PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Short-term weight gain among preschool children in rural Burkina Faso: a secondary analysis of a randomized controlled trial
AUTHORS	Dennis, Elena; Sie, Ali; Ouermi, Lucienne; Dah, Clarisse; Tapsoba, Charlemagne; Zabre, Pascal; Bärnighausen, Till; O'Brien, Kieran; Lebas, Elodie; Keenan, Jeremy; Oldenburg, Catherine

VERSION 1 - REVIEW

REVIEWER	Aulo Gelli IFPRI, USA
REVIEW RETURNED	14-Mar-2019

GENERAL COMMENTS	Thanks for the opportunity to review this well-written paper. The main concern I have on the manuscript is that it doesn't really add much at all in terms of the literature and involves a very small sample of children taking part in a randomised trial on antibiotic, with two measurement points 1 month apart. Yes, food security is associated with weight gain in children during the lean season when diets are generally poor, that's not really a novel finding. In addition, it is not clear how and if some children had received antibiotics between the baseline and subsequent measurement of weigh related outcomes. In addition, as the age range of children is broad (6-59m), and the likely response to health insults, as well as the nutritional needs and food related behaviours (e.g. eating from common pot) will depend on age, there are a lot of confounding factors that the analysis is not considering. Having said that, the paper is well structured and sound. My suggestion is for the editor to think whether the scope of the paper warrants publication- the paper itself is of good quality. Specific comments: -Negative association with health visit in 30 days prior to baseline assessment and wasting after 1 month, this likely suggests that children that were more sick would tend to put on less weight -more common measure of dietary diversity (DD) in young children is measured over 1-d recall (according to WHO IYCF guidelines) and includes 7 food groups, this is the only validated measure of dietary diversity in children, authors may want to examine whether their results are sensitive to their DD indicator specification. -please add detail on whether some children received antibiotics between measurements

REVIEWER	Dr Marko Kerac
	London School of Hygiene & Tropical Medicine, UK
REVIEW RETURNED	01-Apr-2019

GENERAL COMMENTS	Overall this is a well written paper which adds interesting unsurprising data to the questions of which factors are responsible for poor short term weight gain in a nutritionally vulnerable population.
	Since it is a secondary analysis from a larger RCT and does not add much that is new or policy-changing it arguably belongs in a much more specialist rather than general journal.
	If offered a chance to revise there are a number of changes which could significantly improve the paper:
	MAJOR
	- The authors should make a much stronger and more persuasive case for why this analysis matters and what its policy implications are. It is well known (and indeed plainly obvious even without data – hence why a prospective design is not the massive advantage that authors seem to imply.) that food insecurity is a bad thing and is likely to lead to undernutrition. What should/ could be done differently in light of the study findings? Perhaps refence some current policy documents/strategies and say where these are suboptimal / could be changed for the better.
	- The authors correctly describe that power is low due to small numbers of children wasted. There is currently increasing interest in revisiting weight-for-age as a marker of malnutrition (since it captures those who are concurrently wasted and stunted https://www.ennonline.net//fex/59/wastingprevention) Hence I was especially surprised that authors had not done the same analysis (table 1; table 2) looking at underweight (low WAZ) status. This would anyway have been interesting as it adds another dimension of understanding to weight changes; would be especially important here since numbers with low WAZ likely to be larger and hence power may be greater in that analysis.
	Minor
	- Page 6:
	o top: please give reference numbers for ethics committee approvals
	o line 19 – please give more details about the background RCT
	o anthropometry assessment:
	please give more detail and justification of the 'median of three' measures. What if for instance you measured 101cm,101.2cm and 105cm height. Surely taking the average makes no sense as the 105cm measure likely to be the one that is wrong. How often / how large were the differences between measured? For future work, please see much more robust WHO method whereby two independent observers measure and a mean is taken if they two measures are within set limits of agreement (e.g. 0.7cm for height); they both re-measure if the difference is large.
	Was MUAC also taken three times?
	o Line 54 – please describe any 'cleaning criteria' for out of range / extreme measures. (see WHO anthro website for details of what these are) See also here for difference that different cleaning criteria can make https://peerj.com/articles/380/
	- Page 7

	o Please give some more details about the dietary diversity tool that was used. Validated were? Was it a special one for this study or a variant of DHS questionnaire? o Line 47: please clarify how you evaluated BF status. Did you just ask direct about exclusive breastfeeding (a bad way to do this since most mothers know what they should say..) or did you, as say DHS do, ask about other foods/fluids consumed and then calculate whether or not BF was exclusive (as is good practice and more standard in DHS and other surveys) - Results: o Table 1: ☐ Please also present weight-for-age and height/length-for-age and MUAC to give better impression of nutritional status. Also ☐ Whilst correct to give median age, it's much more common and expected in nutritional lit of this kind to present mean rather than median weight, WLZ, WAZ, HAZ, MUAC. This would make it easier to compare these results with others ☐ What exactly do you mean by 'BF status'? What happened to report of whether or not a child had previously been exclusively breastfed. - Discussion – OK but limited. As noted above in 'major comments' need to make much more of the 'so what' from these results. What's unexpected/surprising? What might realistically change of happen differently if policymakers and practitioners know these results?

Thanks for the opportunity to review this well-written paper. The main concern I have on the manuscript is that it doesn't really add much at all in terms of the literature and involves a very small sample of children taking part in a randomised trial on antibiotic, with two measurement points 1 month apart. Yes, food security is associated with weight gain in children during the lean season when diets are generally poor, that's not really a novel finding. In addition, it is not clear how and if some children had received antibiotics between the baseline and subsequent measurement of weigh related outcomes. In addition, as the age range of children is broad (6-59m), and the likely response to health insults, as well as the nutritional needs and food related behaviours (e.g. eating from common pot) will depend on age, there are a lot of confounding factors that the analysis is not considering.

RESPONSE: We appreciate the reviewer's comments. We agree that previous studies have demonstrated associations between food security and weight outcomes. We have addressed specific comments in detail below.

The use of antibiotics post-randomization was uncommon (N=5 children). We note that post-randomization use of an antibiotic could be on the pathway between baseline characteristics and weight gain outcomes (e.g., a mediator if any of the baseline characteristics are predictive of antibiotic use and also of weight gain), which could lead to biased estimates of the predictive value of the baseline characteristics if we adjusted for them. We have included additional details in response to reviewer comments below.

We included a number of baseline characteristics in our models, including age. Our goal was to evaluate predictive factors of weight gain over this short time period (covering approximately 25% of the lean season in Burkina Faso), and all candidate variables were retained in the model. This was intended to be a hypothesis-generating analysis to identify predictive factors for weight gain that could be considered for intervention in the future. We have added to the Introduction (Page 5, Line 33):

"Here, we used data from a randomized controlled trial to evaluate the prospective associations between dietary, socioeconomic, and demographic factors to identify possible modifiable risk factors for short-term weight gain during the beginning of the lean season in a rural area of Burkina Faso."

Having said that, the paper is well structured and sound. My suggestion is for the editor to think whether the scope of the paper warrants publication- the paper itself is of good quality.

RESPONSE: We thank the reviewer for their comments.

Specific comments:

-Negative association with health visit in 30 days prior to baseline assessment and wasting after 1 month, this likely suggests that children that were more sick would tend to put on less weight

RESPONSE: We agree that this is plausible. We have added to the Discussion (Page 13, Line 209):

"Children who had visited a health facility had increased odds of wasting and reduced WHZ at day 35, although this was not statistically significant in multivariable models. This is likely reflective of parents seeking care for malnourished children, and reduced weight gain was likely related to sick children gaining less weight."

-more common measure of dietary diversity (DD) in young children is measured over 1-d recall (according to WHO IYCF guidelines) and includes 7 food groups, this is the only validated measure of dietary diversity in children, authors may want to examine whether their results are sensitive to their DD indicator specification.

