

## A Practical Guide for Multivariate Analysis of Dichotomous Outcomes

James Lee,<sup>1,2</sup>*BSc, MSc, PhD*, Chuen Seng Tan,<sup>2</sup>*BSc, MSc*, Kee Seng Chia,<sup>2</sup>*MBBS, MSc (OM), MD*

### Abstract

**A dichotomous (2-category) outcome variable is often encountered in biomedical research, and Multiple Logistic Regression is often deployed for the analysis of such data. As Logistic Regression estimates the Odds Ratio (OR) as an effect measure, it is only suitable for case-control studies. For cross-sectional and time-to-event studies, the Prevalence Ratio and Cumulative Incidence Ratio can be estimated and easily interpreted. The logistic regression will produce the OR which is difficult to interpret in these studies. In this report, we reviewed 3 alternative multivariate statistical models to replace Logistic Regression for the analysis of data from cross-sectional and time-to-event studies, viz, Modified Cox Proportional Hazard Regression Model, Log-Binomial Regression Model and Poisson Regression Model incorporating the Robust Sandwich Variance. Although none of the models is without flaws, we conclude the last model is the most viable. A numeric example is given to compare the statistical results obtained from all 4 models.**

*Ann Acad Med Singapore 2009;38:714-9*

**Key words:** Alternatives to logistic regression, Cross-sectional studies, Risk ratio vs odds ratio

### Introduction

A dichotomous (2-category) outcome variable is indeed ubiquitous in biomedical research enquiries. Here are some examples:

- (a) a cross-sectional study to compare the prevalence (proportion) of obesity among the adult males and females in Singapore.
- (b) a clinical trial to compare ethnic differences in the 1-year survival among patients with metastatic non-small cell lung cancer treated with Gefitinib (Iressa).
- (c) an epidemiologic study to compare the 2-year mortality rate of lung cancer between those who continue to smoke and those who quit after diagnosis.

The outcome variable (henceforth denoted by Y) in all these examples is dichotomous or binomial: obese or non-obese, surviving or not surviving by the end of 1 year, died or surviving by the end of 2 years. Note that examples (b) and (c) are generally referred to as “time-to-event data with a constant risk period” (1 year, 2 years) or ‘closed cohort studies’.

The usual chi-square test is generally deployed to compare the prevalence proportion (example a) or cumulative incidence rate (examples b and c) among 2 or

more groups. However, this practice can produce highly misleading results because virtually all biological outcomes are multifactorial (affected by many factors, not just the one being investigated). Hence, the apparent difference in the proportion is likely to be confounded by numerous factors, such as gender, age and stage of cancer at time of diagnosis. Indeed, confounding is a virtual guarantee in any non-randomised comparison study, and virtually all aetiologic studies of adverse effects involving human subjects are, and must be, non-randomised.<sup>1,2</sup> It must be noted that confounding can still occur in a randomised comparison study, because random allocation only tends to minimise confounding as the study size increases. For example, if we were to randomly allocate 100 subjects to 2 groups, sex and age will unlikely be totally balanced in the 2 compared groups. However, if we were to randomly allocate 10,000 subjects to 2 groups, sex, age and other potential confounders will be very balanced, and confounding will be negligible.

Clearly, a useful statistical method for comparison studies must have the capability for the adjustment of confounding. For a long time, the so-called stratification and standardisation methods have been widely deployed.<sup>1,2</sup> Because these methods use cross-tabulation, they are

<sup>1</sup> Professor of Biostatistics, University of Hawaii, USA

<sup>2</sup> Centre for Molecular Epidemiology, National University of Singapore, Singapore

Address for Correspondence: Prof Kee Seng Chia, Centre for Molecular Epidemiology, Yong Loo Lin School of Medicine, National University of Singapore, MD3, 16 Medical Drive, Singapore 117597.

Email: cofcks@nus.edu.sg

suitable for adjusting for 1 or at most 2 confounding covariates based on a typical study size. Moreover, a continuous covariate (age, blood pressure) must be grouped into a few classes, thereby incurring residual confounding. Consequently, considerable efforts have been devoted to developing statistical models for the analysis of data from comparative studies. These model-based multivariate statistical methods are capable not only for the adjustment of several confounding covariates, but also the assessment of interaction (effect modification) between factors, based on modest study size.<sup>2,3</sup> Not surprisingly, a cursory look at the journals will attest that model-based multivariate methods are routinely deployed for the analysis of research data. However, not all models are the same, and the wrong choice of model can produce misleading statistical results.

The choice of the model depends on the study design and ‘effect measures’. For the remainder of this paper, we will briefly highlight 3 common ‘effect measures’ (there are others) in medical and epidemiological research and review the various statistical models used to estimate them. The pitfalls and merits of these models will be highlighted and specifically, the limitations of Logistic Regression for the analysis of dichotomous outcomes will be *emphasised*.

**Effect Measures**

In aetiologic research, an ‘effect measure’ quantifies the effect of an exposure or factor X (e.g. gender, active therapy, smoking cessation) on an outcome event Y (obesity, 1-year survival, 2-year mortality). An effect measure can either be a **ratio** (percent of obesity in males divided by that in females), or a **difference**, RD (mortality rate in those who quit minus those who continued smoking). Although in this paper, we will only discuss ratio effect measures, difference effect measures are of particular importance in clinical trials and public health settings where the aim is to evaluate the absolute magnitude of benefit when a new therapy is introduced or when a risk factor is removed.

In many publications, all effect measures are loosely referred to as Relative Risk (RR). Some would restrict the use of RR for cohort studies and the Odds Ratio (OR) for case-control studies. It is important to use the appropriate effect measure based on the study design. In a cross-sectional study, the Prevalence Ratio (PR) is used. In a closed cohort study with time-to-event data with constant risk period (most clinical trials), the Cumulative Incidence Ratio (CIR) is the appropriate effect measure. In an open cohort study, the Incidence Density Ratio (IDR) is used. Odds Ratio (OR) is to be used only for case-control studies.

**Misuses of OR**

As many investigators erroneously misinterpret OR as though it were the RR (IDR, CIR or PR), leading to incorrect

Table 1. Hypothetical Data to Highlight the Difference Between Relative Risk (RR) and Odds Ratio (OR)

Response (Y)	Exposure (X)	
	+	-
+	120 (60%)	40 (20%)
-	80 (40%)	160 (80%)
	<b>200</b>	<b>200</b>

conclusions, it is crucial to highlight the differences between these 2 effect measures. We consider some hypothetical data displayed in Table 1. Let **P+** denote the proportion of positive-Y (death within 2 years) in positive-X (those who continued smoking), and **P-** denote the proportion of positive-Y in negative-X (those who quit smoking). Then

$$RR = P+ / P- = 0.6 / 0.2 = 3$$

$$OR = [P+ / (1 - P+)] / [P- / (1 - P-)] = (0.6 / 0.4) / (0.2 / 0.8) = 6$$

We note that whereas RR has a transparent meaning (the risk of dying within 2 years is, on average, 3 times greater in those who continued smoking than those who stopped), OR, being a ratio of 2 ratios, is completely devoid of intelligibility. Moreover, the numeric value of RR and OR can be quite discrepant. So the question is, why do we use OR at all as an effect measure? The answer is given in the next section.

**Logistic Regression Model**

As noted above, a dichotomous Y variable can have only 2 values. Y is either 0 (alive beyond 2 years) or 1 (death within 2 years). If we denote  $\pi$  as the probability of Y = 1, then Y follows the binomial distribution with  $\pi$  values bounded by 0 and 1, with a binomial variance equal to  $\{\pi \times (1-\pi)\}$ .

The Logistic Regression Model was expressly developed as a multivariate method for binomial Y variable. The model estimates the correct binomial variance, and it deploys the logit link function to constrain the model-predicted  $\pi$  to lie between 0 and 1 (Link functions for various statistical models are expounded in reference 3). Hence in terms of mathematical properties, the Logistic Model is undisputedly the best model for binomial Y, and the merit of any competing model must be judged against the Logistic Model. If the Logistic Model is the best, then why develop competing models in the first place? Because the Logistic Model estimates OR (not RR) as an effect measure.

The Logistic Model was initially adapted for case-control studies (see citations in reference 2) because data from a case-control study can only determine OR.<sup>1</sup> Also, a case-

control study is suitable provided the event is ‘rare’ in the population (say, colon cancer), in which case, OR is a closed approximation of RR. Using the data of Table 1 as an example, if we change  $P^+$  from 0.6 to 0.06, and  $P^-$  from 0.2 to 0.02, then  $OR = 3.13$ , which is very close to  $RR = 3$ . In fact, a case-control study is typically conducted for events with  $\pi$  much lower than 0.02 in the population, so OR and RR are virtually interchangeable. Thus, Logistic Regression is eminently useful for case-control studies only because the numeric value of OR mimics RR. On the other hand, RR can directly be determined from data based on cross-sectional and time-to-event (cohort) studies. Additionally, these studies are practical to undertake only for events that are relatively common. If the event was ‘rare’, we would need to take an inordinately large cross-sectional sample to accrue sufficient cases, or to monitor a large sample of subjects over many years in a time-to-event study. Consequently, OR (an unintelligible effect measure) estimated by Logistic Regression will be numerically highly discrepant than RR (a meaningful effect measure). OR will always be greater than RR if RR is greater than 1 (adverse effect); OR will be smaller than RR if RR is less than 1 (protective effect); and OR approaches RR as  $\pi$  in the population approaches 0 (rare event).

The woeful inadequacy of Logistic Regression for cross-sectional studies was underscored by Lee and Chia in 1993 and 1994 to the Editors of 2 international journals,<sup>4,5</sup> which stimulated considerable discussions over the years<sup>6-21</sup> (and numerous others). Most of the authors agreed that PR is the effect measure of choice, and the few who favoured OR based their contention entirely on mathematical properties, which has doubtful relevance in a practical setting. As will be noted later, the main thrust of the long-standing debate pertains rather to statistical modelling issues, specifically to the variance of the effect measure.

Barros and Hirakata<sup>19</sup> did an online search for cross-sectional and time-to-event studies published in highly reputable international journals, and found Logistic Regression was used in 37 (34%) of the 110 cross-sectional studies, and 10 (22%) of the 45 time-to-event studies. Greenland<sup>22</sup> has demonstrated persuasively that as an effect measure, OR is more defective for time-to-event studies than is generally realised. Deeks and Altman<sup>23,24</sup> provided useful summaries on the pitfalls of OR.

Perhaps the most profound example on the disastrous consequence of misinterpreting OR as though it were RR was a study to compare the percent of physicians’ referrals for cardiac catheterisation between Black and White women.<sup>25</sup> A 7 per cent lower referral rate for Black women (if measured by RR) was erroneously reported as 40 percent (misinterpreting the published OR as though it were RR) in the various news media, including a US television

programme, Nightline. Subsequently, this caused a heated debate on racial issues. The editors of the *New England Journal of Medicine* issued an apology<sup>26</sup> over the mistake. It is no exaggeration to assert that the uncritical application of Logistic Regression and the misinterpretation of OR as RR have, on balanced, incurred greater damage than benefit in biomedical research in the last several decades.

It is gratifying to know that both the *New England Journal of Medicine* and the *American Journal of Epidemiology* officially discourage the use of Logistic Regression (i.e., to report OR) for any study in which RR is ascertainable from the data.<sup>26,27</sup> The recognised need to report RR instead of OR has prompted a publication of a simplistic formula for converting OR to RR,<sup>28</sup> which was severely criticised by others.<sup>20,29</sup> Unbelievably, this naive formula was actually deployed in 56 publications in reputable journals from the end of 1998 to May 2001.<sup>20</sup>

We will now consider alternatives to Logistic Regression. All the models discussed below estimate RR from any comparative study in which this effect measure is ascertainable from the data, viz, cross-sectional study (PR), time-to-event study with constant risk period (CIR). Although none of the models is mathematically perfect, as the Logistic Model is perfect, all the models have undisputed advantages over the Logistic Model for reasons expounded above.

