Identifying clustering of cholera cases using geospatial analysis in Kolkata and surrounding districts: data from patients at tertiary care referral hospitals

Rounik Talukdar,[a,f] Suman Kanungo,[a,f,*] Kei Kitahara,[b,c] Goutam Chowdhury,[a] Debmalya Mitra,[a] Asish Kumar Mukhopadhyay,[a] Alok Kumar Deb,[a] Pallavi Indwar,[a] Biswanath Sharma Sarkar,[e] Sandip Samanta,[d] Basilua Andre Muzembo,[c] Ayumu Ohno,[b,c] Shin-ichi Miyoshi,[b,c] and Shanta Dutta[a]

[a]ICMR - National Institute for Research in Bacterial Infections, Kolkata, West Bengal, India
[b]Collaborative Research Centre of Okayama University for Infectious Diseases at ICMR-NIRBI, Kolkata, West Bengal, India
[c]Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, Okayama, Japan
[d]Dr. B C Roy Post Graduate Institute of Paediatric Sciences, Kolkata, West Bengal, India
[e]Infectious Diseases & Beleghata General Hospital, Kolkata, West Bengal, India

The Lancet Regional Health - Southeast Asia 2024;31: 100510

Published Online xxx

https://doi.org/10.1016/j.lansea.2024.100510

Summary

Background Cholera cases have increased globally across the Eastern Mediterranean, Africa, Southeast Asia, and parts of Europe since early 2024. This study aims to identify cholera hotspots and understand the spatial distribution of cholera in Kolkata and surrounding regions, a key cholera reservoir. Additionally, we examine sociodemographic factors and aspects related to water, sanitation, and hygiene (WASH).

Methods Cholera clusters were detected using kernel density estimation and spatial autocorrelation through Global Moran's-I statistics, with local cluster patterns examined using Local Moran's-I statistics. Cholera cases from August 2021 to December 2023, treated at two tertiary care facilities in Kolkata: Infectious Diseases and Beleghata General Hospital and Dr. B C Roy Post Graduate Institute of Paediatric Sciences Hospital were included. Additionally, through a case–control study, 196 culture-confirmed cholera cases and 764 age/sex-matched neighborhood controls were enrolled, to investigate cholera risk factors.

Findings Spatial analysis revealed a concentration of 196 cholera cases in Kolkata and its surrounding regions of Howrah, Hooghly, and North and South 24 Parganas. Hotspot analysis showed significant clustering in several Kolkata wards (31, 33, 56, 46, 57, 58, 59, 61, 66, 71, and 107), particularly in the northern, central, and east Kolkata wetlands areas (Global Moran's I statistic = 0.14, $p < 0.001$). These clusters had proximity between cases, with a median distance of 187.7 m, and 25.5% of cases as close as 73.9 m apart, suggesting localized transmission. Hotspots were identified with an average distance of 1600 m between them. Local Moran's I analysis found dense "high–high" clusters in these areas ($p < 0.01$), with a mean Moran's I index of 0.3, (range 0.1–4.6). The case–control study revealed that males were more likely to contract cholera, with an adjusted odds ratio of 2.4 ($p < 0.01$). There was no significant association found between cholera infection and sociodemographic factors or various WASH practices.

Interpretation The findings emphasize the importance of targeted interventions, especially in identified hotspots, to mitigate cholera transmission. Addressing Socio-economic, and environmental factors especially improvement in WASH practices may further enhance prevention effects.

Funding The author KK, received funding from the program of the Japan Initiative for Global Research Network on Infectious Diseases, (grant id: JP23wm0125004), from the Ministry of Education, Culture, Sports, Science and Technology in Japan, and Japan Agency for Medical Research and Development.

Keywords: Priority areas for multisectoral interventions; PAMIs; Cluster; Spatial analysis; Hotspot; Cholera

*Corresponding author. ICMR - National Institute for Research in Bacterial Infections, P-33, C.I.T. Road, Scheme XM, Beliaghata, Kolkata 700010, India.

E-mail address: sumankanungo@gmail.com (S. Kanungo).

[f]Equal contributor.

Research in context

Evidence before this study

In 2022, the World Health Organization (WHO) reported a global increase in cholera cases. Without treatment, cholera has a fatality rate of up to 50%, but with proper management, the rate drops below 1%. The disease thrives in regions with limited or no access to safe drinking water and basic sanitation, highlighting poor socioeconomic conditions and negatively impacting trade and tourism. In this study, we searched PubMed with keywords such as "cholera" and "distribution" or "cluster" or "hotspot" and "Kolkata" or "West Bengal" or "India" to identify existing evidence on the distribution, hotspots, and spatial clustering of cholera cases. Most studies focused on estimating the burden using outbreak data, revealing a research gap due to the lack of precise cholera burden data, complicated by underreporting. Notably, only one study was found that presented the spatial distribution of cholera cases using hospitalized patient data, but this was in southern India. Another study employed a machine learning approach to demonstrate the spatio-temporal patterns of cholera in coastal India using similar outbreak-focused data. Two additional studies examined the seasonal variability of cholera in West Bengal and its associations with environmental determinants.

Added value of this study

To the best of our knowledge, this is the first cluster analysis depicting the cholera hotspots in Kolkata and its surrounding districts of West Bengal utilizing confirmed cholera cases admitted in two major tertiary care hospitals from August 2021 to December 2023. The clustering of cases was shown at the level of 144 administrative units of the Kolkata metropolitan corporation (urban wards). The densest cluster was identified in Ward number 59 of Kolkata through kernel density mapping. Optimized hotspot analysis revealed hotspot areas with 95% confidence where enhanced cholera control efforts can be prioritized. Further, we highlighted the localized nature of outbreaks (median distance of 187 m between cases), strengthening the case for targeted interventions in high-risk areas.

Implications of all the available evidence

This study highlights the value of spatial analysis in disease surveillance and strengthens the case for wider adoption to enable outbreak detection and enhance public health preparedness. By geographically focusing cholera vaccination campaigns and resource allocation to the identified high-risk areas, public health authorities can optimize resource use and potentially achieve better cholera control in the regions of Kolkata and its surrounding districts.

Introduction

Cholera continues to pose a significant public health risk in communities lacking safe drinking water, sanitation, and hygiene (WASH).[1] Despite numerous collaborative efforts by low- and middle-income countries to enhance WASH practices for cholera containment, lack of precise reporting and disease surveillance has persistently led to inaccurate estimates of the cholera burden.[2–4] In India, since 2004, the Integrated Disease Surveillance Program (IDSP) has been operational in conducting surveillance and response activities for outbreak-prone diseases, including cholera.[5,6] Despite years of multi-disciplinary research on cholera, understanding its burden and geographical variability across India remains poor. This is due to limited laboratory capacity, excessive dependence on outbreak-based reporting, vague case definitions, and underreporting of cases driven by concerns about negative impacts on trade, tourism, and commerce.[6,7] An analysis of disease surveillance data from India reported a consistent increase in documented cholera outbreaks nationwide. From 1997 to 2006, a total of 68 cholera outbreaks were reported, a figure that went up to 559 between 2009 and 2017.[4,8]

The Gangetic belt of Asia is known to be endemic for cholera. A study conducted in the urban slums of Kolkata during 2003–2004 estimated the burden of cholera through prospective acute, watery, non-bloody diarrheal surveillance set up in five health outposts and two main tertiary referral centers for diarrhea in Kolkata. The surveillance covered a total population of 62,329 individuals and 3386 diarrheal episodes were recorded including repeat visits. Stool culture was conducted for all consenting suspected cholera patients. Among the diarrheal episodes, 4% of culture-confirmed cholera cases were reported.[9] Global enteric multicenter study (GEMS) conducted from 2007 to 2011 reported that the incidence of cholera among children under five in Kolkata was 1.8 per 100 child years.[10] Cholera in Kolkata has increased through outbreaks caused by different serotypes of cholera. These variations include both O1 and O139 serogroups, as well as classical and El Tor biotypes within the O1 group.[11]

The World Health Organization (WHO) -Global Task Force for Cholera Control (GTFCC) charts out roadmaps for countries to reduce cholera-attributable mortality by 90% and eliminate cholera in at least 20 countries by 2030.[12] A comprehensive context-specific tailor-made national action plan for countries is required to achieve this. Drafting such a national action plan, which will act as an operational document necessitates the identification of cholera hotspot areas. In line with this, the GTFCC roadmap also talks about targeted multisectoral interventions, including disease surveillance to the identification of Priority Areas for Multi-sectoral Interventions (PAMIs, or cholera hotspots),[13] targeted implementation of oral cholera vaccine, enhancement in WASH practices, and improvement in case management.[14]

Kolkata, often referred to as the 'homeland of cholera,' has experienced numerous outbreaks in the past.[15–17] Investigations into these outbreaks have consistently highlighted inadequate access to safe drinking water and sanitation facilities, particularly among low-income communities, as a common risk factor for cholera in the region.[9,11] This issue is further exacerbated by the fact that around 62% of the city's population resides in slum areas, where such challenges are more prevalent, creating an environment conducive to the persistence of cholera.[16] In this study, we aimed to identify cholera hotspots and uncover spatial clustering patterns by analyzing culture-confirmed cholera cases treated at two tertiary state medical colleges and referral hospitals, Beleghata Infectious Diseases and Beleghata General (I.D. and B.G.) Hospital and Dr. B C Roy Post Graduate Institute of Pediatric Sciences, in Kolkata, India, between August 2021 and December 2023. These two Hospitals, being specialized centers with a long history of treating patients with diarrheal diseases, including cholera, serve as the main referral facility for infectious diseases in the region, making it an ideal site for cholera surveillance.[9,18] Additionally, we investigated sociodemographic and Water, Sanitation, and Hygiene (WASH) factors that may be associated with infection through a case–control analysis."

Methods

The study was registered and granted ethical clearance from the institutional ethics committee of the Indian Council of Medical Research—National Institute for Research in Bacterial Infections (ICMR-NIRBI, earlier known as ICMR—NICED), Kolkata bearing IEC no. A-1/2019-IEC dated 18th October 2019. All participants provided informed consent for inclusion in the study and for the use of their anonymized data in future scientific research before being enrolled in the surveillance. The cluster analysis was conducted using anonymized surveillance data, ensuring confidentiality. In the case–control study, separate written and verbal consents were obtained from both cases and controls before participation. Verbal consent was used initially to explain the study details to participants, ensuring they understood the purpose, procedures, risks, and benefits of the research. Written consent served as a formal agreement, documenting the participant's informed decision to participate. The study was conducted following the Declaration of Helsinki. Patients or the public were not involved in the conceptualization, conduct, reporting, or dissemination plans of this research. Further the funders also had no role in the study design, data collection, data analysis, interpretation, or writing of the report.

Study design and participants

Acute diarrhea cases with a laboratory-confirmed diagnosis of cholera referred and admitted for treatment and management at Infectious Diseases and Beleghata General (I.D. and B.G.) Hospital and Dr. B C Roy Post Graduate Institute of Paediatric Sciences, Kolkata, India between August 2021 to December 2023 were enrolled upon verbal and written consent in the study. The definition of Acute watery diarrhea was adopted from GTFCC guidelines, as an illness, lasting for less than seven days, with patients suffering from three or more episodes of loose non-bloody watery stools with or without mucous in 24 h.[19]

The two tertiary state medical colleges and referral care hospitals, situated in the nearby vicinity primarily serve the surrounding districts of Kolkata, as well as parts of North and South 24 Parganas, Howrah, and Hooghly districts in West Bengal. For more than two decades, the Indian Council of Medical Research–National Institute for Research in Bacterial Infections (ICMR-NIRBI, earlier known as ICMR-NICED), has been conducting diarrheal disease surveillance within the I.D. and B.G. Hospital campus, where its new facility is situated.[15] Stool samples are generally collected and sent to ICMR–NIRBI for laboratory confirmation. Once diagnosed as cholera positive for any of the strains (O1, and O139), the patients were enrolled in the studies as cases. A semi-structured questionnaire consisting of sociodemographic details, WASH condition, and practices were collected from the enrolled patients in subsequent interviews and follow-up visits at their homes. Global Positioning System (GPS) coordinates were collected for each patient at the follow-up visits. To assess the effects of the determinants like–WASH conditions, practices, and sociodemographic characteristics on the occurrence of cholera, we enrolled neighborhood controls to the identified cases in a 1:4 ratio.

Study settings

Kolkata, a major metropolitan city of eastern India, the capital of the state of West Bengal, is also deemed an artistic and cultural hub of India.[20] With a population of 14.8 million and over 1480 square kilometers, Kolkata municipal corporation is divided into 144 municipal wards as administrative divisions.[21] Further Kolkata is subdivided into sixteen boroughs, each comprising nine to twelve wards.[22] Cholera, while is mostly seen in urban slums, the metropolitan Kolkata homes to many unregistered and registered slum areas in and across the city. One-third of the cities' population stays in slum areas. An urban slum is defined as an urban settlement where habitats are unsuitable for human living due to factors such as dilapidation, overcrowding, poor building design, narrow or poorly arranged streets, inadequate ventilation, lighting, or sanitation, or a combination of these conditions, which pose risks to safety and health.[23] The public healthcare system of the city is serviced through tertiary state medical colleges, subdivisional/district hospitals, and urban municipality corporation health centers.[24] I.D and B.G. Hospital typically specializes in the diagnosis, treatment, and management of various infectious diseases and took a front spot during the pandemic. Cholera has been historically prevalent across the Gangetic belt of Asia. The initial cholera pandemic emerged in the Ganges delta region between 1817 and 1823, extending beyond Asia to European territories along trade routes.[25,26] Following six prior pandemics, cholera has left its mark across America, Africa, and beyond. The seventh and most recent pandemic originated in Indonesia, with the majority of the cases ultimately occurring in Africa.[25] Currently, according to IDSP India, between 2009 and 2017 a total of 559 cholera outbreaks were identified in India. The states of Karnataka, West Bengal, Odisha, Delhi, and Madhya Pradesh have been the most vulnerable in terms of cholera disease incidence.[8,17,27] A nationwide study carried out by ICMR—NIRBI aimed to determine the seroincidence of Cholera in India. The results revealed that 11.7% of the population had experienced a seroincident cholera infection nationwide, with the figure rising to 13.2% in the Western region, which includes West Bengal.[28] Thereby, in this study an attempt was made to identify the cholera hotspots and to understand the spatial clustering of cholera cases across the geographic regions catered by the two tertiary care referral centers in Kolkata, India. Furthermore, ICMR–NIRBI, a renowned institution with a five-decade history of research on cholera and WASH practices, is located within the hospital premises.[29]

Sample size

Assuming an odd of cholera of two among cases with worse WASH conditions compared to those with better WASH conditions, a hypothesized cholera prevalence among controls of 10%, study precision of 5% a minimum sample of 180 (cases), and 720 (controls) needs to be selected.[30] Sample size was calculated in OpenEpi web Version 3.01 Updated 2013/04/06.[31] During the period of the two-year systematic surveillance at the two referral medical college and hospital (I.D. and B.G. and Dr. B C Roy Post Graduate Institute of Pediatric Sciences) a total of 196 participants were diagnosed as culture-confirmed for one of the strains of cholera. These 196 cases were taken for the spatial analysis and were also enrolled as cases in the study. Subsequently, 764 controls were selected.

We plotted the cases occurring from–January to February non-cholera seasons, March–June as summer cholera seasons, and July–December as Monsoon cholera season. This distribution is followed from the methodology charted by Shackleton et al., 2023 where the seasonality was defined based on patterns of occurrence of cholera cases rather than meteorological patterns.[32]

Case and control definitions, analysis and laboratory diagnosis

Primary subjects were all the cases confirmed as culture positive for cholera during the period of the study. We used a combination of age/sex-matched family controls and neighborhood controls to ensure appropriate matching. The primary approach was to select family controls who were age and sex-matched and had not experienced diarrhea in the last two weeks. However, if suitable family controls were not available, we enrolled neighborhood controls to maintain the age/sex-matching criteria. Case-control recruitment was done following a 1:4 ratio. Stool samples were collected from the control after consenting. Logistic regression was conducted to find if the odds of cholera case occurrence have any association with the sociodemographic factors (Gender, education, income, residential area), and different WASH practices (drinking water source used, toilet type, hand washing before food intake) across cases and controls. Statistical analysis was conducted using the software SPSS (IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp).

We included only laboratory-confirmed cholera cases to ensure high specificity in identifying true cholera cases. This reduces the potential for misclassification and provides robust data for accurate spatial clustering analysis. Moreover, the capacity for culture confirmation was readily available (Through ICMR-NIRBI) at the selected referral hospitals, which allowed us to maintain the integrity of the case identification.

Laboratory methodology

Stool or rectal swab samples were collected from diarrhea patients admitted to the I.D. and B.G. Hospital or Dr. B.C. Roy Post Graduate Institute of Paediatric Sciences before antibiotic treatment. Samples were enriched in alkaline peptone water (APW) and cultured on thiosulphate citrate bile-salts sucrose (TCBS) agar to isolate *Vibrio cholerae*. Yellow colonies were sub-cultured on Luria Bertani agar and confirmed via oxidase testing. Oxidase-positive colonies were further tested on triple sugar iron agar, and serotyping was conducted using *V. cholerae* O1 polyvalent, O139, and Ogawa/Inaba monovalent antisera. Non-agglutinating isolates were confirmed as non-O1, and non-O139 using PCR targeting the ompW gene.

