Time Series Analysis of Cholera in Matlab, Bangladesh, during 1988-2001

Mohammad Ali¹, Deok Ryun Kim¹, Mohammad Yunus², Michael Emch³

¹International Vaccine Institute, SNU Research Park, San 4-8 Nakseongdae-dong, Gwanak-gu, Seoul, Korea; ²icddr,b, GPO Box 128, Dhaka 1000, Bangladesh; ³Department of Geography, University of North Carolina-Chapel Hill, Chapel Hill, NC 27599, USA

ABSTRACT

The study examined the impact of \textit{in-situ} climatic and marine environmental variability on cholera incidence in an endemic area of Bangladesh and developed a forecasting model for understanding the magnitude of incidence. Diarrhoea surveillance data collected between 1988 and 2001 were obtained from a field research site in Matlab, Bangladesh. Cholera cases were defined as \textit{Vibrio cholerae} O1 isolated from faecal specimens of patients who sought care at treatment centres serving the Matlab population. Cholera incidence for 168 months was correlated with remotely-sensed sea-surface temperature (SST) and \textit{in-situ} environmental data, including rainfall and ambient temperature. A seasonal autoregressive integrated moving average (SARIMA) model was used for determining the impact of climatic and environmental variability on cholera incidence and evaluating the ability of the model to forecast the magnitude of cholera. There were 4,157 cholera cases during the study period, with an average of 1.4 cases per 1,000 people. Since monthly cholera cases varied significantly by month, it was necessary to stabilize the variance of cholera incidence by computing the natural logarithm to conduct the analysis. The SARIMA model shows temporal clustering of cholera at one- and 12-month lags. There was a 6\% increase in cholera incidence with a minimum temperature increase of one degree celsius in the current month. For increase of SST by one degree celsius, there was a 25\% increase in the cholera incidence at current month and 18\% increase in the cholera incidence at two months. Rainfall did not influence to cause variation in cholera incidence during the study period. The model forecast the fluctuation of cholera incidence in Matlab reasonably well (Root mean square error, RMSE: 0.108). Thus, the ambient and sea-surface temperature-based model could be used in forecasting cholera outbreaks in Matlab.

Key words: Cholera; Climate change; Time series analysis; Matlab; Bangladesh

INTRODUCTION

Cholera is an acute intestinal disease caused by the bacterium \textit{Vibrio cholerae}. Although knowledge about the epidemiology and ecology of cholera has increased during the last two decades, cholera remains a serious problem in many areas of the world. Cholera appears to be influenced by climatic changes (1-3). In some endemic areas, cholera outbreaks have predictable seasonal patterns. Until recently, the reservoirs or sites for survival and multiplication of \textit{V. cholerae} during inter-epidemic periods were unknown. Recent studies have provided more satisfactory explanations on how seasonality and endemicity of cholera are maintained, providing clues about inter-annual variability (4). Lobitz and others hypothesized that rise in the local sea-surface temperature influences the growth of phytoplankton concentrations, and an increase in sea-surface height increases human-\textit{Vibrio} contact by transporting the bacteria into inland waters through tidal intrusion of plankton (5).

Lobitz \textit{et al.} (5) linked remotely-sensed marine environmental data with cholera incidence in Bangladesh and found that sea-surface temperature (SST) was positively associated with cholera cases. They hypothesized that an increase of SST results in replication of phytoplankton populations, which are directly associated with the increase in \textit{V. cholerae} bacteria and are linked spatially and temporally to zooplankton populations. Indirect measurement of ocean chlorophyll concentration (OCC) is also
Huq et al. (8) found significant correlations of water temperature, water depth, rainfall, conductivity, and copepod counts with the occurrence of cholera from the data on four rural areas of Bangladesh. However, the environmental drivers for the occurrence of cholera were not same in all those four areas. The lag periods between increases or decreases in temperature and salinity and occurrence of cholera correlate with biological parameters, e.g. plankton population blooms. Hashizume et al. (9) found that the number of cholera cases increased with both high and low rainfall in the weeks preceding hospital visits in Dhaka, Bangladesh. Lower temperature predicted a lower incidence of cholera in the first 15 weeks of the year, and low rainfall predicted a peak in spring, and high rainfall predicted a peak at the end of the monsoon (10). However, the mechanisms of this seasonality of cholera are still not fully understood, despite long-standing recognition of the bimodal seasonality in Bangladesh. Hashizume and his colleagues (11) observed an exceptionally high SST and sea-surface height (SSH), preceding a sharp increase in the number of cholera patients in Dhaka in 1998, suggesting that SST should be taken into account when building predictive models for cholera, using ocean-climate data. Although attempts to predict the incidence of cholera, using ocean-climate data from preceding months, have been made (12-13), an accurate climate-based prediction of cholera epidemics with a longer lag time has not yet been developed. Because of the serious global consequences of cholera and its sensitivity to climate, the World Health Organization has proposed developing an early warning system for cholera epidemics, using climatic parameters (14).

MATERIALS AND METHODS

Study area

The study site Matlab is located in south-central Bangladesh, approximately 50 km southeast of the capital city Dhaka and is adjacent to where the river Ganges meets the river Meghna, forming the lower Meghna. This is a field research site of the International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b). The river Dhonagoda flows from north to south, bisecting the study area into two approximately equal parts. Numerous canals also exist in the study area. These canals remain dry in the winter and become full of water during the monsoon. An embankment was built alongside the Dhonagoda and the Meghna, which was commissioned in full at the end of 1989. The embankment was built primarily to protect the area from monsoon flooding so that agricultural activities can be carried out throughout the year. The embankment protects 31% of approximately 210,000 people living in the study area from flooding.
The study area has a tropical monsoon-type climate, with a hot and rainy summer and a dry winter. For practical purposes of this study, three seasons were distinguished in Bangladesh: summer (March-June), rainy (July-October), and winter (November-February). January is the coolest month with temperatures averaging near 26 °C (78 °F), and April is the warmest with temperatures ranging from 33 °C to 36 °C (91 °F to 96 °F). The climate is among the wettest in the world. The study area receives more than 2,100 mm of rain a year. Most of the rain occurs during the monsoon (June-September), and there is little rain in the winter (November-February).

Population data

The population database of icddr,b in Matlab is the most comprehensive longitudinal demographic database on a large population in the developing world (15). A health and demographic surveillance system (HDSS) has recorded all vital events of the area’s population since 1966. The population in the study area was approximately 210,000 during the study period.

Cholera data

Diarrhoea surveillance data were obtained for all individuals living in the Matlab study area from 1988 to 2001. Cholera cases were defined as *Vibrio cholerae* isolated from faecal specimens of patients who sought care at treatment centres serving the Matlab population. Cholera cases were aggregated by month. The cholera incidence in a month was calculated as the number of cases in the month divided by the mid-year population of the study area in the year, and the rates are expressed as cases per 1,000 population. Figure 1 shows the fluctuations of monthly cholera incidence during the study period.

Environmental data

We obtained AVHRR (Advanced Very High Resolution Radiometer) satellite-derived SST data for 1985-2001 from NASA’s Jet Propulsion Laboratory. The satellite sensor collects SST data at 4 km area. *In-situ* environmental data include monthly ambient temperature and rainfall that were made available for the study period (1985-2001) from a weather station in Chandpur, which is located just outside the study area. We specifically selected these three environmental variables for this study because these variables are often found to be predictors of the magnitude of cholera in various parts of the world (5,7,15,22).

**Analytical methods**

We analyzed the data by month, incorporating 168 time-points during the study period (1988-2001). The data prior to 1988 were avoided because there was an oral cholera vaccine trial in early 1985 that offered direct (23) and indirect protection (24) to people of the area. A SARIMA model was used for evaluating the effects of environmental variables on cholera transmission and assessing the ability of the model for forecasting cholera transmission in the study area. Using data from 1988 to 2000, we fitted the model to cholera incidence (cases per 1,000 people), and then the fitted model was used in predicting cholera cases for the year 2001.

We used the Box-Jenkins modelling strategy (25) to conduct a time series analysis. First, we evaluated the need for variance-stabilizing transformation by simple inspection of the graph of the untransformed series and practical tool, which is the mean-range plot (the range is plotted against the means for each seasonal period) of the untransformed and some transformed series (e.g. logarithm or square root). If the mean-range plots display a random scatter around a straight line, the logarithm transformation is needed (26). We determined the order of non-seasonal (p,d,q) and seasonal (P,D,Q) autoregressive (AR) parameters (p and P) and moving average (MA) parameters (q and Q) as well as the need for non-seasonal and seasonal differencing (d and D) and seasonal period (s), using the following tools: (i) plot of cholera incidence and unit root test, which assists with the non-seasonal and seasonal differencing; (ii) autocorrelation (ACF) and partial autocorrelation (PACF) functions, which indicate the temporal dependence structure in the stationary time series; (iii) Akaike Information Criterion (AIC), which assists in the goodness-of-fit of the model penalizing for the number of parameters; (iv) Ljung-Box test, which measures the ACF of the residuals; and (v) significance of the parameters.

We estimated the parameters of the SARIMA model by maximum likelihood.

The goodness-of-fit of the model was determined for appropriate modelling, using both time series (ACF and PACF of residuals) and classic tools, which checked for the normality of the residuals (27). Finally, we graphically compared the model’s fitted values with the observed data to check if it, indeed, models the cholera incidence. We assessed the pre-
dictive power of forecasting ability of the method, using the root mean square error (RMSE) criterion.

RESULTS

There were 4,157 cholera cases during the study period, with an average of 1.4 cases per 1,000 people. The characteristics of the data are shown in Table 1. The time series of cholera incidence shows that there was no trend in cholera rates during the study period (Figure 1). Through the mean-range plots for each seasonal period (12 months), we found it necessary to stabilize the variance of cholera incidence by computing its natural logarithm (Figure 2). All further statistical procedures, descriptive and analytical, were performed on the logarithmically-transformed cholera incidence. The plots of the sample ACF describes temporal dependence of cholera incidence at lag 1, 2, and at lag 11, 12, 13, and PACF describes the temporal dependence of cholera incidence at lag 1 and at lag 11 (Figure 3), suggesting that non-seasonal and seasonal parameters are needed in the model to account for the temporal autocorrelation in the data. Upon using the Augmented Dickey-Fuller Unit Root Test and checking ACF and PACF (28,29), we found that the data needed one-month non-seasonal differencing and 12-month seasonal differencing to make the data stationarity. Finally, we fitted the SARIMA \((0,1,2) (0,1,1)_12\) model, which was found to be the best for these data.

| Table 1. Characteristics of the study data, Matlab, 1988-2001 |
|---------------------------------------------------------------|
| Variable (monthly average) |
| Monthly cholera cases (number) | No. of months | Mean | SD | Minimum | Maximum |
| Population (number) | 168 | 209,620 | 6,330 | 199,511 | 219,873 |
| Minimum temperature (in degree celsius) | 168 | 20.98 | 4.79 | 10.30 | 27.30 |
| Maximum temperature (in degree celsius) | 168 | 30.21 | 2.89 | 22.40 | 35.10 |
| Rainfall (in mm) | 168 | 168.65 | 156.36 | 0 | 746.0 |
| Sea-surface temperature (in degree celsius) | 168 | 28.34 | 1.93 | 24.3 | 31.95 |
| SD=Standard deviation |

![Figure 1. Temporal patterns of cholera incidence, Matlab, 1988-2001](image)

![Figure 2. Mean-range plot for cholera incidence per 1,000 population for non-transformed data (left), square-root transformation data (middle), and log-transformed data (right), Matlab, 1988-2001](image)
Inter-correlation among the independent variables, using a 12-month seasonal differencing, shows that minimum and maximum temperatures were significantly correlated with each other ($p<0.01$; Table 2). Maximum temperature was significantly and negatively correlated with rainfall ($p<0.01$). The results of the cross-correlations, using 12-month seasonal differencing, show that cholera incidence was positively associated ($p<0.01$) with minimum temperature in the current month, negatively associated ($p<0.05$) at a lag of three months, and positively associated ($p<0.05$) at a lag of five months (Table 3). Maximum temperature was positively associated ($p<0.05$) with cholera incidence in the current month, negatively associated ($p<0.05$) at a lag of one-month, positively associated ($p<0.05$) at a lag of two months, and again negatively associated ($p<0.05$) at a lag of three months. SST was negatively associated ($p<0.05$) with magnitude of cholera at a lag of two months (Table 3). The data on average maximum temperature did not fit well in our model as seen in a study (30); therefore, we used the average minimum temperature in our model. The results of the SARIMA model showed that minimum temperature during the current month and SST at current month and at 2-month lag significantly influenced ($p<0.01$) cholera incidence (Table 4).

By using the SARIMA model with only minimum temperature variable in the equation, it showed that, with a minimum temperature increase of one degree celsius in the current month, there was a 6% increase in cholera incidence [exp (b) = 1.06, $p<0.01$, caeteris paribus]. For increase of SST by one degree celsius, there was a 25% increase in the cholera incidence at current month and 18% increase of the cholera incidence at two months [caeteris paribus]. This model was used in predicting cholera incidence from 1989 to 2001, and the results are shown in Figure 4. Since we used moving average from the 12-month seasonal differencing in the SARIMA model, the model eventually accounts for the data from 1989. It is apparent that the predicted incidence rates were reasonably well-matched with the observed incidence rate. The RMSE, a measure of the size of residuals from the model that used minimum temperature to predict cholera incidence, was 0.108, which means that if the

---

**Table 2.** Inter-correlations among environmental variables, using 12-month seasonal differencing data, Matlab, 1988-2001

| Variable       | Maximum temperature | Rainfall | Sea-surface temperature |
|----------------|---------------------|----------|-------------------------|
| Minimum温度    | 0.31**              | -0.13    | 0.11                    |
| Maximum温度    | -                   | -0.33**  | 0.07                    |
| Rainfall       | -                   | -        | -0.11                   |
| **p<0.01**     |                     |          |                         |

---

**Table 3.** Cross-correlation coefficients of logarithmic transformation of observed cholera incidence and environmental variables, using 12-month seasonal differencing data, Matlab, 1988-2001

| Variable            | Lag (month) 0 | 1 | 2 | 3 | 4 | 5 |
|---------------------|---------------|---|---|---|---|---|
| Minimum温度         | 0.26**        | -0.01 | 0.04 | -0.15* | -0.01 | 0.16* |
| Maximum温度         | 0.16*         | -0.19* | 0.17* | -0.15* | 0.02 | 0.06 |
| Rainfall            | -0.1          | -0.04 | 0.04 | -0.01 | 0.05 | 0.11 |
| Sea-surface温度     | 0.12          | 0.10 | -0.15* | -0.07 | 0.04 | -0.01 |
| **p<0.05; **p<0.01**|               |     |     |    |    |    |
minimum temperature is used in predicting cholera incidence, roughly two-thirds of the residuals will fall between -0.108 and +0.108. The goodness-of-fit analyses show that there was no significant autocorrelation between residuals at different lags in the SARIMA model (Figure 5). The Shapiro-Wilk test for normality also showed that the residuals were distributed normally (p=0.78).

| Table 4. Regression coefficients of the SARIMA on the monthly cholera incidence, Matlab, 1988-2001 |
| Parameter                                                                 | Estimate (b) | Standard error | p value |
| One-month moving average of cholera incidence (log-transformed)            | 0.406        | 0.079          | <0.0001 |
| Two-month moving average of cholera incidence (log-transformed)            | 0.207        | 0.078          | 0.008   |
| Seasonal moving average (12-month) of cholera incidence (log-transformed) | 0.691        | 0.081          | <0.0001 |
| Minimum temperature at current month                                       | 0.062        | 0.026          | 0.0015  |
| SST at current month                                                       | 0.229        | 0.061          | <0.0001 |
| SST at 2-month lag                                                        | 0.161        | 0.061          | 0.008   |
| Akaike Information Criterion (AIC)                                         |              |                | 362.74  |

Figure 4. Predicted and observed monthly cholera incidence rate, Matlab, 1989-2001

Figure 5. ACF (left) and PACF (right) of residuals from the SARIMA model
DISCUSSION

The results of this study illustrate that fluctuation in the incidence of cholera in Matlab is tempera-
ture-driven. The relationship between both ambient
and sea-surface temperature and amplification
of cholera incidence is well-documented (5-6,8,
30-33). This association likely exists because higher
temperatures facilitate Vibrio reproduction, which
increases concentrations in the aquatic habitat. At
greater concentrations, there is an increased prob-
ability that people living in the area ingest an infec-
tious dose (6). Previous studies (15,33) conducted
in Matlab showed that increased cholera caseload
is associated with higher temperature, lower river
discharge, and higher ocean chlorophyll concentra-
tion. We could not use the data on river discharge
and ocean cholera concentration due to lack of re-
liable datasets. However, as found in a study con-
ducted in Dhaka (9), rainfall has not been found to
be an influencing factor for the outbreak of cholera
in Matlab, suggesting that the causal pathway of
cholera between Dhaka and Matlab is not the same.
The link between cholera outbreaks and higher
temperatures is an important finding in light of
the global warming phenomenon. Climatologists
predict a 1.4 °C to 5.8 °C rise in mean temperatures
over the next 100 years (34). Increased sea tempera-
ture and levels associated with global warming sug-
gest the possibility of increased cholera incidence
in many resource-poor regions of the world.

The SARIMA model, based on temperature, fore-
cast cholera incidence in Matlab reasonably well.
We used the SARIMA model which is appropriate
for analyzing time series data (17); it determines
whether AR or MA terms are needed to correct for
autocorrelation that remains in the seasonally-
differenced data. We found some degree of tempo-
ral autocorrelation in the data, which was removed
from the stationarized series by adding autoregres-
sive terms (lags of the stationarized series) to the
model. The findings of this study, however, lead
the authors to believe that seasonal variation in
temperature may contribute to cholera incidence
in complex ways, presumably through interaction
with some unmeasured environmental or behav-
ioral factors (10).

The results of our study show three peaks in the
cholera incidence in a month of the year 1993,
1994, and 1995. There was an emergence of new
strain of cholera—V. cholera O139—in 1993 (35),
which co-existed with El Tor for some time and
showed a peak in 1993. There are fluctuations of
cholera cases by year as well as variations in the
seasonal peak in Matlab (36). The three peaks are
just the results of the outbreak of V. cholera O139
and variations of cholera cases by year and by sea-
son. Cholera incidence in the study area is usually
higher before and after the monsoon (36,37); thus,
we usually see two peaks in a year. This study also
found that the lag relationship with the outcome
is not linear and one-directional. It is possible that
the cholera caseload can be higher at the month
of increase in temperature and then can be lower
at three months after increase in tempeara in a
country, like Bangladesh.

Limitations

The limitation of this study is that we could not
collect ocean chlorophyll concentration (OCC)
for the study period, which is found to be associ-
ated with the magnitude of cholera in Bangladesh
(16). Also, we could not collect the data on sun-
shine hour, which is found to be associated with
the magnitude of cholera (33). Additionally, since
our forecasting model is based on the symptomatic
cases and the immunity derived from the mild in-
fec tions wanes very rapidly, the inapparent infec-
tions in the study area may change the patterns of
forecasting the disease outbreaks (38).

Conclusions

Albeit with the limited datasets, this paper supports
the premise that the magnitude of cholera in Mat-
lab is temperature-driven, both ambient and sea-
surface temperature. A perfect model for predicting
cholera outbreaks in the area based on only envi-
ronmental variables is difficult because the disease
may have complex relationships with local popula-
tion dynamics, the population's health behaviour,
and health practices. The results of this study may
provide a basis for predicting cholera epidemics in
Bangladesh and have the potential to improve dis-
 ease control in vulnerable areas.

ACKNOWLEDGEMENTS

Funding for the analysis reported in this paper was
provided by the NOAA Oceans and Human Health
Program. The data used in this paper were collected
with the support of icddr,b and its donors which
provide unrestricted support to icddr,b for its oper-
ations and research. Current donors providing un-
restricted support include: Australian International
Development Agency (AusAID), Government of
the People's Republic of Bangladesh, Canadian In-
ternational Development Agency (CIDA), Swedish
International Development Cooperation Agency
(Sida), and the Department for International De-
velopment (DFID), UK.
REFERENCES

1. Colwell RR. Global climate and infectious disease: the cholera paradigm. Science 1996;274:2025-31.
2. Epstein PR, Ford TE, Colwell RR. Marine ecosystems. Lancet 1993; 342:1216-9.
3. Patz JA, Epstein PR, Burke TA, Balbus JM. Global climate change and emerging infectious diseases. JAMA 1996;275:217-23.
4. Koelle K, Pascual M. Disentangling extrinsic from intrinsic factors in disease dynamics: a nonlinear time series approach with an application to cholera. Am Nat 2004;163:901-13.
5. Lobitz B, Beck L, Huq A, Wood B, Fuchs G, Faruque ASG et al. Climate and infectious disease: use of remote sensing for detection of Vibrio cholerae by indirect measurement. Proc Natl Acad Sci USA 2000;97:14384-43.
6. Sack DA, Sack RB, Nair GB, Siddique AK. Cholera. Lancet 2004;363:223-33.
7. Fernández MAL, Bauernfeind A, Jiménez JD, Gil CL, El Omeiri N, Guibert DH. Influenza of temperature and rainfall on the evolution of cholera epidemics in Lusaka, Zambia, 2003-2006: analysis of a time series. Trans R Soc Trop Med Hyg 2009;103:137-43.
8. Huq A, Sack RB, Nizam A, Longini IM, Nair GB, Ali A et al. Critical factors influencing the occurrence of Vibrio cholerae in the environment of Bangladesh. Appl Environ Microbiol 2005;71:4645-54.
9. Hashizume M, Armstrong B, Hajat S, Wagatsuma Y, Faruque ASG, Hayashi T et al. The effect of rainfall on the incidence of cholera in Bangladesh. Epidemiology 2008;19:103-10.
10. Hashizume M, Faruque ASG, Wagatsuma Y, Hayashi T, Armstrong B. Cholera in Bangladesh: climatic components of seasonal variation. Epidemiology 2010;21:706-10.
11. Hashizume M, Faruque ASG, Terao T, Yunus M, Streetfield K, Yamamoto T et al. The Indian Ocean dipole and cholera incidence in Bangladesh: a time-series analysis. Environ Health Perspect 2011;119:239-44.
12. Akanda AS, Jutla AS, Islam S. Dual peak cholera transmission in Bengal Delta: a hydroclimatological explanation. Geophys Res Lett 2009;36:1-6.
13. de Magny GC, Murtugudde R, Sapiano MR, Nizam A, Brown CW, Busalacchi AJ et al. Environmental signatures associated with cholera epidemics. Proc Natl Acad Sci USA 2008;105:17676-81.
14. World Health Organization. Using climate to predict infectious disease epidemics. Geneva: World Health Organization, 2005. 54 p.
15. Emch M, Feldacker C, Yunus M, Streetfield PK, Dinh-Thiem V, Canh do G et al. Local environmental predictors of cholera in Bangladesh and Vietnam. Am J Trop Med Hyg 2008;78:823-32.
16. Helfenstein U. The use of transfer function models, intervention analysis and related time series methods in epidemiology. Int J Epidemiol 1991; 20:808-15.
17. Zhang Y, Bi P, Hiller J. Climate variations and salmonellosis transmission in Adelaide, South Australia: a comparison between regression models. Int J Biometeorol 2008;52:179-87.
18. Yurekli K, Kurunc A, Ozturk F. Application of linear stochastic models to monthly flow data of Kelkit Stream. Ecol Modell 2005;183:67-75.
19. Zhang GP. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing 2003;50:159-75.
20. Allard R. Use of time-series analysis in infectious disease surveillance. Bull World Health Organ 1998;76:327-33.
21. Hu W, Nicholls N, Lindsay M, Dale P, McMichael AJ, Mackenzie JS et al. Development of a predictive model for Ross river virus disease in Brisbane, Australia. Am J Trop Med Hyg 2004;71:129-37.
22. Constantine de Magny G, Guégan JF, Petit M, Cazelles B. Regional-scale climate-variability synchrony of cholera epidemics in West Africa. BMC Infect Dis 2007;7:20.
23. Clemens JD, Sack DA, Harris JR, Van Loon E, Chakraborty J, Ahmed F et al. Field trial of oral cholera vaccines in Bangladesh: results from three-year follow-up. Lancet 1990;335:270-3.
24. Ali M, Emch M, von Seidlein L, Yunus M, Sack DA, Rao M et al. Herd immunity conferred by killed oral cholera vaccines in Bangladesh: a reanalysis. Lancet 2005;366:44-9.
25. Box GEP, Jenkins GM. Time series analysis: forecasting and control. Revised edition. San Francisco: Holden Day, 1976. 558 p.
26. Helfenstein U. Box-Jenkins modelling of some viral infectious diseases. Stat Med 1986;5:37-47.
27. Tabachnick BG, Fidell LS. Using multivariate statistics. 5th ed. Boston: Pearson, 2007. 1,008 p.
28. Dickey DA, Fuller WA. Distribution of the estimators for autoregressive time series with a unit root. J Am Stat Assoc 1979;74:427-31.
29. Dickey DA, Hasza DP, Fuller WA. Testing for unit roots in seasonal time series. J Am Stat Assoc 1984;79:355-67.
30. Reyburn R, Kim DR, Emch M, Khatib A, von Seidlein L, Ali M. Climate variability and the outbreaks of cholera in Zanzibar, East Africa: a time series analysis. Am J Trop Med Hyg 2011;84:862-9.
31. Colwell RR. Infectious disease and environment: cholera as a paradigm for waterborne disease. *Int Microbiol* 2004;7:285-9.

32. Alam M, Sultana M, Nair GB, Siddique AK, Hasan NA, Sack RB *et al.* Viable but nonculturable *Vibrio cholerae* O1 in biofilms in the aquatic environment and their role in cholera transmission. *Proc Natl Acad Sci USA* 2007;104:17801-6.

33. Islam MS, Sharker MAY, Rheman S, Hossain S, Mahmud ZH, Islam MS *et al.* Effects of local climate variability on transmission dynamics of cholera in Matlab, Bangladesh. *Trans R Soc Trop Med Hyg* 2009;103:1165-70.

34. Lipp EK, Huq A, Colwell RR. Effects of global climate on infectious disease: the cholera model. *Clin Microbiol Rev* 2002;15:757-70.

35. Ali M, Emch M, Yunus M, Sack RB. Are the environmental niches of *Vibrio cholerae* O139 different from those of *Vibrio cholerae* O1 El Tor? *Int J Infect Dis* 2001;5:214-9.

36. Longini IM, Jr., Yunus M, Zaman K, Siddique AK, Sack RB, Nizam A. Epidemic and endemic cholera trends over a 33-year period in Bangladesh. *J Infect Dis* 2002;186:246-51.

37. Emch M, Yunus M, Escamilla V, Feldacker C, Ali M. Local population and regional environmental drivers of cholera in Bangladesh. *Environ Health* 2010;9:2.

38. King AA, Ionides EL, Pascual M, Bouma MJ. Inapparent infections and cholera dynamics. *Nature* 2008;454:877-80.