### Modified Cox’s Proportional Hazards Regression Model

Cox’s Proportional Hazards Regression Model (PH) was originally developed for the analysis of time-to-event data with varying risk periods, i.e., a dynamic or opened cohort study with censoring and lost-to-followup.<sup>2</sup> The PH model estimates the Instantaneous Hazard Ratio (HR), which is IDR at the same point in time. For example, if HR of smoking on lung cancer is 2, we say that at any given same moment in time, the likelihood of the occurrence of lung cancer is twice as high in smokers than in non-smokers.

As events (lung cancer) clearly do not occur at the same moment (they occur over different times), the numeric value of HR will lie somewhere between OR and IDR, but if events do occur within the same moment, then HR will be equal to IDR.<sup>21</sup>

For cross-sectional studies, we do not know when the events had occurred for each comparison group. However, we can assume that all the events had occurred within some constant risk period (same time interval, however long or short), thus reducing HR to IDR. A technical discussion of this point is given by Lumley and colleagues.<sup>21</sup>

This was the reason which motivated Lee and Chia<sup>4,5</sup> to propose a modified version of the PH model for cross-sectional studies, as a replacement for Logistic Regression.

They showed that by stipulating a condition of constant risk period in the PH model, the Hazard Ratio reduces to CIR for closed cohort and PR for cross-sectional data. The RR estimated by the Lee-Chia method is in fact the Cumulative Incidence Ratio, (the correct effect measure for time-to-event data in examples b and c), and the Prevalence Ratio, (the correct effect measure for cross-sectional data in example a).<sup>1,5,18</sup>

Thus, if  $CIR = 2$  in example c, we say the likelihood or risk of dying within 2 years is twice as high for those who continue to smoke as compared to those who quit after diagnosis. If PR of men relative to women is 2 in example a, we would say that the likelihood of finding an obese person among men is twice that among women.

Although the Lee-Chia method produces the correct (unbiased) estimate of HR, the method has one notable limitation in that the variance of HR tends to be inflated, resulting in a wider confidence interval than it should be. This is because Lee-Chia assumed  $Y$  to be Poisson with a log link function, when in fact it is Binomial with a logit link function (i.e., Logistic Model). As the Poisson variance is unbounded (variance increases with mean), it is generally greater than the Binomial variance, which is bounded (0 when  $\pi = 0$ , maximum of 0.25 when  $\pi = 0.5$ ), thus resulting in an inflated variance for HR. It should be noted that the Poisson variance approaches the Binomial variance as  $\pi$  approaches 0, so for events of fairly low incidence or prevalence in the population, say  $\pi = 0.2$  or lower, the Lee-Chia estimated variance of HR should be quite acceptable. Nonetheless, this limitation has prompted the genesis of a long-standing debate,<sup>6-21</sup> which we will encapsulate in the following sections.

### Log-Binomial Regression Model

In view of the variance-inflation problem with the Lee-Chia method,<sup>4,5</sup> the Log-Binomial Regression Model<sup>6,7,9,19,20</sup> was suggested as a possible alternative. The model assumed  $Y$  to be Binomial (correct) with a log link function (incorrect). Although this model alleviated the variance-inflation problem, the log link function incurs other setbacks that are more damaging. Unlike the logit link function, which constrains the model-predicted  $\pi$  to fall between 0 and 1, the log link function only constrains  $\pi$  to be  $\geq 0$ , but not  $\leq 1$ . Subjects whose model-predicted  $\pi$  is  $>1$  will be excluded, and consequently, the mean model-predicted probabilities and HR will not be maximum likelihood estimates, i.e., they are biased. This problem is accentuated if  $\pi$  in the population is high, say  $>0.8$ , and if there are 1 or more continuous covariates (e.g., age, blood pressure) included in the model. The side effect of this problem is non-convergence in parameter estimations, that is, HR cannot even be estimated by the statistical programme. These issues are well-recognised.<sup>8,10,16,17</sup>

There have been suggestions on how to lessen the non-convergence problem,<sup>16,30</sup> but others have countered that even if the parameter estimations do converge, there is no guarantee that the HRs are even close to maximum likelihood estimates.<sup>21</sup> Those who defend the use of the Log-Binomial Model claim, without proof, that these problems are negligible if  $\pi$  in the population is not close to 1. They have made no mention about the inclusion of continuous covariates in the model.

It must be noted that the Lee-Chia method<sup>4,5</sup> is not beset by these 2 problems, because the outcome variable in the Cox PH Model is the “instantaneous hazard”, which is not a probability bounded by 0 and 1. Although neither the Lee-Chia method nor the Log-Binomial Model is without flaw, we think the former (yielding an unbiased estimate of RR, but with an inflated variance) is more viable than the latter (yielding a biased estimate of RR, plus the problem of non-convergence) in a practical setting.

### Modified Poisson Regression Model Incorporating the Robust Sandwich Variance

More recently, Zou<sup>31</sup> proposed the modified Poisson Regression Model incorporating the Robust Sandwich Variance (RSV) to statistically compensate for the Poisson variance-inflation problem. Like the PH Model of Lee and Chia,<sup>4,5</sup> Zou also assumed  $Y$  to be Poisson (incorrect) with a log link function. In other words, Lee-Chia<sup>4,5</sup> and Zou<sup>31</sup> both employed an identical model (Poisson) with identical transformation (log link function). In fact, it can be shown that under the condition of constant risk period, the Lee-Chia modified PH model is identical to Zou’s modified Poisson model without RSV, so both models will produce identical HRs and their variances.

Zou’s contribution is his incorporation of RSV<sup>32,33</sup> to the Poisson Model to compensate for the inflated variance of HR that is inherent in the Lee-Chia method, and therein lies the improvement over the Lee-Chia method. Nonetheless, Zou’s method is not mathematically perfect, like the Logistic Model is perfect. Both Lee-Chia and Zou mis-specified the true model, viz, Poisson instead of Binomial, log instead of logit link function.

Note also that RSV is heretofore used in conjunction with Generalised Estimating Equations for the analysis of correlated  $Y$  data based on clustered or repeated-measurement samples, where RSV is estimated as some mathematically complex function of the within-cluster and between-cluster variances.<sup>32-34</sup>

Although Zou incorporated the RSV in the Poisson Model using independent  $Y$  data, he did not elaborate how RSV is estimated from a non-cluster sample. Also, using RSV only statistically compensates for the inflated

Poisson variance. Because RSV is itself estimated based on strong assumptions, there is no certainty that the “RSV compensated variance” of a Poisson variate is equal to the true Binomial variance. The author himself stated that research on this topic is in progress.

In summary, none of the post-Logistic Models discussed here possesses perfect mathematical properties for estimating HR based on dichotomous Y data. However, on balance, Zou’s method is clearly the most viable, and an online search has revealed that the method has indeed been gaining wide acceptance. We too endorse its use.

### A Numeric Example

We now present a numeric example to compare the statistical results obtained from all 4 multivariate models. The data (collected by KSC) came from a cross-sectional study of 124 workers occupationally exposed to cadmium. The data analytic goal is to estimate the crude (unadjusted) and covariates-adjusted effect measure of the duration of exposure to cadmium (X) on blood cadmium level (Y). Duration of exposure is grouped into 3 categories:  $\leq 1$  year, 1 to  $\leq 8$  years,  $> 8$  years. Blood cadmium level is dichotomised into 2 categories:  $> 5 \mu\text{g/L}$  (elevated) and  $\leq 5 \mu\text{g/L}$  (normal).

We note from Table 2 that the Prevalence Ratio (PR) for (1 to  $\leq 8$  years relative to  $\leq 1$  year) is  $56/19 = 2.9$ , and PR for ( $> 8$  years relative to  $\leq 1$  year) is  $85/19 = 4.5$ . The model-predicted effect measures are shown on Table 3. First, we note that both the crude and adjusted ORs estimated by the Logistic Model is grossly discrepant from the corresponding PRs estimated by the other 3 models. If we were to misinterpret OR as PR, as is often done,<sup>25,26</sup> our conclusion would be grossly misleading.

Second, we note that the PRs (crude and adjusted) and

their variances (confidence intervals) as estimated by the Log-Binomial Model is somewhat discrepant from those estimated by the Modified Cox and Poisson with RSV Models. More importantly, the Log-Binomial Model incurred a problem of non-convergence, which prompted the statistical programme to issue a warning that the results may not be trustworthy, thus re-iterating our view that the Log-Binomial Model is not viable.

Third, we note that the crude and adjusted PRs estimated by the Poisson Model with RSV are identical to those estimated by the Modified Cox Model, confirming the fact that both models are identical in estimating the effect measures.

Fourth, we note that the variances are smaller (confidence intervals are narrower) as estimated by the Poisson Model with RSV than those by the Modified Cox Model. This is of course expected, as RSV compensates the inflated Poisson variance in the Modified Cox Model.

All the statistical analyses were carried out by SAS.<sup>35</sup> The SAS program for Zou’s Modified Poisson Regression Model with Robust Sandwich Variance,<sup>31</sup> including a sample dataset and related information documenting the analytical process, can be downloaded from this website: <http://www.med.nus.edu.sg/cof/cme.html>

Table 2. Years of Exposure (risk factor) and Blood Cadmium Level (response) in a Cross-sectional Study of 124 Workers Occupationally Exposed to Cadmium

Blood cadmium level ( $\mu\text{g/L}$ )	Years exposed		
	$\leq 1$	1 to $\leq 8$	$> 8$
“Elevated” ( $> 5$ )	8 (19%)	35 (56%)	17 (85%)
“Normal” ( $\leq 5$ )	34	27	3
Total	42	62	20

Table 3. Crude and Adjusted<sup>1</sup> Effect Measures as Estimated by the Various Multivariate Models

Duration of exposure in years	Logistic (OR)	Log-Binomial (PR) <sup>2</sup>	Modified Cox (PR)	Poisson with RSV (PR)
<b>Unadjusted effect measures</b>				
$\leq 1$	-	-	-	-
1 to $\leq 8$	5.5 (2.2-13.8)	2.9 (1.5-5.7)	2.9 (1.4-6.4)	2.9 (1.5-5.7)
$> 8$	24.1 (5.6-102.5)	4.5 (2.3-8.5)	4.5 (1.9-10.3)	4.5 (2.3-8.5)
<b>Adjusted effect measures</b>				
$\leq 1$	-	-	-	-
1 to $\leq 8$	6.5 (2.3-18.6)	2.6 (1.4-4.9)	3.2 (1.4-7.1)	3.2 (1.6-6.2)
$> 8$	25.8 (5.5-121.5)	3.8 (2.1-6.7)	4.5 (1.9-11.0)	4.5 (2.3-8.9)

( ) 95% confidence intervals

<sup>1</sup> Adjusted for age and gender

<sup>2</sup> Statistical programme issued a warning: “Convergence was questionable”. Therefore these results may not be reliable.

## Conclusion

Although in cross-sectional studies the OR can be computed, it cannot be interpreted as a RR. The most appropriate effect measure for a cross-sectional study is the PR. In biomedical research, the interpretability of an effect measure is far more important than whether it could be computed. The logistic regression was originally adapted for case-control studies as it estimates the OR and should not be used for cross-sectional study design. However, there are statistical challenges in computing the PR and no ideal model exist. The Poisson Regression Model incorporating the Robust Sandwich Variance should be used in cross-sectional studies for estimating the PR.

## REFERENCES

- Rothman KJ, Greenland S. *Modern Epidemiology*. Philadelphia: Lippincott-Raven, 1998.
- Anderson S, Auquier A. *Statistical Methods for Comparative Studies: Techniques for Bias Reduction*. New York: Wiley & Sons, 1980.
- McCullagh P, Nelder JA. *Generalized linear models*. London: Chapman & Hall, 1989.
- Lee J, Chia KS. Estimation of prevalence rate ratios for cross-sectional data: An example in occupational epidemiology. *Br J Indust Med* 1993;50:861-2
- Lee J. Odds ratio or relative risk for cross-sectional data? *Int J Epidemiol* 1994;23:201-3.
- Stromberg U. Prevalence odds ratio v prevalence ratio. *Occup Environ Med* 1994;51:143-4.
- Axelsson O, Fredriksson M, Ekberg K. Use of the prevalence ratio v the prevalence odds ratio as a measure of risk in cross sectional studies. *Occup Environ Med* 1994;51:574.
- Lee J, Chia KS. Use of the prevalence ratio v the prevalence odds ratio as a measure of risk in cross-sectional studies: A reply. *Occup Environ Med* 1994;51:841.
- Zocchetti C. Estimation of prevalence rate ratios from cross-sectional data. *Int J Epidemiol* 1995;24:1064-5.
- Lee J. Estimation of prevalence rate ratios from cross-sectional data: A reply. *Int J Epidemiol* 1995;24:1066-7.
- Axelsson O, Fredriksson M, Ekberg K. Use of the prevalence ratio v the prevalence odds ratio in view of confounding in cross sectional studies. *Occup Environ Med* 1995;52:494.
- Lee J, Chia KS. Prevalence odds ratio v prevalence ratio – a response. *Occup Environ Med* 1995;52:781-2.
- Hughes K. Odds ratios in cross-sectional studies. *Int J Epidemiol* 1995;24:463-4, 468.
- Nurminen M. To use or not to use the odds ratio in epidemiologic analyses? *Eur J Epidemiol* 1995;11:365-71.
- Zocchetti C, Consonni D, Bertazzi PA. Relationship between prevalence rate ratios and odds ratios in cross-sectional studies. *Int J Epidemiol* 1997;26:220-3.
- Skov T, Deddens J, Petersen MR, Endahl L. Prevalence proportion ratios: Estimation and hypothesis testing. *Int J Epidemiol* 1998;27:91-5
- Thompson ML, Myers JE, Kriebel D. Prevalence odds ratio or prevalence ratio in the analysis of cross sectional data: What is to be done? *Occup Environ Med* 1998;55:272-7.
- Case LD, Kimmick G, Paskett ED, Lohman K, Tucker R. Interpreting measures of treatment effect in cancer clinical trials. *Oncologist* 2002;7:181-7.
- 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, Hafner JP. Estimating the relative risk in cohort studies and clinical trials of common outcomes. *Am J Epidemiol* 2003;157:940-3
- Lumley T, Kronmal R, Ma S. Relative risk regression in medical research: Models, contrasts, estimators, algorithms. University of Washington Biostatistics Working Paper Series 293. 2006:24
- Greenland S. Interpretation and choice of effect measures in epidemiologic analyses. *Am J Epidemiol* 1987;125:761-8.
- Deeks J. When can odds ratios mislead? Odds ratios should be used only in case-control studies and logistic regression analyses. *BMJ* 1998;317:1155-6; author reply 1156-7
- Altman DG, Deeks JJ, Sackett DL. Odds ratios should be avoided when events are common. *BMJ* 1998;317:1318.
- Schulman KA, Berlin JA, Harless W, Kerner JF, Sistrunk S, Gersh BJ, et al. The effect of race and sex on physicians' recommendations for cardiac catheterization. *N Engl J Med* 1999;340:618-26
- Schwartz LM, Woloshin S, Welch HG. Misunderstandings about the effects of race and sex on physicians' referrals for cardiac catheterization. *N Engl J Med* 1999;341:279-83; discussion 286-7
- Spiegelman D, Hertzmark E. Easy sas calculations for risk or prevalence ratios and differences. *Am J Epidemiol* 2005;162:199-200.
- 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.
- Greenland S. Model-based estimation of relative risks and other epidemiologic measures in studies of common outcomes and in case-control studies. *Am J Epidemiol* 2004;160:301-5.
- Deddens JA, Petersen MR, Lei X. Estimation of prevalence ratios when PROC GENMOD does not converge. SUGI28: Proceedings of the 28th Annual SAS Users Group International Conference; 2003 Mar 30-Apr 2, Seattle, Washington, United States (Paper 270-28).
- Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol* 2004;159:702-6.
- Royall RM. Model robust confidence intervals using maximum likelihood estimators. *International Statistical Review / Revue Internationale de Statistique*. 1986;54:221-6.
- Lachin JM. *Biostatistical Methods: The Assessment of Relative Risk*. New York: Wiley-Interscience, 2000.
- Diggle PJ, Liang KY, Zeger SL. *Analysis of Longitudinal Data*. Oxford: Oxford University Press, 1994.
- SAS. *Stat 9.1 User's Guide*. Cary, NC: SAS Institute Inc, 2004.