Spatial analysis

After geolocating the co-ordinates of the residence (permanent stay or where the patient has been living continuously at-least for the past year), the co-ordinates were projected in a district shape file of West Bengal in ArcGIS Pro version 3.0.3.[33] The X–Y coordinates were incorporated into a shape file of study area map as point data. Fig. 1 shows the study area map and the distribution of the individual cases across the districts of

North 24 Parganas, South 24 Parganas, Howrah and Hooghly. A Kolkata municipal corporation ward boundary map was imported from ESRI web content and the point features (individual case data) were loaded as a layer in it before spatial analysis.[34] A Universal Transverse Mercator (UTM) projection was used to integrate the point features in the study map. The UTM projection system consists of 60 zones of 6° each making a complete 360°. This projection system is mostly used in cluster analysis as it minimizes the distortion.[35]

The analysis was conducted in three steps: 1. Kernel density mapping in ArcGIS to elucidate the hotspot zones, 2. Spatial autocorrelation/Global Moran's I test to understand if there is any overall clustering in the data, and cluster and outlier analysis (Local Moran's I) or Local Indicator for Spatial Association (LISA) to detect local clustering, 3. Optimized hotspot analysis using Getis-Ord Gi statistics.

Kernel density mapping

A kernel density map uses raster cells to represent geographic areas in the study region. Kernel density estimation calculates the density of cholera cases around each cell by fitting smooth surfaces over data points, decreasing in value with distance. The sum of these surfaces within each cell gives a visual representation of cholera incidence density, with higher values indicating greater case concentrations and lower values showing fewer incidents.[36] We applied the default search radius based on Silverman's Rule of Thumb, which adapts to the spatial variability of the dataset, including outliers. This method finds an optimal balance between capturing high-density clusters and accounting for distant incident cases, ensuring an accurate representation of cholera case distribution across both dense areas and sparsely populated regions in the study area. The default search radius computed was: 1487.6 m, which aligns well with previous literature where a distance around 1000 m, deemed as walkable distance has been used for kernel density estimation.[37–39]

Identifying cholera hotspots

Global Moran's I statistics provided the presence of overall clustering within the incident cholera data. It estimates the level of spatial autocorrelation within a dataset, indicating how neighboring observations relate to each other across space and if the spatial distribution is random or not.[40] Like other correlation coefficients, its range spans from –1 to 1. At p value < 0.05 a value of –1 signifies a complete clustering of dissimilar values, 0 denotes no spatial autocorrelation, and +1 indicates a perfect clustering of similar values.[41,42] Getis ord general Gi* further can identify if the cluster patterns have high or low values.[43,44] Unlike Moran's I, which assesses the

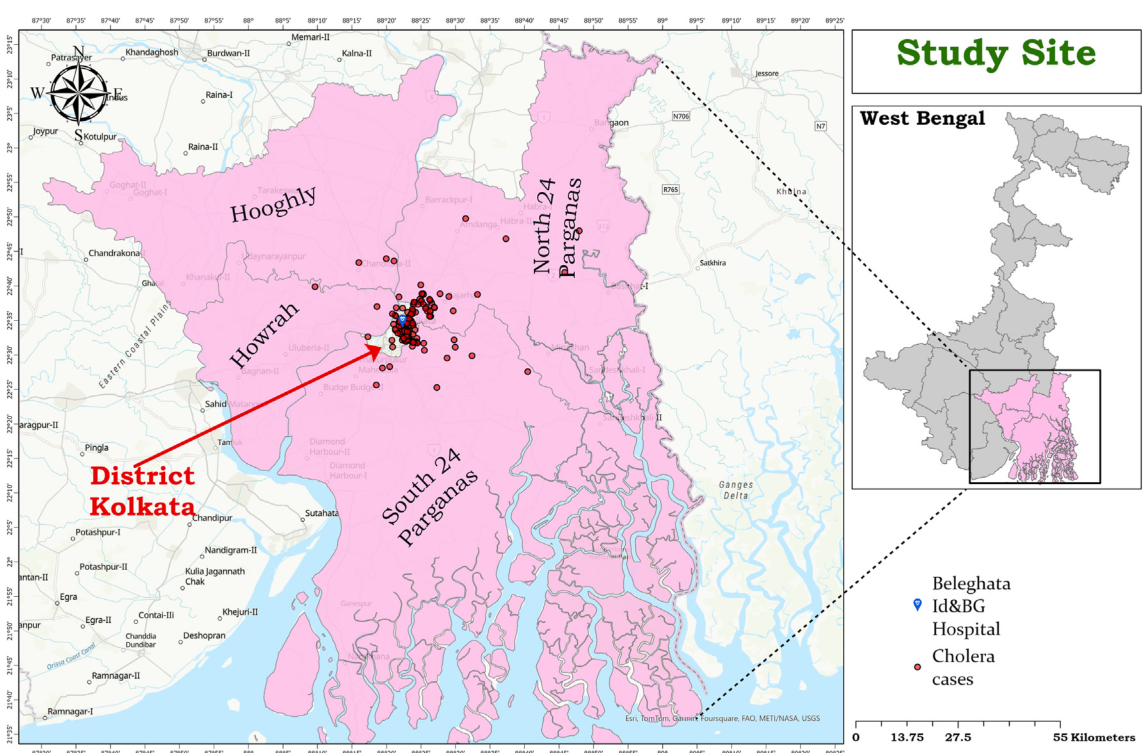

Fig. 1: Distribution of confirmed cholera cases across Hooghly, Howrah, Kolkata, and North, South 24 Parganas districts of West Bengal, referred to I.D and B.G. hospital during the period of August, 2021–December 2023.

overall spatial autocorrelation, Gi* focuses on identifying specific locations that exhibit significantly different values from their neighbors. The Gi* statistics identifies areas with clusters of high or low values within a dataset. A higher positive Z score indicates a more intense clustering of high values (hot pots), while a lower negative Z score signifies a more intense clustering of low values (cold spots).[45,46]

After importing and projecting the point coordinates of the cases as a layer in the Kolkata wards shape file, the point coordinates were integrated with an X–Y tolerance of 30 m. The x-y tolerance refers to the minimum distance between coordinates before they are considered to be at the same location/area.[47] We used 30 m, except for the same household cases, the minimum distance between cases in this study was 13 m. Considering the presence of cases within the same households and the spread of other cases (median distance of 183 m), a tolerance between 15 and 30 m was more suitable. Spatial autocorrelation and other hotspot analyses were conducted on these collected events (integrated point features based on X–Y tolerance). This procedure potentially reduces the overestimation of hotspot areas which is a common phenomenon while analyzing point feature data.[47,48] The collect event procedure further creates weighted data points and a unique feature class containing a field with the sum of all incident cases at each unique location.[47,49–52]

After estimating Global Moran's I, which tells us the presence or absence of spatial clustering, we conducted the cluster and outlier analysis through Local Indicator for Spatial Association (LISA)/Anselin local Moran's I statistics to know where the cluster of cholera cases has occurred.[52,53] Local moran's I can identify the clusters and outliers and categorize them into four types—High–high (HH–the high intensity of cases surrounded by similar high intense incident cases); High-low (HL–high cases surrounded by low number of cases) and similarly Low-high (LH), and Low–low (LL) patterns. High-low and Low-high are also described as outliers.[53,54]

Role of funding source

The funder has no role in the conceptualization, design, data collection, analysis, decision to publish, or preparation of the manuscript.

Results

Out of 3995 cases enrolled through hospital-based systematic surveillance at Infectious Diseases and Beleghata General Hospital and patients attending diarrhea treatment unit in Dr. B C Roy Post Graduate Institute of Paediatric Sciences Hospital, 2790 cases were non-pathogenic. Bacterial infections were found in 836 cases, 340 had viral and 96 cases had parasitic infections. Some cases had overlapping involvement of bacteria, viruses, and/or parasites. Out of the cases with bacterial infections, 196 had culture-confirmed cholera, and 33 cases were due to non-cholera vibrio infections. Adenovirus, norovirus, astrovirus, and rotavirus were found among the 340 viral diarrheal episodes. Among the bacterial cases (except cholera and non-cholera vibrio infections) *Aeromonas*, Enteroaggregative *Escherichia coli*, Enteropathogenic *E. coli*, Enterotoxigenic *E. coli* (ETEC), campylobacter, *Salmonella*, and *Shigella* were mainly found. In this analysis, 196 culture-confirmed cholera cases were enrolled from August 2021 to December 2023. Most cholera cases (176/196; 89.8%) were enrolled from I.D and B.G. hospital located in ward 33 of the Kolkata municipal corporation. Fig. 1 describes the study area map and the spread of the individual cholera cases across the districts of North and South 24 Parganas, Kolkata, Hooghly, and Howrah. Beleghata Id being situated in the district of Kolkata (ward no. 33), most cases seemed to be concentrated in the districts of Kolkata.

