Reviewer's report

Title: Associations Between Open Drain Flooding and Pediatric Enteric Infections in the MAL-ED Cohort in a Low-Income, Urban Neighborhood in Vellore, India

Version: 0 Date: 06 Jan 2019

Reviewer: Sari Kovats

Reviewer's report:

This is an interesting paper that uses a dataset in a neglected population (children in urban slums) to examine the importance of sanitation and drainage in the transmission of diarrhoeal disease.

In the background section, the summary of the literature on child exposure to faecal sludge is extensive but focussed on the microbiological evidence.

The paper would benefit from a better discussion of the epidemiological evidence. Is it really correct that there have been no previous studies that have linked exposure or proximity to open drains to enteric infections in children? It would be useful to discuss what is known about sanitation-related determinants of enteric infection. Presumably lack of toilet has been shown to have a significant effect? but what about type of toilet and related infrastructure? There has also been several studies of diarrhoeal disease episodes and rainfall and temperature factors- these should also be reviewed in the background section.

It would also be useful to have a clearer discussion on pathogen loads in flood water vs open drain water by season. The discussion indicates that flood water has a relatively lower pathogen load and therefore its not clear why this exposure would lead to higher infection rates?

This paper would benefit from some clearly stated hypotheses.
A description of the climate/rainfall patterns for this population/location is missing (e.g. figure showing monthly rainfall for the data collection period, with indications of the data collection dates). The description of the rainfall data in the methods is inadequate. Is this modelled data or station data? How representative are the data? Is the exposure variable monthly total rainfall used in the analysis? The description (text and table 1) is very confusing and needs to be re-written. If the stool sample was taken early in the month, then matching by month is not appropriate and lag should be considered. Surely, daily data could have been obtained and the total for the 30 days prior to stool collection would be a better exposure variable?

Temperature is also a determinant of enteric infections, especially bacterial ones, and so should be included as a confounder.

The study design and the rationale for the model is difficult to follow. It would help of the model was explicitly described. The main outcome of the analysis is at least one enteric pathogen in a childs stool at a given sampling event. Table 1 should include all the variables used in the dataset.

Overall, the lack of consideration of other exposure routes and key risk factors makes the results very difficult to interpret. The use and description of the weather (not climate) factors is also not appropriate and impossible to interpret.

There is clearly some confounding between poverty and toilet ownership, exposure to drain water and exposure to flood water- and these associations should be described in the data (page 8).

The household flood exposure data (based on self-reported assessments) is not useful unless linked in time to the illness event. In some cases, the flood may have occurred after the illness? On page 15 the authors state that the data were collected after the stool collection -and so there may be issue in the recall? The validity of this variable needs to be discussed. The timing of the data collection needs be clearly stated in the methods section. Further, how is "contact" defined. Does flooding (drains and house) occur every year with the monsoon? There is an important difference between annual seasonal monsoon flooding and an extreme events (1 in 100 year rainfall). There needs to be some definition of "heavy" or extreme rainfall otherwise the results are not easy to interpret.
The value of the mapping is unclear. The exposure variables generated by mapping and cluster analysis are not described or was the mapping for illustration only, and the exposure data are derived from the household surveys?

Minor comments.

The authors rely heavily on referring to other papers and methods. It would be more useful if this paper included the relevant information in the methods.

The term flooding is used ambiguously. The description of "flooding" variables on page 9 is confusing and needs to be improved. What is aspatial flooding?

Methods. Why control of household sanitation? Surely a key question would be how does having a toilet affect the role of contact with drain water. What is the level of contact with drain water for persons without toilet- surely it is also very high (reportedly >80% across the study population)?

The text is repetitive in places and needs editing to improve clarify- particularly when discussing the exposure variables.
Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

No

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

No

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

No

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I recommend additional statistical review

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

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?
If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

'I declare that I have no competing interests'

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal