Title: The effect of rainfall on the incidence of cholera in Bangladesh.

Article Type: Original Article

Section/Category:

Corresponding Author: Dr. masahiro hashizume, MD PhD

Corresponding Author's Institution: Nagasaki University

First Author: Masahiro Hashizume, M.D. PhD

Order of Authors: Masahiro Hashizume, M.D. PhD; Ben Armstrong, PhD; Shakoor Hajat, PhD; Yukiko Wagatsuma, M.D. Dr. PH; A.S.G Faruque; Taiichi Hayashi; David A Sack

Abstract: Background: The incidence of cholera in Bangladesh shows clear seasonality, suggesting that weather factors could play a role in its epidemiology. We estimated the effects of rainfall on the incidence of cholera in Dhaka, Bangladesh.

Methods: We examined time-series patterns of the weekly number of hospital visits due to cholera in relation to weekly rainfall from 1996 to 2002. We used Poisson regression models, adjusted for seasonal variation, between-year variation, public holidays and temperature. The role of river level on the rainfall-cholera relationship was also examined by incorporating river-level terms into the models.

Results: The weekly number of cholera cases increased by 14% (95 % confidence interval = 10.1% - 18.9%) for each 10-millimeter increase above the threshold of 45 millimeters for the average rainfall, over lags 0−8 weeks. Conversely, the number of cholera cases increased by 24% (10.7% - 38.6%) for a 10-millimeter decrease below the same threshold of average rainfall, over lags 0−16 weeks. River level partly explained the association between high rainfall and the number of cholera cases.

Conclusions: The number of cholera cases increased with both high and low rainfall in the weeks preceding hospital visits. These results suggest that factors associated with river level are on the causal pathway between high rainfall and incidence of cholera.

ABSTRACT

Background: The incidence of cholera in Bangladesh shows clear seasonality, suggesting that weather factors could play a role in its epidemiology. We estimated the effects of rainfall on the incidence of cholera in Dhaka, Bangladesh.

Methods: We examined time-series patterns of the weekly number of hospital visits due to cholera in relation to weekly rainfall from 1996 to 2002. We used Poisson regression models, adjusted for seasonal variation, between-year variation, public holidays and temperature. The role of river level on the rainfall-cholera relationship was also examined by incorporating river-level terms into the models.

Results: The weekly number of cholera cases increased by 14% (95 % confidence interval = 10.1% - 18.9%) for each 10-millimeter increase above the threshold of 45 millimeters for the average rainfall, over lags 0−8 weeks. Conversely, the number of cholera cases increased by 24% (10.7% - 38.6%) for a 10-millimeter decrease below the same threshold of average rainfall, over lags 0−16 weeks. River level partly explained the association between high rainfall and the number of cholera cases.

Conclusions: The number of cholera cases increased with both high and low rainfall in the weeks preceding hospital visits. These results suggest that factors associated with river level are on the causal pathway between high rainfall and incidence of cholera.

Cholera remains a major public health problem in many areas of the world, particularly in Bangladesh, India, and countries in Africa and South America. *Vibrio Cholerae*, the bacterium that causes the disease, is known to inhabit rivers, estuaries and coastal waters.[1] Outbreaks of cholera can be initiated when *V. Cholerae* is present in drinking water in sufficient numbers to constitute an infective dose. Incidences of cholera are dynamic in endemic regions and show regular seasonal cycles. A bimodal annual distribution is observed in Bangladesh,[2] while single annual peaks can be seen elsewhere (eg, South America[3] and India[4]). This clear seasonality suggests that weather factors play a role, which is likely through multiple pathways. With growing concerns about global climate change, many studies have focused on the associations between climate variability and fluctuations in the incidence of cholera. There is considerable evidence for a role of El Niño-Southern Oscillation, a major source of interannual climate variability, in the interannual variation of endemic cholera in Bangladesh.[5-8] Sea surface temperature in the Bay of Bengal shows a bimodal annual cycle similar to the seasonal pattern of cholera in Dhaka, Bangladesh.[9] A role of climate factors related to water levels such as rainfall has also been invoked to explain seasonality of cholera since early times.[2] Floods and droughts can affect the concentration of the bacterium in the environment, survival of the bacterium (through effects on salinity,[10] pH or nutrient concentrations) human exposure to the pathogen, sanitary conditions, and susceptibility to disease.[2] However, there have been few studies that have quantified the impact of short-term climate variability on the incidence of cholera. This is of interest in its own right, as well as potentially clarifying of the pathways from climate to the seasonal epidemics of the disease.[2]

This study aimed to investigate the relationship between short-term variations in climate – particularly rainfall – and incidence of cholera in Dhaka, Bangladesh. To gain some insight into possible causal pathways linking rainfall to the occurence of cholera, the association between river levels and incidence of cholera was also investigated. Population factors that might affect vulnerability to the effect of rainfall on incidence of cholera were also investigated.

METHODS

Hospital surveillance

The primary outcome for this study was the weekly number of patients with cholera who visited the International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B) Dhaka Hospital. This hospital serves an urban population of approximately 10 million[11] and provides free treatment for more than 100,000 cases of diarrhea each year. Every 50th patient has been enrolled in the surveillance system since 1996. For each patient, stool was microbiologically examined to identify any enteric pathogens. We retrieved information on the date of hospital visit and the pathogens identified from each stool specimen, during a 7-year period (January 1996 to December 2002). Information on age, sex, socioeconomic status (educational level and roof structure of the house), and hygiene and sanitation practices (drinking water source, distance to the water source and type of toilet) for each patient was also abstracted to investigate potentially vulnerable groups. These were selected as potential markers of vulnerability because socioeconomic status[12,13] and hygiene and sanitation practices[14,15] have been identified as important determinants of cholera in Bangladesh. A patient was classified as a cholera case when *V. cholerae* of the

serogroup O1 or O139 was identified from the stool specimen, regardless of the presence of other pathogens.

Meteorologic and river level data

We obtained data on daily rainfall and on maximum and minimum temperature in Dhaka from the Bangladesh Meteorological Department. The daily river level of Brigonga River at Mill Barrack in Dhaka was recorded by the Bangladesh Water Development Board. The weekly mean for maximum temperature and maximum river level and the total weekly rainfall were calculated from the daily records.

Statistical analysis

Statistical methods are summarized here and described in detail in the appendix (available with the online version of this article). We examined the relationship of the number of cholera cases per week with rainfall, river level and temperature, using generalized linear Poisson regression models allowing for overdispersion.[16] To account for seasonality in the incidences of cholera that are not directly due to the weather, we included in the model Fourier terms up to the sixth harmonic (a time span of approximately two months). Indicator variables for the years of the study were incorporated into the model to allow for long-term trends and other variations among the years. An indicator variable for public holidays was incorporated into the model to control bias in the event that holidays affected access to the hospital, as was suggested in a previous time-series study in the UK.[17] To allow for autocorrelations, an autoregressive term at order one was incorporated into the models.[18]

Models for rainfall

4

From exploratory analyses, existing literature, and considerations of interpretational difficulty with very long lags, we considered lags of up to 16 weeks for the delay in the effect of rainfall on the number of cholera cases. In our initial analyses, we fit a natural cubic spline (3 df)[19] to (a) the average rainfall over lags 0−16, and (b) the average over 0−8 weeks and 9−16 weeks, as separate splines that were simultaneously included in the model. We also included temperature as a natural cubic spline (3 df) in all models to control confounding , with lag 0−4 weeks, as suggested by previous studies.[20,21] Because initial analyses of rainfall suggested a broad "U" shape, we then fitted a double-thresholds model, comprising linear terms for rainfall above and below "high" and "low" thresholds respectively, with no association, (ie flat bottom of "U") in between.[22] Guided by the spline analyses, the low and high rainfall terms were based on the 0−16 and 0−8 week averages, respectively. The thresholds were estimated by maximum likelihood. An increase or decrease in the number of cholera cases that were associated with a 10-millimeter increase or decrease in a given measure of rainfall above or below the thresholds, estimated as coefficients from the regression model, was reported as percentage change.

Using the simple thresholds model, we then examined lag effects in more detail, by fitting linear unconstrained distributed lag models, comprising terms for low and high rainfall at each lag up to the previous 16 weeks.[22] The simple linear-thresholds model also allowed investigation of the modification of rainfall effects by patient characteristics (for example socioeconomic status), by fitting the models separately to incidence series according to their characteristics.

Models for river level

5

Because we hypothesised that river level is one of the causal pathways between rainfall and incidence of cholera, river level was not included in the initial analyses; the relationship between river level and the number of cholera cases was examined in later analyses. Specifically, we fit a natural cubic spline (3 df) to the average river level over lags 0−4 weeks and incorporated this into a model comprising the same confounders included in the model for rainfall. (The lag period was set at 0−4 weeks because preliminary analyses showed rainfall to predict river level most strongly after a delay of four weeks.) Finally, the relationship between rainfall and the number of cholera cases was re-estimated, adjusted for river level by incorporating the river-level terms into the model for rainfall, in order to clarify what component of the rainfall-cholera association was through factors associated with elevated river level.

We carried out several sensitivity analyses, including repeating the same procedures using weekly average for daily minimum (instead of maximum) river level or temperature. All statistical analyses were carried out using Stata 9.0 (Stata Corporation, College Station, Texas).

RESULTS

Between 1996 and 2002, 3807 of the sampled hospital visits were due to cholera. The mean ± SD weekly number of cholera patients was 10.5 ± 8.8. Half of the cholera patients were children younger than 15 years (Table 1). More than 80% lived in household with non-concrete roof, and approximately half of the patients used sanitary toilet and tap water for drinking. These figures are comparable with the

general population in Dhaka city (73% for non-concrete roof, 46% for tap water and 55% for sanitary toilet).[23] Seasonal variation in hospital visits due to cholera showed the well-known bimodal seasonality – peaking before the monsoon (high rainfall period) and at the end of the monsoon (Figure 1), with a trough in the middle of the monsoon. There was thus little obvious tracking of seasonal patterns of incidence of cholera with rainfall.

Relationship with rainfall

   Figure 2 shows the relationship between the number of cholera cases and rainfall, before and after adjustment for seasonal variation, between-year variation, public holiday and temperature. Before adjustment, there appears to be an increase in cholera cases with higher rainfalls at lag 0−8 (Fig. 2A) and 0−16 weeks (Fig. 2B). After adjustment, there is a positive slope with high rainfall for a lag of 0−8 weeks (Fig. 2C); this is true of both high and low rainfall for a lag of 0−16 weeks (Fig. 2D). For the linear-thresholds model, the maximum likelihood estimate of the threshold for high rainfall and low rainfall was 45 millimeters (95% CI = 36 - 53) for the average rainfall over a lag of 0−8 and 0−16 weeks, respectively. The total number of "high-rainfall" weeks, according to this definition, was less than that of low-rainfall weeks (152 vs 208). For a 10-millimeter increase above the rainfall threshold (45 mm), the number of cholera cases increased by 14% (95% CI = 10.1% - 18.9%) after controlling for seasonal and between-year differences, public holidays and temperature. For a 10 millimeter-decrease below the rainfall threshold, the number of cholera cases increased by 24% (11.3% - 38.9%).

Figure 3 shows the effects of high or low rainfall (the slopes of the thresholds model) at different lags using the distributed lag model. The "high-rainfall" effect was observed at shorter lags (1 to 5 weeks) (Fig. 3A). In contrast, the "low-rainfall" effects persisted for longer, (Fig. 3B) with increased risk through most of lags 1 to 16 weeks. These findings confirm our original choice of lags of 0−8 weeks for high rainfall and 0−16 weeks for low rainfall. For both "high rainfall" and "low rainfall," the risk of a hospital visit due to cholera was similarly elevated for all of the subgroups examined: i.e. there was no evidence for the difference in the effects of "high rainfall" and "low rainfall" by socioeconomic status, hygiene and sanitation practice and other characteristics examined, and thus no evidence for heightened vulnerability to these factors in the sub-groups examined.

Relationship with river level

Figure 4 shows the relationships between the number of cholera cases and river level before and after adjustment for seasonal variation, between-year variation, public holiday and temperature. Before adjustment, there appears to be an increase in cholera cases when the river level is higher than 4−5 meters (Fig. 4A). This pattern was similar after adjusting for seasonal variation, between-year variation, public holiday and temperature (Fig. 4B). The estimated threshold was 4.0 meters (95% CI = 3.9 - 4.2): for a 1 meter increase above the threshold, the number of cholera cases increased by 137% (109% - 170%). The estimate was similar when daily minimum river level was used instead of maximum river level.

Relationship with rainfall adjusted for river level

The risk-response relationships between rainfall and the number of cholera cases after adjusting for river level, seasonal variation, between-year variation, public holidays and temperature are shown in **e**Figure 1 (available with the online version of the paper). The positive slope shown for a lag of 0−8 weeks declined after adjustment for river level, but was not completely eliminated (**e**Fig. 1A effect of "high rainfall") . The positive slope shown for a lag of 0−16 weeks substantially declined after adjusting for river level (**e**Fig. 1B effect of "low rainfall"). After adjusting for river level, we estimated that the effect of high rainfall was decreased to 8.3% (3.1% - 13.8%) by using the same threshold as that used before adjusting for river level. Nevertheless, the evidence for an association remained very strong. In addition, the effect of "low rainfall" decreased to 12.4% (0.2% - 26.2%) after adjusting for river level.

Relationship with temperature

Although temperature was not the focus of this study, we considered the broad form of the relationship between number of cholera cases and temperature. Figure 5 shows the relationship before and after adjustment for seasonal variation, between-year variation, public holiday and rainfall. Before adjustment, there appears to be an increase in the number of cholera cases with higher temperatures (Fig. 5A), while after adjustment, the positive relationship with temperature levelled off at approximately 30°C (Fig. 5B). The relationship was largely unchanged when daily minimum temperature was used instead of maximum temperature (results not shown).

DISCUSSION

High rainfall was associated with increased hospital visits due to cholera over the next eight weeks. River level, or factors associated with it, was found to be an important part of the causal link between high rainfall and the incidence of cholera. Low rainfall and high temperature were also associated with increases in cholera cases.

Most previous studies of temporal variation in cholera and weather in Bangladesh have sought to explain inter-annual or seasonal patterns.[5-9,24] In contrast, this study looked at how weather explained short-term associations after controlling for seasonal and inter-annual patterns. Thus our results draw on different information than the previous studies, and are complimentary to them. Our results are not subject to confounding bias by factors that might explain inter-annual or seasonal patterns, although our results are subject to other factors that cause departures from typical seasonal patterns.

We considered whether the short-term associations we found can explain the well-known bimodal seasonal pattern in cholera in Bangladesh. Broadly, the high-rainfall effect over lags 0−8 weeks explains some but not all of the second peak, and the low rainfall effect over lags 0−16 weeks explains some but not all of the first peak. However, even after adjusting for these rainfall effects, a bimodal pattern remains. Thus the associations we report here appear to be in addition to whatever factors explain the typical seasonal pattern, rather than simply a component of it. Our results are therefore not contradictory to those proposing alternative explanations of the typical seasonal pattern.[5,9] The same is probably true of most factors that explain inter-annual variation.

The finding of the adverse effect of high rainfall on the incidence of cholera is nevertheless in contrast with a previous suggestion that high rainfall reduces the occurrence of cholera during the monsoon.[2] The possible mechanisms of the decrease in cholera during the monsoon were proposed as the dilution in the concentration of pathogens in aquatic environments, and the reduction of water salinity to below the optimum requirements for pathogen survival as the result of increased river discharge due to monsoon rains.[2] However, this seasonal coincidence may be a consequence of other unmeasured environmental or behavioral factors, which are more (or less) often observed in the monsoon season. However, identification of the specific attributable factors is not possible because many environmental variables of interest show regular seasonal cycles. The current study aimed to examine the effect of rainfall on the incidence of cholera controlling for such unmeasured seasonal factors.

Although this study did not directly investigate the mechanisms of the association between high rainfall and the incidence of cholera, the results point toward several hypotheses. Heavy rain leads to flooding, which may affect water and sanitation systems and thereby promote the intake of contaminated water. The presence of a threshold in an exposure-response curve for both rainfall and river level against the number of cholera cases, may support a hypothesis for the pathway: high rainfall → flooding (overflowing water) → exposure to contaminated water with *V. cholerae,* as flooding is likely to happen only after certain amount of rainfall. Ingestion of a few copepods, which carry a high concentration of *V. cholerae*, can initiate an infection[25] and this occurs more frequently when exposed to untreated water during flooding. However, we found no evidence for the modification of the rainfall effect by drinking

11

water source. Although drinking water contamination is not expected to be as pronounced in Dhaka (where the main drinking water sources are tube-wells and taps, rather than surface water), surface water might nevertheless have some impact as it is used for washing and bathing in Dhaka. Investigations of more detailed pathways of the rainfall-cholera relationship, particularly the role of drinking water quality, would be of interest.

The threshold of river level (4.0 m) for the increase in the number of cholera cases was below the flood-danger level of the major river (6.0 m). This suggests that the observed increase in cholera cases occurred in the absence of overflowing river water. Although we measured water levels only of the major river, there are a number of smaller rivers, ponds and lakes scattered among the communities in Dhaka that perhaps flood even if the river does not.

Rainfall can affect not only human exposure to the pathogen and sanitary conditions, but also the growth of the pathogen and its survival, through the effect of salinity,[10] pH or nutrient concentrations in aquatic environments.[25] High rainfall increases the levels of insoluble iron, which improves the survival of *V. cholerae* in aquatic environments.[25] Moderate levels of iron also increase expression of the cholera toxin.[25] It has also been suggested that high rainfall might wash away the vibriophages that prey on *V. cholerae* in water, leading to the epidemics of cholera.[26] However, the time lag between when rain falls and when *V. cholerae* have multiplied enough to successfully infect a host is not clear; thus, it is not obvious whether the effect of high rainfall on incidence of cholera found in this study can be linked to these pathways. In order to provide more direct evidence for the pathways, it will be

necessary to quantify the level of *V. cholerae* in aquatic environments at the same time as measuring rainfall and river levels.

This study also found that low rainfall might explain any increase in cholera cases. Low rainfall might influence the incidence of cholera through changes in water supply and hygiene behaviors.[2,27] Although severe scarcity of drinking water would not be expected in Dhaka, surface water might nevertheless have some impact; a certain proportion of people of Dhaka rely on surface water for washing and bathing, which could consequently increase the risk of contamination and exposure to cholera during low rainfall – in part because the likelihood of multiple uses in a body of water may increase when the amount of surface water is scarce.[25]

Our finding of a positive relationship between ambient temperature and number of cholera cases is broadly in accordance with previous studies.[20,28] This relationship might be due to the promotion of growth and multiplication of *V. cholerae*, which directly influences the abundance and toxicity of *V. cholerae* in aquatic environments; alternatively, it might be due to an indirect influence on pH levels as a result of increased aquatic plant or algal growth when the temperature is warm.[25] It has been suggested that the high correlation between surface temperature in the Bay of Bengal and the outbreak of cholera is due to warm waters along the coast, coupled with plankton blooms driven by warm ocean temperatures, which are favorable conditions for multiplication of *V. cholerae*.[9,25]

There may be concerns about a possible effect of immunity because the population at risk changes over time. Infection with *V. cholerae* O1 or O139 increases immunity

to reinfection, but it is not known for how long.[29] However, as population immune status is not likely to change quickly, it seems unlikely that this would obscure the short-term (within 16 weeks) dependence of cholera on the factors investigated in this study.

The observed association between rainfall and the number of cholera cases may be specific to places characterized by low land with bodies of water that are vulnerable to flooding. The association may also be greatly dependent on the degree of hygiene and sanitation in an area. The large estuary of Bangladesh provides a suitable environmental reservoir for *V. cholerae* and these topographic characteristics may be unique to Bangladesh. Therefore, our findings may not pertain to other places.

The World Health Organization listed cholera as a candidate for developing early warning systems because of cholera's large global burden and climate sensitivity.[30] The results of this study can contribute to the base of information for predicting epidemics, and therefore has the potential to improve disease control. Expected increases in temperature, changes in precipitation patterns, and increased flooding in Bangladesh[31] give particular focus to the results, although the short-term associations reported here should not be directly extrapolated to changes in climate over decades.

In conclusion, this study found that the number of cholera cases increased with high and low rainfall in the weeks preceding hospital visits, even after potential confounding by seasonal variation, between-year variation, public holidays and temperature are taken into consideration. The association between rainfall and incidence of cholera may be explained, in part, by the effect of rainfall on river levels.

REFERENCES

1.   Colwell RR, Kaper J, Joseph SW. Vibrio cholerae, Vibrio parahaemolyticus,
      and other vibrios: occurrence and distribution in Chesapeake Bay. *Science.*
      1977;198:394-396.

2.   Pascual M, Bouma MJ, Dobson AP. Cholera and climate: revisiting the
      quantitative evidence. *Microbes Infect.* 2002;4:237-245.

3.   Codeco CT. Endemic and epidemic dynamics of cholera: the role of the
      aquatic reservoir. *BMC Infect Dis.* 2001;1:1.

4.   Jesudason MV, Balaji V, Mukundan U, Thomson CJ. Ecological study of
      Vibrio cholerae in Vellore. *Epidemiol Infect.* 2000;124:201-206.

5.   Bouma M, Pascual M. Seasonal and interannual cycles of endemic cholera in
      Bengal 1891-1940 in relation to climate and geography. *Hydrobiologia.*
      2001;460:147-156.

6.   Colwell RR. Global climate and infectious disease: the cholera paradigm.
      Science. 1996;274:2025-2031.

7.   Pascual M, Rodo X, Ellner SP, Colwell R, Bouma MJ. Cholera dynamics and
      El Nino-Southern Oscillation. *Science.* 2000;289:1766-1769.

8.   Rodo X, Pascual M, Fuchs G, Faruque AS. ENSO and cholera: a
      nonstationary link related to climate change? *Proc Natl Acad Sci U S A.*
      2002;99:12901-12906.

9.   Lobitz B, Beck L, Huq A, et al. Climate and infectious disease: use of remote
      sensing for detection of Vibrio cholerae by indirect measurement. *Proc Natl
      Acad Sci U S A.* 2000;97:1438-1443.

10.   Miller CJ, Drasar BS, Feachem RG. Cholera and estuarine salinity in Calcutta
      and London. *Lancet.* 1982;1:1216-1218.

11. Bangladesh Bureau of Statistics. *Statistical Pocketbook of Bangladesh 2001*. Dhaka, Bangladesh: Statistics Division, Ministry of Planning. Government of the People's Republic of Bangladesh, 2002.

12. Emch M. Diarrheal disease risk in Matlab, Bangladesh. *Soc Sci Med.* 1999;49:519-530.

13. Ali M, Emch M, Donnay JP, Yunus M, Sack RB. The spatial epidemiology of cholera in an endemic area of Bangladesh. *Soc Sci Med.* 2002;55:1015-1024.

14. Hughes JM, Boyce JM, Levine RJ, et al. Epidemiology of eltor cholera in rural Bangladesh: importance of surface water in transmission. *Bull World Health Organ.* 1982;60:395-404.

15. Sommer A, Woodward WE. The influence of protected water supplies on the spread of classical-Inaba and El Tor-Ogawa cholera in rural East Bengal. *Lancet.* 1972;2:985-987.

16. McCullagh P, Nelder J. *Generalized linear models*. London: Chapman and Hall, 1989.

17. Bentham G, Langford IH. Environmental temperatures and the incidence of food poisoning in England and Wales. *Int J Biometeorol.* 2001;45:22-26.

18. Brumback B, Burge HA, Neas LM, et al. Transitional regression models, with application to environmental time series. *J Am Stat Assoc.* 2000;95:16-27.

19. Durrleman S, Simon R. Flexible regression models with cubic splines. *Stat Med.* 1989;8:551-561.

20. Speelmon EC, Checkley W, Gilman RH, et al. Cholera incidence and El Nino-related higher ambient temperature. *JAMA.* 2000;283:3072-3074.

21. Checkley W, Epstein LD, Gilman RH, et al. Effect of El Nino and ambient temperature on hospital admissions for diarrheal diseases in Peruvian children. *Lancet.* 2000;355:442-450.

22. Armstrong B. Models for the relationship between ambient temperature and daily mortality. *Epidemiology.* 2006;17:624-631.

23. Bangladesh Bureau of Statistics. *Population Census 1991, National Series, Vol. III, Urban area report of Bangladesh*. Dhaka: Bangladesh Bureau of Statistics., 1991.

24. Koelle K, Rodo X, Pascual M, Yunus M, Mostafa G. Refractory periods and climate forcing in cholera dynamics. *Nature.* 2005;436:696-700.

25. Lipp EK, Huq A, Colwell RR. Effects of global climate on infectious disease: the cholera model. *Clin Microbiol Rev.* 2002;15:757-770.

26. Faruque SM, Naser IB, Islam MJ, et al. Seasonal epidemics of cholera inversely correlate with the prevalence of environmental cholera phages. *Proc Natl Acad Sci U S A.* 2005;102:1702-1707.

27. Rogers. L. The periodicity of cholera. Letter to the Lancet. *Lancet.* 1928;June 20.

28. Huq A, Sack RB, Nizam A, et al. Critical factors influencing the occurrence of Vibrio cholerae in the environment of Bangladesh. *Appl Environ Microbiol.* 2005;71:4645-4654.

29. Chin. J. *Control of Communicable Diseases Manual*. 17th ed. Washington, DC: American Public Health Association, 2000.

30. World Health Organization. *Using climate to predict infectious disease epidemics*. Geneva: WHO, 2005.

31.   Watson R, Zinyowera M, Moss R. *The Regional Impacts of Climate Change: An Assessment of Vulnerability. Special Report of IPCC Working group II.* Cambridge: Cambridge University Press, 1997.

FIGURE LEGENDS

FIGURE 1. Seasonal variation in the number of cholera cases per week and meteorological and river level data in Dhaka, 1996−2002.

FIGURE 2. Relationship between the number of cholera cases and rainfall (shown as a 3 d.f. natural cubic spline) before (A, average rainfall over lags of 0−8 and B, 0−16 weeks) and after (C, average rainfall over lags of 0−8 and D, 0−16 weeks) adjustment for seasonal variation, between-year variations, public holidays and temperature. RR represents the relative risk of cholera (scaled against the mean weekly number of cholera cases). The center line in each graph shows the estimated spline curve, and the upper and lower lines represent the 95% confidence limits.

FIGURE 3. Percent change (and 95% CIs) in number of cholera cases for "high" rainfall (A, per 10 mm increase above threshold of 45 mm) and "low" rainfall (B, per 10 mm decrease below threshold of 45 mm) at each lag (unconstrained distributed lag models).

FIGURE 4. Relationship between the number of cholera cases and the average river level over a lag of 0−4 weeks (shown as a 3 d.f. natural cubic spline), before (A) and after (B) adjustment for seasonal variation, between-year variation, public holiday and temperature. RR is the relative risk of cholera (scaled against the mean weekly number of cholera cases). The center line in each graph shows the estimated spline curve, and the upper and lower lines represent the 95% confidence limits.

FIGURE 5. Relationship between the number of cholera cases and the average temperature over a lag of 0−4 weeks (shown as a 3 d.f. natural cubic spline) before (A) and after (B) adjustment for seasonal variation, between-year variation, public holidays and rainfall. RR represents the relative risk of cholera (scaled against the mean weekly number of cases). The center line in each graph shows the estimated spline curve, and the upper and lower lines represent the 95% confidence limits.

Figure %

Figure &

A

B

C

Average rainfall (Lag 0-8 weeks) (mm)

D

Average rainfall (Lag 0-16 weeks) (mm)

Figure ' A

B

Figure (

A

B

Average river level (Lag 0-4 weeks) (m)

Figure 5

A

B

Average temperature (Lag 0-4 weeks) (° C)
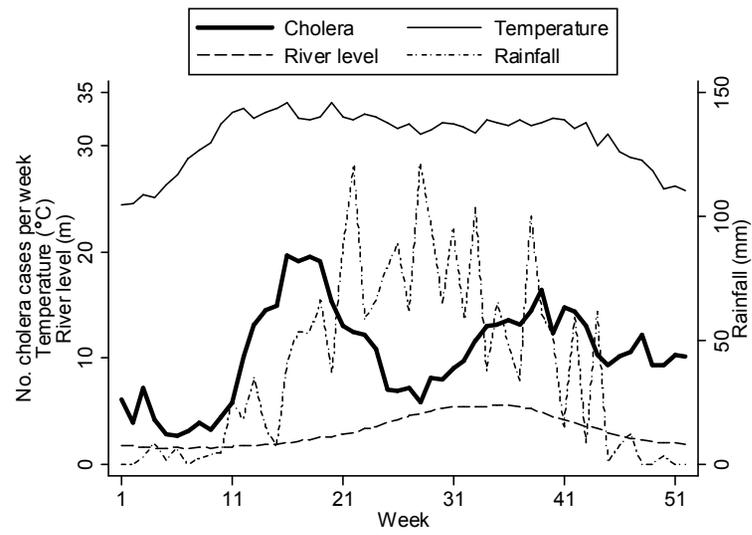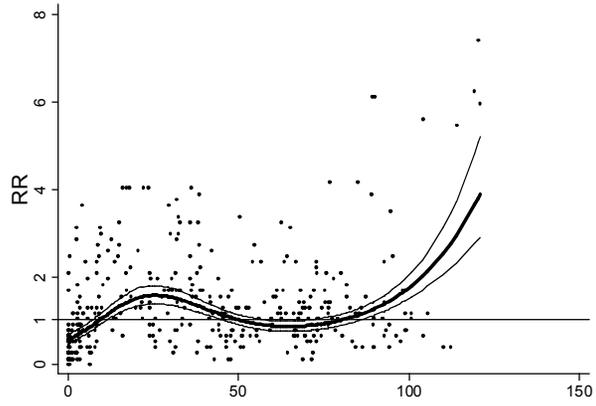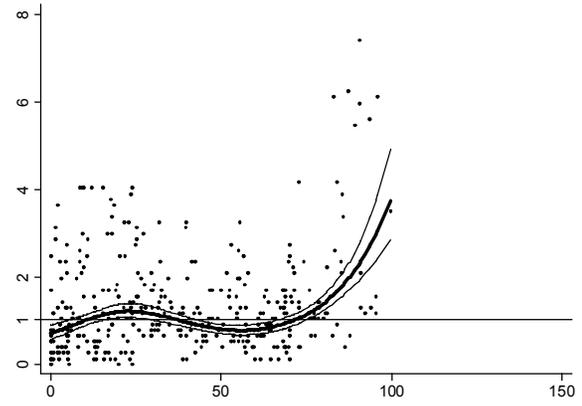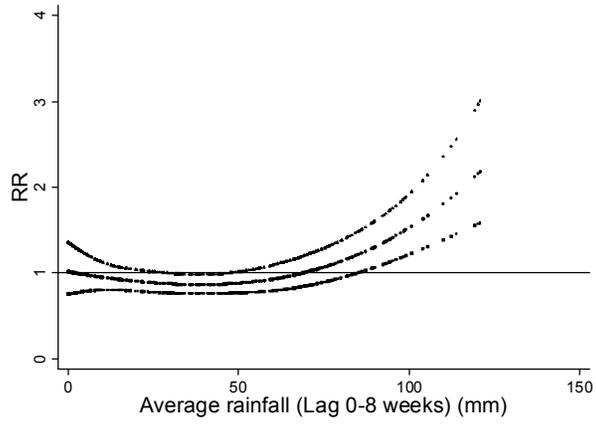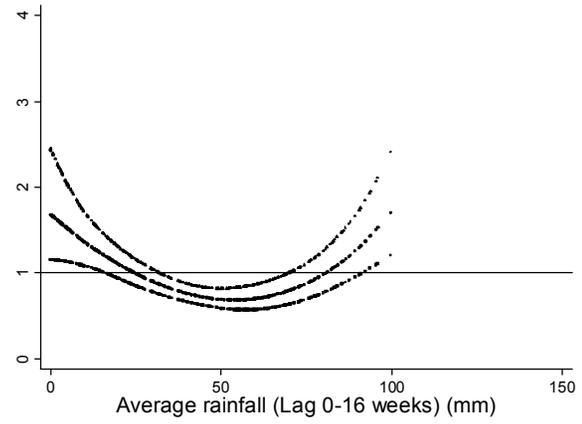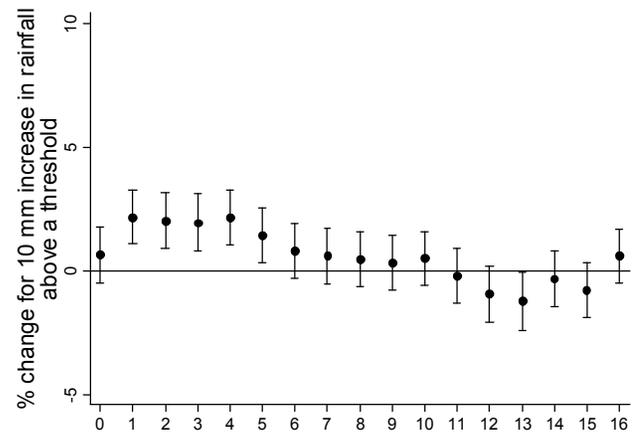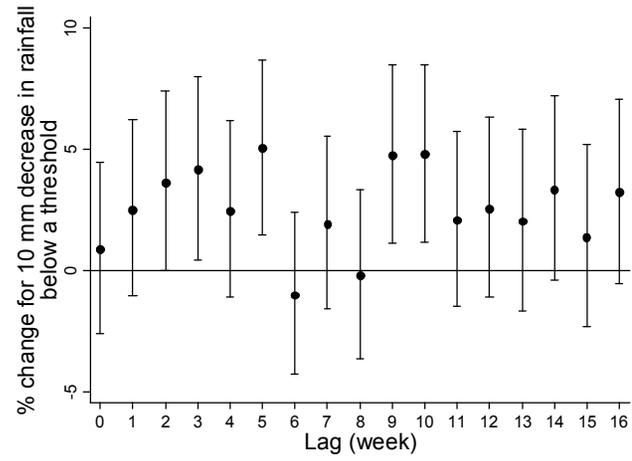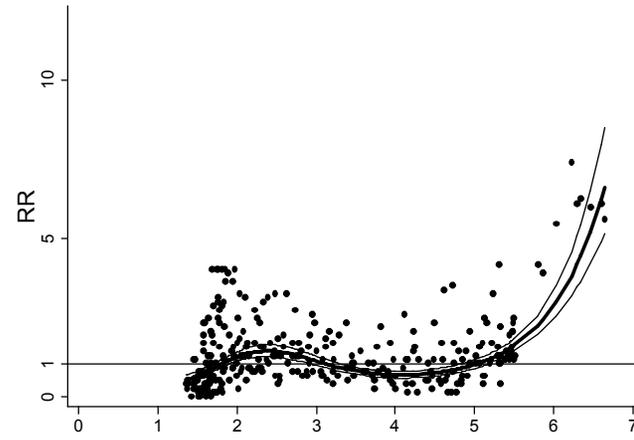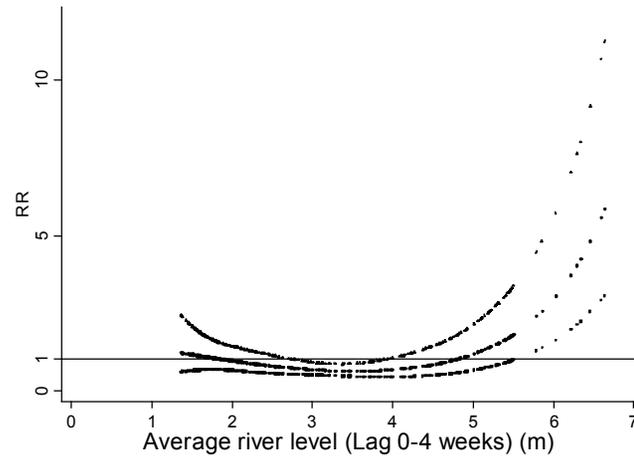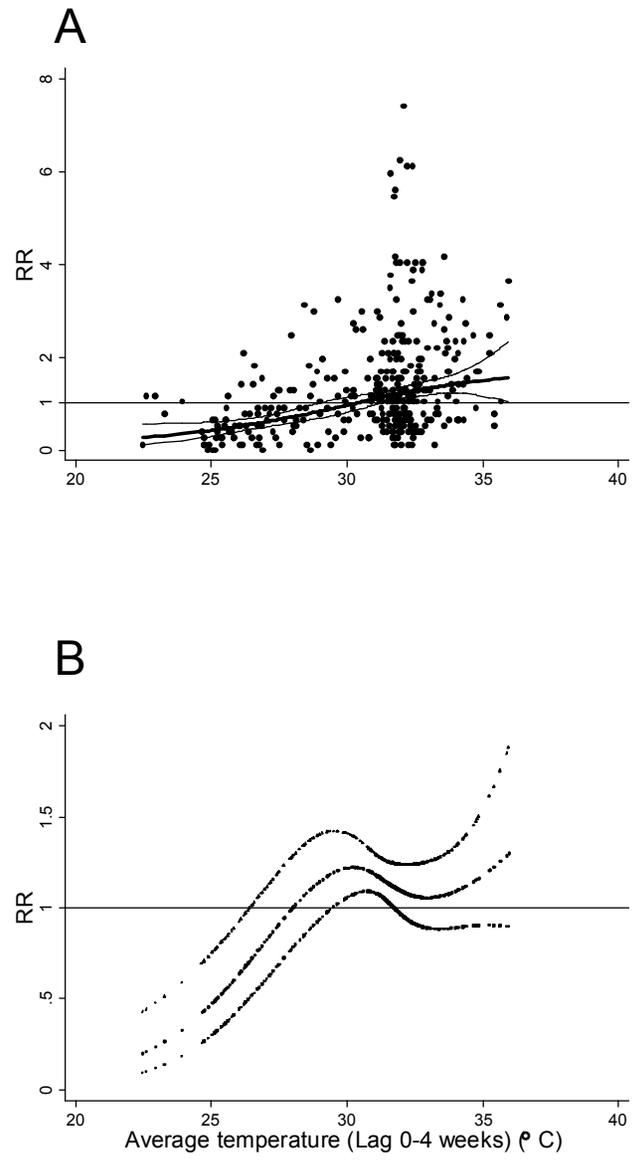
Table 1: Characteristics of cholera cases in ICDDR,B Dhaka Hospital (1996-2002).

		No.	
		Cases	(%)
		(n=3807)	
Characteristics			
Sex			
	Female	1733	(45.5)
	Male	2074	(54.5)
Age (years)			
	<15	1914	(50.3)
	15-29	979	(25.7)
	30+	914	(24.0)
Socioeconomic status			
Education*			
	None	2222	(58.4)
	Informal or <6 yrs	624	(16.4)
	6+ yrs	948	(24.9)
Roof structure			
	Not concrete	3181	(83.6)
	Concrete	626	(16.4)
Hygiene and sanitation			
Drinking water source			
	Tube well	1749	(45.9)
	Tap water	2045	(53.7)
Distance to water source (m)			
	> 5	2610	(68.6)
	≤ 5	1189	(31.2)
Type of toilet†			
	Nonsanitary	1879	(49.4)
	Sanitary	1928	(50.6)

*Mother's educational level for patients younger than 15 years and patient's educational level for those 15 years or older.

†Nonsanitary includes dug hole, open pit, hanging and no fixed place. Sanitary includes sanitary and semi-sanitary toilets.

Models for rainfall

Terminology:

The variables " rain" and "temp" indicate mean weekly amount of rainfall and average weekly temperature in each lag, respectively. NS indicates a natural cubic spline function. Fourier represents Fourier (trigonometric) terms. i.year represents indicator variables of year. i.holiday represents an indicator variable for weeks that include public holidays. $(x)^+ = x$ if $x > 0$, otherwise $= 0$.

Confounder terms (in all models for rainfall and river level):

(confounders) $= \alpha + time(Fourier, 6\ harmonics/year) + i.year + i.holiday + NS(temp_{0-4}, 3\ df)$

Model 1: spline for rainfall over lags 0−16 weeks (Fig. 2D):

$log[E(Y)] = NS(rain_{0-16}, 3\ df) + (confounders)$

Model 2: splines for rainfall over lags 0−8 and 9−16 weeks (Fig. 2C):

$log[E(Y)] = NS(rain_{0-8}, 3\ df) + NS(rain_{9-16}, 3\ df) + (confounders)$

Model 3: double-threshold model (Text slope estimates):

$log[E(Y)] = \beta_{low}(low\_threshold - rain_{0-16})^+ + \beta_{high}(rain_{0-8}, -high\_threshold)^+ + (confounders)$

The choice of thresholds was based on maximum likelihood estimation for the rainfall over a grid of all possible integer values within a range indicated on the rainfall-cholera graphs, constrained for interpretability so that $r_l = r_h$ where unconstrained estimates gave $r_l > r_h$. Likelihood profile confidence intervals (CIs) for the threshold were calculated as the thresholds for which deviance of the model was 3.84 more than the minimum.

Model 4: distributed lag model for high and low rainfall (Fig. 3)

$log[E(Y)] = \Sigma \beta_{low,l}(low\_threshold\text{-}rain_{i\text{-}l})^+ + \beta_{high}(rain_{i\text{-}l}\text{-}high\_threshold)^+ +$

$(confounders)$

Models for river level

Model 5 (Fig. 4B):

$Log[E(Y)] = \alpha + NS(river\ level_{0\text{-}4},\ 3\ df) + (confounders)$

Model 6 (**e**Fig. 1B)

$Log[E(Y)] = (Model\ 2) + NS(river\ level_{0\text{-}4},\ 3\ df)$

Models for temperature

Confounder terms:

$(confounders) = \alpha + time(Fourier,\ 6\ harmonics/year) + i.year + i.holiday + NS(rain_{0\text{-}8},$

$3\ df) + NS(rain_{9\text{-}16},\ 3\ df)$

Model 7: spline for temperature over lags 0−4 weeks (Fig. 5B):

$log[E(Y)] = NS(temp_{0\text{-}4},\ 3\ df) + (confounders)$

Test of heterogeneity

The heterogeneity of coefficients across Poisson regressions carried out on independent

sub-sets in each category of interest (table 1) was tested by using the chi-squared statistic

proposed by DerSimonian and Laird in the context of meta-analysis.[32]

Reference List

32.   DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials

1986;7:177-88.
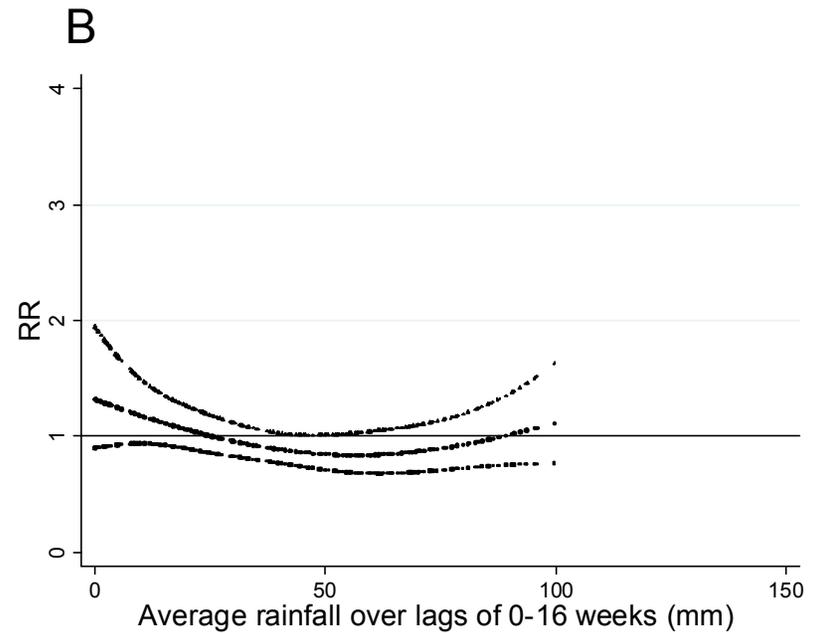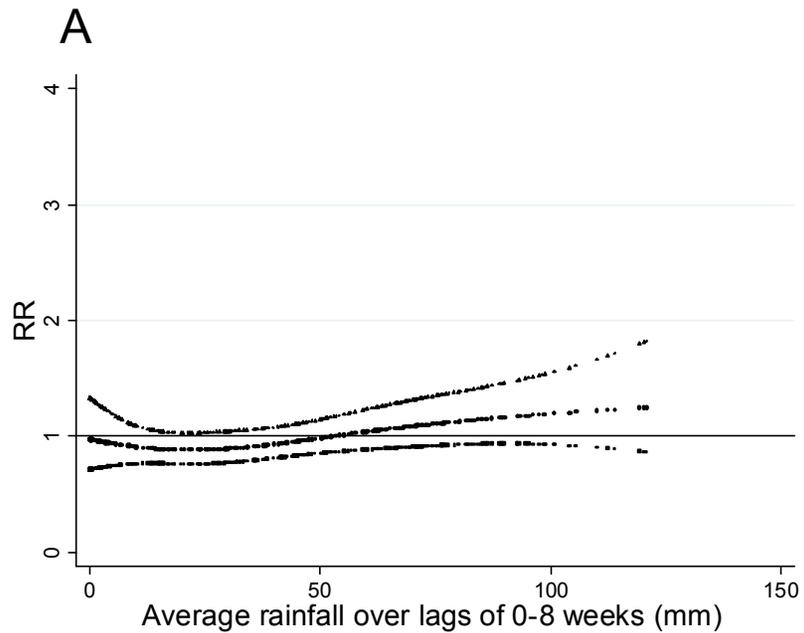
eFigure 1. Relationship between the number of cholera cases and rainfall (shown as a 3 d.f. natural cubic spline) over a lag of 0-8 (A) or 0-16 weeks (B), after adjustment for river level, seasonal variation, between-year variation, public holidays and temperature. RR represents the relative risk of cholera (scaled against the mean weekly number of cholera cases). The center line in each graph shows the estimated spline curve, and the upper and lower lines represent the 95% confidence limits.