Author's response to reviews

Title: On exploring and ranking risk factors of child malnutrition in Bangladesh using multiple classification analysis

Authors:

Kakoli Bhowmik (kakolistat@hotmail.com)

Sumonkanti Das (sumon_148@yahoo.com)

Version: 1 Date: 04 Dec 2016

Author’s response to reviews:

Reviewer reports:

Reviewer #1:

A good manuscript on an interesting subject, please consider the followings:

1. Please remove abbreviations from abstract, also all abbreviations all over the manuscript should be revised as some of them are confusing.

Response: In the abstract, MCA is replaced by multiple classification analysis. Regarding revise of the abbreviation, an example would be good to understand. However, we have tried to be consistent using the abbreviation. GLM is used for “Generalized linear model” and a new abbreviation is added for Binary logistic regression model. BLogM is used for Binary logistic regression model to avoid confusion. The whole paper has been revised by replacing GLM by BLogM accordingly.

2. The study methodology is NOT complete. The inclusion and exclusion criteria should be more clear.

Response: It would be good with some example to revise. We are apologizing not to understand about inclusion and exclusion criteria. Later we mentioned why we elaborate the methodology section.

3. The results and discussion are good, except some English errors and please separate from each other in 2 sections.

Response: Regarding English correction, we tried our best.
Since “Conclusion” is the last section, we tried to provide the results with their discussion in one section. If it is compulsory for publication, we will separate them.

4. Many sections in introduction and discussion need to be summarized by 50%

Response: We tried to do this in discussion part instead of introduction part. According to second reviewer we need to provide more explanation in discussion. So we tried to adjust both the reviewers’ comments.

5. The English of the manuscript needs revision. There are many grammatical and syntax errors in the manuscript.

Response: We have tried to correct the grammatical errors where we found. Also reviewer 2 indicated some errors and we have corrected them.

Reviewer #2: General comments

* This is an interesting and useful study as many authors have conducted logistic and linear regression analysis to identify predictors of stunting/under nutrition. The MCA analysis presented her presents a potentially useful statistical technique to order the most important predictors of undernutrition.

The overall purpose of the paper should be more clear. A hypothesis or at least a purpose statement should be included in the introduction. It seems that the goal of the article is to compare MCA, linear and logistic regression analysis approaches for identifying predictors of malnutrition, and 2) to argue that MCA is a preferential method for identifying predictors. The conclusion does state the author's main point and the study's purpose…This purpose needs to be more clearly presented at the outset of the paper.

Response: According to the reviewer’s comment, I added a line at the end of the introduction section as below:

“Thus major goals of the article are (1) to compare linear and logistic regression analysis approaches to MCA approach theoretically and empirically for identifying predictors, and (2) to show how the MCA approach provides some additional information over the other approaches for identifying predictors of malnutrition.”

* The authors do not adequately refer to and explain their tables in the results sections. These are complex and need more thorough explanation.

Response: Now we refer all tables and explains clearly in the corresponding paragraphs of Results and Discussion section. Please see in the revised version with correction.
* The table footnotes or parentheses should explain more of what the predictor variables represent. For example, family size small, middle, and large should indicate the # of people that classify households in each category.

Response:

Table 1: “Child’s weight” is replaced by “Child’s birth weight”

Table 2: “Family size” definition is included in Table 2 as Table 1.

* The percentage correct outcome in table 3 is hard to understand. What is the gold standard by which this is being corrected?

Response: There is no gold standard but it is expected the higher the proportion the higher the goodness-of-fit. The table shows a comparison between the observed and predicted values and shows more than 65% cases are predicated by the fitted LM and BLogM models. These are clearly described in methodology section and referred the article [27] Hosmer and Lemeshow (2000).

* Table 4 seems superfluous and unrelated to the main objective of the paper.

Response: This Table is provided for supporting the opposite results found due to urban-rural residence of the children. If you look at the results in Table 1 and Table 2 for the variable “Place of residence”, you find that the unadjusted results such as UPM is higher (UPP is lower) for Urban children but its adjusted value APM is lower (APP is higher). It is expected that the urban children will have better access to proper health facilities but when other explanatory variables are combine, the expected results are reversed. This is the main reason of showing the Table 4.

* The tables need footnotes. The sample size should be noted in each table and significant results should be marked with a symbol.

Response: Sample sizes are given in the footnote of Table 1 & 2. Significant results are also marked with one, two, three stars’ symbol in Table 1 & 2. In Table 5, the DFs of F-statistics also show the sample size. So there is no need to mention again the sample sizes in Table 5.

