Author’s response to reviews

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

Authors:
Emmanuel Grellety (Emmanuel.Grellety.Bosviel@ulb.ac.be)
Michael Golden (Mike@pollgorm.net)

Version: 3 Date: 04 Jun 2018

Author’s response to reviews:

Dear editor,

We have been asked to split our submission into 3 different submissions with different numbers. The utility of submitting them together previously was so that the reviewers could have the papers together and get a general overview of the three papers and how then complement each other; this has been achieved.

Consequently we now copy this response to your comments to each of the 3 papers in the cluster.
Numbers: NUTJ-D-17-00157R2, NUTJ-D-18-00169 and NUTJ-D-18-00168

We would ask that they all be accepted (or not) and published together at the same time. If any of the papers requires revision we would ask that the others be held until they can all be published together.
Below we include a point-by-point response to the comments and questions raised by reviewer 4 and yourself to each of the 3 papers in this cluster.

In terms of your understandable comment “consider the volume of tables and figures, both main and supplemental”. We would point out that the data that we have added since the previous review has all been inserted because of the comments and requests of the reviewers. That being the case it does not seem that we should now remove these additional analyses which your reviewers considered to be important in order to shorten the papers.

We do realise that this is a contentious topic and many of the arguments rest upon statistical principles that may not be familiar to some of your readers. We presume that this is why it has taken about one whole year to review the papers. We do note that your reviewers are in agreement and have not raised any of the points that are now raised by yourself. We presume that this is related to the reviewers being knowledgeable about this subject whilst you are looking at the manuscripts from the point of view of the general reader. It is important that general readers will be able to clearly follow the data, the resulting arguments and are made aware of the sort and quality of the assertions that have been advanced by others that have resulted in major changes in policy by many Governments and NGOs. It is important that general readers, particularly those responsible for policy, are able to appreciate the manuscripts without getting bogged down by trivial or tangential considerations. Therefore we really do thank you for making the effort to make the papers more easily understood by a non-expert reader.

We have addressed each of your points with either a change or a justification. It will be good if the more contentious statements raise the interest of the general reader to explore some of the references and arguments further and allow them to reach their own opinion about the interpretation of the data, the potential pit-falls in these types of analyses, the policy implications and the ethics of abandoning the use of WHZ entirely. Where we decline to make a suggested edit, we hope you will appreciate the reasons we give. We trust that the papers with these
clarifications and minor corrections are now acceptable for publication. After a year in gestation we would be grateful for a speedy decision.

Yours faithfully,

Emmanuel Grellety & Michael H. Golden

Author’s response to reviews:

Reviewer #4: The reviewer thanks the authors for addressing the comments. All three manuscripts have been much improved. The reviewer has the following minor comments:

Manuscript I:

Table S1: The reviewer would like to thank the authors for correcting the errors made in the initial submission. The following comment is just a detail, but the reviewer would like to know why there were slight differences in the totals number of children in Table 2 and Table S1. For example, there were 16531 children in M-whz in Table 2, whereas there were 16530 children in M-whz in Table 1S. Shouldn't the total numbers be exactly the same in both tables?

R4 A1: Thank you very much: The table S1 has been corrected (table 2 was correct).

Manuscript II:

Table 2: Please, could you add a sentence in the manuscript explain why the Fisher's exact tests were one-tailed as opposed to two-tailed (as done for Chi-squared tests)?
R4 A2: Thank you for this comment; as you know Fisher’s exact test is not normally used with very large numbers of subjects because of the nature of the computation. We found that with some of the analyses we repeatedly got “stack overflow” errors when we tried to compute with a two-sided test. As we were asked to include Fisher’s exact test during a previous review of manuscript we complied by including this with a single sided test.

Editor comments:

PAPER I

• p2 – please add comment on the nature of the data/population studied (line 10)

E A1: This has been done.

• p3, line 27 – there is uncertainty in this statement that needs to be stated e.g. previous analyses may be affected by Simpson’s paradox?

E A2: Undoubtedly this has happened with some of the previous analyses (as we show in manuscript II). Nevertheless, we have changed “makes” to “may make”.

• p4, line 43 – please clarify to explain why prevalence overestimated – and inclusion of incidence would increase the burden

E A3: Prevalence is always overestimated because no data set is without random errors. Random errors always increase the SD. If the random errors are trivial the overestimation of prevalence will be marginal. This is fully explored in the reference given where the magnitude is demonstrated with a Monte-Carlo simulation and also by an analysis of surveys from West Africa. However, as this introductory statement in assessment of the magnitude of the potential problem is tangential to the main thrust of this paper in our opinion the reference is sufficient.
The length of time a child is malnourished has a mean of somewhere around 3 months, so the incidence (per year) is around 4 times the prevalence at any point in time – this also varies quite widely by country and seasonally, but annual incidence is always greater than prevalence for any illness with a mean duration of less than 12 months. All we are trying to indicate in this section of the introduction is that SAM is a major and underestimated problem.

• p5, line 65 – obsolete seems very strong without supporting evidence – have they been superseded?

E A4: Yes, they have been superseded. Everyone concerned with assessment of child growth is aware of this definitive change. This has been clearly stated by WHO in the reference given [14]. It is fully accepted by the international community. The WHO contacted national health authorities from 219 countries. By April 2011, 125 countries had fully adopted the new WHO standards, another 25 were considering their adoption, and 30 mainly industrialised countries had not yet adopted them (most have since then adopted these standards). The countries with patients included in this paper have officially followed WHO’s advice and declared the old reference obsolete. Notable exceptions were the USA and the UK who are using their own standards although there is pressure for them to change. All national protocols for treating SAM and all International NGOs and UN agencies use WHO standards exclusively. Among the industrialized countries, the standards have since been endorsed in 15 European countries. In the United States, the Centers for Disease Control and Prevention, the National Institutes of Health, and the American Academy of Paediatrics have endorsed the use of the WHO growth standards. Changing growth references to the international standards presents major training and logistic problems for governments which accounts for the delays in full implementation.

• p15, line 320 very high CFR?

E A5: Thank you, “high” was missing. This has been rectified.
• Figure 4 needs explanation (also applies to Paper II): the relative risk is potentially hard to interpret as both criteria are markers for increased mortality and the lower RR to signify higher mortality may be challenging to understand without explanation. It would be helpful to the reader to set out an example, and also to add comment to the legend or footnote – term ‘favours’ could be amended

E A6: These are the natural logarithms of the risks of death comparing the mortality with MUAC and with WHZ (not with the general population). The solid line represents equality. The dotted line represents the mean of the comparison. Where the line becomes positive it means that WHZ has a higher mortality rate than the MUAC mortality rate – “favours WHZ” means that the risk of death for a child with a low WHZ is greater than for a child with a low MUAC and vice versa.

This is an absolutely standard method of presenting a forest plot and the term “favours xxx” has been used by many other authors. Nevertheless, we have added to the legend:

“In each of the forest plots “favours WHZ” indicates that the Odds ratio for death is higher in children with WHZ <-3Z than with a MUAC of <115mm; “favours MUAC” indicates that the Odds ratio for death of children with a MUAC <115mm is higher than those with WHZ <-3Z.”

We hope that this will remove the uncertainty when readers not familiar with meta-analyses read our paper.

• p18, line 394 ‘beauty’?

E A7: Yes – “Beauty” is correct. It is even given in the title of the paper that we quote and has been widely misquoted to justify the contention that children with a low WHZ are healthy (there is no doubt that this is a misquotation given a careful reading of the paper as there is no mention of wasting anywhere in the paper). - Bogin et al. “Leg Length, Body Proportion, and Health: A Review with a Note on Beauty” Int J Environ Res Pub Health 2010, 7: 1047-1075. http://www.mdpi.com/1660-4601/7/3/1047/htm

• p18, line 395 – please justify ‘this contention is without evidence’
E A8: There are no data to support the contention that children with a WHZ< -3Z are healthy anywhere in the literature, perhaps this is best illustrated by those asserting that such children are healthy having to misquote a review which does not address this issue at all for they themselves cannot find any evidence to support this contention. On the other hand there are many papers which show the opposite – i.e. that these children are at considerable risk of death and we have referenced 12 of these publications. Therefore this statement in our and your reviewers’ opinions does not need additional justification.

• p18, line 403 – comment on Briend’s authorship of a supporting paper seems a personal comment – please amend

E A9: This is an important point – it is not a personal attack. It is inserted to show that the person who drafted the paper claiming that children with a low WHZ are healthy has elsewhere published data that completely contradicts this contention. As such it is fully justified to bring this contradiction to the notice of your readers. We also note that it has not been commented upon by your reviewers who are aware of the papers involved.

• p19, line 407 – ‘misquote’ is a judgement and could be expressed differently – as explained in the text that follows

E A10: But he does quote these papers to support his contention that “long legs” is the main reason why children have a low WHZ, whereas the papers he quotes show that this is not the case as the authors of the original papers conclude. That makes it a misquotation by definition.

Nevertheless, we have changed this to read:

“whereas the only papers that have addressed this issue show this to be, at best, a minor contributor [32,42,79] and the original authors are clear that long legs do not explain the discrepancy between WHZ and MUAC”.

• p19, line 417 – ‘also’ is not needed in this statement
E A11: This is a further contention by Briend et al so that “also” is appropriate here. Perhaps this is simply a difference in style – if you want we could start the sentence with “Furthermore”, but “also” is succinct.

• p19, line 428 – Simpson’s paradox will not be familiar to readers – could a clearer example be included here to help explain? – the classical examples [82] given on p20 are challenging for the reader to understand

E A12: We agree that the readers may not be familiar with this paradox, although it is well known to statisticians. That is why we have explained it fully in the results section and have even inserted a simple explanatory (table 5) table that can be copied into a spreadsheet for the reader to play with if s/he wants to understand the phenomenon. Because it may be challenging we have gone to these lengths to try to make it comprehensible. We do not know how we can be clearer. To simply repeat the explanation we have given in the results in the discussion would be a tautology that would simply lengthen our paper even more than it is at the moment.

• p22, line 496 – should this be ‘not given’?

E A13: Thank you, you are correct and this has been changed.

• p22, line 513 – please delete ‘and praise’

E A14: This has been deleted. As a scientist you must agree that if a paper is quoted we are entitled to assume that the person quoting the paper has read it, understood it and agrees with its conclusions (or not if this is clear from the text). The author quoted here indeed praised this paper. We can only conclude that the author quoting the paper did not read or understand it, or just read the abstract or simply ignored the flaws. It seems that he praised it because the conclusions agreed with his preconceived notions. Nevertheless, we have removed the “and praise” as it is not really necessary.
• p27, line 610 – the data presented in the paper address and inform the controversy re diagnosis of SAM – comments such as ‘some within the nutritional community has been deceived’ are unhelpful and the authors’ judgement – please amend

E A15: But this is the truth. It is precisely what has happened. The nutrition community are indeed unfamiliar with statistics and the flawed analysis has deceived them into believing the rhetoric of the advocates of MUAC-only programs. We have not apportioned any blame. But it is really excessive to claim that severely malnourished children with a WHZ<-3 are healthy. Your reviewers understand that there has been exceptionally strong advocacy, and rude comments made on ENN-Net against anyone who questions them.

PAPER II

• line 69 – please provide supporting evidence for ‘morphed into the only tool...’ or revise

E A16: We have added references and changed the sentence to read:

“What started as a simple screening tool for the community identification and as an alternative to WHZ has changed into the primary tool to be used to diagnose SAM with advocacy to extend the suppression of WHZ assessment universally [22,29,30,32].”


• figure 1 – was the number of additional records identified n=6?

E A17: This arises because one of the papers (Garenne 2009) contained an analysis of two datasets. That means that there were 5 additional papers/reports obtained – but they contributed 6 datasets. This is explained clearly in the methods.

• line 142 – please include brief statement to explain mathematical coupling so that reader does not need to refer to companion paper

E A18: We included the sentence below to explain briefly mathematical coupling. This is a direct quote from Archie’s paper – we cannot better his definition which we consider to be clear, unambiguous and concise.

“Mathematic coupling occurs where “one variable directly or indirectly contains the whole or part of another, and the two variables are analysed using standard statistical techniques” [37,38].”


• line 151 – please explain ‘redundancies’

E A19: This was a shorthand way of indicating that there was duplicate use of the same raw data which had been analysed in a different way. We have changed the sentence to read:

“Each of the duplicate datasets were included in an initial analysis by subgroup comparing the effect of using different diagnostic cut-off points to analyse the data and demonstrate the effect of using different references on the derived outcomes.”
• line 166 – and define DOI

E A20: This is a term used in meta-analysis. We have changed the sentence to read:

“Tests for heterogeneity were generated for each analysis by two methods (funnel plots and Doi plots). In the funnel plot, the ORs were plotted against the standard error, while in Doi plots, the ORs were plotted against z-score.”

• additional file S2 – please check for balance in commentary on individual studies throughout this file – for example whilst paper 9 (Chiabi et al) has significant imitations and low quality score (Additional table 1); comment on whether it ‘should have been refused by anonymous reviewers’ is unhelpful; paper 13 comment on the observed difference in deaths is speculative and ‘grossly biased’ seems inappropriate; whilst paper 14 does not make the comparison needed, its conclusions may not necessarily be biased?; paper 17 ‘discredited algorithm with gross inaccuracy’ (and ‘discredited’ also in the text) – please amend

E A21: These commentaries have been amended.

• p11 line 207 – should this refer to 1 to 7?

E A22: Thank you, this has been changed.

• figures 3 - 6 – please add more detail to explain the forest plot (as above); the relative risk is potentially hard to interpret as both criteria are markers for increased mortality and the lower RR to signify higher mortality may be challenging to understand without explanation. It would be helpful to the reader to set out an example, and also to add comment to the legend or footnote

E A23: Similar to response E A6 we have added the following to the end of the first forest plot legend.

“In each of the forest plots “favours WHZ” indicates that the Relative Risk for death is higher in children with WHZ <-3Z than with a MUAC of <115mm; “favours MUAC” indicates that the Relative Risk for death of children with a MUAC <115mm is higher than those with WHZ <-3Z.”
• p14, line 290; p16, line 335 & 346 – please justify/explain use of ‘reliable’

E A 24: Here we are talking about the relative reliability of the various studies.

- p14 line 290: study 4 is from a secondary analysis of a prospective RCT (randomised controlled trial). The assessment of a study’s reliability is ranked according to the probability of bias and RCTs have the highest level of reliability because they are designed to be unbiased and have less risk of systematic errors compare to the others studies. This is well understood by all who undertake systematic literature reviews and is outlined in the instructions to Cochran reviewers and all the systematic reviews undertaken by WHO. Further justification is therefore not required when an RCT is compared to other types of study.

- p16, line 335: the use of ‘reliable’ is explained and justified by the following sentence (p16, line 341 to 344) in terms of their reliability to address the question posed in the papers: “These 7 reports are the only reports where children with S-whz and S-muac were separated in the analyses to eliminate the effect of mathematical coupling [64,65], oedematous children were excluded, the age range of the children was from 6 to 60 months and the WHO recommended criteria for diagnosis of SAM were used.”

- p16, line 346: we have justified this on the basis (line 350) that this group should not have an ascertainment bias whereas the reports from patient cohorts in both IPFs and OTPs have been criticised because they “do not represent the community” (line 455). However we did omit to state that this specific group is represented in figure 5. This has been rectified.

• p19/20 – please consider whether text could be revised (diagnostic bias section) – ‘2009 paper must not continue to be used....despite this paper used obsolete criteria’. In particular, explanation of why criteria are obsolete is not given and appears subjective.

E A 25: Please see our response to your query above at E A4. The WHO has officially declared to all governments of the World Health Assembly that the old growth references are redundant and should no longer be used. They have even attempted to change the Global Database of Child Malnutrition to convert the old surveys using NCHS standards to the WHO standards so that all
historical data can be compared with recent data using the new standards. Obsolete is the correct term to use. We have demonstrated 1) in fig 4 the difference in the standards, 2) in the comparison of the same raw data analyses using different standards and 3) in our meta-analysis comparing standards that the use of obsolete standards leads to different results than when the WHO standards are used. The demonstration of the magnitude of the disparity by comparison of Garenne et al’s reports is the point of this section and is the first time such an analysis has been made; the demonstration is important.

There is reason to believe that the use of obsolete standards 3 years after the WHO standards were introduced was deliberate in order to give the impression that the author wished to obtain. We have refrained from making any comment concerning this apparent manipulation in this paper as indeed that would constitute a personal attack which are trying to avoid as far as possible.

• p20 – ‘misleadingly quoted’ – please reword

E A 26: We have shown the limitations of the data. These data have been repeatedly quoted and used to change public health policy which, in our opinion, puts hundreds of thousands of children at risk. We do contend that those setting public health policy have been misled – that is what our data shows and it appears that your referees agree.

• p25, p553 – ‘justify’ appears sufficient rather than ‘justify and praise’ – ‘found wanting’ is unhelpful – please specify the principal limitation

E A 27: ‘and praise’ has been deleted and ‘found wanting’ has been replaced by ‘found sufficiently problematic that they should not be used to guide policy decisions’

PAPER III

• p2, line 10 – please delete ‘in excel’; line 20 – please add ‘estimated’ to the minimum number of deaths and link to the analyses presented; line 21 – should this refer to deaths among children excluded?
E A 28: p2, line 10 - ‘in excel’ has been removed; line 20 & 21 – “ineligible for treatment” and “excluded from treatment” are exactly the same but as we have used excluded elsewhere we agree that this is a useful change. The sentence now reads:

“Globally, the estimated minimum number of deaths that would occur among children excluded from treatment in our analyses is 300,000 annually.”

• p4, lines 51/52 – please rephrase ‘adamant’ and ‘waste of effort’

E A29: This is precisely what the advocates for a MUAC-only program are repeatedly and strongly advocating for – adamant is the correct description of their campaign. They are very strident in their statements and have themselves stated that research to assess WHZ in the community is a “waste of effort” – this is a quote from the references given. We have therefore now put that in as a quote.

• p5, line 81 – please rephrase ‘flawed’

E A 30: We have replaced this by “severely biased”. Nevertheless, the data are statistically flawed and that is the correct term to use when there is such severe mathematical coupling – and it would appear that your statistical editor is in agreement with this (from the data we have presented at least).

• p6, lines 82-86 – the argument regarding the nature of the mathematical models and their interpretation is well-made – but unnecessary then to state that there are ‘flaws’ in previous research or to emphasise ‘appropriate’ analysis in relation to the authors’ own work. It is the authors’ opinion that the criticisms of Briend et al are incorrect – and should be phrased as such – please re-word

E A31: We have demonstrated that there are indeed flaws in the previous research and we stand by that statement. In terms of the criticisms of Briend et al, we have added “in our opinion”.
• p6, line 87 – please add supporting evidence for this statement

E A32: Many agencies and several national governments have ceased attempting to identify and treat any children with SAM diagnosed by WHZ. If you wish, we could put into yet another supporting file the following examples:

1) The Philippine Government, with technical support from UNICEF, officially launched the National Guidelines for the Management of SAM in late 2015 and is now scaling up interventions in 17 provinces. These Guidelines state that MUAC is an independent admission criterion and can be solely used to admit and discharge patients. However, if capacity and resources exist, WHZ can also be used [1].

2) The Nigerian 2011 National CMAM guideline states [2]: ‘WHZ is an independent admission criterion to identify SAM in addition to MUAC and bilateral oedema. Assessment of WHZ is more time consuming. If resources (staffing, training and height board equipment) are adequate, WHZ measures can be taken, but if not, and if numbers of admissions are high, it is recommended, to just use MUAC and oedema for admission and discharge and in addition monitor weight changes at each visit to the health facility’(13, p20). This approach has been promoted by the Government and partners during the scale up of services across the country and in particular, during the ongoing 2016-2017 emergency response in the North East.

3) The Cambodian National Guidelines for the Management of Acute Malnutrition (2011) state that in addition to Community Health Workers using MUAC only anthropometric screening, ‘hospital and health centre staff will use MUAC and check for oedema. Weight-for-height (W/H) will be used only for inpatient care.’ (14, p8) [3].

4) In Pakistan, the National CMAM Guideline, since 2013, has allowed for MUAC and oedema only programming [4]. UNICEF is currently supporting a number of programmes using such criteria. For example, in Khyber Pakhtunkhwa, Balochistan and Punjab.

5) In Sudan, there has been a move to simplify protocols in order to support scale up, after an evaluation by Valid and UNICEF in 2015. The 2015 National Nutrition Program Outpatient Management of Severe Acute Malnutrition at Primary Health Care Level (PHC) Operating Guide for Health Workers endorses MUAC and oedema only programming [5].
6) The Bangladesh National CMAM Guidelines (2011) do not include WHZ as an admission criterion [6]. Instead, SAM with complications, for inpatient care, is defined as MUAC<115mm; symptoms of dehydration and oedema (any grade); SAM without complications, for outpatient admission, is defined as MUAC <115mm with appetite and no medical complications.

The above list is not exclusive but provides some examples of national Government policies that are promoting MUAC and oedema only programming. Moreover some donors as ECHO have regional initiative agenda to promote the MUAC only in West Africa.

Since as early as 2012, various NGOs have been piloting MUAC and oedema only programming. This following list is not exclusive but includes some examples of programmes previously or currently implemented using MUAC & oedema only criteria.

1) Médecins Sans Frontières (MSF) in Chad, Central African Republic, DRC, Haiti, India, Madagascar, Malawi, Mali, Nigeria, Somalia, Sudan, South Sudan, Uganda, etc.

2) Alliance for International Medical Action (ALIMA) in Burkina Faso, Cameroon, Central African Republic, Chad, DRC, DRC, Guinea, Mal, Niger, Nigeria, South Sudan.

3) Action Against Hunger Spain & US in Nigeria, Pakistan, Nepal, etc.

4) UNICEF in Jordan.


6) International Rescue Committee (IRC) in South Sudan.

7) Goal in South Sudan.

The following non-exhaustive list of research projects are all aimed at developing and extending the MUAC-only policy:

1) ALIMA is conducting a non-comparative study in Burkina Faso on improved evidence of the effectiveness of approaches to treat acute malnutrition with a single 'MUAC-only' protocol and a single product, namely, RUTF [7].

2) IRC, MSF and Action Against Hunger are implementing a MUAC-only based programme in South Sudan and Kenya from 2017 as part of a study examining the feasibility of a combined protocol to treat MAM and SAM (ComPas) [8].
3) In South Sudan, IRC is currently conducting a research study of the integration of SAM treatment into iCCM by low-literate CHWs. Low-literacy tools have been developed.

4) MSF, in partnership with a group of nutrition experts, has developed, and is currently field testing, a “Universal MUAC” (uniMUAC) strap which assesses the nutritional status of both children and adults [9].

Moreover, the CORTASAM was established in 2016 by an NGO (and has no official standing) as the “first-ever technical panel of independent experts on acute malnutrition” and their first recommendation is that: "MUAC should be used as the primary tool for the detection, diagnosis, and discharge of acute malnutrition in children 6-59 months of age in the community." (https://www.nowastedlives.org/advisory-group/).

A web site 'No wasted lives' has been built to promote the MUAC-only: https://www.nowastedlives.org/

References:


This is far from an exhaustive list of what is happening and the extent to which MUAC-only programming has already been widely adopted. Hence the urgency in producing data and the counter arguments.

• p11, line 208 – should last % be 83% for Senegal (Fig 2)?
E A33: Yes, you are correct, thank you.

• p12, line 233 – these are modelled figures – please specify
E A34: Indeed we clearly state that they are estimates…

• p14, paragraph starting at line 287 – text needs amendment to align with the data presented; ‘faulty statistics’ ‘completely false assertions’ go beyond the data presented. The findings of this paper make the point very clearly that in an attempt to have a simple screening criterion for SAM there is a risk that many undernourished children will be missed. It is unhelpful (and risks being inflammatory) to suggest that there are factions e.g. ‘those subscribing to a MUAC-only policy’ and that these researchers are content with the significant numbers of deaths in children (e.g.
humanitarian perspective p13) that may result. The data presented in this paper are modelled and the implications of these analyses needs to be discussed in an objective and balanced way.

E A35: We disagree with the fact that we are going beyond the data presented. This is a discussion and not a simple repetition of our data. We discuss the public statements that have been made on the relationship between WHZ and MUAC which is germane to the whole thrust of our papers. We discuss the implications of our 3 manuscripts for public health policy and we put the counter-arguments to the propositions that have been advanced in favour of a MUAC-only policy. It is appropriate to discuss these assertions where we consider them to be incorrect and to give cogent reasons for our considerations. Furthermore, we contend that it is useful to address the question of why this situation has arisen. This is mainly because we do not wish to bring any criticism against those who have accepted and acted upon advice of the leading advocates for the MUAC-only programs. They mostly are not statisticians and have never heard of Simpson’s paradox or mathematical coupling, are not au fait with the problems of ROC curve analysis and have only been presented with CFR data rather than attributable fractions which are used by most public health professionals to determine policy.

We make it perfectly clear that we are not against the use of MUAC, quite the reverse. That is quite a different thing from complete suppression of WHZ.

We and others have repeatedly pointed out to those suppressing the use of WHZ entirely that a MUAC-only policy will lead to unnecessary deaths of large numbers of children and the only response we get is the assertion that they are healthy so do not need treatment. That is why we have collected the empirical data and done this analysis of the largest number of SAM children collected from many countries and programs to improve the external validity of our conclusions, conducted a detailed and exhaustive review of the literature and examined the effect of case load on potential numbers of deaths that will be missed. These analyses for the first time show that this is already a real threat to the children living in developing countries.

Nevertheless, we have replaced “faulty statistics” with “based upon the data presented in the first two papers of this series [16,17] as well as the present paper it is clear that an inappropriate statistical strategy has been exclusively used in the past using ROC curve analysis of entire populations to compare the relative CFRs [15]”.
“Completely false” has been replaced by:

“Fourth, repeated assertions that it is safe to ignore S-whz children because they are healthy when there are no data to support this contention and abundant data to show that these children are at high risk of death.”

We decline to rewrite the comment on misquotation and criticism of any data that does not support the proposition, if you take a view of the ENN-net posts on this topic you will see that this statement is correct.

• lines 320-321 – please revise the text accordingly – terms such as ‘deceived’ and ‘seduced’ are not justified

E A 36: We have replaced “deceived” by “misled” and do not think that a milder term is appropriate given the current situation. We agree that “seduced” is inappropriate and have replace this with:

“They are also attracted by the ease and low cost of MUAC screening; these practical aspects are clearly advantages to be considered.”