

RESEARCH METHODOLOGY SERIES

Overestimation of risk ratios by odds ratios in trials and cohort studies: alternatives to logistic regression

Mirjam J. Knol PhD, Saskia Le Cessie PhD, Ale Algra MD PhD, Jan P. Vandenbroucke MD PhD, Rolf H.H. Groenwold MD PhD

Logistic regression analysis, which estimates odds ratios, is often used to adjust for covariables in cohort studies and randomized controlled trials (RCTs) that study a dichotomous outcome. In case-control studies, the odds ratio is the appropriate effect estimate, and the odds ratio can sometimes be interpreted as a risk ratio or rate ratio depending on the sampling method.¹⁻⁴ However, in cohort studies and RCTs, odds ratios are often interpreted as risk ratios. This is problematic because an odds ratio always overestimates the risk ratio, and this overestimation becomes larger with increasing incidence of the outcome.⁵ There are alternatives for logistic regression to obtain adjusted risk ratios, for example, the approximate adjustment method proposed by Zhang and Yu⁵ and regression models that directly estimate risk ratios (also called “relative risk regression”).⁶⁻⁹ Some of these methods have been compared in simulation studies.^{7,9} The method by Zhang and Yu has been strongly criticized,^{7,10} but regression models that directly estimate risk ratios are rarely applied in practice.

In this paper, we illustrate the difference between risk ratios and odds ratios using clinical examples, and describe the magnitude of the problem in the literature. We also review methods to obtain adjusted risk ratios and evaluate these methods by means of simulations. We conclude with practical details on these methods and recommendations on their application.

Misuse of odds ratios in cohort studies and RCTs

An odds ratio is calculated as the ratio of the odds of the outcome in the patients with the treatment or exposure and the odds of the outcome in the patients without the treatment or exposure. The risk ratio, also referred to as the relative risk, is calculated as the ratio of the risk of the outcome in these two groups. In this article, we illustrate,

by means of two empirical examples, that use of odds ratios in cohort studies and RCTs can lead to misinterpretation of results.

Clinical example 1: cohort study

A cohort study evaluated the relation between changes in marital status of mothers and cannabis use by their children.¹¹ Use of cannabis was reported by 48.6% of the participants at age 21. Table 1 presents the crude and adjusted odds ratios as reported in the paper for one to two changes in maternal marital status and the risk of cannabis use, and for three or more changes in maternal marital status and the risk of cannabis use. We calculated the corresponding crude and adjusted risk ratios (Table 1) based on the data provided in the article. The odds ratios and risk ratios were quite different: a modest increase of the risk by 50% (adjusted risk ratio is 1.5) was observed, whereas the “risk” seemed more than doubled when the odds ratio was interpreted as a risk ratio (adjusted odds ratio is 2.3).

Clinical example 2: RCT

In an RCT, 101 patients with spinal cord compression caused by metastatic cancer were ran-

Competing interests:
Please see end of article.

This article has been peer reviewed.

Correspondence to:
Dr. Mirjam J. Knol,
m.j.knol@umcutrecht.nl

CMAJ 2012. DOI:10.1503/cmaj.101715

KEY POINTS

- Odds ratios, often used in cohort studies and randomized controlled trials (RCTs), are often interpreted as risk ratios but always overestimate the risk ratio.
- We evaluated alternatives for logistic regression to obtain adjusted risk ratios to determine which method performed best in estimating the correct risk ratio and confidence interval.
- The Mantel-Haenszel risk ratio method, log-binomial regression, Poisson regression with robust standard errors, and the doubling-of-cases method with robust standard errors gave correct risk ratios and confidence intervals.
- To avoid any misinterpretation of odds ratios, adjusted risk ratios should be calculated and presented in cohort studies and RCTs.

This is the first in an occasional series that examines controversial aspects of research methods and reporting.

domly assigned to groups receiving surgery followed by radiotherapy, or radiotherapy alone.¹² The primary outcome was the ability to walk, which occurred in 70.3% of the patients. The authors stratified their results for ability to walk at baseline and presented a Mantel–Haenszel odds ratio of 6.2 (95% confidence interval 2.0–19.8) in their abstract. Based on the numbers presented in the paper, we calculated the Mantel–Haenszel risk ratio and also the crude odds ratio and risk ratio. These results are presented in Table 2. The difference between the odds ratio and risk ratio is very large, especially for the stratified odds ratio and risk ratio (6.26 v. 1.48). Readers could easily mistake the presented odds ratio as a risk ratio, which would lead to strong misinterpretation of the results.

Frequency of this problem in the literature

To verify how frequent these problems are, we did a survey of published cohort studies ($n = 75$)

and RCTs ($n = 288$).¹³ About one-third of cohort studies used logistic regression to adjust for baseline variables, and 40% of these presented odds ratios that deviated more than 20% from the approximate underlying risk ratio. Only about 5% of RCTs used logistic regression to adjust for baseline variables; however, about two-thirds of these presented odds ratios that deviated more than 20% from the risk ratio. The odds ratios deviate more often in RCTs, presumably because the frequency of the outcomes is more often large in RCTs.

Alternatives to logistic regression to estimate adjusted risk ratios

We found eight methods to estimate adjusted risk ratios in the literature (Table 3^{5,7-9,14-19}). The Mantel–Haenszel risk ratio method is straightforward and gives a weighted risk ratio over strata of covariables.^{14,15} This method is only practica-

Table 1: Results of an observational cohort study that assessed the effect of changes in maternal marital status on use of cannabis in children¹¹

Variable	Changes in maternal marital status		
	Nil $n = 2306$	1–2 $n = 584$	≥ 3 $n = 118$
Cannabis use ever, no. (%)	1035 (44.9)	346 (59.2)	80 (67.8)
Crude risk ratio (95% CI)	1.0 (reference)	1.3 (1.2–1.4)	1.5 (1.3–1.7)
Crude odds ratio (95% CI)	1.0 (reference)	1.8 (1.5–2.1)	2.6 (1.7–3.8)
Adjusted risk ratio* (95% CI)	1.0 (reference)	1.3 (1.2–1.4)	1.5 (1.2–1.6)
Adjusted odds ratio* (95% CI)	1.0 (reference)	1.7 (1.4–2.0)	2.3 (1.5–3.4)

CI = confidence interval.
*Adjusted for sex of the child, mother's age, family income, maternal and child mental health at five years, maternal substance use at five years, and frequency of change in marital status between the 5- and 14-year follow-up.

Table 2: Results for a randomized controlled trial on the effect of surgery on ability to walk in 101 patients with spinal cord compression caused by metastatic cancer¹²

Variable	Result	Surgery and radiotherapy $n = 50$	Radiotherapy alone $n = 51$
Walking at baseline	Walking at follow-up	32	26
	Not walking at follow-up	2	9
Not walking at baseline	Walking at follow-up	10	3
	Not walking at follow-up	6	13
Crude risk ratio (95% CI)		1.48 (1.13–1.93)	1.00 (reference)
Crude odds ratio (95% CI)		3.98 (1.56–10.17)	1.00 (reference)
Stratified risk ratio* (95% CI)		1.48 (1.16–1.90)	1.00 (reference)
Stratified odds ratio* (95% CI)		6.26 (1.98–19.75)	1.00 (reference)

CI = confidence interval.
*Stratified for walking at baseline.

ble if adjusting for a small number of categorical covariables (i.e., continuous covariables first need to be categorized). Log-binomial and Poisson regression are generalized linear models that directly estimate risk ratios.^{7,8} The default standard errors obtained by Poisson regression are typically too large; therefore, calculation of robust standard errors for Poisson regression may be needed to obtain a correct confidence interval around the risk ratio.⁹ The other four methods use odds ratios or logistic regression to estimate risk ratios. The Zhang and Yu method is a simple formula that calculates the risk ratio based on the odds ratio and the incidence of the outcome in the unexposed group.⁵ The doubling-of-cases method concerns changing the data set in such a way that logistic regression yields a risk ratio instead of an odds ratio.¹⁷ Again, calculation of robust standard errors may be needed to obtain a correct confidence interval around the risk ratio.¹⁸ Lastly, the method proposed by Austin uses the predicted probabilities obtained from a logistic regression model to estimate risk ratios.¹⁹ A recent review article of methods to estimate risk ratios and risk differences in cohort studies illustrated several of these eight methods using empirical data.²⁰

Simulation study

We conducted a simulation study to evaluate which of these eight methods performed best with regard to estimating the correct risk ratio and confidence interval. We also compared the estimated risk ratios with the odds ratio obtained with logistic regression. Details of the methods and results of the simulations are described in Appendix 1, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.101715/-/DC1. In this section, we summarize the main findings of the simulations in a simple situation (dichotomous determinant and outcome, and one continuous confounder) (Figures 1 and 2 in Appendix 1). Results for more complex situations (multiple dichotomous or continuous confounders) were essentially the same.

As expected, the odds ratio obtained with logistic regression overestimated the risk ratio importantly. This overestimation increased with increasing incidence of the outcome, increasing exposure effect and increasing amount of confounding. The method of Zhang and Yu also overestimated the risk ratio, although the overestimation was less pronounced than in logistic regression. This overestimation also increased with increasing incidence, increasing exposure

Table 3: Eight methods to estimate adjusted risk ratios that have been described in the literature

Mantel-Haenszel method to estimate a risk ratio ^{14,15}	A Mantel-Haenszel risk ratio is calculated by taking a weighted average of risk ratios in strata of covariables, where the weight depends on the size of the strata.
Log-binomial regression ⁸	Log-binomial regression is a generalized linear model with a log link and a binomial distribution. It is similar to logistic regression, except that the link function is a log link instead of a logit link, hence providing risk ratios instead of odds ratios.
Ordinary Poisson regression ⁷	The data are fitted with a generalized linear model with a log link and a Poisson distribution. This approach yields correct estimates of the risk ratio, but the obtained standard errors are in general too large.
Poisson regression with robust standard errors ⁹	Robust standard errors are estimated with a procedure known as sandwich estimation ¹⁶ to account for the incorrect assumption of Poisson distributed outcomes in the Poisson regression approach.
Method proposed by Zhang and Yu ⁵	A risk ratio is calculated based on the odds ratio and the incidence of the outcome in the unexposed group.
Doubling-of-cases method, proposed by Miettinen ¹⁷	Miettinen proposed to include those with the outcome twice in the data set, i.e., once with the outcome and once without the outcome. Then the odds ratio that is estimated by logistic regression analysis in the "new" cohort is in fact the risk ratio of the "original" cohort: the odds ratio is an exact estimation of the risk ratio. This solution is akin to the case-cohort study with a sampling at baseline of 100%. However, the obtained standard errors are too large.
Doubling-of-cases method with robust standard errors ¹⁸	Robust standard errors are estimated with a procedure known as sandwich estimation ¹⁶ to adjust for the too-large standard errors obtained by the doubling-of-cases method proposed by Miettinen.
Method proposed by Austin ¹⁹	This method uses logistic regression analysis to estimate individual probabilities of having the outcome if a subject would have been either exposed or unexposed. A risk ratio is then calculated by taking the ratio of the means of these probabilities.

effect and increasing amount of confounding. The method proposed by Austin underestimated the risk ratio in case of a large exposure effect and a large incidence of the outcome. The Mantel–Haenszel risk ratio method performed well in all situations, except in the situation with moderate confounding, where it slightly overestimated the true risk ratio. This was due to residual confounding because we simulated a continuous confounder and categorized the confounder into quintiles to calculate the Mantel–Haenszel risk ratio.

Log–binomial regression, Poisson regression with robust standard errors, and the doubling-of-cases method with robust standard errors all yielded correct risk ratios and confidence intervals in all situations of our simulations. However, all of these methods have potential disadvantages with particular data sets that could force the investigator to discard some methods and prefer another method, according to the data at hand. A disadvantage of log–binomial regression is that the model does not converge in certain situations (i.e., the model cannot find a solution and therefore the risk ratio cannot be calculated). These convergence problems mainly come up if several continuous covariates are included in the model and if the incidence of the outcome is high. Poisson regression with robust standard errors does not have this problem but has the disadvantage that the model may yield individual predicted probabilities above 1. Probabilities above 1 are not a problem if the only interest is in obtaining a valid risk ratio. If the interest is also in the individual predicted probabilities of disease, for example in prognostic or diagnostic research, probabilities above 1 may be problematic. A disadvantage of the doubling-of-cases method with robust standard errors, which has neither of these problems, is that it requires some manipulation of data before the analyses can be performed. Furthermore, the calculation of the robust standard error in the doubling-of-cases approach is not available in standard statistical software packages and demands expertise to program.

Recommendations for clinical researchers

We showed in the clinical examples and simulations that an odds ratio can substantially overestimate the risk ratio. In fact, both are correct, but when an odds ratio is interpreted as a risk ratio, serious misinterpretation with potential consequences for treatment decisions and policy-making can occur, as illustrated by the two clinical examples. Therefore, any misinterpretation of odds ratios should be avoided with calculation

and presentation of adjusted risk ratios in both cohort studies and RCTs. Also, if adjustment for baseline covariates is not done, which is often the case in RCTs, the risk ratio is the preferred measure of association in case of dichotomous outcomes.²¹ Note that in case–control studies, the odds ratio is the appropriate effect estimate and the odds ratio can be interpreted as a risk ratio or rate ratio depending on the sampling method.^{1–4} Of course, if data of cohort studies or RCTs are collected so that a time-dependent analysis is possible, Cox regression yielding hazard ratios is recommended because it estimates relative hazards and does not involve problems related to odds ratios.

There are several valid methods to estimate adjusted risk ratios. In a situation with only one or two categorical covariables, for example, to take into account stratified randomization in an RCT (example 2), we recommend use of the simple Mantel–Haenszel risk ratio method. This method can be easily applied by using Rothman’s spreadsheet *Episheet* (can be downloaded from <http://krothman.byethost2.com/>). In a situation with more covariables or continuous covariables, we recommend use of log–binomial regression. If log–binomial regression does not converge, Poisson regression with robust standard errors can be applied. Both methods are easy to perform in standard statistical software packages, including SAS, Stata, R and SPSS^{22,23} (see Appendix 2, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.101715/-/DC1, for codes). If the Poisson method is in turn problematic because individual probabilities have to be estimated and those estimates become larger than 1 for some individuals, there may be no other solution than the doubling-of-cases method with robust standard error estimation, but this needs extra programming and statistical expertise. In line with other commentators,^{7,10} we discourage the use of the Zhang and Yu method, despite its ease of application and its appealing conceptual simplicity.

Conclusion

In this paper we have shown the problems of using odds ratios as an approximation of risk ratios in cohort studies and RCTs. Researchers, reviewers and journal editors should be aware of potential misinterpretation of odds ratios, especially when the incidence of the outcome is large. The problem often arises when researchers use logistic regression to adjust for potential confounders. Misinterpretation of odds ratios should be avoided by calculating adjusted risk ratios. Journal editors and statistical reviewers can play

an important role in encouraging researchers to present risk ratios instead of odds ratios in cohort studies and RCTs.

