Author’s response to reviews

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

Authors:

David Berendes (dberendes@cdc.gov)
Juan Leon (juan.leon@emory.edu)
Amy Kirby (agk1@cdc.gov)
Julie Clennon (jclenno@emory.edu)
Suraja Raj (suraja.jeya.raj@emory.edu)
Habib Yakubu (hyakubu@emory.edu)
Katharine Robb (karobb@emory.edu)
Arun Kartikeyan (dr.arunks@gmail.com)
Priya Hemavathy (priya.hemavathy82@gmail.com)
Annai Gunasekaran (annaimbavellore@gmail.com)
Sheela Roy (sheelaroy1@gmail.com)
Ben Ghale (benowr@live.com)
J Senthil Kumar (senthilcmc2008@gmail.com)
Venkata Mohan (venkat@cmcvellore.ac.in)
Gagandeep Kang (gkang@cmcvellore.ac.in)
Christine Moe (clmoe@emory.edu)

Version: 1 Date: 26 Feb 2019

Author’s response to reviews:

We note that all line numbers correspond to lines in the ‘track changes’ version of the manuscript. Please see attachment that also includes responses (below).
Reviewer reports:

Sari Kovats, Ph.D. (Reviewer 1): 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.

- We greatly appreciate all of the reviewer’s comments and time spent in reviewing this paper.

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?

- We appreciate the reviewer’s comment. While studies (e.g. Barreto et al. 2007 and others) have indirectly evaluated sewerage interventions as compared to open drainage on self-reported diarrhea, the ability to collect more sensitive data on enteric infections generally—which encompass both symptomatic and asymptomatic infections—is a recent development as a result of enhanced lab capacity, test sensitivity, and the growing recognition of the importance of subclinical infections in longer term pediatric health outcomes. The MAL-ED study is the first of its kind to enlist longitudinal follow-up of young children for such outcomes using robust laboratory testing of both routine (asymptomatic) and diarrheal (symptomatic) stool. To-date, most estimates of the direct effects of open drains on health have been quantitative microbial risk assessments (QMRAs) modeling risk of infection after exposure to drain water.

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?

- We appreciate the reviewer’s comment. As discussed above, the uniqueness of the enteric infection data means that studies focusing on enteric infection as an outcome are relatively new. We have added a reference to our previous study from this population that did test ownership of toilets and observed significant associations with reduced bacterial and protozoal infections, but not viral infections or diarrhea (reference [15], line 82-83). Studies on the type of toilet and related infrastructure have been limited due to the lack of availability of data on enteric infections, though several large-scale trials (e.g. the Maputo Sanitation trial) may have such data in coming years.
There has also been several studies of diarrhoeal disease episodes and rainfall and temperature factors- these should also be reviewed in the background section.

-We appreciate the reviewer’s comment and agree that the literature on rainfall and temperature interactions with diarrhea is rich. We have clarified the focus of the analysis, per subsequent comments of the reviewer, and removed references to ‘climate’ to keep the text focused on rainfall. Further, as the purpose of the analysis was to examine rainfall as limited to its effects in causing overflows of urban drains, we have added discussion of rainfall as a risk factor for enteric infections, and its complex relationship with urban infrastructure, to the introduction (lines 123-134).

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?

-We appreciate the reviewer’s comment. To clarify, we hypothesize that exposure to drain flooding (increased rainfall in areas with drains that were reported to flood) may have been sufficient—by itself—to cause infection. We have modified the discussion to represent this (lines 370-393).

This paper would benefit from some clearly stated hypotheses.

-We appreciate this comment and have added hypothesis statements to the background (lines 121-124).
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?

We appreciate the reviewer’s comments. Unfortunately, as this was an additional subanalysis (i.e. not pre-designed at the start of the projects) using data from the MAL-ED and SaniPath projects, rainfall data was not collected in real-time and therefore we are limited to publicly-available retrospective data, which is aggregated monthly by the Indian government for the Vellore district. The rainfall is from doppler weather radar at Chennai, a nearby city. As our main goal was to estimate whether drain flooding events were associated with incidence of enteric infections in children, we used the monthly data on rainfall (which was at district-level) and geographical clusters of reported flooding locations to our best approximation of flooding. Thus, the interaction of rainfall incidence and spatial clusters of reported flooding—testing in the models in this paper—represents the best estimate of heavy flooding conditions. We have included text on the limitations of this approach in the discussion (lines 426-434).

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

We appreciate the reviewer’s comment and agree that temperature is an important driver of enteric infections when examined in more general models aiming to estimate more ‘upstream’ effects of climate on incidence of enteric infection. However, given the environmental exposure pathway being examined—rainfall causing overflowing/flooding of open drains with feces in them, and associated exposures—it would appear that temperature would be an ‘upstream’ driver of rainfall (e.g. acting through rainfall) in this directed acyclic graph, and thus controlling for temperature would be duplicative of testing our main effects (e.g. detrimental to testing associations with rainfall/drain flooding) rather than accounting for residual confounding.
The study design and the rationale for the model is difficult to follow. It would help if the model was explicitly described. The main outcome of the analysis is at least one enteric pathogen in a child's stool at a given sampling event. Table 1 should include all the variables used in the dataset.

-We appreciate the reviewer’s comment. Table 1 represents all variables used in the analysis, with the exception of the outcome data on enteric infections, which is presented in-text in the results and has already been described in-detail in tables in another publication (15).

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.

-We appreciate the reviewer’s comments. As mentioned earlier, we have clarified the goal of the analysis overall at the end of the introduction. Further, because the goal of the study was to test whether open drain flooding was associated with enteric infections, we were not attempting to describe all potential exposure pathways—fingers, flies, food, etc.—and were not conducting a full/comprehensive risk factor assessment. We have explained this further in the background section (lines 121-124) and added the requested hypothesis statements, which we agree helps clarify the goals of the analysis.

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

-We appreciate the reviewer’s comment. We agree that poverty and toilet ownership could confound the exposure variables (drain flooding, as constructed from reported flooding and rainfall), thus all multivariable models are adjusted for these factors (see footnote in Table 3).
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?

As described above, the goal of the analysis was to estimate drain flooding using self-reported data on habitual drain flooding (to identify households or geographic areas that would be susceptible to drain flooding) and rainfall data (to approximate times when shallow open drains would overflow their boundaries and flood). Rainfall data was linked in time (month) to when stool specimens were collected and analyzed. Self-reported data on habitual drain flooding was collected at a single point in time, but the question was phrased to have the respondent indicate if the drains in the area ever flood. As discussed in the limitations, there are some static assumptions required for the self-reported data, but we believe these to be robust in that no large infrastructure projects that would have fundamentally altered the drain flow or flooding were undertaken during the period of the MAL-ED or SaniPath study data collection. A sensitivity analysis of changes in toilet infrastructure in reference 15 did not yield meaningful changes in sanitation infrastructure (lines 420-421).

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.

Please see our response earlier about the how the self-reported data were used. Discussion of the need for static assumptions with regard to the exposure data has also been added in the limitations.

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.

We thank the reviewer for this comment. The timing of data collection has been described in the methods (lines 139-145). Contact with drains was referred to as any physical contact between the respondent and drains or floodwater. As described earlier and in the limitations, incidence of drain flooding could not be ascertained directly, thus the combination of rainfall data and reported locations where drains flood was used to approximate this, as described in previous responses.
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?

-We have added further detail to the methods (lines 186-213 & 228-237) to further explain the mapping. Briefly, searching for spatial clusters of reported drain flooding allowed us to estimate the locations of low-lying areas where drain flooding may occur. These clusters were, in fact, locations towards the ‘ends’ of drain flow patterns, and thus made sense as areas where drains could flood in the neighborhood. When these data are combined with rainfall data (as in interaction terms in our statistical models), we can test whether there is a specific association between living in that area/cluster of drain flooding and enteric infection that varies (increases) with rainfall, which would be indicative of a flood-carried risk, and thus approximate drain flooding that we were unable to estimate directly during the study.

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.

-We appreciate the reviewer’s comment, and have described the methods and data from other papers in more detail.

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?

-We have addressed these comments in our response to requests for further definitions of 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)?

-We appreciate the reviewer’s comment. The specific question of the association between toilet ownership and enteric infection is the subject of another, previously published, analysis that we conducted (ref 15). We controlled for toilet ownership to attempt to mitigate exposures to fecal waste from the household’s discharge alone, in order to separate the potential contributions of one’s own toilet from the open drainage throughout the neighborhood.

The text is repetitive in places and needs editing to improve clarify- particularly when discussing the exposure variables.

-We have elaborated on the exposure variables and clarified their definitions in the methods.