Reviewer’s report

Title: Assessing the risk factors before pregnancy of preterm births in Iran: a population-based case-control study

Version: 1 Date: 18 Mar 2017

Reviewer: Hassan Ba'Aqeel

Reviewer's report:

This manuscript may have a message that is not coming across because of English language issues. The manuscript was drafted in Persian and translated into English. A reviewer struggles to understand the text (within the framework of scientific methods and scientific reporting). The understanding may not be what the authors intended. This reviewer is of the opinion that authors should not be penalized for reporting in a none native language. However, the authors are strongly advised to revise the manuscript with the help of English language scientific editor (a colleague or professional proof reader).

I will use the title of the manuscript as an example of how the English usage may have distorted the intention of the authors.

"Characteristics and risk factors pre pregnancy of preterm births in some provinces of Iran"

The title indicate that the article reports on 1) characteristics and 2) risk factors of preterm birth. The manuscript did not include any data on the characteristics of preterm birth (PTB) such as severity by gestational age, medically indicated versus spontaneous, proportions with anomalies, NICU admission etc. The term "pre pregnancy" seems to refer to the risk factors. Is the intention of the authors is to assess risk factors before pregnancy? One of the factors assessed is "supplement consumption during pregnancy"! Do the authors mean "prenatal" instead of "pre pregnancy"?

Based on my understanding of the current text, below are some of the issues that need to be clarified:

* The study subjects were recruited from mothers referred to "Health care centers". A description of what these centers mean in term of services provided. For example, do they provide delivery and neonatal care services? The term means different things in different countries.

* I quote from page 5, lines 71-72 "The case group was defined as mothers with history of preterm birth whereas the control group included mothers without history of preterm birth."
Elsewhere in the manuscript it is mentioned and I quote "Data were collected through interview, line 73. It is not clear wither mothers with history of PTB i.e. not pregnant now and had PTB in the past, were identified in centers and interviewed about their past PTB or the authors meant mothers who delivered in the centers and the birth was preterm? The implication of significant bias is obvious. In page 6, lines 108-109 and I quote "Preterm birth was the outcome variable which was ascertained through questioning the exact gestational age at the time of birth", case definition is the sine qua non of a case-control design. In this manuscript, there is a cut off value for gestational age to classify a subject as PTB/case or term birth/control. How did the authors ascertained gestational is missing from the manuscript? This is a major drawback.

* The so called "cluster sampling" is as follows; out of 31 provinces 9 were selected, out of each province four cities were selected i.e. 36 cities. In each city four centers were selected i.e. 144 centers. From lines 82-88 in page 5, I understand that ten cases were selected from each center and controls were selected from a random sample of mothers referred to the centers. Therefore, with 1:3 ratios, there will be around 40 subjects from each centers giving a sample size of 40x144= 5760 subjects. While, total sample size calculation was stated as 740. The authors elected to add more subjects to be able to test 14 independent variables in the regression model and gave a final sample size of 1300 subjects. The analysis was reported for 2463 subjects. I believe that these inconsistencies may be a reflection of an English language issues. This does not serve the efforts of the authors well.

* In developing the multiple logistic regression model, it is important to describe how the independent variables were selected to be entered into the model? This is missing from the manuscript.

* Although mentioned in the methodology section, goodness of fit of the regression model is not reported. Also, the predictive power of the model was not reported.

* In the tables row percentages were given where columns percentages is more informative to the reader.

**Are the methods appropriate and well described?**
If not, please specify what is required in your comments to the authors.

No

**Does the work include the necessary controls?**
If not, please specify which controls are required in your comments to the authors.

No
Are the conclusions drawn adequately supported by the data shown? If not, please explain in your comments to the authors.

Unable to assess

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review? If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I recommend additional statistical review

Quality of written English
Please indicate the quality of language in the manuscript:

Not suitable for publication unless extensively edited

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons
CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal