OPEN PEER REVIEW

Open Peer Review (OPR)

As a new feature for progressing towards transparency, we decided to open a new window for all of our editors as well as authors titled "**Peer Review**". We hope by this new facility, our reviewers will be more motivated and authors will be more satisfied with the review process. We believe that publishing our peer review reports could make a transparent and clear environment for all our efforts within a journal, but not all reviewers tend to publish their comments.

What is "Open Peer Review" process?

An "Open Peer Review" process is making the details of all review process (including reviewers, associate editors, and EICs comments) as "Public" as it is agreed by EIC, Authors, and reviewers.

Advantages of "Open Peer Review" process

- More transparency, constructiveness and tactful comments of the peer review process: leads to an increase in the quality of reviews
- More motivations for all involved roles in the review process
- Authors' satisfactions from the review process: Increases honesty between authors and reviewers
- Education of both authors and new students
- Prevents reviewers from following their individual agendas and leading to the detection of reviewers' conflicts of interests

You can find out more at:

Peer Review

Journal of Comprehensive Pediatrics
Article DOI: https://doi.org/10.5812/compreped.108575
Published in: Journal of Comprehensive Pediatrics: 12(1) ; e108575

Peer Review Report for "Predictive Factors of Infant Mortality Using Data Mining in Iran"

Author(s): Mahmoud Hajipour, Niloufar Taherpour, Haleh Fateh, Ebrahim Yousefi, Koorosh Etemad, Fatemeh Zolfizadeh, Abdolhalim rajabi, Tannaz Valadbeigi, Mehrabi Yadollah

Review Timeline:

Submit Date:	16 Aug 2020
Revised Date:	6 Jan 2021
Accepted Date:	12 Jan 2021

Revision (0)

Here, you can see the **Reviewers**, **Associate Editors** and **EICs'** comments from the beginning to the end of the revision process.

REFEREE : Reviewer — 1 Sep 2020

Dear AE,
Some comments are written in the text and the article is attached.
The authors should pay attention.
Thanks.

REFEREE : Reviewer — 13 Dec 2020

Dear AE,
The study is reviewed carefully. According to comments and revisions by the author,

this manuscript is accepted.
Thanks

REFEREE : Associate Editor

17 Dec 2020

Dear EIC,
The manuscript is studied, and there are some comments which should be addressed by the author. According to the comments, the document needs major revision.
Kind regards,

REFEREE : EIC

17 Dec 2020

Dear Author,
The reviewers' comments are mentioned. Please correct the details as mentioned above carefully according to reviewers' request.
Thanks.

OPEN PEER REVIEW

Revision (1)

Reply to Reviewers

Ideally, the reviewing process can significantly improve the submitted manuscripts by allowing the authors to take into account the advice of reviewers. Author(s) must reply to all reviewers' comments in a separate Word file, point by point. A "**Reply to Reviewers**" document is submitted along with revised manuscript during submission of revised files, summarizing the changes that the authors made in response to the reviewers' comments. The responses to reviewers' comments specifies how the authors addressed each comment the reviewers made.

You can read the authors' responses to the reviewers' comments in the next page.

4 Jan 2021

Dear Editor-in-chief

Journal of Comprehensive Pediatrics

Greetings

With many thanks to the reviewers and the associate editor for their thoughtful critiques of our manuscript (ID: 108575), entitled "**Predictive factors of infant mortality using data mining in Iran"** the revised version of the manuscript is attached for your consideration.
We have considered comments and suggestions raised by the respected reviewer and edited the manuscript accordingly. Point-by-point response to comments are detailed below. The edited manuscript is submitted along with this letter.

Yadollah Mehrabi

Corresponding author

Department of Epidemiology, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran
Email: mehrabi@sbmu.ac.ir Tel: +98(21)22431993 Fax: +98(21)22432036
Address: School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Daneshjoo Blvd, Evin, Tehran, IR. IRAN. Postcode: 1983535511.

Detailed responses to reviewers comments on the manuscript:

1. What has been the specialty and field of work of the authors?

Response: The authors of this manuscript are in the field of Biostatistics, Epidemiology, health information technology and health care management science.

2. What was the reason for using this method and what is the advantages of this method?

Response: Nowadays data mining methods are widely used in analysis of health information. Researchers use modern models of data mining approaches such as multi-dimensional databases, machine learning, soft computing, artificial neural network (ANN), support vector machine (SVM), data visualization, etc. beside the traditional statistical methods such as logistic regression. Off course, if an accurate model with high ability is used in data analysis, the results can be more reliable for policy makers and planners to improve the health system. It should be kept in mind that applying such models requires appropriate skills. Since reduction of child mortality is an important National and International goal, we employed data mining techniques to assess the risk factors of infant mortality.