References

- Greenland S, Thomas DC. On the need for the rare disease assumption in case-control studies. *Am J Epidemiol* 1982;116:547-53.
- Greenland S, Thomas DC, Morgenstern H. The rare-disease assumption revisited. A critique of "estimators of relative risk for case-control studies". *Am J Epidemiol* 1986;124:869-83.
- Knol MJ, Vandenbroucke JP, Scott P, et al. What do case-control studies estimate? Survey of methods and assumptions in published case-control research. *Am J Epidemiol* 2008;168:1073-81.
- Miettinen O. Estimability and estimation in case-referent studies. *Am J Epidemiol* 1976;103:226-35.
- Zhang J, Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. *JAMA* 1998;280:1690-1.
- Barros AJ, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. *BMC Med Res Methodol* 2003;3:21.
- McNutt LA, Wu C, Xue X, et al. Estimating the relative risk in cohort studies and clinical trials of common outcomes. *Am J Epidemiol* 2003;157:940-3.
- Robbins AS, Chao SY, Fonseca VP. What's the relative risk? A method to directly estimate risk ratios in cohort studies of common outcomes. *Ann Epidemiol* 2002;12:452-4.
- Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol* 2004;159:702-6.
- McNutt LA, Hafner JP, Xue X. Correcting the odds ratio in cohort studies of common outcomes. *JAMA* 1999;282:529.
- Hayatbakhsh MR, Najman JM, Jamrozik K, et al. Changes in maternal marital status are associated with young adults' cannabis use: evidence from a 21-year follow-up of a birth cohort. *Int J Epidemiol* 2006;35:673-9.
- Patchell RA, Tibbs PA, Regine WF, et al. Direct decompressive surgical resection in the treatment of spinal cord compression caused by metastatic cancer: a randomised trial. *Lancet* 2005;366:643-8.
- Knol MJ, Duijnhoven RG, Grobbee DE, et al. Potential misinterpretation of treatment effects due to use of odds ratios and logistic regression in randomized controlled trials. *PLoS ONE* 2011;6:e21248.
- Greenland S, Rothman KJ. Introduction to stratified analysis. In: *Modern epidemiology*, 3rd ed. Philadelphia (PA): Lippincott, Williams & Wilkins; 2008. p. 258-82.
- Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 1959;22:719-48.
- Royall RM. Model robust confidence intervals using maximum likelihood estimators. *Int Stat Rev* 1986;54:221-6.
- Miettinen O. Design options in epidemiologic research. An update. *Scand J Work Environ Health* 1982;8(Suppl 1):7-14.
- Schouten EG, Dekker JM, Kok FJ, et al. Risk ratio and rate ratio estimation in case-cohort designs: hypertension and cardiovascular mortality. *Stat Med* 1993;12:1733-45.
- Austin PC. Absolute risk reductions, relative risks, relative risk reductions, and numbers needed to treat can be obtained from a logistic regression model. *J Clin Epidemiol* 2010;63:2-6.
- Austin PC, Laupacis A. A tutorial on methods to estimating clinically and policy-meaningful measures of treatment effects in prospective observational studies: a review. *Int J Biostat* 2011;7: Article 6. DOI: 10.2202/1557-4679.1285.
- Groenwold RH, Moons KG, Peelen LM, et al. Reporting of treatment effects from randomized trials: a plea for multivariable risk ratios. *Contemp Clin Trials* 2011;32:399-402.
- Lumley T, Kronmal R, Ma S. *Relative risk regression in medical research: models, contrasts, estimators, and algorithms: UW biostatistics working paper series 2006*. Seattle (WA): University of Washington; 2006. Working paper 293.
- Spiegelman D, Hertzmark E. Easy SAS calculations for risk or prevalence ratios and differences. *Am J Epidemiol* 2005;162:199-200.

Competing interests: Mirjam Knol's institution has received a grant from Top Institute Pharma. Ale Algra's institution has received speaker fees and funding for participation in interna-

tional advisory board meetings from Boehringer Ingelheim, and has grants or grants pending for cerebrovascular research from Netherlands Heart Foundation, Trombostichting Nederland, Netherlands Organisation for Scientific Research and Netherlands Organisation for Health Research and Development. Ale Algra has received funding for accommodation from the European Stroke Conference for chairing sessions and grading abstracts, and is a principal investigator of the European/Australasian Stroke Prevention in Reversible Ischaemia Trial, which received financial support from Boehringer Ingelheim for post-hoc exploratory analyses of the trial data. None declared by Saskia Le Cessie, Jan Vandenbroucke or Rolf Groenwold.

Affiliations: From the Julius Center for Health Sciences and Primary Care (Knol, Algra, Groenwold), University Medical Center Utrecht, Utrecht, the Netherlands; and the Departments of Clinical Epidemiology (Le Cessie, Algra, Vandenbroucke) and Medical Statistics and BioInformatics (Le Cessie), Leiden University Medical Center, Leiden, the Netherlands

Contributors: All of the authors conceived and designed the analysis. Mirjam Knol, Saskia Le Cessie and Rolf Groenwold analyzed and interpreted the data. Mirjam Knol and Rolf Groenwold drafted the article, which Saskia Le Cessie, Ale Algra and Jan Vandenbroucke revised. All of the authors approved the final version of the article.

CANADIAN MEDICAL ASSOCIATION 145th ANNUAL MEETING
145^e ASSEMBLÉE ANNUELLE DE L'ASSOCIATION MÉDICALE CANADIENNE

Yellowknife, Northwest Territories | August 12–15, 2012
Yellowknife (Territoires du Nord-Ouest) | du 12 au 15 août 2012

Registration and travel information:
CMA Meetings and Travel Management
800 663-7336 or
613 731-8610 x2383
Fax: 613 731-8047

Information au sujet de l'inscription et des arrangements de voyage :
Gestion des conférences et voyages de l'AMC
800 663-7336 ou
613 731-8610, poste 2383
Télécopieur : 613 731-8047

Hosted by the Northwest Territories Medical Association
Votre hôte : l'Association médicale des Territoires du Nord-Ouest

