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: 2 Date: 02 Jun 2017

Author’s response to reviews:

Dear Pierluigi Marzuillo
BMC Nutrition

Subject: Submission of Revised article NUTN-D-16-00114R1

Dear Sir/Madam,

Many thanks to the reviewer and the editors for their valuable comments and directions. We have followed the reviewers’ directions to modify and revised the our manuscript "On exploring and ranking risk factors of child malnutrition in Bangladesh" (NUTN-D-16-00114R1). The responses are given below with the comments of the reviewers.

Reviewer reports:

Scott Ickes (Reviewer 2):

The introduction of the paper is overly long and has an unnecessary explanation of regression methods; this is a widespread method used in public health/nutrition research and the specific model construction, but not purpose, can be described in the methods, not the introduction. Also, I agree with an earlier reviewer who suggested to separate results and discussion. As this is a quantitative article, it is non-convention to keep these as one section, even with a separate conclusion section.

Responses: (1) We have tried to reduce the Introduction section as per suggestion. However, I briefly describe the problems of the linear and logistic regression models and then highlights the MCA method to focus my research goals. I detail the methods in methods section and consider the reviewers’ suggestion when I squeeze them. (2) We have prepared two sections Results and Discussion, and reduced the Conclusion section into one paragraph.
In general, the writing style is still informal in some places, even after the review. While generally written with good grammar, there are a few places where grammar needs revision.

Responses: We tried to convert the informal sentences to be formal. As for example, “Note” is replaced by “It is noted that”, “see” is replaced by “please see”, and so on.

The model is explained in the introduction, but this is methods material. Further, the "nourish" and "malnourish" language is unclear. Specify the height-for-age Z score cut point.

Responses: (1) We have reduced the explanation of the models in Introduction section, and explain the methods as brief as possible in the Methods section. (2) For reducing confusion on the child malnutrition definition, we put a line in the fifth paragraph of the Introduction Section - “The most commonly used anthropometric measures are height-for-age Z-score (HAZ), weight-for-age Z-score (WAZ), and weight-for-height Z-score (WHZ), which are on interval scale. A child is considered as malnourished if any of his/her anthropometric index is less than -2.00 standard deviation (SD).”

The earlier reviewer comment about inclusion and exclusion criteria is very important and was not addressed. This means that the authors need to specify how they selected their final sample from the original DHS full dataset. Were only last born children included?

Responses: We have tried to follow the Demographic and Health Survey (DHS) technique to consider the children aged under five at the interview date. They were not restricted according to their birth order. In DHS, the children who are aged under five and available in the surveyed household are considered for getting anthropometric information particularly height and weight. Among the children, some were absent during the survey, some refused to provide information, some measurements were not plausible. We provide this information in the Study Materials subsection as below:

“In total about 8281 children under age five at the interview date were selected for measuring height and weight, however measurements were completely collected from 7826 children (few were absent or refused to provide height and weight), of which 7647 children had plausible anthropometric information for calculating anthropometric measures. As the 2011 BDHS data, the same children data has been utilized in this study.”

Per earlier reviewer comments, the tables need footnotes and better headings. In the current revision, interpretation of Tables 2-4 is very difficult as standalone tables.

Responses: The Table captions are modified to make them standalone. These are discussed in later comments.

The paper lacks a clear conceptual framework. What was the theoretical approach? Why were the variables used actually selected? The authors provide some clarity towards the end of the discussion in the ranking of factors associated with malnutrition, but the concept of immediate (child), household, and maternal level variables. It is not clear what this study adds beyond what other published studies about risks factors for malnutrition in Bangladesh have demonstrated.
Responses: (1) The study mainly shows how can we identify and then rank the significant risk factors, which are already found as the risk factors in previous child malnutrition studies of Bangladesh, via multiple classification analysis. So, we think to select the common risk factors of child malnutrition based on the studies of Bangladeshi under five children, instead of searching more potential explanatory variables. We believe we cover most of the significant predictors of child malnutrition. (2) The study adds beyond other published articles on risk factors of child malnutrition is that how the risk factors are significantly associated with both nutrition anthropometric measures (HAZ in this study) and nutrition status (either HAZ < -2.00 or HAZ ≥ -2.00) via two multivariate statistical techniques. As far we know there is no study to deal both types of dependent variable simultaneously for identifying the risk factor of child nutrition. In addition, the article will encourage the researchers to rank the risk factors after identifying them as significant contributor to child malnutrition.

