Participatory Disease Surveillance for Early Detection of Cholera-like Diarrheal Diseases Outbreaks in Rural Villages in Malawi - a prospective cohort.

Mariana Valerio, Beverly Laher, John Phuka, Guilherme Lichand, Daniela Paolotti, Onicio Leal Neto

Submitted to: JMIR Public Health and Surveillance
on: June 01, 2023

Table of Contents

Original Manuscript **5**
Supplementary Files **28**
Figures **29**
Figure 1 ..... **30**
Figure 2 ..... **31**
Figure 3 ..... **32**
Figure 4 ..... **33**
Figure 5 ..... **34**
Figure 6 ..... **35**
Figure 7 ..... **36**
Related publication(s) - for reviewers eyes onlies **37**
Related publication(s) - for reviewers eyes only 0 ..... **38**
Related publication(s) - for reviewers eyes only 0 ..... **38**

Participatory Disease Surveillance for Early Detection of Cholera-like Diarrheal Diseases Outbreaks in Rural Villages in Malawi - a prospective cohort.

Mariana Valerio[1] MD; Beverly Laher[2] BSc, MSc; John Phuka[2] PhD; Guilherme Lichand[3] PhD; Daniela Paolotti[4] PhD; Onicio Leal Neto[5] PhD

[1]Global Studies Institute University of Geneva Geneva CH
[2]Kamuzu University of Health Sciences Lilongwe MW
[3]Department of Economics University of Zurich Zurich CH
[4]Institute for Scientific Interchange Turin IT
[5]Department of Computer Science ETH Zurich Zurich CH

Corresponding Author:
Onicio Leal Neto PhD
Department of Computer Science
ETH Zurich
Universitatstrasse 6
Zurich
CH

Abstract

Background: Cholera outbreaks are complex and influenced by environmental factors, socio-economic conditions, and population dynamics, leading to limitations in traditional surveillance methods. In Malawi, cholera is considered an endemic disease. Its epidemiological profile is characterized by seasonal patterns, often coinciding with the rainy season when contamination of water sources is more likely. However, the current outbreak has extended to the dry season, having deaths reported in all 29 districts since March 2022, and it has been considered the worst outbreak of the 10 past years.

Objective: The study aims to evaluate the feasibility and outcomes of participatory surveillance using Interactive Voice Response technology for early detection of cholera-like diarrheal disease outbreaks in Malawi.

Methods: The longitudinal cohort study followed up for 24 weeks, including 740 households in rural settings in Malawi. The survey tool was designed to have 10 symptom questions collected every week. The proxies' rationale was related to exanthematic, ictero-hemorragica for endemic diseases or events, diarrhea and respiratory/targeting acute diseases or events and Diarrhea and respiratory/targeting seasonal diseases or events. This present work will focus only on the CLDD as a proxy for gastroenteritis and cholera. The definition of CLDD utilized in this study was the following: reports that informed Diarrhea AND Fever OR Vomiting/Nausea.

Results: During the period of the study, our data comprises 16,280 observations, achieving an average of 35% of weekly participation rate. Maganga TA showed the highest average of completed calls, 144.83 (SD = 10.587), while Ndindi TA showed an average of completed calls of 123.66 (SD = 13.176). Participation rates were slightly higher at the beginning of the study and presented a discrete decay over time, thanks to the sensitization activities rolled out at the CBCCs level. On the attack rates for CLDD, the study found close rates between Maganga TA and Ndindi TA, showing 16% and 15%, respectively.

Conclusions: Participatory surveillance has proven to be of high value for the early detection of epidemics, and IVR technology is a promising approach for disease surveillance in rural villages in Africa, where access to healthcare and traditional disease surveillance methods may be limited. The study highlights the feasibility and potential of IVR technology for timely and comprehensive reporting of disease incidence, symptoms, and behaviors in resource-limited settings.

(JMIR Preprints 01/06/2023:49539)
DOI: https://doi.org/10.2196/preprints.49539

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Participatory Disease Surveillance for Early Detection of Cholera-like Diarrheal Diseases Outbreaks in Rural Villages in Malawi - a prospective cohort.

INTRODUCTION

Cholera-like diarrheal diseases (CLDD), refer to a group of diarrheal illnesses that share similar pathogenesis and symptoms with cholera but are caused by agents other than the *Vibrio cholerae*, such as Escherichia coli, and some strains of Campylobacter, Yersinia, Aeromonas and other pathogens [1]. Typically, CLDD are responsible for milder and shorter episodes of diarrhea than cholera, but a comprehensive clinical assessment, including laboratory testing, is essential for accurate diagnosis, especially when occurring in cholera endemic areas like Malawi [1–3]. Its epidemiological profile is characterized by seasonal patterns, often coinciding with the rainy season when contamination of water sources is more likely [4–6]. Therefore, the global burden of diarrheal disease lies heavily in places constantly plagued by disruptive environmental and socio-economic conditions able to affect water supply and sanitation infrastructure, most of them located in Sub-Saharan Africa, where the cross-border cholera outbreaks, lack of testing capacity and the surveillance systems limitations impose an extra layer of complexity to the situation [7–14]. In December 5, 2022 cholera was declared as a public health emergency, and it has been considered the country's worst outbreak of the 10 past years [15]. The raise of cases expected during the wet season has extended to the dry season, having deaths reported in all 29 districts since March 2022 [15,16]. The situation started to be alarming after the flooding caused by tropical storm Ana and Cyclone Gombe (January and March 2022, respectively) [17]. These phenomena led to population displacement, WASH infrastructure disruption and were worsened by old problems such as the lack of preparedness to refrain the disease, going from a reliable and geographically distributed early detection system to shortages in Oral Cholera Vaccine (OCV), and treatments due to multiple outbreaks within the African Region [2,5,15,17–19]. As of April 4, 2023, the World Health Organization (WHO) has reported 160,756 cases of suspected cholera in the African Region with a case-fatality ratio (CFR) of 2.1%. Malawi has been hit particularly hard, accounting for 35% (56,763) of the total cases and 52% (1,722) of the total fatalities, and keeping a steadily high CFR of above 3% [16]. Unfortunately, the true burden of cholera is likely higher than reported figures due to underreporting and lack of access to healthcare in many areas [7,8].

In line with urgent needs for comprehensive action, it is crucial to focus on viable strategies that can complement traditional epidemiological surveillance and foster community engagement [20]. In this field, participatory surveillance (PS) has proven to be of extreme value for the early detection of epidemics as this system engages the community in a bidirectional way, able to capture strategic data from the community processing the knowledge acquired and giving the information back in nearly real-time [21–23]. This approach involves the active participation of community members, health workers, and other stakeholders in the surveillance process, including reporting and monitoring of disease incidence, symptoms, and behaviors. PS has demonstrated applicability in different scenarios and diseases, such as influenza, cholera, Covid-19, Zika and others [21,24–31]. One of the main advantages of PS is that it overcomes several challenges associated with traditional disease surveillance methods. Participatory surveillance empowers communities to take ownership of their health and well-being, allowing for more comprehensive and timely reporting of cases [32,33].

Participatory surveillance using Interactive Voice Response (IVR) technology is a promising approach for disease surveillance in rural villages in Africa for collecting timely and comprehensive data on disease incidence, symptoms, and behaviors in resource-limited settings [34,35]. Additionally, IVR technology is cost-effective and has the potential to improve disease surveillance and control efforts in low-income settings [35–37]. Furthermore, it is an easy-to-use technology that requires minimal training and can reach a wide range of people, including those in remote and rural areas, where access to healthcare and traditional disease surveillance methods may be limited.

The aim of this study is to evaluate the feasibility and outcomes of low-cost, high- frequency, and high-quality data collection through PS for early detection of cholera-like diarrheal diseases (CLDD) outbreaks in Malawi, implementing a system to identify its early signals.

METHODS

This longitudinal prospective cohort study followed up for 24 weeks 740 households in rural settings in Malawi. It was rolled out in two Traditional Authorities (TAs) located in Salima District, the central region of Malawi (Figure 1). The choice of the TAs took into consideration the broader context of a larger study - the Child Development Study, an initiative to leverage high-frequency data collection and novel technologies for understanding child development in low-income settings.

Hence, including logistical considerations and resource availability, which made these TAs suitable candidates for inclusion. The College of Medicine Research and Ethics from Malawi – COMREC (ref. Number P_11_20_3202) and the Human Subjects Committee of the Faculty of Economics, Business Administration and Information Technology at the University of Zurich (OEC IRB# 2018-046) approved its realization. A total of 4,051 households were enrolled in the Child Development Study. From this total, a sample randomization was performed, selecting 740 households from both TAs areas, representing 3,743 household members, including 2,393 children. For the purpose of this study, it is imperative to highlight the definition of household in Malawi, that is “one or more persons, related or unrelated, who make common provisions for food and who regularly take their food from the same pot and/or share the same grain house (nkhokwe) or pool their incomes together for the purpose of purchasing food” [38].

The recruitment strategy was implemented by the Kamuzu University of Health Sciences (KUHeS) local team, together with trained enumerators responsible for collecting the households’ phone numbers, conducting a baseline survey, and gathering metadata. All these interactions, including the sensitization campaigns promoted throughout the study took place at the Community Based Childcare Centers (CBCCs), as the inclusion criteria for the main study required at least one child attending the CBCCs.

All participants signed the informed consent prior to enrollment, and were compensated for their time spent reading and responding to messages in accordance with COMREC guidelines. Additionally, all phone costs were covered by the project.

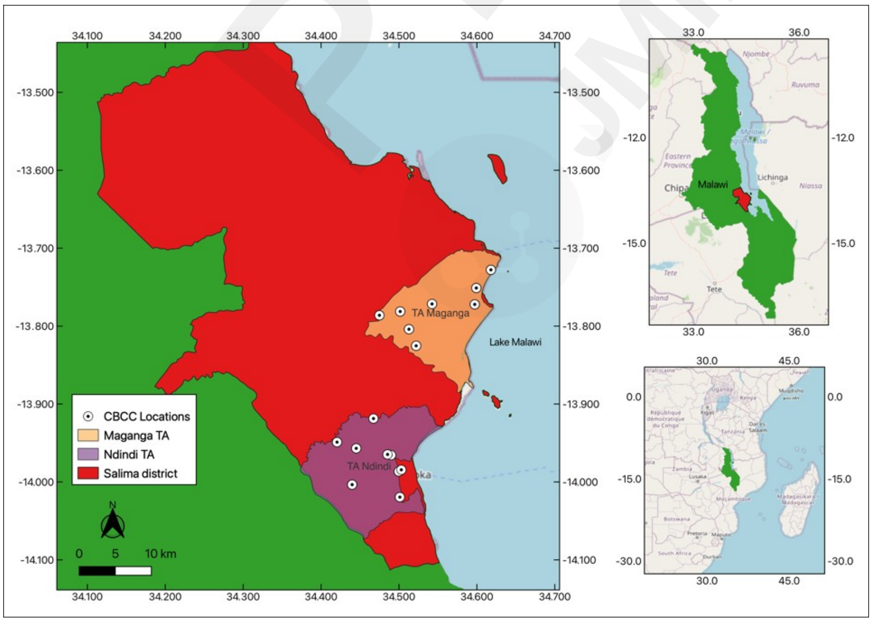


Figure 1. Study area highlighting the location of Salima district, Maganga TA, Ndindi TA and location of CBCCs.

Data collection via The Interactive Voice Response (IVR)

The survey instrument was designed to have 10 questions about symptoms to be answered every week by the same household member, over 24 weeks (Table 1). In addition, the design allowed all sorts of phone devices to be included, with participants required to simply type, press or dial the numbers indicated by the voice message to respond with Yes or No. The IVR went live from 18th July 2022 to 8th January 2023. Calls were conducted in Chichewa, the primary language spoken in Malawi. The survey content was pretested with in-country representatives of KUHeS to reduce events of loss in translation and misinterpretation. The survey was asking whether anyone in the household had experienced specific symptoms in the past 7 days, including: *fever*, *headache*, *joint pain*, *vomiting/nausea*, *jaundice*, *chills*, *body ache*, *bleeding*, and *diarrhea*. Additionally, participants were asked if anyone in the household had undergone a malaria test during the same period, with response options including: "*took a test with the negative result*" ; "*did not take the test*"; and "*took a test with a positive result*".

Table 1. List of questions and answers collected. Adjustments and the translation were made in Chichewa keeping the meaning of each symptom.

Question	Answer
Did you or anybody in your house has experienced fever in the last 7 days?	If yes, dial 1, if no dial 2.
Did you or anybody in your house has experienced headache in the last 7 days?	If yes, dial 1, if no dial 2.
Did you or anybody in your house has experienced joint pain in the last 7 days?	If yes, dial 1, if no dial 2.
Did you or anybody in your house has experienced vomiting/nausea in the last 7 days?	If yes, dial 1, if no dial 2.
Did you or anybody in your house has experienced jaundice in the last 7 days?	If yes, dial 1, if no dial 2.
Did you or anybody in your house has experienced chills in the last 7 days?	If yes, dial 1, if no dial 2.
Did you or anybody in your house has experienced body ache in the last 7 days?	If yes, dial 1, if no dial 2.
Did you or anybody in your house has experienced bleeding in the last 7 days?	If yes, dial 1, if no dial 2.
Did you or anybody in your house has	If yes, dial 1, if no dial

experienced diarrhea in the last 7 days?	2.
Did you or anybody in your house take malaria test in the last 7 days?	If took the test with negative result (dial 1); If did not take the test (dial 2); and If took the test with a positive result (dial 3).

The proxies and, consequently the questions used in the survey were selected based on the literature [39,40]. Due to the necessity to repeat the same questions every week, special attention was given to have the questions capable of gathering as much information as possible while avoiding participants' fatigue, enhance adherence. For this reason, the survey used proxies for malaria, rash, respiratory diseases and CLDD. The proxies' rationale was related to exanthematic, ictero-hemorragica for endemic diseases or events, diarrhea and respiratory/targeting acute diseases or events and Diarrhea and respiratory/targeting seasonal diseases or events. This study will specifically focus on reports related to CLDD by the occurrence of *Diarrhea AND Fever OR Vomiting/Nausea*. Other proxies and their associated syndromes will be explored in future studies.

The household heads received calls from Monday to Friday, between 8 am to 10 am (A retry pattern of 2 calls with 2 hours intervals) and follow-up calls in the evening between 5:30 pm and 7:30 pm, running with this pattern throughout the week. An extract, transform and load (ETL) process was performed weekly to feed in a dashboard for data visualization of compiled results in near-real time.

Sensitization campaigns and participants engagement

Sensitization campaigns were conducted to raise awareness among participants before the calls started officially. Additionaly, during the development of the study the research field team organized meetings at each CBCCs used as a reference point for the study (Figure 1). Based on the eligibility criteria, all households involved in the study had at least one child who regularly attended the CBCCs. To ensure widespread coverage and engagement, we conducted sensitization campaigns at the CBCC locations. This approach made it easier for participants to access the information and encouraged their active involvement. The campaigns were held at various times and on different days to maximize participation and reach as many people as possible. They were facilitated by an

Assistant Environmental Health Officer (AEHO), a social welfare representative, and the Association of Early Childhood Development Personnel in Malawi (AECDM). These campaigns aimed to raise awareness and inform participants about the process of calls and interactions, addressing any concerns or questions they had regarding any technical issue related to the use of the mobile phone in this specific context, to reinforce the fact that the repetitive calls were not a span or a mistake, and to discuss the importance of preventive health measures. The sensitization campaigns framework is described in table 2.

Table 2. Sensitization campaigns framework aiming participants engagement during the prospective cohort.

Description	Purpose
Demonstrative live call	To introduce the IVR system to potential participants by giving them a live demonstration of how it works. This allows participants to become familiar with the technology and understand how to use it.
Presentation explaining in simple words how the data would be collected and the relevance of the study.	To explain the purpose of the study and the importance of the data being collected. This helps participants understand how their participation will contribute to the success of the project and how it will benefit child development and health.
Clarification and reinforcement that the calls would come weekly from the same number, addressing the same questions.	To ensure that participants understand the frequency and consistency of the calls they will receive. This reinforces the importance of their participation and helps to build trust and rapport with the participants.
Presentation explaining the importance of preventive and early detection of outbreaks and empowering the participants as key collaborators to the success of such initiatives	To motivate and empower participants to take an active role in the project. By understanding the importance of their participation in preventing and detecting outbreaks, participants are more likely to engage fully with the IVR system and provide accurate information.
Reinforcement about the possibility of calling back in case the participant missed a call.	To ensure that participants understand that they can still participate even if they miss a call. By providing clear instructions on how to do so and reinforcing that there is no cost involved, participants are more likely to remain engaged with the project.
Enquirement for the preferable time spot to receive the calls.	To allow participants to choose a convenient time for the weekly calls. By accommodating the schedules of participants, the project is more likely to receive accurate and consistent data.
Opening time for participant's questions and further requests for clarifications.	To provide participants with an opportunity to ask questions and seek further clarifications about the project. This helps to build trust and rapport with the participants and ensures that they are fully informed about the purpose and methodology of the study.

Statistical analysis

We computed the household attack rate as an indication of how many households were under risk over the studied period. We assumed household reports as a proxy for the risk of presenting any of those symptoms related to CLDD. The attack rate is expressed as following:

$$Household\ attack\ rate = \frac{N_{hpr}}{T_{hh}} \times 100$$

where hpr is the household positive reports (number of true reports on the given proxy), and T_{hh} the total number of households in the study.

We utilized a polynomial generalized additive model (GAM) to analyze a time-series dataset, giving that the independent variable was time. The covariates were CLDD reports, diarrhea-like reports, fever-like reports, and vomiting/nausea-like reports, and they were chosen as non-parametric due to the uncertainty relationship with the outcome.

The GAM is a flexible regression technique that allows for nonlinear relationships between the predictor variables and the response variable. To evaluate the performance of the polynomial GAM, we calculated R squared, Deviance explained, Generalized Cross-Validation (GCV), and Scale estimate. R squared values range from 0 to 1, with higher values indicating a better fit of the model to the data. The formula presented below is used to estimate the coefficient of determination:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

where SS_{res} is the sum of squared residuals and SS_{tot} is the total sum of squares.

Deviance explained measures the reduction in deviance, which is a measure of the difference between the observed data and the fitted values, by the model. Deviance-explained values range from 0 to 1, with higher values indicating a better fit of the model to the data. The formula is described as follow:

$$Deviance_{explained} = \frac{(Deviance_{null} - Deviance_{model})}{Deviance_{null}}$$

where $Deviance_{null}$ is the deviance of a null model with no predictors, and $Deviance_{model}$ is the deviance of the model being evaluated.

The GCV score is the average of the squared differences between the predicted values and the actual values. To estimate GCV score it was utilized the following formula:

$$GCV = \frac{\frac{1}{N}\sum_{i=1}^{n}[y_i - \hat{y}_i]}{(1-h_i)^2}$$

where N is the sample size, y_i is the observed response variable for the i^{th} observation, $\hat{y}_i$ is the predicted response variable for the i^{th} observation, and h_i is the hat matrix element for the i^{th} observation.

Finally, the Scale estimate is a measure of the residual variance in the model. It represents the standard deviation of the residuals from the model and is used to assess the goodness of fit of the model. The formula can be described as follow:

$$Scale\,estimate = \frac{\sqrt{\sum res^2}}{(n-p)}$$

where n is the sample size, p is the number of predictors in the model, and res are the residual differences between the observed response variable and the predicted response variable.

A kernel estimator was used to evaluate the spatial density of the events. This technique is a non-parametric method that can estimate the probability density function (PDF) of a random variable based on observed data. In the case of CLDD reports, the PDF would represent the distribution of cases across a population, which is an important factor in understanding the spread of the disease and identifying high-risk areas. The formula is defined as:

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right)$$

where x is the point at which the estimate is being made, x_i represents the i^{th} observation in the data set, n is the number of observations in the data set, h is the bandwidth parameter, and K is the kernel function.

In this study, to safeguard participant privacy and avoid disclosing sensitive information, we utilized the coordinates of CBCCs. Notably, CBCCs cater to multiple villages that are usually close to the facility. The study's included households had at least one child attending the CBCC.

Consequently, given the dynamic social landscape within these villages, the CBCC locations serve as a suitable proxy for the frequently visited vicinities by households. These measures ensure the confidentiality of participants while allowing for accurate spatial analysis within the study's scope.

RESULTS

During the period of the study, our data comprises 16,280 observations, achieving an average of 35% of weekly participation rate. Maganga TA showed the highest average of completed calls, 144.83 (SD = 10.587), while Ndindi TA showed an average of completed calls of 123.66 (SD = 13.176). The participation rates were slightly higher at the beginning of the study and showed a discrete decline over time, with no significant drop, possibly influenced by the sensitization efforts implemented at the CBCCs level. Figures 2 displays the participation rates for each week throughout the entire period; Figure 3 brings these data disaggregated by TA area.

On the attack rates for CLDD, the study found close rates between Maganga TA and Ndindi TA, showing 16% and 15%, respectively (table 5). While Fever-like reports could be taken as more sensitive than Diarrhea-like reports and Vomiting/Nausea-like reports. However, using the proxy for CLDD might have helped to envelop the sensitivity of the signal, decreasing the potential number of false positives.

Considering the aggregated view, CLDD showed a consistent signal over time (R^2 = 0.681404) and was even stronger when isolating diarrhea-like symptom and fever-like symptom, as seen in table 3. When breaking down by TA area, Maganga (table 4) presents a better performance than Ndindi (table 5). At the time series distribution, the highest activity at the beginning of the cohort could be explained by the initial efforts and fresh recall of participants about the project. After 5 weeks of reports, the signal started increasing again, which might be related to the occurrence of the cholera outbreaks that hit that region during those months. The first report of cholera case in the Salima district was found on September 14th, two days after all signals showed a spike, as seen in Figure 5. Even though official reports were scarce to inform this occurrence, it was detected in a hyperlocal media website the confirmation of this first case of cholera in the Salima district [41].

In the following weeks, the peak of cholera cases in the Maganga TA area was detected, where an ascending pattern in the diarrhea-like reports could be seen [42]. Three weeks before this peak, CLDD reports were at the highest level when this area is looked isolated.

Regarding the spatial distribution of the CLDD reports, even considering a homogeneous distribution of these events over the regions, it is possible to verify a concentration close to the lake vicinity. The Salima District Council locally issued a Cautionary Statement on Cholera on October 13, 2022, aiming to provide the populace with guidance. In this statement, they reported: “Salima district has registered a total of 54 cases as of 13 October 2022 with TA Maganga along the lake shore, being the most affected with 45 cases” [43].

That can corroborate the ability of the proxies in pointing out a trend of possible cholera outbreak in areas close to important risk areas.

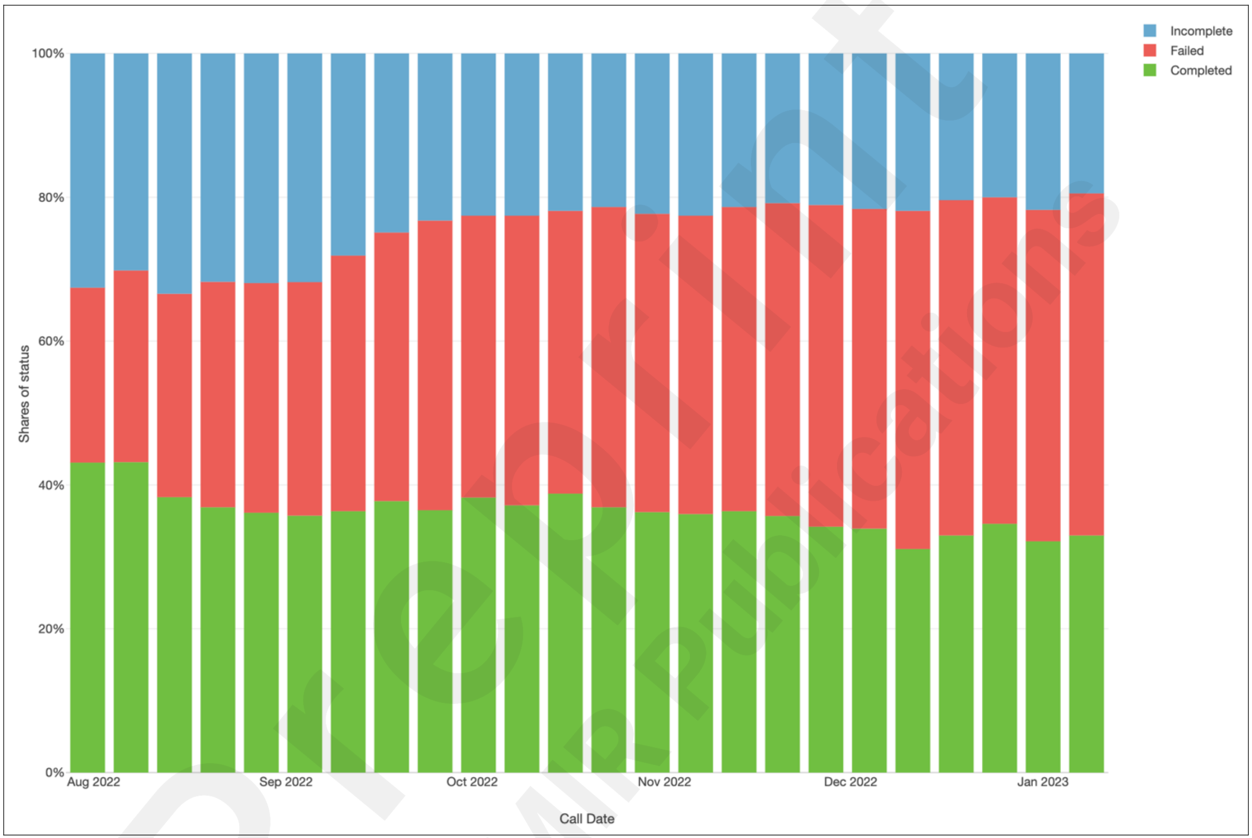


Figure 2. Participation profile rate according to the status of the IVR call aggregated.

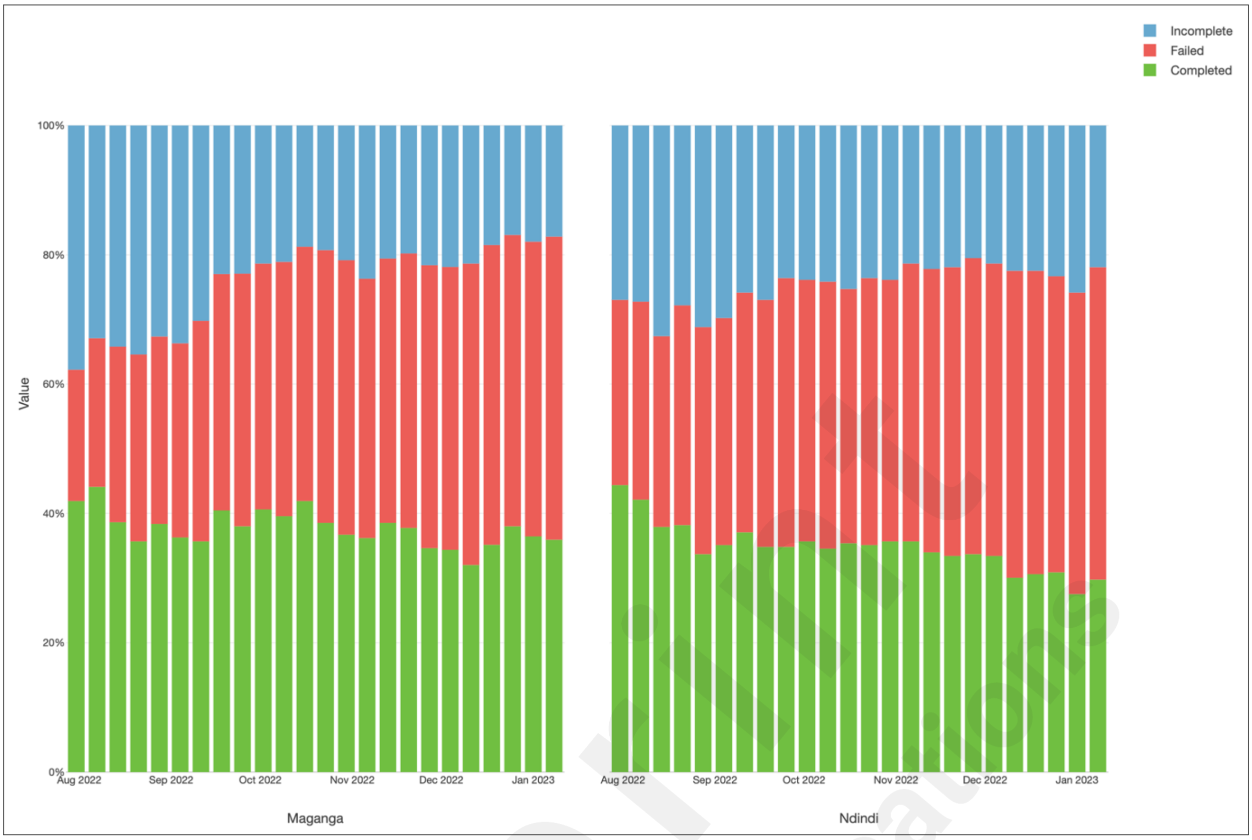


Figure 3. Participation profile rate according to the status of the IVR call disaggregated by TA areas.

Table 3. Attack rates at household level based on the signals report during the study.

Location	Signal			
	Diarrhea-like reports	Fever-like reports	Vomiting/Nausea-like reports	CLDD reports
Maganga	0.1398	0.2043	0.1327	0.1699
Ndindi	0.1203	0.2059	0.1224	0.1539

Table 4. Statistical summary for the time-series of signals captured by the IVR strategy, considering the aggregated data for Salima district.

Signal	R^2	Deviance explained	GCV	Scale estimate
CLDD	0.681404	0.774625	0.0624334	0.0423252
Diarrhea-like reports	0.831161	0.885469	0.0422496	0.0274658
Fever-like reports	0.92951	0.954495	0.0420778	0.0260313
Vomiting/Nausea-like reports	0.587607	0.692355	0.081265	0.0580976

Table 5. Statistical summary of GAM for the time-series of signals captured, considering the disaggregated data for Maganga TA and Ndindi TA areas.

TA area	Signal	R^2	Deviance explained	GCV	Scale estimate
Maganga	CLDD	0.596143	0.69602	0.020795	0.015
	Diarrhea-like reports	0.711076	0.796532	0.0256	0.017277
	Fever-like reports	0.899293	0.9338	0.016462	0.0103705
	Vomiting/Nausea-like reports	0.537684	0.632479	0.0247917	0.0188871
Ndindi	CLDD	0.421066	0.576651	0.0414793	0.0290681
	Diarrhea-like reports	0.711076	0.796532	0.0256	0.017277
	Fever-like reports	0.785284	0.851777	0.0343511	0.0227253
	Vomiting/Nausea-like reports	0.408624	0.560045	0.047594	0.0304854

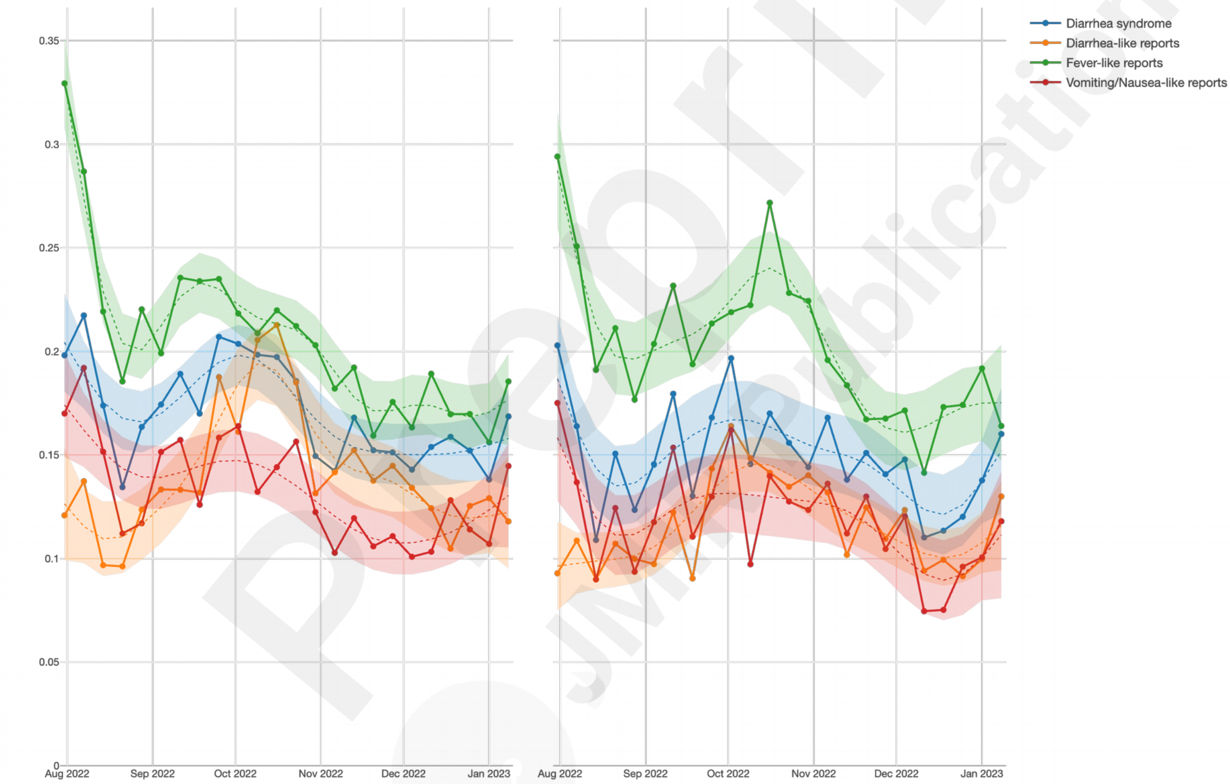


Figure 4. Time-series visualization of the signals including the trend line (GAM) of CLDD, Diarrhea-like reports, Fever-like reports and Vomiting/Nausea-like reports for Maganga TA and Ndindi TA. On y-axis the percentage of reports which answered positive for the symptoms or syndrome.

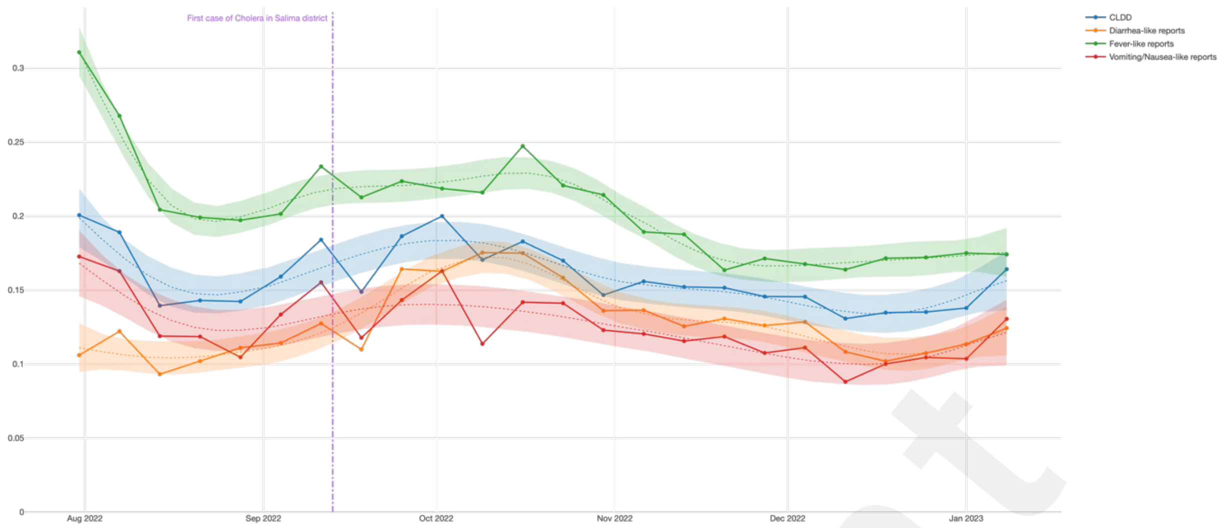


Figure 5. Time-series visualization of the signals including the trend line (GAM) of CLDD, Diarrhea-like reports, Fever-like reports and Vomiting/Nausea-like reports aggregated by Salima district. It includes also the date of the first Cholera case reported in the district. On y-axis the percentage of reports which answered positive for the symptoms or syndrome.

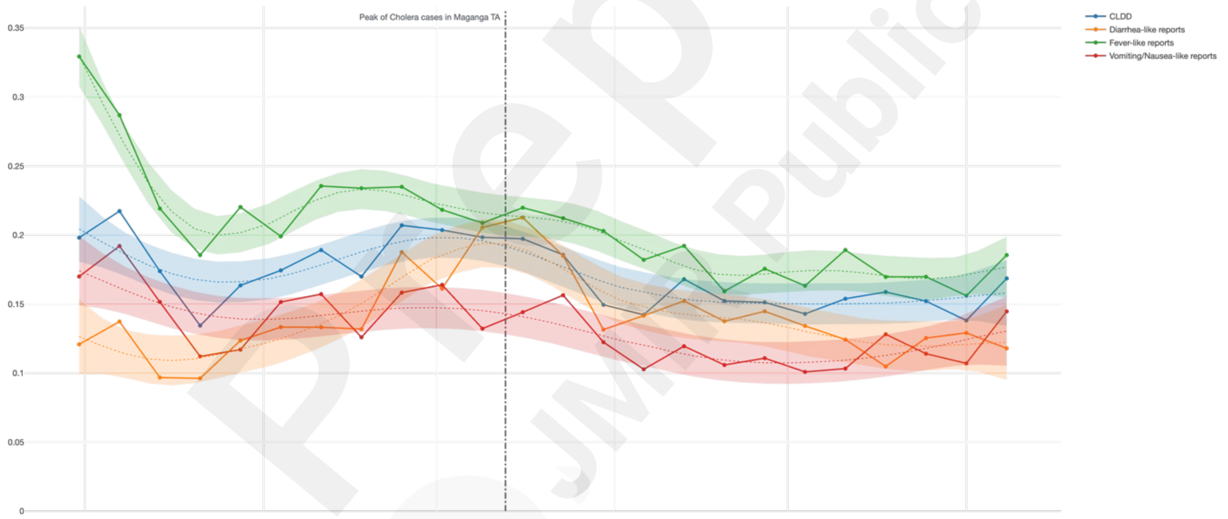


Figure 6. Time-series visualization of the signals including the trend line (GAM) of CLDD, Diarrhea-like reports, Fever-like reports and Vomiting/Nausea-like reports aggregated for Maganga TA. It includes also the peak of cholera cases in that TA area. On y-axis the percentage of reports which answered positive for the symptoms or syndrome.

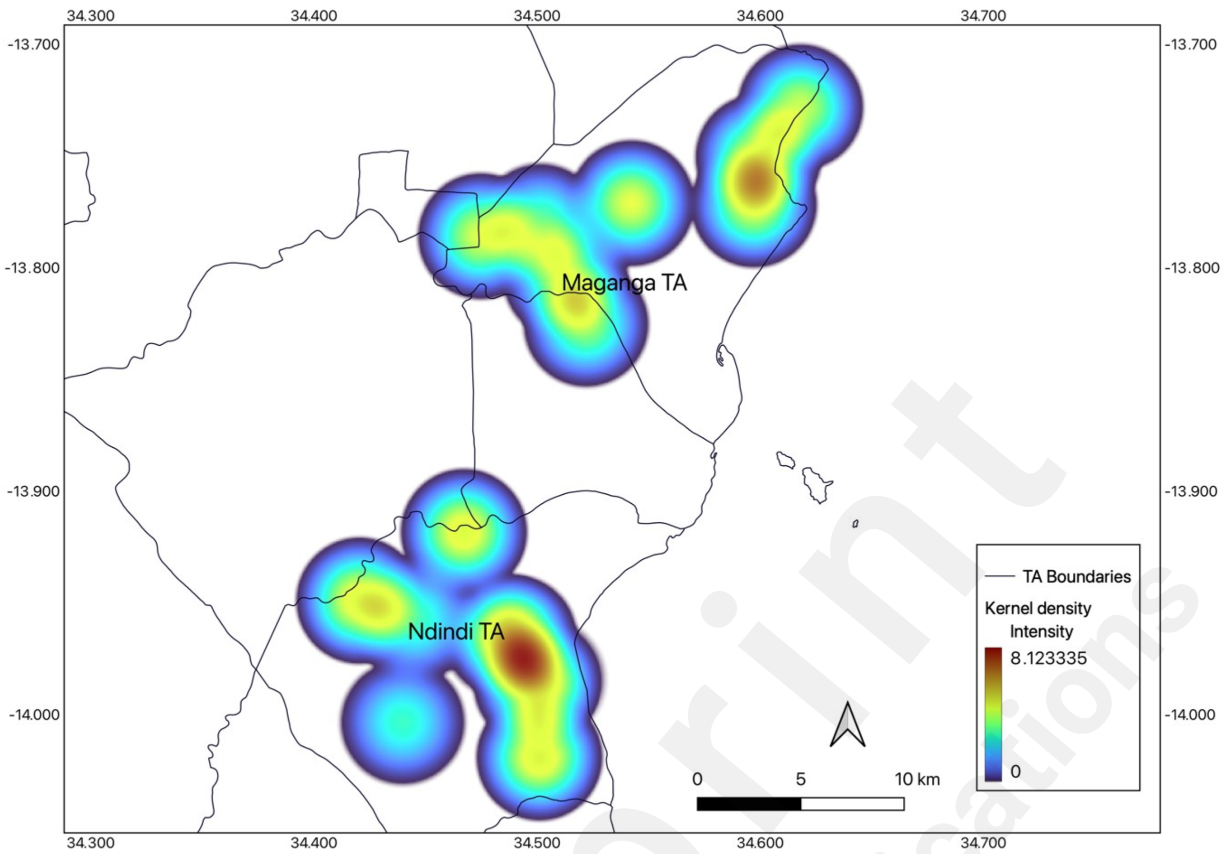


Figure 7. Kernel density estimator for the CLDD reports in both Maganga TA and Ndindi TA areas aggregating the whole period of the study.

DISCUSSION

In this study, we implemented a PS system using IVR in two districts of Malawi and assessed its feasibility, acceptability, and effectiveness in detecting early signals of CLDD outbreaks. During the 24 weeks of follow up, the system captured 16,280 observations and the average of the weekly participation rate was 35%, proving the strategy's feasibility and acceptability, according to the literature [24,44–46].

In summary, our findings demonstrate that this method might be effective in identifying CLDD with a notable and consistent signal being captured over time (R^2 = 0.681404). The signal showed a significant increase coinciding with cholera outbreaks in the region. This pattern was notably observed with the first cholera case in the Salima district, detected shortly after a spike in our CLDD data (Figure 5). Despite limited official reporting, this outbreak was confirmed through hyperlocal media sources. Furthermore, a subsequent analysis highlighted a peak in cholera cases in the Maganga TA area, preceded by a rise in CLDD reports of diarrhea-like symptoms underscoring the potential of CLDD in early outbreak detection and response facilitation. The local media released a 'Breaking News' report on October 15, 2022, highlighting the Cautionary Statement on Cholera, 13th October 2022 issued by the Salima District Council. Signed by the Director of Health and Social Services, the statement aimed to provide guidance to the population. It mentioned a total of 54 cases of cholera in Salima, emphasizing that TA Maganga accounted for 45 cases. The statement also underscored that the regions most affected were those in close proximity to the lake [43].

Traditional surveillance systems can have limitations in detecting early signals of outbreaks, particularly in resource-limited settings [7,10,14]. Therefore, PS has been proposed as an additional approach that leverages the local knowledge and resources of communities to identify early signals of outbreaks.

The successful implementation of a participatory surveillance strategy needs to take into consideration particularities of the site's geolocation, epidemiological profile, seasonality and the available telecommunication infrastructure [47]. For this particular study, Maganga and Ndindi encompass a vast territorial area bordered by Malawi Lake, the ninth-largest freshwater lake on Earth [48,49]. The lake plays a vital role in the region's economy, demographic density, and especially, in the definition of two different seasons - the dry and wet seasons [48]. Thus, during the wet season between December and April, the population expects floods followed by disruptions in the WASH infrastructure, and consequently, an increase in overall diarrhea cases [16,48].

Moreover, in 2019 there were more than five million mobile subscriptions, tending to a gradual increase in mobile penetration over the years, predominantly through pre-paid subscriptions [50]. The growth tendency of the mobile penetration could be confirmed as the reports pointed out approximately 12.27 million mobile subscribers at the beginning of 2022 [51]. Regarding the internet coverage, the International Telecommunication Union (ITU), showed that only 10% of the Malawi population was using the internet until 2019 [49]. Although the tendency of internet penetration is also growing, reports from 2022 showed that almost 80% of the country's population remained offline and the internet penetration is smaller out of the cities [52]. Bearing that in mind, models of PS done through mobile applications such as the Flu Near You, the AfyaData, or the Guardians of Health platform would still not provide a relevant engagement in the most remote areas of Malawi, and the choice of a self-report survey via IVR waves may be the most appropriate strategy for the setting [33,40,44,53,54].

Participatory Surveillance (PS) strategies have been applied worldwide showing positive results in different settings and events to trace pandemics such as Zika, Influenza, and Covid-19, as it has the potential to anticipate outbreaks or to give a quicker overview and guide the authorities on where the hotspots of that determined disease are [25,53–55]. Despite limitations in the field of technology and communication, PS strategies can rely on affordable systems that are successfully put into practice in areas considered to be under-resourced, such as Malawi [26]. By taking advantage of Interactive Response Voice Response (IVR) Design, it is possible to have in place a less expensive, flexible, scalable, and reliable system that captures data voluntarily and provides information that is not possible to capture by using traditional surveillance methods. Hence, empowering communities to take an active role in anticipating disease outbreaks even in settings where internet coverage is limited [34,56,57].

Our study presents several limitations. Regarding the setting for this sort of study design, population displacement and changes in the household configuration affected the consistency of the persons being reported inside that household phone number. Constant phone number changes, resulting in losing contact with an entire household. Pitfalls in the adherence rate over the weeks due to the fact of receiving the same call with the same questions every week can lead one to believe that the call is spam, and then start avoiding to answer that specific phone number. Selection bias is also an important limitation that should be mentioned. Not all individuals may have access to a mobile phone, and those who have may differ from those who do not have the device in terms of their socioeconomic status, education level, or other factors that may affect their likelihood of participating in the study. This reality could result in an unrepresentative sample and affect the generalizability of the results. Another limitation is that the quality and accuracy of the data collected through IVR may be affected by factors such as poor network coverage, low battery, or other technical issues that may prevent individuals from completing the survey or lead to missing data. Additionally, individuals may not always report their symptoms accurately or truthfully, which could affect the validity of the data collected through IVR. Validating participatory surveillance data with traditional sources in low-income settings is also challenging due to the scarcity of disease surveillance data. In many low-income settings, disease surveillance systems may be under-resourced or underdeveloped, resulting in delays in the collection, analysis, and reporting of disease data. As a result, it can be challenging to validate the accuracy of the data collected through participatory surveillance with traditional sources. This sum of constraints can lead to a lack of confidence in the accuracy of the participatory surveillance data and limit its usefulness for public health decision-making. Moreover, traditional sources of disease data, such as hospital records or laboratory test results, may not always be available or accessible in low-income settings, further complicating the validation process. Therefore, while participatory surveillance has the potential to provide valuable data on disease trends and outbreaks in low-income settings, its usefulness and advantages may be limited by the lack of timely and reliable validation data from traditional sources.

Conclusion

Interactive voice response (IVR) systems have the potential to facilitate participatory surveillance in low-income and low-resource countries by enabling efficient and cost-effective data collection. In this study's approach, weekly automated phone calls were made to a representative sample of the population over a six-month period, during which participants answered a consistent set of 10 questions related to cholera and malaria and associated risk factors. In low-income countries and settings where a low technological resource is faced, participatory surveillance using this method can be particularly useful in preventing CLDD outbreaks for several reasons. Firstly, the timely detection of cases via IVR due to the rapid identification of suspected CLDD cases within communities, providing public health officials with real-time data to respond promptly and contain potential outbreaks. Secondly, by using a standardized set of questions, the IVR system ensures that data collected across different participants and time points are consistent and comparable, thus improving the reliability of the surveillance system. Additionally, the use of phone-based surveys enables the collection of data from geographically dispersed and hard-to-reach populations, overcoming logistical barriers typically encountered in low-resource settings. Finally, by longitudinally monitoring the same set of participants over six months, the IVR-based participatory surveillance system can capture temporal trends and identify emerging risk factors, enabling targeted and context-specific interventions to prevent and control disease outbreaks.

Data Availability

Data related to this study is available upon request. Interested parties can contact the corresponding author to access the datasets, which are shared in accordance with applicable privacy and ethical guidelines.

Use of Generative AI

Generative AI was employed solely for proofreading purposes in this paper, focusing on grammar and phrase concordance, since part of the authors are non-native English speakers. It was not used to generate any content or intellectual material.

References

1. Abdulkadir SA. Cholera and cholera-like diarrhoeal diseases: Why the differences in severity? *Med Hypotheses*. 1991;34(3):278-281. doi:10.1016/0306-9877(91)90225-N

2. Msyamboza KP, Kagoli M, M'bang'ombe M, Chipeta S, Masuku HD. Cholera outbreaks in Malawi in 1998-2012: social and cultural challenges in prevention and control. *J Infect Dev Ctries*. 2014;8(06):720-726. doi:10.3855/jidc.3506

3. Sack DA, Sack RB, Nair GB, Siddique AK. Cholera. *The Lancet*. 2004;363(9404):223-233. doi:10.1016/S0140-6736(03)15328-7

4. Khonje A, Metcalf CA, Diggle E, et al. Cholera outbreak in districts around Lake Chilwa, Malawi: Lessons learned. *Malawi Med J*. 2012;24(2):29-33. doi:10.4314/mmj.v24i2

5. Swerdlow DL, Malenga G, Begkoyian G, et al. Epidemic cholera among refugees in Malawi, Africa: treatment and transmission. *Epidemiol Infect*. 1997;118(3):207-214. doi:10.1017/S0950268896007352

6. Kamanula JF, Zambasa OJ, Masamba WRL. Quality of drinking water and cholera prevalence in Ndirande Township, City of Blantyre, Malawi. *Phys Chem Earth Parts ABC*. 2014;72-75:61-67. doi:10.1016/j.pce.2014.09.001

7. Zuckerman JN, Rombo L, Fisch A. The true burden and risk of cholera: implications for prevention and control. *Lancet Infect Dis*. 2007;7(8):521-530. doi:10.1016/S1473-3099(07)70138-X

8. Ali M, Nelson AR, Lopez AL, Sack DA. Updated Global Burden of Cholera in Endemic Countries. *PLoS Negl Trop Dis*. 2015;9(6):e0003832. doi:10.1371/journal.pntd.0003832

9. Lessler J, Moore SM, Luquero FJ, et al. Mapping the burden of cholera in sub-Saharan Africa and implications for control: an analysis of data across geographical scales. *The Lancet*. 2018;391(10133):1908-1915. doi:10.1016/S0140-6736(17)33050-7

10. Bwire G, Mwesawina M, Baluku Y, Kanyanda SSE, Orach CG. Cross-Border Cholera Outbreaks in Sub-Saharan Africa, the Mystery behind the Silent Illness: What Needs to Be Done? *PLOS ONE*. 2016;11(6):e0156674. doi:10.1371/journal.pone.0156674

11. Taylor DL, Kahawita TM, Cairncross S, Ensink JHJ. The Impact of Water, Sanitation and Hygiene Interventions to Control Cholera: A Systematic Review. *PLOS ONE*. 2015;10(8):e0135676. doi:10.1371/journal.pone.0135676

12. Codeço CT. Endemic and epidemic dynamics of cholera: the role of the aquatic reservoir. *BMC Infect Dis*. 2001;1(1):1. doi:10.1186/1471-2334-1-1

13. Paye NG, Miller RHB, Zhou C, Weamie SJY, Davies TA. Systematic Study and Meta-Analysis

of Risk Factors Associated with Cholera Outbreaks in Africa: A Review. *Adv Infect Dis*. 2021;11(2):240-260. doi:10.4236/aid.2021.112022

14. Joseph Wu TS, Kagoli M, Kaasbøll JJ, Bjune GA. Integrated Disease Surveillance and Response (IDSR) in Malawi: Implementation gaps and challenges for timely alert. Uthman O, ed. *PLOS ONE*. 2018;13(11):e0200858. doi:10.1371/journal.pone.0200858

15. United Nations - Malawi. Malawi Cholera Flash Appeal 2023 (February-June 2023). Published March 9, 2023. https://malawi.un.org/en/222346-malawi-cholera-flash-appeal-2023

16. World Health Organization African Region. *Cholera in the WHO African Region: Weekly Regional Cholera Bulletin*.; 2023. https://www.afro.who.int/publications/cholera-who-african-region-weekly-regional-cholera-bulletin-29-march-2023-cloned

17. World Health Organization. Disease Outbreak News - Cholera, Malawi. who.int. Published February 9, 2023. https://www.who.int/emergencies/disease-outbreak-news/item/2022-DON435

18. Connolly MA, Gayer M, Ryan MJ, Salama P, Spiegel P, Heymann DL. Communicable diseases in complex emergencies: impact and challenges. *The Lancet*. 2004;364(9449):1974-1983. doi:10.1016/S0140-6736(04)17481-3

19. Odevall L, Hong D, Digilio L, et al. The Euvichol story – Development and licensure of a safe, effective and affordable oral cholera vaccine through global public private partnerships. *Vaccine*. 2018;36(45):6606-6614. doi:10.1016/j.vaccine.2018.09.026

20. Global Task Force on Cholera Control (GTFCC). *ROADMAP 2030 A New Global Strategy to Reduce Cholera Deaths and Stop Cholera Transmission*. https://www.gtfcc.org/about-cholera/roadmap-2030/

21. Wójcik OP, Brownstein JS, Chunara R, Johansson MA. Public health for the people: participatory infectious disease surveillance in the digital age. *Emerg Themes Epidemiol*. 2014;11(1):7. doi:10.1186/1742-7622-11-7

22. McNeil C, Verlander S, Divi N, Smolinski M. The Landscape of Participatory Surveillance Systems Across the One Health Spectrum: Systematic Review. *JMIR Public Health Surveill*. 2022;8(8):e38551. doi:10.2196/38551

23. Suy Lan C, Sok S, Chheang K, et al. Cambodia national health hotline - Participatory surveillance for early detection and response to disease outbreaks. *Lancet Reg Health West Pac*. 2022;29:100584. doi:10.1016/j.lanwpc.2022.100584

24. Leal-Neto O, Egger T, Schlegel M, et al. Digital SARS-CoV-2 Detection Among Hospital Employees: Participatory Surveillance Study. *JMIR Public Health Surveill*. 2021;7(11):e33576. doi:10.2196/33576

25. Pagliari C, Vijaykumar S. Digital Participatory Surveillance and the Zika Crisis: Opportunities and Caveats. Churcher TS, ed. *PLoS Negl Trop Dis*. 2016;10(6):e0004795. doi:10.1371/journal.pntd.0004795

26. Neto OL, Haenni S, Phuka J, et al. Combining Wearable Devices and Mobile Surveys to Study Child and Youth Development in Malawi: Implementation Study of a Multimodal Approach.

JMIR Public Health Surveill. 2021;7(3):e23154. doi:10.2196/23154

27. Mahmud AS, Chowdhury S, Sojib KH, et al. Participatory syndromic surveillance as a tool for tracking COVID-19 in Bangladesh. *Epidemics*. 2021;35:100462. doi:10.1016/j.epidem.2021.100462

28. Guenin MJ, Nys HMD, Peyre M, et al. A participatory epidemiological and One Health approach to explore the community's capacity to detect emerging zoonoses and surveillance network opportunities in the forest region of Guinea. *PLoS Negl Trop Dis*. 2022;16(7):e0010462. doi:10.1371/journal.pntd.0010462

29. Garg S, Bhatnagar N, Gangadharan N. A Case for Participatory Disease Surveillance of the COVID-19 Pandemic in India. *JMIR Public Health Surveill*. 2020;6(2):e18795. doi:10.2196/18795

30. Ngwa MC, Liang S, Mbam LM, et al. Cholera public health surveillance in the Republic of Cameroon-opportunities and challenges. *Pan Afr Med J*. 2016;24:222. doi:10.11604/pamj.2016.24.222.8045

31. Ohene SA, Klenyuie W, Sarpeh M. Assessment of the response to cholera outbreaks in two districts in Ghana. *Infect Dis Poverty*. 2016;5(1):99. doi:10.1186/s40249-016-0192-z

32. Wittwer S, Paolotti D, Lichand G, Leal Neto O. Participatory Surveillance for COVID-19 Trend Detection in Brazil: Cross-sectional Study. *JMIR Public Health Surveill*. 2023;9:e44517. doi:10.2196/44517

33. Karimuribo ED, Mutagahywa E, Sindato C, et al. A Smartphone App (AfyaData) for Innovative One Health Disease Surveillance from Community to National Levels in Africa: Intervention in Disease Surveillance. *JMIR Public Health Surveill*. 2017;3(4):e94. doi:10.2196/publichealth.7373

34. Gibson DG, Pereira A, Farrenkopf BA, Labrique AB, Pariyo GW, Hyder AA. Mobile Phone Surveys for Collecting Population-Level Estimates in Low- and Middle-Income Countries: A Literature Review. *J Med Internet Res*. 2017;19(5):e139. doi:10.2196/jmir.7428

35. Daftary A, Hirsch-Moverman Y, Kassie GM, et al. A Qualitative Evaluation of the Acceptability of an Interactive Voice Response System to Enhance Adherence to Isoniazid Preventive Therapy Among People Living with HIV in Ethiopia. *AIDS Behav*. 2017;21(11):3057-3067. doi:10.1007/s10461-016-1432-8

36. Byonanebye DM, Nabaggala MS, Naggirinya AB, et al. An Interactive Voice Response Software to Improve the Quality of Life of People Living With HIV in Uganda: Randomized Controlled Trial. *JMIR MHealth UHealth*. 2021;9(2):e22229. doi:10.2196/22229

37. Greenleaf AR, Mwima G, Lethoko M, et al. Participatory Surveillance of COVID-19 in Lesotho via Weekly Calls: Protocol for Cell Phone Data Collection. *JMIR Res Protoc*. 2021;10(9):e31236. doi:10.2196/31236

38. Mwanandi LG, Mwanandi C. Impact of Family Structure on Household Economy in Malawi: A Case of Chinsapo 2 and Area 49, an Urban Setting in Lilongwe City in Malawi.

39. Leal Neto O, Dimech GS, Libel M, et al. Saúde na Copa: The World's First Application of Participatory Surveillance for a Mass Gathering at FIFA World Cup 2014, Brazil. *JMIR Public Health Surveill*. 2017;3(2):e26. doi:10.2196/publichealth.7313

40. Neto OL, Cruz O, Albuquerque J, et al. Participatory Surveillance Based on Crowdsourcing During the Rio 2016 Olympic Games Using the Guardians of Health Platform: Descriptive Study. *JMIR Public Health Surveill*. 2020;6(2):e16119. doi:10.2196/16119

41. Salima District registers cholera case. The Atlas Malawi. Published September 15, 2022. Accessed May 5, 2023. https://atlasmalawi.com/salima-district-registers-cholera-case/

42. Kulinji.com by Gladys Nthenda. Nkhata Bay leads with 1,002 Cholera cases. Kulinji.com. Published October 17, 2022. Accessed May 5, 2023. https://kulinji.com/article/news/health/2022/nkhata-bay-leads-1002-cholera-cases

43. Salima District Council. Breaking news today- Cautionary Statement on Cholera, 13th October 2022. Facebook Love Community Radio 104.5 FM - Salima. Published October 15, 2022. Accessed October 14, 2022. https://www.facebook.com/lovefmsa/photos/a.671824986502056/1866965190321357/?type=3

44. Smolinski MS, Crawley AW, Baltrusaitis K, et al. Flu Near You: Crowdsourced Symptom Reporting Spanning 2 Influenza Seasons. *Am J Public Health*. 2015;105(10):2124-2130. doi:10.2105/AJPH.2015.302696

45. Dalton CB, Carlson SJ, Butler MT, Feisa J, Elvidge E, Durrheim DN. Flutracking weekly online community survey of influenza-like illness annual report, 2010. *Commun Dis Intell Q Rep*. 2011;35(4):288-293.

46. Koppeschaar CE, Colizza V, Guerrisi C, et al. Influenzanet: Citizens Among 10 Countries Collaborating to Monitor Influenza in Europe. *JMIR Public Health Surveill*. 2017;3(3):e66. doi:10.2196/publichealth.7429

47. Smolinski MS, Crawley AW, Olsen JM, Jayaraman T, Libel M. Participatory Disease Surveillance: Engaging Communities Directly in Reporting, Monitoring, and Responding to Health Threats. *JMIR Public Health Surveill*. 2017;3(4):e62. doi:10.2196/publichealth.7540

48. UNESCO WHC. Lake Malawi National Park. UNESCO World Heritage Centre. Accessed December 19, 2022. https://whc.unesco.org/en/list/289/

49. Lake Malawi/Niassa/Nyasa | AGLI. Accessed December 19, 2022. https://africangreatlakesinform.org/article/lake-malawiniassanyasa

50. Handforth C Wilson M. *GSMA Digital Identity Country Report: Malawi 2019*.; 2019:24. https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2019/02/Digital- Identity-Country-Report.pdf

51. Valev N. *The Global Economy, Mobile Phone Subscribers: Malawi*. The Global Economy https://www.theglobaleconomy.com/Malawi/Mobile_phone_subscribers/

52. Kemp S. *Digital Report 2022: Malawi*.; 2022. https://datareportal.com/reports/digital-2022-Malawi

53. Drew DA, Nguyen LH, Steves CJ, et al. Rapid implementation of mobile technology for real-time epidemiology of COVID-19. *Science*. 2020;368(6497):1362-1367. doi:10.1126/science.abc0473

54. Paolotti D, Carnahan A, Colizza V, et al. Web-based participatory surveillance of infectious diseases: the Influenzanet participatory surveillance experience. *Clin Microbiol Infect*. 2014;20(1):17-21. doi:10.1111/1469-0691.12477

55. Guerrisi C, Turbelin C, Blanchon T, et al. Participatory Syndromic Surveillance of Influenza in Europe. *J Infect Dis*. 2016;214(suppl_4):S386-S392. doi:10.1093/infdis/jiw280

