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

<table>
<thead>
<tr>
<th>TITLE (PROVISIONAL)</th>
<th>Testing the socio-economic and environmental determinants of better child-health outcomes in Africa: a cross-sectional study among nations</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTHORS</td>
<td>Annamalay, Alicia; Heft-Neal, Sam; Wagner, Zachary; Le Souëf, Peter</td>
</tr>
</tbody>
</table>

VERSION 1 – REVIEW

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Corey Chivers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>University of Pennsylvania Health System, United States</td>
</tr>
<tr>
<td>REVIEW RETURNED</td>
<td>27-Mar-2019</td>
</tr>
</tbody>
</table>

| GENERAL COMMENTS    | The authors have presented a novel analysis of country-scale variation in child health outcomes and their potential socio-economic correlates. The data preparation analysis is clearly presented and appears to be carefully conducted and reported. Overall, the manuscript provides clearly reported findings, however, it is not clear that conclusions are justified by the results. My specific concern is that the authors conclude that “Child health was lowest when water quality, improved sanitation, air quality, and environmental performance were lowest”. While the results of the main statistical approach (boosted regression trees) do suggest these to be important factors, the alternative models described in the supplementary material suggest that GDP and Governance explain by far the most variance in outcomes. On line 339, they conclude “These results are all broadly consistent between the boosted regression trees, structural equation models, and the general linear mixed effects models”. As a reader, this conclusion does not seem to be supported by the information presented in the tables (main and supplementary). It is also not clear why the authors chose to present the boosted regression trees in the main manuscript over the other two modeling approaches. Minor issues: - From a study design perspective, the two main weaknesses are the temporal aspect, and the within country scale variance. The temporal limitation is given in the summary and stressed in the discussion. However, with respect to within country variance, the authors mention only a related limitation, specifically that they “cannot identify sub-national patterns of success of project-based health interventions.”. The authors could highlight the inability to identify potentially important within-country variance (potentially a Gini coefficient) which may be an important driver of child health outcomes. |

- With respect to population density, the authors mention "proximate population density" (line 231) and "As we predicted, increasing household size, which we argue is a better indicator of the population-density conditions experienced by the average child than country-level population density itself, was correlated with worsening child health" (line 388). However, there is no mention of this specific prediction earlier on, nor is it made clear up front that household size is a proxy for proximate population-density, not landscape-scale density (which can be inverted - as in the high fertility rural vs low fertility urban areas).

- It looks like household size is a 2nd order effect. While high-density households have a more negative effect than low-density, the best outcomes in the middle (although mildly). If this is a correct reading of figure 2, it may be worth mentioning in the results and discussion.

- Did the authors bootstrap re-sample 1,000 (line 58) or 10,000 (line 67) times?

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Victor Chima</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obafemi Awolowo University, Nigeria</td>
<td></td>
</tr>
</tbody>
</table>

| REVIEW RETURNED | 29-Mar-2019 |

| GENERAL COMMENTS | The outcomes were defined but there is need for clarity in understanding deeper issues responsible for the outcomes. Although, these might not be solely determined by quantitative data but adding causal variables into the analysis might help get some of the factors responsible for the outcome. Are the discussion and conclusions justified by the results? If some causal variables are examined. The discussion might proffer region, country or cluster specific intervention to to address the identified factors in the analysis. If these are done, the study would enable government, stakeholders and funders use specific program initiatives proposed by the study in improving child health outcomes |

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Louis Niessen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liverpool School of Tropical Medicine / Johns Hopkins School of Public Health</td>
<td></td>
</tr>
<tr>
<td>I do integrated mathematical modelling of global change. It includes child health and determinants.</td>
<td></td>
</tr>
</tbody>
</table>

| REVIEW RETURNED | 08-May-2019 |

| GENERAL COMMENTS | The study describes the contribution of multi-sectoral determinants of child health. This is a very timely and appropriate study and a contribution to the SDG agenda and wider approaches to human/child health. It is well presented and the conclusions are based on the findings. I normally give a lot of comments to improve the paper, yet this is a first one in years that I suggest to publish without major revisions. The BMJ editors will have to decide whether the methodologies are appropriate to be part of the main paper, given the general health audience. One major point |
The Discussion needs to address the weaknesses and limitations of an statistical ecological approach in the epi sense. The study produces correlates and hence there are always questions about causalities. Timewise all determinants and child health indicators move into the same directions.

**VERSION 1 – AUTHOR RESPONSE**

**Reviewer 1**

*My specific concern is that the authors conclude that “Child health was lowest when water quality, improved sanitation, air quality, and environmental performance were lowest”. While the results of the main statistical approach (boosted regression trees) do suggest these to be important factors, the alternative models described in the supplementary material suggest that GDP and Governance explain by far the most variance in outcomes. On line 339, they conclude “These results are all broadly consistent between the boosted regression trees, structural equation models, and the general linear mixed effects models”. As a reader, this conclusion does not seem to be supported by the information presented in the tables (main and supplementary). It is also not clear why the authors chose to present the boosted regression trees in the main manuscript over the other two modeling approaches.*

**RESPONSE:** The simplest response here is that we opted to present and rely on the results derived from the boosted regression trees because they are by far the most robust of all the methods we applied. Indeed, boosted regression trees easily handle non-linearities, they have no convergence or tolerance issues, and n-order interactions are permitted (in our case, we opted for the most parsimonious first-order interactions because beyond n = 1, interpretation becomes difficult). There are several other advantages. We have now justified this better in the Methods section with the following added text:

> “The latter setting considers only first-order interactions and combines these effects if present into the relative influence scores (see below). Boosted regression trees are particularly suited to the data and relationships we hypothesized because such hybrid statistical/machine-learning techniques do not assume an appropriate data model to build the regression trees like purely statistical techniques do; instead, they apply an algorithm to learn the relationship between the predictors (socio-economic variables) and the response (child health)². This places emphasis on predictive performance rather than on the underlying model per se.¹ Other advantages of boosted regression trees over many other statistical techniques is that they can easily combine predictor variables of any type, results are unaffected by different scales of measurement, results are insensitive to outliners, they can accommodate missing data in the predictors, they automatically model interactions, and the fitting process is stochastic such that it generally improves predictive performance of the final model.³ ³.*

The real question then is why we also chose to include the results from both structural equation models and generalised linear mixed-effects models at all. While inferior in most ways to boosted regression trees, each of the other approaches offers at least one advantage: (i) for structural equation models, the covariance structure among the hypothesised predictors can be accounted for
directly, and (ii) mixed-effects models explicitly account for non-independence of data within explicit categories designated as random effects (i.e., spatial non-independence in our cases).

That said, the structural equation models failed dismally to capture the inter-country variance in child-health indicators as demonstrated by their universally low goodness-of-fit (Tables S1-S6). As such, we cannot trust the specific, relative rankings of the dominant predictors to the same degree as for the boosted regression trees.

While the GLMM results are better, we still have the same issues regarding the inferior fitting capacity of GLMM relative to boosted regression trees, and the obvious advantages of using machine-learning for the type of data we analysed.

What is therefore important here is not maintenance of the exact predictor rankings among all three techniques; instead, it is the general conclusion that the ‘main’ predictors identified by the boosted regression trees also show up in the top-ranked models of the other two approaches. This is generally what occurred. Improved water and sanitation, GDP, household size, and environmental performance are still maintained in both the structural equation models and GLMM results, even though the boosted regression trees did not explicitly account for inter-predictor covariance (the advantage of structural equation models) or potential spatial autocorrelation (the advantage of random effects in the GLMM). Thus, our results are demonstrably robust.

We have therefore added the following text to the Methods to clarify this idea:

“Model results were broadly similar between the GLMM and boosted regression trees, although there were some differences in terms of relative ranking of the most influential predictors (Supplementary Information). However, given the lack of a strong goodness-of-fit for the structural equation models, and the clear advantages of boosted regression trees over GLMM in particular, we contend that the boosted regression trees used to rank the relative importance of the predictors considered provide more robust and credible results than those derived from the other methods. The results of the structural equation models and GLMM merely provide additional confidence that the results arising from the boosted regression trees are not compromised by potential spatial autocorrelation (GLMM) or explicit relationships between predictor variable (structural equation models).”

As well as the following text at the end of the Results:

“These results are all broadly consistent between the boosted regression trees, structural equation models, and the general linear mixed-effects models (i.e., the dominant predictors appeared within the main, top-ranked models for each method), although some differences in predictor rankings were apparent (see Supplementary Information).”
From a study design perspective, the two main weaknesses are the temporal aspect, and the within country scale variance. The temporal limitation is given in the summary and stressed in the discussion. However, with respect to within country variance, the authors mention only a related limitation, specifically that they “cannot identify sub-national patterns of success of project-based health interventions.” The authors could highlight the inability to identify potentially important within-country variance (potentially a Gini coefficient) which may be an important driver of child health outcomes.

RESPONSE: Agreed. This is an important point, so we have added the following sentence to the end of the second paragraph of the Discussion:

“Nor can our data identify potentially important within-country (sub-national) relationships that might operate at finer spatial scales.”

With respect to population density, the authors mention “proximate population density” (line 231) and “As we predicted, increasing household size, which we argue is a better indicator of the population-density conditions experienced by the average child than country-level population density itself, was correlated with worsening child health” (line 388). However, there is no mention of this specific prediction earlier on, nor is it made clear up front that household size is a proxy for proximate population-density, not landscape-scale density (which can be inverted - as in the high fertility rural vs low fertility urban areas).

RESPONSE: We agree that more justification is needed, but we had already included this hypothesis already in the last paragraph of the Introduction:

“… we hypothesized that demographic competition could in itself compromise child health given the negative relationships observed between health performance (children and adults) and household size (or number of siblings and/or population density)4-9; we therefore predicted a negative correlation between child health and household size for African nations.”

The problem with using gross, country-averaged values of population density as a correlate to explain variation in a host of phenomena, is, as the reviewer rightly points out, potentially fraught because of the non-Gaussian behaviour of this variable across space (e.g., the rural/urban divide mentioned). Thus, spatial variation in population density is probably best used at subnational levels7. Another problem with using gross population density is that it does not capture the age structure of a population if the actual living conditions of people varies markedly from one country to another. For example, high-density living could in fact mean low household size and high resource availability per child if there is a higher adult-to-child ratio compared to another situation with lower overall population density, but a low adult-to-child ratio (meaning higher household density).
Indeed, high-dependency households (dependency defined as “the total number of children under age 16 in a household”) had higher proportions of stunted and underweight children than lower-density households in Ghana^4, with the effect of the number of children in the household mediated by nucleation (i.e., the tendency toward households comprised of two parents and their own children) and conditioned on the wealth status of the household^4. Further, household crowding increased mortality from measles in West African communities^10. These phenomena likely arise because in extended-family households, resources are more likely to flow from children to adults, whereas in more nucleated families the resource flow tends to flow more from parents to children^11.

We have therefore added the following justification to the last paragraph of the section ‘Environmental and socio-economic data’ in the Methods where we elaborate on the hypothesis between household size and child health:

“It looks like household size is a 2nd order effect. While high-density households have a more negative effect than low-density, the best outcomes in the middle (although mildly). If this is a correct reading of figure 2, it may be worth mentioning in the results and discussion.”

RESPONSE: We do not think that this can be interpreted from Fig. 2D, mainly because inferring nonlinearity from such a response is most likely a visual artefact of examining only the median trend. When put into context of the bootstrapped confidence intervals, it is impossible to identify any nonlinearity with confidence. We therefore think it is best not to speculate beyond the overall decline in child health with increasing household size.

Did the authors bootstrap re-sample 1,000 (line 58) or 10,000 (line 67) times?
RESPONSE: Apologies. The bootstraps were based on 1000 samples. Now correctly stated.

Reviewer 2

The outcomes were defined but there is need for clarity in understanding deeper issues responsible for the outcomes. Although, these might not be solely determined by quantitative data but adding causal variables into the analysis might help get some of the factors responsible for the outcome. If some causal variables are examined.

RESPONSE: Nowhere do we claim causality. If we had access to experimental data at finer spatial scales, then the question of ‘causal’ might be able to be addressed. There is certainly no way to determine causality per se at the among-nation scale of investigation we present, because there is no way to control for all extenuating circumstances across an entire continent. Rather, we established all of our hypotheses based on finer-scale relationships (most of them being correlational as well) to determine whether they also explain variation in child health at the continental scale. We clearly articulated this concept in the Introduction when we stated:

“… it is unclear if generalizable ‘rules of thumb’ regarding the relative effectiveness of different interventions can be formulated across greater spatial scales.”

and

“Our aim was to test hypotheses regarding the principal correlates of child-health performance among African nations based on previous evidence collected at finer spatial scales.”

What is essential here, and indeed, the basis for all broad-scale scientific investigation, is the consistency with which phenomena hold across different spatial (and temporal scales). Thus, in cases where experimental manipulation to determine causality is difficult or impossible, confirmation of relationships based on different datasets, scales of investigation, and under different circumstances all act to increase the strength of inference from correlation toward causation.

To make this clearer, we have carefully gone through the manuscript and avoided terminology that could be confused as inferring causality (including changing the word ‘determinants’ to ‘correlates’ in the title), although we maintain the directionality implied in the hypothesized relationships. We have also compiled a new Table (new Table S1 in the Supplementary Information and Results) that summarizes the African evidence for child health-correlates across different spatial scales.
The discussion might proffer region, country or cluster specific intervention to address the identified factors in the analysis. If these are done, the study would enable government, stakeholders and funders use specific program initiatives proposed by the study in improving child health outcomes.

**RESPONSE:** As mentioned above, we have provided a new summary table (new Table S1) in the Supplementary Results and Information describing the various global, regional, sub-national and other scales of previous studies examining the relationships between child-health metrics and various environmental and socio-economic variables. We hesitate to be too prescriptive to individual national or regional governments in this regard though because of the potential scale mismatch between our results and effective interventions. As we originally stated in the Conclusion section:

“… the complex interaction of environmental and socio-economic conditions that influence child-health outcomes in Africa makes generalizations regarding the prioritization of interventions difficult, and the most effective, regionally specific interventions will necessarily be highly variable”

Reviewer 3

*The Discussion needs to address the weaknesses and limitations of a statistical ecological approach in the epi sense. The study produces correlates and hence there are always questions about causalities. Timewise all determinants and child health indicators move into the same directions.*

**RESPONSE:** We have addressed the issue of ‘causality’ versus ‘correlation’ in our response to Reviewer #2; however, we agree that some more discussion of the limitations is warranted. In particular, we acknowledge specifically that our results are likely to be temporally dependent, especially with respect to the relative importance of the different correlates examined, as we stated in the Discussion:

“We stress that these relationships are derived from the most recent data and therefore represent the present snapshot in time; thus, they do not necessarily reflect past child health outcomes and socio-economic conditions in African nations\(^2\) that might have been historically more important than those today (e.g., food supply).”

Unfortunately, we are restricted to the most recent period to investigate these relationships, mainly because many of the correlates lack comparable temporal replicates (e.g., many of the variables making up the environmental performance index, such as ecological footprint, megafauna conservation index, comparable IUCN Red List data, forest loss).
But to make these and other limitations clearer, we have therefore added the following text to the Discussion:

“Of course, another limitation of an analysis of this spatial scale is that even with statistical evidence for the hypothesized relationships, we cannot necessarily infer causality. However, that nearly all of the hypothesized relationships derived mainly from smaller-scale studies examining single correlates were supported at the continental scale contributes to the strength of inference of consistent responses across multiple spatial scales.”

References

Correction: Testing the socioeconomic and environmental determinants of better child-health outcomes in Africa: a cross-sectional study among nations


This article was previously published with an error.

Zia Mehrabi is not an author and has been removed from the author list.

Open access. This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.