

Vulnerability to climate-driven pathogen-specific diarrheal diseases: secondary analysis from the Global Enteric Multicenter Study and Vaccine Impact on Diarrhea in Africa follow-up

Corresponding Author: Dr Megan Kowalczyk

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

Thank you for the opportunity to review the manuscript, "Vulnerability to Climate-Driven Pathogen-Specific Moderate-to-Severe Diarrheal Disease: Results from GEMS and VIDA." This is a generally well-written study examining how vulnerability factors may modify associations between climate drivers and the risk of pathogen-specific moderate-to-severe diarrhea (MSD), focusing on rotavirus, Shigella, Cryptosporidium, and enterotoxigenic E. coli (ETEC). Please see my comments below:

Major comments

Introduction

- Consider adding a short paragraph (or a few sentences) noting that diarrheal etiology is heterogeneous and that climate-related mechanisms may differ by pathogen (e.g., temperature, rainfall, water contamination, person-to-person transmission dynamics). This will help readers interpret why climate associations might vary across rotavirus vs Shigella/ETEC vs Cryptosporidium.
- Please add 1–2 sentences defining vulnerability.
- It would help to briefly explain how "adaptive capacity factors" differ conceptually from the asset-based wealth index. This could be done in the Introduction and/or in the Discussion.

Methods

- Please describe how the composite sensitivity index was constructed (methods).
- Lines 71-75: It is unclear what was aggregated and at what level. For example:
 - o Are counts aggregated as PCR-confirmed pathogen-specific MSD cases per day, or all MSD regardless of pathogen, or something else?
 - o For adaptive capacity factors (e.g., car ownership), are these measured at the individual/household level and then aggregated to site, or used as individual-level predictors in a case-level model?
 - o Please clarify what "site level" means (country?).
- Line 92: Prior to what exactly, symptom onset, enrollment/sample collection, PCR detection?
- The two-stage approach can introduce bias/overconfidence if selection and estimation use the same data. Please clarify whether you used sample splitting, cross-validation, or any strategy, or acknowledge this limitation.
- Climate predictors (especially multiple lags and precipitation metrics) are often highly correlated, please clarify how you handled correlated predictors in the RF stage.
- Categorizing continuous exposures into tertiles can reduce power and impose arbitrary cutpoints. Did you consider adding a sensitivity analysis using alternative cutpoints (e.g., quartiles) or a continuous non-linear specification (e.g., splines)?
- The manuscript defines effect modification as differences in relative risks between strata. Typically, there is a formal test (e.g., interaction terms and an interaction p-value or CI for contrasts). Please clarify, whether interaction terms were included in a model, or whether stratified models were compared descriptively, and how uncertainty in the difference between strata was quantified.

- The narrative around vaccination as an effect modifier is interesting but unclear. Please explicitly describe how vaccination was operationalized whether comparisons are within the same surveillance platform or across GEMS vs VIDA, and whether children in GEMS and VIDA are the same cohort (my understanding is they are separate studies/samples).

Minor comments

Introduction

- Lines 31-32: Please add a citation for the statement “Rotavirus incidence increases when temperatures decrease.”
- Lines 30-36: The discussion of rotavirus vaccination feels abrupt. Consider moving it after a brief pathogen overview and climate-pathogen mechanism discussion, or add a transition sentence linking vaccination to vulnerability/adaptive capacity.
- Line 41: Please verify whether “2008-2018” should be “2007-2018.”

Methods

- Lines 78-80: If the vaccination aim is a key study objective, consider stating it consistently in the Introduction and framing it as a specific sub-aim, rather than introducing it mid-methods.
- Lines 102-104: If sensitivity is always assessed as an effect modifier regardless of the RF results, please describe sensitivity as a pre-specified modifier and RF as exploratory screening for other modifiers.
- Line 108: Please clarify what this adjustment is (study indicator? offset?) and how it enters the model.

Results

- Tables: Please define acronyms (e.g., GEMS, VIDA, etc.) in table footnotes.
- Figure 1: Adding year labels to the x-axis (or indicating the calendar years covered) would improve interpretability.
- Consider adding a supplementary table with interaction terms or contrasts (with CIs).

Discussion

- Lines 296-303: Please clarify the vaccination modification interpretation and whether GEMS vs VIDA involve the same participants.
- Lines 313-314: For the statement “coinfections are not expected to alter the relationship...,” please provide a citation.

Thank you.

Reviewer #2

(Remarks to the Author)

This work looked into the role of social vulnerability on climate-driven pathogen-specific diarrheal diseases. The topic is of importance. However, the methods of this work do not seem sophisticated enough.

1. General: Descriptions are vague and confusing.
2. General: Weather and climate are two different concepts. Based on the descriptions of data, it seems that the authors were looking into the weather variability, not the climate variability. The authors should be cautious about these terminologies.
3. Background – Line 5: The description of “increased variability in ...” is misleading. Readers may misinterpret the exposure of interest to be weather variability.
4. Background: The title of this work and the beginning of the introduction seem to emphasize the idea of “Pathogen-specific.” However, after two paragraphs on social vulnerability, such an idea was not highlighted sufficiently in the last paragraph of background when describing the research aims.
5. Background – role of rotavirus vaccination: The authors listed this factor among the vulnerability factors. This is okay, but should be more clearly stated when first introducing this concept (Line 30). Now the paragraph regarding rotavirus vaccination is a bit abrupt.
6. Methods – diarrhea data: Was there an age restriction on the participants of the two studies? The starting sentence of the background section made me wonder if the target population was under-five children.
7. Methods – diarrhea data – Line 54: What did the author mean by “8-9 cases of MSD”?
8. Methods – diarrhea data – control selection: what was the case-control ration?
9. Methods – vulnerability factors: The description of the primary caregiver’s education is neutral, not consistent with the definition of adaptive capacity. The authors may consider something like “Primary care giver has an education level of at least primary”.
10. Methods – weather – daily maximum temperature: The authors should elaborate on the selection of temperature metric. In other words, why did the authors not use daily mean or minimum temperature.
11. Methods – weather – time window: First, the selection of time window should be moved to the statistical analysis section, not here. Second, the lagged patterns are usually context-dependent. Regarding the lag issue, the cited reference only described that “Authors used a range of exposure definitions including average, maximum, and minimum temperatures over daily, weekly, biweekly, and monthly resolutions, and considered a variety of different temporal lags between exposure and outcome.” Not sure how did the authors conclude a time window of three weeks.
12. Methods – weather – drought and flood: The introduction of the impact of drought and flood is abrupt. Why did the authors look into these two specific events? Were these two considered as the main exposures or just confounders?
13. Methods – Statistical Analysis: Why did the authors use separate RF models to identify the key climate drivers and adaptive capacity factors? What were the advantages of such a practice compared to including all the climate and adaptive capacity factors in the same model?
14. Methods – Statistical Analysis: The authors should elaborate on the reason of using RF to select key adaptive capacity factors. It is reasonable to use RF to select climate factors, which were exposures of interest that may be linked to the

outcomes. However, adaptive capacity factors were considered as effect modifiers in this study. In this case, an influential factor related to the outcomes (i.e., the classification target of RF) is not necessarily an effect modifier.

15. Methods – Statistical Analysis – nonlinear associations: Why did the authors use tertiles instead of splines (e.g., natural cubic splines) to capture the non-linearity?

16. Methods: Sensitivity analyses were missing, e.g., dfs for long-term trends and seasonality, hyperparameter settings in RF.

17. Methods: Multiple pathogens and multiple vulnerability factors were investigated in this study. How did the authors consider the issue of multiple comparisons?

Reviewer #3

(Remarks to the Author)

This manuscript addresses an important question about effect modification of pathogen-weather relationships by household-level demographic vulnerabilities. It uses valuable multi-site data across three African countries and two separate studies which span a 10-year period. It employs a promising and appropriate analytical approach. However, substantial clarification is required regarding methodological transparency (weighting, index construction, model specification), definition of effect modification, and consistency between figures and text. Some interpretations extend beyond the evidence presented and should be tempered.

Major Comments

1. Cohort characteristics and pathogen distribution

Lines 48-54: Please include the age range (or distribution) of children enrolled in each cohort. Age is strongly associated with pathogen distribution in pediatric diarrhea studies, and differences in enrollment age across sites or studies/time periods could influence interpretation of pathogen-specific findings.

2. Weighting procedure

Lines 59-63: Please provide additional detail on how the weighting was implemented. How was it operationalized in the raw data or incorporated into the statistical models? Did the weighting account for pathogen-specific patterns? For example, if there were a localized outbreak of a specific pathogen (e.g. ETEC) during the beginning of the sampling window, resulting in the maximum number of cases being met, and thus resulting in a temporary dominance of that pathogen, how would this influence the weighting scheme? Is there any adjustment to account for short-term clustering that might affect the apparent absence or reduction of other pathogens during that period? Clarifying this would help the reader understand how the weighting approach handles temporal heterogeneity and pathogen-specific fluctuations.

3. Sensitivity index construction

Line 67-68: Table S1 does not provide any information about a composite sensitivity index by individual. Please provide sufficient detail for how the composite sensitivity index was created, including:

- The formula used
- Whether variables were standardized
- How scaling was performed
- Whether weights were applied

Without this information the index cannot be evaluated or reproduced.

4. Aggregation of adaptive capacity variables

Line 74-75: The description of aggregating adaptive capacity variables at the site level is unclear. Could you clarify 1) how site level proportions were calculated, 2) whether these were used as a scaling factor, and 3) how they were aggregated. It seems from the analysis that these 'adaptive capacity' variables were not aggregated and instead evaluated individually, which is part of the confusion. Providing a table with site-level values and how they were incorporated into the models would improve transparency.

5. Weather data matching

Line 86-87: How were the cases matched to the weather variables? Was it based on the clinic location or household location? If clinic-based, what is the size of the catchment area for the clinic and what is the typical travel distance to the clinic? Clarifying this is important for interpreting exposure misclassification.

6. Resolution of weather data

Line 88: It is unclear how 0.05-degree grid resolution was obtained from a 0.25-degree resolution dataset? If higher-resolution interpolation of ERA5 core dataset was used, please cite that and clarify the resolution available in that dataset. If not, please edit the text accordingly (including line 323).

7. Modeling approach

Lines 105-106: Please clearly state 1) the model type (e.g. GLM, GAM, etc.), 2) the outcome variables (counts, rates, risk ratios, etc), and 3) the link function. The repeated use of the word 'risk' suggests incidence rate ratios, but it should be explicitly stated here in the methods.

Lines 110-111 & Lines 151-153: For the tertile analysis, please clarify model specification. Were the 2nd and 3rd tertiles compared against a reference group (1st tertile)?

8. Selective interpretation of between-study differences

Lines 186-190: This paragraph focuses primarily on the rotavirus findings when comparing GEMS to VIDA. However, there are several additional instances where risk estimates differ between studies (e.g. increased risk of ETEC with high or moderate rainfall in The Gambia during VIDA but not GEMS) that are not addressed. Expanding this section to include a broader comparison of pathogen-specific differences across studies would provide a more balanced interpretation. Doing so may also help contextualize the extent to which observed differences are attributable to vaccination versus stochastic variation or other unmeasured factors. A more comprehensive discussion of between-study heterogeneity would strengthen the interpretation of these findings.

9. Definition and interpretation of effect modification

Lines 191-196: The definition of effect modification is unclear. Is modification defined by:

- Statistical significance of an interaction term from your models?
- Visual interpretation of Figure S2?
- Directional differences?

For several instances, the description in the text does not appear to match Figure S2. For example, for the Kenya plot and lines 192-194: Fig S2 shows moderate and higher temperatures resulted in an increased risk for the low sensitivity group for ETEC, but posed no difference in risk in high-sensitivity individuals. For rotavirus, both high and moderate temps were null for both the high and low sensitivity groups. This is contradictory to what you have written which is "In Kenya, increasing temperature was associated with higher risk of ETEC but had a null impact on Rotavirus and Cryptosporidium among cases with lower sensitivity."

Additionally, if groups differ only in statistical significance but show the same direction of effect, this may reflect differences in sample size rather than true modification.

10. Interpretation of adaptive capacity variables

Lines 201-213: I found this section challenging to interpret mechanistically. For variables such as improved sanitation, electricity, and refrigerator ownership (which may correlate with electricity access), I can envision plausible causal pathways through food storage, water quality, or pathogen exposure. However, I'm less clear on the hypothesized mechanisms by which transportation or television ownership would modify the relationship between weather and pathogen-specific MSD. As you suggest in the discussion, it seems possible that many of these 'adaptive capacity' indicators might be capturing broader socioeconomic status (SES). Have you considered constructing a composite SES index from these variables and evaluating effect modification on that summary measure? A composite approach might simplify interpretation and help distinguish between specific mechanistic effects and broader SES gradients. In addition, because the analysis is conditioned on clinic attendance, I was wondering whether you considered the possibility of differential care-seeking by transportation access. Specifically, could limited transportation access reduce the likelihood that less severe diarrhea cases present to the clinic, thereby increasing the MSD cases among children from households without transportation.

Do you have data on care-seeking behavior within these cohorts that would allow you to assess whether transportation predicts clinic attendance independent of severity, or whether mild diarrhea cases are captured similarly across transportation groups? Understanding this might help clarify whether the observed interaction reflects true effect modification or potential differences/selection bias in case ascertainment.

11. Resilience interpretation

Lines 338-341: This statement appears somewhat inconsistent with earlier findings showing that specific vulnerability factors modify MSD risk for particular pathogens and study sites. These results suggest that improving even a single adaptive capacity feature could potentially enhance resilience in certain contexts. It may be helpful to reconcile this broader statement with the site- and pathogen-specific effect modification results presented earlier in the manuscript.

12. Over interpretation

Line 252: The statement attributing variability to 'differences in immunity' is not supported by the data presented here. I suggest removing the mention of immunity here as the remainder of the paragraph is a coherent discussion the site-specific differences in weather patterns. At the very least, rephrase it to indicate that 'differences in immunity' is speculative.

Lines 336-348: This section extrapolates beyond the presented results and reads more as a policy-advocacy than data-driven interpretation. I recommend tightening this discussion and anchoring conclusions more directly in the study's empirical findings.

Minor / Clarification Comments

Lines 90-93: The described data is not included in Table S3.

Lines 108-111: I recommend you move this sentence up to be the last sentence of the previous paragraph. The transition from the RF models to this paragraph was difficult to follow. Including this explanation of how the RF model results were used (e.g. feature selection for the subsequent model) provides clarity for the reader.

Line 115: Please cite R and all packages used.

Line 191: I suggest rewording to avoid implying all climate drivers were modified.

Lines 263-265: This statement appears to be inconsistent with the plot from Figure S2. Please revisit this analysis and interpretation.

Lines 315-316: Table S5 reports model accuracy, not MDAs.

Please be careful throughout the manuscript with the use of the term 'climate' which typically in the scientific framework

implies long-term data collected over a 30+ year timeframe. Anything shorter is considered studies of 'weather.' Specific examples where 'climate' should be changed to 'weather' include lines 321, 331, 336, but there are many others.

References 8 & 9 are duplicates.

Please ensure all figures include detailed captions sufficient for independent interpretation. Figure captions should describe the variables shown, units, modeling context, and any abbreviations used.

Figure 1. Panel titles are confusing. Is panel A truly a plot of the daily total across all grid cells? It would make more sense to report the daily average of total precipitation across all grid cells so that the reader has a sense of the amount of precipitation at any given area within the study site (not the sum of all precipitation recorded at each grid cell in the study area). Same for Panel B – is this showing the average maximum temp across all grid cells or is it just showing the max of any single grid cell in the study area?

Figure 2: Add numeric y-axis labels. The current labels are repetitive with the color legend and don't provide any additional information.

Figure S1. Remove grey background to improve readability.

Table 1: Clarify whether precipitation averages (daily total, cumulative 1-week, cumulative 2-week) include zeros from the dry season. Consider reporting seasonal stratification.

Table 3: Clarify meaning of NA values. Are these cases where the model failed to converge, had sparsity issues, etc?

Table S1: Please define 'young' (e.g. under 5 years old).

In accordance with the journal's AI policy, I confirm that the manuscript was not uploaded into any generative AI tool. I used a generative AI tool solely to assist with refining the tone and clarity of my written comments. All scientific evaluation, interpretation, and conclusions expressed in this report are entirely my own.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

Thank you to the authors for carefully considering my comments and suggestions. The authors have addressed all of my comments.

Reviewer #2

(Remarks to the Author)

Thanks for addressing the comments. Some points need further clarification.

1. Methods – weather – daily maximum temperature: The citation seems not support the authors' statement that "maximum temperature was a stronger predictor of diarrheal disease". The original paper has two relevant descriptions about maximum temperature. First, for all-cause diarrhoea, "Pooled estimates based on maximum and minimum temperature were more variable than estimates based on mean ambient temperature". Second, for bacterial diarrhoea, "Pooled estimates were stable across mean, maximum and minimum temperature ...". Although the association was a bit higher for max temperature in Figure 4, but there was no significant difference from mean or min temperature and this is for bacterial diarrhoea only. The authors should justify the use of maximum temperature further or add a sensitivity analysis using mean or min temperature.
2. Methods – weather – time window: The use of 21 days as the max lag is still not well justified. Additionally, why is the starting lag 4 not 3 for Shigella. Lag 3-21 days are 19 days, not 18. More importantly, the interpretation of single lag is difficult because temperatures are highly auto correlated. For example, in Figure 3, for Kenya-Cryptosporidium, it is hard to explain why temperature at lag 10-11 and lag 13 were important, while lag 12 is not selected.
3. Methods – weather metrics: Not clear why temperatures were included on a daily basis while precipitations were included on a weekly basis. Would a larger number of temperature variables make it more likely to be selected in the random forest model?
4. Methods – weather: The authors should add a citation for this statement "Based on the well-established associations between temperature, precipitation, cumulative rainfall, and dry conditions with diarrheal disease". Furthermore, the authors should more clearly state the research aim in the introduction and clarify the exposure of interest.

Minor

1. Line 86: The comma between "three" and "age strata" should be removed?

2. Supplementary material – Figure S1: The title is broken by the figure.
3. Figures and tables: “intermittency” should be more clearly stated as “precipitation intermittency” or “rainfall intermittency”

Reviewer #3

(Remarks to the Author)

Thank you for the constructive edits you have incorporated into the updated manuscript, all of my concerns have been addressed. I have no further recommendations, but wanted to note that the updated Fig. 1 y-axis seems to be missing the 'a' in Total.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

REVIEWER COMMENTS

We greatly appreciate the detailed comments, recommendations, and thoughts from all three reviewers. We have addressed each comment below and feel that their reviews have led to a substantial improvement of the manuscript regarding framing of pathogen specific impacts, vulnerability factors, clarity of methods, formal tests of interaction and interpretation of our results. Please find below point by point responses to reviewer comments in bold/italic.

Reviewer #1 (Remarks to the Author):

Thank you for the opportunity to review the manuscript, “Vulnerability to Climate-Driven Pathogen-Specific Moderate-to-Severe Diarrheal Disease: Results from GEMS and VIDA.” This is a generally well-written study examining how vulnerability factors may modify associations between climate drivers and the risk of pathogen-specific moderate-to-severe diarrhea (MSD), focusing on rotavirus, Shigella, Cryptosporidium, and enterotoxigenic E. coli (ETEC). Please see my comments below:

Major comments

Introduction

- Consider adding a short paragraph (or a few sentences) noting that diarrheal etiology is heterogeneous and that climate-related mechanisms may differ by pathogen (e.g., temperature, rainfall, water contamination, person-to-person transmission dynamics). This will help readers interpret why climate associations might vary across rotavirus vs Shigella/ETEC vs Cryptosporidium.

Thank you for this suggestion, we agree that this addition would help readers interpret the varying associations. We have added the following text to the background: “Recent work in SSA found that rotavirus, Cryptosporidium, Shigella and enterotoxigenic E. coli producing heat stable toxin (ETEC) are the most common causes of moderate-to-severe diarrheal disease in children under 5^{4,5}, and each have distinct transmission pathways and seasonal patterns. Rotavirus is a viral pathogen, associated with increased risk during dry and cool periods² and a U-shaped relationship with rainfall.⁶ Cryptosporidium is a zoonotic parasitic protozoan that is associated with lower temperatures and increased rainfall, due to the ideal conditions of oocyst survival.^{2,7–11} Shigella and ETEC are bacterial pathogens and are positively associated with increases in both temperature and precipitation, although ETEC risk decreases above a temperature threshold.^{12–15}”

- Please add 1–2 sentences defining vulnerability.

We agree that a definition of vulnerability would increase the readability, we have added the following text to the background: “An individual’s vulnerability to climate change is defined as the predisposition to be adversely affected and is comprised of two components, sensitivity and adaptive capacity.¹⁷ Sensitivity refers to factors that make an individual more susceptible to the adverse health outcomes whereas adaptive capacity refers to factors that decrease the risk of and ability to cope with the adverse effects of climate change.¹⁷”

- It would help to briefly explain how “adaptive capacity factors” differ conceptually from the asset-based wealth index. This could be done in the Introduction and/or in the Discussion.

Thank you for this suggestion. Asset-based wealth and socioeconomic indices were explored but showed poor predictive ability in this setting, as a result we focused on individual predictors of adaptive capacity. This has been clarified in the methods: “Previous work in this population has found low predictive accuracy of composite wealth indices, so for the purpose of this study we focused on identifying individual adaptive capacity factors associated with risk.”

Methods

- Please describe how the composite sensitivity index was constructed (methods).

We agree that this description was missing, this has been added to the methods: “We developed a composite sensitivity index at the individual level, using a linear combination of raw sensitivity factors: number of individuals living and sleeping in the household, children under 5 in the household, and household ownership of animals, then normalized using min-max median normalization (Table S1).”

- Lines 71-75: It is unclear what was aggregated and at what level. For example:
 - Are counts aggregated as PCR-confirmed pathogen-specific MSD cases per day, or all MSD regardless of pathogen, or something else?

Thank you for pointing out the lack of clarity, this section has been clarified: “This analysis focuses on daily number of MSD cases qPCR positive for the most common pathogens, rotavirus, *Shigella* spp., *Cryptosporidium* spp., or enterotoxigenic *E. coli* (ETEC), separately, based on stool samples taken at the day of enrollment, aggregated to the country level.^{25,26} To estimate daily incidence rates for each study country site for input into statistical models, a number of weighting steps were performed: The study

protocol limited MSD case enrollment to no more than 9 MSD cases per fortnight, however contemporaneous data were collected on number of eligible MSD cases that presented to study clinics. In order to account for fluctuations in the number of eligible cases, case counts were weighted by the ratio of eligible MSD cases that presented at study clinics to enrolled cases for a given two-week period and rounded to discrete counts. Additionally, qPCR was performed for all enrolled MSD cases and controls in VIDA but was limited to a random subset of MSD cases and their first control in GEMS; moreover, the proportion of GEMS participants with qPCR varied over time. To account for this, we weighted daily cases by the inverse proportion of enrolled participants (cases and controls) that had qPCR performed by week.”

o For adaptive capacity factors (e.g., car ownership), are these measured at the individual/household level and then aggregated to site, or used as individual-level predictors in a case-level model?

We agree that this was not clearly stated, this has been clarified in the methods:

“We grouped individuals by their adaptive capacity factors and the sensitivity index to obtain daily case counts for each pathogen at each adaptive capacity and sensitivity level by country. For sensitivity, we dichotomized individual sensitivity values into low and high based on the median sensitivity level for each country and assigned cases as either having high or low sensitivity (Table S2). We then grouped cases of MSD by their sensitivity level and aggregated to cases per day by sensitivity at the country level. Similarly, individual cases of MSD were grouped by their adaptive capacity factors to get the number of cases per day with and without the adaptive capacity factor at the country level.”

o Please clarify what “site level” means (country?).

We agree that site and country were used interchangeably, we have clarified this throughout the methods.

• Line 92: Prior to what exactly, symptom onset, enrollment/sample collection, PCR detection?

We have clarified that this is enrollment in the text: “we included cumulative precipitation in the 1 to 2 weeks prior to enrollment as well as precipitation intermittency (the number of days since the last daily rainfall total of 1 mm).”

- The two-stage approach can introduce bias/overconfidence if selection and estimation use the same data. Please clarify whether you used sample splitting, cross-validation, or any strategy, or acknowledge this limitation.

Thank you for this comment. We use the first stage of modeling, random forest, to identify key climate drivers and adaptive capacity factors. We split the data into training and testing data, as defined in supplemental table S4. We then utilized the topmost influential climate drivers as the exposure in individual models and assessed effect modification by the identified adaptive capacity.

- Climate predictors (especially multiple lags and precipitation metrics) are often highly correlated, please clarify how you handled correlated predictors in the RF stage.

We agree that climate drivers are correlated and this was the reason that we utilized random forest as the first stage of our modeling. Random forest is ideal for identifying important predictors from a set of highly correlated predictors, so this was not a concern at this stage. The topmost influential climate drivers from RF were examined in separate regression models if they were correlated, specifically we allowed temperature and precipitation to be in the same model, but precipitation and intermittency and other precipitation variables were kept in separate models.

- Categorizing continuous exposures into tertiles can reduce power and impose arbitrary cutpoints. Did you consider adding a sensitivity analysis using alternative cutpoints (e.g., quartiles) or a continuous non-linear specification (e.g., splines)?

We agree with the reviewer's point, and we initially attempted to include splines, but this unfortunately led to model fit issues. Utilizing tertiles allowed us to split the data into comparable sized groups to improve model fit and convergence while allowing us to compare across levels of climate drivers, this has been added to the text: "We included the topmost influential climate drivers based on mean decrease in accuracy (MDA) from the climate driver RF and categorized variables into tertiles by site to account for nonlinear associations, 2nd and 3rd tertiles were compared against the first tertile as a reference group. This was done to ensure that there were sufficiently sized groups of observations for comparisons and to ensure model convergence."

- The manuscript defines effect modification as differences in relative risks between strata. Typically, there is a formal test (e.g., interaction terms and an interaction p-value

or CI for contrasts). Please clarify, whether interaction terms were included in a model, or whether stratified models were compared descriptively, and how uncertainty in the difference between strata was quantified.

Thank you for this comment, we agree a formal test could improve the manuscript. Interactions were tested, but due to the use of zero-inflated models type III p-values are not reliable. Additionally, presenting contrasts is quite complex given the tertiles for each of the maximum of three climate drivers, by two adaptive capacity levels, four pathogens, and three countries. To address this we have conducted likelihood ratio tests comparing base models to models with an interaction term for each climate driver separately, supplemental table S7. We have added this to the methods text: “We assessed effect modification by vulnerability factors, defined as differential incidence rate ratios between strata and a significant likelihood ratio test comparing the interaction model to the base model, 95% confidence intervals were used to characterize uncertainty.”

- The narrative around vaccination as an effect modifier is interesting but unclear. Please explicitly describe how vaccination was operationalized whether comparisons are within the same surveillance platform or across GEMS vs VIDA, and whether children in GEMS and VIDA are the same cohort (my understanding is they are separate studies/samples)

Thank you for pointing out that this was not clear, we have clarified this more in the methods: “We aimed to examine whether introduction of rotavirus vaccination into the routine infant immunization schedule between the two studies was a modifying adaptive capacity factor in the association between climate drivers and rotavirus infection by comparing the estimates of association between the two studies.”

Participants in GEMS and VIDA were not the same. While the two studies were performed in the same geographic catchment area and each study targeted children 0-59 months of age, the chance of a child being enrolled in both studies is minimal because VIDA began 4 years after GEMS and cases were enrolled from children seeking care at a study health facility. Vaccination status was verified by examining the child’s vaccination card or the vaccination center register.

Minor comments

Introduction

- Lines 31-32: Please add a citation for the statement “Rotavirus incidence increases when temperatures decrease.”

This has been added: “Rotavirus incidence increases when temperatures decrease², and therefore it is possible that lower rotavirus disease incidence resulting from

vaccination may lead to increasing burden of diarrhea due to other causes driven by climatic variability.¹⁵” **We have also added additional text in the background regarding the associations between rotavirus and climate drivers:** “Rotavirus is a viral pathogen, associated with increased risk during dry periods² and a U-shaped relationship with rainfall.⁶”

- Lines 30-36: The discussion of rotavirus vaccination feels abrupt. Consider moving it after a brief pathogen overview and climate-pathogen mechanism discussion, or add a transition sentence linking vaccination to vulnerability/adaptive capacity.

We agree and have added a transition sentence: “One such vulnerability factor that has yet to be explored, is vaccines for enteric pathogens. Specifically, the introduction of rotavirus vaccination has implications for climate-sensitive diarrheal disease burden that has not been fully explored.”

- Line 41: Please verify whether “2008-2018” should be “2007-2018.”

Thank you for catching this, this has been corrected in the text: “Using over a decade of data (2007-2018) from three”

Methods

- Lines 78-80: If the vaccination aim is a key study objective, consider stating it consistently in the Introduction and framing it as a specific sub-aim, rather than introducing it mid-methods.

Thank you for this suggestion, we have added this to the aim sentence: “Using over a decade of data (2007-2018) from three countries across Sub-Saharan Africa (The Gambia, Kenya, and Mali), we aim to examine the role of vulnerability factors, including the introduction of the rotavirus vaccine, on the association between climate drivers and the risk of pathogen-specific moderate-to-severe diarrhea (MSD), focusing on rotavirus, Shigella, Cryptosporidium, and enterotoxigenic E. coli (ETEC).”

- Lines 102-104: If sensitivity is always assessed as an effect modifier regardless of the RF results, please describe sensitivity as a pre-specified modifier and RF as exploratory screening for other modifiers.

A great suggestion, this has been added to the text: “Sensitivity was a pre-specified modifier and was not included in the RF screening.”

- Line 108: Please clarify what this adjustment is (study indicator? offset?) and how it

enters the model.

Thank you, we are including a time variable defined as time since the first day of GEMS adjusted for the time gap between GEMS and VIDA. This has been clarified in the methods text: “To account for long-term time trends and seasonality, time since the first day of the GEMS study was included as a natural cubic spline with four degrees of freedom, based on model fit, removing the 4-year gap between the GEMS and VIDA studies.”

Results

- Tables: Please define acronyms (e.g., GEMS, VIDA, etc.) in table footnotes.

Thank you for the suggestion, definition of acronyms have been added to table footnotes.

- Figure 1: Adding year labels to the x-axis (or indicating the calendar years covered) would improve interpretability.

Thank you for this suggestion, we have updated the labels in the x-axis to be clearer.

- Consider adding a supplementary table with interaction terms or contrasts (with CIs).

Thank you for this comment, we agree a formal test could improve the manuscript. Interactions were tested, but due to the use of zero-inflated models type III p-values are not reliable. Additionally, presenting contrasts is quite complex given the tertiles for each of the maximum of three climate drivers, by two adaptive capacity levels, four pathogens, and three countries. To address this, we have conducted likelihood ratio tests comparing base models to models with an interaction term for each climate driver separately, supplemental table S7. We have added this to the methods text: “We assessed effect modification by vulnerability factors, defined as differential incidence rate ratios between strata and a significant likelihood ratio test comparing the interaction model to the base model, 95% confidence intervals were used to characterize uncertainty.”

Discussion

- Lines 296-303: Please clarify the vaccination modification interpretation and whether GEMS vs VIDA involve the same participants.

Thank you for this suggestion. GEMS and VIDA did not involve the same participants, please refer to the Methods section where the data is described. No child in GEMS received the rotavirus vaccine, however we did obtain rotavirus vaccination status in VIDA and know vaccine coverage among enrolled children in VIDA. Since this analysis is not conducted at an individual level and there was widespread uptake of the vaccine we did not incorporate vaccination status data. We have added that individual vaccine uptake is not a part of this study, text in the discussion: “Rotavirus risk moved toward the null due to temperature in Kenya and due to precipitation in both The Gambia and Mali as hypothesized, potentially due to widespread uptake of rotavirus vaccines, although not modeled at an individual level.”

• Lines 313-314: For the statement “coinfections are not expected to alter the relationship...,” please provide a citation.

Given that we ran separate models for each pathogen, and individual having a co-infection would likely not alter the association with climate drivers since we are modeling each pathogen separately. Additionally, research has found that co-infections such as rotavirus and ETEC occur more frequently than expected, suggesting there is no competition, this citation has been added.

Thank you.

Reviewer #2 (Remarks to the Author):

This work looked into the role of social vulnerability on climate-driven pathogen-specific diarrheal diseases. The topic is of importance. However, the methods of this work do not seem sophisticated enough.

1. General: Descriptions are vague and confusing.

Thank you for this comment. We have worked to improve clarity and specificity in our descriptions throughout the revised manuscript.

2. General: Weather and climate are two different concepts. Based on the descriptions of data, it seems that the authors were looking into the weather variability, not the climate variability. The authors should be cautious about these terminologies.

We completely agree, and we have revised the text to specify weather or climate driver when talking about the models.

3. Background – Line 5: The description of “increased variability in ...” is misleading. Readers may misinterpret the exposure of interest to be weather variability.

Thank you for this comment. Our paper does not just look at the evolving time-averages, but we do include analysis of weather variability such as intermittency. Additionally, there is a precedent from previous literature on this topic has characterize these relationships this way.

4. Background: The title of this work and the beginning of the introduction seem to emphasize the idea of “Pathogen-specific.” However, after two paragraphs on social vulnerability, such an idea was not highlighted sufficiently in the last paragraph of background when describing the research aims.

We agree that the addition of pathogen-specific text would help frame the research better. We have added text on heterogeneity of diarrheal etiology and their associations with climate drivers to the second paragraph of the background: “Recent work in SSA found that rotavirus, *Cryptosporidium*, *Shigella* and enterotoxigenic *E. coli* producing heat stable toxin (ETEC) are the most common causes of moderate-to-severe diarrheal disease in children under 5^{4,5}, and each have distinct transmission pathways and seasonal patterns.^{6–8} Rotavirus is a viral pathogen, associated with increased risk during dry periods² and a U-shaped relationship with rainfall.⁶ *Cryptosporidium* is a zoonotic parasitic protozoan that is associated with lower temperatures and increased rainfall, due to the ideal conditions of oocyst survival.^{2,7–11} *Shigella* and ETEC are bacterial pathogens and are positively associated with increases in both temperature and precipitation, although ETEC risk decreases above a temperature threshold.^{12–15}”

5. Background – role of rotavirus vaccination: The authors listed this factor among the vulnerability factors. This is okay, but should be more clearly stated when first introducing this concept (Line 30). Now the paragraph regarding rotavirus vaccination is a bit abrupt.

We agree and have added a transition sentence to improve clarity: “One such vulnerability factor that has yet to be explored, is vaccines for enteric pathogens. Specifically, the introduction of rotavirus vaccination has implications for climate-sensitive diarrheal disease burden that has not been fully explored.”

6. Methods – diarrhea data: Was there an age restriction on the participants of the two

studies? The starting sentence of the background section made me wonder if the target population was under-five children.

This is correct and has been added to the pathogen-specific diarrhea data methods paragraph: “Both studies measured the population-based incidence, etiology, and outcomes of moderate-to-severe diarrheal disease (MSD) among children under 5 seeking care at sentinel health facilities serving censused populations.”

7. Methods – diarrhea data – Line 54: What did the author mean by “8-9 cases of MSD”?

The methods of the GEMS and VIDA studies aimed to enroll 8-9 cases of moderate-to-severe diarrheal disease in each of 3 age strata (0-11, 12-23, and 24-59 months) every two-week period, the original studies are cited and this is clarified in the text: “Both studies aimed to enroll 8 to 9 children in each of three, age strata (0-11, 12 – 23 and 24 – 59 months) every two weeks throughout the study along with 1-3 randomly-selected controls matched to the index case by age, sex, time of presentation and residence. ^{4,5,24} MSD (cases) was defined as new and acute onset diarrhea accompanied by either dehydration, dysentery, or hospitalization. ^{4,5,24} The study protocol was limited to enrolling up to 9 MSD cases per fortnight, however data were collected on number of eligible MSD cases that presented to clinics.”

8. Methods – diarrhea data – control selection: what was the case-control ration[al]?

Thank you for this question. The original goal of the GEMS and VIDA case control studies was to measure the population-based attributable incidence, etiology and outcomes of moderate to severe diarrheal disease (lines 61-71). Therefore, the choice of case-control was to determine which pathogens were significantly associated with moderate to severe diarrheal disease and estimate the attributable fraction and incidence, i.e., the number of pathogen-specific episodes that could be prevented with an effective intervention. The original studies are cited. Additionally, we clarified in the methods text that 1-3 age-matched controls were recruited for every case: “Both studies aimed to enroll 8 to 9 children in each of three, age strata (0-11, 12 – 23 and 24 – 59 months) every two weeks throughout the study along with 1-3 randomly-selected controls matched to the index case by age, sex, time of presentation and residence. ^{4,5,24} MSD (cases) was defined as new and acute onset diarrhea accompanied by either dehydration, dysentery, or hospitalization. ^{4,5,24} ”

9. Methods – vulnerability factors: The description of the primary caregiver’s education

is neutral, not consistent with the definition of adaptive capacity. The authors may consider something like “Primary care giver has an education level of at least primary”.

We agree and have specified the education level in the text: “We assessed the following adaptive capacity factors: improved drinking water ⁵, improved sanitation facility⁵, maternal education level of at least primary, electricity, and household ownership of a television, telephone, refrigerator, agricultural land, radio, boat, car, cart, scooter, or bike.”

10. Methods – weather – daily maximum temperature: The authors should elaborate on the selection of temperature metric. In other words, why did the authors not use daily mean or minimum temperature.

We appreciate this comment and have added clarification to the methods: “Daily maximum temperature was chosen based on preliminary analysis demonstrating maximum temperature was a stronger predictor of diarrheal disease compared to minimum and mean temperature. ²⁹”

11. Methods – weather – time window: First, the selection of time window should be moved to the statistical analysis section, not here. Second, the lagged patterns are usually context-dependent. Regarding the lag issue, the cited reference only described that “Authors used a range of exposure definitions including average, maximum, and minimum temperatures over daily, weekly, biweekly, and monthly resolutions, and considered a variety of different temporal lags between exposure and outcome.” Not sure how did the authors conclude a time window of three weeks.

Thank you for this suggestion. Since we included all lags in the modeling approaches, we find it appropriate to keep this in the data section. We agree that lag patterns are context and pathogen dependent, this is why we considered lags from 3 days up to three weeks to account for the pathway laid out in the cited metanalysis from climate driver to pathogen incidence. We have also added that we did this to account for various incubation periods among the pathogens in the text: “Given the well-established lagged association between weather and diarrheal disease, we considered 18 different lags for daily temperature and precipitation ranging from three days to 3 weeks, due to the length of various incubation periods and exposure-infection pathways (Table S3).²”

12. Methods – weather – drought and flood: The introduction of the impact of drought and flood is abrupt. Why did the authors look into these two specific events? Were these two considered as the main exposures or just confounders?

We agree that this came across as abrupt and have adjusted the introduction of the climate drivers in the methods text: “Based on the well-established associations between temperature, precipitation, cumulative rainfall, and dry conditions with diarrheal disease, we focused on those climate drivers. Additionally, research has established a lagged association between weather and diarrheal disease, therefore we considered 18 different lags for daily temperature and precipitation ranging from three days to 3 weeks, due to the length of various incubation periods and exposure-infection pathways (Table S3).² Further, to assess the impact of dry periods and rainy periods, we included cumulative precipitation in the 1 to 2 weeks prior to enrollment as well as precipitation intermittency (the number of days since the last daily rainfall total of 1 mm).²⁹”

13. Methods – Statistical Analysis: Why did the authors use separate RF models to identify the key climate drivers and adaptive capacity factors? What were the advantages of such a practice compared to including all the climate and adaptive capacity factors in the same model?

Thank you for this clarifying question. There are a variety of reasons for running two separate RF models. First, the RF's were done on two different scales – daily gridcell level (climate drivers) and individual level (adaptive capacity factors) as they are intended to identify two different things, this is included in the text: “First, two separate random forest (RF) classification models were run to identify 1) key climate drivers (exposure) at the gridcell level and 2) key adaptive capacity factors (potential effect modifiers) at the individual level for each pathogen and site. Sensitivity was a pre-specified modifier and was not included in the RF screening. Models were designed to classify 1) days with and without pathogen-specific MSD at the 0.05-degree grid and 2) individual level cases with and without pathogen-specific MSD.” **The climate drivers RF identifies key exposures at the community level whereas the adaptive capacity RF identifies potential effect modifiers at the individual case level as a result these need to be two separate models. Additionally, we did not seek to identify which weather variables were important after adjusting for adaptive capacity or vice versa, which would have been the result of a single RF model.**

14. Methods – Statistical Analysis: The authors should elaborate on the reason of using RF to select key adaptive capacity factors. It is reasonable to use RF to select climate factors, which were exposures of interest that may be linked to the outcomes. However, adaptive capacity factors were considered as effect modifiers in this study. In this case, an influential factor related to the outcomes (i.e., the classification target of RF) is not necessarily an effect modifier.

We agree, RF on adaptive capacity was not used to determine a factor was an effect modifier. Instead, it was used to screen for potential effect modifiers by identifying adaptive capacity factors that predict pathogen specific MSD. Given that effect modifiers need to be causally related to the outcome, the identified predictors were candidates for being effect modifiers of the relationship between climate drivers and pathogen specific MSD risk. Additionally, adaptive capacity factors that do not independently alter risk of pathogen specific MSD will not be effect modifiers. This logic is supported by the conceptual framework put forth by VanderWeele, Tyler J.; Robins, James M. Four Types of Effect Modification: A Classification Based on Directed Acyclic Graphs. *Epidemiology* 18(5):p 561-568, September 2007. | DOI: 10.1097/EDE.0b013e318127181b

15. Methods – Statistical Analysis – nonlinear associations: Why did the authors use tertiles instead of splines (e.g., natural cubic splines) to capture the non-linearity?

We assessed splines and tertiles, tertiles fit the model better and ensured that we had three groups with comparable numbers of observations. We have added this to our methods text: “We included the topmost influential climate drivers based on mean decrease in accuracy (MDA) from the climate driver RF and categorized variables into tertiles by site to account for nonlinear associations, 2nd and 3rd tertiles were compared against the first tertile as a reference group. This was done to ensure that there were sufficiently sized groups of observations for comparisons and to ensure model convergence.”

16. Methods: Sensitivity analyses were missing, e.g., dfs for long-term trends and seasonality, hyperparameter settings in RF.

We agree that we did not specify the sensitivity analyses that were done. We did look at various df's for time and found that df of 4 fit consistently across all models and had the best fit compared to df of 2,3, 5, or 6, this has been added to the text: “To account for long-term time trends and seasonality, time since the first day of the study was included as a natural cubic spline with four degrees of freedom, based on model fit, adjusted for the time between the GEMS and VIDA studies. This was determined after testing a spline with degrees freedom of 2,3, 4, 5, 6, and 14 to ensure model convergence and maximize model fit for all sites and pathogens.” **We also selected variables from the two RF models based on the mean decrease in accuracy across ~400 models with varying hyperparameter settings as specified in table S4 and in the text:** “Models were designed to classify 1) days with and without pathogen-specific MSD at the 0.05-degree grid and 2) cases with and without pathogen-specific MSD. Both were run ~ 400X varying hyperparameters and the data was split

into training and testing subsets (Table S4)..... We included the topmost influential climate drivers based on mean decrease in accuracy (MDA) from the climate driver RF and categorized variables into tertiles by site to account for nonlinear associations, 2nd and 3rd tertiles were compared against the first tertile as a reference group.”

17. Methods: Multiple pathogens and multiple vulnerability factors were investigated in this study. How did the authors consider the issue of multiple comparisons?

Individual models were run for one pathogen in one country, with one primary climate exposure and then stratified. Within model the issue of multiple comparisons does not need to be corrected for, across models we have revised the results and discussion to focus on trends seen across countries, climate drivers, and pathogens to address between model multiple comparisons. We have added this as a limitation to the discussion text: “Finally, our results may reflect spurious relationships due to multiple comparisons, given the number of climate drivers and vulnerability stratifications tested. However, we focused on interpreting trends across strata, country, pathogen, climate driver, and adaptive capacity factors to reduce this bias.”

Reviewer #3 (Remarks to the Author):

This manuscript addresses an important question about effect modification of pathogen-weather relationships by household-level demographic vulnerabilities. It uses valuable multi-site data across three African countries and two separate studies which span a 10-year period. It employs a promising and appropriate analytical approach. However, substantial clarification is required regarding methodological transparency (weighting, index construction, model specification), definition of effect modification, and consistency between figures and text. Some interpretations extend beyond the evidence presented and should be tempered.

Major Comments

1. Cohort characteristics and pathogen distribution

Lines 48-54: Please include the age range (or distribution) of children enrolled in each cohort. Age is strongly associated with pathogen distribution in pediatric diarrhea studies, and differences in enrollment age across sites or studies/time periods could influence interpretation of pathogen-specific findings.

Thank you for this comment, we agree and have added that the case control studies recruited equal number of cases across three age strata (0-11, 12-23, and 24- 59 months) to the methods text: “GEMS and VIDA measured the population-based incidence, etiology, and outcomes of moderate-to-severe diarrheal disease (MSD) among children under 5 seeking care at sentinel health facilities serving censused populations. Both studies aimed to enroll 8 to 9 children in each of three age strata (0-11, 12 – 23 and 24 – 59 months) every two weeks throughout the study along with 1-3 randomly-selected controls matched to the index case by age, sex, time of presentation and residence. ^{4,5,24} MSD (cases) was defined as new and acute onset diarrhea accompanied by either dehydration, dysentery, or hospitalization.^{4,5,24}” And we have added detail to the discussion specifying that seasonal trends for rotavirus in GEMS were comparable across age groups in a recent publication: “Additionally, we did not further group children by age as was done in GEMS and VIDA. Recent work on the seasonality of rotavirus in GEMS found comparable associations between season and pathogen among children younger and older than 24 months of age, but differences in the most dominant pathogen.¹⁴”

2. Weighting procedure

Lines 59-63: Please provide additional detail on how the weighting was implemented. How was it operationalized in the raw data or incorporated into the statistical models? Did the weighting account for pathogen-specific patterns? For example, if there were a localized outbreak of a specific pathogen (e.g. ETEC) during the beginning of the sampling window, resulting in the maximum number of cases being met, and thus resulting in a temporary dominance of that pathogen, how would this influence the weighting scheme? Is there any adjustment to account for short-term clustering that might affect the apparent absence or reduction of other pathogens during that period? Clarifying this would help the reader understand how the weighting approach handles temporal heterogeneity and pathogen-specific fluctuations.

Thank you for this feedback. The weighting that was done has to do with case enrollment limits, not all eligible cases were enrolled, as well as qPCR tested samples, not all samples were tested with qPCR. This weighting procedure was done to the raw data prior to inclusion into the statistical analysis. These weights indirectly account for pathogen-specific patterns such as outbreaks during the sampling window. We have clarified the reason for weighting the data in the methods: “To estimate daily incidence rates for each study location for input into statistical models, a number of weighting steps were performed: The study protocol was limited to enrolling up to 9 MSD cases per fortnight, however data were collected on number of eligible MSD cases that presented to clinics. In order to account for fluctuations in the number of eligible cases, case counts were weighted by the ratio of

eligible MSD cases that presented at study clinics to enrolled cases for a given two-week period and rounded to discrete counts. Additionally, not all samples from enrolled cases had qPCR performed, and the proportion of participants with qPCR varied over time. To account for this, we weighted daily cases by the inverse proportion of enrolled participants (cases and controls) that had qPCR performed by week.”

3. Sensitivity index construction

Line 67-68: Table S1 does not provide any information about a composite sensitivity index by individual. Please provide sufficient detail for how the composite sensitivity index was created, including:

- The formula used
- Whether variables were standardized
- How scaling was performed
- Whether weights were applied

Without this information the index cannot be evaluated or reproduced.

We agree that we did not include sufficient information on the creation of the sensitivity index, we have added this to the methods: “We developed a composite sensitivity index at the individual level, using a linear combination of raw sensitivity factors then normalized using min-max median normalization (Table S1) For sensitivity, we dichotomized individual sensitivity values into low and high based on the median sensitivity level for each country and assigned cases as either having high or low sensitivity (Table S2). We then grouped cases of MSD by their sensitivity level and aggregated to cases per day by sensitivity at the country level.”

4. Aggregation of adaptive capacity variables

Line 74-75: The description of aggregating adaptive capacity variables at the site level is unclear. Could you clarify 1) how site level proportions were calculated, 2) whether these were used as a scaling factor, and 3) how they were aggregated. It seems from the analysis that these ‘adaptive capacity’ variables were not aggregated and instead evaluated individually, which is part of the confusion. Providing a table with site-level values and how they were incorporated into the models would improve transparency.

We agree that this section was not clear. We created daily counts of cases for each pathogen by sensitivity level, presence and absence of each individual AC factor, this is clarified in the methods text: “We grouped individuals by their individual adaptive capacity factor levels and the sensitivity index to obtain daily case counts for each pathogen at each adaptive capacity and sensitivity level by country. For sensitivity, we dichotomized individual sensitivity values into low and high based on the median sensitivity level for each country and assigned cases as either having high or low

sensitivity (Table S2). We then grouped cases of MSD by their sensitivity level and aggregated to cases per day by sensitivity at the country level. Similarly, individual cases of MSD were grouped by their adaptive capacity factors to get the number of cases per day with and without the adaptive capacity factor at the country level.”

5. Weather data matching

Line 86-87: How were the cases matched to the weather variables? Was it based on the clinic location or household location? If clinic-based, what is the size of the catchment area for the clinic and what is the typical travel distance to the clinic? Clarifying this is important for interpreting exposure misclassification.

Thank you for this comment, we had latitude and longitude of participants households so we were able to directly match the gridcells to the participants homes, this has been clarified in the text: “We considered daily maximum temperature (C), and daily total precipitation (mm) matched to cases household latitude and longitude at the 0.05-degree grid resolution.”

6. Resolution of weather data

Line 88: It is unclear how 0.05-degree grid resolution was obtained from a 0.25-degree resolution dataset? If higher-resolution interpolation of ERA5 core dataset was used, please cite that and clarify the resolution available in that dataset. If not, please edit the text accordingly (including line 323).

We agree that this was unclear and have clarified that we regridded the ERA5 data using linear interpolation to match the resolution of CHIRPS, this has been clarified in the text: “We regridded the ERA5 data using linear interpolation to match the 0.05 degree resolution of CHIRPS.” We believe this is a valid approach since precipitation was sourced at a 0.05 degree resolution, and is more sensitive to spatial scale, and we do not believe there would have been meaningful differences in our results if we had used an alternative temperature data source.

7. Modeling approach

Lines 105-106: Please clearly state 1) the model type (e.g. GLM, GAM, etc.), 2) the outcome variables (counts, rates, risk ratios, etc), and 3) the link function. The repeated use of the word ‘risk’ suggests incidence rate ratios, but it should be explicitly stated here in the methods.

