Original Report

Are the Environmental Niches of *Vibrio cholerae* O139 Different from Those of *Vibrio cholerae* O1 El Tor?

Mohammad Ali, PhD;* Michael Emch, PhD;† Mohammad Yunus, MBBS;* and Richard Bradley Sack, PhD, MD‡

ABSTRACT

Background: *Vibrio cholerae* are known to be normal inhabitants of surface water. However, the environmental niches of the different strains of cholera are not well known, and therefore, populations at risk for cholera outbreaks cannot be clearly identified.

Methods: This study identifies environmental risk factors for cholera caused by *V. cholerae* O1 El Tor and O139 and environmental niches of the two strains present in Matlab, a cholera endemic area of Bangladesh. The study year was 1993, the year that the O139 strain first appeared in the study area. Patients who had either strain of cholera identified in a laboratory were included in the study. A geographic information system was used to map the household locations of the patients, to describe the human sanitary environment and population density, and to address potential anthropogenic and environmental risk factors of the disease. Spatial point pattern and exploratory spatial data analysis techniques were used to define the environmental niches of the two cholera strains.

Results: The study suggests the niches of O1 El Tor and O139 strains of *V. cholerae* appear to be similar, based on common environmental risk factors.

Conclusions: The results of this study support a theory that O1 El Tor could possibly be replaced by the newer O139 strain in the future.

Key Words: *cholera, environment, spatial analysis*, Vibrio cholerae *O139*

Int J Infect Dis 2001; 5:214–219.

The emergence of a new strain of cholera, *Vibrio cholerae* O139, and its rapid spread in Bangladesh and neighboring countries, and a few cases in other parts of the world, has raised the question, Does this new strain have the potential to cause a pandemic?[1-10] Referring back to the emergence of El Tor, the strain was initially recognized as a mild cholera-like disease in an Indonesian village in 1937 and was confined there for approximately 2 decades.[11] In 1959, it was detected in Thailand, and by 1963, it had extended to India and Bangladesh in pandemic form. It then spread throughout the world, attacked more than half a million people, and claimed 5000 lives within 18 months.[12] Evidence suggests that classic cholera has been entirely replaced by El Tor. The replacement has occurred because the environmental niches where the vibrios live are the same for the two strains.[13] They also share similar ecologic environments and patterns of transmission.[14] *Vibrio cholerae* O139 has spread even more rapidly than did O1 biotype El Tor. This leads to the question of whether El Tor will be replaced by O139 in the future.

An environmental study on cholera conducted in Bangladesh has shown that *V. cholerae* O139 is present in ponds, lakes, and rivers.[15] Identifying the niches of O139 requires defining the areas of increased risk of contracting the disease. Once these areas are known, the focus may shift to investigating the characteristics of the areas, in an effort to identify environmental risk factors of the disease. In an earlier study, the environmental risk factors of the two previous biotypes of cholera (classic and El Tor) were found to be the same.[16] The authors set out to determine whether the environmental risk factors of O139 are the same as for the other two strains of cholera, and if so, where O139 cholera is being contracted relative to where El Tor is contracted.

The predominant mode of transmission for the O139 vibrios is through water.[15] Its rapid spread in populations, pattern of infection, and ability to survive in aquatic environments, suggest that O139 is more infectious, more virulent, and in an ecologic sense, a more robust organism, whose pandemic potential appears to be significant.[17] A comparative analysis of the environmental niches of O139 and another strain of cholera (El Tor) would reveal whether the niches of the two strains are the same or not. Eventually, such a study would show where the two

*Centre for Health and Population Research, Dhaka, Bangladesh; †Department of Geography, Portland State University, Portland, Oregon; ‡Johns Hopkins University, Baltimore, Maryland.

Address correspondence to Dr. Mohammad Ali, International Vaccine Institute, Kwanak PO Box 14, Seoul, Korea 151-600. E-mail: mali@ivi.int.

strains of cholera create increased risk of the disease in an endemic situation. This study used a geographic information system (GIS) to describe the spatial and environmental epidemiology of O139 and O1 El Tor strains and investigated the differences in environmental risk factors for the two strains of cholera to hypothesize future trends of the O139 strain of cholera.

DATA AND METHODS

The study was conducted in Matlab, a field research area of the Centre for Health and Population Research (ICDDR,B). Matlab, which is 184 km^2, is an area in central Bangladesh that has endemic cholera. The Dhonagada River bisects the study area, dividing it into two nearly equal parts. In 1988, an embankment was completed adjacent to the river to allow year-round agricultural activities inside the embankment and to provide flood protection for local communities. The 210,000 people living in the study area reside in clusters of patrilineally related households called *baris*. Initiated in 1966, a demographic surveillance system (DSS) recorded all vital demographic events of the study area population. In 1994, a GIS was implemented and the DSS components were integrated within the framework of a spatial information system. The GIS facilitated investigations of local-level spatial variability of health and environmental phenomena within the study area. This study integrated DSS data on population, migrations, and socioeconomic status within the spatial framework provided by the GIS. There is also a diarrhea hospital in the surveillance area and patients' records are maintained in a hospital surveillance system. Diarrheal stool samples are collected for all patients admitted from the DSS area, and they are tested for enteric pathogens in the laboratory. The study extracted data on patients with cholera from the hospital surveillance database. This study analyzed 1993 cholera data, the year of the O139 epidemic, acquired from the

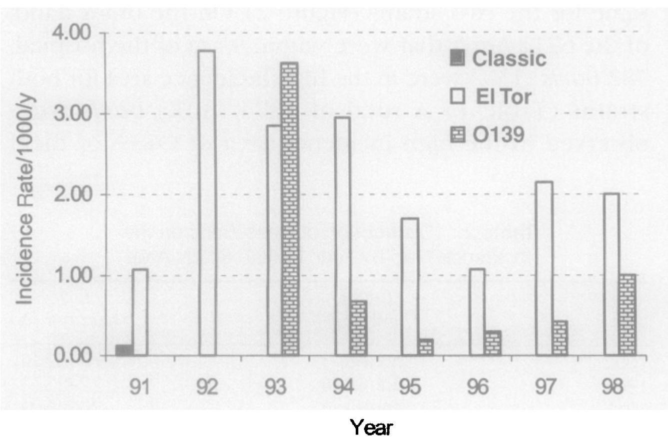

Figure 1. Temporal pattern of incidence rate of cholera in Matlab study area, 1991–1998.

hospital records (Figure 1). A raster GIS was used to conduct the spatial analysis for this study. The spatial resolution of the pixels was set to 30 meters to represent a *bari* by a single pixel. Each data set including cholera morbidity, population, migrations, educational status, and sanitation condition was rasterized, and the data became the attributes of the pixels.

Cholera Incidence

Exploratory spatial data analysis methods were used to define the high incidence areas for both O139 and O1 El Tor.[18,19] A low-pass spatial filtering technique was employed within the raster GIS to obtain a spatial moving average rate of the disease incidences.[20,21] This method removes the dependency of the data at specific points, and transforms the rates into a spatially continuous form, which is a better reflection of the actual disease surface.[22] Since the distribution of the study population varies in space, instead of using a fixed area, this study used a fixed population size to compute the moving average.[23] The size of the population that was used to calculate the moving average was 35 people, because that was determined to be the optimum number to remove random noise (e.g., inaccurate population records, differences in population structure within *baris*, mislocated cases) while retaining local level variation in the data. Since this differential filtering method was used to compute the incidence rates, the targeted population was obtained by choosing varying sizes of filters ranging from 3×3 to 21×21 pixels. The spatial moving average incidence rate surfaces were used to define high incidence areas for the two strains of the disease.

Kriging,[24] a method that fits spatial data to a linear variogram model, was used to interpolate the data at regularly spaced intervals (150×150 m). The spatial distribution of the *baris* prompted the choice of that specific interval, because it provided enough local-level variation while reducing random noise in the data. The interpolated surfaces were then used to define the areas of cholera disease incidences. Since few patients who lived further than 9 km from the hospital came to the hospital for treatment, the interpolated surfaces had several holes in that part of the study area. Therefore, the areas beyond 9 km were excluded from the incidence maps. One-fourth of the highest incidence area of the disease surface was considered to be the high incidence area. The high incidence area had a threshold of 6.1 per 1000 people for O139 and 4.1 per 1000 for El Tor.

Independent Variables Hypothesized To Be Related to Cholera Incidence

A raster surface of educational status was created by using a spatial filter on the proportion of people having at least 4 years of secular education. A constant filter size (7×7 pixels) was chosen, because it was assumed that there

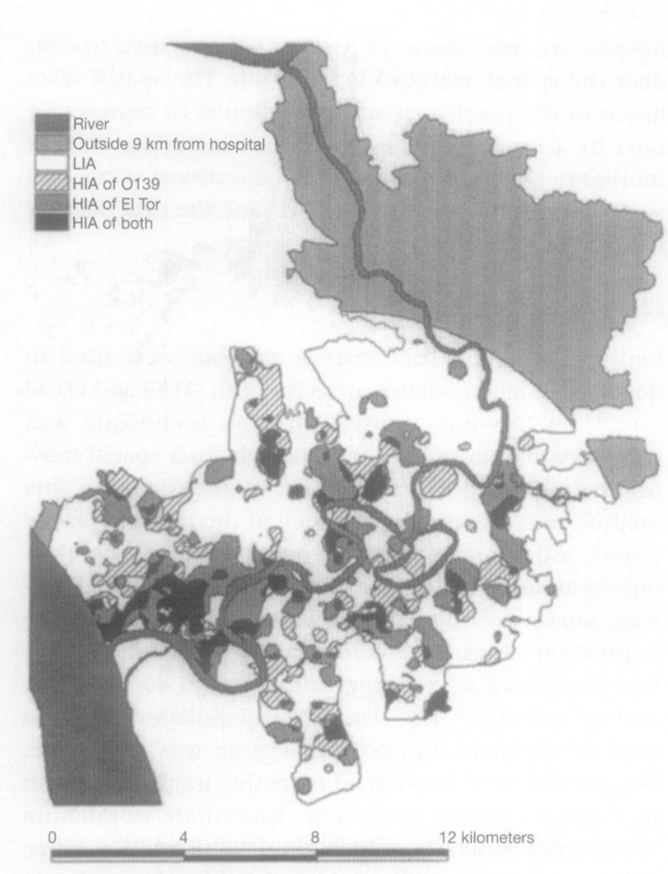

Figure 2. Cross-classification map of the high incidence areas of O139 and El Tor, 1993 Matlab study area. HIA = high incidence area; LIA = low incidence area.

would be no influence from neighbors' educational status beyond that distance. The filter also assumed that neighbors' influences diminish with increased distances. Therefore, the status of neighbors' *baris* was weighted by a distance decay term ($e^{-\text{spatial lag}}$), which yielded rapidly diminishing values with increased spatial lag. An edge correction term, based on the proportion of filter pixels that falls within the study area, was also introduced in the computation, to remove boundary effects. The same method was applied while creating the surface of poor sanitation status. Latrine types (sanitary or unsanitary) used by the *bari*-members were used to create the surface. Latrines in which feces could overflow to nearby areas were defined as unsanitary. The sanitation surface was created by summing the number of people who used unsanitary latrines in the 7- × 7-pixel area, weighted by the distance decay term.

A surface of population density was created based on the density of population around the *baris*. The fixed 7- × 7-pixel filter with unitary kernel values was used to aggregate the population surrounding the *baris*. The edge correction term transformed the data into the density of population per unit area. Similarly, a surface of the number of people who migrated into the study area was created. A map of the embankment was created as a binary

Table 1. Cross-Classification of the *Baris* in Low and High Incidence Areas of El Tor and O139, Matlab Study Area, 1993

	El Tor		
O139	*Low Incidence Area (n = 4495)*	*High Incidence Area (n = 1717)*	*Total (n = 6212)*
Low incidence area	3403	935	4338
(% within O139)	(78.4)	(21.6)	
(% within El Tor)	(75.7)	(54.5)	
High incidence area	1092	782	1874
(% within O139)	(58.3)	(41.7)	
(% within El Tor)	(24.3)	(45.5)	

map with attribute values of 0 representing pixels inside the embankment and values of 1 representing pixels outside embankment. Rivers and canals were defined as the waterbodies. A distance matrix image was created from this waterbody map and then the distance to the nearest waterbody for each *bari* was computed by superimposing the *bari* image onto the distance image.

Statistical Methods

The data were arranged by *baris* for analyzing environmental risk factors of cholera associated with both strains, O139 and O1 El Tor. The data created within the GIS environment were extracted at the *bari* level to perform the statistical analysis. Both simple and multiple regression were used. Simple linear regression was used to describe the bivariate relation between cholera morbidity and each of the independent variables introduced in this article. Stepwise multiple regression was used to retain only those factors in the final equation that have a significant relation with cholera morbidity.

RESULTS

The high incidence areas for O139 and O1 El Tor were compared using cross-classification analysis. The analysis revealed that 35% of the high incidence areas are the same for the two strains (Figure 2). On the other hand, of the 6212 *baris* that were within 9 km of the hospital, 782 *baris* (13%) were in the high incidence area for both strains (Table 1). A total of 1874 (30%) *baris* were observed in the high incidence area of O139; of them

Table 2. Distribution of Case-*Baris* on the Incidence Map by Year, Matlab Study Area

Year	*O139 Number (%)*	*El Tor Number (%)*
1993	356 (68.1)	279 (82.8)
1994	44 (46.3)	122 (33.7)
1995	11 (34.4)	79 (35.1)
1996	19 (42.2)	60 (37.3)
1997	21 (32.8)	122 (41.1)
1998	50 (38.5)	101 (36.7)

Table 3. Descriptive Statistics of the Study Variables, 1993 (n = 6212)

Variables	*Mean*	*SD*
Incidence rate of O139/1000/y	4.68	6.16
Incidence rate of El Tor/1000/y	3.67	8.10
Educational status*	0.82	0.55
Poor sanitation condition*	1.38	0.67
Density of migrations (per unit area)	25.52	26.98
Density of population (per unit area)	172.33	122.81
Distance to the nearest waterbody (m)	168.08	170.43

*Arbitrary units (see the text for descriptions).

42% fell within the area of high O1 El Tor incidence as well. The *baris* with O139 and O1 El Tor cases were overlaid with the respective incidence maps beyond 1993, and the results of this comparison are shown in Table 2. The results show that in each of the years, a large proportion of the case-*baris* fell within the high incidence areas for both the strains. The median distance of those *baris* that did not fall within the high incidence areas was only 200 meters.

Table 3 summarizes the study variables. The table shows that the incidence rates for O139 and O1 El Tor were 4.68 and 3.67 per 1000 population, respectively, for 1993. Approximately 57% of the *baris* were outside the embankment, as obtained from the database. Table 4 shows the results of the simple linear regression. Areas outside the embankment and near surface water are significantly associated with the incidence rates for both O139 and O1 El Tor. A 41% lower incidence rate for O1 El Tor (obtained from the constant and coefficient of the regression equation) was found outside the embankment compared with the rate inside the embankment, and the difference in the rates is statistically significant ($P < 0.01$). For O139, the rate was also significantly lower (10%) outside the embankment. No significant associations were observed between the incidence rate of O139 and educational status or sanitation condition. However, these two variables did show significant associations with O1 El Tor incidence. Higher densities of migrations and population were found to be significantly associated with higher O139 incidence. These two variables did not show any impact on the disease incidence rate for O1 El Tor.

Results of the multiple regression analysis are shown in Table 5. After controlling for multicollinearity among the independent variables, the flood-control embankment, population density, and distance to surface water were all risk factors for O139 incidence. Density of migrations was not retained as a risk factor in the final equation because of its strong association with population density ($r = 0.77$). All of the risk factors for O1 El Tor found in the simple regression analysis were retained in the final equation of the multiple regression analysis.

DISCUSSION

This study used spatially scaled data analyzed within a GIS to identify the risk factors for high cholera incidence for both O139 and O1 El Tor strains. It also defined high incidence areas for both strains, which were coexisting in the endemic area in 1993. It was found that the risk factors related to the physical environment (proximity to surface water and whether a person lived within the flood control embankment) were the same for both strains of cholera in 1993. The absence of natural flushing by floodwater during the monsoon season may have changed the hydrologic dynamics within the embankment, resulting in increased salinity,[25] thus creating a suitable environment for both strains of cholera and, therefore, an increase in the disease incidence. The anthropogenic risk factors for the two strains were different. This study found that high population density is a risk factor for O139 whereas poor educational status and poor sanitation are risk factors for El Tor. Since O139 is a new strain, the population does not have any previous immunity, as they do to El Tor, which has been around for a long time. This could be one of the explanations for the differences in education and sanitation between rates of El Tor and O139. It may be noted that there is a rela-

Table 4. Results of the Simple Regression Analysis

Factor	*Constant (a)*	*Coefficient (b)*	*Standard Error b*	*t-statistic b*	*Significance b*
Dependent variable:incidence rate of O139					
Outside embankment	4.977	−0.516	.158	−3.265	.001
Educational status	4.663	0.02039	.142	0.143	.886
Poor sanitation condition	4.448	0.167	.117	1.430	.153
Density of migrations	4.373	0.01201	.003	4.155	.000
Density of population	4.132	0.003176	.001	5.004	.000
Distance to the nearest waterbody	4.911	−0.001376	.000	−3.004	.003
Dependent variable:incidence rate of El Tor					
Outside embankment	4.828	−1.999	.207	−9.675	.000
Educational status	4.457	−0.954	.187	−5.111	.000
Poor sanitation condition	2.528	0.830	.154	5.399	.000
Density of migrations	3.766	−0.003628	.004	−0.952	.341
Density of population	3.450	0.001296	.001	1.549	.121
Distance to the nearest waterbody	3.995	−0.001915	.001	−3.177	.001

Table 5. Results of the Multiple Regression Analysis

Factor	*Coefficient (b)*	*Standard Error b*	*t-statistic b*	*Significance b*
Dependent variable:incidence rate of O139				
(constant)	4.592	.177	25.876	.000
Outside embankment	–0.440	.159	–2.762	.006
Density of population	0.003073	.001	4.843	.000
Distance to the nearest waterbody	–0.001122	.000	–2.431	.015
Dependent variable:incidence rate of El Tor				
(constant)	4.958	.336	14.738	.000
Outside embankment	–1.719	.216	–7.974	.000
Higher educational status	–0.958	.186	–5.154	.000
Poor sanitation condition	0.529	.158	3.347	.001
Distance to the nearest waterbody	–0.001122	.001	–2.314	.021

tion between high population density and poor sanitation. Thus, high population densities are an indirect cause of high disease incidence.

All previous epidemiologic studies that identified socioenvironmental risk factors in the study area did not use a spatial research framework. The methods used in this study were more robust than those used in previous cholera studies, because GIS technology made it possible to measure environmental variables more appropriately.[26] Thus, it could be confidently determined that the risk factors that were identified were not spurious.

The results of the cross-classification analyses of both maps and tables reveal that the high incidence areas for both strains were close to each other, indicating the close proximity of the origin of the disease for both strains. The results of the analysis also indicate that the definition of high incidence areas is more robust over time, since the cases in the subsequent years were either within or near high incidence areas.

Differences in the anthropogenic risk factors of the two cholera strains may provide a reason that the incidence areas were somewhat different. However, the observed risk factors provide some evidence that surface water near areas with high population densities is the niche for both strains of cholera in an endemic situation. In these settings, water is contaminated rapidly because of poor sanitation conditions, thus making the water in these areas a suitable environment for the vibrios. The authors previously found that the environmental niches of classic and El Tor were the same,[13] and the two strains coexisted only temporarily in the same area.

CONCLUSIONS

Cholera remains one of the great epidemic diseases of the tropical world, and it has spread throughout the world.[27] The findings from an endemic area of Bangladesh show that although the number of cases of O139 decreased after the outbreak in 1993, it has been steadily increasing since 1996. In contrast, a decrease in the number of cholera cases caused by El Tor was noticed since the O139 outbreak. Biochemical and genetic analyses showed that *V. cholerae* was closely related to the El Tor biotype, and it was suggested that it might be a mutant form of that biotype.[17] The pattern of infection by the O139 serotype and the ability of the organism to survive in aquatic environments suggest that O139 could be, in an ecologic sense, a more robust organism.[17] This indicates that it is a suitable organism to spread throughout the world. Since the environmental niches of the two strains appear to be the same, it is hypothesized that cholera is changing its form to O139, and it may replace El Tor in the future.

ACKNOWLEDGMENTS

This research was funded by Belgian Administration for Development Cooperation and the Centre for Health and Population Research (ICDDR,B), which is supported by countries and agencies that share its concern for the health problems of developing countries. Current donors providing unrestricted support include the aid agencies of the Governments of Australia, Bangladesh, Belgium, Canada, Japan, The Netherlands, Sweden, Sri Lanka, Switzerland, the United Kingdom, and the United States of America; international organizations include United Nations Children's Fund.

REFERENCES

1. Shimada T, Nair GB, Deb BC, Albert MJ, Sack RB, Takeda Y. Outbreak of *Vibrio cholerae* non-O1 in India and Bangladesh. Lancet 1993; 341:1346–1347.
2. Cholera Working Group (Albert MJ, Ansaruzzaman M, Bardhan PK, Faruque ASG, Faruque SM, Islam MS, Mahalnabish D, Sack RB, Salam MA, Siddique AK, Yunus M, Zaman K). Large epidemic of cholera-like disease in Bangladesh caused by *Vibrio cholerae* O139 synonym *Bengal*. Lancet 1993; 342:387–390.
3. Cheasty T, Rowe B, Said B, Frost J. *Vibrio cholerae* serogroup O139 in England and Wales. BMJ 1993; 307:1007.
4. Morris JG Jr and the Cholera Laboratory Task Force. *Vibrio cholerae* O139 Bengal. In: Wachsmuth K, Blake PO, Olsvik O, eds. *Vibrio cholerae* and cholera. Washington, DC: ASM Press, 1994:95–102.
5. Attapattu M. First isolation of *Vibrio cholerae* O139 (Bengal strain) from a Sri Lankan patient. Ceylon Med J 1994; 39:192–193.

6. Tay L, Goh KT, Lim YS. *Vibrio cholerae* O139 'Bengal' in Singapore. J Trop Med Hyg 1994; 97:317-320.
7. Sachdeva V, Khanna KK, Singh M, Singh J, Kumari S, Verghese T. Widespread emergence of *Vibrio cholerae* O139 in India. Southeast Asian J Trop Med Public Health 1995; 26:342-346.
8. Siddique AK, Akram K, Zaman K, Mutsuddy P, Eusof A, Sack RB. *Vibrio cholerae* O139: How great is the threat of a pandemic? Trop Med Int Health 1996; 1:393-398.
9. Dalsgaard A, Nielsen GL, Echeverria P, Larsen JL, Schonheyder HC. [The first case of *Vibrio cholerae* O139 in Denmark]. Ugeskr Laeger 1996; 158:5169-5171. [In Danish].
10. Dalsgaard A, Serichantalergs O, Forslund A, Pitarangsi C, Echeverria P. Phenotypic and molecular characterization of *Vibrio cholerae* O1 isolated in Samutsakorn, Thailand, before, during, and after the emergence of *V. cholerae*. Epidemiol Infect 1998; 121:259-268.
11. Barua D. History of cholera. In: Barua D, Greenough WB III, eds. Cholera. New York, London: Plenum Medical, 1992:1-36.
12. Epstein PR. Climate, ecology, and human health. Consequences 1997; 3(2). http://www.gcrio.org/consequences/vol3no2/climhealth.html.
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 2000 (in press).
14. Shears P. Cholera. Ann Trop Med Parasitol 1994; 88:109-122.
15. Islam MS, Hasan MK, Miah MA, et al. Isolation of *Vibrio cholerae* O139 *Bengal* from water in Bangladesh. Lancet 1993; 342:430.
16. Ali M, Emch M, Donnay JP, Yunus M, Sack RB. Identifying environmental risk factors of endemic cholera: a raster GIS approach. Health Place (submitted).
17. Dalsgaard DA, Larsen JL. [Is *Vibrio cholerae* serotype O139 a potential cause of a new pandemic?] Ugeskr Laeger 1995; 157:280-283. [In Danish].
18. Bithell JF. An application of density estimation to geographical epidemiology. Stat Med 1990; 9:691-701.
19. Bailey TC, Gatrell AC. Interactive spatial data analysis. Harlow, UK: Longman, 1995.
20. Meijerink AMJ, de Brouwer HAM, Mannaerts CM, Valenzuela C. Introduction to the use of geographic information systems for practical hydrology. Enschede, The Netherlands: ITC, 1994.
21. Watkins CD, Sadun A, Marenka S. Modern image processing: warping, morphing, and classical techniques. San Francisco: Morgan Kaufmann, 1993.
22. Rushton G. Improving the geographic basis of health surveillance using GIS. In: Gatrell AC, Loytonen M, eds. GIS and health (GISDATA Series 6). London: Taylor & Francis, 1998:63-79.
23. Talbot TO, Kulldroff M, Teven PF, Haley VB. Evaluation of spatial filters to create smoothed maps of health data. Stat Med 2000; 18:2399-2408.
24. Oliver MA, Webster R. Kriging: a method of interpolation for geographical information systems. Int J Geogra Inf Syst 1990; 4:313-332.
25. Siddique AK, Baqui AH, Eusof A, et al. Survival of classic cholera in Bangladesh. Lancet 1991; 337:1125-1127.
26. Ali M, Emch M, Ashley C, Streatfield PK. Implementation of a medical geographic information system: concepts and uses. J Health Popul Nutr 2001; 19:100-110.
27. WHO (1999). Cholera 1998. Wkly Epidemiol Rec 1999; 74:257-269.