Understanding the Impact of Rainfall on Diarrhea: Testing the Concentration-Dilution Hypothesis Using a Systematic Review and Meta-Analysis

Alicia N. M. Kraay,1 Olivia Man,2 Morgan C. Levy,3,4 Karen Levy,1 Edward Ionides,2 and Joseph N. S. Eisenberg2

1Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA
2Department of Epidemiology, University of Michigan–Ann Arbor, Ann Arbor, Michigan, USA
3Scripps Institution of Oceanography, University of California San Diego, La Jolla, California, USA
4School of Global Policy and Strategy, University of California San Diego, La Jolla, California, USA
5Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, Washington, USA

BACKGROUND: Projected increases in extreme weather may change relationships between rain-related climate exposures and diarrheal disease. Whether rainfall increases or decreases diarrhea rates is unclear based on prior literature. The concentration-dilution hypothesis suggests that these conflicting results are explained by the background level of rain: Rainfall following dry periods can flush pathogens into surface water, increasing diarrheal incidence, whereas rainfall following wet periods can dilute pathogen concentrations in surface water, thereby decreasing diarrhea incidence.

OBJECTIVES: In this analysis, we explored the extent to which the concentration-dilution hypothesis is supported by published literature.

METHODS: To this end, we conducted a systematic search for articles assessing the relationship between rain, extreme rain, flood, drought, and season (rainy vs. dry) and diarrheal illness.

RESULTS: A total of 111 articles met our inclusion criteria. Overall, the literature largely supports the concentration-dilution hypothesis. In particular, extreme rain was associated with increased diarrhea when it followed a dry period (incidence rate ratio (IRR) = 1.26; 95% confidence interval (CI): 1.05, 1.51), with a tendency toward an inverse association for extreme rain following wet periods, albeit nonsignificant, with one of four relevant studies showing a significant inverse association (IRR = 0.911; 95% CI: 0.771, 1.08). Incidences of bacterial and parasitic diarrhea were more common during rainy seasons, providing pathogen-specific support for a concentration mechanism, but rotavirus diarrhea showed the opposite association. Information on timing of cases within the rainy season (e.g., early vs. late) was lacking, limiting further analysis. We did not find a linear association between nonextreme rain exposures and diarrheal disease, but several studies found a nonlinear association with low and high rain both being associated with diarrhea.

DISCUSSION: Our meta-analysis suggests that the effect of rainfall depends on the antecedent conditions. Future studies should use standard, clearly defined exposure variables to strengthen understanding of the relationship between rainfall and diarrheal illness. https://doi.org/10.1289/EHP6181

Introduction

Climate change is expected to affect health conditions with known environmental determinants, including diarrheal disease (Ebi 2017). Diarrheal disease is already one of the leading causes of death in children under 5 years of age, and the World Health Organization (WHO) estimates that between 2030 and 2050, climate change will cause an additional 48,000 deaths per year from diarrhea alone (WHO 2018). Some of this expected increase in risk is related to projected increases in extreme weather events (including flooding and drought). Experts anticipate increased destabilization of the water cycle by the end of this century, with nonuniform changes in precipitation globally that result in regionally specific increases or decreases in total rainfall and the frequency of heavy rainfall (IPCC 2014). Areas with tropical climates are particularly likely to experience changes in the intensity and frequency of rainfall, and these regions already have some of the highest rates of diarrheal illness (Patz et al. 2005; Kolstad and Johansson 2011; GBD Diarrheal Diseases Collaborators 2017). However, studies have reported different directions in the association between rainfall and diarrheal disease; some studies have shown positive associations, and others have shown inverse associations. Additionally, both drought and flood have been associated with increased diarrhea (WHO 2003). Improved understanding of the effects of rainfall extremes and related weather events is necessary to prepare for and address public health in this uncertain future.

Carlton et al. (2014) and Levy et al. (2009) suggested that these conflicting findings may be partially explained by the background level of rain, proposing a “concentration-dilution hypothesis.” Specifically, a lack of rain may cause pathogens to accumulate in the environment (creating concentration conditions). Therefore, heavy rainfall can increase the risk of diarrheal disease by flushing pathogens into surface water, delivering them in one concentrated dose. However, during wet periods, rainfall may regularly flush environmental pathogens into water sources (such that pathogens do not appreciably accumulate in environmental sources), creating a dilution effect (Carlton et al. 2014). We hypothesize that extensions of this concentration-dilution hypothesis can be used to relate other climatic variables to diarrheal disease.

Specifically, we expect that the following:

1. Extreme rain is a risk for diarrhea following dry periods through the flushing of pathogens into the environment, and protective of diarrhea following a wet period via the dilution of pathogens.
2. Flooding leads to increased diarrhea via direct flushing of fecal material into the environment and by overwhelming sanitation and/or flooding infrastructure, but floods of longer duration may dilute pathogens and lower risk.
3. Drought concentrates pathogens in the environment, leading to higher risk of diarrhea.
4. Rainfall has a stronger positive effect on diarrhea in arid climates by concentrating pathogens (because pathogens may accumulate in the environment during drier times of the year). However, the effect of rain on diarrhea is inconsistent in the tropics due to variable rainfall patterns.
5. Overall, diarrhea is more common in rainier seasons. However, within rainy seasons, risk of diarrhea is concentrated either at the start of the rainy season or during the
opportunities to target the timing and location of public health

Two recent systematic reviews (Levy et al. 2016, Carlton et al.

Methods

We used the literature search strategy defined by Levy et al.

Eligibility

We included studies in the review if they met the following inclu-

Lags

Studies reported multiple associations per study for many reasons,

Data extraction and Variable Definitions

Data were extracted from each study on exposure category as
described in the text, outcome category as described in the text,
description of the association, point estimate with the scale or
metric type, mathematical model, and the authors’ hypothesized
mechanism.

We categorized associations as having extreme rain exposures
if authors reported rain percentiles (such as rainfall exceeding the
90th percentile), severe rainstorms (such as monsoons, typhoons)
without a description of a subsequent flood event, or anomalous
rain. For studies using percentile cut points for defining extreme
rain events, all studies defined these percentiles based on local
rainfall patterns. The reference period used to define normal rain-
fall patterns ranged from 1 to 30 y. Because the rainfall patterns
among the regions included in this review varied widely, the
absolute value of each threshold varied even when the same per-
centile was used. For example, for the 90th percentile, cut points
ranged from 52 mm in a week (averaging to 7.4 mm/d) up to
56 mm in a 24-h period. More details regarding the exposure def-
inition for each study can be found in Table S2.

Seasonal categories were considered when authors mentioned
cumulative rain over a seasonal period, when authors mentioned
variations of rainy season descriptors, or when studies had
monthly categories that could be combined into seasonal groups.

We categorized all seasonal data as falling in rainy seasons or
nonrainy seasons and did not separately analyze locations where
rainfall peaked two or more times during the year. For point esti-
mates that were not reported as a rate, we also extracted the
length of the rainy and dry seasons to allow us to calculate the
rate ratio, so that the pooled rate ratio would be unbiased by time.

We considered flood exposures when authors described flood-
related outbreaks of diarrhea, flooding frequency, or extremely
elevated river levels. Rainfall exposures encompassed studies
that measured cumulative rainfall over some period of time or
water level. Finally, author-described droughts were used to
define the drought exposure.

We included all associations published in articles and supple-
ments that met our inclusion criteria, regardless of whether that
association was identified as the final model. When the authors
reported heterogeneity of effects (for example, the effect of
extreme rain on diarrhea varied by season), we used the most
stratified estimate. Therefore, we report multiple associations per
study (range: 1–435 associations). For associations that could be
categorized in multiple climate exposure groups, we selected the
most extreme exposure. For example, a study that described an
extreme rainfall event that resulted in a flood would be catego-
ized as a flood because, due to duration and intensity, the result-
ing flood was likely to have a greater impact on health outcomes.

These studies were evaluated on a case-by-case basis. Complete
results from our extraction can be found in the Excel Table S1
(for all associations). Excel Tables S2–S5 show extraction results
for studies included in the meta-analysis for Extreme rain (Excel
Table S2), flood (Excel Table S3), season (Excel Table S4), and
rain (Excel Table S5), which includes some additional data for
each exposure that were not relevant to all exposure types (for
example, the cut point used to define extreme rain events was rel-
evant only for extreme rain exposures).

Environ. Health Perspect. 128(12) December 2020
Environmental Health Perspectives

126001-2

128(12) December 2020
Lagged exposure–outcome pairs were for extreme rain, flood, or rain. Because we did not know a priori which lags were most relevant for each exposure or location and did not wish to bias our results by picking associations based on their significance and effect size, we retained information from all available lags, exploring the relevance of the lag between exposure and outcome as a separate analysis. We accounted for nonindependence associated with having multiple lags within the same model (described below in the meta-analysis section).

**Confounding**

Although we were unable to adjust published estimates for additional variables beyond those provided in the text of each article, we extracted a list of all variables that were controlled for in the statistical model for each association included in our review. Authors controlled for a variety of variables, but we expected that temperature would be the most likely variable to bias the overall results. As a sensitivity analysis, we re-ran the main effect models, comparing results from studies that adjusted for temperature with those that did not.

**Effect Modifiers**

We specified the following effect modifiers a priori as potential sources of heterogeneity: study design, range of lags considered, frequency of measurements, climate zone, study location, urban vs. rural location, primary water source, community vs. hospital based, identified pathogens, and included age groups. These variables were used to identify sources of variation in the published estimates and to provide a statistical test for possible reasons why the relationship between rain and diarrhea might be context specific. When possible, we selected estimates that were stratified by the prespecified effect modifiers. Because the threshold used to define extreme rain events varied across studies, we also tested to see whether varying the threshold value used to define the cut point affected the strength of the rain–diarrhea association.

For extreme rain, we also extracted whether or not a given extreme rain association stratified by prior rainfall conditions. The definition of prior rain conditions was based on how the authors described their study context. A total of four extreme rain articles stratified their extreme rain associations by prior rain conditions: Carlton et al. 2014, Chhetri et al. 2017, Bush et al. 2014,
and Mertens et al. 2019. Of these four studies, Carlton et al. defined prior rain conditions based on average rainfall in the prior 8-wk period, which was then converted into tertiles (2013). Chhetri et al. defined prior rain conditions based on whether there were 30 or more dry days in the prior 2-month period. For this article, extreme rain events occurring after a period with less than 30 d of rainfall was coded as following “moderate” rain because of the lack of data on the amount of rainfall characterizing such a period. Bush et al. defined prior rain conditions based on the season in which the extreme rainfall occurred: premonsoon (assumed to be a dry period), early monsoon (assumed to be a moderate period), and late monsoon (assumed to be a wet period). Mertens et al. defined prior rain based on rain occurring over the prior 60 d, which was converted into tertiles.

We extracted study design features, such as overall frequency of data collection and range of lags used, as potentially relevant effect modifiers because of the short time scale over which rainfall runoff processes operate. We hypothesized that studies with lower frequencies of data collection would be more likely to obtain null results, whereas studies with daily or weekly measurement lags would report a positive effect more consistently. Climate zones and seasons were extracted because the concentration–dilution hypothesis would predict that rainfall is most likely to be a risk when it either follows a drier season or occurs in an arid climate, where rainfall is low more of the time, allowing pathogens to accumulate in the environment. The climate zone of each study location was coded using the Köppen–Geiger classification method (Kottek et al. 2006). The climate zone was assigned using the centroid of the location described in the text, or specific latitude/longitude coordinates if provided. Because relatively few studies were present for each climate zone, we aggregate zones into two groups: climates with seasonally consistent precipitation (AF: equatorial rainforest; Cf: warm temperate climate, fully humid; or Df: snow climate, fully humid) vs. seasonally varying precipitation (all other climate zones) prior to analysis.

Rural vs. urban and water source locations were extracted because these locations might generate different associations between rainfall and diarrhea and might represent different risk factors. For example, rural locations are less likely to have access to improved sanitation or improved water sources and likely have fewer impervious surfaces, so they might be more vulnerable to concentration–dilution processes at the start of rainy season due to pathogen flushing. It follows that we would expect the effect of rain on diarrhea to be stronger in locations with unimproved sanitation or water; however, it is possible that flood or extreme rain can affect risk regardless by overwhelming the infrastructure. For example, additional water from both extreme rain and flood may overpower combined sewer systems, designed to collect rainwater runoff and wastewater, causing backups in the environment and spreading of potential pathogens (Hata et al. 2014). Because urban areas are more likely to have access to piped water, their waterborne diarrhea risk might be higher after combined sewer overflow events that generally occur later in rainy seasons. Urban areas may also have higher levels of baseline risk from direct transmission due to their higher population density and may have higher coverage of impervious surfaces that accelerate runoff.

We also extracted the World Bank Human Development Index for each location (World Bank 2020), because the level of socioeconomic development might influence population vulnerability to rainfall events. Pathogen taxa were extracted because previous studies have shown that rain may be a risk factor for bacterial diarrhea but have not yet demonstrated this association for viral or protozoan diarrhea (Levy et al. 2016; Mertens et al. 2019). Finally, age was extracted because children are at a higher risk of diarrheal disease than adults are (UNICEF 2009), may interact with their environment in different ways than adults do (Medgyesi et al. 2018), and are more susceptible to different types of pathogens than adults are (Walker et al. 2010; Kotloff et al. 2013). We also extracted hospital vs. community-based studies as a possible indication of disease severity, but there was little variation within each exposure group (i.e., the majority of studies that measured seasonal exposures were conducted in hospital-based settings).

**Quality Assessment**

We adapted the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) scale to our specific exposure variables to assess the quality of the studies that met our inclusion criteria (Atkins et al. 2004). Our quality scale was adapted from a previous review on a similar topic, and points were awarded based on the following categories: hypothesis, length of data collection, multiple estimates and risk of bias, exposure data source, exposure definitions, exposure-specific model criteria, outcome data source, and model stratification (Carlton et al. 2016). We graded study outcomes for their quality using a point system based on a set of eight criteria, detailed in Table S1. GRADE scores were applied to every relevant outcome reported by a study and scores ranged from 0 to 9 points, with higher scores indicating higher quality. GRADE Scores were assigned for each association that was extracted (including those associations that were only included in the qualitative synthesis). The GRADE score for each association is shown in Excel Table S1, along with the subscore for each category.

A.N.M.K. and O.M. completed data extraction and quality rankings. Disagreements were discussed until a consensus was reached. J.N.S.E. provided a third evaluation as needed.

**Qualitative Analysis**

When possible, we assessed the direction of the association between each climate exposure and diarrhea. We described associations that overlapped the null as ‘neutral.’ We also included in our classification studies that reported an association as ‘significant’ and described the direction of the association but did not provide a point estimate. We compared the qualitative findings from our review with those from a prior review that investigated the impact of flood and extreme rain on diarrhea (Levy et al. 2016). Additionally, we summarized the mechanisms authors described as possible reasons for their findings. We pre-specified four categories of explanations that were directly related to the concentration-dilution hypothesis (concentration, dilution, concentration and dilution, or other). Among the other mechanisms, we developed other categories of mechanisms based on the general themes described in the studies. We examined these descriptions at both the association and study level.

**Quantitative Analysis and Meta-Analysis**

The two criteria needed to be included in the meta-analysis were that a) the association of interest approximated the rate ratio; and b) either the standard error (SE) or information needed to calculate the variance was available. For rainfall exposures, only associations using continuous rain as the exposure were included, as there were too few associations with other exposure definitions to conduct a meta-analysis.

When possible, we converted all measures of association to the rate scale, using the rate ratio as the effect measure of interest. Where data were presented on the risk scale but counts and relative follow-up time were available, we used these data to calculate the rate ratio. We also included associations on the risk or odds scale that would be expected to approximate the rate ratio. Briefly, associations would be expected to approximate rate if
either: a) the disease was rare in all exposure groups (such that this risk and rate are comparable for short follow-up durations); b) data to calculate the risk ratio came from an evenly spaced time series where the only assumption required for comparability was a constant population size; or c) data came from a case-control study where some type of time-matching was used [so that the odds ratio (OR) directly approximates the rate ratio], d) the OR was reported, but the data did not come from a case-control study and the outcome was rare (<10%) in all exposure groups.

Our meta-analysis included 1,209 associations, 315 unique exposure-outcome groups (that either contained one association or varied only by lag), and 60 studies. Because the log link was used in our meta-analysis models, point estimates and variances with a value of zero could not be included. To address this issue, we added a small constant ($1 \times 10^{-5}$) to point estimates and variance values that were reported as being exactly equal to zero. This type of modification has been described previously (Berry 1987). All regression models were weighted by inverse variance.

### Averaged Lag Models

For all averaged models, we took the average effect estimate ($\theta$) and SE for associations that varied only by lag group, using the geometric mean. These models therefore considered a total of 315 rows of pooled data. We then used meta-regression models that weighted each estimate of effect by its inverse variance (weighting by GRADE score was also considered as a sensitivity analysis, described in more detail below). Residual variation between studies was accounted for using random effects, with a random intercept being included for each study. A sample model equation for estimate in lag group $j$ in study $i$ is shown below. Because the relationship between the outcome and its predictors would only be expected to be linear on the log scale, we took the log of each effect estimate prior to running the regression model. To obtain the results without effect modification, the overall intercept term (without other explanatory factors) was exponentiated, producing the average rate ratio across all studies.

$$ \log(\theta_{ij}) = \beta_0 + b_i. $$

The $b_i$ terms are random intercepts for study, which are assumed to be uncorrelated with the residual errors. To obtain the regression modeling results, other covariates were added to the same regression model. The model was coded such that the group of interest was set as the reference, and the rate ratio for each group was then derived by exponentiating the overall intercept term. For example, to assess the effect of prior rain conditions, the following model was used:

$$ \log(\theta_{ij}) = \beta_0 + \beta_1 \text{PriorModerate}_i + \beta_2 \text{PriorWet}_i + b_i. $$

The global test of moderators provided the $p$-value for the significance of the interaction. In the equation above, the exponentiated $\beta_0$ provided the estimate for extreme rain events following dry periods, and the model was re-run with changing the reference group to obtain overall effect estimates for extreme rain following a wet period and following a moderate period.

We used these averaged regression models to evaluate the association between the four climate exposures of interest (rainfall, extreme rain, flooding, rainy season) and diarrhea. To account for the fact that in some cases multiple associations were extracted that only varied by lag, we calculated the average effect size and corresponding variances (using the geometric mean) and used the average effect size and variance in our random effects models, including a random effect for study. We used the R package metafor, version 2.4-0 (R Development Core Team, version 4.0.2) and specified a compound symmetry covariance structure.

To identify how the relationship between the four climate exposures of interest and diarrhea varied, we added prespecified effect modifiers (tested independently, one at a time) to the regression model and estimated the effect estimate for each group. In many cases we were limited by both the small number of relevant studies and the precision of those studies when formally testing for effect modification. For this reason, we describe results both for statistically significant effect modification and for those with strong effect sizes where the overall $p$-value for statistical interaction was marginal. We did not do formal tests for effect modification in cases where there were fewer than three association groupings for a given exposure category.

As a sensitivity analysis, we also re-ran all final models (both main effects and effect modifiers) using a) clustered SE models; and b) fixed-effects models (described below).

### Clustered SE Models

As a sensitivity analysis to account for variability in associations by lag while also capturing the dependent structure, we also ran regression models including all associations, including those that varied only by lag, but added a second random effect for lagged group, thereby accounting for the two-level correlation structure. The overall rate ratio was estimated for each exposure using the averaging method including random effects for study but without any covariates (total of four models). We also checked for commonalities across studies in terms of the most relevant lag for each exposure by estimating pooled associations at the lag level.

For clustered SE models, we included all 1,209 associations. As with the averaged models, each estimate was weighted by its inverse variance. We then used random effects models with random intercepts for both study $i$ and lag group $j$ to assess the relationship between the outcome (using the outer|inner coding method in metafor with a compound symmetry structure). A sample model equation for estimate in lag group $j$ in study $i$ is shown below.

$$ \log(\theta_{ij,k}) = \beta_0 + \beta_1 + \beta_{ij}. $$

### Fixed Effects Models

Fixed effects models included all lag groups (averaging associations across lags before analysis, for 315 data rows). However, unlike the average effect models, no random effects were included.

### Results

A total of 111 studies and 1,963 associations were included in our qualitative synthesis. The majority of studies’ exposures were categorized as rainfall ($n=51$), followed by season ($n=37$), flooding ($n=26$), and heavy rain ($n=19$). We identified only one study describing drought, so we did not consider this exposure any further. For all four remaining climate exposures, authors often reported multiple associations per exposure category and the majority of associations reported were neutral, even among studies that also reported positive or inverse associations (Table 1). Overall, 32.4% of articles (36/111) had only one estimate. The remaining 75 articles reported more than one association. No “season” associations considered a lag, but 36 articles considered different lags between exposure and outcome, which is 40% of flood, extreme rain, and rain articles (36/90). For all four exposures combined, 21.6% (24/111) of articles assessed associations for more than one geographic location, 20.7% (23/111) provided data for more than one outcome/pathogen, 18% (20/111)
provided data for more than one exposure category (i.e., both season and extreme rain), and 13.5% (15/111) of articles provided information for more than one way of operationalizing the exposure (i.e., using the 95th and the 99th percentile to define extreme rain events within the same study). Smaller percentages of studies assessed population prior climate conditions (9.0%, 10/111), age (6.3%, 7/111), or water source (2.7%, 3/111). Some authors presented associations stratified by multiple variables (e.g., lags and geographic location); hence percentages do not sum to 100% (see Excel Table S6 for more details about the different variables defining unique association groups for each article). Studies are located in countries representing a range of development conditions and most continents, except Antarctica (Figure 2). However, most studies were conducted in temperate or tropical climates and relatively few studies were from climate zones where precipitation commonly falls in the form of snow.

In general, studies varied widely in both the frequency of data collection and the overall length of the time series considered (Table S2). Studies describing extreme rain exposures at different lags predominantly used daily data (85.3%), whereas studies for flood and rain used more variable temporal units. For flood, 57% used daily data, 20% used weekly data, and 23% used monthly data. For rain, 53% used monthly data, 33% used weekly data, 5% used daily data, and 9% used some other temporal summary method. The length of time series used for each study varied widely, but the majority of studies had long time series. The average time-series length was 6 y for extreme rain, 9 y for rain, 5 y for flood, and 5 y for season. See Table S2 for more details.

Table 1. Direction of associations across articles (overall section) and for individual articles.

| Overall | Extreme rain | Flood | Rain | Season |
|---------|--------------|-------|------|--------|
| Inverse associations | 3.2% (24/739) | 4.3% (32/752) | 11.7% (39/333) | 32.3% (21/65) |
| Neutral associations | 69.3% (512/739) | 79.7% (599/752) | 66.7% (222/333) | 36.9% (24/65) |
| Positive associations | 27.5% (203/739) | 16.1% (121/752) | 21.6% (72/333) | 30.8% (20/65) |
| Articles with positive associations | | | | |
| For articles with multiple associations | 91.7% (11/12) | 82.4% (14/17) | 61.8% (21/34) | 56.3% (9/16) |
| Articles with single associations | 50% (2/4) | 75% (3/4) | 50% (7/14) | 66.7% (6/9) |
| Articles with positive associations | | | | |
| (combining articles with single and multiple estimates) | | | | |
| Current literature review | 81% (13/16) | 81% (17/21) | 58% (28/48) | 60% (15/25) |
| Articles with positive associations in Levy et al. 2016 | 71% (10/14) | 76% (19/25) | — | — |

Note: —, no data.

aAt the association level (combining across different articles).

bAt the article and exposure level. This section combines associations across different lag groups for the same exposure within the same article. For example, an article that reported different extreme rain associations at different lags and for different prior rain conditions would be combined in the extreme rain column.

