Author’s response to reviews

Title: "Individual and household risk factors of severe acute malnutrition among under-five children in Mao, Chad: a matched case-control study"

Authors:
Jovana Dodos (wash1@actioncontrelafaim.org)
Chiara ALTARE (chiara.altare@gmail.com)
Mahamat BECHIR (mahamatbechir@gmail.com)
Mark MYATT (mark@brixtonhealth.com)
Brigitte PEDRO (bpedro@unicef.org)
Francois BELLET (fbellet@unicef.org)
Jean LAPEGUE (jlapegue@actioncontrelafaim.org)
Joachim PETERS (joachim.peeters@hotmail.fr)
Mathias ALTMANN (altmannmathias@yahoo.fr)

Version: 2 Date: 12 May 2018

Author’s response to reviews:

COVER LETTER – Response to reviewers

Charles Plomos

Journal’s Editorial Office (JEO)

ARCHIVES OF PUBLIC HEALTH

May 11th, 2018
Dear Mr Charles Plomos,

Re: Manuscript reference No. AOPH-D-18-00003

Please find attached a revised version of our manuscript "Individual and household risk factors of severe acute malnutrition in children under five: a matched case-control study in Mao, Chad" which we would like to resubmit for publication as a Research Article in ARCHIVES OF PUBLIC HEALTH.

The reviewers’ comments were insightful and enabled us to improve the quality of our manuscript. In the following pages are our point-by-point responses to each of the comments of the reviewers.

Revisions in the text are shown using the Track Changes feature for any additions and deletions. We hope that the revisions in the manuscript and our accompanying responses will be sufficient to make our manuscript suitable for publication in ARCHIVES OF PUBLIC HEALTH.

We shall look forward to hearing from you at your earliest convenience.

Yours sincerely,

Jovana Dodos, MPH

Address: 14/16 Boulevard Douaumont - CS 80060 - 75854 PARIS CEDEX 17

Ph: +381 (0) 64 266 45 34

E-mail: wash1@actioncontrelafaim.org
Reviewer two

1. Reviewer two writes:

Method: you ended up with two models; how about the confounding effect of individual factor on household factors? You need a third model which includes both factors at a time.

Analysis: how did you control the effect of individual factors on household level factors?

Thank you for raising this question. Logistic regression (and other multivariable methods) is most frequently used to identify independent effects. This means being able to identify real associations and apparent associations that are due to confounding.

What happens when we have all variables (individual + household) in one model is that risk factors at one level can cloak / mask risk factors at another level. We have seen, for example, that recent diarrhoea masks poor household sanitation. This does not mean that the two risk factors are not on the causal pathway. We would not be way off the truth to have:

poor sanitation --> diarrhoea --> SAM

In this case, poor sanitation is on the causal pathway and the association:

poor sanitation --> SAM

is a real rather than an apparent association.
Poor sanitation may not, when we have diarrhoea in the model, be an independent association:

poor sanitation --\(\rightarrow\) SAM

expands to:

poor sanitation --\(\rightarrow\) diarrhoea --\(\rightarrow\) SAM

as we "fill in" the causal pathway and we see only:

diarrhoea --\(\rightarrow\) SAM

We risk confusing "independent" with "causal".

Having models at two levels gives us some idea of the causal pathway (we fill in the links with biological plausibility) and yields much more information. We believe this approach is good enough for what we want to do (i.e. identify intervention targets). For example, diarrhoea and vomiting are significant risk factors at the individual level. These results suggest a general WASH (water, sanitation, hygiene) intervention and indicates an ORS (Oral Rehydration Salts) promotion intervention.

We can also see from the household level analysis that lack of a latrine and failure of the carer to wash hands after defecation are significant household risk factors. A WASH intervention is now
indicated (alongside ORS promotion) as we have more evidence. We might also be able to concentrate the WASH intervention on hygiene promotion (as part of an Infant and Young Child Feeding intervention) and latrine building. The third - all variables at both levels model tells us less about intervention targets.

Also, putting all variables in one model raises the issue of sparse data bias, due to the fact that most cases had comorbidities at admission (diarrhoea, fever, vomiting, cough), excluding the possibility to introduce other risk factors in the model.