Thank you, we have specified this in the text: “We assessed effect modification by vulnerability factors using zero inflated Poisson models, defined as differential incidence rate ratios between strata, 95% confidence intervals were used to characterize uncertainty.”

Lines 110-111 & Lines 151-153: For the tertile analysis, please clarify model specification. Were the 2nd and 3rd tertiles compared against a reference group (1st tertile)?

Thank you for this suggestion, we have clarified this in the methods text: “We included the topmost influential climate drivers based on mean decrease in accuracy (MDA) from the climate driver RF and categorized variables into tertiles by site to account for nonlinear associations, 2nd and 3rd tertiles were compared against the first tertile as a reference group. This was done to ensure that there were sufficiently sized groups of observations for comparisons and to ensure model convergence.”

8. Selective interpretation of between-study differences

Lines 186-190: This paragraph focuses primarily on the rotavirus findings when comparing GEMS to VIDA. However, there are several additional instances where risk estimates differ between studies (e.g. increased risk of ETEC with high or moderate rainfall in The Gambia during VIDA but not GEMS) that are not addressed. Expanding this section to include a broader comparison of pathogen-specific differences across studies would provide a more balanced interpretation. Doing so may also help contextualize the extent to which observed differences are attributable to vaccination versus stochastic variation or other unmeasured factors. A more comprehensive discussion of between-study heterogeneity would strengthen the interpretation of these findings.

We agree that expanding this section beyond rotavirus would strengthen the interpretation and have done so in the text: “The risk of pathogen specific MSD due to climate drivers changed between GEMS and VIDA in all sites (Table S7). Risk of rotavirus due to temperature increased in VIDA in both The Gambia and Kenya. Similarly, risk of rotavirus due to cumulative rainfall decreased in both Mali and Kenya in VIDA (Figure 4). The risk of Cryptosporidium due to high cumulative rainfall increased in VIDA across all sites. In Kenya and The Gambia, risk of ETEC due to temperature decreased from GEMS to VIDA. Finally, there was a decreased risk of Shigella, Cryptosporidium and ETEC due to moderate intermittency in VIDA for The Gambia and Mali (Figure 4).”

9. Definition and interpretation of effect modification

Lines 191-196: The definition of effect modification is unclear. Is modification defined by:

- Statistical significance of an interaction term from your models?
- Visual interpretation of Figure S2?

- Directional differences?

For several instances, the description in the text does not appear to match Figure S2. For example, for the Kenya plot and lines 192-194: Fig S2 shows moderate and higher temperatures resulted in an increased risk for the low sensitivity group for ETEC but posed no difference in risk in high-sensitivity individuals. For rotavirus, both high and moderate temps were null for both the high and low sensitivity groups. This is contradictory to what you have written which is “In Kenya, increasing temperature was associated with higher risk of ETEC but had a null impact on Rotavirus and Cryptosporidium among cases with lower sensitivity.”

Additionally, if groups differ only in statistical significance but show the same direction of effect, this may reflect differences in sample size rather than true modification.

We defined effect modification as differential estimates by strata and a significant likelihood ratio comparing the interaction model to the base model. We also focused on results that follow a trend either across countries, pathogens, or a particular climate driver to account for multiple comparisons. This has been added to the methods, “We assessed effect modification by vulnerability factors using zero inflated Poisson models, defined as differential incidence rate ratios between strata and a significant likelihood ratio test comparing the interaction model to the base model, 95% confidence intervals were used to characterize uncertainty.” and the effect modification results have been edited accordingly.

10. Interpretation of adaptive capacity variables

Lines 201-213: I found this section challenging to interpret mechanistically. For variables such as improved sanitation, electricity, and refrigerator ownership (which may correlate with electricity access), I can envision plausible causal pathways through food storage, water quality, or pathogen exposure. However, I’m less clear on the hypothesized mechanisms by which transportation or television ownership would modify the relationship between weather and pathogen-specific MSD. As you suggest in the discussion, it seems possible that many of these ‘adaptive capacity’ indicators might be capturing broader socioeconomic status (SES). Have you considered constructing a composite SES index from these variables and evaluating effect modification on that summary measure? A composite approach might simplify interpretation and help distinguish between specific mechanistic effects and broader SES gradients.

Thank you for this comment. We believe that there is a clear mechanism between access to transportation and being climate resilient, those with access to transportation have better access to care. This is directly addressed as a part of our research question, and we found evidence that transportation ownership modifies the effect of climate drivers on pathogen specific MSD suggesting that

transportation ownership increases access to care resulting in increased climate resilience. While we agree that a composite-based SES index may be easier to interpret, we did explore composite based indices, but they showed poor predictive ability in this setting, as a result we focused on individual predictors of adaptive capacity as proxies. This has been clarified in the methods: “Previous work in this population has found low predictive accuracy of composite wealth indices, so for the purpose of this study we focused on identifying individual adaptive capacity factors associated with risk.”

In addition, because the analysis is conditioned on clinic attendance, I was wondering whether you considered the possibility of differential care-seeking by transportation access. Specifically, could limited transportation access reduce the likelihood that less severe diarrhea cases present to the clinic, thereby increasing the MSD cases among children from households without transportation. Do you have data on care-seeking behavior within these cohorts that would allow you to assess whether transportation predicts clinic attendance independent of severity, or whether mild diarrhea cases are captured similarly across transportation groups? Understanding this might help clarify whether the observed interaction reflects true effect modification or potential differences/selection bias in case ascertainment.

We agree that this could be a concern. As a result, we explored the coverage of transportation among the cases of MSD, as presented in the final row of Table 1. We found a high proportion of cases lived in households that had access to a form of transportation ranging from 49.5% of VIDA cases in Kenya, 80% of GEMS cases in Mali, and 94% of VIDA cases in The Gambia, at the lowest. Similarly, when looking at both cases and controls together, although not presented in this paper, we saw high coverage of transportation access across all sites and time periods. However, it is possible that we would be more likely to detect cases with transportation as they’d be more likely to seek care and we do not have the ability to differentiate this with our data. Ultimately, care-seeking by transportation access is not the focus of the questions posed in this work.

11. Resilience interpretation

Lines 338-341: This statement appears somewhat inconsistent with earlier findings showing that specific vulnerability factors modify MSD risk for particular pathogens and study sites. These results suggest that improving even a single adaptive capacity feature could potentially enhance resilience in certain contexts. It may be helpful to reconcile this broader statement with the site- and pathogen-specific effect modification

results presented earlier in the manuscript.

We agree that this could be misleading, this statement has been revised accordingly, “Second, while individual interventions may have profound effects on risk of pathogen-specific diarrhea, resilience to climate-linked MSD broadly emerges not from single one-size fits all intervention (e.g., improved drinking water) but from a network of infrastructure, public health interventions and socioeconomic factors that mediate multiple pathogen exposures.”

12. Over interpretation

Line 252: The statement attributing variability to ‘differences in immunity’ is not supported by the data presented here. I suggest removing the mention of immunity here as the remainder of the paragraph is a coherent discussion the site-specific differences in weather patterns. At the very least, rephrase it to indicate that ‘differences in immunity’ is speculative.

We agree and have removed this statement: “Associations between temperature and precipitation and risk of pathogen-specific MSD varied by site, likely because of site-specific variations in weather patterns.”³⁴

Lines 336-348: This section extrapolates beyond the presented results and reads more as a policy-advocacy than data-driven interpretation. I recommend tightening this discussion and anchoring conclusions more directly in the study’s empirical findings.

Thank you for this comment. We believe that this section contextualizes the policy and public health implications of the data driven results and is beneficial to this manuscript.

Minor / Clarification Comments

Lines 90-93: The described data is not included in Table S3.

Thank you for catching this, this reference to Table S3 has been removed.

Lines 108-111: I recommend you move this sentence up to be the last sentence of the previous paragraph. The transition from the RF models to this paragraph was difficult to follow. Including this explanation of how the RF model results were used (e.g. feature selection for the subsequent model) provides clarity for the reader.

We appreciate this suggestion and understand how this could be useful to readers. We have subsequently edited the methods text accordingly: *“The modifying effect of vulnerability factors on climate drivers was assessed through a two-stage modeling approach – identifying key climate drivers and potential effect modifiers and modeling the association and testing effect modification.....To model the association between climate drivers identified using random forests, and pathogen specific MSD we utilized daily level case counts, aggregated to the site level, to fit zero inflated Poisson models for each pathogen and site separately.”*

Line 115: Please cite R and all packages used.

We have added a citation for R, listed and cited all packages used in the text: *“All analysis was conducted using R packages randomForest and pscl.^{30–32”}*

Line 191: I suggest rewording to avoid implying all climate drivers were modified.

