Reviewer’s report

Title: Current and potential role of grain legumes on protein and micronutrient adequacy of the diet of rural Ghanaian infants and young children: using linear programming

Version: 0 Date: 20 Nov 2018

Reviewer: Elaine Ferguson

Reviewer’s report:

General Comments:

The strengths of this study are the high-quality dietary data collected from a relatively large sample of children in Northern Ghana; and the authors modelled essential amino acids in addition to micronutrients, which is a unique contribution to the linear programming literature. However, the introduction needed strengthening, including recognition of study weaknesses. Please check the interpretation of references; and consider including a reference by Suri et al (2014), FNB 35:372-81 that models protein quality, using linear programming analyses, for legume and cereal-based complementary foods in Ghana. From my perspective, there were also a few issues with the analyses, and the authors need to fully address the apparent discrepancy between their findings on protein quality and quantity and the literature presented in the introduction. The manuscript would also benefit from a thorough editorial review.

Specific Comments:

Lines 25-26 - delete "largest" and "most strongly" because legumes will offer the benefits to all societies not just emerging economies and future food demand increases are not necessarily the strongest in emerging economies.

Lines 32- 33 - reword these sentences slightly i.e, delete "best". The modelling process does not guarantee it's the "best" FBRs to cover nutrient adequacy, because the selection of final FBRs is somewhat subjective. Also - it should be "linear programming analysis" not "linear modelling"

Line 35 - indicate if its 20g per day of cooked or uncooked legumes; and if uncooked whether they are mature/dry or immature/fresh.

Line 60 - is maize the most commonly consumed staple globally? Rice? Perhaps reword to "one of the most commonly consumed....." or "a commonly consumed...". Similarly, in lines 72-73 reword to "one of the highest prevalence…". Many non-African countries have high prevalence rates of stunting, especially South Asia; for example, India, Nepal, Afghanistan and Pakistan.

Line 66-67 - reword this sentence, as the meaning is not clear.

Line 68 - should this be "production" instead of "productivity"?
Lines 70-78 - rewrite this paragraph in a more nuanced fashion. The evidence presented is based on weak data. For example, evidence that protein intakes are inadequate in Africa, is not based on estimated dietary intake data but instead food disappearance data in relation to the protein requirements of an adult man. These weaknesses can also be considered when interpreting the results from this study in the discussion.

Line 77 - reference 22 is a commentary not a description of a trial; and reference 23 only describes a study that will be done it does not present results from the study.

Lines 78 and 89 - check interpretation of the references. For example, does reference 16 show a decline in the estimated prevalence of inadequate micronutrients in sub-Saharan Africa or instead show a decline in micronutrient densities. These analyses also used food balance sheets rather than consumption data; please modify if necessary.

Line 97 - reword the second objective slightly to "... identify a set of FBRs that will improve..." because "best" is difficult to define.

Line 98 - reword from "improve adequacy" to reduce "the number of ", because by definition "problem nutrients" are inadequate.

Lines 120-123 - unclear as written. Reword. Also - 100 children were not selected for each of the 4 groups as stated (lines 114-15). The sample size for the non-breastfed children is very small and well below the specified criteria of 100 children. Consider excluding the results for these 29 children. For linear programming analyses, the number of foods in the food lists will be low and food portion sizes will not be estimated robustly. These factors will limit the comparability of the results of this group to the other groups.

Line 143 - state what will be adjusted for day-to-day variation.

Lines 191-193 - for children under 12 months of age, with a few exceptions (for which there is an EAR), nutrient recommendations are Adequate Intakes. Population level dietary adequacy should not be assessed by calculating the percentage of children below the AI, as the AI could be well above actual requirements.

Line 220 - please check and correct - as far as I am aware 15% absorption is not assumed by iZiNCG for an unrefined diet.

Line 222 - change verb tense "meet" to "met"

Line 227-228 - modify - module 3 is maximizing and minimizing nutrients not diets.

Line 235 - was a cost constraint introduced or was this merely one of the criteria used to select the FBRs?

Line 252 - typo - should be 75th not 75ht
Line 278 - breast milk intakes were not quantified so it is not possible to definitively state energy intakes were below daily requirements. Perhaps these children had above average breast milk intakes? On the other hand, the high percentage of wasted children indicates energy intakes were inadequate, which would strengthen the point. Reword.

Line 280-283 -for children under 12 months of age the percentage at risk is over-estimated i.e., the percentage below the AI where the AI is an intake that likely exceeds requirements. Percentages at risk should not be estimated for the younger children. Instead compare nutrient densities to desired nutrient densities for all breastfed children to identify nutrients that are likely low in their diets; which will also allow interpretation across age groups.

Lines 307-309 - as far as I can tell these analyses did not control for age or breastfeeding status. Older children were more likely to consume legumes and would have higher energy and nutrient intakes, which makes it essential to control for these factors in the analyses.

Line 314 - delete the first line as it is not necessary i.e., it is not a result.

Line 319 - there seems to be some mixing of terminology for "best" diets. Module 2 were the nutritionally "best" diets and module 3 were "best-case scenario" diets. Change "best" diets to either "maximised" or "best-case-scenario" diets for module III.

Line 322 - were the "worst-case scenario" diet values for thiamine or vitamin B6 >70% when run with no FBR? If yes, then its okay to substitute the amino acids for these nutrients. If no, then either test the FBRs selected, for these two micronutrients, to ensure they are above 70% in the worst-case scenario analyses or choose alternative micronutrients to substitute, for example, vitamins A and C for the breastfed children.

Table 5 - state whether it's the number of servings per day or per week.

Line 364, 375 and 395 - typos? Should be IYC?

Line 395 - missing a "were" ie, …products were….

Lines 399-401 - are EAA at stake based on the results of the study? Consider modifying this statement. Also - do the authors have any insights into why the results from their study differed from others? Were legumes consumed by children in the other populations?

Lines 402-411 - for me, the interpretation of the results in this section - at least as worded - assumes the results, for protein, were incorrect. Even though its true, protein requirements may be elevated because of malabsorption, etc, this also applies to micronutrients. Based on the results of this study i.e., a high percentage of children had inadequate intakes of micronutrients but generally adequate intakes (both quantity and quality) of protein, would not the interpretation be that they are at high risk of micronutrient deficiencies but are only at risk of inadequate protein intakes if requirements are elevated? A critical discussion of these results in relation to the protein literature presented in the introduction would strengthen the discussion.
Line 419 - for lysine the % below daily requirements was only 17%. If the daily requirements are equivalent to recommended daily allowances, then 17% is an over-estimation of the percentage at risk of low intakes; so delete lysine from the list of AA which are a concern for the non-breastfed children.

Lines 423-425 - was the percentage of stunted children statistically higher in this group than the other groups? If not, then delete this statement. The sample size of non-breastfed children was very small i.e., this apparent difference might merely reflect sampling

Lines 460-463 and lines 491-492 - consider rewording, because the cut-off of 70% in the worst-case scenario analyses is an arbitrary choice i.e. cannot state it does not improve adequacy. Instead, these FBRs did not achieve the criteria selected to define a low risk of inadequate intakes for all children in the population.

493 - delete "more-or-less"

Line 590 - check this reference. It does not appear to be correct as written.

Table 2 - were the percentage at risk of inadequate iron intakes estimated using the full probability approach? Also - what percentage absorption was assumed for iron - 5% was used for modelling?

Table 3 - the SD are high. Should the contribution of legumes to energy and nutrient intakes be presented as medians and inter-quartile ranges instead of means given the high SDs? Also, for the daily legume intake estimates, specify in the footnote whether these intakes are only for consumers or for all children?

Level of interest
Please indicate how interesting you found the manuscript:

An article whose findings are important to those with closely related research interests

Quality of written English
Please indicate the quality of language in the manuscript:

Not suitable for publication unless extensively edited

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 organization 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