2. Reviewer two writes:

Conclusion: "Our study identified integrated management of childhood illnesses as a main priority." This is overrated.

We removed this from the revised version of the manuscript.

3. Reviewer two writes:

Line 26; "Despite all the efforts made by the Government....." What nutritional interventions/strategies are in place for under children in Chad?

Thank you for this suggestion. We added existing in-country efforts to the background section.

4. Reviewer two writes:
Line 32: GAM or SAM; it may be editorial problem.

This is not an editorial problem. In our manuscript we use both GAM, referring to Global Acute Malnutrition and SAM, referring to Severe Acute Malnutrition.

5. Reviewer two writes:

Line 28: "Prevalence of acute malnutrition at the national level increased compared to 2016; …" from which magnitude to which magnitude. It is good if you mention the burden.

Good suggestion. We added additional information on trends and changes in SAM prevalence at the national level.

6. Reviewer two writes:

Line 40: "…none of these studies assessed individual and household risk factors…” This is not true; what about the factors you mentioned in background: line 41-44.

You are right. We removed this from the revised version of the manuscript.

7. Reviewer two writes:
Provide some justifications why the study was done in the background.

Thank you for this suggestion. We elaborated more on the added value of conducting this study in the revised version of the manuscript.

7. Reviewer two writes:

Selection of controls: were they selected from the community or health facility? At the community, who made the screening?

We specified in the text, data collection and instruments section that we recruited cases from one health centre and matched them with two controls, recruited from cases’ neighbourhood. “Cases were recruited during the admission days to the OTP in Mao health centre… Controls were recruited from the same neighbourhood where cases reside. After completing the interview with the case caretaker, data collectors visited the closest neighbouring households to identify eligible children for the controls. Anthropometric measurements of the controls were conducted at their household by the two team members in charge of this task...”

8. Reviewer two writes:

Selection: what was your sampling technique for both cases and controls?

Who were the data collectors; their profession.

We elaborated on the selection of cases and controls and sampling in data collection and instruments section, pages 6 and 7. As proposed by the reviewer, we added additional information on data collectors’ profession, we quote: “Data collectors were recruited locally from a pool of enumerators who had previously worked with ACF and ASRADD”.

9. Reviewer two writes:

While you describe your population, it is good if you show the matching variables in the result. Table 2: if age was one of the matching variables, it should not be included in the analysis. Matching is one of controlling confounding effect.

This is true. In our study, we did a matched analysis and age was a matching criteria and we wanted to show that we have accounted for age in the model. You are right, it makes no sense to have age as a separate term in the model. We did not include age in multivariable model, thus it was not included in the stepwise multivariable analysis on individual risk factors.

10. Reviewer two writes:

Recent morbidity in the household what type of morbidity? What questions were participants asked?

We explained what “morbidity in the household” refers to in the Household-level risk factors section. We quote: “Recent morbidity in the household was defined as presence of symptoms (fever, diarrhoea, vomiting, cough) during the past 15 days in other family members other than a child.”, at the bottom of a page 8.

11. Reviewer two writes:

Hand washing after defecation/ using toilet, absence of toilet in the household", versus "hand washing behaviour score" are not these variables collinear? What the VIF (variance inflation factor) for these data?
Thank you for pointing this out. We calculated how much the variance of an estimated regression coefficient is increased because of collinearity. In our study dataset, none of the variables had VIF > 4, so we do not have to worry about multicollinearity in our analysis.

We added this information in the method section of the manuscript, we quote: “Multicollinearity was assessed using variance inflation factor (VIF). None of the variables had VIF > 4.”

12. Reviewer two writes:

Paragraph 4, Line13: "our study, we found a significant association between diarrhea and family dish" such finding should be first included in the result section.

This has been removed from the revised version of the manuscript.

13. Reviewer two writes:

"Overmatching might have occurred and hindered the identification of risk factors associated with the matching criteria (proximity)." Make sure matching variables are not included in the analysis.

We matched cases and controls on their location (neighbourhood) because our objective was to identify risk factors at household and individual levels, not at the community level. While this strategy was appropriate to our objective, we acknowledge that this might have introduce overmatching. Neighbourhood/ location variables were not included in the analysis.
Matching on neighbourhood matches on a large number of “implicit” factors (e.g. altitude, climate, food-economy, ethnicity, socio-economic status, extended family, religion, etc.). This means that the association between SAM and these factors could not be, and was not, explored. It may also mean that risk factors that may also associated with location (e.g. water source) may not be identified. Matching on age and sex, however, expands the “search radius” required to find controls. Having more than one matched control per case has the same effect. These procedures help to reduce problems of overmatching when neighbourhood controls are used.

14. Reviewer two writes:

"Programs focusing on women's empowerment….." Is this supported by your data?

This has been removed from the revised version of the manuscript.

Reviewer Four

1. Reviewer four writes:

Title: It is better to correct as follow "Individual and household risk factors of severe acute malnutrition among under-five children in Mao, Chad: a matched case-control study"

Thank you for the suggestion; we revised the title accordingly.

2. Reviewer four writes:
Abstract: The abstract is well done, but it does not include all-important components of an abstract. For example, the method section of this abstract does not show the method of data entry and analysis. In addition, in line 25 it says multivariate analysis please write as multivariable analysis. Moreover, in line 19 please add a comma between year and month.

We added data entry method to the abstract. Using multivariable instead of multivariate has been synchronized throughout the manuscript.

3. Reviewer four writes:

Please use Oxford English throughout the document. E.g line 35 insert comma before and

Be consistent when you write terms, therefore please write age in months or in years throughout the document. Line 46 women not women's

Lines 35 and 46 corrected as per reviewer’s comment. Document is formatted English (United Kingdom), writing age in months synchronized throughout the document.

4. Reviewer four writes:

Keywords: The number of key words should be between 3-5, but you put more than six key word words.

We reduced the number of key words, thanks for pointing it out.
5. Reviewer four writes:

BACKGROUND: In this section, the authors should describe about the magnitude of SAM in worldwide, and the magnitude of SAM in developed countries. The authors should also explain about the impact of SAM on economical and psychological without intervention. In addition, what measurements the Chad government has done to alleviate the existing problem. Moreover, what were the effects of the interventions? Finally, what is the gap that you intended to fill? Generally, the paper needs extensive grammatical edition thought-out the document. Moreover, at the beginning please define what is SAM???

We added information on SAM trends, consequences and existing in-country efforts to the background section, as well as we elaborated more on the added value of conducting this study.

6. Reviewer four writes:

Methods: Why you use 15% none response rate??? Because your study is case control.

In the data, collection toll you said a structured questioner was used. Please specify whether interviewer administered or self-administered. In line 29 data collectors were trained. For how many days???

We clarified in the text the fact that the increase in sample size was due to availability of resources. It was not due to foreseen non-response. We added missing information on the tool as per reviewer’s suggestion.

7. Reviewer four writes:
The major limitation of this paper is on the statistical analysis. Under this section, the authors explained that data were entered and analysed using EPI-info???. In addition, the authors used conditional logistic regression, which is quite appropriate for matched-case control, but the authors did not consider cluster effect. The authors should do reanalysis by considering a two stage hierarchical model (a two stage multilevel analysis) to see the household level effect and the individual level effect of SAM.

We thank the reviewer for this suggestion. However, we think the multilevel analysis is not appropriate to our study for the following reasons…

In our study, we did not use a cluster sampling method. We recruited cases from one health centre and matched them with two controls, recruited from cases' neighbourhood.

We do have factors at different levels (i.e. individual and household). We approached this by, in effect, treating the data as two separate matched case-control studies:

(1) individual risk factors,

(2) household risk factors.

The first study has individual cases as cases of SAM. The second study has cases as households containing a case of SAM. For both studies the matching is at the individual case level. We do not think this is wrong as we look at the effect of household level exposures on children of similar ages while controlling for environmental and other factors subsumed under "place of residence".
We think this is a useful approach as we can see (e.g.) that diarrhoea and vomiting are significant individual risk factors at the individual level. These results suggest a general WASH (water, sanitation, hygiene) intervention and indicates an ORS (Oral Rehydration Salts) promotion intervention. We can also see from the household level study that lack of a latrine and failure of the carer to wash hands after defecation are significant household risk factors. A WASH intervention is now indicated (alongside ORS promotion) as we have more evidence. We might also be able to concentrate the WASH intervention on hygiene promotion (as part of an Infant and Young Child Feeding intervention) and latrine building.