3. What was the reason for choosing these 8 provinces?

Response: In this stratified cluster random sampling, 8 out of 31 provinces from North, East, West, South and center of Iran were selected. Then 5 districts were chosen again from North, East, West, South and Center of each selected province. Thereafter we randomly selected one urban and one rural health center from each sampled district.

4. How was the sample size determined? How did people fit into these two groups?

Response: According to literature review, considering mother's education <5 years as a risk factor (P_0 = 0.58, P_1 = 0.68, $Z_{0.95}$ = 1.96, $Z_{(1-\beta)}$ = 0.90, design effect = 2) [1] and using sample size formula for comparison of proportions, the sample size was estimated as 508 for each of case and control groups. By taking design effect=2 into account, sample size for each group reached to 1016. We added some more samples in order to replace the probable incomplete data. The sample size was sufficient considering > 80% power of tests.

1.Chaman R, Alami A, Emamian MH, et al. Important Risk Factors of Mortality Among Children Aged 1-59 Months in Rural Areas of Shahroud, Iran: A Community-based Nested Case-Control Study. *Int J Prev Med*. 2012;3(12):875–879. doi:10.4103/2008-7802.104859

After coordination and obtaining the consent of Universities of Medical Sciences and affiliated health centers, we gathered data from health records about infant mortality. The case group included mothers who had lost an infant due to disease or congenital disorder. The control group included mothers who had at least a live child with the age of less than one and had not experienced under one-year-old child death so far. Subjects in control group were randomly (based on code of health records) selected from mothers referred to health centers for their children's vaccination (page 4, lines 78 to 86)

5. Which titles have been collected by these check lists? If the checklist is standard, by whom, in what year and how was it designed? How was demographic information collected?

Response: Checklists included data of demographic status and disease or disorders history of mothers and their children (page 4, lines 85 to 86). The checklist was created by researchers. Its items were extracted from related published literature and past reports. Then discussed by the research team before embedding each variable in the checklist. In fact, it is not a published standard checklist, so we changed the word "standard" to "researcher made" checklist in the manuscript (line 84, page 4).

6. Which items have been asked about interviews? Under what conditions were the interviews conducted?

Response: Participants with poor reading or writing skills were interviewed face to face by a trained interviewer to complete the checklist. This issue was added to the manuscript (lines 83 to 85, page 4).

7. Why are formulas written?

Response: We have added the formulas for easier understanding the indices by readers; but we do not insist on this, so the formulas can be removed if the respected referees/editors believe that they are redundant.

8. When is the outcome variable expected to occur?

Response: This was a case control study, so the outcomes have been occurred prior to this study. Their past exposure to suspected factors were compared to those without outcome (control group).

9. There are instruments that measure mortality in infants and have been translated in Iran and the psychometric properties of these instruments have been confirmed. Like "SNAPPE-II Score System" and "Clinical Risk Index for Babies (CRIB-II) Scoring System" To strengthen the discussion, the authors can refer to these tools and factors predicting infant death.

Response: With many thanks to the respected reviewer suggestions, we have not used those instruments in our study since SNAPPE- II score system have been designed for neonatal intensive care unit (NICU) and CRIB-II is also for clinical use, while our study has carried out in health centers.

10. In the discussion, the results of the studies are mentioned, which is in line with the results of this study, but there is no conclusion from this alignment. What has been the innovation of this research?

Response: The innovation of the current study is to apply modern methods of data mining in detecting important risk factors of infant mortality. The set of factors that have been found in this study are somehow different from that of traditional statistical analysis such as logistic regression. This is mentioned in lines 206-209 of the manuscript.

11. What were the limitations of the study? What suggestions are there?

Response: Missing data due to incomplete checklist and response bias due to unwillingness of participants to report status with strong social stigma among women such as smoking were the limitation of this study. This paragraph is added to the end of discussion. (page 8, lines 210 to 212).

12. What weight (table 1)?

Response: Definitions of Normal, low and high birthweight were added to table 1and 3.Normal weight at birth : 2500 -4000 g, Low birth weight : <2500 g and high birth weight >4000g.

13. This division has been done on what basis? (age of getting pregnant)

Response: Categorization of pregnancy age groups are based on national protocol given by the Health deputy.

14. Did they have a consanguineous marriage or not?

Response: Yes, 33.6 and 25.6 percent of the case and control groups had consanguineous marriage, respectively. This was corrected in tables 1 and 3.

OPEN PEER REVIEW

Revision (1)

REFEREE : Reviewer — 11 Jan 2021

Dear AE,
It's a good article, and with its novelty, it could be the basis for future research.
Thanks.

REFEREE : Associate Editor — 12 Jan 2021

Dear EIC
Based on the authors' answers to the reviewers' comments, the reviewer(s) have recommended publication.
Kind regards.

REFEREE : EIC — 12 Jan 2021

Dear Author,
The document is reviewed carefully by the reviewers. According to their comments and your revisions as requested, this manuscript is accepted.
Thanks.