, Cook RJ. A comparison of several point estimators of the odds ratio in a single 2 x 2 contingency table. Biometrics 1991; 47: 795-811
17. Haldane JBS. The estimation and significance of the logarithm of a ratio of frequencies. Ann Hum Genet 1956; 20: 309-11
18. Jewel NP. On the bias of commonly used measures of association for 2 x 2 tables. Biometrics 1986; 42: 351-8
19. Gart JJ, Zweifel JR. On the bias of various estimators of the logit and its variance with application to quantal bioassay. Biometrika 1967; 54: 181-7
20. Greenland S, Schwartzbaum JA, Finkle WD. Problems due to small samples and sparse data in conditional logistic regression analysis. Am J Epidemiol 2000; 151: 531-9
21. Breslow NE, Day NE. Statistical methods in cancer research, vol II: the analysis of cohort studies. Lyon: IARC, 1987
22. Mathsoft Inc. S-Plus, version 5.1. 1999
23. SAS Institute Inc. SAS, version 6.12. Cary (NC), 1999
24. Doll R, Hill AB. Smoking and carcinoma of the lung: preliminary report. BMJ 1950; II: 739-48
25. Psaty BM, Koepsell TD, LoGerfo JP, et al. Beta-blockers and primary prevention of coronary heart disease in patients with high blood pressure. JAMA 1989; 261: 2087-94
26. Hu S, Hertz-Picciotto I, Siemiatycki J. When to be skeptical of negative studies: pitfalls in evaluating occupational risks using population-based case-control studies. Can J Public Health 1999; 90: 138-42
27. Lee-Feldstein A. A comparison of several measures of exposure to arsenic: matched case-control study of copper smelter employees. Am J Epidemiol 1989; 129: 112-24
28. Jarup L, Pershagen G. Arsenic exposure, smoking, and lung cancer in smelter workers: a case-control study. Am J Epidemiol 1991; 134: 545-51
29. Czeizel AE, Rockenbauer M, Sorenson HT, et al. A population-based case-control teratologic study of furazidine, a nitrofuran-derivative treatment during pregnancy. Clin Nephrol 2000; 53: 257-63
30. Thomas DG. Exact confidence intervals for the odds ratio in a 2x2 table. Appl Stat 1971; 10: 105-10
31. Stata Corporation. Stata 7.0. Texas: College Station, 2001

Correspondence and offprints: Dr *Peter C Austin*, Institute for Clinical Evaluative Sciences, Ontario, 2075 Bayview Avenue, Toronto, G-160, M4N 3M5, Canada.
E-mail: peter.austin@ices.on.ca