

# Spatiotemporal modelling of cholera and implications for its control, Uvira, Democratic Republic of the Congo

Ruwan Ratnayake (0000-0003-4978-6668)<sup>1\*</sup>, Jacqueline Knee (0000-0002-0834-8488)<sup>1</sup>, Oliver Cumming (0000-0002-5074-8709)<sup>1</sup>, Jaime Mufitini Saidi<sup>2</sup>, Baron Bashige Rumedeka<sup>2</sup>, Flavio Finger (0000-0002-8613-5170)<sup>3</sup>, Andrew S. Azman (0000-0001-8662-9077)<sup>4</sup>, W. John Edmunds (0000-0002-9179-2917)<sup>1</sup>, Francesco Checchi (0000-0001-9030-5382)<sup>1,^</sup>, Karin Gallandat (0000-0001-6412-1483)<sup>1,^</sup>

1. London School of Hygiene and Tropical Medicine, London, UK
2. Ministère de la Santé Publique, Division Provinciale de la Santé du Sud-Kivu, Zone de Santé d'Uvira, Uvira, South Kivu, Democratic Republic of the Congo
3. Epicentre, Paris, France
4. Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

<sup>^</sup> These senior authors contributed equally to this article.

\* Corresponding author: Ruwan Ratnayake, London School of Hygiene and Tropical Medicine, Keppel Street, London, UK WC1E 7HT ([ruwan.ratnayake@lshtm.ac.uk](mailto:ruwan.ratnayake@lshtm.ac.uk))

**Article Summary Line:** We find evidence of recurrent small-scale clustering of cholera infections in Uvira, Democratic Republic of the Congo and discuss the implications for control.

**Running Title:** Cholera clustering and control in Uvira

**Keywords:** cholera/epidemiology; Democratic Republic of the Congo; outbreaks/prevention & control; spatio-temporal analysis; transmission dynamics, *Vibrio cholerae*

## ABSTRACT

The African Great Lakes region including Eastern Democratic Republic of the Congo is a hotspot for cholera transmission. We evaluated the local and global clustering of cholera using 5 years (2016—2020) of suspected cases positive by rapid diagnostic test in Uvira, South Kivu to detect spatiotemporal clusters and the extent of zones of increased risk around cases. We detected 26 clusters (mean radius 652m and mean duration 24.8 days) which recurred annually in three locations and typically preceded seasonal outbreaks. We found a 1100m zone of increased infection risk around cases during the 5 days following clinic attendance for the 2016—2020 period and a 600m radius risk zone for 2020 alone. These risk zone sizes correspond with the area typically used for targeted intervention in the Democratic Republic of the Congo. Our findings underscore the value of the site-specific evaluation of clustering to guide targeted control efforts.

## INTRODUCTION

Cholera outbreaks continue to impact communities that lack access to safe water and adequate sanitation.(1-4) In these communities, cholera's relatively high reproduction number and short median incubation period (1.4 days, 95% credible interval 1.3—1.6)(1, 5) mean that an initial cluster can rapidly propagate across space. During outbreaks, household fecal-oral transmission through contaminated water, food, and fomites and direct contact becomes substantial and therefore, interventions to prevent infection of household contacts can reduce household transmission.(1, 6-8) Spatiotemporal clustering patterns around affected households have also demonstrated the propagation of transmission between neighboring households.(9, 10) To attenuate and possibly contain transmission during outbreaks based on this natural clustering, case-area targeted interventions (CATI), consisting of an early, multisectoral response within a 100—500m area around case-households, have been proposed.(11-13) CATIs typically include water, sanitation, and hygiene (WASH) interventions to improve water quality and safety (i.e., point of use water treatment, safe drinking water storage containers) and promote hygiene practices like handwashing, antibiotic chemoprophylaxis, and sometimes, oral cholera vaccination (OCV).(13) CATIs with WASH have been a major component of response strategies in Haiti and Yemen(13, 14) while CATIs with WASH and OCV have been used to suppress small outbreaks after mass vaccination campaigns in Juba, South Sudan and Kribi, Cameroon.(15, 16) In the Democratic Republic of the Congo (DRC), similar area-targeted interventions have included the

distribution of hygiene kits to case-households and the targeting of WASH interventions to a 500m radius around the most recent cases.(8, 17)

Studies in urban Kalemie, DRC and N'Djamena, Chad have estimated the zone of increased risk of infection around incident cases of at least 200m within the first 5 days after case presentation and in rural Matlab, Bangladesh, up to 450m within the first 3 days after case presentation.(9, 10) As CATIs and other targeted interventions become part of routine public health practice(14), more insight is required into the size and duration of the spatiotemporal zones of increased infection risk required to achieve a substantive impact on transmission, particularly in endemic areas. The size of the zone is likely influenced by factors that determine the strength of community transmission including population density, immunity, vaccination coverage, access to safe water and sanitation, and timeliness of the response.(3, 13)

Cholera has been endemic in the African Great Lakes Region including Eastern DRC since at least 1978 and now contributes substantially to the global cholera burden.(2, 18-22) The *V. cholerae* O1 sublineage AFR10 was introduced from South Asia to East Africa in the late 1990s, and has been driving transmission across the region during the last two decades.(23, 24) In Uvira, South Kivu, DRC, cholera is endemic with stable transmission punctuated by seasonal outbreaks.(18) Using an enhanced clinical surveillance system with rapid diagnostic testing that was setup in Uvira's cholera treatment units to support an impact evaluation of water supply infrastructure improvements(25), we investigate the location, timing, and annual

prediction of spatiotemporal clustering and to estimate the extent of spatiotemporal zones of increased cholera risk around incident cases in an endemic setting.

## **METHODS**

### **Setting**

Uvira is a town of approximately 280,000 located on the shore of Lake Tanganyika, an internationally-designated transmission hotspot where suspected cholera cases are reported year-round.(18-20) Seasonal emergence of cholera in Uvira is driven by seasonal exposure to aquatic reservoirs of *V. cholerae* in lakeside waters and person-to-person transmission, excess rainfall linked to fecal contamination of water sources, interruption of water supply and conflict and forced displacement.(18, 20, 26-28) Several city-wide interventions have been implemented including a water supply infrastructure program to improve the production and supply of piped drinking water for which construction started in 2018 and mass vaccination that took place from July to October 2020.(25)

### **Data sources**

We used a line list of suspected cases (i.e., passing  $\geq 3$  loose or watery stools in 24 hours) who received care between 2016 and 2020 at either the Uvira General Hospital's cholera treatment center (CTC) or the Kalundu cholera treatment unit (CTU, which opened in July 2019). Since April 2016, as part of an evaluation of a water supply infrastructure improvement program, rectal swabs collected from suspected cases have been systematically tested using an RDT (Crystal® VC O1/O139, Arkray Healthcare Pvt. Ltd, Gujarat, India) after a 6 hour enrichment period in alkaline peptone water at ambient temperature.(25, 29) The pooled sensitivity and specificity estimates for enriched

Crystal® VC RDTs are 83% (95% CI 67—92) and 98% (95% CI 94—99).(30) We extracted data for 2016—2020, including date of admission, completion of RDT and result, and avenue of residence (i.e., a small census enumeration area of mean size of 1177 (range 180—5711) persons based on population sizes estimated from 2017 official records.(31) The mean avenue area is 0.08 km<sup>2</sup> and minimum, maximum, and mean distances between avenue centroids are 0.0437m, 12.4km, and 3.1km, respectively. We based the main analyses on enriched RDT-positive cases given the presence of systematic testing in the CTC/CTU and assessments done at the CTC between 2017—2018 where only 40% of suspected cases were confirmed by polymerase chain reaction.(32)

### **Descriptive analysis and seasonal decomposition**

We described suspected cases using incidence per 10,000 population, proportion tested with RDT and proportion tested that were RDT-positive. To identify the timing of the cholera season, we analyzed the seasonal decomposition of the weekly incident series of RDT-positive cases with seasonal and trend decomposition using LOESS STL (locally estimated scatterplot smoothing). This method decomposes the time series into trend, seasonal, and random error components based on a two-week trend window and fixed seasonal pattern, and uses the additive model as the seasonal trends appeared relatively constant over time.(33) Missing data were integrated over the case-series with multiple imputation using chain equations.

## **Methods for spatiotemporal clustering**

We used two different methods to measure spatiotemporal clustering for different phenomena. The *space-time scan statistic* describes local clustering, or the expected density of cases at specific locations within a given area.(34) This gives the timing and locations where cases cluster, exceeding their expected density. The *tau statistic* ( $\tau$ ) describes global clustering, or the overall tendency for cases to occur near other cases in space and time.(35) This suggests the geographic and temporal extents of the zones of increased infection risk. See Appendix Text 1 for the mathematical formulation of these statistics.

### **Local clustering to identify recurrent locations and timing of seasonal outbreaks**

We used the space-time scan statistic to retrospectively detect the presence and location of spatiotemporal clusters. We conducted the analysis for the entire period (2016—2020) and by year. A relative risk (RR) compares the observed versus expected number of cases inside and outside of a cluster. Poisson distribution of the cases per avenue was assumed. To find the most likely cluster, candidate clusters were ordered by a log-likelihood ratio (LLR) where the cluster with the largest LLR is the least likely to be due to chance and therefore, the most likely cluster. The significance of each cluster was evaluated using Monte Carlo simulation to compare the original dataset with 999 random replicates produced under the null hypothesis.

We examined the entire dataset (i.e., a retrospective scan). We restricted the temporal and spatial windows to capture brief time periods (7—60 days) and a radius

that included  $\leq 10\%$  of the population at-risk. To capture clustering that persisted across years, we also used a longer temporal window (7—365 days) for 2016—2020.

To explore whether the space time scan statistic produced signals that preceded outbreaks, we conducted prospective scans of each of the clusters that were detected retrospectively. This was done to detect the earliest warning signal that indicates when that cluster would have first been detected. We simulated repeated prospective scans on the date of the retrospective cluster start day and each successive day (up to 4 weeks later) and calculated the median and IQR of the delay where a prospective scan would have first detected the cluster from the date produced by the retrospective scan that used more case data, and the cluster size at first detection. We visualized on the epidemic curve the timing of the first day of each retrospective cluster. To explore where transmission predominated, we calculated the proportion of the years that the avenue was included in any cluster from 2016—2020, ranging from 0 (not included in any cluster) to 5 (included in a cluster every year).(36)

### **Global clustering to inform the boundaries of risk**

We estimated the tau statistic ( $\tau$ ) for the entire period (2016—2020) and annually to quantify the spatial extent of the risk zone around an index case.(35) As the dataset only contained the date of the visit to the CTC/CTU as opposed to the date of symptom onset, this represented the risk of developing medically-attended disease, which we assumed indicates severe dehydration/diarrhea (compared with mild dehydration/diarrhea). This approach defines clustering in terms of how likely it is that any pair of cases are potentially transmission-related, within a given distance between



the cases. Accordingly, we first classified each pair of cases as potentially transmission-related if their dates of presentation were within 0-4 days of each other (approximately one serial interval).<sup>(5)</sup>  $\tau$  is the RR that an individual in the population within a given distance band ( $d_1, d_2$ , e.g., 100m, 150m) from an incident case also becomes a case that is potentially transmission-related, compared to the risk of any individual in the population becoming a potentially transmission-related case.  $\tau > 1$  indicates evidence of clustering within the given distance band.

As we lacked individual household locations of cases,  $\tau$  reflects the spatial scale of the avenues. We estimated  $\tau$  with a moving window of 50m computed every 10m at distances from 420m (as 5% of inter-avenue centroids fell below this value) to 2500m (approximate width of Uvira). We calculated the 95% CIs using the 2.5th and 97.5th quantiles from 1000 bootstrap replicates. We evaluated  $\tau$  over a 5-day window which included the date of case presentation, and a 4-day window which excluded the date of case presentation to account for a more realistic response started the day after.<sup>(9)</sup> To smooth the artefactual fluctuations resulting from the resolution of the data and smaller sample size of annual datasets, we calculated a moving average over the previous 10m. We defined the high-risk zone around incident cases as the radius up to which the moving average's lower 95% CIs crossed 1.0 for  $\geq 30$  consecutive meters. We defined another elevated-risk zone around incident cases as the radius up to which the moving average point estimate crossed 1.0 for  $\geq 30$  consecutive meters. To explore the potential biases from using centroids compared to household locations, we conducted a simulation study where we randomly assigned household locations within each case's

avenue and then estimated  $\tau$  using a lower distance range (75—2500m) (Appendix Text 2, Figures 1—4).

For both the scan statistics and  $\tau$ , we carried out sensitivity analyses using all suspected cases (i.e., RDT-positive and negative cases, and untested cases that met the suspected case definition).

All analyses were carried out in R software (v. 4.1.2) using the `rsatscan` (v. 1.0.5) and `IDSpatialStats` (v. 0.3.12) R packages to calculate the scan statistic (with SaTScan™ v. 10.0.2 software) and  $\tau$ .(37-40) Ethical approval was provided by the London School of Hygiene and Tropical Medicine (#10603-5) and the University of Kinshasa School of Public Health (#ESP/CE/173B/2022).

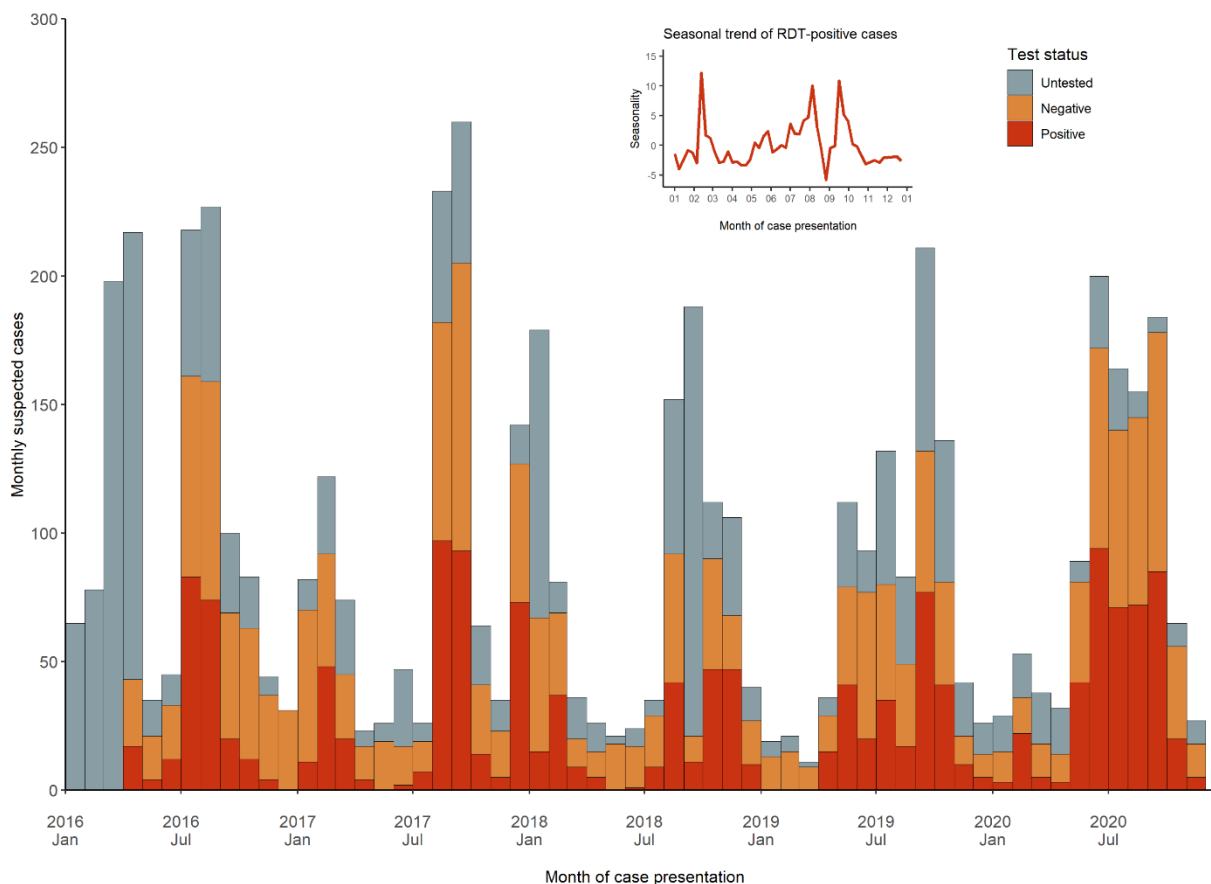
## RESULTS

5,447 suspected cases were recorded from 2016 to 2020. 3,456 (63.4%) of the 5,447 suspected cases were tested, of which 1,493 (43.2%) of the 3,456 tested cases were RDT-positive (Figure 1). Testing was not done when RDTs were stocked out, patients were admitted at nighttime and discharged by morning, patients refused, or the technician was ill. Percent positivity among those tested ranged from 36.6% to 46.9% (Table 1). Stable, seasonal transmission was observed with seasonal outbreaks typically beginning in the dry season (at the end of June/July to early October), followed by lower transmission in the rainy season (October—March/April) (Appendix Figures 5, 6). In some years, multiple peaks were seen in February and in the second half of the year

between August and October (Figure 1). Peaks typically reached 80—100 weekly suspected cases. Earlier transmission starting in March was seen in 2019 and 2020.

**Table 1.** Description of testing among suspected cholera cases, Uvira, 2016 to 2020

	2016	2017	2018	2019	2020
No. suspected cases	1341	1134	1000	922	1050
No. suspected cases per 10,000 population	47.9	40.5	35.7	32.9	37.5
No. (%) suspected cases that were RDT-tested	617 (46.0)	857 (75.6)	533 (53.3)	597 (64.8)	852 (81.2)
No (%) RDT-positive cases among RDT-tested	226 (36.6)	374 (43.6)	233 (43.7)	260 (43.6)	400 (46.9)



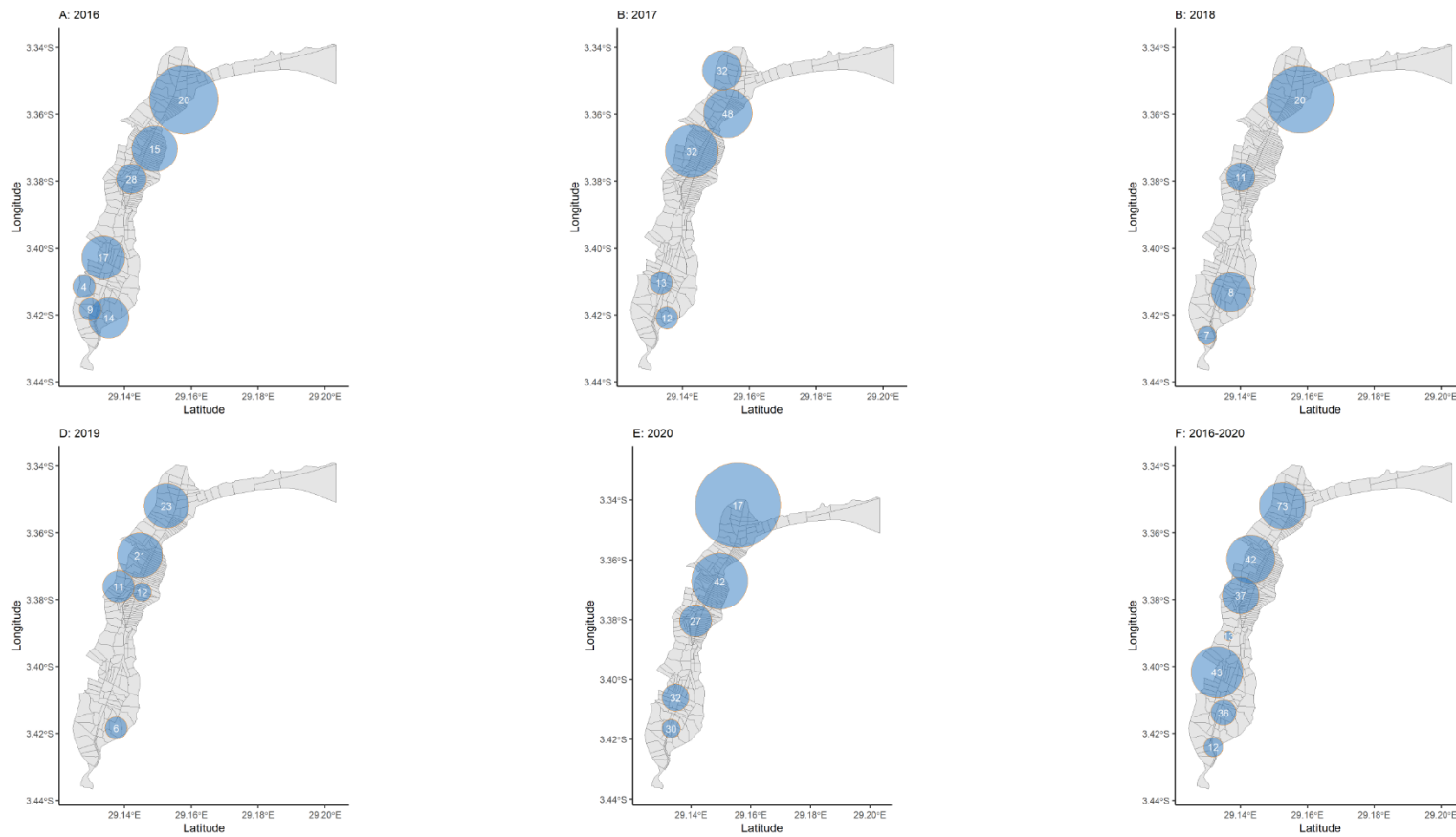
**Figure 1.** Cholera, Uvira, 2016—2020: Epidemic curve of monthly suspected cases by test status(25) with seasonal trend of monthly RDT-positive cases (inset)

Twenty-six spatiotemporal clusters were detected (Table 2). The mean cluster radius was 652m (range, 308—1582), mean size was 20 cases (range, 4—48), and mean duration was 24.8 days (range 1—58). The annual comparison of clustering demonstrated clustering in similar locations each year (Figure 2). The date of the first day of a retrospectively detected cluster usually (though not always) anticipated a surge in transmission over the next weeks, for all seasonal outbreaks except for early 2016 and 2017 when there were few cases tested (Figure 3, top). The median delay to the early warning signal (i.e., the number of days between retrospective detection date with all available data and the earliest prospective detection date) was 1 day (IQR 0—3) with a maximum delay of 23 days (Table 2). The median cluster size at signal detection was 3 cases (IQR 2—7) with a maximum size of 21 cases. Persistent clustering was seen in the north-central and southern areas, some distance away from the CTC (Figure 3, bottom). The 2016—2020 scan did not show clusters persisting between years but found larger clusters of 175—226 cases in the same locations. Sensitivity analyses of suspected cases found more clusters (N=32) of a similar mean radius and range (590m, range 270—1557) and larger mean size and range (37 cases, range 2—130) and longer duration (27.8 days, range 5—59), in similar locations (Appendix Table 2, Appendix Figure 7).

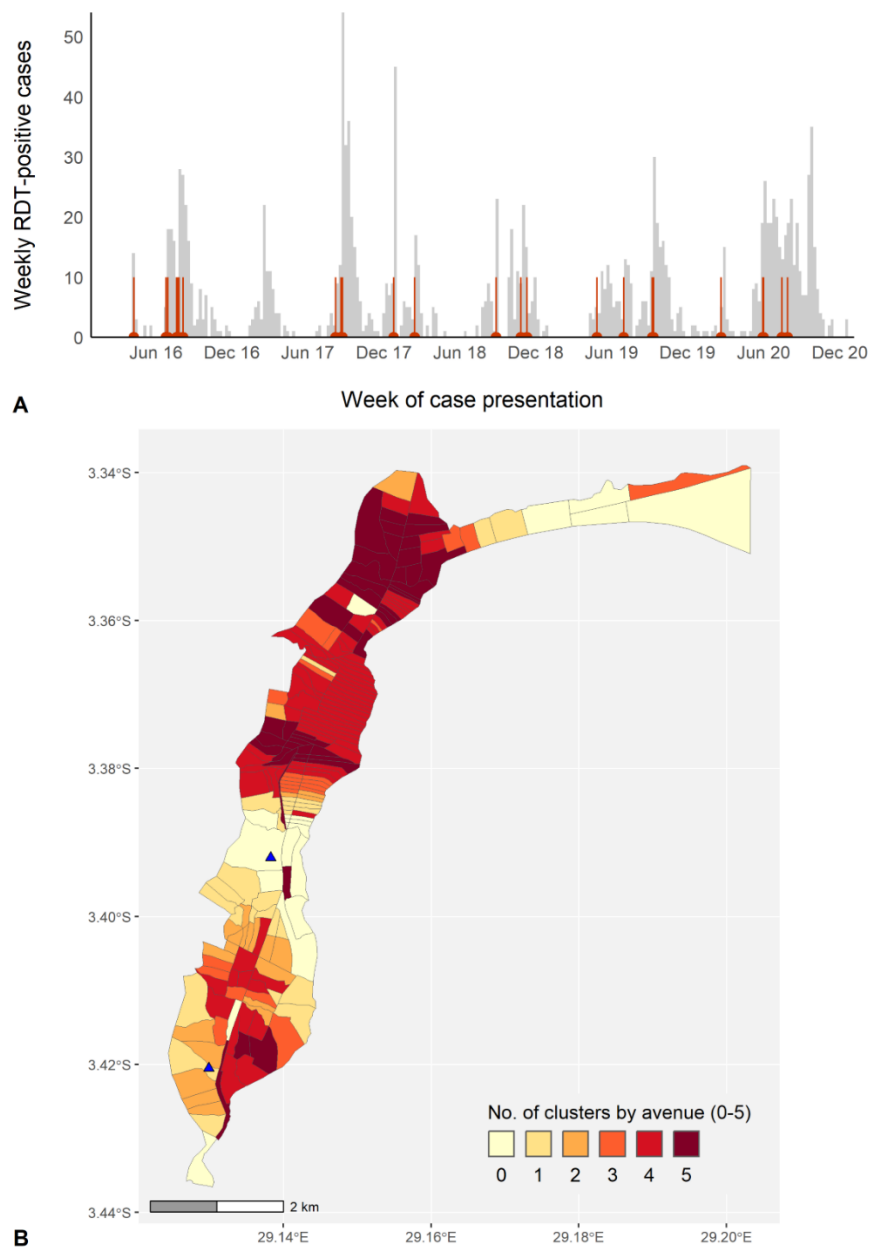
**Table 2.** Spatiotemporal clusters of RDT-positive cholera cases detected through annual scanning at the avenue level, Uvira, 2016—2020

Year	No.	Cases observed: expected	Population at-risk	RR <sup>†</sup>	Cluster radius (m)	Cluster start date (mm/dd)	Cluster duration (days)	Signal delay (days) <sup>‡</sup>	Size at signal (cases)
2016	1	20:1	30553	20.9	1140	08/05	18	8	11
	2	28:3	34232	10.5	497	06/25	48	0	2
	3	17:1	30758	13.8	717	07/22	23	5	12
	4	15:1	31240	11.9	758	06/29	23	1	4
	5	4:0	6579	344.4	376	04/09	1	0	3
	6	14:2	30082	8.8	668	07/21	30	0	3
	7	9:1	27452	12.6**	368	07/26	14	3	4
2017	1	48:4	51012	13.0	811	08/07	40	2	2
	2	32:2	43992	16.4	657	08/20	23	1	13
	3	32:4	49794	7.7	880	08/23	44	0	2
	4	13:1	51016	16.4	378	12/24	7	0	2
	5	12:2	50635	7.6**	368	08/23	15	12	2
2018	1	20:1	28884	26.6	1116	10/26	13	6	9
	2	11:1	31204	22.7	475	02/13	7	0	3
	3	8:0	25148	40.6	662	08/28	3	0	4
	4	7:0	17345	18.6**	308	11/10	10	1	3
2019	1	23:1	33751	18.6	743	09/10	18	1	7
	2	21:3	33162	9.0	755	09/07	35	0	12
	3	12:1	16210	12.3	309	04/27	29	1	2
	4	11:1	16495	13.2	527	09/07	24	0	2
	5	6:0	15001	27.8**	368	06/30	6	0	2
2020	1	42:6	60378	7.8	1048	07/29	58	2	3
	2	27:3	42423	8.7	599	07/15	46	23	21
	3	17:1	56029	19.1	1582	02/20	9	0	2
	4	30:5	63207	6.5	343	05/30	46	2	6
	5	32:6	63593	5.8	501	06/01	55	4	6

\* p-value < 0.001 ≥ p-value < 0.01; \*\* p-value < 0.001. † RR, relative risk. ‡ Signal delay indicates the number of days between retrospective detection date with all available data and the earliest prospective detection date.



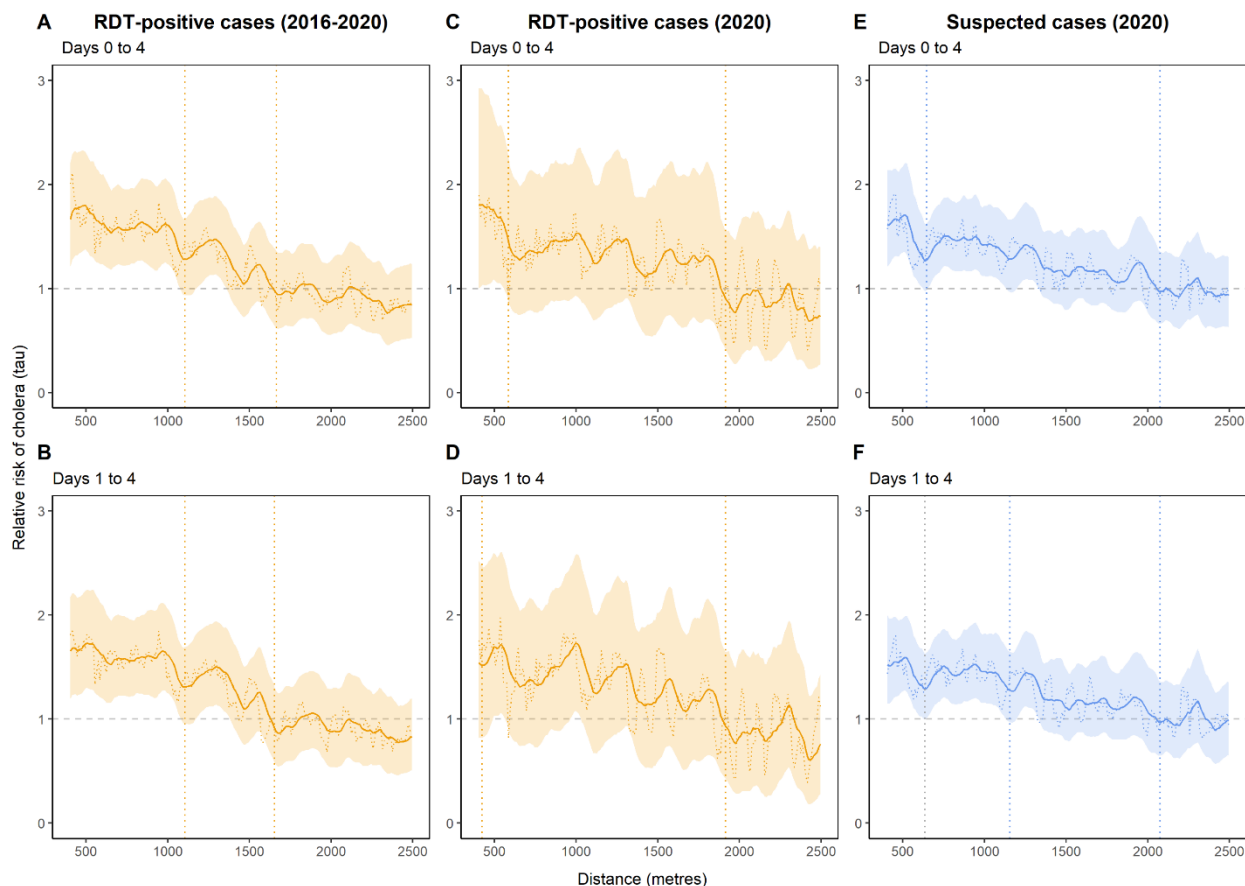
**Figure 2.** Spatial distribution of spatiotemporal clusters of RDT-positive cholera cases at the avenue level, Uvira, 2016—2020. Clusters have a relative risk  $>1$ ,  $p < 0.05$ . The size of the orange circle depicts the spatial radius and the number of cases, in white.



