Author's response to reviews

Title: A comparison of two methods for estimating odds ratios: Results from the National Health Survey

Authors:

Enayatollah Bakhshi (Bakhshi@razi.tums.ac.ir)
Mohammad Reza Eshraghian (eshraghian@yahoo.com)
Kazem Mohammad (mohamadk@tums.ac.ir)
Behjat Seifi (behins@yahoo.com)

Version: 2 Date: 11 October 2008

Author's response to reviews: see over
Dear Natalie Pafitis,
Assistant Editor

Thanks for sending our manuscript to the reviewers and also sending their valuable comments to us. We now have revised our manuscript according to their comments. We have changed the method section of our manuscript. We hope that you will now find our manuscript suitable for publication in the BMC Medical Research Methodology. Thank you very much for your prompt review.

Sincerely Yours
K Mohammad
Corresponding Author
Author's response to reviews
MS: 1347357845210737
Title: A comparison of two methods for estimating odds ratios: Results from the National Health Survey
Authors:
   Enayatollah Bakhshi(Bakhshi@razi.tums.ac.ir)
   Mohammad R Eshraghian(meshraghian@yahoo.com)
   Kazem Mohammad(mohamadk@tums.ac.ir)
   Behjat Seifi(behins@yahoo.com)

Author's response to reviews: see over
Title: A comparison of two methods for estimating odds ratios: Results from the National Health Survey

Version: 1 Date: 1 August 2008
Reviewer: Joanna Stewart

Reviewer's report:
General comments
While this paper is simply applying the method proposed by Moser and illustrating its benefits through example as done by Moser, it does add another practical demonstration of the method and exposes it to a wider audience. As such, although not novel it would be of interest to many readers. There does, however, seem to be a little tension between whether the paper is primarily intended to demonstrate the utility of Moser’s method via an example, or to report the factors related to BMI in the Iranian study. Given the journal it has been submitted to it may be more appropriate to reduce some of the detail of the results of the study.

*The reviewer is correct and we have deleted the second and third paragraphs of results. We have also deleted some of the detail of the results. For example:

- Each 1-year increase in age had at least 0.26% and at most 0.32% increase in the odds of obesity in women.
- Each 1-year increase in education had at least 0.06% and at most 0.11% decrease in the odds of obesity in women.

It is also interesting to note that the association between economic index level and obesity was greater in lower-middle level(discussion).

It could also be considered whether the explanation of Moser's proposed method, which has been taken from his paper, could be streamlined and the reader referred to Moser’s paper.

*The reviewer is correct. We have changed the method section of our manuscript and the following statements now appear in the section of the method:

Overview of WDICH
The WDICH method overcomes some of the disadvantages of logistic regression model [11]. The linear regression model can be stated as follows:

\[ y_i = x_i \beta + e_i \]

Where \( e_i \) is random error term with mean 0 and variance \( \sigma^2 > 0 \); \( e_i \) and \( e_j \) are uncorrelated so that the covariance \((e_i, e_j) = 0 \) for all \( i, j \); \( i \neq j \). Moser and Coombs supposed that the random terms \( e_i \) follow a logistic distribution and explanatory variables \( x_i \) follow a discrete uniform distribution. They provided an estimate of the same odds ratio parameter as the DICH method, but without loss of information [11]. The estimates obtained from WDICH are more efficient than those from the logistic model [11]. They also carried out an extensive simulation study to evaluate the robustness of this conclusion to changes in the distributions of \( e_i \) and \( x_i \) [11]. The reliability of these simulation results is assessed in this paper.
Major Comments
1. Some of the construction of the sentences needs to be improved. There are also many instances of the omission of small words, making the sentences hard to read.

Some examples are the second sentence of the abstract, the first sentence of the results section of the abstract, the conclusion in the abstract. These are only examples. There are many more throughout the paper.

*The reviewer is correct and the following statement now appears in the background section of the abstract:
That means the loss of information.

* The following statement now appears in the results section of the abstract:
The odds ratio estimates changed only slightly over the two methods.

*The following statement now appears in the conclusion section of the abstract:
If responses are continuous then the "without dichotomizing" method is certainly more useful than the "dichotomizing" method and leads to more precise estimation of odds ratios.

*We have changed some sentences. For example:
1. on p 4, $x_{1-j}$ has been changed to $x_{-1, j}$;
2. on p 4 “critical point C” has been changed to a lower-case “c”;
4. on p 5, “and an estimated” has been changed to “and an estimate”;
5. p 5: “the P x 1 vector” has been changed to a lower-case “p”.

It is noticeable that we have changed the method section of our manuscript.

2. I do not believe the sigma should be present as a multiplier of the denominator of equation 1.

*There is a typographical error. The reviewer is correct and we have deleted the sigma in the denominator of Eq (1).

3. As the outcome of interest in this paper is BMI it does not seem relevant to provide a test of significance of the difference in age and education in those from an urban and rural region in this paper. (table 1) It also is not appropriate to present tests of significance of the difference in obesity level in urban and rural regions for each of the levels of the other explanatory variables separately.(table
2) Both table 1 and 2 are useful to describe the population so that generalisation of the results of the study can be assessed. Given that the clinical interpretation of the results of the study is being included in this paper, the contents of table 1 and 2 are useful, without their p values. However table 2 would be more useful displaying the sample numbers in each category, rather than the numbers who are obese. The numbers obese on their own are uninterpretable. The numbers in the cells can be calculated using the % obese but it would be much simpler to quote these in place of the number obese.