We do not have groups / levels as is usually understood in multilevel modelling (MLM). With MLM we are interested in careful assessment of the effects of group / level membership. We are completely uninterested in looking at the effect of set ID number since the matching criteria are treated, by design, as confounders that we want to ignore. When we match like this we exclude any consideration of the effects of matching criteria.

8. Reviewer four writes:

Moreover, from line 33-35 it says, "Variables that were significant in bivariate analysis were considered". Therefore, what does mean significant please specify the p-value that you used in the variable selection process?

In case and control proportion, how much proportion you used 1:1 or 1:2 or1:3

We did mention control to case ratio, which is 2:1, in Sample size calculation section, line 60 (bottom of a page 6 in the improved version of the manuscript). We also specified significance level for the variable selection process, lines 34-36: “Variables that were significant in bivariate analysis were considered for inclusion in the multivariable conditional logistic regression models. A p value = 0.05 was considered to be significant” and lines 46-47: “A stepwise
backwards elimination approach with a significance level of $p = 0.05$ was used to build the final models.”

9. Reviewer four writes:

Discussion: This section of the research did not give any meaning to the finding. The authors simply put the result, but they should interpret the result and compare this finding with other previous research output, studied either in the country or other countries, WHO or country guideline. Moreover, they should come up with possible justifications. Their finding has no possible justification for the discrepancy. The discussion is weak and does not explain the importance of the study as well.

We have revised the discussion section to include further references to existing studies, possible explanation and discussion points.

10. Reviewer four writes:

In the discussion part, the researchers explained the whole result part. In addition, the authors discussed non-significant findings, for example to our surprise; we did not observe a significant association between point-of-use water quality and SAM. Please avoid such and the like issues. The discussion should be presented in related to the objective. Therefore, I recommend to the authors to modify the discussion based on our suggestions.

We do discuss one result that is not significant as this factor was expected to play a major role. We reduced the section on this.

11. Reviewer four writes:
The limitation section was omitted, which is a very important issue in a case control study (recall bias is inevitable).

We discussed potential limitations of this study in the last paragraph of the discussion section, lines 4-15, page 13. We quote: “While this is the first study on the household-level risk factors for SAM in the Kanem region, it has some limitations. The study relied on participants’ self-reported data, which is prone to recall bias and tendency of respondents to report socially desirable behaviours. Retrospective information tracking is one of weaknesses of case control studies. However, due attention was given to respecting study procedures, including training of data collectors, taking anthropometric measurements and supervision throughout data collection period to minimize expected bias. Overmatching might have occurred and hindered the identification of risk factors associated with the matching criteria (proximity).”

In addition to what was already written, we added some other limitations we could think of to the revised version of the manuscript, page 17, we quote:

“Matching on neighbourhood matches on a large number of “implicit” factors (e.g. altitude, climate, food-economy, ethnicity, socio-economic status, extended family, religion, etc.). This means that the association between SAM and these factors could not be, and was not, explored. It may also mean that risk factors that may also associated with location (e.g. water source) may not be identified. Matching on age and sex, however, expands the “search radius” required to find controls. Having more than one matched control per case has the same effect. These procedures help to reduce problems of overmatching when neighbourhood controls are used.

In this study design, the ascertainment of exposure is done after the outcome. This means results lack the element of temporality and therefore causality cannot be implied. Recruiting cases at the health centre may have introduced a self-selection bias, as cases who sought care at the health
centre may be different than cases who did not. SAM symptoms may have alerted caretaker to go to the health centre, overestimating the odds ratio between recent morbidity and SAM. Nevertheless, this was considered a suitable approach to investigate a rare condition such as SAM. In fact, recruiting cases in the community would have been a long and costly process, and recruiting controls in the health centre was not appropriate to our objective of matching on the place of residence (village/neighbourhood).

Finally, the data is context specific and special consideration must be taken on extrapolating the findings to other contexts.”

12. Reviewer four writes:

Conclusion: This section is relatively well written