**Figure 3.** Cholera, Uvira, 2016—2020: (A) Epidemic curve showing weekly number of RDT-positive cholera cases based on week of presentation and start dates of 26 clusters (red vertical lines), (B) Cluster persistence within avenues for RDT-positive cases showing the number of years affected by clustering within avenues. The cholera treatment center (top) and unit (bottom) are marked with blue triangles.

Among RDT-positive cases from 2016—2020, within the first 5 days after a case presented for care, the high-risk zone extended from the case residence to 1105m and the risk remained elevated up to 1665m (maximum moving average  $\tau = 1.8$ , 95% CI 1.4—2.3, Figure 4A). During days 1—4, when a response can be more realistically mobilized, the risk zones remained similar (Figure 4B). Examining RDT-positive cases in 2020 alone, the high-risk zone extended to a smaller radius of 585m, and the risk remained elevated up to 1915m (maximum moving average  $\tau = 1.8$ , 95% CI 1.0—2.9, noting the wider confidence intervals for the smaller 2020 dataset, Figure 4C). During days 1—4, the risk zones were 425m and 1915m, respectively (maximum moving average  $\tau = 1.7$ , 95% CI 1.1—2.6, Figure 4D). In the sensitivity analysis of suspected cases from 2020, the trends remained like RDT-positive cases in 2020 (Figure 4C, 4D versus 4E, 4F). For suspected cases, during days 1—4, these risk zones were 1155m and 2075m, respectively (maximum moving average  $\tau = 1.6$ , 95% CI 1.3—2.0, though a drop in the lower CI was observed at 635m, as marked by first vertical dashed line in Figure 4F). Results by year for RDT-positive cases showed lower ranges in the radii of the high-risk zone (425m across all years except 2017 where it is 875m) and the elevated zone (1125—1485m) (Appendix Figure 8). Using simulated individual household locations (from 75—2500m for days 0—4), the results were similar to the main analysis of centroid locations with a moving average  $\tau \geq 2.0$  measured from 75m to 275m (maximum moving average  $\tau = 2.4$ , 95% CI 1.7—3.3), a similar high-risk zone radius of 1415m, and a similar descending trend in risk over distance, central tendencies and correlation coefficients (Appendix Figure 4, Appendix Table 1).





**Figure 4.** Cholera, Uvira, 2016—2020: Moving average estimates of  $\tau$  (relative risk) and 95% CIs (solid line and shading) with point estimates (dashed horizontal line) for days 0—4 (panels A, C, E) and days 1—4 (panels, B, D, F), for RDT-positive cases (in orange) and suspected cases (in blue) for cholera in 2016—2020 (panels A, B) and (panels C, D, E, F), using 1000 bootstrap samples. The vertical dashed lines indicate the spatial extent of the zone of high-risk where the lower 95% CI crossed 1.0 for  $\geq 30\text{m}$  consecutively (first line) and zone of elevated risk where the point estimate crossed 1.0 for  $\geq 30\text{m}$  consecutively (second line).

## DISCUSSION

We provide insight into clustering dynamics from one of the world's most burdensome cholera hotspots. Elucidating clustering patterns in an endemic setting can specify where intensive transmission is occurring early on, within small areas. Our results suggest that targeted interventions can take advantage of this natural clustering to mitigate seasonal outbreaks before they enlarge. Cities like Uvira, Goma and Bukavu are thought to regularly seed regional outbreaks; investigating transmission routes and coordinating prevention and control strategies there can have substantial public health benefits.(18)

The two clustering methods produced aligned results. When evaluating the 2016—2020 period, a 1105m high-risk radius around RDT-positive cases within the 5 days after case presentation was estimated with a  $\tau \leq 1.8$ . The risk zone for 2020 showed a 600m high-risk radius with a  $\tau \leq 1.8$ . These radii are consistent with those estimated for Matlab, Bangladesh (500m, within 4—6 days post-presentation, RR = 1.9).(10) In Uvira, an elevated-risk radius up to 2000m around cases demonstrates the persistence of risk in a densely-populated city. The risk zones remained intact after a 1-day delay wherein it is realistic to launch a response. The  $\tau$  estimates are higher than estimates from N'Djamena (220m) and Kalemie (330m), likely reflecting propagation among neighboring households, but it should be noted that we do not include distances <420m.(9) To explore this omission further, the simulation of household locations produced similar risk zones and an initial increase in risk from 75—275m equivalent to  $\tau < 2.5$ . There may be additional epidemiological differences including increased

environmentally-mediated transmission, immunity, population density, and mobility related to seasonal fishing and trading.

The space-time scan statistic demonstrated a mean radius (650m) among the 26 spatiotemporal clusters which emerged in the south and central-north areas. This is similar to the  $\tau$  high-risk zone for 2020 (600m). The start date of the retrospectively-detected cluster acted as an alarm that usually preceded the onset of seasonal outbreaks. The early warning signal for these clusters was delayed by a median of 1 day (whereas delays of >1 week for 3 clusters are not feasible for rapid response)

Our study's main strength is the use of high-specificity RDT-positive cases as compared to suspected cases alone. Previous analyses have relied on suspected cases which may overestimate risk due to inclusion of other diarrheal pathogens.(9-11) Our sensitivity analyses based on suspected cases showed similar results but with likely false positive clusters detected and a larger radius. In Uvira, the CTC/CTU may not be the main source of care for diarrhea, particularly when mild. In a 2021 community survey of health-seeking practices among Uvira residents, most persons with any diarrhea in the past week (70%) reported that they first visited pharmacies for care, rather than CTCs (4%).(41) Use of medically-attended cases introduces a potential bias of including only moderate to severely-ill cases, and therefore missing the transmission from milder cases. This is mitigated if medically attended cases represent a random proportion of all cholera cases. A major limitation is that the spatial resolution is based on avenue centroids, not household locations. This misses household transmission and case-pair distances <420m, where 5% of distances fell. Our simulations of household

locations showed qualitatively similar trends, with higher  $\tau$  at a smaller radius. Other limitations of the  $\tau$  statistic include limited power to detect true risk areas using narrow distance bands where the sample size of related pairs is small.(42) The  $\tau$  trendline and its sampling error are not smooth, as the clustering algorithm is recalculated every 50m and the moving average recalculated every 10m. Given that annual estimates are based on <500 RDT-positive cases, evaluation of the minimum number of cases needed for reliable  $\tau$  estimation is needed.(42) For both  $\tau$  and the space-time scan statistic, a circular radius has reduced sensitivity to detect the true geographical extent of noncircular clustering or an outbreak (i.e., an outbreak along the coastline, as might be the case in Uvira given its lakeside position), though detection appears unaffected.(34)

The results can inform control measures for seasonal outbreaks. The mapping of persistent clustering can be used to prioritize persistent high-recurrence hotspots for preventative measures, where transmission occurs early and predictably. Aiming for high coverage and uptake of preventative interventions in these areas can reduce exposure, reinfection, and transmission. Daily prospective scanning for local clustering could aid in early cluster detection across Uvira.(38) The radii of 100—500m used for a CATI strategies in DRC(17, 43) are justified by these findings and could perhaps be enlarged further. A 600—1105m radius of infection risk would include several thousand persons and would be logistically prohibitive to cover rapidly (i.e., within the 3 to 5 day risk window). However, CATIs may be considered for early containment of (a) potential zones of infection around new cases in less affected areas that fall outside of the known high-recurrence areas, (b) small outbreaks among lakeside communities in hotspots

which may seed larger outbreaks(18), and (c) sporadic cases after mass vaccination(15). Building on previous and current studies(9, 10) and operational experience(17, 43), a 200—600m radius can be used to narrow down areas for CATIs where transmission is likely.

## **Funding**

RR is funded by a Doctoral Foreign Study Award from the Canadian Institutes of Health Research (Award no. DFS-164266 to RR). The trial on which this study sources its data was co-funded by the French Agency for Development (Ref. No. EVA/364-2015) and the Veolia Foundation (Ref. No. 13/14 HD 1123).

## **Acknowledgements**

We thank the Uvira Health Zone and CTC/CTU collaborators for the support provided to testing and data collection, often under difficult circumstances. We thank John Giles, University of Washington, for advice on implementing the IDSpatialStats package. Last but not least, we thank all the patients who agreed to participate in the main trial.

**Data availability:** All data and code produced are available online at

[https://github.com/ruwanepi/Uvira\\_spatiotemporal](https://github.com/ruwanepi/Uvira_spatiotemporal)

## References

1. Kanungo S, Azman AS, Ramamurthy T, Deen J, Dutta S. Cholera. *Lancet*. 2022 Apr 9;399(10333):1429-40.
2. WHO. Cholera, 2021. *Weekly Epidemiological Record (WER)*. 2022;37:453-64.
3. Zheng Q, Luquero FJ, Ciglenecki I, Wamala JF, Abubakar A, Welo P, et al. Cholera outbreaks in sub-Saharan Africa during 2010-2019: A Descriptive Analysis. *Int J Infect Dis*. 2022 May 20.
4. Shannon K, Hast M, Azman AS, Legros D, McKay H, Lessler J. Cholera prevention and control in refugee settings: Successes and continued challenges. *PLoS Negl Trop Dis*. 2019 Jun;13(6):e0007347.
5. Azman AS, Rudolph KE, Cummings DA, Lessler J. The incubation period of cholera: a systematic review. *J Infect*. 2013 May;66(5):432-8.
6. Weil AA, Khan AI, Chowdhury F, Larocque RC, Faruque AS, Ryan ET, et al. Clinical outcomes in household contacts of patients with cholera in Bangladesh. *Clin Infect Dis*. 2009 Nov 15;49(10):1473-9.
7. George CM, Monira S, Sack DA, Rashid MU, Saif-Ur-Rahman KM, Mahmud T, et al. Randomized Controlled Trial of Hospital-Based Hygiene and Water Treatment Intervention (CHoBI7) to Reduce Cholera. *Emerg Infect Dis*. 2016 Feb;22(2):233-41.
8. D'Mello-Guyett L, Cumming O, Bonneville S, D'hondt R, Mashako M, Nakoka B, et al. Effectiveness of hygiene kit distribution to reduce cholera transmission in Kasai-Oriental, Democratic Republic of Congo, 2018: a prospective cohort study. *BMJ Open*. 2021;11(10).

9. Azman AS, Luquero FJ, Salje H, Mbaibardoum NN, Adalbert N, Ali M, et al. Micro-Hotspots of Risk in Urban Cholera Epidemics. *J Infect Dis.* 2018 Aug 24;218(7):1164-8.
10. Debes AK, Ali M, Azman AS, Yunus M, Sack DA. Cholera cases cluster in time and space in Matlab, Bangladesh: implications for targeted preventive interventions. *Int J Epidemiol.* 2016 Oct 27.
11. Finger F, Bertuzzo E, Luquero FJ, Naibei N, Toure B, Allan M, et al. The Potential impact of case-area targeted interventions in response to cholera outbreaks: A modeling study. *PLoS Med.* 2018 Feb;15(2):e1002509.
12. Michel E, Gaudart J, Beaulieu S, Bulit G, Piarroux M, Boncy J, et al. Estimating effectiveness of case-area targeted response interventions against cholera in Haiti. *Elife.* 2019 Dec 30;8.
13. Ratnayake R, Finger F, Azman AS, Lantagne D, Funk S, Edmunds WJ, et al. Highly targeted spatiotemporal interventions against cholera epidemics, 2000-19: a scoping review. *Lancet Infect Dis.* 2021 Mar;21(3):e37-e48.
14. UNICEF, Bulit G, Ramos M. Response to Cholera Outbreaks: Case-Area Targeted Interventions (CATI) and Community Outbreak Response Teams (CORT) Guidelines. New York, USA: UNICEF; 2020.
15. Ouamba JP, Fouda Mbarga N, Ciglenecki I, Ratnayake R, Tchiasso D, Finger F, et al. Implementation of targeted cholera response activities, Cameroon. *Bulletin of the World Health Organization.* 2023;101(03):170-8.
16. Parker LA, Rumunu J, Jamet C, Kenyi Y, Lino RL, Wamala JF, et al. Neighborhood-targeted and case-triggered use of a single dose of oral cholera vaccine



in an urban setting: Feasibility and vaccine coverage. PLoS Negl Trop Dis. 2017

Jun;11(6):e0005652.

17. Bompangue D, Moore S, Taty N, Impouma B, Sudre B, Manda R, et al.

Description of the targeted water supply and hygiene response strategy implemented during the cholera outbreak of 2017-2018 in Kinshasa, DRC. BMC Infect Dis. 2020 Mar 18;20(1):226.

18. Bompangue Nkoko D, Giraudoux P, Plisnier PD, Tinda AM, Piarroux M, Sudre B, et al. Dynamics of cholera outbreaks in Great Lakes region of Africa, 1978-2008. Emerg Infect Dis. 2011 Nov;17(11):2026-34.

19. Rebaudet S, Sudre B, Faucher B, Piarroux R. Environmental determinants of cholera outbreaks in inland Africa: a systematic review of main transmission foci and propagation routes. The Journal of infectious diseases. 2013 2013-Nov-01;208:S46-54.

20. Ingelbeen B, Hendrickx D, Miwanda B, van der Sande MAB, Mossoko M, Vochten H, et al. Recurrent Cholera Outbreaks, Democratic Republic of the Congo, 2008-2017. Emerg Infect Dis. 2019 May;25(5):856-64.

21. Birmingham ME, Lee LA, Ndayimirije N, Nkurikiye S, Hersh BS, Wells JG, et al. Epidemic cholera in Burundi: patterns of transmission in the Great Rift Valley Lake region. Lancet. 1997 Apr 5;349(9057):981-5.

22. Debes AK, Shaffer AM, Ndikumana T, Liesse I, Ribaira E, Djumo C, et al. Cholera Hot-Spots and Contextual Factors in Burundi, Planning for Elimination. Trop Med Infect Dis. 2021 May 11;6(2).

23. Weill FX, Domman D, Njamkepo E, Tarr C, Rauzier J, Fawal N, et al. Genomic history of the seventh pandemic of cholera in Africa. *Science*. 2017 Nov 10;358(6364):785-9.
24. Hounmanou YMG, Njamkepo E, Rauzier J, Gallandat K, Jeandron A, Kamwiziku G, et al. Genomic Microevolution of *Vibrio cholerae* O1, Lake Tanganyika Basin, Africa. *Emerg Infect Dis*. 2023 Jan;29(1):149-53.
25. Gallandat K, MacDougall A, Jeandron A, Saidi JM, Rumedeka B, Malembaka EB, et al. Improved water supply infrastructure to reduce acute diarrhoeal diseases and cholera in Uvira, Democratic Republic of the Congo: results and lessons learned from a pragmatic trial. *OSF Preprints*. 2023 July 16, 2023.
26. Jeandron A, Saidi JM, Kapama A, Burhole M, Birembano F, Vandeveld T, et al. Water supply interruptions and suspected cholera incidence: a time-series regression in the Democratic Republic of the Congo. *PLoS Med*. 2015 Oct;12(10):e1001893.
27. Bompangue D, Giraudoux P, Handschumacher P, Piarroux M, Sudre B, Ekwanzala M, et al. Lakes as source of cholera outbreaks, Democratic Republic of Congo. *Emerg Infect Dis*. 2008 May;14(5):798-800.
28. Charnley GEC, Jean K, Kelman I, Gaythorpe KAM, Murray KA. Association between Conflict and Cholera in Nigeria and the Democratic Republic of the Congo. *Emerg Infect Dis*. 2022 Dec;28(12):2472-81.
29. Jeandron A, Cumming O, Rumedeka BB, Saidi JM, Cousens S. Confirmation of cholera by rapid diagnostic test amongst patients admitted to the cholera treatment centre in Uvira, Democratic Republic of the Congo. *PLoS One*. 2018;13(8):e0201306.

30. Muzembo BA, Kitahara K, Debnath A, Okamoto K, Miyoshi SI. Accuracy of cholera rapid diagnostic tests: a systematic review and meta-analysis. *Clin Microbiol Infect.* 2022 Feb;28(2):155-62.
31. Town of Uvira. Official census data; 2018.
32. Williams C, Cumming O, Grignard L, Rumedeka BB, Saidi JM, Grint D, et al. Prevalence and diversity of enteric pathogens among cholera treatment centre patients with acute diarrhea in Uvira, Democratic Republic of Congo. *BMC Infect Dis.* 2020 Oct 9;20(1):741.
33. Cleveland R, B., Cleveland W, S., McRae J, E., Terpenning I. STL: A Seasonal Trend Decomposition Procedure Based on Loess. *Journal of Official Statistics.* 1990;6(1):3-73.
34. Kulldorff M, Heffernan R, Hartman J, Assuncao R, Mostashari F. A space-time permutation scan statistic for disease outbreak detection. *PLoS Med.* 2005 Mar;2(3):e59.
35. Lessler J, Salje H, Grabowski MK, Cummings DA. Measuring Spatial Dependence for Infectious Disease Epidemiology. *PLoS One.* 2016;11(5):e0155249.
36. Cleary E, Boudou M, Garvey P, Aiseadha CO, McKeown P, O'Dwyer J, et al. Spatiotemporal Dynamics of Sporadic Shiga Toxin-Producing *Escherichia coli* Enteritis, Ireland, 2013-2017. *Emerg Infect Dis.* 2021 Sep;27(9):2421-33.
37. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2023.

38. Kulldorff M, Information Management Services Inc. SaTScan™ v10.0.2: Software for the spatial and space-time scan statistics. 2021.
39. Kleinman K. rsatscan: Tools, Classes, and Methods for Interfacing with SaTScan Stand-Alone Software. 2017 September 14, 2017 [cited 2022; Available from: <https://github.com/Kenkleinman/rsatscan>]
40. Giles JR, Salje H, Lessler J. The IDSpatialStats R Package: Quantifying Spatial Dependence of Infectious Disease Spread. *The R Journal*. 2019;11(2):308-27.
41. Gallandat K, Hutchins C, Malembaka EB, Jeandron A, Saidi JM, Rumedeka BB, et al. Process evaluation of an urban piped water supply infrastructure improvement programme in a cholera hotspot. *OSF Preprints*. 2023 July 23, 2023.
42. Pollington TM, Tildesley MJ, Hollingsworth TD, Chapman LAC. The Spatiotemporal Tau Statistic: a Review. *arXiv [preprint]*. 2019.
43. Ratnayake R, Peyraud N, Ciglenecki I, Gignoux E, Lightowler M, Azman AS, et al. Effectiveness of case-area targeted interventions including vaccination on the control of epidemic cholera: protocol for a prospective observational study. *BMJ Open*. 2022 Jul 6;12(7):e061206.