Author’s response to reviews

Title: A Comparison of Two Methods for Estimating Prevalence Ratios

Authors:

  Martin R Petersen (mrp1@cdc.gov)
  James A Deddens (jad0@cdc.gov)

Version: 2 Date: 20 December 2007

Author’s response to reviews:

Dear Dr. Bucceri,

Thank you for considering the manuscript “A Comparison of Two Methods for Estimating Prevalence Ratios” by Martin R. Petersen and James A. Deddens for publication in BMC Medical Research Methodology. We have modified the manuscript taking into account your comments and those of the 3 reviewers. Our point-by-point responses are given below.

Editor

1. We strongly encourage you to include an Acknowledgements section between the Authors’ contributions section and Reference list. Please acknowledge anyone who contributed towards the study by making substantial contributions to conception, design, acquisition of data, or analysis and interpretation of data, or who was involved in drafting the manuscript or revising it critically for important intellectual content, but who does not meet the criteria for authorship. Please also include their source(s) of funding. Please also acknowledge anyone who contributed materials essential for the study.

Authors should obtain permission to acknowledge from all those mentioned in the Acknowledgements.

Please list the source(s) of funding for the study, for each author, and for the manuscript preparation in the acknowledgements section. Authors must describe the role of the funding body, if any, in study design; in the collection, analysis, and interpretation of data; in the writing of the manuscript; and in the decision to submit the manuscript for publication.

Response: Because we had no data to collect, we did the work ourselves. However, we do appreciate the excellent comments of the BMC reviewers and have included an Acknowledgements section thanking them. We also added the following sentence to the Authors’ Contributions section: “The authors performed this work as part of their official duties as employees of the National Institute for Occupational Safety and Health.”

2. Please also ensure that your revised manuscript conforms to the journal style (http://www.biomedcentral.com/info/ifora/medicine_journals). It is important that
Response: We went through the above referenced checklist and made the necessary changes.

Reviewer 1: Thomas Behrens

Major Compulsory Revisions:

1. Background: The authors discuss the modeling of a logistic regression against the background of the rare disease assumption. Two issues should be more clearly pointed out: The rare disease assumption does not apply to case-control-studies in general where the disease under investigation is either rare or the assumption is not required because controls are selected by incidence density sampling. Choosing between OR and PR may represent a problem in cross-sectional studies where the studied disease is oftentimes not rare, and the calculation of the OR may lead to false perceptions if interpreted as a relative risk.

Response: We have now modified the first sentence to correctly restrict our paper to cross-sectional studies. We mention it again later in the Background to remind the reader. We have also added the sentence “Thus, any author or reader, who considers exposure to be related to a four-fold increase in the chances of getting the disease, would be substantially overestimating the effect of the exposure.” to the paragraph discussing the consequence of using the OR when one is interested in the PR.

2. Methods: Since every article should stand on its own, the authors should present in more detail how the exact estimates were yielded instead of merely citing an earlier publications (p. 8).

Response: Following our mentioning the exact method, we added the text “This was accomplished using the macro supplied by Deddens et al.[12] Briefly the macro finds the point on the boundary, and it restricts the search for the MLE to parameters which force the likelihood through this point.”

3. The finding that in multivariable analyses the statistical significance of estimates depended on the choice of model should be further discussed with respect to the problem of stepwise regression techniques.

Response: We addressed this in response to reviewer 3’s comment #1. We show that different independent variables are necessary for logistic than for log-binomial, which would affect stepwise or any regression. The purpose of the paper is to compare the log-binomial and Robust Poisson methods, but we now mention this difference between functional forms.

4. The statement on p. 16, 1st paragraph should be modified. A model is not `incorrect` in itself (neither would be a logistic regression). The results may only give a false perception if interpreted as PR or RR. For the same reason, the authors should apply more caution when stating that results of a model were
Response: The functional form of the model may be correct or incorrect depending on how the dependent variable is related to the independent variable in the population. For real data, one may not be able to tell which form is correct, which we mention in the Discussion preceding the sentence in question. We consider the distribution together with the functional form as the model. Thus the Robust Poisson is always an incorrect model because the data can only consist of zeros and ones, which means that the distribution cannot be Poisson. However, this part of the Discussion was substantially revised based on Reviewer 3’s comment #1. The discussion of bias referred to the simulations, is for a case where the true model is known. The bias is the difference between the estimate and the true (known) parameter. We modified the sentence to make it clear that we were talking about the simulations (where the true parameter was known).

5. In accordance with the published literature, the authors recommend a log-linear binomial regression to model prevalence data. In contrast to recent publications that favor the robust Poisson, the authors recommend their own COPY method when the model fails to converge. However, we had considerable problems to apply the COPY method to real data sets with many variables since the method used considerable computational time and memory (unlike the authors’ statement on p. 17 that 10,000 copies of a data set would be a feasible approach). Therefore, we would still prefer the Robust Poisson in the case of non-convergence. It is noteworthy that in the present manuscript most results were very similar when using the one or other method and only in some malign cases, results between the two methods differed.

Response: We have modified the sentence to suggest this possible difficulty. (One thing that can help is to run the models on a data set which only includes variables in the model.) We have also recently tested a modification involving weights instead of physical copies due to Lumley et al.[new reference #31]. On the 20-30 data sets on which we tested it, the results were exactly the same as those using physical copies. This should be feasible for any reasonable number of copies. We have added text in 2 places in the Discussion concerning this.

6. Citation # 28 is a reviewer’s report for another article which does not qualify as a proper publication and should therefore not be cited.

Response: This reference has been removed.

Minor Essential Revisions:

7. Abstract: The first sentence should include in cross-sectional studies.

Response: We have now done this.

8. The statement that the article compares two of the better methods is judgmental and should be avoided.
Response: We changed “better” to “newer”.

9. Add “vaso-“ (‘20 of 39 observations were constricted”) (p.11, 2nd paragraph)
Response: We changed this in several places.

Reviewer 2: Robert Gibberd

Major Compulsory Revisions:

1. The Poisson model is claimed to be less biased for prevalence = 0.7, but this is not the case for high slope, and overall Table 1 does not readily support this claim.
Response: We agree and have changed the text.

Discretionary Revisions:

2. The binomial model provides a better model of a proportion than the Poisson, as claimed in the paper, but the link function (logarithm) may not be the most appropriate: logistic or complementary log-log may be better. However, large samples are required to distinguish which link function is best, but intuitively, the log link is not likely to be as appropriate as the other two, especially for proportions approaching 1.0.
Response: This may often be true, and in response to Reviewer 3’s comment #1, we now discuss this fact in the text as it relates to log-binomial versus logistic.

Reviewer 3: Christopher Leigh Blizzard

Major Compulsory Revisions:

General
This is a well-written and interesting paper on a statistical method that deserves to be given greater attention.

1. However I have two related misgivings.

Firstly, the authors propose to use the COPY method when the log binomial model does not converge. They justify this with reference to a small real data example in previous published work (12,23,28) but do not use the data simulations in this paper to test the properties of the COPY solution in the circumstances that the log binomial model (without the COPY modification) has not converged. That is an opportunity lost, I think.

In this paper the authors provide three additional real data examples. Each has a preponderance of data values corresponding to event (y = 1) at large values of
the predictor ($x_\ell$). These are high leverage values in the log binomial model and, in my opinion, the analyst would be well advised not to attempt fit a log binomial model to any of these real data examples. The authors state that close examination of the death penalty data indicates that either the logistic regression model or the log binomial regression model fits the data reasonably well. I examined the data, and reached a different conclusion. The logistic model does fit the data moderately well, but the log binomial model (with the COPY modification) provides a solution that is rather less satisfactory.

Is it sound, then, to propose that the COPY method be used when the log binomial model does not converge? Is lack of convergence of the log binomial model (without the COPY modification) a signal that the log binomial model does not fit the data particularly well, and that the solution that will be provided by the log binomial model (with the COPY modification) may not fit the bulk of the data very well?

Secondly, the log binomial model (with the COPY modification) may not be suitable for confounder analysis. My suspicion was aroused by the death penalty example. The authors noted that different conclusions would be reached depending on the estimation method used. The outcome of receiving a death penalty is significantly related to the race of the defendant in the results of Robust Regression ($P=0.010$), and in the results of logistic regression ($P=0.006$), but not using the COPY method ($P=0.140$).

This is unsatisfactory. On adjustment for covariates, the COPY method results in a greatly reduced estimate of the coefficient of the covariate for race of the defendant relative to the estimate of its estimated standard error. Why is this?

The table shows the results of adjusting the estimate of the coefficient of the binary (0/1) covariate for race (Black Defendant) by the scaled covariate for culpability (Culpability). The results for the COPY method are markedly different to those for the Robust Poisson of logistic methods. The estimated coefficient is reduced using the COPY method, but increased using the Robust Poisson method or logistic regression.

<table>
<thead>
<tr>
<th>Model and covariate</th>
<th>COPY (c=1000)</th>
<th>Robust Poisson</th>
<th>Logistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Univariable (unadjusted)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black Defendant</td>
<td>0.254</td>
<td>0.255</td>
<td>0.386</td>
</tr>
<tr>
<td>Culpability</td>
<td>0.474</td>
<td>0.539</td>
<td>1.116</td>
</tr>
<tr>
<td>Multivariable (adjusted)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black Defendant</td>
<td>0.192</td>
<td>0.386</td>
<td>1.277</td>
</tr>
<tr>
<td>Culpability</td>
<td>0.467</td>
<td>0.544</td>
<td>1.237</td>
</tr>
</tbody>
</table>

The correlations between the binary (0/1) indicator variable for the death penalty (Death), the binary race indicator (Black Defendant) and the scaled covariate for culpability (Culpability) are shown below. The binary race indicator (Black
Defendant) and the scaled covariate for culpability (Culpability) are negatively associated, but each is positively associated with the binary outcome indicator (Death):

Death  Black  Defendant  Culpability
Death  1.000
Black  Defendant  0.091  1.000
Culpability  0.664  -0.096  1.000

The estimated association between the binary race indicator (Black Defendant) and the binary indicator variable for the death penalty (Death) should increase on adjustment for the scaled covariate for culpability (Culpability). It does in the Robust Poisson solution, and in that of the logistic regression model. It does not in the COPY method solution.

Similarly the estimated association between the scaled covariate for culpability (Culpability) and the binary indicator variable for the death penalty (Death) should increase on adjustment for the binary race indicator (Black Defendant). It does in the Robust Poisson solution, and in that of the logistic regression model. It does not in the COPY method solution.

The Robust Poisson method (and the logistic regression model) produces the expected pattern of change in the estimated coefficient of binary indicator variable for race (Black Defendant). The solution produced by the COPY does not.

This example indicates that the solutions produced by the COPY (c = 1000) method may not be suitable for change-in-parameter-estimate confounder analyses. The COPY solutions have to be discounted in this example, and perhaps in others. The question is whether or not this is a widespread problem.

Response: We did save the estimates from the simulations. When a substantial number of the models didn’t converge, both the log-binomial (COPY) and the Robust Poisson methods gave overestimates on the models which didn’t converge. On their convergent counterparts, both methods usually gave underestimates. The simulations represent cases where the fitted model has the correct functional form (and the true parameters are known). It might be interesting to do simulations and see how often one erroneously changed the functional form of the model based on the data, both when the model converged and when it didn’t, but it is not the purpose of the simulations that we performed. We were interested in comparing the MLEs for the correct model to the Robust Poisson approximation.

The reviewer makes a good point about the fit of the models for the death penalty data. There is a curvature in the log-binomial model which is not present in the logistic model. When a quadratic term for culpability is added to the log-binomial and Robust Poisson models, the term is significant, and the same conclusions are reached by all 3 methods. We have re-written the text concerning this
example and also discuss differences between the logistic and log-binomial models.

The analgesic example also had a significant quadratic term for age, but the statistical test is probably not valid because there were too many terms in the model. We have removed some terms and re-analyzed that data, but there are still too many terms to test for the quadratic effect of age. We now mention the restriction that the number of terms should be no more than one-tenth the number of events and suggest that for a set of data, one model (logistic or log-binomial) might be more parsimonious than the other, which could be an advantage to using that model.

The death penalty data provide an example where one can fit the log-odds to a simpler model, or fit the log-probability to a more complex model. Our purpose for this paper, however, is to compare the log-binomial method to the Robust Poisson method when one decides to fit a log-binomial model. We do not believe that the lack of convergence of the log binomial model is necessarily a signal that the log binomial model does not fit the data particularly well. All of the simulations generate data which comes from a log-binomial model, but the log-binomial model failed to converge on many of these data sets. Not knowing this, an analyst might switch models if the model didn’t converge, but it would be the wrong thing to do.

We greatly appreciate the reviewer’s detective work in discovering the reason for the differences between the logistic and Robust Poisson methods compared to the log-binomial method. Our solution to this was to add a quadratic term for culpability to the log-binomial and Robust Poisson models. We include a sentence saying that confounding, as well as interactions and linearity, will be different in logistic and log-binomial models. It is not clear whether there is a problem with confounder analysis for the log-binomial method. It could be that the log-binomial model is more sensitive to the omission of a term than the Robust Poisson. Even with the quadratic term omitted, however, Culpability would probably be retained in the model because it was a confounder and changed the coefficient for black defendant (albeit in the wrong direction for the log-binomial method), or because it was highly significant. In addition the correlation between black defendant and culpability was only 0.1, so one might expect only a small change in the coefficient for black defendant. The smallest change occurs for the log-binomial method. It is an interesting topic for future study. We have added a sentence at the end of the Conclusions section suggesting this future research.

Minor Essential Revisions:

2. Some justification of the sample size of n = 100 should be provided.

Response: The following sentence was added to the Methods section: “This sample size was chosen because it was felt to be large enough for large sample properties to hold, but not so large that both methods would have power too high for comparison.”
3. I had no difficulty implementing the COPY method in SAS 9.1 software, but encountered problems in achieving the same in Stata 10 software. My implementation using Stata’s glm command provided a solution matching the one given in the paper for the vaso-constriction example data, but required very good starting values (correct to four digits) to be provided, and did not provide convergent solutions for either the neuralgia or death penalty datasets. I think that a note about possible limitations to the applicability of the COPY method is required.

Response: SAS will almost always converge for the expanded data set if one simply uses intercept=-4 as a start value, which (almost always) makes the starting parameters inside the parameter space. This apparently doesn’t work for all programming languages. In response to our letter to the editor, Blizzard and Hosmer indicate that Stata’s ml command may solve the problem, but they also say that the glm command is intended to be the first choice and that it doesn’t always solve the problem. Other software may be better or worse. We mention this and add references at the end of the Discussion.

Discretionary Revisions:

5. You use the term ‘size’ to refer to the probability of type I error. This was unfamiliar to me, and may be to readers also.

Response: We have now added a footnote to the table.

Sincerely,

Martin R. Petersen

and

James A. Deddens