Author’s response to reviews

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

Authors:

Maryam Soltani (maryam.soltani52@yahoo.com)
Hamid Reza Tabatabaee (tabatabaee@sums.ac.ir)
Shahin Saeidinejat (saeedish1@mums.ac.ir)
Marzieh Eslahi (marzieh.eslahi@gmail.com)
Halime Yaghoobi (Lemen_2012@yahoo.com)
Ehsan Mazloumi (Ehsan.mazloumi1@gmail.com)
Abdolhalim Rajabi (rajabiepid@gmail.com)
Ali Ghasemi (aria_201323@yahoo.com)
Naeimeh Keyghobadi (naeimehkeyghobady@yahoo.com)
Mostafa Enayatrad (mostafaenayatrad@gmail.com)
Abed Noori (Abed.noori@gmail.com)
Seyyed Aliasghar Hashemi (hashemishiz@gmail.com)
Fatemeh Zolfizadeh (fa.9712@yahoo.com)
Sepideh Mahdavi (S.mahdavi88@gmail.com)
Tannaz Valadbeigi (Tanazvaladbeigi@gmail.com)
Koorosh Etemad (etemadk@gmail.com)
Ali Taghipour (Taghipour1@mums.ac.ir)
Cirruse Salehinasab (Salehinasab@yahoo.com)
Mahmoud Hajipour (m.hajipour.13@gmail.com)

Version: 2 Date: 18 Apr 2017
Author’s response to reviews:

We would like to thank both reviewers for their insightful comments on the paper, as these comments led us to an improvement of the work. Our revisions reflect all reviewers’ suggestions and readers’ comments. Detailed responses to reviewers are given below. We marked the changes in yellow type script in our new submitted manuscript.

Answer to dear reviewer Isaiah Awintuen Agorinya

Comment1. In the table 1, column percentage should be considered to make comparison between cases and controls possible for variables.

Answer: it is corrected and highlighted

Comment 2. There is no proof of use of chi-squared test in the tables as stated in 112-113.

Answer: univariate and multivariate logistic regression model was used. It is highlighted in the text

Comment 3. Line 113, ".....the association between ....". the T should be capitalized

Answer: it was corrected and highlighted in the text.

Comment 4. From Table 2. The odds ratios for the variables that were not significant in the multivariate logistic model have not been shown in the table. It is important to have these results displayed in the table as part of the output from the logit models.

Answer: insignificant variables were omitted and significant variables are shown in a table.

Comment 5. Description of demographics characteristics have not been included in the results section. Column percentages should be computed for the demographics and the results discussed.

Answer: results of descriptive analysis are discussed in the results. Table of results is corrected and highlighted.

Answer to dear reviewer Hassan Ba'Aqeel
Comment 1. 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 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"?  
Answer: we appreciate your comment on the title and changed it to: Assessing Risk factors pre pregnancy of preterm births in Iran. Some risk factors such as severity gestational age, medically indicated versus spontaneous, proportions with anomalies, NICU admission have not been assessed yet. The word pre pregnancy is used according to some reasons: in Iran acid folic is prescribed 3 month before pregnancy for women who intend to become pregnant, literature review on previous studies on preterm birth, maternal factors affect the occurrence of still birth. On the other hand supplement consumption was omitted from the model since it was not associated with stillbirth after controlling for confounders.

Comment 2. 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.  
Answer: Health care center is defined in the text and highlighted.

Comment 3. 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 the authors ascertained gestational did is missing from the manuscript? This is a major drawback.
Answer: we agree your idea, that if it was aimed to assess the history of preterm birth the study would be subjected to recall bias.

The definition of case and control group is corrected and highlighted in the text. The case group was defined as women who had preterm birth in recent pregnancy and the control group was defined as women who had full-term birth in recent pregnancy. This study was conducted in 2016. According to literature review and CDC, preterm birth was defined as delivery before completing 37 weeks of pregnancy. Delivery between the 37 – 42 weeks of pregnancy were defined as control group. Data were collected based on mothers health record, if it was insufficient interview with the mother was carried out.

Comment 4. 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.

Answer: Study population were recruited through multistage cluster sampling method. In the first stage regarding geographical divisions of Iran, 9 clusters (provinces) were randomly selected. In the second stage, in each of the 9 clusters (provinces) 4 clusters (cities) were randomly selected from the north, the south, the east, the west and the central area. In each city 2 health care centers (1 urban and 1 rural healthcare centers) were randomly selected. In each center 40 questionnaires were filled, 10 for cases and 30 for controls. Sample size was determined using mothers age more than 35 as an independent risk factor, 80% power of study and \( \alpha = 0.05 \). Total sample size was 700 for cases and 2100 for controls. (The 1:3 ratio was considered to improve power of study.). Incomplete questionnaires were ignored so the final sample size included 688 cases and 1795 controls. At the end of study power analysis was carried on data and final power was higher than 80%. Regarding high prevalence of preterm birth in Iran it seems possible to have larger sample size. However increasing sample size could be associated with information bias resulted from decreased accuracy of data and other administrative problems. We are not aimed to compare provinces from which data are collected.
Comment 5. 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.

Answer: all variables with p<0.03 (Kleinbaum book) were entered in to the multivariate logistic regression model to control for confounding effect.

Comment 6. 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.

Answer: Hosmer lemeshow test for goodness of fit was used. The value 0.52 indicated goodness of fit for the model. This value is highlighted in the text.

Comment 7. In the tables row percentages were given where columns percentages is more informative to the reader.

Answer: corrections are made and highlighted.