RESPONSE: We have recategorized the dietary diversity measurement into the 7-group measure as specified by the WHO as recommended by the reviewer and re-run all analyses with the updated measure. Results did not change qualitatively after re-categorized. We have also included a citation for the measure (Arimond et al, J Nutr, 2004). We have edited the Methods (Page 7, Line 92):

"The answers were made into a composite dietary diversity score by categorizing the food groups into 7 unique food groups, including starch, vitamin A-rich foods, other fruits and vegetables, animal protein (e.g., meat, eggs, poultry, fish), legumes, dairy, and fat (e.g., oil, butter, other fat).[3] We then summed the number of food groups reported for each child by the caregiver."

-please add detail on whether some children received antibiotics between measurements

RESPONSE: We have added to the Results (Page 10, Line 154):

"Caregivers of five children reported that their child received antibiotics outside of the study treatment during the course of the study."

And the Discussion (Page 13, Line 228):

"Few children were given antibiotics outside of the study treatment during the course of the study. Such antibiotic use may be influenced by baseline characteristics and could potentially be a mediator of any effect of baseline characteristics on nutritional outcomes."

Reviewer:

Reviewer Name: Dr Marko Kerac

Institution and Country: London School of Hygiene & Tropical Medicine, UK

Please state any competing interests or state 'None declared': None declared

Overall this is a well written paper which adds interesting unsurprising data to the questions of which factors are responsible for poor short term weight gain in a nutritionally vulnerable population.

RESPONSE: We thank the reviewer for their comments.

Since it is a secondary analysis from a larger RCT and does not add much that is new or policy-changing it arguably belongs in a much more specialist rather than general journal.

If offered a chance to revise there are a number of changes which could significantly improve the paper:

MAJOR

- The authors should make a much stronger and more persuasive case for why this analysis matters and what its policy implications are. It is well known (and indeed plainly obvious even without data – hence why a prospective design is not the massive advantage that authors seem to imply.) that food insecurity is a bad thing and is likely to lead to undernutrition. What should/ could be done differently in light of the study findings? Perhaps refence some current policy documents/strategies and say where these are suboptimal / could be changed for the better.

RESPONSE: We thank the reviewer for their comments. Cross-sectional analyses are often limited by reverse causality, because temporality cannot be established (and thus, some potential "predictor" variables may actually be influenced by the outcome). Although we agree that the findings are relatively obvious (food insecurity is bad), our overall goal was to identify independent predictors of weight gain as hypothesis generating, rather than testing a specific hypothesis about food security and weight gain. We also believe that having data-supported evidence to support claims or hypotheses, particularly evidence arising from prospective data which have the advantage of being able to establish temporality, are important for informing the development of interventions and supporting policy, even when relationships seem obvious. These results suggest that food insecurity prior to the lean season, above other potentially modifiable risk factors, is a major contributor to adverse nutritional outcomes in children during the lean season. We have clarified the overall goals of the analysis in the Introduction (Page 5, Line 31):

"Cross-sectional studies are limited by inability to determine temporality, and potential predictors may be influenced by outcomes of interest. Here, we used data from a randomized controlled trial to evaluate the prospective associations between dietary, socioeconomic, and demographic factors to identify possible modifiable risk factors for short-term weight gain during the beginning of the lean season in a rural area of Burkina Faso."

We have also added to the Discussion (Page 5, Line 31):

"The results of this study suggest that food insecurity, above and beyond other potential risk factors, is an important potentially modifiable risk factor for adverse nutritional outcomes. These findings underscore the importance of prioritizing policies related to improving food security in areas with seasonal malnutrition, as experiences of food insecurity immediately before the beginning of the lean

season may predispose children to worse outcomes during the course of the lean season. Interventions addressing food insecurity prior to the lean season, not only during the lean season, may help improve outcomes for children during this vulnerable time."

- The authors correctly describe that power is low due to small numbers of children wasted. There is currently increasing interest in revisiting weight-for-age as a marker of malnutrition (since it captures those who are concurrently wasted and stunted https://urldefense.proofpoint.com/v2/url?u=https-3A__www.ennonline.net__fex_59_wastingprevention&d=DwIFaQ&c=iORugZls2LlYyCAZRB3XLg&r=vPuVsbhJaqV4m3VEIoqDsZLIg8Ve2JJPhc2e-x3TsM8&m=vigeV_y7SBaOtXsWpKz6gBS0nfiZ55t0UHz50xPJEXk&s=FWBJdIyCYfnjl_17o5-BYpZBPPgsCwPG-Sg4hoDncsI&e=) Hence I was especially surprised that authors had not done the same analysis (table 1; table 2) looking at underweight (low WAZ) status. This would anyway have been interesting as it adds another dimension of understanding to weight changes; would be especially important here since numbers with low WAZ likely to be larger and hence power may be greater in that analysis.

RESPONSE: We appreciate the reviewer's suggestion to look at WAZ. We have now included an analysis looking at WAZ (continuously and dichotomized as underweight), which we have now included as Table 4. The results of these analyses were largely null. We have added to the Results (Page 11, Line 178):

"Table 4 lists bivariate and multivariable models for the association between candidate predictor variables and WAZ and underweight four weeks after baseline. Age was significantly associated with WAZ in the multivariable model (mean difference -0.005 SD per one-month increase in age, 95% CI -0.009 to -0.0008, P=0.02). No other candidate predictors were statistically significantly associated with WAZ or underweight."

Minor

- Page 6:

o top: please give reference numbers for ethics committee approvals

RESPONSE: We have added protocol numbers for the ethical committee approvals to the Methods.

• line 19 – please give more details about the background RCT

RESPONSE: We have added additional details about the background RCT as well as relevant references to the primary trial papers (Page 6, Line 60):

"After the baseline assessment, children were randomized in a 1:1:1:1 fashion to a 5-day course of placebo, amoxicillin, azithromycin, or cotrimoxazole [10]. All treatments were directly observed by study staff and administered as pediatric oral suspension. Children were followed for 35 days from enrollment for anthropometric outcomes [11]."

o anthropometry assessment:

please give more detail and justification of the 'median of three' measures. What if for instance you measured 101cm,101.2cm and 105cm height. Surely taking the average makes no sense as the 105cm measure likely to be the one that is wrong. How often / how large were the differences between measured? For future work, please see much more robust WHO method whereby two independent observers measure and a mean is taken if they two measures are within set limits of agreement (e.g. 0.7cm for height); they both re-measure if the difference is large.

RESPONSE: We appreciate the reviewer's suggestion for measurement in future studies. We used the median rather than the mean in this analysis because we agree that use of the mean would lead to over-weighting of outliers (e.g., the measurement of 105cm in the reviewer's example). The median in the reviewer's example would be 101.2cm, which is much more likely to be the truth given the data from the other measurements. We have added to the Methods (Page 7, Line 73):

"The median of the three measurements was used to avoid undue influence of outlying or implausible values."

Was MUAC also taken three times?

RESPONSE: MUAC was taken only a single time. We have clarified in the Methods (Page 7, Line 74):

"MUAC was measured a single time."

o Line 54 – please describe any 'cleaning criteria' for out of range / extreme measures. (see WHO anthro website for details of what these are) See also here for difference that different cleaning criteria can make https://urldefense.proofpoint.com/v2/url?u=https-3A__peerj.com_articles_380_&d=DwIFaQ&c=iORugZls2LlYyCAZRB3XLg&r=vPuVsbhJaqV4m3VEIo qDsZLIg8Ve2JJPhc2e-x3TsM8&m=vigeV_y7SBaOtXsWpKz6gBS0nfiZ55t0UHz50xPJEXk&s=w3Qfmd2lGgyE1eT7SKl1GQu HUPolncIFVK-sMv74Ygk&e=

RESPONSE: For consistency with previous published work, this analysis used the same criteria that was used in the primary anthropometry analysis, e.g., children with implausible weight change values were excluded from the analysis. We appreciate that different criteria can cause differences in results.

We chose a weight gain-based criteria (rather than based on Z-scores) because this was the primary outcome from the trial, and we elected to maintain this criterion in the present analysis because it was chosen prior to any data analysis by study arm or otherwise. We have added to the Methods (Page 9, Line 126) and cited the primary anthropometry analysis:

"Children with implausible weight changes between baseline and one-month measurements (gained or lost more than 2 kg) were assumed to be data entry errors (for example, the wrong child was measured), and were excluded from analyses."

- Page 7

o Please give some more details about the dietary diversity tool that was used. Validated were? Was it a special one for this study or a variant of DHS questionnaire?

RESPONSE: It was a variant of a DHS questionnaire. We have edited the analysis of the dietary diversity tool in response to Reviewer 1 and have added additional details about the tool, along with a citation for the tool (Arimond et al, J Nutr, 2004).

"The answers were made into a composite dietary diversity score by categorizing the food groups into 7 unique food groups, including starch, vitamin A-rich foods, other fruits and vegetables, animal protein (e.g., meat, eggs, poultry, fish), legumes, dairy, and fat (e.g., oil, butter, other fat).[3] We then summed the number of food groups reported for each child by the caregiver."

• Line 47: please clarify how you evaluated BF status. Did you just ask direct about exclusive breastfeeding (a bad way to do this since most mothers know what they should say..) or did you, as say DHS do, ask about other foods/fluids consumed and then calculate whether or not BF was exclusive (as is good practice and more standard in DHS and other surveys)

RESPONSE: We asked if the child was currently breastfeeding and if so if it was exclusive. Due to the age range of the children, we did not expect that most children would be exclusively breastfeeding (and only one mother reported that her child was exclusively breastfeeding). We have added to the Methods (Page 7, Line 85):

"Breastfeeding status was determined by asking the caregiver if the child was currently breastfeeding, and if so if the child was exclusively breastfeeding."

- Results:

o Table 1:

☐ Please also present weight-for-age and height/length-for-age and MUAC to give better impression of nutritional status. Also

RESPONSE: We have added WAZ, WLZ, HAZ, and MUAC to the table.

☐ Whilst correct to give median age, it's much more common and expected in nutritional lit of this kind to present mean rather than median weight, WLZ, WAZ, HAZ, MUAC. This would make it easier to compare these results with others

RESPONSE: We have changed the medians to means for the anthropometric indicators, including WLZ, WAZ, HAZ, and MUAC at baseline in Table 1.

☐ What exactly do you mean by 'BF status'? What happened to report of whether or not a child had previously been exclusively breastfed.

RESPONSE: Only a single child was reported to be exclusively breastfeeding. The measurement in Table 1 is any breastfeeding. We did not report exclusive breastfeeding here because it was so infrequent. We have clarified this in the table.

- Discussion – OK but limited. As noted above in 'major comments' need to make much more of the 'so what' from these results. What's unexpected/surprising? What might realistically change of happen differently if policymakers and practitioners know these results?

RESPONSE: We have attempted to add more "so what" to the Discussion, balancing our interpretation of the findings with the limitations of the study. We appreciate the reviewer's suggestion to improve on concrete conclusions of the study, although we would like to be cautious about over-interpretation (as with any study). We have added to the Discussion (Page 12, Line 200) as well as substantially reworked several parts of the Discussion:

"The results of this study suggest that food insecurity, above and beyond other potential risk factors, is an important potentially modifiable risk factor for adverse nutritional outcomes. These findings underscore the importance of prioritizing policies related to improving food security in areas with seasonal malnutrition, as experiences of food insecurity immediately before the beginning of the lean season may predispose children to worse outcomes during the course of the lean season. Interventions addressing food insecurity prior to the lean season, not only during the lean season, may help improve outcomes for children during this vulnerable time."