Figure 2. Map of associations and studies included in the qualitative synthesis for (A) extreme rain, (B) flood, (C) rain, and (D) season. Each point corresponds to a study. The one study and association with the climatic exposure of drought was located in Tuvalu and is excluded from the map.
Of the 111 studies that met our inclusion criteria, 51 articles were excluded from the quantitative synthesis (Figure 1) but were included in the qualitative summary. The main reason for exclusion from the quantitative analysis was lacking information on variance, which was needed for weighting. Several studies reported associations that could not be approximated to a rate ratio due to noncomparable sampling methods and reporting of risk for diseases that were not rare. For flood, extreme rain, and season, too few studies reported data on the linear scale to analyze these associations separately. Thus, only studies reporting estimates on the logarithmic scale were included in the quantitative analysis. For the rainfall exposure, definitions were not comparable for many studies, and ranged from average monthly precipitation to total runoff. This meta-analysis focuses on studies that measured continuous rainfall because it was the only category with a sufficient number of associations. Of the continuous rainfall exposures, we selected only associations that could approximate rate.

A qualitative summary of our findings and how they relate to our original hypothesis can be found in Table 2. An overall summary of point estimates from all studies from our extraction and from those considered previously by Levy et al. (2016) is in Table 1. Most articles reported at least one positive association for both extreme rain (81%) and flood (81%), similar to what Levy et al. found previously. For rain and season, the literature was less consistent, with 58% of rain articles presenting at least one positive association and 60% of season articles.

When evaluating author-hypothesized mechanisms, we found that the majority of studies described some “other” mechanism not directly related to concentration–dilution processes (Table 3). Among the “other” explanations identified across all studies (including those that indicated that concentration–dilution processes might also be a driver), most common explanations included direct exposure to contaminated water; statistical or methods issues; inadequate water, sanitation, and hygiene (WASH) infrastructure/practices; and treatment failure. Many authors also described contamination of drinking water after rainfall due to inadequate sanitation, which provided a mechanism for contamination of drinking water apart from runoff processes. A “concentration” process was the second most common explanation provided by authors. Very few authors indicated that dilution might play a role in decreasing diarrhea. In addition, many authors did not provide any explanation for their findings (more than 22% of all studies, 25/111).

**Extreme Rain**

We included 364 estimates from 13 studies in our quantitative synthesis of the effect of extreme rain on diarrhea. Studies measuring extreme rain tended to have high GRADE scores [mean: 7.70, standard deviation (SD): 0.74]. Without considering effect modification by the prior level of rain, pooled estimates suggest no statistically significant association between extreme rain and diarrhea (Table 4). However, extreme rain was a risk when preceded by a dry period (incidence ratios (IRR) = 1.26, 95% CI: 1.05, 1.51) but not when following a moderately wet or wet period. The point estimate tended to be lowest when extreme rain followed a wet period, but this pattern was statistically significant for only one study (Carlton et al. 2014). For extreme rain, the strongest associations were between 0–2 wk after extreme rain events. Within 0–1 wk and 1–2 wk post exposure, the point estimates were similar, suggesting that daily data provided no additional information over a weekly analysis (Figure S1). Forest plots of dry to moderate prior rain level and wet prior rain level modifying the relationship between extreme rain and diarrhea can be found in Figures S2–S3.

Additionally, the effect of extreme rain was stronger and statistically significant among studies that defined extreme rain based on a storm event compared with studies that used a percentile cut point to define extreme rain events (IRR = 2.51, 95% CI: 2.03, 3.10). Although the percentile cut points used to define extreme rain events were locally specific, and the rainfall amount corresponding to this cut point varied widely, the numeric value (magnitude) of the cut point corresponding to the extreme rain event definition did not modify the association between extreme rain and diarrhea ($p = 0.76$). The two studies defining extreme rain events based on storm events did not provide a numeric rainfall value for each storm (Mukhopadhyay et al. 2019; Kang et al. 2015). One of the articles using the 90th percentile also did not provide the numerical value of rainfall corresponding to that threshold (Wu et al. 2014). For the remaining articles, the numerical rainfall value corresponding to each cut point was 16.82 mm for the 90th percentile (range: 16.82–16.82), 15.0 mm for the 90th percentile (range: 15.0–15.0), and 15.66 mm for the 90th percentile (range: 15.66–15.66). Thus, the numerical depth of rainfall corresponding to each cut point was similar on average, except for studies using the 90th percentile. The effect of pathogen type on the relationship between extreme rain and diarrhea could not be assessed because only one study measured pathogen-specific diarrhea; all other studies reported on all-cause diarrhea. Other hypothesized effect modifiers did not appear to have meaningful heterogeneity between subgroups. Because GRADE scores were generally high for nearly all associations, results were similar when analysis was restricted to the highest-quality studies (Excel Table S2, Table S3, Table S4).

**Season**

We identified 62 associations from 24 articles in our meta-analysis for season exposures. Studies that measured season exposures tended to have lower and more variable GRADE scores (mean: 4.58, SD: 0.967; Excel Table S4). In general, the rate of diarrhea was higher during rainy seasons, but the overall association was not statistically significant (IRR for all pathogens = 1.16; 95% CI: 0.981, 1.24 (Figure S4); however, this association varied by pathogen, with both bacterial and parasitic diarrhea being more common in rainy seasons (IRR for bacteria = 2.70; 95% CI: 1.20, 4.90; IRR for parasites = 2.76; 95% CI: 1.32, 5.77). Diarrhea caused by viruses tended to have inverse associations, albeit nonsignificant (IRR = 0.844; 95% CI: 0.530, 1.35). When viral diarrhea was limited to studies of rotavirus, the association was strong and inverse (IRR = 0.461; 95% CI: 0.291, 0.729). The remaining two viral diarrhea estimates were for norovirus and one (Becker-Dreps 2017) was strongly positive (IRR = 2.94; 95% CI: 1.70, 5.10) and the other (Hossain et al. 2019. Clinical Infectious Diseases) was strongly inverse (IRR = 0.406; 95% CI: 0.228, 0.724). For parasitic diarrhea, point estimates were strongest for diarrhea caused by Cyclospora (Figure S4). There was no clear difference by pathogen for bacterial diarrhea (Figure S4).

In addition, the rainy season peak was most striking and statistically significant in rural areas (IRR = 1.55; 95% CI: 1.02, 2.36) and in low-income countries (IRR = 1.81; 95% CI: 1.15, 2.85). Studies of upper- and upper middle-income countries had the strongest point estimates but wider confidence intervals. Forest plots for all three exposures can be found in Figures S4 (associations by pathogen), S5 (associations by urbanization level), and S6 (associations by income level). When restricting to studies with higher GRADE scores, our point estimate was more positive but remained statistically insignificant with wider confidence intervals (Table S3; Table S4). Forest plots for all season associations in Table 4 are in Figures S4–S6.
Table 2. Summary of the concentration-dilution processes related to different climate exposures, and findings from the literature across studies.

| Exposure | Original generated hypothesis | Mechanism | Supported by the literature? | Additional findings from the literature |
|----------|-------------------------------|-----------|-----------------------------|----------------------------------------|
| Extreme rain | Heavy rain is a risk following dry periods but protective following wet periods | Rainfall-runoff processes | Partially; heavy rainfall was a risk following dry periods | (i) Weekly data appears adequate to assess this association (no appreciable heterogeneity within weeks) (ii) Positive associations tended to be clustered in the first 2 weeks after the event. (iii) Studies that used storm-based definitions of extreme rain had the strongest positive associations |
| Flood | (1) Overall increased risk (2) Weaker and/or protective effect at long lags | (1) Floods overwhelm infrastructure (2) Runoff processes lead to dilution | (1) Yes (2) Unable to test | (1a) These associations were strongest for bacterial and all-cause diarrhea (b) The strongest positive associations were 4-7 days after flood events (2) Studies often did not report the length of flood events in the literature |
| Rain | Stronger effects in arid climates | Rain in an arid climate is more likely to concentrate pathogens | Unable to test | No clear linear association between rain on diarrhea; however, several studies reported non-linear associations, with increased risk at both low and high levels of rainfall |
| Season | Risk will be highest at the start of the rainy season | Rainfall can concentrate pathogens at the start of the season | Partially | Bacterial diarrhea was more common during the rainy season, whereas other pathogens had similar risk throughout the year, suggesting that concentration might be more important for bacterial diarrhea. Although several studies reported that incidence was highest earlier in the rainy season, there was insufficient data to formally test whether the increased risk of rain on pathogen specific diarrheal disease was confined to the beginning of the rainy season. This may arise from minimal studies in arid regions and fewer rainfall events in arid regions, both resulting in lower power. |
We were unable to assess our hypothesis about the timing of risk during the rainy season because too few studies were conducted in arid regions (in which rainfall is different between the rainy and dry seasons), giving our results low power. Additionally, authors typically reported cases for each season in aggregate, making it impossible to determine when during the rainy season the highest risk of diarrhea was experienced. Articles that did provide information about the timing of the seasonal peak in diarrhea observed different peak timing: three articles reported an early rainy season peak, and three articles reported a mid/late rainy season peak instead.

### Flooding

Our search resulted in 699 associations from 14 articles that assessed the relationship between floods and diarrhea. These studies had high GRADE quality scores (average: 7.71, SD: 0.69; Excel Table S3). For flooding, one association had an extremely large point estimate and variance score (IRR = 214, SE = 12.2), suggestive of data sparsity (the flood association from Saulnier et al. 2018 for Chhouk district at 1-month lag). We therefore did not include this estimate in our analysis. Although the overall effect of flood on diarrhea was not statistically significant, our point estimate was relatively strong (IRR = 1.56; 95% CI: 0.913, 2.67) (Table 4). When associations with a GRADE score of less than 7 were excluded, the association was attenuated and remained non-significant (IRR = 1.22; 95% CI: 0.979, 1.51; Tables S3 and S4). However, the association was positive and statistically significant among articles that adjusted for temperature (IRR = 1.03) (Table 4; Figure S9). However, the association between flood and season was not statistically significant for any pathogen (Table 4). The association was strongest between 4 d and 1 wk after flooding event and then decayed through the second week after flooding. There was a second increase in the association between flooding and diarrhea incidence about 4 wk after flooding (Figure S8). We were unable to test the effect of flood duration on diarrhea because authors did not report the length of floods.

### Rainfall

We included 84 associations from 15 articles to evaluate the relationship between rainfall and diarrhea. Studies that assessed rainfall exposures had medium quality GRADE scores and larger standard deviations relative to other exposures (mean: 6.49, SD: 1.28; Excel Table S5). Rainfall had no linear association with diarrhea incidence and did not appear to vary by any of the effect modifiers considered: IRR = 0.998 (95% CI: 0.967, 1.03) (Table 4; Figure S9). However, five articles reported nonlinear associations. Three of these articles reported a U-shaped association between rain and diarrhea, with higher incidence at both low and high levels of rainfall (Fang et al. 2019; Dunn and Johnson 2018; Ikeda et al. 2019). One study reported highest diarrhea risk at moderate rainfall (Chowdhury et al. 2018; Johnson 2018; Ikeda et al. 2019). One study reported highest diarrhea risk at moderate rainfall (Chowdhury et al. 2018; Johnson 2018; Ikeda et al. 2019), and another study found excess risk at the highest rainfall levels (Uejio et al. 2014). There was also a wide range of definitions used for the rainfall exposure, including average monthly rainfall, average daily rainfall, monthly total rainfall, monthly total precipitation, cumulative weekly rain, average rain over the prior 7 d, and average rain over the prior 15 d. Because few studies were conducted in arid climates, we were not able to assess whether the impact of rainfall varied for arid compared with nonarid regions.
Comparison with Fixed-Effects and Clustered Random-Effects Models

We re-ran all models in Table 4 using clustered SEs and fixed-effects models (see Tables S6 and S7). In general, the clustered SEs models were extremely similar to the averaged models. Although the confidence intervals for the fixed-effects model analysis were narrower, the results were broadly consistent between the two methods. The biggest difference was for flood, where overall point estimates became attenuated towards the null using the fixed-effects models, but the point estimate between flooding and bacterial diarrhea became statistically significant.

Comparison with Articles with High GRADE Scores

When aggregating to unique association groupings (i.e., averaged across lags), we also averaged the GRADE score for all associations included in that grouping. As a sensitivity analysis, we ran regression models restricted to association groupings with the highest average GRADE scores (using the top quartile of scores) for each exposure and the results were similar (see Tables S3 and S4).

Confounding by Temperature

As a sensitivity analysis we re-ran the main effects models comparing results adjusting for temperature with those that did not. Overall, 33 of the 60 articles included in the quantitative synthesis considered confounding by temperature. Whether and how each association in our literature review adjusted for temperature is detailed in Excel Tables S1–S5, and the method of adjustment used for each article are shown in Table S8. All articles that considered potential confounding by temperature directly adjusted their regression models for temperature. Most articles used an average temperature variable (21 studies). Other temperature variables considered were apparent temperature instead, which combines information on both temperature and humidity (two studies), maximum temperature (two studies), both average temperature and maximum temperature (one study), minimum and maximum weekly temperature (one study), monthly average of the maximum and minimum temperature (one study), categorical variable for above or below the monthly average (one study), maximum monthly temperature (one study), land surface temperature (one study), anomalous temperature, number of hot days and average temperature (one study), and a cubic spline for temperature (one study).

None of the seasonal associations included in our meta-analysis controlled for temperature (Excel Table S4). For rainfall, results were similar for studies that controlled for temperature in comparison with those that did not (Table S5). For extreme rain, the point estimate was stronger for studies that did not control for temperature, but the point estimate was not statistically significant in either group (Table S5). When restricted to studies that adjusted for temperature, floods had a positive and significant association with diarrhea (see Table S5; IRR = 1.23, 95% CI: 1.04, 1.47).

Impact of Climate Zone

We initially expected that climate zone would modify the association between each exposure and diarrhea. There was only one study for rain exposures conducted in a climate with seasonally consistent precipitation, so we were not able to assess the impact of rainfall climate on the relationship between rain and diarrhea. For the other three climate exposures, there was no statistically significant difference by climate zone (Table S9). The p-value was marginal for season exposures, with studies in areas with seasonally varying precipitation tending to have stronger point estimates (p = 0.06; Table S8).

Publication Bias

Figure S10 shows funnel plots for the review, which are commonly used to assess publication bias. Symmetry and location with respect to the peak of the funnel are used to assess bias, with asymmetry suggesting that associations of a particular direction are more likely to be published. Studies with lower precision are expected to have more variable point estimates, creating a funnel shape (Sedgwick 2013). Overall, publication bias appeared to be minimal for all four associations. For flooding, the majority of associations were clustered around the peak of the funnel, and associations with the largest SEs were all from the same study. Most associations fell within the 95% confidence band. For season, few studies fell within the 95% confidence interval and associations were spread out with varying IRRs and standard errors, indicating low overall precision. Extreme rain associations predominantly fell within the 95% confidence interval but were slightly asymmetric, with the distribution of overall associations generally being slightly right tailed, and with associations with stronger point estimates being more likely to be published relative to the average. Flood, season, and rain funnel plots were generally symmetric with overall effect lines nearing the value of one.

Discussion

Our systematic review and meta-analysis suggest that published studies generally support the concentration mechanism hypothesis, but there is less evidence for the dilution hypothesis. In particular, we found that prior rain modified the effect of extreme rain on diarrhea directly (Table 4). In locations that were previously dry or had moderate rain, extreme rain was associated with increased rates of diarrhea (concentration). We also found partial support for our flooding hypothesis, that flooding may provide an overall increased risk of diarrhea by overwhelming infrastructure. Diarrhea was more common in rainy seasons, particularly for bacterial and parasitic diarrhea. For rotavirus, this association was reversed, with increased incidence in dry seasons. A sudden increase in rain, following a prior dry season, may flush the pathogens into the environment, supporting a concentration mechanism. The association between rainfall and diarrhea appeared to be nonlinear, with higher risk at both low and high levels of rainfall.

Our study complements the prior review by Levy et al. (2009) by considering effect modification by other relevant covariates. Although we were not able to include all articles previously included in Levy et al. due to differences in inclusion criteria, our findings are broadly consistent and add to the growing body of work connecting rainfall and diarrhea.

Although our quantitative analysis provided supporting evidence, few authors considered concentration–dilution processes as hypothesized mechanisms for their results (Table 3). Among all studies combined, concentration was the most common explanation for study results by authors (34% of studies). Dilution was mentioned by only a very small number of studies (4%). Inadequate WASH infrastructure in general and sanitation in particular, along with direct exposure to contaminated water, were also common explanations for findings. These different explanations may all play a role in determining the complex relationship between rain and diarrhea disease. Climatic factors, such as those that relate to the concentration-dilution hypothesis, are more distal risk factors, whereas direct exposure to contaminated water and inadequate infrastructure are more proximal risk factors. Many authors did not articulate an explanation for their findings (22%), suggesting that follow-up research is needed to understand these associations.

When using season as an exposure, disease rates also tended to be higher in the rainy season. During the rainy season, there is
likely more standing water susceptible to contamination. The reason that this risk was restricted to studies examining bacterial and parasitic diarrhea is uncertain. The difference might occur because bacterial diarrhea is predominantly transmitted through contaminated water, whereas rotavirus (the most common viral pathogen in our literature review) can also be transmitted through direct contact or fomites, which may be less affected by rainfall. Moreover, this finding is consistent with prior literature (Levy et al. 2016). This result may have been driven by temperature variations by season, but we were unable to control for temperature due to the nature of the meta-analysis (i.e., summarizing effects across studies reporting on the exposure of interest as rainy vs. dry season). Nevertheless, this result suggests that concentration-dilution processes might be especially relevant in locations where bacterial and parasitic diarrhea predominates. We were also not able to formally assess how this association varied within a season, which could help provide further insight into how pathogens become concentrated during the season. A peak of diarrhea early in the rainy season might suggest initial flushing of pathogens following a previously dry period, whereas a peak later in the season might suggest that the local infrastructure was overwhelmed by persistent rainfall. Follow-up studies that provide more detailed information about the timing of excess risk within a season, controlling for temperature, are needed to help address this gap. However, the generally positive associations between both extreme rain following dry periods and flooding with diarrhea suggest that concentration processes do commonly occur during rainy seasons.

Additional sources of heterogeneity may be related to how studies define climate exposure variables. Extreme rain exposures varied widely in terms of the threshold over which rain was deemed to be an extreme exposure, with the most extreme "storm"-based threshold having the strongest associations with risk of diarrhea. For rain exposures, we found high heterogeneity in how studies defined the exposure, which made synthesizing findings across studies difficult. For example, rainfall has been defined as average monthly rainfall, average yearly rainfall, total monthly rainfall, and maximum water level, among others. Results from different studies were therefore not always comparable.

The fact that authors tended to control for different variables in their analyses is also a source of uncertainty. Although we could not adjust published estimates for potential confounders without access to the raw data, the fact that studies that adjusted for temperature had similar point estimates to those that did not suggests that the main effects were not biased by this covariate. However, it is possible that simple adjustment for temperature is not the best approach in locations where temperature and rainfall are highly colinear. In such contexts, alternative methods, assessment of joint effects of changing temperature and rainfall might be a better approach (Ureña-Castro et al. 2019).

Although this review specifically explored the concentration-dilution hypothesis as it relates to flushing of pathogens, these associations might also be explained by immunity resulting from other sources of diarrhea seasonality. Because most diarrheal infections are incompletely immunizing, concentration of risk at the start of the rainy season might also result from waning of seasonally acquired immunity from the prior diarrhea season, with reduced risk later in the season resulting from increased exposure. It is likely that these empirical associations are the results of multiple processes, of which rainfall-related pathogen concentration is one.

Additionally, the concentration-dilution hypothesis may operate differently in different climates zones. Although climate zone was not a statistically significant effect modifier for any of the exposures considered, it is possible that there was some misclassification of climate exposure for associations covering geographic areas, where more than one climate zone might be represented. Such misclassification may have left us unable to detect significant differences by zone due to those analyses being underpowered. Although we included all articles returned by our literature search, few studies were available for northern climates, where precipitation less commonly falls as rain and more commonly falls as snow. In such locations, vulnerability of pathogen flushing may be related to seasonal shifts in temperature, with increased risk following spring snowmelt, as others have suggested (Harper et al. 2011). Similar patterns might also be observed in locations at high altitude, where snow is more likely.

The relationship between rainfall and diarrhea outcomes has important implications for public health interventions and population vulnerability. Areas that have high levels of bacterial and parasitic diarrhea may want to consider more seasonally specific health preparations than areas where other types of pathogen specific diarrhea predominate. For example, stockpiling cholera vaccines prior to the start of the rainy season, when cholera outbreaks are more common, may enable rapid response in the event of an outbreak (Poncin et al. 2018). If vaccines for other pathogens are developed, similar timing considerations could be useful to minimize disease risk. Additionally, public health practitioners might predict a spike in diarrhea cases directly after extreme rain events, particularly when those events occur after a dry period. In such situations, public health officials might prepare for higher demands for oral rehydration solutions or antibiotics to treat patients effectively. As a preemptive measure, officials could conduct targeted campaigns prior to the start of the rainy season to empty or switch pits for pit latrines and to also distribute soap and chlorine to improve access to WASH at the time when and where burden is likely to be the greatest, particularly in rural regions. Although point estimates for overall diarrhea seasonality are similar for high-income and low-income settings, the absolute burden of diarrhea is generally far higher in low-income settings, so the absolute difference in risk between rainy and dry seasons is likely to be especially pronounced. Given that low-income settings typically have lower access to safe WASH, prioritizing rural regions in low-income settings for interventions targeted at improving baseline WASH infrastructure is important to reduce risk. Using associations between climate exposures and health outcomes, public health practitioners can time interventions, prioritize areas at greatest need, and forecast demand for health services for future changing climates.

Combining information about these general patterns in risk with local understanding about vulnerability to rainfall flushing could provide more certain predictions. For example, less rainfall may be needed to flush pathogens into a local water supply for a city reliant on untreated surface water. In contrast, a rural or peri-urban location also reliant on untreated surface water for drinking water, but with relatively little impervious surfaces, may not face the same risk because the local environment provides a physical mechanism for absorbing pathogens transported by excess water. Although we found that the rainy season pattern of diarrheal diseases was not statistically significant for urban areas, the point estimate was strong, suggesting that other factors, such as local infrastructure, may determine whether or not rainfall increases incidence of diarrhea disease. The threshold for water needed to flush pathogens may also depend on the adaptive capacity of the region, which is likely to be shaped at least in part by typical regional rainfall patterns. This tendency may explain why the categorization of extreme or anomalous rain events for a given location had a stronger apparent effect on diarrhea risk than the magnitude of the rainfall associated with those same events.

The quality of our review is largely dependent on the quality of studies included, which appears to be related to the type of
climate exposure examined. Extreme rain exposures had the highest GRADE scores, followed by flood and rain exposures with medium-quality scores. Season exposures had the lowest-quality scores. The most common methodological problems across all four study categories that led to a reduction in GRADE score were a) the lack of a clear a priori hypothesis; and b) not specifying a final model for the analysis. Both of these weaknesses reflect an underdevelopment of theory relating rainfall to diarrhea. This lack of theory is also evident in the tendency for many authors to fail to provide an explanation for their results. For season, nearly all studies failed to clearly define and justify the seasonal categories used, which in part reflects the complex nature of seasonality and variation in the timing of seasonal rain from year to year. More studies are needed that clearly tie the timing of diarrhea to the onset vs. peak of seasonal rainfall. Because changes in the seasonality and timing, not just intensity and frequency, of rainfall are expected with climate change, studies evaluating the effects of seasonality exposures will be important.

The lack of rainfall–diarrhea theory is somewhat understandable because rainfall may affect diarrhea disease risk through multiple pathways, all of which have different lags. This finding is consistent with findings that there were no consistent statistical differences between lag duration and diarrhea; i.e., we did not identify a statistically significant time lag at which the exposure had a consistent relationship to the outcome. In many locations, the timing of effect between exposure and disease may be the result of many factors, including the pathogen incubation period, the route of exposure, and the structure of the surveillance system (Lo Iacono et al. 2017). For example, rainfall may increase exposure to pathogens through mechanisms other than contaminated drinking water. For all of these reasons, the time between exposure and outcome may be context specific and depend on the primary mechanism of exposure in a particular region.

Future studies that combine environmental sampling data with climate exposures and health outcomes would be useful in helping to distinguish between competing hypotheses for these associations. For example, a recent study that combined water quality sampling, climate data, and environmental sampling found evidence to support the concentration-dilution hypothesis in an area dependent on groundwater, suggesting that the associations are likely mediated through pathways beyond contaminated drinking water (Mertens et al. 2019), likely associated with sanitation. Such studies could also help clarify why diarrhea is more common during rainy seasons and whether spikes in pathogen concentration in water supplies tend to occur early or later during rainier seasons. Additional studies that consider different types of environmental sampling could provide data that would help clarify the interpretation of these results.

The overlapping nature of rain-related exposures also adds uncertainty to this review, particularly for comparing flood exposures (which could potentially be classified as extreme rain or even season in some cases). Logistically, there may not be clear distinctions between events, and it is possible to have concurrent extreme rain and flooding. In locations with seasonally varying precipitation, flood exposures may be capturing overall seasonality because flooding events may be more likely to occur in the rainy season, the coupling of which could result in dilution of pathogens. These complex patterns of heterogeneity might explain why the overall main effect of flood was not statistically significant.

In many cases, selection of measurement scale among many studies appeared to obscure potential causal effects, with several studies reporting estimates on the risk scale. In particular, associations often changed significance and direction depending on whether or not estimates were measured on the risk or the rate scale (see Supplementary Excel Tables S1–S5). Focusing on the risk scale can provide misleading results because positive associations may be purely reflective of differences in follow-up time. For example, if a similar number of cases occur during a rainy season that is only 2 months long compared with a dry season that is 10 months long, the risk ratio would suggest a null association, even though the rate ratio, which accounts for differences in observation time between the two periods, would suggest a strong effect. For this reason, we recommend that the risk scale not be used to describe climatic effects, particularly when the outcome is not rare or the duration of follow-up is highly variable between groups.

Understanding the public health implications of climate change is an integral part of producing effective, regionally specific public health policy. Progress in targeted public health policy development and interventions requires a mechanistic understanding of the relationship between climate and disease. Here we provide insights into the context-specific relationships between rain and diarrhea that help resolve the mixed results present in the literature. The concentration-dilution hypothesis provides an important theoretical underpinning to guide further research on connections between rainfall and waterborne disease. For example, authors should carefully consider how climatic variables are defined, include definitions that support comparability across studies, and target evaluation of a specific mechanism hypothesized to connect the exposure and the outcome.

Studies investigating associations with rainfall more generally should consider effect modification by prior rain level and the shape of the association, to allow for assessment of nonlinear interaction effects. Using continuous rainfall as the exposure of interest rather than categorizing rainfall a priori would help researchers to better assess these nonlinear relationships. Comparisons of risk of diarrhea in rainy and dry seasons should also clearly specify the months that define each season and, if possible, note which time in the rainy/dry season risk is highest. Studies of extreme rain should also specify both the numerical threshold used to define extreme events and the percentile, which would allow researchers to evaluate whether it is the absolute depth of rainfall or its value relative to typical patterns for the region that shapes risk. Consistently reporting whether or not an extreme rain event led to subsequent flooding would also be useful. For flooding, clarifying the timing of the flood relative to rainfall just prior to the flood event would also be useful to determine whether the flood was more likely to concentrate or dilute pathogens. Making raw climate and incidence data available, even in aggregate form to protect the privacy of individuals, would also allow researchers to conduct follow-up analyses more flexibly. Carefully considering potential confounding by temperature and including information about how this problem was handled would also be helpful. Such improvements in study designs can improve our ability to predict how specific climate exposures, and projected changes in those exposures, affect diarrhea, and in turn provide insight into public health intervention design and implementation now and in the future.

Acknowledgments

This work was supported by the Models of Infectious Diseases and Agent Study program within the National Institute of General Medical Sciences (grant U01 GM110712 to A.N.M.K. and J.N.S.E.); the National Institute of Allergy and Infectious Diseases (grant R01 AI050038 to A.N.M.K., O.M., and J.N.S.E., grant 1K01AI103544 to KL, grant R01AI125842 to MCL); and the National Institute of Food and Agriculture Integrated Food Energy Water Systems (grant #TF-1639318 and 1316536 to M.C.L.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the funders.
across a megacity of the developing world. Proc Natl Acad Sci USA 113(15):4092–4097, PMID: 27035945, https://doi.org/10.1073/pnas.1518977113.

McKerlie SA, Bahamouni KM, Kim H, Honda Y, Prentick M, Iddings S, et al. 2015. Assessment of climate-sensitive infectious diseases in the Federated States of Micronesia. Trop Med Health 43(1):28–40, PMID: 25895151, https://doi.org/10.2149/mh.2014-17.

McLer IJ, Lai CM, Buettner PG, Gager P, Chan VS, Hashizume M, et al. 2016. Diarrheal diseases and climate change in Cambodia: environmental epidemiology and opportunities for adaptation. Asia Pac J Public Health 28(7):576–585, PMID: 27485998, https://doi.org/10.1177/1010539516649102.

Medgyesi DN, Brogan JM, Sewell DK, Creve-Coeur JP, Kwong LH, Baker KK. 2018. Where children play: young child exposure to environmental hazards during play in transitioning internally displaced persons community in Haiti. Int J Environ Res Public Health 15(8):1646, PMID: 30081480, https://doi.org/10.3390/ijerph15081646.

Menten T, Babirakhrishnan K, Ramaswamy P, Rajkumar R, Ramaprabha P, Durairaj N, et al. 2019. Associations between high temperature, heavy rainfall, and diarrhoea among young children in rural Tamil Nadu, India: a prospective cohort study. Environ Health Perspect 127(4):47004, PMID: 30986088, https://doi.org/10.1289/EHP3711.

Moher D, Liberati A, Tezlaflf J, Altman DG, The PRISMA Group. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 6(7):e1000097, PMID: 19281972, https://doi.org/10.1371/journal.pmed.1000097.

Mukabutera A, Thomson D, Murray M, Basinga P, Nyirazinyoye L, Atwood S, et al. 2015. Assessment of climate-sensitive infectious diseases in the mountainous informal settlements of Huye town, Rwanda. SpringerPlus 5(1):908–911, https://doi.org/10.1186/s40064-016-2858-3.

Mukhopadhyay AK, Deb AK, Chowdhury G, Debnath F, Samanta P, Saha RN, et al. 2019. Associations between high temperature, heavy rainfall, and diarrhoea risk among children under the age of 5 years in Mbour, Senegal. BMC Public Health 19(1):1, PMID: 30986055, https://doi.org/10.1186/s12889-019-7654-0.

Mukherjee S, Lai CM, Buettner PG, Gager P, Chan VS, Hashizume M, et al. 2016. Diarrheal diseases and climate change in Cambodia: environmental epidemiology and opportunities for adaptation. Asia Pac J Public Health 28(7):576–585, PMID: 27485998, https://doi.org/10.1177/1010539516649102.

Mukhopadhyay AK, Deb AK, Chowdhury G, Debnath F, Samanta P, Saha RN, et al. 2019. Associations between high temperature, heavy rainfall, and diarrhoea risk among children under the age of 5 years in Mbour, Senegal. BMC Public Health 19(1):1, PMID: 30986055, https://doi.org/10.1186/s12889-019-7654-0.

Mukhopadhyay AK, Deb AK, Chowdhury G, Debnath F, Samanta P, Saha RN, et al. 2019. Associations between high temperature, heavy rainfall, and diarrhoea risk among children under the age of 5 years in Mbour, Senegal. BMC Public Health 19(1):1, PMID: 30986055, https://doi.org/10.1186/s12889-019-7654-0.

Mukhopadhyay AK, Deb AK, Chowdhury G, Debnath F, Samanta P, Saha RN, et al. 2019. Associations between high temperature, heavy rainfall, and diarrhoea risk among children under the age of 5 years in Mbour, Senegal. BMC Public Health 19(1):1, PMID: 30986055, https://doi.org/10.1186/s12889-019-7654-0.

Mukhopadhyay AK, Deb AK, Chowdhury G, Debnath F, Samanta P, Saha RN, et al. 2019. Associations between high temperature, heavy rainfall, and diarrhoea risk among children under the age of 5 years in Mbour, Senegal. BMC Public Health 19(1):1, PMID: 30986055, https://doi.org/10.1186/s12889-019-7654-0.

Mukhopadhyay AK, Deb AK, Chowdhury G, Debnath F, Samanta P, Saha RN, et al. 2019. Associations between high temperature, heavy rainfall, and diarrhoea risk among children under the age of 5 years in Mbour, Senegal. BMC Public Health 19(1):1, PMID: 30986055, https://doi.org/10.1186/s12889-019-7654-0.

Mukhopadhyay AK, Deb AK, Chowdhury G, Debnath F, Samanta P, Saha RN, et al. 2019. Associations between high temperature, heavy rainfall, and diarrhoea risk among children under the age of 5 years in Mbour, Senegal. BMC Public Health 19(1):1, PMID: 30986055, https://doi.org/10.1186/s12889-019-7654-0.

Mukhopadhyay AK, Deb AK, Chowdhury G, Debnath F, Samanta P, Saha RN, et al. 2019. Associations between high temperature, heavy rainfall, and diarrhoea risk among children under the age of 5 years in Mbour, Senegal. BMC Public Health 19(1):1, PMID: 30986055, https://doi.org/10.1186/s12889-019-7654-0.

Mukhopadhyay AK, Deb AK, Chowdhury G, Debnath F, Samanta P, Saha RN, et al. 2019. Associations between high temperature, heavy rainfall, and diarrhoea risk among children under the age of 5 years in Mbour, Senegal. BMC Public Health 19(1):1, PMID: 30986055, https://doi.org/10.1186/s12889-019-7654-0.

Mukhopadhyay AK, Deb AK, Chowdhury G, Debnath F, Samanta P, Saha RN, et al. 2019. Associations between high temperature, heavy rainfall, and diarrhoea risk among children under the age of 5 years in Mbour, Senegal. BMC Public Health 19(1):1, PMID: 30986055, https://doi.org/10.1186/s12889-019-7654-0.
Wang H, Di B, Zhang T, Lu Y, Chen C, Wang D, et al. 2019. Association of meteorological factors with infectious diarrhea incidence in Guangzhou, southern China; a time-series study (2006–2017). Sci Total Environ 672:7–15, PMID: 30954825, https://doi.org/10.1016/j.scitotenv.2019.03.330.

Wang P, Goggins WB, Chan EYY. 2018. A time-series study of the association of rainfall, relative humidity and ambient temperature with hospitalizations for rotavirus and norovirus infection among children in Hong Kong. Sci Total Environ 643:414–422, PMID: 29940452, https://doi.org/10.1016/j.scitotenv.2018.06.189.

Wangdi K, Clements ACA. 2017. Spatial and temporal patterns of diarrhoea in Bhutan 2003–2013. BMC Infect Dis 17(1):507, PMID: 28732533, https://doi.org/10.1186/s12879-017-2611-6.

WHO (World Health Organization). 2003. Climate Change and Human Health ~ Risks and Responses: Summary. Geneva, Switzerland: WHO.

WHO. 2018. Climate Change and Health, https://www.who.int/news-room/fact-sheets/detail/climate-change-and-health [accessed 1 December 2020].

World Bank. 2020. Data: countries and economies. https://data.worldbank.org/country [accessed 1 December 2020].

Wu J, Yunus M, Ali M, Escamilla V, Emch M. 2018. Influences of heatwave, rainfall, and tree cover on cholera in Bangladesh. Environ Int 120:304–311, PMID: 30107291, https://doi.org/10.1016/j.envint.2018.08.012.

Wu J, Yunus M, Streatfield PK, Emch M. 2014. Association of climate variability and childhood diarrhoeal disease in rural Bangladesh, 2000–2006. Epidemiol Infect 142(9):1859–1869, PMID: 24169897, https://doi.org/10.1017/S095026881300227X.

Xu X, Ding G, Zhang Y, Liu Z, Liu Q, Jiang B. 2017. Quantifying the impact of floods on bacillary dysentery in Dalian City, China, from 2004 to 2010. Disaster Med Public Health Prep 11(2):190–195, PMID: 27229186, https://doi.org/10.1017/dmp.2016.90.

Yadav R, Somashekar D, Sodha SV, LasersonKF, Venkatesh S, Chauhan H. 2019. Post-flood rapid needs assessment in Srinagar city, Jammu and Kashmir state, India, September, 2014. Disaster Med Public Health Prep 13(2):133–137, PMID: 29559022, https://doi.org/10.1017/dmp.2018.21.

Yomwan P, Cao C, Rakwatin P, Suphamitmongkol W, Tian R, Saokarn A. 2015. A study of waterborne diseases during flooding using Radarsat-2 imagery and a back propagation neural network algorithm. Geomatics, Natural Hazards and Risk 6(4):289–307, https://doi.org/10.1080/19475705.2013.853325.

Yusof N, Hamid N, Ma ZF, Laswenko RM, Mohammad WMZW, Collins DA, et al. 2017. Exposure to environmental microbiota explains persistent abdominal pain and irritable bowel syndrome after a major flood. Gut Pathog 9:75, PMID: 29255480, https://doi.org/10.1186/s13099-017-0224-7.

Zhang F, Liu Z, Gao L, Zhang C, Jiang B. 2016. Short-term impacts of floods on enteric infectious disease in Qingdao, China, 2005–2011. Epidemiol Infect 144(15):3278–3287, PMID: 27312685, https://doi.org/10.1017/s0950268816001084.