We agree and have reworded this sentence: *“Sensitivity modified the association between some climate drivers and infection risk for all pathogens and sites.”*

Lines 263-265: This statement appears to be inconsistent with the plot from Figure S2. Please revisit this analysis and interpretation.

Good catch – this has been corrected to be in line with S2: *“Sensitivity modified the association between some climate drivers and infection risk for all pathogens and sites. In Kenya, increasing temperature was associated with higher risk of ETEC, Shigella and Cryptosporidium but had a null impact on Rotavirus among cases with lower sensitivity. In The Gambia, individuals with high sensitivity had twice the risk of Rotavirus, in periods with high cumulative rainfall. The association between Shigella and high temperature in Mali was higher among cases with high sensitivity levels (Figure S2).”*

Lines 315-316: Table S5 reports model accuracy, not MDAs.

Correct, this reference has been removed: *“The MDA for climate drivers in RF models was small for almost all pathogens and sites.”*

Please be careful throughout the manuscript with the use of the term ‘climate’ which typically in the scientific framework implies long-term data collected over a 30+ year timeframe. Anything shorter is considered studies of ‘weather.’ Specific examples where

'climate' should be changed to 'weather' include lines 321, 331, 336, but there are many others.

Thank you for this comment, we have reviewed the text and have changed wording to either state weather or climate drivers – as both are accepted ways to state these short-term trends and observations.

References 8 & 9 are duplicates.

This has been corrected.

Please ensure all figures include detailed captions sufficient for independent interpretation. Figure captions should describe the variables shown, units, modeling context, and any abbreviations used.

Thank you for this comment, we have ensured that all figure captions are sufficient for independent interpretation.

Figure 1. Panel titles are confusing. Is panel A truly a plot of the daily total across all grid cells? It would make more sense to report the daily average of total precipitation across all grid cells so that the reader has a sense of the amount of precipitation at any given area within the study site (not the sum of all precipitation recorded at each grid cell in the study area). Same for Panel B – is this showing the average maximum temp across all grid cells or is it just showing the max of any single grid cell in the study area?

We agree and have subsequently updated figure 1.

Figure 2: Add numeric y-axis labels. The current labels are repetitive with the color legend and don't provide any additional information.

We agree and have recreated figure 2 to correct this.

Figure S1. Remove grey background to improve readability.

Thank you for this suggestion, we have explored other backgrounds to improve readability. We have decided that readability is the best with a gray background. Without it, the black gridlines are distracting and without gridlines it is hard to interpret.

Table 1: Clarify whether precipitation averages (daily total, cumulative 1-week, cumulative 2-week) include zeros from the dry season. Consider reporting seasonal

stratification.

Thank you for pointing this out, the precipitation averages presented in table 1 are measuring different things. For the daily total, we are averaging across the entire time period and site, as such we have included zeros and accounted for them appropriately. For the cumulative rainfall, these are measurements of the sum of the rainfall in 1 week and 2 week periods, and include zeros, we are then presenting the average over the entire study area and period. We did not focus on seasonality in this analysis, as this has been done in previous literature, we additionally have a measure of days since the last rainfall, capturing the dry seasons during this study.

Table 3: Clarify meaning of NA values. Are these cases where the model failed to converge, had sparsity issues, etc?

Thank you for this comment, this has been clarified in table 2 and 3.

Table S1: Please define 'young' (e.g. under 5 years old).

This has been defined.

In accordance with the journal's AI policy, I confirm that the manuscript was not uploaded into any generative AI tool. I used a generative AI tool solely to assist with refining the tone and clarity of my written comments. All scientific evaluation, interpretation, and conclusions expressed in this report are entirely my own.

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

Thank you to the authors for carefully considering my comments and suggestions. The authors have addressed all of my comments.

Reviewer #2 (Remarks to the Author):

Thanks for addressing the comments. Some points need further clarification.

1. Methods – weather – daily maximum temperature: The citation seems not support the authors' statement that "maximum temperature was a stronger predictor of diarrheal disease". The original paper has two relevant descriptions about maximum temperature. First, for all-cause diarrhoea, "Pooled estimates based on maximum and minimum temperature were more variable than estimates based on mean ambient temperature". Second, for bacterial diarrhoea, "Pooled estimates were stable across mean, maximum and minimum temperature ...". Although the association was a bit higher for max temperature in Figure 4, but there was no significant difference from mean or min temperature and this is for bacterial diarrhoea only. The authors should justify the use of maximum temperature further or add a sensitivity analysis using mean or min temperature.

Thank you for this thoughtful comment. We agree that the citation alone does not support this claim. However, we also found that maximum temperature was a stronger predictor than minimum and mean in preliminary analysis, which we have now clarified in the text. Additionally, the cited paper as you pointed out mentions that pooled estimates were stabled across minimum, maximum and mean temperature suggesting that results are less sensitive to the choice of temperature measurement. The cited paper also states that estimates were more variable using maximum and minimum temperature, and this variability, as demonstrated by our results in Kenya is key to model fit.

2. Methods – weather – time window: The use of 21 days as the max lag is still not well justified. Additionally, why is the starting lag 4 not 3 for Shigella. Lag 3-21 days are 19 days, not 18. More importantly, the interpretation of single lag is difficult because temperatures are highly auto correlated. For example, in Figure 3, for Kenya-Cryptosporidium, it is hard to explain why temperature at lag 10-11 and lag 13 were important, while lag 12 is not selected.

Thank you for this question. Regarding lag periods we selected the starting lag days based on the incubation period of the pathogen and for shigella the average

is closer to 4 days than 3 resulting in this choice, this is outlined in Kotloff et al. (reference 41) as cited. We have indicated in the text that lags were pathogen-specific based on biology, but we deemed it unnecessary to include details previously published extensively. We have corrected the lag day period:

“Additionally, research has established a lagged association between weather and diarrheal disease, therefore we considered 19 different lags for daily temperature and precipitation ranging from three days to 3 weeks, due to the length of various incubation periods and exposure-infection pathways (Table S3).^{2,6,12,41} We agree with the reviewer that temperatures are highly auto correlated, and a single lag may not actually be the most important lag day. This was the justification for running >400 random forest models and choosing the temperature lag with the highest average MDA value to assess the role of vulnerability factors in modifying the association between pathogen specific diarrheal disease and the identified climate driver. We acknowledge that the chosen lags may not be more influential than other lags in our discussion text: “The MDA for climate drivers in RF models was small for almost all pathogens and sites. Therefore, it is possible that the climate drivers at the specific lags chosen may not be more influential than other lag options.”

3. Methods – weather metrics: Not clear why temperatures were included on a daily basis while precipitations were included on a weekly basis. Would a larger number of temperature variables make it more likely to be selected in the random forest model?

Thank you for the clarifying question. Both temperature and precipitation are included on a daily basis from 3 to 21 days. To look at cumulative precipitation impacts we additionally aggregated precipitation to weekly basis to look at this. As a result, both temperature and precipitation are well represented in the random forest model.

4. Methods – weather: The authors should add a citation for this statement “Based on the well-established associations between temperature, precipitation, cumulative rainfall, and dry conditions with diarrheal disease”. Furthermore, the authors should more clearly state the research aim in the introduction and clarify the exposure of interest.

Thank you for this suggestion, we have added a citation to the suggested statement. In the introduction we state that the exposure of interest is climate drivers: “Using over a decade of data (2007-2018) from three countries across Sub-Saharan Africa (The Gambia, Kenya, and Mali), we aim to examine the role of

vulnerability factors, including the introduction of the rotavirus vaccine, on the association between climate drivers and the risk of pathogen-specific moderate-to-severe diarrhea (MSD), focusing on rotavirus, Shigella, Cryptosporidium, and enterotoxigenic E. coli producing heat stable toxin (ETEC).” Since the first stage of our modeling approach is to identify the key climate drivers for each pathogen and site, keeping the exposure of interest as climate drivers is important since we know that climate drivers differ by site and pathogen.

Minor

1. Line 86: The comma between “three” and “age strata” should be removed?

This has been removed.

2. Supplementary material – Figure S1: The title is broken by the figure.

This has been corrected.

3. Figures and tables: “intermittency” should be more clearly stated as “precipitation intermittency” or “rainfall intermittency”

Thank you for this suggestion, unfortunately increasing the length of labels within the figure (i.e. from Moderate Intermittency to Moderate Precipitation Intermittency) decreases the readability of all of our figures. To address this, intermittency is defined in each figure and table legend.

Reviewer #3 (Remarks to the Author):

Thank you for the constructive edits you have incorporated into the updated manuscript, all of my concerns have been addressed. I have no further recommendations but wanted to note that the updated Fig. 1 y-axis seems to be missing the 'a' in Total.

Thank you for all of the thoughtful comments and feedback you provided in your review and thank you for catching this mistake, this has been corrected.