Finally, please make the public health significance of this paper more clear.

Response: The aim of this paper is to show how the LM, GLM and MCA approaches can be implemented for identifying predictors of child malnutrition. We focus MCA would be a better approach for getting the answers what are also available from LM and GLM model. Thus the MCA approach can be easily implemented in other public health related research (such diarrhoea prevalence), particularly dealing with either interval or nominal scale response variable. A line is given at last of the conclusion section.
“The application of MCA approach and its findings in this study suggest that MCA (with/out LM or BLogM) can be used in other public health studies (such as episode of diarrhoeal diseases or acute respiratory infection) with the aim of exploring and ranking the risk factors.

Specific comments

* The conclusion of the paper in the abstract is confusing. Do they authors mean that I order to use MCA they must be relevant to policies.

Response: We tried to say the policy makers can order the possible significant predictors of child malnutrition via MCA and then prioritize their interventions according to the order.

* Recommend a slight revision to ensure suitability for publication in English. E.g. in line 39 "the developing" should just be "developing"

Response: Revised as “the world’s developing nations” instead of “the developing world”.

* Line 49 - four-fifth needs to be four-fifths

Response: Revised as per suggestion

* Lines 11-12 on page 5. What do the authors mean by Bangladesh till having a high prevalence of chronic malnutrition? Compared to whom?

Response: According to WHO threshold of “high prevalence: 30-39%” of stunting. Thus a line is added with the line: “according to WHO threshold of “high prevalence: 30-39%”.

* Line 48-50. The authors list refs 14-20 (a lot of references) simply to say that logistic regression is used to determine risk factors for stunting.

Response: These references are used for several purposes: uses of logistic regression, dichotomous response variable and selection of explanatory variables. We are interested to show these references and so we are requesting to keep them.

* Page 4, line 11-12 malnourished, not malnourish. Response: Revised as per suggestion

* Line 43-44 "say" should be e.g. Response: Revised as per suggestion

* The introduction is a written informally, e.g "the rest of the article is organized…” should not describe what is expected, but should introduce the objectives of the article. The introduction should set up the need for the study and not serve as a table of contents for the paper.
Response: The paragraph is for suitability of the readers not given as content of a book. Also it is a standard format for a paper. So we wish to keep it. These lines are given after focusing on the goals of the paper.

* The methods section begins with an explanation of a statistical model. This section needs more revision to include sub-headings such as "Data source", "Statistical models", etc.

Response: Study materials are given for “Data Source”. The statistical methodologies are described separately for the ease of readers.

* The methods section is very long and not suitable for a general nutrition audience as it includes a large amount of statistical jargon. It would be helpful to clarify the study question that the authors sought to answer. Was the objective to compare results from MCA and logistic and linear regression? If so, to what end and for what purpose? This should be made explicit in the introduction or the end of the methods section. Did the authors have a hypothesis?

Response: The main reason for more illustration of the three regression approaches (LM, BLogM, MCA) is to show how they are developed under what assumptions, and what are the interpretation of the regression coefficients. How the MCA models are providing almost similar results of linear model and logistic model are illustrated in the methodology part and then these are attempted to prove numerically in the results section as well.

It is aimed to clearly describe (in simplified way) all the 4-types of model (LM, BLogM, IS-MCA & NS-MCA) to the target readers who are basically not pure statistician. Since the purpose is to compare these models, we tried to focus the main properties including types of response variable, assumptions, model fitting techniques, model diagnostics, and interpretation of model parameters. To avoid the statistical jargon, all the three-types of models are separately described with mathematical notation.

* Page 14, line 32- research is not plural.

Response: “researches” changed by “research”.

* Create separate results and discussion. The authors are suggesting reasons for their results here, which seems inappropriate. Please simply state the results and discuss them separately.

Response: Since conclusion is another section, we tried to provide the results with their discussion in one section. We tried to provide possible reasons of the findings with results section.

Considering the note of the reviewer, we removed the reasons from text due its irrelevance with the objectives of the study. The deleted parts are shown in the revised file with correction.

* Page 17, line 53- should say "and" instead of "but" because a higher z score and lower probability of stunting are the same thing.
Response: Changed as per suggestion.

* Table 2 includes child's weight. I believe this should be birth weight.

Response: Table 1 & 2: “Child’s weight” is replaced by “Child’s birth weight”