The conclusions do provide some clarify towards the purpose and unique contribution of the study comparing risks factors using the logistic and MCA ranking), but the purpose of the study should come earlier in the methods. Further, the authors need to explain how their findings extend, contradict, or corroborate with the other findings that have examined a similar question in Bangladesh.

Responses: (1) The objectives and the contribution of the paper have been highlighted before the Methods section as below:

“The aim of this study is to show how the MCA can provide similar results as linear and logistic regression analyses and in addition can rank the predictors considered in the model specification. More specifically, the major goals of this article are (1) to determine risk factors of child malnutrition considering both interval scale HAZ and the nominal scale nutrition status variable as response variables (2) to compare linear and logistic regression analysis approaches to MCA approach empirically for identifying predictors of child malnutrition in Bangladesh, and (3) to show how the MCA approach provides some additional information over linear and logistic regression approaches.”

(2) In the Discussion section we have explained how the findings of this study contradicts and corroborate with the other studies related to child malnutrition. The main aim of this study is to show how the MCA can provided some extra information over the traditional approach of exploring the risk factors only. Also, we provided a line given below at the last paragraph of conclusion, how the idea of the article can be used in other public health related issues.

“The application of MCA approach and its findings in this study also suggest that MCA (with/out LM or BLogM) can be used in other public health studies like episode of diarrhoeal diseases or acute respiratory infection with the aim of exploring and ranking the risk factors.”

Table 1 title is also unclear. Title is again circular. Should read something mean predicted HAZ according to socio-demographic factors. The column heading also needs to be more specific to describe HAZ. IS-MCA is the method, but it needs to describe the variable that is measured.

Responses: To make the Table 1 standalone, the title of the Table has been changed as below:
“Estimated regression coefficients of linear regression model (LM) for Height-for-Age Z-score (HAZ), and mean predicted HAZ (both unadjusted, UPM and adjusted, APM) calculated from interval scale multiple classification analysis (IS-MCA) model by different socio-demographic factors, BDHS 2011”.

Also, the footnote “R indicates Reference Category” has been removed by providing the reference category for each variable with the variable name as: Child’s age in months (Reference Category: <12), Child’s birth weight (in kg) (Reference Category: Large), and so on.

The last two column headings are also modified to “Estimated regression coefficient of LM” and “Mean predicted HAZ from IS-MCA”.

Table 2 title is circular. The authors describe logistic regression of nutritional status by nutritional status- what is the predictor? Sub-heads are missing in the table. For example, it appears that age categories are listed first, but there is no title to indicate these. Instead, this table appears to be demonstrating logistic regression of nutritional status by socio-demographic factors.

Responses: To make the Table 2 standalone, the title has been changed as below:

“Estimated regression coefficients of binary logistic regression model (BlogM) for child malnutrition status defined as height-for-age Z-score less than -2.00, the corresponding odds ratios (ORs), and predicted proportion of malnourished children (both unadjusted, UPP and adjusted, APP) from nominal scale multiple classification analysis (NS-MCA) model, BDHS 2011”.

The last two column headings are also modified to “Estimated regression coefficient of BlogM and OR” and “Predicted proportion from NS-MCA”.

Table 3 - The definitions of "nourish" and malnourish need to be defined. These are not conventional. Does this refer to height-for-age Z scores < -2? It is unclear. Also, the percentage correct is also not clear.

Responses: To make the Table 3 standalone, the title of the Table has been changed as below:

“Correct classification rate of children nutrition status based on height-for-age Z-score (HAZ) as either malnourish (HAZ < -2.0) or nourish (HAZ ≥ -2.0) from linear regression (LM) and logistic regression (BLogM) models, and the overall correct classification rate of children nutrition status by LM and BLogM, BDHS 2011”. Also, a note has given as footnote - “Note: Correct classification rates are row wise percentages”.

Note that we switch Table 3 by the old Table 5 (now it is Table 4)

Table 4 - Education status should indicate if this is mother's highest level of education received. It is not clear why Table 4 is needed. What new information does this contribute?
Responses: To make the Table 4 (which is one Appendix Table 1) standalone, the title has been changed as below:

“Distribution of mothers by household (HH) wealth status, mother’s highest level education status and residential place, BDHS 2011”. Also, a note has given as footnote - “Note: Percentages are row wise”.

This Table is now considered as “Appendix Table 1”, since this table is not directly related to the research objectives. This table is used to show how cross-classification of mother education and household wealth status vary by residential place. The table shows that in rural area the richer households have less illiterate mothers (as well as higher number of educated mothers) compared to urban areas. In this article, this circumstance is indicated as one of the reasons why the urban children are more likely to be malnourished than the rural children when other predictors are considered in the model.

The abbreviations for linear and logistic regression are not necessary and do not all value to the table, instead they add confusion. The tables need footnotes in order to be standalone pieces of evidence. As they read now, the results cannot be interpreted with the text, and even then they are difficult to interpret. For example, percent correct is not a convention.

Responses: We have tried to modify the Tables as well as in the texts to make them understandable. The “percent correct” has been changed by “Correct Classification by LM (%)” and “Correct Classification by BLogM (%)” accordingly, and “Overall Percentage” by “Overall correct classification rate (%)” in the new Table 4.

Dianjianyi Sun (Reviewer 3):

This paper provided essential evidence for child malnutrition in Bangladesh, and compared results from GLM and MCA. However, several problems still need to be carefully addressed.

1. From abstract to the conclusion, the paper has reached to 6400 words. It must be shortened into 50~70%, especially in the background and methods.

Responses: We have tried to make it short as much as possible. Now the number of words reduced to about 5300 which is around 80% of 6400 words.

2. I believe it is more of public health important to address the risk factors and their ranking for child malnutrition. However, there is a lack of results or discussion on this (e.g., how the child, household, and mother-level variables influence? Are the results consistent or inconsistent with previous studies? Why?). Either GLM or MCA is just a statistical tool for exploring. If the main purpose of this paper is to compare GLM and MCA, why it has been submitted to BMC nutrition, rather than a statistical journal. Therefore, more results and discussion should be added on the topic of risk factors and their ranking for child malnutrition.
Responses: (1) We have discussed the results more elaborately in Results section, and (2) discussed their consistency with previous studies in the Discussion section. (3) Since we compared the MCA with linear and logistic models empirically not theoretically, and our aim was to highlight the MCA as an alternative to both linear and logistic models with some extra benefits, we feel it’s better to focus to the public health researchers (instead of pure statisticians) who usually use logistic models for determining the risk factors of any public health issue. (4) We have reduced the Methods section and highlight the Results and Discussion section as per both the reviewers.

3. It is necessary to separate results and discussion.

Responses: We have prepared two sections Results and Discussion, and reduced the Conclusion section into one paragraph.

4. In the abstract, too wordy in background, whereas insufficient contents in results.

Responses: Since the article is mainly based on a comparison of MCA with linear and logistic models, the background takes the maximum place to show the main aim of the study. I hope the abstract (reduced to 200 words from about more than 250 words) now is smooth to read. Though the result section looks insufficient, it mainly provides the actual findings of the study in two sentences.

5. Please delete all the basic statistical concepts of Linear Regression Analysis and Logistic Regression Analysis in the methods section.

Responses: We have tried to delete the irrelevant statistical concepts of both linear and logistic regression models. Also, the MCA methodology has been reduced.

Overall, we have also tried to improve the writing of the article. Please take necessary steps for publication of the article.

Regards,

Kakoli Rani Bhowmik and Sumonkanti Das