* The reviewer is correct. We have deleted the both table 1 and 2. We have added the new table (table 1). We have added this sentence to the first paragraph of the Results: Distribution of age, BMI, education, marital status, economic index and smoking are shown in table 1 in order to make the data presentation complete.

4. P10 3rd paragraph. A comparison has not been made between lower middle and the 2 higher classes and so it should not be commented on. Also the difference in the estimates of the odds ratios are much too small compared to their confidence intervals to have any indication of a difference

* The reviewer is correct and we have deleted the sentence.

Minor Essential Revisions
1. There is a typo in the subscripts of the OR on page 4.

* Change made as indicated by the reviewer

2. P8 4th line - education is not in table 2

* The reviewer is correct. According to “general comments”, we have deleted this paragraph

3. P9 1st dot point – compared to non smokers – also as smoking status is categorical rather than numeric it is not sensible to refer to an inverse relationship with smoking status

* The reviewer is correct and the following statements now appear in the first paragraph of the P9: Non smoker women were more obese than smokers. Obesity odds ratio was 0.69 (95 percent CI: 0.553-0.856) for smoker women compared to non smokers.

4. P10 – 1st line - more obese

* Change made as indicated by the reviewer

Discretionary revisions
1. P7 second to last sentence under statistical methods – would be clearer if it
said confidence interval rather than just interval.

* Change made as indicated by the reviewer

2. Table 3 may be easier to read if the odds ratios calculated in the 2 different ways were under separate heading rather than with the dichotomised one in brackets. It would be helpful if the direction (which was the numerator and which the denominator) was given for the CI ratio. If p values are to be quoted anywhere this is the table where they would be informative (from the combined model with the one for economic status for the combined variable, not for the individual contrasts). It probably would be helpful to state the comparator for the categorical variables in this table too.

- Change made as indicated by the reviewer (new table 2).
Most health surveys collect both categorical and continuous measures; very often, the continuous measures are dichotomized so that odds ratio measures of association can be calculated and reported. For example, obesity is measured using BMI, a continuous trait, but we usually dichotomize BMI at 30 kg/m² as a working definition of ‘obese’.

In this manuscript, the authors use data from the National Health Survey of Iran to demonstrate the advantages of a recently published novel method of estimating odds ratios directly from the continuous underlying trait. The authors use relative confidence interval width to show the gain in precision from using the novel method.

First, it should be noted that there is no new methodology in this manuscript. Most of the methods section (p 4-6) is copied almost verbatim, with attribution, from the seminal paper of Moser & Coombs (Stat Med 2004; 23: 1843). The authors then use these methods, and the more traditional methods based on dichotomization, to estimate odds ratios for obesity among Iranian women from the 1999-2000 health survey.

The results, unsurprisingly, reveal precision advantages for the Moser & Coombs method, which is the objective of this manuscript to demonstrate. I think this paper is a neat and tidy application of an important new biostatistical method, and deserves publication as a short report.

I am not sure of the value of lengthy repetition of formulae and results from the original Stat Med paper, particularly since there are some typos and errors which invalidate the algebra. I list the ones I spotted here:

1. the first # in the denominator of Eq (1) does not appear in the corresponding Eq (2) of Moser & Coombs (p 1845);

   * The reviewer is correct and we have deleted the sigma in the denominator of Eq (1).

2. on p 4, x subscript (1-,j) should read x subscript(-1, j);

   * Change made as indicated by the reviewer.

3. “critical point C” on p 4 should have a lower-case “c”;

   * Change made as indicated by the reviewer.
4. there is a prime missing from the very last regression equation on p 4;

* Change made as indicated by the reviewer.

5. on p 5, “and an estimated” should read “and an estimate”;

* Change made as indicated by the reviewer.

6. p 5: “the P x 1 vector” should have a lower-case “p”.

* Change made as indicated by the reviewer.

It is noticeable that we have changed the method section of our manuscript.

I think the manuscript could be improved by excising all this algebra, simply by citing Moser and Coombs, and by presenting the authors’ own valuable application of the novel method and demonstration of its superior precision. The authors may also want to comment on the interpretation of the two odds ratios they calculate.

*The reviewer is correct and the following statements now appear in the section of the method:

**Overview of WDICH**

The WDICH method overcomes some of the disadvantages of logistic regression model [11]. The linear regression model can be stated as follows:

\[ y_i = x_i \beta + e_i \]

Where \( e_i \) is random error term with mean 0 and variance \( \sigma^2 > 0 \); \( e_i \) and \( e_j \) are uncorrelated so that the covariance \( (e_i, e_j) = 0 \) for all \( i, j \neq j \). Moser and Coombs supposed that the random terms \( e_i \) follow a logistic distribution and explanatory variables \( x_i \) follow a discrete uniform distribution. They provided an estimate of the same odds ratio parameter as the DICH method, but without loss of information [11]. The estimates obtained from WDICH are more efficient than those from the logistic model [11]. They also carried out an extensive simulation study to evaluate the robustness of this conclusion to changes in the distributions of \( e_i \) and \( x_i \) [11]. The reliability of these simulation results is assessed in this paper.

Does the Moser & Coombs method assume a linear-logistic relation between the exposure variable (eg age) and the outcome (BMI)? What are the consequences of using M&C method if the underlying continuous relationship is not log-linear?

*No. This method doesn’t assume a linear-logistic relation between the exposure variable and the outcome. The method assumes a linear relation between the explanatory variable and the response.