PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Protocol for the economic evaluation of a community-based intervention to improve growth among children under two in rural India (CARING trial)
AUTHORS	Skordis-Worrall, Jolene; Sinha, Rajesh; Kumar Ojha, Amit; Sarangi, Soumen; Nair, Nirmala; Tripathy, Prasanta; Sachdev, Harshpal; Bhattacharyya, Sanghita; Gope, Rajkumar; Rath, Shibanand; Rath, Suchitra; Srivastava, Aradhana; Batura, Neha; Pulkki-Brännström, Anni-Maria; Costello, Anthony; Copas, Andrew; Saville, Naomi; Prost, Audrey; Haghparast-Bidgoli, Hassan

VERSION 1 - REVIEW

REVIEWER	Kaushik Bose Vidyasagar University, India
REVIEW RETURNED	22-Apr-2016

GENERAL COMMENTS	This is a very good manuscript and merits early publication.

REVIEWER	Virginia Wiseman London School of Hygiene and Tropical Medicine, United Kingdom AND University of New South Wales Australia I currently work on a research project with one of the authors
REVIEW RETURNED	06-May-2016

GENERAL COMMENTS	**Protocol for the economic evaluation of a community-based intervention to improve growth among children under two in rural India (CARING trial)** This is a well written protocol that clearly explains aims and methods for the economic evaluation to accompany the CARING trial in India. I have a few suggestions on how to further strengthen the protocol and ensure its usefulness to readers of BMJ Open. It is important that full protocols for economic evaluations such as this are published – currently many are merged with the main trial protocol and consequently lack the detail required to be of real benefit to others working in the field. *Abstract* • The use of the word 'retrospectively' in the abstract is unclear. Is this data being collected after implementation or evaluation of the interventions? Please clarify. • A stated aim is to 'understand the fiscal space for

	investment in early child growth' (also highlighted under 'strengths'). How will this be done? Will, for example, the team undertake Budget Impact Analysis to assist in addressing this aim? • Suggest rewording of last bullet point under 'limitations' to: 'Community based interventions would require careful piloting and modification for use in other resource poor settings.' *Background* • Check wording of first sentence paragraph 3. • Para 5 should read '…a new community-based health worker engaged in improving feeding…' • Both here and when describing interventions, more detail on what is meant by 'undernutrition-focused participatory learning and action cycle' is needed. I realise that this information has been published elsewhere and it can be challenging reaching a balance between information provided in different protocols but interventions are key to each of these documents and right now, there is insufficient information on this aspect of the intervention. I suggest adding 1 or 2 sentences here. • Was the literature review (page 6) targeting all types of economic evaluations or just CEAs? • Of the studies listed/discussed on page 6, it is difficult to gauge robustness of the data – can authors add a little detail on study design? Again, this does not have to be detailed information but it would be useful when trying to interpret findings. *Methods* • It would be helpful to include rates of mortality and undernutrition in para 2 page 8. • Para 3 page 8 – define 'MTCs/NRCs' • Again, the main trial protocol has been published elsewhere but some basic information on the trial design would be useful (stratification, etc.). Also, was there a logic model showing expected effect of intervention on primary and secondary outcomes (including cost-effectiveness)? • Unclear whether the 60 clusters were randomly selected from across both districts (top page 9). • See earlier point about explaining 'participatory learning and action cycle' (para 2 page 9). • Unclear whether VHSNCs is equivalent to current practice – please explain. • It is stated on page 10 that the CEA will only be undertaken

	if there is a statistically significant effect. How about if the intervention is relatively cheap but not quite (almost) effective. Wouldn't policy makers be interested in this result? • The BIA (page 10) is rather brief. Where is the detailed utilisation data coming from? Analytical methods? • The CEA will use individual patient-level data on costs and effects from the cluster-randomized trial. Can the authors explain what statistical methods will be used to address issues such as correlation between costs and effects at the individual-level and cluster-level or intra-cluster correlation in costs and effects? Methods for covariate adjustment in the cost-effectiveness analysis? • Discounting both costs and effects? • Example of 'step down 'approach to costing (page 12) would be useful. • Suggest adding a table that briefly summarises each type of data needed for CEA, source of that data and sample size (i.e. indicate if data coming from sub sample or all participants). • I would like to see secondary outcomes listed somewhere and not just primary. • Page 13 – consumption and expenditure survey is described on page 13. Which instrument will be used? Has it been validated in this context? • Please explain/justify sub-samples (e.g. n=300 for C&E survey and n=120 for time use survey). • How confident are the authors that changes in C&E between 2 time points can be linked to your intervention? What statistical measures will be undertaken to minimise bias here? • The discussion of affordability at the top of page 15 is a little vague and it would be helpful if the authors could pin down measures to be used and steps involved. Not all readers will be familiar with MCDA – what data is needed? • Estimates of cost-effectiveness are inevitably associated with some degree of uncertainty. In addition to the point estimate for cost-effectiveness, a confidence interval around the estimate provides the decision-maker with a measure of the degree of uncertainty. Will cost-effectiveness acceptability curves be calculated/useful in this context? *Minor* • When using bullet points there are several bullet points missing. • Check spacing between paragraphs • Abstract - overuse of word 'robust' under section titled 'strengths'.

	• Page 8 – suggest re-wording: 'We will conduct a cost-effectiveness analysis alongside CARING, a…' • Page 11 – check wording second sentence under 'Programme related costs' • Finally, it is important for the authors to describe if/how their instruments and findings will be made accessible to the public.

REVIEWER	Jianduan Zhang Tongji Medical College, Huazhong University of Science and Technology China
REVIEW RETURNED	26-May-2016

GENERAL COMMENTS	Re: bmjopen-2016-012046, entitled "Protocol for the economic evaluation of a community-based intervention to improve growth among children under two in rural India (CARING trial)." Under-nutrition in childhood remains prevalent in some parts of the world and is a severe public health threat due to its immediate and long-term adverse outcomes. India has the highest health burden related to under-nutrition problems; therefore, numerous intervention programs targeting under-nutrition problem have been going on in the country. Through such economic evaluation as proposed in the study, the authors aim to understand the cost-effectiveness and cost-utility of the programs, which is of significant importance for the policy makers to make the best out of public health investment especially in the resource-limited setting. Overall, the proposal was well-written, and scientifically appeared sound. Here are several questions that the authors might want to consider for improvement. The growth of children is multifactorial, and it is important to assure the improvement in growth, if any, was resulted from the intervention programs before the cost-effectiveness was to be analyzed. What strategies will the authors use to rule out the improvement in child growth which might attribute to other factors, i.e., confounding factors? However, there is a lack of description in how the information is to be collected or how the calculation to be done. Detailed information is needed. There are six specific objectives of this proposed study as specified on page 7. Accordingly, in the Method section, the authors might also need to be specific e.g., how the indicators are to be obtained, for example, how the cost of the intervention to the health system is to be measured? The description of the method was lengthy and not easy to follow. A flow chart shall assist in the storytelling for a better understanding of the evaluation process.

REVIEWER	Devika Suri University of Wisconsin-Madison / Tufts University, USA
REVIEW RETURNED	22-Jul-2016

GENERAL COMMENTS	This manuscript outlines the protocol for a cost-effectiveness

	evaluation of a community-based intervention aiming to improve the growth of children under 2 in rural India. The results of this study will help fill a void in the evidence base of cost-effectiveness analysis in the context of interventions addressing health of young children in developing countries. The use of a societal perspective will help to capture costs at all levels, not just those of the provider. Furthermore, publishing the protocol, as the authors note, will help other researchers considering/planning CE analysis as well as improve standards and increase comparability of findings from different trials (such as through use of consistent indicators and cost data collected). To this end, the authors could further consider making their cost data collection instruments and analysis plan publicly available. Overall, this protocol is an extremely thorough, well-referenced and important addition to the research literature in this field. I look forward to the results of the cost-effectiveness evaluation of the CARING trial. Just a few specific questions/edits: Page 6 – "Waters et al [23] trialed **a** nutrition education programme**s** in Peru" A single program or multiple programs? Page 10 – "This cost-effectiveness analysis will only be conducted if a significant impact on the primary outcome is observed in the trial." How do you define significant impact; do you mean if the study group z-scores are *statistically* significantly different? Page 10 – How will the benefit incidence analysis be incorporated? Will there be a stratified analysis, or will the multidimensional poverty score be included in regression analysis, for example? Page 13 – Will potential differences in seasonality be captured in the comprehensive household consumption and expenditure survey? (maybe not within a household since data collection points are a year apart but among households?) How was the sub-sample size arrived at? (also for the time use survey)

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Kaushik Bose
Institution and Country: Vidyasagar University, India

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

Please leave your comments for the authors below
This is a very good manuscript and merits early publication.

Authors' Response: Many thanks for your positive review

Reviewer: 2
Reviewer Name: Virginia Wiseman

Institution and Country: London School of Hygiene and Tropical Medicine, United Kingdom AND University of New South Wales Australia

Please state any competing interests or state 'None declared':
I currently work on a research project with one of the authors

Please leave your comments for the authors below

Protocol for the economic evaluation of a community-based intervention to improve growth among children under two in rural India (CARING trial)
This is a well written protocol that clearly explains aims and methods for the economic evaluation to accompany the CARING trial in India. I have a few suggestions on how to further strengthen the protocol and ensure its usefulness to readers of BMJ Open. It is important that full protocols for economic evaluations such as this are published – currently many are merged with the main trial protocol and consequently lack the detail required to be of real benefit to others working in the field.

Authors' Response: Many thanks for your positive comments

Abstract
- _The use of the word 'retrospectively' in the abstract is unclear. Is this data being collected after implementation or evaluation of the interventions? Please clarify.

Authors' Response: These data are being collected toward the end of the trial implementation period and before the end of the evaluation. The word retrospectively has now been removed from the abstract.

- _A stated aim is to 'understand the fiscal space for investment in early child growth' (also highlighted under 'strengths'). How will this be done? Will, for example, the team undertake Budget Impact Analysis to assist in addressing this aim?

Authors' Response: We had intended to conduct a generalised fiscal space assessment as described in Tandon and Cashin (2010) and Heller (2006) (see reference 44 and 45 of the paper). This method is not significantly different from a Budget Impact Assessment but is accepted and understood both within and out-with the field of health spending. For clarity, the generalized fiscal space method is now explicitly mentioned on page 15 of the revised paper.

- _Suggest rewording of last bullet point under 'limitations' to: 'Community based interventions would require careful piloting and modification for use in other resource poor settings.'

Authors' Response: Thank you for your suggestion. We now have revised the sentence accordingly.

Background
- _Check wording of first sentence paragraph 3.

Authors' Response: This sentence has been revised to improve clarity.

• _Para 5 should read '…a new community-based health worker engaged in improving feeding…'

Authors' Response: Revised accordingly, thank you.

• _Both here and when describing interventions, more detail on what is meant by 'undernutrition-focused participatory learning and action cycle' is needed. I realise that this information has been published elsewhere and it can be challenging reaching a balance between information provided in different protocols but interventions are key to each of these documents and right now, there is insufficient information on this aspect of the intervention. I suggest adding 1 or 2 sentences here.

Authors' Response: Thank you for highlighting this omission. The following text has now been added to this section to better explain the participatory learning and action cycle:

"The participatory learning and action cycle involves monthly group meetings for women. The new community-based health worker facilitates the meetings and a problem-solving approach is used to help groups identify community level health and nutrition problems, and find locally feasible strategies to address the problems identified."

• _Was the literature review (page 6) targeting all types of economic evaluations or just CEAs?

Authors' Response: The literature review was targeting all type of economic evaluations, including both partial (cost analyses) and full-economic evaluation (CEA, CUA, CBA) analyses.

• _Of the studies listed/discussed on page 6, it is difficult to gauge robustness of the data – can authors add a little detail on study design? Again, this does not have to be detailed information but it would be useful when trying to interpret findings.

Authors' Response: We now have added more information on the design of the mentioned studies on page 6 and 7.

Methods
• _It would be helpful to include rates of mortality and undernutrition in para 2 page 8.

Authors' Response: We have added information on child mortality and undernutrition on page 8.

• _Para 3 page 8 – define 'MTCs/NRCs'

Authors' Response: We have added a brief definition of MTCs/NRCs and removes the acronym in favour of the full name: Malnutrition Treatment Centres. These are specialized health facilities that treat severe acute malnutrition. This revision now falls on page 9.

• _Again, the main trial protocol has been published elsewhere but some basic information on the trial design would be useful (stratification, etc.).

Authors' Response: The following text has now been added on page 9 to provide more detail on the trial design:
"…a total of 120 purposively selected geographical clusters, with an estimated total population of 121,531, were randomised to the two trial arms. The 60 clusters in the intervention arm received home visits, participatory learning and action cycle meetings, and support to Village Health Sanitation and Nutrition Committees as described further below. The 60 clusters in the control arm received only the support to Village Health Sanitation and Nutrition Committees. Study participants are pregnant women identified in the third trimester of pregnancy and their children (n = 2520). The randomisation,

which took place in July 2013, was stratified by district and by number of hamlets per unit."

Also, was there a logic model showing expected effect of intervention on primary and secondary outcomes (including cost-effectiveness)?

Authors' Response:

The following theory of change is shown in the trial protocol paper (Nair et al, 2015). We feel it is too detailed for inclusion in this paper and, as it is an open access publication, the interested reader can easily access the figure. Unfortunately cost effectiveness was not a secondary outcome of the CARING trial.

(Nair et al, 2015)

- _Unclear whether the 60 clusters were randomly selected from across both districts (top page 9).

Authors' Response: The clusters were purposively selected on the basis of population size (c. 1000 population). We have now clarified this in the text.

- _See earlier point about explaining 'participatory learning and action cycle' (para 2 page 9).

Authors' Response: Please see the response to the earlier comment, this is now expanded upon earlier in the paper.

- _Unclear whether VHSNCs is equivalent to current practice – please explain.

Authors' Response: VHSNCs are responsible for current practice but are not current practice in and of themselves. They act more as commissioning and monitoring bodies at a grass roots level.

- _It is stated on page 10 that the CEA will only be undertaken if there is a statistically significant effect. How about if the intervention is relatively cheap but not quite (almost) effective. Wouldn't policy makers be interested in this result?

Authors' Response: In the event of a trial outcome that is not statistically significant, the cost of the trial will be reported, as will the cost per unit of delivery. We cannot however, report cost effectiveness if no effect is found.

- _The BIA (page 10) is rather brief. Where is the detailed utilisation data coming from? Analytical methods?

Authors' Response: The title and wording of this section has been changed to refer instead to 'equity impact' rather than 'benefit incidence analysis'. The equity impact analysis will not use detailed utilization data but will instead decompose the gains from the intervention, by socio-economic status group. Apologies for any confusion caused by our terminology.

- _The CEA will use individual patient-level data on costs and effects from the cluster-randomized trial. Can the authors explain what statistical methods will be used to address issues such as correlation between costs and effects at the individual-level and cluster-level or intra-cluster correlation in costs and effects? Methods for covariate adjustment in the cost-effectiveness analysis?

Authors' Response: Measurement of the trial effect will be adjusted for clustering as described in Nair et al (2015). The majority of the costs collected for this trial are neither individual nor cluster level costs and will therefore not be subject to the risk of clustering effects. Only health-seeking costs will be subject to the risk of clustering. These costs will be adjusted for clustering using either the cluster-adjusted bootstrap or the Huber-White robust estimator of variance. While the number of clusters per arm may permit the conduct of the preferred bootstrap method, the final selection of the bootstrap versus the Huber-White method will need to be determined by the skewness of the cost data collected. (Flynn and Peters, 2004)

Flynn, T.N. and Peters, T.J. Use of the bootstrap in analysing cost data from cluster randomised trials: some simulation results, BMC Health Services Research, 2004, 4:33, DOI: 10.1186/1472-6963-4-33

- _Discounting both costs and effects?

Authors' Response: Both costs and effects will be discounted as mentioned in page 15, second paragraph.

- _Example of 'step down 'approach to costing (page 12) would be useful.

Authors' Response: The step down method is detailed later in that paragraph however, it may not be clear that this is what the text is describing. For clarity, the text has been edited as follows: "Using a step-down method, the main worksheets for entering data allocate costs to one of the following categories: staff, material, capital, and joint costs. Costs are also divided into start-up and implementation costs, and between intervention, monitoring and evaluation and research costs."

- _Suggest adding a table that briefly summarises each type of data needed for CEA, source of that data and sample size (i.e. indicate if data coming from sub sample or all participants).

Authors' Response: Thank you for your suggestion. We have prepared a table including information on perspective, data needed, source of data and sample size.

- _I would like to see secondary outcomes listed somewhere and not just primary.

Authors' Response: This trial has a very lengthy list of secondary outcomes as follows. The length of this list prohibits its inclusion in this paper but the list of trial outcomes can be provided in a web appendix if required.

Secondary outcomes of the CARING Trial (Nair et al 2015)

Anthropometry
- Mean weight for height z score at 18 months
- Mean weight for age z score at 18 months
- Mean MUAC z score at 18 months
- % of children who are stunted at 18 months
- % of children who are underweight at 18 months
- % of children who are wasted at 18 months
- Mean birth weight
- Change in weight from birth to 18 months
- Change in height from birth to 18 months
- Mean length for age z score at 72h, 3, 6, 9 and 12 months
- Mean weight for height z score at 72h, 3, 6, 9 and 12 months
- Mean weight for age z score at 72h, 3, 6, 9 and 12 months

• Mean MUAC z score at 6, 9 and 12 months
• Mean maternal MUAC in third trimester of pregnancy
• Mean maternal BMI 9 months after delivery

Health and nutrition in pregnancy
• % mothers receiving 3+ ANC by qualified provider
• % mothers receiving minimum IFA supplements and two TT injections
• % mothers with self-reported symptoms of anemia, vaginal bleeding or malaria
• % mothers seeking care for any of the above
• % mothers who delivered with a skilled attendant
• % mothers and children who received one or most postnatal check-up(s)
• n of months since previous birth
• Mean age at marriage
• Mean number of meals per day
• % mothers with minimum dietary diversity

Infant and Young Child Feeding
• % infants breastfed within one hour of birth
• % infants exclusively breastfed until 6 months
• % infants who started complementary foods at six months
• % infants still breastfed at one year
• % children given food from four or more groups at 6, 9, 12 and 18 months
• % children given minimum meal frequency at 6,9, 12 and 18 months
• % children given iron-rich foods at 6, 9, 12 and 18 months
• % children given a source of protein at 6, 9, 12 and 18 months
• Care-giving and care-seeking during illness
• % children with diarrhea, cough, fever in past 2 weeks
• % children receiving appropriate care during illness episode (fluid replacement for
• diarrhea and continued feeding for all illnesses)
• % children for whom care was sought from appropriate provider
• % children who received appropriate treatment from qualified provider

Infection control
• % children who received BCG, OPV3, DTP3, measles and Hep. B, Vitamin A dose in last 6 m (children 6-18 m)
• % of children who received deworming in the past 6 months
• % mothers who keep soap in the household
• % mothers who keep water container covered
• % mothers reporting handwashing with soap before feeding a child
• % mothers reporting handwashing with soap after helping a child to defecate
• % mothers reporting handwashing with soap after defecation
• % of mothers who report placing child feces in latrine or burying them

Mortality
• Infant mortality rate (per 1000 livebirths)

Receipt of ICDS entitlements
• % pregnant women who received Take Home Rations from AWW in previous month
• % mothers who received THR for their child from AWW in previous month
• % children weighed by AWW in previous month
• % mothers who received nutritional counselling by AWW in previous month
• % mothers who have received entitlement through Odisha's Mamta Scheme

Maternal mental health and decision-making
• % mothers with K10 score > 16 (indicating moderate to severe psychosocial distress)
• % mothers with a lot or complete control over decisions about their own healthcare
• % mothers with a lot or complete control over decisions about major household purchases
• % mothers with a lot or complete control over decisions about daily household purchases
• % mothers who have money of their own and over which they have complete control over

• _Page 13 – consumption and expenditure survey is described on page 13. Which instrument will be used? Has it been validated in this context?

Authors' Response: The consumption and expenditure survey was designed using Consumer Expenditure survey of National Sample Survey Office (NSSO) government of India and world Bank Survey of Living Conditions in India. The survey was pre-tested in both control and intervention areas.

• _Please explain/justify sub-samples (e.g. n=300 for C&E survey and n=120 for time use survey).

Authors' Response: Please see our detailed response to Reviewer 4 at the end of this document. That text provides a detailed description of the sub-sample selection and controls for seasonality in consumption. In our opinion, the text is too lengthy for inclusion in the paper. However, we could provide this text as a Web Appendix if the Editors would like to make it available. Similarly, if the Editors feel the value of the text outweighs its length, it can be incorporated into the main text of the article. We leave this decision to the Editors and the Reviewers.

• _How confident are the authors that changes in C&E between 2 time points can be linked to your intervention? What statistical measures will be undertaken to minimise bias here?

Authors' Response: The consumption study aims to compare the change in expenditure on food over this period, between intervention and control households. Change in expenditure is preferable to a measure of absolute consumption at a single time period, because households were not randomized to the intervention. As such, absolute consumption levels may be deemed endogenous to the trial effect. Change in consumption can be assumed to be an exogenous outcome. The consumption data will be analysed using a random effects panel model to account for intra-household variation, with controls for clustering. Price inflation will be controlled using data from a market survey of the price of common food items, conducted concurrently with the consumption survey. For any residual bias, we rely on the trial design.

• _The discussion of affordability at the top of page 15 is a little vague and it would be helpful if the authors could pin down measures to be used and steps involved. Not all readers will be familiar with MCDA – what data is needed?

Authors' Response: We accept that this section is vague and apologize for the lack of explicit detail on each of the listed methods. This is largely because of the number of measures we intend to use to measure the affordability of the trial. As we explain: "A wide range of affordability measures have been selected, in part, due to the paucity of evidence on the cost-effectiveness of comparable interventions." Providing methodological detail on all four of the methods listed would create a lengthy additional section to this paper. We have added a phrase briefly describing the fiscal space method and we already reference sources where would be further detail on these methods can be found. We would be willing to write a more detail section on the measurement of affordability if the Editors feel

the value of the text would outweigh its length. We leave this decision to the Editors and the Reviewers

- _Estimates of cost-effectiveness are inevitably associated with some degree of uncertainty. In addition to the point estimate for cost-effectiveness, a confidence interval around the estimate provides the decision-maker with a measure of the degree of uncertainty. Will cost-effectiveness acceptability curves be calculated/useful in this context?

Authors' Response: Thank you for this useful suggestion. We will calculate cost effectiveness acceptability curves and now say as much on page 15.

Minor

- _When using bullet points there are several bullet points missing.

Authors' Response: This should now be corrected

- _Check spacing between paragraphs

Authors response: We have tried to ensure this is more uniform but will check again once the paper is formatted for publication – assuming it is accepted for publication.

- _Abstract - overuse of word 'robust' under section titled 'strengths'.

Authors response: We have removed two of three uses of the word 'robust'.

- _Page 8 – suggest re-wording: 'We will conduct a cost-effectiveness analysis alongside CARING, a…'

Authors response: Reworded as follows; "A cost-effectiveness analysis of the CARING trial, a cluster randomised…"

- _Page 11 – check wording second sentence under 'Programme related costs'

Authors response: Reworded as follows; "Cost data are entered regularly into the tool, which is adapted each year to reflect the changing cost structure of the trial at different phases of activity."

- _Finally, it is important for the authors to describe if/how their instruments and findings will be made accessible to the public.

Authors response: Findings will be published in open access publications and instruments, including the cost-capture Excel tool will be made freely available on our University Department's website.

Reviewer: 3

Reviewer Name: Jianduan Zhang

Institution and Country: Tongji Medical College, Huazhong University of Science and Technology, China

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

Please leave your comments for the authors below:
Under-nutrition in childhood remains prevalent in some parts of the world and is a severe public health threat due to its immediate and long-term adverse outcomes. India has the highest health burden related to under-nutrition problems; therefore, numerous intervention programs targeting under-nutrition problem have been going on in the country. Through such economic evaluation as proposed in the study, the authors aim to understand the cost-effectiveness and cost-utility of the programs, which is of significant importance for the policy makers to make the best out of public health investment especially in the resource-limited setting. Overall, the proposal was well-written, and scientifically appeared sound. Here are several questions that the authors might want to consider for improvement.

The growth of children is multifactorial, and it is important to assure the improvement in growth, if any, was resulted from the intervention programs before the cost-effectiveness was to be analyzed. What strategies will the authors use to rule out the improvement in child growth, which might attribute to other factors, i.e., confounding factors? However, there is a lack of description in how the information is to be collected or how the calculation to be done. Detailed information is needed.

Authors' Response: Thank you for both the positive comments and your constructive question. You raise an important issue about the need to control for confounding when calculating the effect of the trial. In short, the randomized controlled trial design is the ideal method to control for confounding. The cluster randomized controlled design of the CARING Trial is briefly described on pages 5 and 8 of our paper, where a reference to the main trial protocol paper (Nair et al 2015) can also be found. In response to your comments, the description of the trial design has been enhanced and additional text can now be found on page 9.

The calculation of the primary and secondary effects/outcomes are described briefly page 10 of our paper. A fuller description of the trial methodology and the measurement of the trial effect falls outside of the remit of the Economic Evaluation Protocol and has been published elsewhere:
Nair, Nirmala, et al. "Participatory women's groups and counselling through home visits to improve child growth in rural eastern India: protocol for a cluster randomised controlled trial." BMC public health 15.1 (2015): 1.

There are six specific objectives of this proposed study as specified on page 7. Accordingly, in the Method section, the authors might also need to be specific e.g., how the indicators are to be obtained, for example, how the cost of the intervention to the health system is to be measured?

Authors' Response: The text on page 12 and 13 explains in detail how the cost of the intervention to the health system will be measured and how indicators will be obtained. However, this section is headed "Provider Costs" rather than "Health System Costs", which may have been the source of confusion and we apologize for our lack of clarity in this regard.

This heading has now been changed to "Provider or Health Systems Costs" to be more consistent with the wording of the objectives. The first line of the section on page 12 has also been edited slightly to clarify that the methods described below pertain to the measurement of health systems costs.

The description of the method was lengthy and not easy to follow. A flow chart shall assist in the storytelling for a better understanding of the evaluation process.

Authors' Response: Thank you for your suggestion, we have constructed Table One to provide an overview of the steps in the cost data collection. This can be found on page 12 of the revised draft.

Reviewer: 4
Reviewer Name

Devika Suri

Institution and Country

University of Wisconsin-Madison / Tufts University, USA

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

Please leave your comments for the authors below
This manuscript outlines the protocol for a cost-effectiveness evaluation of a community-based intervention aiming to improve the growth of children under 2 in rural India. The results of this study will help fill a void in the evidence base of cost-effectiveness analysis in the context of interventions addressing health of young children in developing countries. The use of a societal perspective will help to capture costs at all levels, not just those of the provider. Furthermore, publishing the protocol, as the authors note, will help other researchers considering/planning CE analysis as well as improve standards and increase comparability of findings from different trials (such as through use of consistent indicators and cost data collected). To this end, the authors could further consider making their cost data collection instruments and analysis plan publicly available. Overall, this protocol is an extremely thorough, well-referenced and important addition to the research literature in this field. I look forward to the results of the cost-effectiveness evaluation of the CARING trial.

Authors' Response: Thank you for the positive comments and the suggestion that we make our tools publically available. Once the findings are published, we will make the tools freely available (without data) on our website: www.ighe.org

Just a few specific questions/edits:
Page 6 – "Waters et al [23] trialed a nutrition education programmes in Peru" A single program or multiple programs?
Authors' Response: This was a single program and the typing error has now been corrected, thank you.

Page 10 – "This cost-effectiveness analysis will only be conducted if a significant impact on the primary
outcome is observed in the trial." How do you define significant impact; do you mean if the study group
z-scores are statistically significantly different?
Authors' Response: The impact of the trial would be considered significant if the mean length-for-age z-scores in the intervention group were significantly greater than in the control clusters and/or if the number of cases of stunting averted is significantly different from zero. Apologies for the lack of clarity, a statement to this effect has now been added on page 10.

Page 10 – How will the benefit incidence analysis be incorporated? Will there be a stratified analysis, or
will the multidimensional poverty score be included in regression analysis, for example?
Authors' Response: The analysis will be stratified by socio-economic status, where a household's multidimensional poverty score determines socio-economic status. The paragraph on page ten has

been edited slightly to improve clarity

Page 13 – Will potential differences in seasonality be captured in the comprehensive household consumption and expenditure survey? (maybe not within a household since data collection points are a
year apart but among households?) How was the sub-sample size arrived at? (also for the time use survey)

Authors' Response:

The following text provides a detailed description of the sub-sample selection and controls for seasonality in consumption. In our opinion, this text is too lengthy for inclusion in the paper. However, we could provide this text as a Web Appendix if the Editors would like to make it available. Similarly, if the Editors feel the value of the text outweighs its length, it can be incorporated into the main text of the article. We leave this decision to the Editors and the Reviewers.

Consumption Study

The consumption study had a sample of 300 households, with equal numbers from West Singhbhum district in the state of Jharkhand, and Keonjhar district in the state of Odisha. The sample size of 300 households was calculated using data from the Consumer Expenditure Survey for the year 2009-10, conducted by India's National Sample Survey Office (NSSO). These data were used to calculate a households' average expenditure on food staples and protein, as a proportion of total expenditure. Using a two-sample comparison of means with 90% power to detect change at the 95% confidence level and accounting for 10% attrition, a minimum sample of 270 was required. A sample of 300 households was thus selected to ensure that the minimum sample would be achieved.

Participating households with babies aged three months, were randomly selected from the main trial sample. The random sampling was stratified by intervention/control and district. For cases of attrition, new cases were selected following the same sampling standard and procedure.

To account for variation in seasonal consumption in inter-household analyses, the sample was selected in three 'tranches' of 100 over the first year. These three randomly selected tranches of 100 households were selected at 4 month intervals and data collection for that sample was then spread over the intervening period i.e. every month, data for 25 households were collected. Each household was interviewed twice over a two-year period. To control for seasonal variation in intra-household consumption, a household was interviewed in the same month in each year.

The consumption study aims to compare the change in expenditure on food over this period, between intervention and control households. Change in expenditure is preferable to a measure of absolute consumption at a single time period, because households were not randomized to the intervention. As such, absolute consumption levels may be deemed endogenous to the trial effect. Change in consumption can be assumed to be an exogenous outcome.

Time Use Study

The time use study had a sample of 120 households, with equal numbers from West Singhbhum, district of the state of Jharkhand and Keonjhar, district of the state of Odisha.

To select the time use sub-sample, mothers participating in the main trial were first stratified by the

age of their children (13-18 month) and their tribal status. After stratification, 30 households were randomly selected from the intervention and control areas in each of the two study districts.
The time use study was a cross-sectional survey conducted in the month of June 2016 in both districts. The survey was conducted on days not normally set aside for festival/religious activities and respondents were asked to reference the most recent 'normal day' in their responses.

VERSION 2 – REVIEW

REVIEWER	Virginia Wiseman London School of Hygiene and Tropical Medicine, UK & University of New South Wales, Australia I am a co-investigator on another study with one of the authors.
REVIEW RETURNED	08-Sep-2016

GENERAL COMMENTS	The authors have one a very good job at addressing my comments. These changes should help ensure that the paper is of maximum value to those wishing to design similar studies in other settings. A couple of minor issues: Methods – The manuscript says: 'The randomisation, which took place in July 2013, was stratified by district and by number of hamlets per unit.' Unclear what a 'unit' represents. Is it a cluster? Methods – My comment about circumstances under which a CEA would be undertaken were unclear. My question is: If the intervention is effective but the effect is not statistically significant then it may still be useful to policy-makers to see ICERs especially if the intervention is relatively cheap? I had requested that some justification be given for sub-samples (e.g. n=300 for C&E survey and n=120 for time use survey). Web appendix might be useful but not essential – defer to editor. Table 1 – this is very helpful to interpreting the methods. Suggest indicating those samples that are random and those that are non-random - this is currently done for some samples and not others. Ethics – Approval was obtained in June 2013 for the main trial. Please confirm that approval is still current and covers collection of cost data for the economic evaluaiton.

REVIEWER	Jianduan Zhang Huazhong University of Science and Technology
REVIEW RETURNED	03-Sep-2016

GENERAL COMMENTS	To reduce the prevalence of stunting in children is one of the major purposes of the trial and the authors addressed the severity of stunting problem in the background section. In order to stay consistent, it would be more sensible if the authors can alter the indicator of stunting with underweight in the study setting section.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2
Reviewer Name: Virginia Wiseman

Institution and Country: London School of Hygiene and Tropical Medicine, United Kingdom AND University of New South Wales Australia

Please state any competing interests or state 'None declared':
I currently work on a research project with one of the authors

Please leave your comments for the authors below

A couple of minor issues:

Methods – The manuscript says: 'The randomisation, which took place in July 2013, was stratified by district and by number of hamlets per unit.' Unclear what a 'unit' represents. Is it a cluster?

Authors' response: Yes, it is the number of hamlets per cluster. Apologies for the ambiguity, this has now been corrected on page 9 of the text.

Methods – My comment about circumstances under which a CEA would be undertaken were unclear. My question is: If the intervention is effective but the effect is not statistically significant then it may still be useful to policy-makers to see ICERs especially if the intervention is relatively cheap?

Authors' response:

We fully understand the tension to which the reviewer is referring – time and effort are invested in the generation of cost data and wherever possible, those cost data should be used to inform policy. In our opinion however, any effect that is not statistically significant cannot be considered to be greater than zero with any statistical confidence. Any spending on a program that yields an effect no greater than zero but incurs positive expense can surely not be considered cheap? Perhaps this tension warrants an article of its own to frame the debate around the use of non-significant results in cost-effectiveness analyses and guide best practice guidelines? If the reviewer would like to collaborate on such a paper, we would be keen to work together.

Within this protocol however, we have sought to clarify the specific conditions under which we will conduct cost-effectiveness analyses for stunting, wasting, underweight and infant mortality on p.10. If ICERS cannot be calculated for any outcomes, total and average cost data will be presented along with a discussion of any likely benefits of implementation that may not be represented in the statistical or epidemiological data. The revised text reads as follows:

Primary outcome

"The primary outcome of the CARING trial is children's length-for-age z-scores (17). The trial was powered to detect a 0.15 SD difference in LAZ between intervention and control at the 0.05 significance level. A greater sample size would be required to detect a realistic change in the prevalence of stunting: our sample size allows us to detect a 13% reduction in stunting prevalence, which is an unfeasible effect for a 24 months intervention without food supplementation. Cost-effectiveness analyses focusing on cases of stunting averted would be useful to inform policy-makers' decisions. We therefore propose to conduct a cost-effectiveness analysis for cases of stunting

averted if there is evidence of impact on the prevalence of stunting in children at 18 months at the 0.1 significance level or below. Using these z-scores, we will calculate the prevalence of stunting among children at 18 months and calculate the number of cases of stunting averted. Cases averted will be calculated as the difference between the expected and the actual number of cases using the adjusted odds ratio."

Secondary outcomes

"The trial has a number of secondary outcomes described in detail in the trial protocol (17). Cost-effectiveness analyses will be conducted for any reduction in three key secondary outcomes – underweight, wasting, and infant mortality – where a significant impact on these outcomes is observed at the 0.1 level.
Incremental cost-effectiveness ratios (ICERs) will thus be calculated for wasting, underweight and infant mortality, if a significant effect is observed at a p-value of 0.1 or less.. Estimates of life years saved will be calculated for infant mortality averted by the intervention if a significant finding is observed. Although no agreed DALY weight currently exists for undernutrition in any form, expected 'DALYs averted' will be modelled using mortality averted by the intervention if significant, and evidence from the wider literature about morbidity gains from averting stunting, wasting and underweight in young children (31)."

I had requested that some justification be given for sub-samples (e.g. n=300 for C&E survey and n=120 for time use survey). Web appendix might be useful but not essential – defer to editor.

Authors' response:

We had provided the following text in response to this request previously. If this does not provide the level of detail required, please don't hesitate to let us know. We defer to the Editorial team about the inclusion of this text in the main paper or a web appendix:

"The following text provides a detailed description of the sub-sample selection and controls for seasonality in consumption. In our opinion, this text is too lengthy for inclusion in the paper. However, we could provide this text as a Web Appendix if the Editors would like to make it available. Similarly, if the Editors feel the value of the text outweighs its length, it can be incorporated into the main text of the article. We leave this decision to the Editors and the Reviewers.

Consumption Study

The consumption study had a sample of 300 households, with equal numbers from West Singhbhum district in the state of Jharkhand, and Keonjhar district in the state of Odisha. The sample size of 300 households was calculated using data from the Consumer Expenditure Survey for the year 2009-10, conducted by India's National Sample Survey Office (NSSO). These data were used to calculate a households' average expenditure on food staples and protein, as a proportion of total expenditure. Using a two-sample comparison of means with 90% power to detect change at the 95% confidence level and accounting for 10% attrition, a minimum sample of 270 was required. A sample of 300 households was thus selected to ensure that the minimum sample would be achieved.

Participating households with babies aged three months, were randomly selected from the main trial sample. The random sampling was stratified by intervention/control and district. For cases of attrition, new cases were selected following the same sampling standard and procedure.

To account for variation in seasonal consumption in inter-household analyses, the sample was selected in three 'tranches' of 100 over the first year. These three randomly selected tranches of 100 households were selected at 4 month intervals and data collection for that sample was then spread over the intervening period i.e. every month, data for 25 households were collected. Each household was interviewed twice over a two-year period. To control for seasonal variation in intra-household consumption, a household was interviewed in the same month in each year.

The consumption study aims to compare the change in expenditure on food over this period, between intervention and control households. Change in expenditure is preferable to a measure of absolute consumption at a single time period, because households were not randomized to the intervention. As such, absolute consumption levels may be deemed endogenous to the trial effect. Change in consumption can be assumed to be an exogenous outcome.

Time Use Study

The time use study had a sample of 120 households, with equal numbers from West Singhbhum, district of the state of Jharkhand and Keonjhar, district of the state of Odisha.

To select the time use sub-sample, mothers participating in the main trial were first stratified by the age of their children (13-18 month) and their tribal status. After stratification, 30 households were randomly selected from the intervention and control areas in each of the two study districts.
The time use study was a cross-sectional survey conducted in the month of June 2016 in both districts. The survey was conducted on days not normally set aside for festival/religious activities and respondents were asked to reference the most recent 'normal day' in their responses.

Table 1 – this is very helpful to interpreting the methods. Suggest indicating those samples that are random and those that are non-random - this is currently done for some samples and not others.

Authors' response: Thank you for the suggestion, we have added further detail to Table One.

Ethics – Approval was obtained in June 2013 for the main trial. Please confirm that approval is still current and covers collection of cost data for the economic evaluaiton.

Authors' response: We can confirm that the Ethics approval is still current and covers collection of the cost data for the economic evaluation.

Reviewer: 3

Reviewer Name: Jianduan Zhang

Institution and Country: Tongji Medical College, Huazhong University of Science and Technology, China

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

Please leave your comments for the authors below:

To reduce the prevalence of stunting in children is one of the major purposes of the trial and the

authors addressed the severity of stunting problem in the background section. In order to stay consistent, it would be more sensible if the authors can alter the indicator of stunting with underweight in the study setting section.

Authors' Response:

Thank you for identifying this inconsistency. We have now provided local measures of stunting for the study setting.

The CARING trial aims to improve growth in children as measured by length for age z-scores. We have attempted to clearly specify what would be done for the primary outcome (mean for length-for-age z scores at 18 months) and secondary outcomes of special interest (the % of children who are stunted; wasted; underweight at 18 months; infant mortality) on pp.10-11.Stunting is a manifestation of undernutrition, which is also captured by underweight and wasting, two of the trial's secondary outcomes. This is why we propose to calculate ICERs for secondary anthropometric outcomes like wasting and underweight, which are marks of undernutrition and related to growth.

We hope you will find this clearer and more consistent with the text in the background and the study setting