The proximity of the cases was assessed. The median distance between the cases was 187.7 m, whereas 25.5% of the cases were located within proximity of 73.9 m (quartile 1) of each other. The minimum distance of cases found was zero meters. Six such cases were found, and those were cases within the same households. Fig. 2A contains distance analysis between cholera cases. We plotted the cholera cases across different cholera seasons from January to December as described in the methods section. As evident from Fig. 3 (cholera cases from 2021 to 23 plotted across months), most of the cholera cases were concentrated in the summer seasons, followed by another second peak in the monsoon seasons.

Characteristics of the study population

Among the 196 cases 68.2% were men, and among the 764 controls recruited 48.5% were women respectively. The mean (standard deviation/SD) age of the participants with confirmed cholera diagnosis was 27.9 (20.0) and among control, it was 30.0 (16.3) years. Supplementary Table S1 and Fig. 4 describe the characteristics of case and control participants. We found males had higher odds of contracting cholera than women in our dataset with an unadjusted odd of 2.0 at p-value < 0.01. When adjusted against the residential area, income quartile, education, and the different WASH factors the adjusted Odds (confidence interval/95% CI) became 2.4 (95% CI 1.6, 3.5) at p-value < 0.01.

No statistically significant association was observed between the level of education and cholera cases. Participants with lower levels of education, including those who were illiterate (unadjusted OR 1.0, 95% CI 0.4, 2.9) and those with secondary or lower education (unadjusted OR = 1.0, 95% CI 0.4, 2.8), did not exhibit notably different odds of cholera when compared to participants with a master's degree. No significant difference in

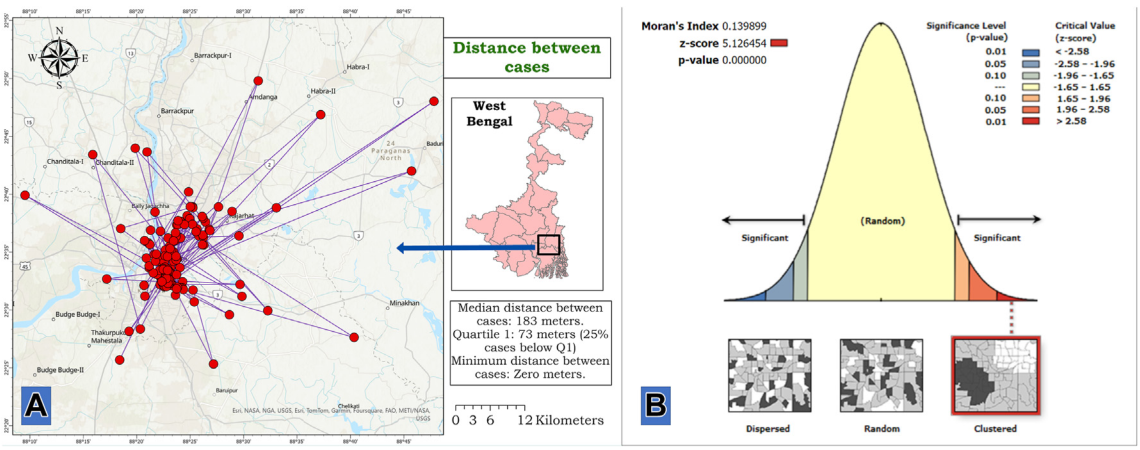

Fig. 2: Distance/proximity analysis of the cases (A) and Spatial autocorrelation (Global Moran's I index) (B).

cholera odds was observed across income quartiles compared to the lowest income group (<10,000 INR). Odds ratios for higher income quartiles ranged from 0.9 to 1.0, (95% CI 0.5, 1.7) and p-values above 0.05, indicating no statistically significant relationship. The analysis of drinking and household water sources revealed no significant association between cholera and different water sources. Also, neither hand hygiene practices before cooking or eating nor the use of toilet type showed a statistically significant association with cholera cases in this analysis. Though non-significant, those who "sometimes" or "never" washed their hands before cooking had slightly higher odds of cholera compared to those who "always" washed their hands (unadjusted OR = 1.04 and 1.02, respectively). Similarly, for toilet use, participants who used flush toilets (either to septic systems, pits, or piped sewers) had lower odds of cholera compared to those who practiced open defecation. The odds ratios for the different types of flush toilets were around 0.5, which indicates a potential trend towards lower cholera cases among those using improved sanitation, even though the confidence intervals were wide and the results were not statistically significant. The detailed regression result is given in Supplementary Table S1.

Kernel density mapping

A smoothed representation of data from a list of cholera incidence points was produced by using the kernel density tool in ArcGIS Pro (ESRI, 2021). This improves the visual clarity and comprehension. This tool uses a kernel function to create smoothly tapered surfaces

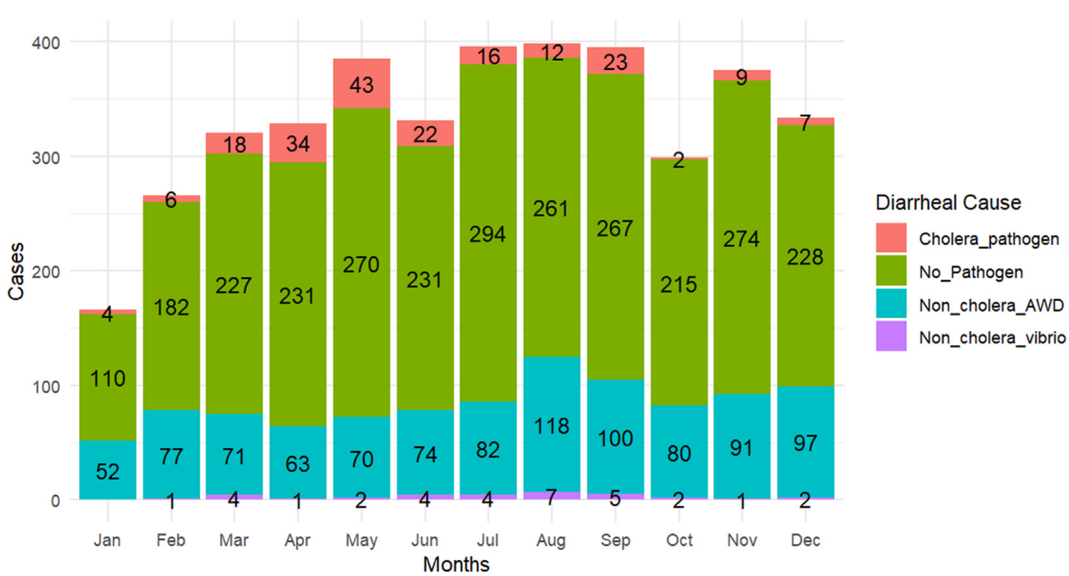

Fig. 3: Cholera cases across different months from January to December (2021–2023) in summer and monsoon seasons.

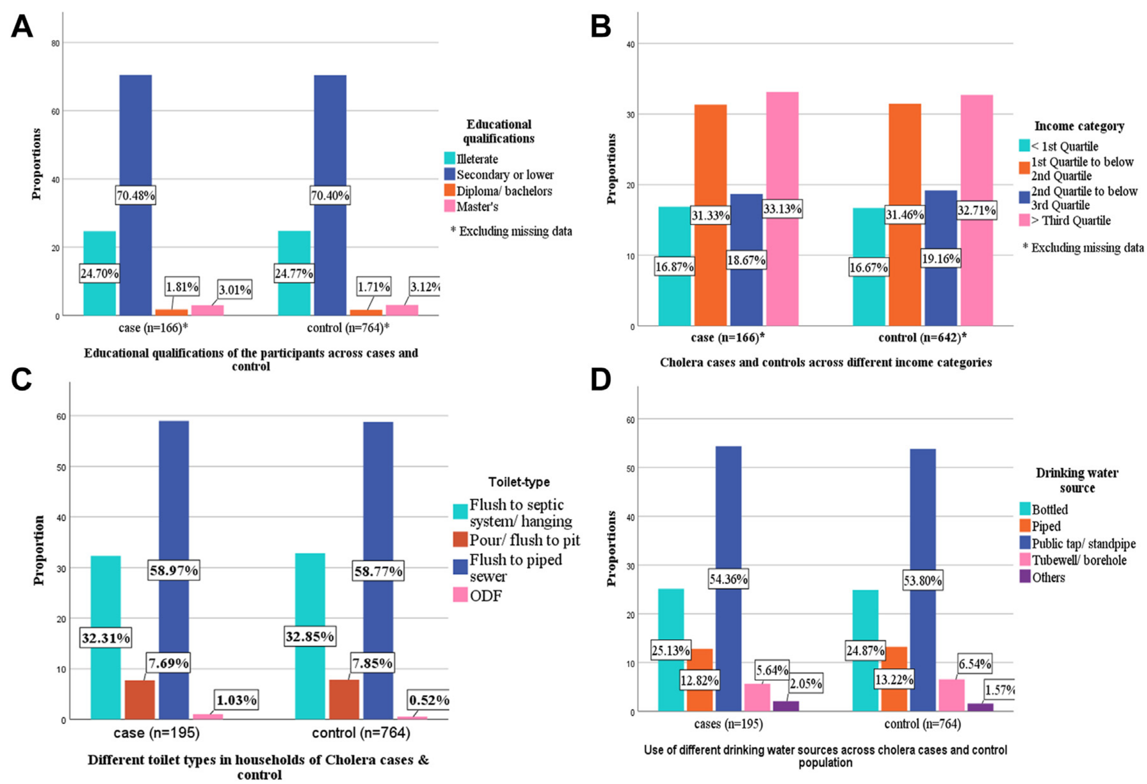

Fig. 4: Participant characteristics and WASH practices description across cases and controls: Educational qualification - (A); Income categories - (B); Use of different Toilet types - (C); Use of different drinking water sources - (D).

around each cholera case, by computing the magnitude per unit area based on the cholera incidence locations. Fig. 5 describes a medium to high intensity clusters noted across ward numbers 31, 36, 56, 58, 59, 64, 65, and 66. These areas incorporate parts of the Northern and central Kolkata region, east Kolkata wetlands east Kolkata wetlands, Gobra, Tangra, Topsia, etc.

Hotspot analysis

Fig. 5 describes the spatial autocorrelation report of the cholera incident cases across the four districts of West Bengal. A global Moran's statistics of 0.14, at p-value < 0.001, and a high z score of 5.1 indicates the presence of clustering (Fig. 2B). While conducting the optimized hotspot analysis, the Gi Z score ranged from −0.9 to 4.8 (p-value < 0.001). Areas designated as hotspots with 99% confidence had a mean (SD) Gi Z score of 3.8 (0.3). The highest Gi values were observed through wards 31, 33, 56, 46, 57, 58, 59, 61, 66, 71, and 107 of the Kolkata municipal corporation. These wards are mostly located in boroughs 6,7 and 8. Fig. 5 describes the hotspots with 95% and 99% confidence across the wards of Kolkata. Local Moran's I statistics showed areas of HH cluster with Moran index >0, at p-value < 0.01 (Fig. 6). The mean (SD) local Moran index of HH clusters was 0.3 (0.5) at p-value < 0.01. Corresponding local Moran index Z score ranged from 0.1 to as high as up to 4.6. Fig. 6 describes the clusters identified using Local Moran's I statistics. In optimized hotspot analysis the mean (SD) intercluster distance between clusters with 99% confidence found was 1600 (0.99) meters (Fig. 5). The median distance was 1300 m, with 25% of clusters falling within 852 m of each other.

Discussion

This study aimed to identify cholera hotspots and to understand the spatial clustering pattern among the cholera cases referred to I.D and B.G. hospital and Dr. B C Roy Post Graduate Institute of Paediatric Sciences Hospital, Kolkata from August 2021 to December 2023. We utilized different spatial cluster analysis tools. An assumption that this analysis followed was, that cholera transmission occurs at the household level or within the communities.[55,56] Global Moran's I indicated significant spatial clusters across the urban slums in districts of Howrah, Hooghly, North and South 24 Parganas. Cholera hotspots were identified as areas with the most significant clustering in the district of Kolkata across areas in KMC wards 29, 46, 48, 54, 55, 56, 58, 59, 107, 108, etc (Figs. 5 and 6). These urban wards are mostly situated in the Northern and central Kolkata region, east

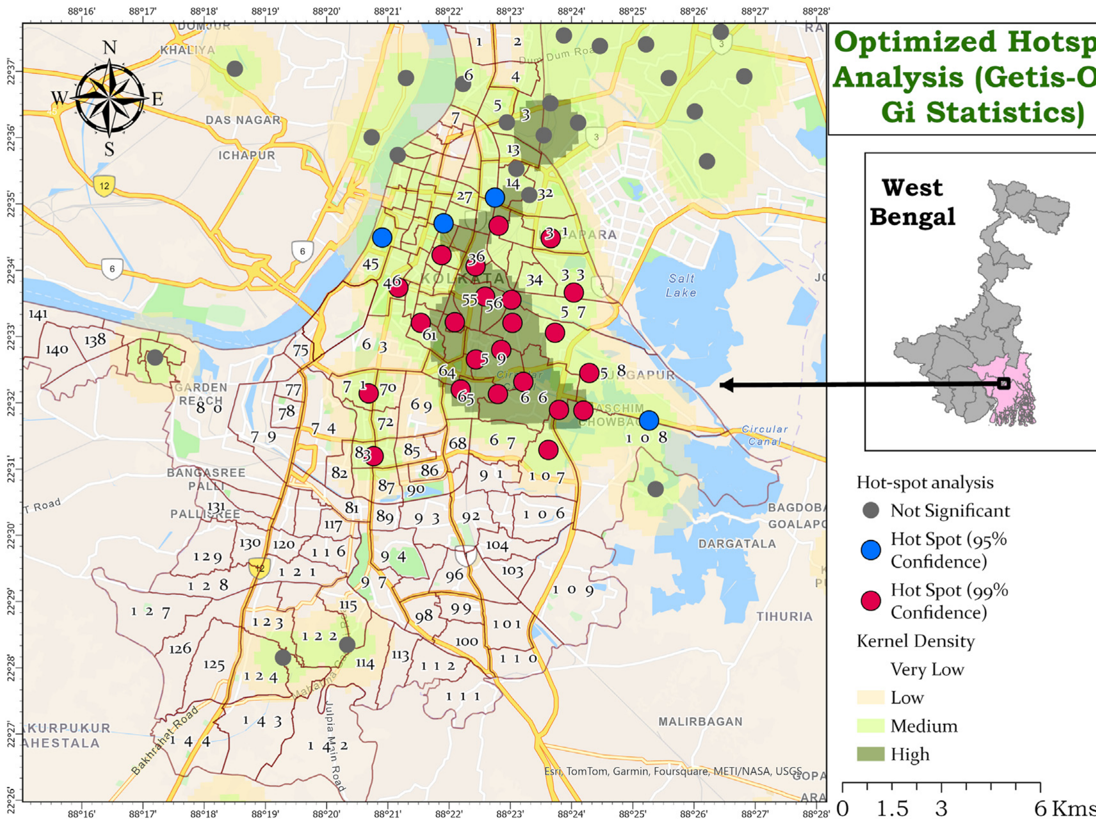

Fig. 5: Optimized hotspot analysis and Kernel density-based clustering showing confirmed hotspots with high Gi Z score values (Red circle—99% confidence, mean Gi Z score 3.89, standard deviation 0.35, p value < 0.001).

Kolkata wetlands, etc., and were spread across boroughs 6,7, and 8. Mukherjee et al., 2022, geospatially represented the population density and urban slums in the different urban wards in Kolkata. They showed that wards 29, 46, 54, 55, 56, 58, 59, 107, and 108 have a higher density of urban slums and population density in the region. Cholera clusters were also found in these urban wards.[57]

Most cases (25%) were in proximity to each other and were within less than 73 m. Except for five cases within the same household, the closest distance of cases was 13 m from each other. The median distance was 187 m in our study sample. A similar study conducted by Rokosky et al. in Kathmandu Valley, Nepal assessed cholera cases that occurred during an outbreak. They reported that the median average distance between each of the cases was 790 m, and the cluster was defined using a 1000-m radius of cholera incidence.[58] The Proximal density in our analysis was much smaller than in the Kathmandu study. This may be attributable to the general endemic nature of cholera in the Gangetic belt, more dense population (24,000 vs. 20,000 per square kilometer[59]). Moreover, among the cholera cases in our study, most (45.2% and 33.1%) were from urban non-slum and urban slum areas. Cases from rural areas were the least (21.7%). The seasonality of cholera cases with two prominent peaks found in our data (March–June: summer cases; July–December: Monsoon cases) was similar to the trends found through analysis of 21 years of cholera cases from the I.D and B.G hospital, Kolkata between 1999 and 2019, by a study done by Shackleton et al., 2023 (Fig. 3).[32]

A study by Osei et al. assessed the temporal trend and spatial interaction between cholera incident cases in Kumasi-Ghana during the 2005 outbreaks. They reported that maximum significant interaction was noted within 1000 m of the epidemic center, and most interactions were noted within 300 m.[38] Comparable estimates can be seen in our study with 25% of the cases being situated within 73 m from each other, and 50% being within 183 m. These presentations hint towards local and household-level transmission. Spread beyond primary transmission is important in these proximity scenarios.[38] Further in our study 25% of the high confidence (99%) clusters were within 852 m of each other indicating a probable intercluster transmission (Fig. 5). The kernel density map (Figs. 5 and 6) shows a circular pattern of cholera incidence dense regions followed by

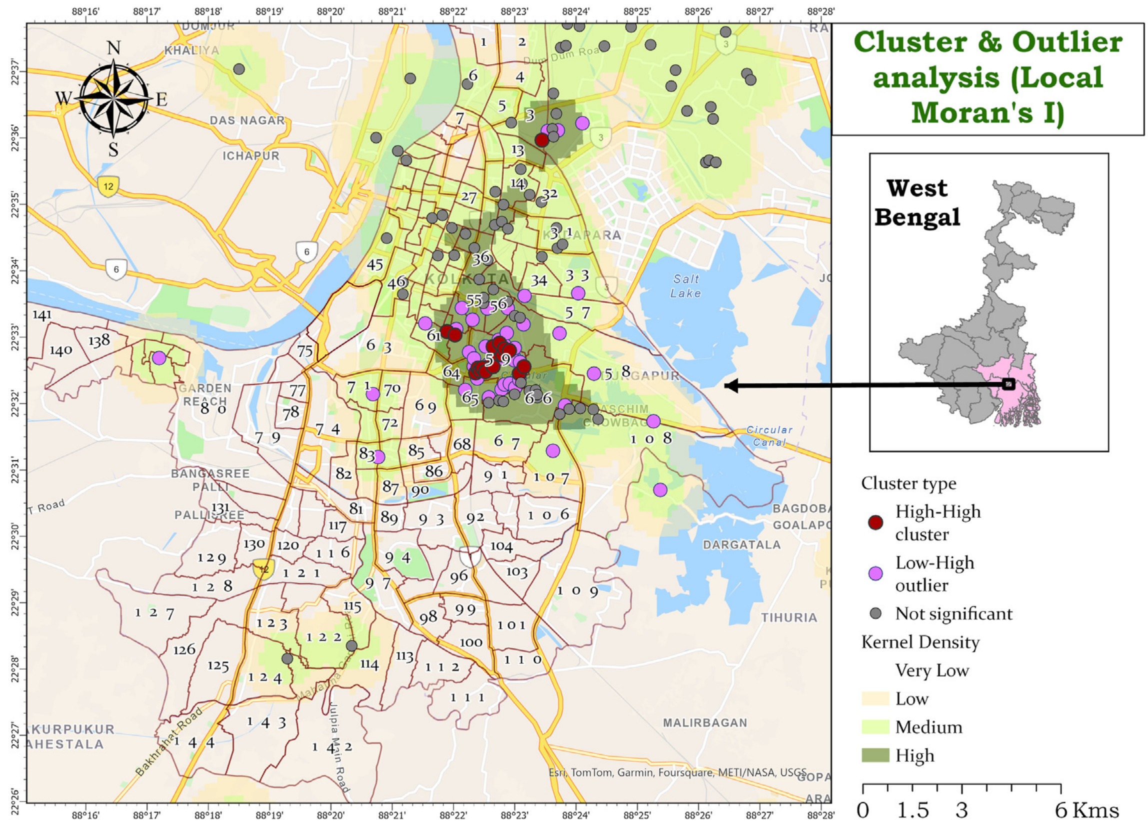

Fig. 6: Cluster and outlier analysis and Kernel density-based clustering showing areas of dense case incidences surrounded by intense clusters (red circles) across wards of Kolkata metropolitan corporation.

consecutive light intensity zones. The Kathmandu study has depicted a ring vaccination strategy to address outbreaks and deploy vaccines in emergency scenarios. The ring radius was 1 km in their study setup.[58] A similar reactive vaccination strategy can be adhered to in our study setup. A 2017 ICMR policy brief had also advised the deployment of available and licensed oral cholera vaccine (OCV), in high-priority settings (areas with frequent cholera outbreaks) on an exploratory basis.[60]

Similar clustering in our study was noted in mat lab Bangladesh. A study by Craig et al. identified cholera cluster scale was 250 m in endemic regions of Matlab, Bangladesh.[61] Similarly, Moreno et al. reported the presence of cholera clusters as small as 50 m and as large as 2340 m.[62] In our analysis, global moran's I index of 0.14 and z score of 5.1 indicated the presence of spatial cluster. Cluster outlier analysis revealed a high z score for local moran's I, and getis ord Gi* statistics, indicating clusters nearby.[52,63] A study by Shackleton et al. depicted a bi-annual pattern of cholera case incidence in Kolkata, culminating in a total of four seasonal peaks across two years.[32] Similar peak has been prevalent during our study period August 2021 to December 2023 and contributed to the dense clustering of cholera cases across the four districts of West Bengal.

Logistic regression in our case–control data revealed that males had 2.4 higher odds of contracting cholera than women in this dataset. Conflicting evidence exists on the gender distribution of cholera cases. A hospital-based study conducted in Assam, India between 2003 and 2013, revealed that cholera was predominantly present in females in most of the years except in 2003 and 2007, males had higher occurrences of the disease. In endemic regions, cholera is mostly seen among women than men, whereas in non-endemic regions it does not follow any specific trend.[64] Regarding the effects of WASH practices on cholera occurrences we found marginally increased odds of chances of cholera in practices like open defecation, using public tap/standpipe or piped water sources which generally run close to sewage systems in urban areas, for drinking or household chores, irregular hand washing, etc. but all these associations were statistically insignificant. Islam et al. conducted a study utilizing data from a 2009 oral cholera vaccine (OCV) trial in Kolkata, India.[65] The study aimed to assess if the inherent fluctuations in water, sanitation, and hygiene (WASH) conditions within an urban slum environment could predict the risk of cholera. They characterized improved WASH conditions through a composite index comprising four binary variables: access to a safe and clean water source for regular use, access to safe drinking water, use of private or shared flush toilets, and consistent handwashing with soap after defecation. The research revealed that living in households with better WASH conditions was linked to a 30% decrease in cholera risk over five years. However, when examined individually, these factors did not demonstrate significant enhancements.[65] Overall WASH practices in our study were found to fare well (Supplementary Table S1), similar to the findings in two consecutive surveys done in 2018 and 2019 among the participants of an enteric fever surveillance network across two wards of Kolkata, by Kanungo et al.[66] Adjunct to the state of WASH conditions, endemicity, environmental changes play a hefty role in cholera occurrences in and around Kolkata. A study by Mukhopadhyay et al. depicted a scenario of an increase in cholera incidence following waterlogging after the monsoon season in August 2015. A geospatial projection of cases showed that most cholera cases were surrounding logged water canals through the Kolkata municipal wards.[15]

GTFCC charts out three steps to identify PAMIs. Firstly, in the preparatory phase, data through the past five to fifteen years is collected and consolidated at a specified administrative level for National Cholera Plan (NCP) implementation, necessitating collaboration with regional authorities to obtain epidemiological and vulnerability indicators. Secondly, NCP Operational Geographic Units are scored using a numeric priority index based on four key indicators: incidence, mortality, persistence, and, if feasible, cholera test positivity. Thirdly, country stakeholders validate a final list of PAMIs by establishing a priority index threshold, with consideration given to contextual knowledge, particularly in areas where the reliability of the index is questionable. This ensures that interventions are targeted effectively in areas most affected by cholera, facilitating the comprehensive implementation of the NCP.[13,67] The strength of our study, though the analysis spanned only two years, serves as an initial stepping stone in identifying priority areas of multisectoral interventions (PAMIs).

The major limitation of our study was that the cholera incident cases spanned around 2-year periods and occurred at different time points. Though we complemented our spatial analysis with different techniques and applied robust checks in terms of X–Y tolerance of incident cases, we did not account for transmission dynamics, which are critical in understanding the spread of cholera. Economic status in our analysis independently had no significant effect on the occurrence of cholera, thereby we did not check the effect of occupation on the occurrence of cholera. We further did not present separate quantitative findings regarding the different organisms involved for all the diarrheal episodes (pathogenic—1205; non-pathogenic—2790 among 3995 total cases enrolled in the surveillance) and their distribution across space and time, as that is out of the scope of this manuscript and would be taken up as a separate exercise pertaining to the surveillance. We acknowledge that Kolkata is served by thirteen public and private medical college hospitals, and while some cases may have been missed, the majority of severe diarrheal cases are referred to I.D. and B.G. Hospital for further management and confirmation as these hospitals are specialized centers with a long history of treating patients with diarrheal diseases, including cholera, and they serve as key referral facilities for infectious diseases in the region.[9,18] Thus, we believe the two hospitals provide robust insight into cholera clustering in the region. However, we recognize the potential bias in limiting our analysis to these two facilities, as patients may also present to other clinics or might not seek treatment at all. The use of both age/sex-matched family and neighborhood controls allowed us to improve matching while reducing the risk of overestimating associations due to shared exposures in family controls. However, we acknowledge that both family and neighborhood controls may have similar socioeconomic and WASH conditions, potentially masking associations between exposures and cholera infection. Further, we did not look for determinants like immune status, or evidence of prior exposure through vibriocidal or other markers. Future efforts would be undertaken to incorporate these epidemiological checks and barriers.

Conclusion

This study identified significant spatial clustering of cholera cases in Kolkata and surrounding districts, with notable concentrations in urban slum areas. The proximity of cases, with a median distance of 187 m, highlights potential localized transmission. Utilizing spatial analysis tools, such as kernel density mapping and hotspot analysis, offers insights for targeted interventions, including vaccination strategies. Comparison with other endemic regions underscores variations in clustering patterns and suggests tailored approaches for containment. Understanding these spatial dynamics is crucial for effective cholera control and prevention efforts in densely populated areas like Kolkata. Moreover, the GTFCC encourages the identification of Priority Area Mapping Interventions (PAMIs) to advance toward a tailored national cholera action plan. Our study serves as the initial milestone in the Gangetic Delta region for pinpointing operational geographic units of cholera hotspots for the national cholera action plan and subsequently designating them as PAMIs.

Contributors

Conceptualization: SK, KK, SD; *Data curation:* SK, RT, GC, DM, AKD, PI; *Formal Analysis:* RT; *Investigation:* AKM, AKD, AO, SM, PI, SS, BSS; *Methodology:* SK, KK, RT, SD, BAM, SM; *Project administration:* SK, KK, SD, SS, BSS; *Resources:* BAM, AO, KK, SM, RT, SD; *Software:* RT, DM, KK, GC; *Supervision:* SK, KK, SD, AKD, AKM; *Validation:* RT, DM,

BAM, AO, SM, PI; ***Visualization:*** RT, SK, KK; ***Writing—original draft:*** SK, RT; ***Writing—review and editing:*** SK, RT, KK, SD, AKD, AKM.

All the contributing authors have critically reviewed the final version of the manuscript.

Data sharing statement

The data collected and analyzed in this research project are available upon reasonable request from the corresponding author through email communications. We are committed to promoting transparency and reproducibility in research and encourage interested parties to reach out for access to the data.

Editor note

The Lancet Group takes a neutral position with respect to territorial claims in published maps and institutional affiliations.

Declaration of interests

All authors declare no competing interest.

Acknowledgements

We acknowledge the field workers, and supervisors (Mr. Prasun Kumar Dey, Mr. Abhijit Guha, Mr. Sahadeb Das, Mr. Bijay Parbat, Mr. Swapan Bera, Mr. Suvankar Mondal, Mr. Biplab Jash, Mr. Debjit Das, Mr. Himangshu kr. Ran, Mr. Prabir Sinha, Ms. Tumpa Mondal, Mr. Alok Banerjee), involved in this project, who has been instrumental in field level surveillance activities throughout the project. We also want to extend our sincere gratitude to Mr. Debobrata Sadhukhan, GIS Expert - Land and Development Office, MOHUA, New Delhi, India, for sharing his invaluable technical insights on the spatial analysis methods employed in this study. Further we want to convey our gratitude to the patients and the community consenting people for their co-operation in making this project successful.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.lansea.2024.100510.

References

1 Lopez AL, Dutta S, Qadri F, et al. Cholera in selected countries in Asia. *Vaccine*. 2020;38(Suppl 1):A18–A24. https://doi.org/10.1016/j.vaccine.2019.07.035.

2 Azman AS, Lessler J, Luquero FJ, et al. Estimating cholera incidence with cross-sectional serology. *Sci Transl Med*. 2019;11(480):eaau6242. https://doi.org/10.1126/scitranslmed.aau6242/suppl_file/aau6242_sm.pdf.

3 Diep TT, Jensen O, Van Thuong N, et al. Sero-evaluation of immune responses to Vibrio cholerae in a postelimination setting. *Open Forum Infect Dis*. 2020;7(5):ofaa136. https://doi.org/10.1093/ofid/ofaa136.

4 Panda S, Chatterjee P, Deb A, et al. Preventing cholera in India: synthesizing evidences through a systematic review for policy discussion on the use of oral cholera vaccine. *Vaccine*. 2020;38(Suppl 1):A148–A156. https://doi.org/10.1016/j.vaccine.2019.07.029.

5 MoHFW, GOI. Idsp :: integrated disease surveillance programme (IDSP). n.d. Available from: https://idsp.mohfw.gov.in/. Accessed August 23, 2024.

6 Gupta SS, Ganguly NK. Opportunities and challenges for cholera control in India. *Vaccine*. 2020;38:A25–A27. https://doi.org/10.1016/j.vaccine.2019.06.032.

7 Saha GK, Ganguly NK. Spread and endemicity of cholera in India: factors beyond the numbers. *J Infect Dis*. 2021;224(Supplement_7):S710–S716. https://doi.org/10.1093/infdis/jiab436.

8 Muzembo BA, Kitahara K, Debnath A, et al. Cholera outbreaks in India, 2011–2020: a systematic review. *Int J Environ Res Public Health*. 2022;19(9):5738. https://doi.org/10.3390/ijerph19095738.

9 Sur D, Deen JL, Manna B, et al. The burden of cholera in the slums of Kolkata, India: data from a prospective, community based study. *Arch Dis Child*. 2005;90(11):1175–1181. https://doi.org/10.1136/adc.2004.071316.

10 Kotloff KL, Nataro JP, Blackwelder WC, et al. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *Lancet*. 2013;382(9888):209–222. https://doi.org/10.1016/S0140-6736(13)60844-2.

11 Mandal S. Cholera epidemic in and around Kolkata, India: endemicity and management. *Oman Med J*. 2011;26(4):288–289. https://doi.org/10.5001/omj.2011.71.

12 Global Task Force on Cholera Control (GTFCC), Surveillance Working Group. *Interim guidance document on cholera surveillance*. 2017.

13 GTFCC–WHO. *Technical guidance - identification of priority areas for multisectoral interventions (PAMIs) for cholera control*; 2023. Available from: https://www.gtfcc.org/resources/identification-of-priority-areas-for-multisectoral-interventions-pamis-for-cholera-control/. Accessed April 13, 2024.

14 Ahmed MU, Baquilod M, Deola C, et al. Cholera prevention and control in Asian countries. *BMC Proc*. 2018;12(Suppl 13):62. https://doi.org/10.1186/s12919-018-0158-1.

15 Mukhopadhyay AK, Deb AK, Chowdhury G, et al. Post-monsoon waterlogging-associated upsurge of cholera cases in and around Kolkata metropolis, 2015. *Epidemiol Infect*. 2019;147:e167. https://doi.org/10.1017/S0950268819000529.

16 Bhunia R, Ramakrishnan R, Hutin Y, et al. Cholera outbreak secondary to contaminated pipe water in an urban area, West Bengal, India, 2006. *Indian J Gastroenterol*. 2009;28(2):62–64. https://doi.org/10.1007/s12664-009-0020-5.

17 Chatterjee P, Kanungo S, Bhattacharya SK, et al. Mapping cholera outbreaks and antibiotic resistant Vibrio cholerae in India: an assessment of existing data and a scoping review of the literature. *Vaccine*. 2020;38(Suppl 1):A93–A104. https://doi.org/10.1016/j.vaccine.2019.12.003.

18 Nair GB, Ramamurthy T, Bhattacharya MK, et al. Emerging trends in the etiology of enteric pathogens as evidenced from an active surveillance of hospitalized diarrhoeal patients in Kolkata, India. *Gut Pathog*. 2010;2:4. https://doi.org/10.1186/1757-4749-2-4.

19 GTFCC, World Health Organization. *Public health surveillance for cholera, guidance document*. 2024.

20 Chitty T. *Kolkata: a city of arts, culture and cuisine*; 2018. Available from: https://www.cnbc.com/2018/11/09/kolkata-a-city-of-arts-culture-and-cuisine.html. Accessed April 6, 2024.

21 Kolkata Municipal Corporation. *Basic statistics of Kolkata*; 2016. Available from: https://www.kmcgov.in/KMCPortal/jsp/KolkataStatistics.jsp. Accessed October 14, 2024.

22 Majumdar S. A study of dengue disease data by GIS in Kolkata city: an approach to healthcare informatics. In: Suresh A, Paiva S, eds. *Deep learning and edge computing solutions for high performance computing*. Cham: Springer International Publishing; 2021:35–46. https://doi.org/10.1007/978-3-030-60265-9_3.

23 Kundu N. *"Understanding Slums: Case Studies for The Global Report on Human Settlements", The Case of Kolkata, India*. 2003.

24 Government of West Bengal. Kolkata municipal corporation. n.d. Available from: https://www.kmcgov.in/KMCPortal/jsp/BoroughCommitte.jsp. Accessed April 6, 2024.

25 *History.com editors. Cholera | epidemiology, pathophysiology, spread & history*; 2023. Available from: https://www.history.com/topics/inventions/history-of-cholera. Accessed April 14, 2024.

26 Kanungo S, Azman AS, Ramamurthy T, et al. Cholera. *Lancet*. 2022;399(10333):1429–1440. https://doi.org/10.1016/S0140-6736(22)00330-0.

27 Ali M, Sen Gupta S, Arora N, et al. Identification of burden hotspots and risk factors for cholera in India: an observational study. *PLoS One*. 2017;12(8):e0183100. https://doi.org/10.1371/journal.pone.0183100.

28 Kanungo S, Nandy RK, Talukdar R, et al. Analysis of cholera risk in India: insights from 2017–18 serosurvey data integrated with epidemiologic data and societal determinants from 2015–2019. *PLoS Neglected Trop Dis*. 2024;18(9):e0012450. https://doi.org/10.1371/journal.pntd.0012450.

29 ICMR. *National Institute of cholera and enteric diseases: profile*; 2021. Available from: https://www.niced.org.in/niced/NICED-Profile.htm#DirectorsDesk. Accessed April 13, 2024.

30 Endris AA, Tadesse M, Alemu E, et al. A case-control study to assess risk factors related to cholera outbreak in Addis Ababa, Ethiopia, July 2016. *Pan Afr Med J*. 2019;34:128. https://doi.org/10.11604/pamj.2019.34.128.17997.

31 Dean A, Sullivan K, Soe M. *OpenEpi: open source epidemiologic statistics for public health*. 2013.

32 Shackleton D, Economou T, Memon FA, et al. Seasonality of cholera in Kolkata and the influence of climate. *BMC Infect Dis*. 2023;23(1):572. https://doi.org/10.1186/s12879-023-08532-1.

33 ESRI. ArcGIS Pro: 2D, 3D & 4D GIS mapping software | esri India. n.d. Available from: https://www.esri.in/en-in/products/arcgis-pro/overview. Accessed October 14, 2024.
34 ESRI India. *Kolkata municipal corporation ward boundaries - overview*; 2022. Available from: https://www.arcgis.com/home/item.html?id=22935cc469ee42358cb0512cc0c2b734. Accessed April 7, 2024.
35 GIS-Geography. *How universal transverse mercator (UTM) works*; 2016. Available from: https://gisgeography.com/utm-universal-transverse-mercator-projection/. Accessed April 6, 2024.
36 ESRI. *How kernel density works—ArcGIS Pro | documentation*; 2021. Available from: https://pro.arcgis.com/en/pro-app/3.1/tool-reference/spatial-analyst/how-kernel-density-works.htm. Accessed April 7, 2024.
37 Madullu MT, Thomas DSK, Nyanza EC, et al. Spatial distribution of suspected and confirmed cholera cases in Mwanza City, Northern Tanzania. *PLoS Glob Public Health*. 2023;3(1):e0001261. https://doi.org/10.1371/journal.pgph.0001261.
38 Osei FB, Duker AA. Spatial dependency of V. cholera prevalence on open space refuse dumps in Kumasi, Ghana: a spatial statistical modelling. *Int J Health Geogr*. 2008;7:62. https://doi.org/10.1186/1476-072X-7-62.
39 Silverman BW. *Density estimation for statistics and data analysis*. New York: Routledge; 2017.
40 Zhang M, Gurung A, Anglewicz P, et al. Acceptance of COVID-19 vaccine among refugees in the United States. *Public Health Rep*. 2021;136(6):774–781. https://doi.org/10.1177/00333549211045838.
41 Cheruiyot K. Detecting spatial economic clusters using kernel density and global and local Moran's I analysis in Ekurhuleni metropolitan municipality, South Africa. *Reg Sci Policy Pract*. 2022;14(2):307–327. https://doi.org/10.1111/rsp3.12526.
42 ESRI. *How spatial autocorrelation (global moran's I) works—ArcGIS Pro | documentation*; 2021. Available from: https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-how-spatial-autocorrelation-moran-s-i-spatial-st.htm. Accessed April 7, 2024.
43 Mailman school of public health. *Hot spot spatial analysis | columbia public health*; 2016. Available from: https://www.publichealth.columbia.edu/research/population-health-methods/hot-spot-spatial-analysis. Accessed March 30, 2024.
44 Smith, Goodchild, Longley. *Geospatial analysis*. 6th ed. 2021.
45 Jana M, Sar N. Modeling of hotspot detection using cluster outlier analysis and Getis-Ord Gi* statistic of educational development in upper-primary level, India. *Model Earth Syst Environ*. 2016;2(2):60. https://doi.org/10.1007/s40808-016-0122-x.
46 Songchitruksa P, Zeng X. Getis–ord spatial statistics to identify hot spots by using incident management data. *Transport Res Rec*. 2010;2165(1):42–51. https://doi.org/10.3141/2165-05.
47 ESRI. XY tolerance (environment setting)—ArcGIS Pro | documentation. n.d. Available from: https://pro.arcgis.com/en/pro-app/latest/tool-reference/environment-settings/xy-tolerance.htm. Accessed April 3, 2024.
48 SSP Innovations. *Get from A to B more quickly: changing your XY tolerance*; 2013. Available from: https://sspinnovations.com/blog/get-b-more-quickly-changing-your-xy-tolerance/. Accessed October 14, 2024.
49 ESRI. *Integrate (data management)—ArcMap | documentation*; 2018. Available from: https://desktop.arcgis.com/en/arcmap/latest/tools/data-management-toolbox/integrate.htm. Accessed April 12, 2024.
50 ESRI. Collect events (spatial statistics)—ArcGIS Pro | documentation. n.d. Available from: https://pro.arcgis.com/en/pro-app/3.1/tool-reference/spatial-statistics/collect-events.htm. Accessed April 12, 2024.
51 Packt. The collect events tool - spatial analytics with ArcGIS. n.d. Available from: https://subscription.packtpub.com/book/programming/9781787122581/6/ch06lvl1sec26/the-collect-events-tool. Accessed April 12, 2024.
52 Anselin L. Local indicators of spatial association—LISA. *Geogr Anal*. 1995;27(2):93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x.
53 Osadebey M, Pedersen M, Arnold D, et al. Local indicators of spatial autocorrelation (LISA): application to blind noise-based perceptual quality metric index for magnetic resonance images. *J Imaging*. 2019;5(1):20. https://doi.org/10.3390/jimaging5010020.
54 Zhang P, Yang S, Dai S, et al. Global spreading of Omicron variant of COVID-19. *Geospatial Health*. 2022;17(s1). https://doi.org/10.4081/gh.2022.1083.
55 Sugimoto JD, Koepke AA, Kenah EE, et al. Household transmission of Vibrio cholerae in Bangladesh. *PLoS Neglected Trop Dis*. 2014;8(11):e3314. https://doi.org/10.1371/journal.pntd.0003314.
56 Phelps MD, Simonsen L, Jensen PKM. Individual and household exposures associated with cholera transmission in case–control studies: a systematic review. *Trop Med Int Health*. 2019;24(10):1151–1168. https://doi.org/10.1111/tmi.13293.
57 Mukherjee S, Sundberg T, Sikdar PK, et al. An integrated quantitative assessment of urban water security of a megacity in the global South. *Front Water*. 2022;4:834239. https://doi.org/10.3389/frwa.2022.834239.
58 Roskosky M, Ali M, Upreti SR, et al. Spatial clustering of cholera cases in the Kathmandu Valley: implications for a ring vaccination strategy. *Int Health*. 2021;13(2):170–177. https://doi.org/10.1093/inthealth/ihaa042.
59 Population Division–United Nations. World population prospects. n.d. Available from: https://population.un.org/wpp/. Accessed April 14, 2024.
60 Panda S, Chatterjee P, Deb AK, et al. *A policy brief: oral cholera vaccines—worth a shot?* 2017.
61 Craig M. Time-space clustering of Vibrio cholerae 01 in Matlab, Bangladesh, 1970-1982. *Soc Sci Med*. 1988;26(1):5–13. https://doi.org/10.1016/0277-9536(88)90040-8.
62 Ruiz-Moreno D, Pascual M, Emch M, et al. Spatial clustering in the spatio-temporal dynamics of endemic cholera. *BMC Infect Dis*. 2010;10(1):51. https://doi.org/10.1186/1471-2334-10-51.
63 Getis A, Ord JK. The analysis of spatial association by use of distance statistics. *Geogr Anal*. 1992;24(3):189–206. https://doi.org/10.1111/j.1538-4632.1992.tb00261.x.
64 Ilic I, Ilic M. Global patterns of trends in cholera mortality. *Trop Med Infect Dis*. 2023;8(3):169. https://doi.org/10.3390/tropicalmed8030169.
65 Islam MT, Im J, Ahmmed F, et al. Better existing water, sanitation, and hygiene can reduce the risk of cholera in an endemic setting: results from a prospective cohort study from Kolkata, India. *Open Forum Infect Dis*. 2023;10(11):ofad535. https://doi.org/10.1093/ofid/ofad535.
66 Kanungo S, Chatterjee P, Saha J, et al. Water, sanitation, and hygiene practices in urban slums of eastern India. *J Infect Dis*. 2021;224(Supplement_5):S573–S583. https://doi.org/10.1093/infdis/jiab354.
67 Demlie YW, Moore S, Dunoyer J, et al. Comparison of analysis methods to classify cholera hotspots in Ethiopia from 2015 to 2021. *Sci Rep*. 2024;14(1):7377. https://doi.org/10.1038/s41598-024-56299-5.