Reviewer’s report

Title: Severely malnourished children with a low weight-for-height have a higher mortality than those with a low mid-upper-arm-circumference: I. Empirical data demonstrates Simpson's paradox

Version: 1 Date: 26 Nov 2017

Reviewer: Dominique Roberfroid

Reviewer's report:

Paper I:

This is an important paper as it confirms that children with a low WHZ are at increased risk of death, whereas many nutrition programs around the globe screen malnutrition with MUAC only.

MAJOR COMMENTS:

- The authors mention "The data from all the IPFs, OTPs and SFCs were combined to give three separate datasets as individual facilities did not contain sufficient deaths to allow for meaningful statistical analysis". Simply pooling results from various studies/settings is inadequate, because such approach does not account for the clustering of data (and one can expect high intra-correlation within clusters given the differences in countries, years, service organisation, refeeding strategies, etc…).

There are two options here. If the authors have data aggregated per country at their disposal, they should present a meta-analysis (this can be done for both case fatality rates and relative risks) where results of individual studies are weighted according to sample size and number of events. A forest plot would be a very visual way of showing the variations in CFR/RR across country. Heterogeneity of the case fatality rate across studies (even within strata of mode of treatment) should be investigated (possibly by meta-regression). If the authors have individual participant data at their disposal, which seems to be the case, then multilevel statistical models (random effect:
country) with adjustment for covariates (age, sex, …) should be used. Such statistical models would also allow formally testing interactions (e.g. interaction between mode of treatment and the relative risk of death).

- Not enough information is provided on the characteristics of the refeeding programs per country: year, admission criteria, refeeding protocol, co-morbidities, etc…Such information would be useful to interpret the results, e.g. the variations in fatality rates across settings. If the authors have such information at their disposal, they should insert a table with study characteristics.

- The proportion of defaulters is substantial in some settings. If they have that information at their disposal, the authors should provide the characteristics of defaulters (how much did they differ from remainders?). If not, this should be discussed as a study limitation.

- Why are both relative risks and odds ratios reported? The authors should opt for only one measure as they convey the same information.

- It is a merit of the study to present results stratified by the presence of oedema. However, as nutritional oedema is considered a stand-alone criterion of SAM, the relative risk of K-muac, K-whz and K-both should be computed in comparison with Kwash, not with M-muac. One could also wonder if the stratification by presence of oedema should be reported at all in this paper, as 1. the objective of the paper is about comparing mortality risk associated with a low MUAC vs. a low WHZ; 2. WHZ is not an accurate indicator of malnutrition in the presence of oedema; 3. nutritional oedema is considered a stand-alone criterion of SAM. The authors could consider concentrating on oedema-free children in this paper, which would help focusing the message and the discussion.

- The discussion section contains very interesting points. However, this section needs to be better structured (for example the Simpson paradox is explained at length in one section, and then referred to again in the section on ROC curves). On line 427, the authors state "Apart from confounding, co-morbidity, bias and the stochastic nature of prognostic models there is another
major problem with the analysis of the data for WHZ and MUAC: that of "mathematical coupling". The discussion could be structured along those elements.

-On line 439, the authors state "but also confounding due to the presence of oedema, HIV, convulsions, measles and other biases that affect children with a low MUAC and WHZ differently." The authors should discuss how these confounding factors may have affected their own results.

MINOR COMMENTS:

- In the methods section, the authors report on data management (age, oedema, sex not recorded). It could be useful, as an indicator of data quality, to report if lacking information affected some particular datasets.

- Table 3a-3b (significance levels) should be integrated with table 2

- Some typos remaining, e.g. "complimentary" (l415), "is" (instead of "in"; l483)

Paper II:

On the basis of below elements, I consider that the paper is not (yet) in a sufficient stage of preparation. The authors might consider turning the document into a concept paper or a position paper given their stated objective, or to transfer the main elements in the discussion of paper I

MAJOR COMMENTS:

- The paper does not apply the PRISMA statement recommendations for reporting systematic review (http://www.equator-network.org/reporting-guidelines/prisma/), which makes a correct review of it quite impossible. The authors are invited to fulfill the criteria of the PRISMA statement.

- The objective of the review is unclear "to examine the evidence in the published literature that is relied upon to assert that a MUAC-only program is ethical". The authors should redefine their
objective. Is it to pinpoint the reasons why MUAC is presented as a better predictor of death in community programs, in contrast to what they observed in therapeutic program in their paper I?

- The rationale to include altogether very different types of study is not straightforward. As the authors themselves state "the data from individual datasets were combined, despite the fact that different diagnostic criteria were used in many of the studies, and similarly analysed. [...]The heterogeneity of the standards used, the admixture of oedematous cases and the failure to account for confounding, such as TB, HIV, and non-nutritional conditions, makes amalgamation of data from the different studies problematic and the absolute numbers computed and compared should be considered to be approximate at best". The authors could have focused their review on studies with current standards and allowing disintegrating MUAC-only, WHZ-only and MUAC-WHZ, in children without oedema, and exclude other studies as inappropriate for this study.

- Although referring to published literature, the data presented in the companion paper I of this series and from an unpublished presentation were also included in the analysis. The former accounts for a substantial proportion of the information included. It is therefore not surprising that the authors state that "conclusions drawn from our empirical data (paper I) are supported by the published reports…".

- Results and discussion are mixed together. There is no presentation of overall results. The discussion is organized in themas and only the results of some studies are picked up and discussed. The statement: "In general the mortality rate was higher in those children fulfilling the WHO2006 WHZ criterion than the MUAC criterion alone" (l 18-19) is not substantiated by the results presented in table 3. Moreover, as in paper I, to reach the total the authors simply added the numbers from individual studies instead of making a meta-analysis (which accounts for the respective weights of the individual studies).

MINOR COMMENTS

-Referring the reader to studies highlighted with different colors in the tables makes the reading extremely complex
Paper III:

MAJOR COMMENTS:

-The paper convey an important message for policy-making: both the mortality risk associated with the each nutritional indicators and the case-load associated with each indicator must be considered. This said, this is a general rule (the authors could refer to the concept of Population Attributable Risk Fraction), although it might be useful to emphasize it given the terms of the current debate around MUAC vs. WHZ.

-There are some apparent contradictions in the methods section. First, the authors state "The ratio of S-muac: S-whz: S-Both reported in the patient data differed significantly from that found in representative community surveys of malnourished children. For that reason the ratios of the number of children in the S-muac, S-whz and S-Both categories reported in papers I and II were not used in any calculation." (l 142-145). They also state in paper I that "The ascertainment bias indicates that the data do not reflect the children with SAM in the community and disproportionally describes the experience of more severely affected children than are generally found during a community survey" (1489-491). Notwithstanding, the authors used the mortality rate of those children in their computation. Second, they present datasets from the literature as "flawed in various ways", but still used the figures from these datasets.

-The authors state that "The mortality risk associated with a low MUAC is additive to the risk associated with a low WHZ when children have both deficits" (l 299). In view of results in paper I, the effect seems rather multiplicative.

-The discussion section is rather a plea to use both WHZ and MUAC for diagnosing SAM (which is a fully correct conclusion based on results of paper I), and could be easily integrated in the discussion section of paper I.

MINOR COMMENTS:

-What means "inadequate # deaths" in table 4?
Level of interest

Please indicate how interesting you found the manuscript:

An article of importance in its field

Quality of written English

Please indicate the quality of language in the manuscript:

Needs some language corrections before being published

Declaration of competing interests

Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organisation that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below.
If your reply is yes to any, please give details below.

'I declare that I have no competing interests'

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report
including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal