Excess risk in infant mortality among populations living in flood prone areas in Bangladesh: A cluster-matched cohort study over three decades, 1988-2017

François Rerolle
UCSF
Benjamin Arnold
UCSF
Tarik Benmarhnia (tbenmarhnia@ucsd.edu)
UCSD

Research Article

Keywords:

Posted Date: November 3rd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2231972/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

The Ganges-Brahmaputra-Meghna river basin, running through Tibet, Nepal, Bhutan, Bangladesh and northern India, is home to more than 618 million people. Annual monsoons bring extensive flooding to the basin, with floods predicted to be more frequent and extreme due to climate change. Yet, evidence regarding the long-term impacts of floods on children’s health is lacking. In this analysis, we used high-resolution maps of recent large floods in Bangladesh to identify flood prone areas over the country. We then used propensity score techniques to identify, among 58,945 mothers interviewed in six DHS population-based surveys throughout Bangladesh, matched cohorts of exposed and unexposed mothers and leverage data on 150,081 births to estimate that living in flood prone areas was associated with an excess risk in infant mortality of 5.9 [2.8;8.9] additional deaths per 1000 births compared to living in non-flood prone areas over the 30-year period between 1988 and 2017. Finally, drawing on national-scale, high-resolution estimates of flood risk and population distribution, we estimated an excess of 170,136 [80,742;256,645] infant deaths over the past 30 years were attributable to living in flood-prone areas in Bangladesh, with marked heterogeneity in attributable burden by subdistrict and heterogeneity in excess risk over time, with a risk difference of 9.2 [2.9;15.5], 1.5 [-2.6;5.7] and 8.9 [3.7;14.2] per 1000 births for the 1988–1997; 1998–2007 and 2008–2017 decades respectively. Our approach demonstrates the importance of measuring longer-term health impacts from floods and provides a generalizable example for how to study climate-related exposures and long-term health effects.

Significance Statement

As a result of climate change, populations are predicted to experience more frequent and extreme flooding events in coming decades. This is particularly true in Bangladesh, a country with low elevation and where annual monsoons bring extensive flooding to the Ganges-Brahmaputra-Meghna river basin. Our study estimated that infants born to mothers living in flood prone areas had a 13% higher chance of dying within their first month compared to matched infants born to mothers living outside flood prone areas of Bangladesh in the past 30 years. Our approach highlighted important differences across space and over time and provides a generalizable framework for how to study climate-related exposures and long-term health effects.

Introduction

Anthropogenic CO₂ emissions and ecosystems degradation have increased the occurrence and magnitude of extreme weather events worldwide¹. Climate-sensitive events such as heat waves, extreme precipitation, and floods impose large economic damages to human populations and directly threaten their health and survival². Among extreme climate events, large floods have increased worldwide with recent severe examples in Bangladesh³⁴, South Africa⁵ or Pakistan⁶ in 2022. The Ganges-Brahmaputra-Meghna river basin, running through Tibet, Nepal, Bhutan, Bangladesh and northern India, is home to more than 618 million people (2010 data) and regularly floods during the annual monsoon⁷. Even in the
most optimistic scenarios considered by the Intergovernmental Panel on Climate Change (IPCC), populations living in the region are predicted to experience more frequent and extreme flooding events in coming decades.

Accurately assessing the human and financial toll in the aftermath of a disaster is essential, but for populations living in flood prone areas and that are repeatedly exposed to large floods, longer term health effects are plausible but understudied. In this context, children are a high-risk subgroup and flooding has been associated with adverse birth and child health outcomes through mechanisms that act over both short term (e.g., exposure to contaminated water) and longer term (e.g., social displacement) timescales. Most studies of health impacts from floods have focused on single events and to our knowledge no studies have assessed how effects have changed over time among populations that are consistently vulnerable to floods. Measuring excess health risks associated with living in flood prone areas, and how such risks have evolved in recent decades provides essential information for population preparedness and contributes estimates of health burden attributable to extreme climate events.

In low-resource settings, disproportionally affected by climate change, such long-term analyses are particularly difficult because of scarce environmental and health data. Large and nationally representative surveys such as Demographic Health Surveys (DHS) provide a unique opportunity to create high-resolution, longitudinal time series for infant mortality through combining birth history information across surveys. Satellites increasingly measure the earth’s environment at a fine spatio-temporal resolution and joining data from novel, remotely sensed climate layers to geo-located DHS survey data creates a new opportunity for epidemiologic studies of climate-related exposures on long-term health outcomes in low- and middle-income settings.

We conducted a long-term study of infant mortality risk associated with living in flood-prone areas of Bangladesh, a country at the intersection of three major rivers flowing from the Himalaya where two thirds of the territory’s elevation lies at or below 5 meters. Every monsoon season — May to October — heavy rains lead to floods that often have devastating consequences for millions of people. In May 2022 for instance, 7.2 million people were affected by floods in nine districts and half a million were displaced. Infant mortality is an extreme outcome of climate-related exposures, but it is a useful outcome in this context because it likely integrates both acute and longer-term effects from floods. Infant mortality is a relatively rare outcome even in high mortality populations, so estimating excess mortality risk associated with environmental exposures requires large cohorts monitored over long periods to ensure sufficient deaths and capture sufficient heterogeneity in exposure. We estimated excess risk in infant mortality associated with living in flood-prone areas of Bangladesh using a matched cohort design, reconstructed over a 30-year period from six, nationally representative Demographic and Health Surveys (DHS) and high-resolution Global Flood Database derived from earth observation data and public health reports.

Results
Identifying flood prone areas throughout Bangladesh

The Global flood database mapped seven country-wide floods that occurred in Bangladesh between 2000 and 2018. We derived flood prone areas throughout Bangladesh using the percentage of days flooded during the seven floods (Fig. 1a). Across the 34% of the country classified as flood-prone in the main analysis, there was a range of severity as measured by days flooded over the seven events (Fig. 1b). In combination with world population data, we estimated that 44 million people lived in Bangladesh’s flood prone areas in 2020.

Reconstructing infant mortality cohorts over 30 years

Of the 60,678 women interviewed in one of the six last DHS surveys conducted in Bangladesh (1999, 2004, 2007, 2011, 2014 and 2017), 58,945 (97.1%) with complete data were included in the analysis. DHS clusters provided near uniform coverage of both flood prone and non-flood prone areas (Fig. 1c). With data on 150,081 births, we used children’s exact date of birth and age at death to reconstruct infant mortality time series between September 1988 and September 2017. Over the 30-year period, infant mortality declined substantially in Bangladesh, from 75 per 1000 live births in 1988 to 25 per 1000 live births in 2017, and populations living in flood prone areas consistently had higher mortality compared with those living in non-flood-prone areas (Fig. 1d).

Socio-demographic characteristics differed between the 43,677 (74.1%) interviewed women living in non-flood prone areas and the 15,268 (25.9%) women living in flood prone areas. Women in flood prone areas tended to have lower levels of education, were poorer as measured by an asset-based wealth index, and lived in more rural areas at lower population density with worse sanitation conditions and more rudimentary floor materials (Table 1). Temperature and rainfall were similar across the study populations (Table 1). There was progressive improvement in overall socio-demographic characteristics over the study period (Table A1).

We matched all mothers living in flood-prone areas to mothers living outside flood-prone areas using a propensity score estimated from characteristics in Table 1 (Methods). After matching, standardized mean differences between groups were < 0.1 for all characteristics, indicating adequate covariate balance (Figure A1).
Table 1
Population characteristics in flooded and non-flooded areas (unmatched samples) measured across six DHS surveys, 1999–2017. All estimates except raw frequencies account for survey sampling weights.

<table>
<thead>
<tr>
<th></th>
<th>Non-flood prone area (N = 43677)</th>
<th>Flood prone area (N = 15268)</th>
<th>Overall (N = 58945)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Place of residence</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>27196 (72%)</td>
<td>12546 (87%)</td>
<td>39742 (76%)</td>
</tr>
<tr>
<td>Urban</td>
<td>16481 (28%)</td>
<td>2722 (13%)</td>
<td>19203 (24%)</td>
</tr>
<tr>
<td><strong>Highest level of education</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No education</td>
<td>11401 (28%)</td>
<td>5147 (34%)</td>
<td>16548 (30%)</td>
</tr>
<tr>
<td>Primary</td>
<td>13407 (31%)</td>
<td>5071 (33%)</td>
<td>18478 (31%)</td>
</tr>
<tr>
<td>Secondary</td>
<td>15065 (34%)</td>
<td>4359 (28%)</td>
<td>19424 (33%)</td>
</tr>
<tr>
<td>Higher</td>
<td>3804 (8%)</td>
<td>691 (4%)</td>
<td>4495 (7%)</td>
</tr>
<tr>
<td><strong>Wealth index quintile</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poorest</td>
<td>6144 (14%)</td>
<td>3274 (22%)</td>
<td>9418 (17%)</td>
</tr>
<tr>
<td>Poorer</td>
<td>7177 (17%)</td>
<td>3187 (21%)</td>
<td>10364 (18%)</td>
</tr>
<tr>
<td>Middle</td>
<td>7825 (18%)</td>
<td>2950 (20%)</td>
<td>10775 (19%)</td>
</tr>
<tr>
<td>Richer</td>
<td>8513 (20%)</td>
<td>2519 (17%)</td>
<td>11032 (19%)</td>
</tr>
<tr>
<td>Richest</td>
<td>9817 (20%)</td>
<td>1975 (12%)</td>
<td>11792 (18%)</td>
</tr>
<tr>
<td>Not measured in 1999 survey</td>
<td>4201 (10%)</td>
<td>1363 (9%)</td>
<td>5564 (10%)</td>
</tr>
<tr>
<td><strong>Number of children</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>2.91 (± 1.75)</td>
<td>3.17 (± 1.93)</td>
<td>2.98 (± 1.81)</td>
</tr>
<tr>
<td><strong>Age of mother at first birth (years)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>17.8 (± 3.23)</td>
<td>17.7 (± 3.12)</td>
<td>17.7 (± 3.20)</td>
</tr>
<tr>
<td><strong>Source of drinking water</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improved</td>
<td>42733 (98%)</td>
<td>14790 (98%)</td>
<td>57523 (98%)</td>
</tr>
<tr>
<td>Unimproved</td>
<td>944 (2%)</td>
<td>478 (2%)</td>
<td>1422 (2%)</td>
</tr>
<tr>
<td><strong>Sanitation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improved</td>
<td>26337 (59%)</td>
<td>7724 (50%)</td>
<td>34061 (56%)</td>
</tr>
<tr>
<td>Unimproved</td>
<td>17340 (41%)</td>
<td>7544 (50%)</td>
<td>24884 (44%)</td>
</tr>
<tr>
<td></td>
<td>Non-flood prone area (N = 43677)</td>
<td>Flood prone area (N = 15268)</td>
<td>Overall (N = 58945)</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>----------------------------------</td>
<td>-----------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td><strong>Sanitation facilities shared with other HH</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>23552 (52%)</td>
<td>7940 (51%)</td>
<td>31492 (52%)</td>
</tr>
<tr>
<td>Yes</td>
<td>11210 (27%)</td>
<td>4242 (28%)</td>
<td>15452 (27%)</td>
</tr>
<tr>
<td>Not measured in 1999 and 2004 surveys(^1)</td>
<td>8915 (21%)</td>
<td>3086 (20%)</td>
<td>12001 (21%)</td>
</tr>
<tr>
<td><strong>Main floor material</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finished</td>
<td>14173 (30%)</td>
<td>3058 (19%)</td>
<td>17231 (27%)</td>
</tr>
<tr>
<td>Rudimentary</td>
<td>29504 (70%)</td>
<td>12210 (81%)</td>
<td>41714 (73%)</td>
</tr>
<tr>
<td><strong>Main wall material</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finished</td>
<td>31279 (69%)</td>
<td>11091 (75%)</td>
<td>42370 (71%)</td>
</tr>
<tr>
<td>Rudimentary</td>
<td>12398 (31%)</td>
<td>4177 (25%)</td>
<td>16575 (29%)</td>
</tr>
<tr>
<td><strong>Main roof material</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finished</td>
<td>41753 (95%)</td>
<td>14528 (96%)</td>
<td>56281 (95%)</td>
</tr>
<tr>
<td>Rudimentary</td>
<td>1924 (5%)</td>
<td>740 (4%)</td>
<td>2664 (5%)</td>
</tr>
<tr>
<td><strong>Land surface temperature (°C)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>24.6 (± 0.589)</td>
<td>24.4 (± 0.504)</td>
<td>24.5 (± 0.571)</td>
</tr>
<tr>
<td><strong>Rainfall (mm)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>2460 (± 720)</td>
<td>2450 (± 766)</td>
<td>2460 (± 733)</td>
</tr>
<tr>
<td><strong>Population density (per km(^2))</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^1\)Wealth index and sanitation sharing weren't reported in all DHS questionnaires. As matching was conducted within DHS wave, missingness here is not an issue.

**Excess infant mortality associated with living in flood-prone areas**
We estimated infant mortality over the study period among children born to women living in flood prone areas and their matched controls living in non-flood prone areas, and from the infant mortality rates estimated the risk difference and risk ratio of death by age 30 days (Methods). Over the 30-year period, there was an excess 5.9 deaths [95% CI: 2.8 to 8.9] per 1,000 births among children born in flood prone areas compared to the matched cohort born outside flood prone areas, corresponding to a 13% [6%; 21%] relative increase (Fig. 2, Table A2, Table A3, Figure A5).

Heterogeneity In Excess Risk Over Time And Seasons

We found heterogeneity when testing for subgroup differences (p = 0.03) across strata defined by decades and seasons (Fig. 2, Table A2). Over time, we found an effect in the earlier 1988–1997 decade (RD = 9.2 [2.9; 15.5] per 1000 births) that disappeared during the 1998–2007 decade (RD = 1.5 [-2.6; 5.7] per 1000 births) but came back in the most recent 2008–2017 decade (RD = 8.9 [3.7; 14.2] per 1000 births). Aggregated over the three decades, we found a similar excess risk in infant mortality of 7.3 [2.6; 12.0] and 4.7 [0.6; 8.7] per 1000 births during the rainy and dry seasons respectively (p = 0.4 on the Cochran Q test for heterogeneity). Stratifying across both decade and season highlights that the causal effect in the earlier 1988–1997 decade used to concentrate in the rainy season (p = 0.02 on the Cochran Q test for heterogeneity) whereas in the most recent 2008–2017 decade, the excess risk in infant mortality happened over both the rainy and dry seasons (p = 0.94 on the Cochran Q test for heterogeneity). In the appendix, Table A2 presents estimates in Fig. 2; Figures A2, A3 and A4 show the full results from the meta-analyses; and Table A3 and Figure A5 summarize the results on the risk ratio scale.

Small area estimation of excess mortality burden

We combined estimates of excess risk with WorldPop population estimates to generalize estimates of excess mortality over the country (Methods), highlighting marked fine-scale heterogeneity in populations most affected (Fig. 3a). Heterogeneity in excess infant deaths reflects the locations of flood-prone areas (Fig. 1a) and actual population settlement patterns. We aggregated 1km² projections of excess infant deaths by subdistrict to generate small area estimates and identify most affected regions throughout Bangladesh (Fig. 3b & 3c). Four subdistricts stand out with excess mortality attributable to living in flood prone areas above 4 per 1000 deaths while the 15 least affected subdistricts had excess infant mortality below 1 per 1000 births. Aggregating over the entire country, we estimated a total of 170,136 [95% CI: 80,742 to 256,645] excess infant deaths attributable to living in flood prone areas between 1988 and 2017 in Bangladesh.

Methods

Identifying high resolution flood prone areas

The Global Flood Database provides high resolution maps of 913 large flood events that have happened around the world since 2000, providing a unique opportunity to measure flood exposure from
GPS coordinates alone and identify which communities are living in flood-prone areas or not. The number of flooded days during any of the seven country-wide floods available in the global flood database\textsuperscript{12} that happened in Bangladesh (2002, 2003, 2004, 2007, 2010, 2010 and 2017), was extracted, summed, and divided by the total maximum number of flooded days across all seven flood events. The resulting layer gives, for every 250m\textsuperscript{2} pixel in Bangladesh, the percent number of days it was flooded during these seven flood events and can be leveraged to determine flood prone areas. Specifically, we classified as flood prone area any pixel where the percent number of flooded days was strictly higher than zero. In sensitivity analyses, other thresholds were considered, based on quartiles of the distribution of percent number of flooded days among pixels flooded at least once in one of the seven flood events.

**Infant mortality**

The six last demographic health surveys (DHS) conducted in Bangladesh – 1999\textsuperscript{14}, 2004\textsuperscript{15}, 2007\textsuperscript{16}, 2011\textsuperscript{17}, 2014\textsuperscript{18} and 2017\textsuperscript{19} - included GPS coordinates and were used to extract infant mortality time series in flood prone areas and non-flood prone areas. Each survey wave is a random cluster survey of households across the country that collects, among other things, the birth date of any of the household's children -from interviewed women younger than 49 years old - as well as the age at death for those no longer alive at the time of interview. The proportion of children that died before reaching 1 month, including stillbirths, was used to estimate monthly infant mortality time series in the 30 years leading up to the last DHS survey wave, i.e between September 1987 and September 2017. Data was restricted to singleton births and births happening before mothers lived where they were interviewed were excluded.

**Matching analyses**

Every mother that gave birth in the flood prone area was matched to a mother that gave birth in the non-flood prone area with similar socio-economics and environmental characteristics. Propensity score matching with the nearest neighbor algorithm without replacement was used, within each DHS survey wave, on the variables presented in Table 1: region, urban/rural place of residence, mother's birth date, age of mother at first birth, total number of children, highest level of education, wealth index, source of drinking water, sanitation, floor, roof, and wall materials, whether the toilets are shared with other households as well as 2015 average land surface temperature\textsuperscript{20}, precipitation\textsuperscript{21} and population density\textsuperscript{22}. SES variables were directly extracted from the questionnaire while environmental covariates, extracted at cluster locations by the DHS program, were downloaded when requesting access to the DHS data. Including sampling weights did not improve covariates balance and were therefore discarded for matching but accounted for in statistical analyses below.

**Primary analysis: a cluster-matched cohort study**

Matched infant mortality time series in flood prone areas and non-flood prone areas were aggregated by season to compute infant mortality for every dry and rainy season of the 30 years between September 1987 and September 2017. The risk difference, risk ratio and odds ratio and their standard errors for these 60 time periods were empirically computed and fed to a meta-analysis with random effects,
stratified by decade (1988–1997 vs 1998–2007 vs 2008–2017) and season (dry vs rainy). Stratified and overall estimates represent the adjusted causal effect (here our main estimand is the Average Treatment among the Treated: ATT) of living in flood prone areas on infant mortality. The analysis was repeated for all thresholds used to define flood prone areas in sensitivity analyses (Figure A6). Using the standard definition, the rainy season consisted of the May to October months, further supported by precipitation data (Figure A7).

We combined overall estimates of excess risk with high-resolution population data to compute small-area estimates of excess mortality across Bangladesh. We used WorldPop13 1km² resolution population estimates from 2003, the median of our study period, and considered an average annual birth rate of 25 per 1000 people (World Bank data) to calculate excess mortality with the following formula:

\[
\text{excessmortality} = \text{population} \times \text{birthrate} \times \text{studyperiodlength} \times \text{riskdifference}
\]

Finally, high-resolution projections were aggregated both over the entire country and by subdistrict to identify most affected regions of Bangladesh.

**Sensitivity analysis: individually matched analysis**

In a secondary analysis, instead of aggregating births within flooded and non-flooded areas, we used conditional logistic regression to directly estimate the odds ratio of living in a flood prone area compared to living in a non-flood prone area on infant mortality conditioning on each matched pair of mothers, with pairs matched using the propensity score approach described above. Using this alternative analysis approach, the stratified meta-analysis of causal effects for each of the 60 seasons provided very similar results as the primary analysis (Figure A8).

**Discussion**

In this analysis that included 150,081 births from 58,945 mothers in population-based surveys throughout Bangladesh, living in flood prone areas was associated with an excess risk in infant mortality of 5.9 [2.8; 8.9] additional deaths per 1000 births compared to living in non-flood prone areas over the 30-year period, 1988–2017. Drawing on national-scale, high-resolution estimates of flood risk and population distribution, we estimate an excess of 170,136 [80,742; 256,645] infant deaths over the past 30 years were attributable to living in flood-prone areas in Bangladesh, with marked heterogeneity in attributable burden by subdistrict (Fig. 3) and heterogeneity in excess risk over time, with a risk difference of 9.2 [2.9; 15.5], 1.5 [2.6; 5.7] and 8.9 [3.7; 14.2] per 1000 births for the 1988–1997; 1998–2007 and 2008–2017 decades respectively (Fig. 2).

Over the study period, excess risk in infant mortality was reduced substantially between the 1988–1997 and 1998–2007 decades, despite massive flood events in 2004 and 2007, yet excess mortality risk increased during the final 10-year period of the analysis, 2008–2017. Although Bangladesh has made remarkable progress to reduce infant mortality rates over the past three decades (Fig. 1d), IPCC has
forecasted increased flood hazard in the country, and our analysis suggests that heightened mortality associated with living in flood prone areas could, in part, jeopardize these gains.

Our results are consistent with the literature that documents adverse health outcomes among populations exposed to flooding events but goes further by characterizing their evolution over time. Previous studies have reported increased mental health risk\(^{23,24}\), infectious diseases\(^{9,25,26}\) or birth outcomes\(^{10,27}\) and vulnerability\(^{28,29}\) following flooding events but seldom are the studies interested in long-term effects of such exposure and none focus on repeated exposure. The matched cohort analysis used here demonstrated a heightened infant mortality risk among Bangladesh’s flood prone areas over the past thirty years that does not appear to have been driven by any single flood event based on fine-scale estimates of temporal heterogeneity in excess risk (Figure A2). Our study was not designed to determine the exact causal mechanisms and further case studies should investigate why this population’s vulnerability changed over time and disentangle short- and long-term mechanisms. In particular, the absence of differences in excess risk between the rainy and dry seasons in the most recent 2008–2017 decade, whereas all floods keep happening during the rainy season because of monsoon, suggests lasting effects on populations health beyond the immediate flood events that may not have been at play in the earlier 1998 – 1997 decade when most of the excess risk was during the rainy season (Fig. 2).

Estimating causal effects of extreme weather events on child health presents methodologic challenges especially because exposure to such events cannot be randomized. Study designs that leverage the apparent randomness of extreme weather events as natural experiments require that health effects are localized close to the event in space and time and that high resolution data are available around spatial and temporal event boundaries. Furthermore, such studies typically focus on single events as such designs are not applicable to long-term effects related to floods.

Yet, studying the long-term impacts of such climate-sensitive exposures is important because the population affected by a single, extreme flood is likely to be consistently vulnerable to flooding and is thus more likely to have experienced the effects of previous floods. Many climate-related exposures beyond floods, such as droughts or temperature extremes, face the same challenge of consistently vulnerable populations that face repeated exposures such that health effects may not be limited to periods immediately adjacent to single events. In this study, we proposed a design to overcome such limitations and study long-term impacts. We anticipate that the approach developed here, which combines remotely sensed data with DHS to identify matched, counterfactual groups for populations that are consistently vulnerable climate-related exposures, will generalize to other countries and other extreme weather events such as heat waves, droughts, heavy precipitation, or wildfires.

Beyond the potential for residual, unmeasured confounding this analysis had some additional limitations. The Global flood database uses remote sensing to map water surfaces around flood events notified in the DFO database, itself based on emergency reports and news coverage. Cloud coverage impeding remote sensing and possible biases in the DFO database could limit the representability of the
seven country-wide flood events extracted from the Global flood database, resulting in measurement bias in the exposure. We addressed this issue by conducting sensitivity analyses where flood prone areas were defined more conservatively using higher thresholds on the percent number of days flooded, which provided similar results. In addition, to protect anonymity, GPS coordinates in the DHS data are provided with a random offset which could result in non-differential misclassification in the exposure, which if anything, would bias results towards the null. Another limitation from the DHS survey is that it records birth histories only from women younger than 49 years old, which means that infant mortality estimates for months closer to interviews may be more representative as earlier estimates miss data from women that were younger than 49 at the time but above 50 at the time of interview. Since DHS surveys were conducted every 3 to 4 years in Bangladesh, this is unlikely to be a major source of error but use of this general approach in other settings may need to restrict their sample of mothers to avoid this potential source of measurement error. Finally, there are several possible mechanisms that could link living in a flood prone area with infant mortality, including post-disaster mental health effects, environmental contamination leading to respiratory infections or diarrheal diseases or longer-term effects affecting nutrition and socio-economics resources within the flooded community. Since these mechanisms could vary geographically and over time, generalizing the effects beyond the present study requires a strong assumption of consistency in the treatment effect, which could be unlikely. This does not harm the validity of the present estimates but suggests that new research to clarify the mechanisms for how living in flood-prone areas increases infant mortality could further guide actionable interventions.

In conclusion, this study provided high-resolution estimates of infant mortality attributable to living in flood prone areas of Bangladesh and evaluated how effects changed over 30 years. The study provides a generalizable example for methods to study climate-related exposures and longer-term health effects and demonstrates the importance of measuring longer-term impacts in addition to acute impacts immediately following extreme events.

Declarations

Data Availability

All the data used is publicly available, but access needs to be requested to DHS data and therefore cannot be deposited on public repositories. All other data were deposited here: https://osf.io/vrfmz/ (DOI 10.17605/OSF.IO/VRFMZ). All code is available on Github at https://github.com/FrancoisRerolle/floods-bangladesh

Author contributions

FR, BA and TB designed the study. FR conducted the analysis. FR, BA and TB wrote the manuscript.

Acknowledgments
This work was supported in part by the National Institute of Allergy and Infectious Diseases (R01-AI158884, R01-AI166671, PI: Arnold).

**Competing interests**

No competing interests.

**References**

8. Skea, J., Shukla, P. & K\il\k\cs, \cSiir. Climate Change 2022: Mitigation of Climate Change. (2022).

Appendix

An Appendix (including Figures A1-A8 and Tables A1-A3) is not available with this version.

Figures
Figure 1

Map of the percent number of days flooded across seven country-wide flood events available in the Global Flood database (a) and its distribution among areas flooded on at least 1 days (b). Locations of clusters included in 1999, 2004, 2007, 2011, 2014 and 2018 DHS waves (c) and infant mortality rate in populations living in flood prone and non-flood-prone areas of Bangladesh, 1988-2017 (d). Points
represent seasonal averages with locally weighted regression smoothed lines (span = 0.1) super-imposed. May to October months defined the rainy seasons.

Figure 2

Risk differences in infant mortality (deaths per 1,000 births) comparing flood prone areas to non-flood prone areas of Bangladesh overall and stratified by decade and season.
Figure 3

Excess infant deaths attributable to living in flood prone areas between 1988 and 2017 at 1km$^2$ resolution (a) and excess infant mortality per 1000 births attributable to living in flood prone areas between 1988 and 2017 aggregated by subdistricts (zilas) (b & c). Figure A9 includes 95% confidence intervals on subdistrict-level estimates.