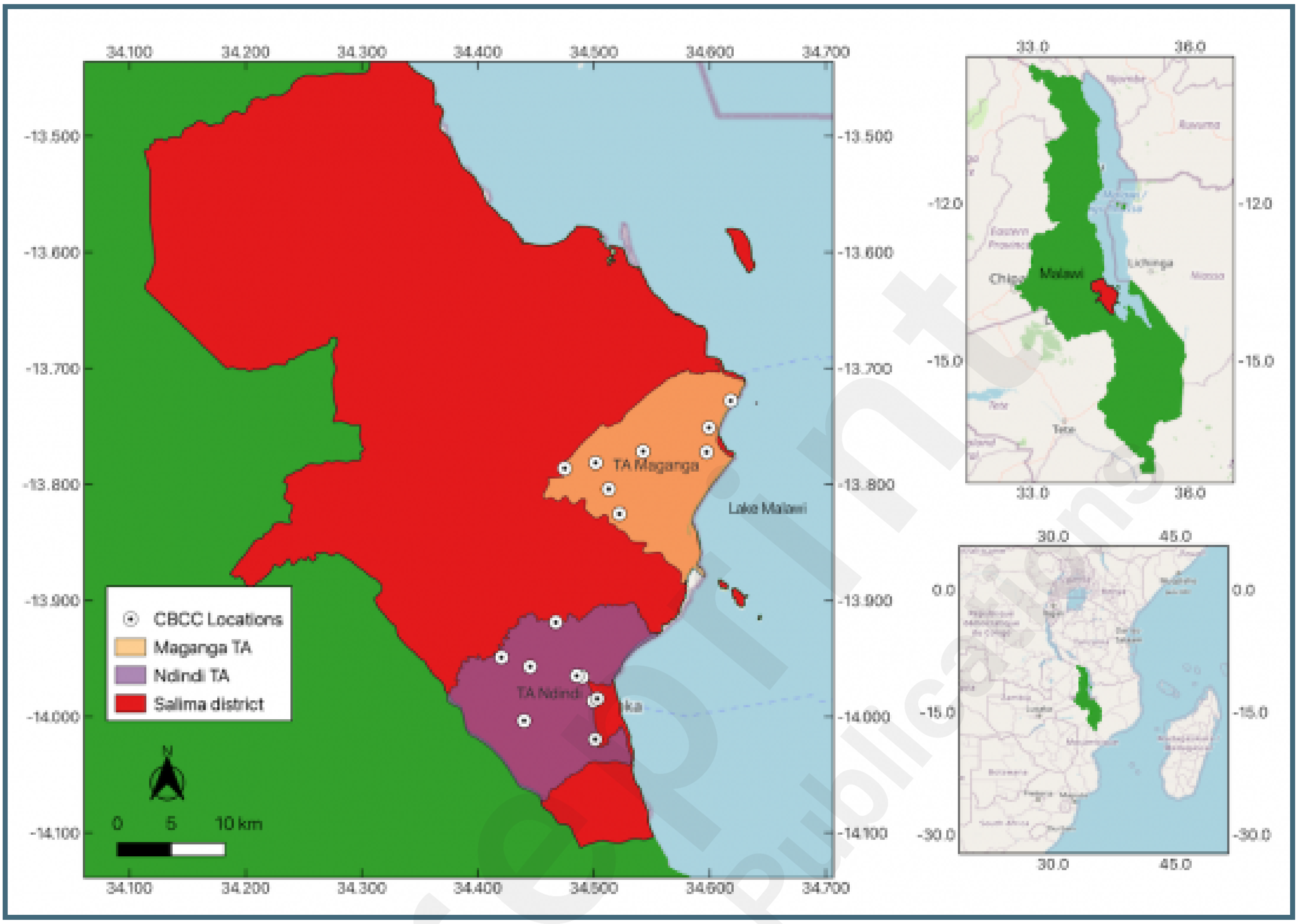
56. Klabbers RE, Muwonge TR, Pham P, et al. Leveraging interactive voice response technology to mitigate COVID-19 risk in refugee settlements in Uganda: Lessons learned implementing "Dial-COVID" a toll-free mobile phone symptom surveillance and information dissemination tool. *PLOS ONE*. 2023;18(1):e0279373. doi:10.1371/journal.pone.0279373

57. Lester J, Paige S, Chapman CA, et al. Assessing Commitment and Reporting Fidelity to a Text Message-Based Participatory Surveillance in Rural Western Uganda. *PloS One*. 2016;11(6):e0155971. doi:10.1371/journal.pone.0155971

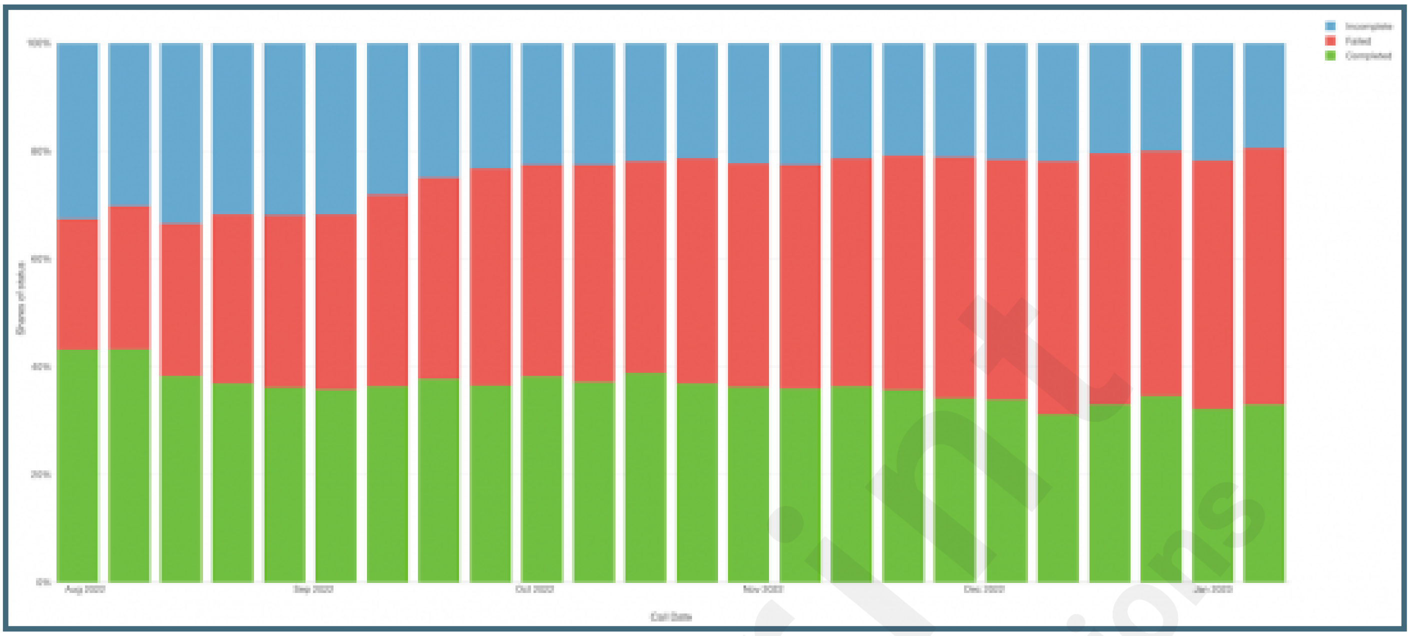
Supplementary Files

Figures

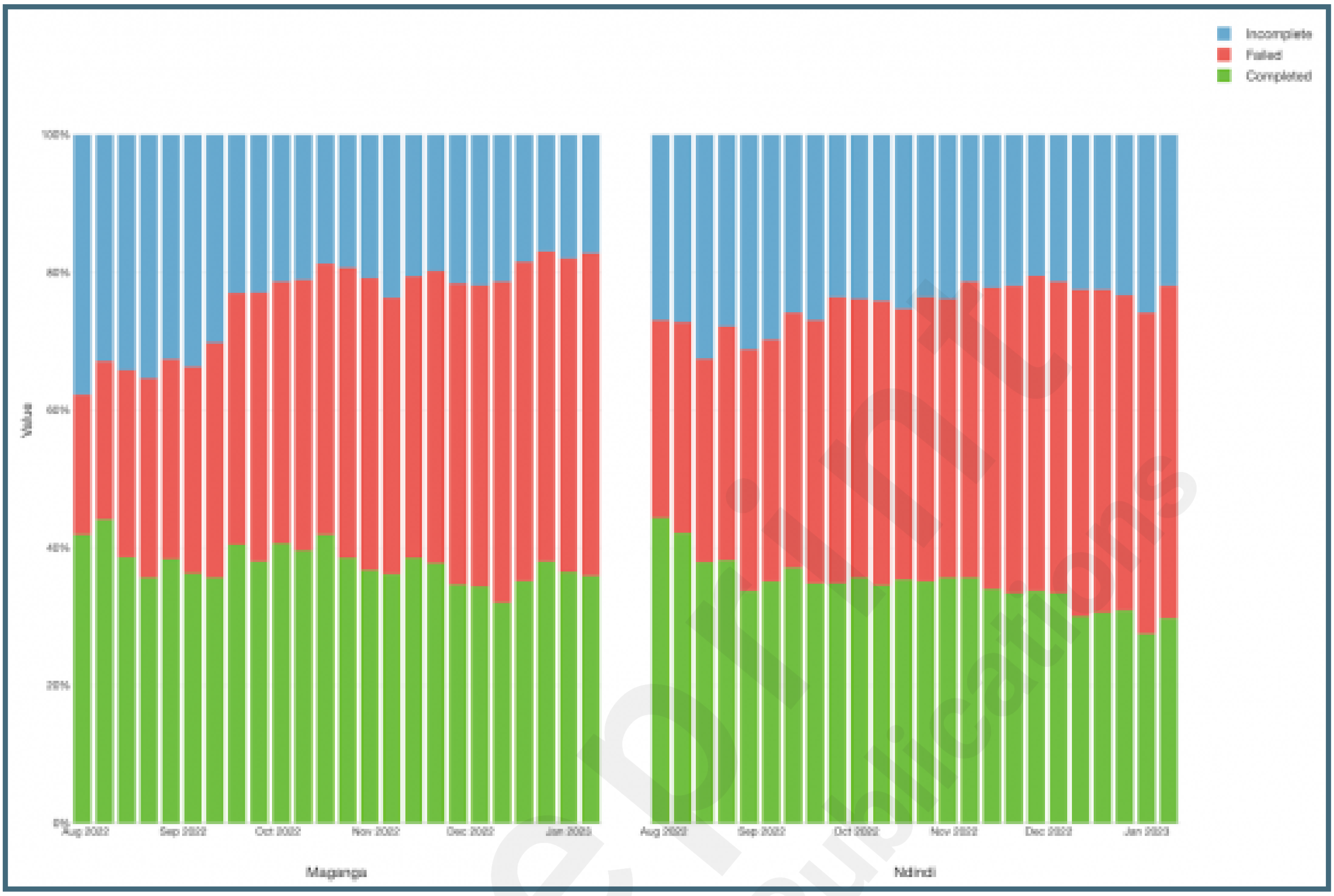
Study area highlighting the location of Salima district, Maganga TA, Ndindi TA and location of CBCCs.



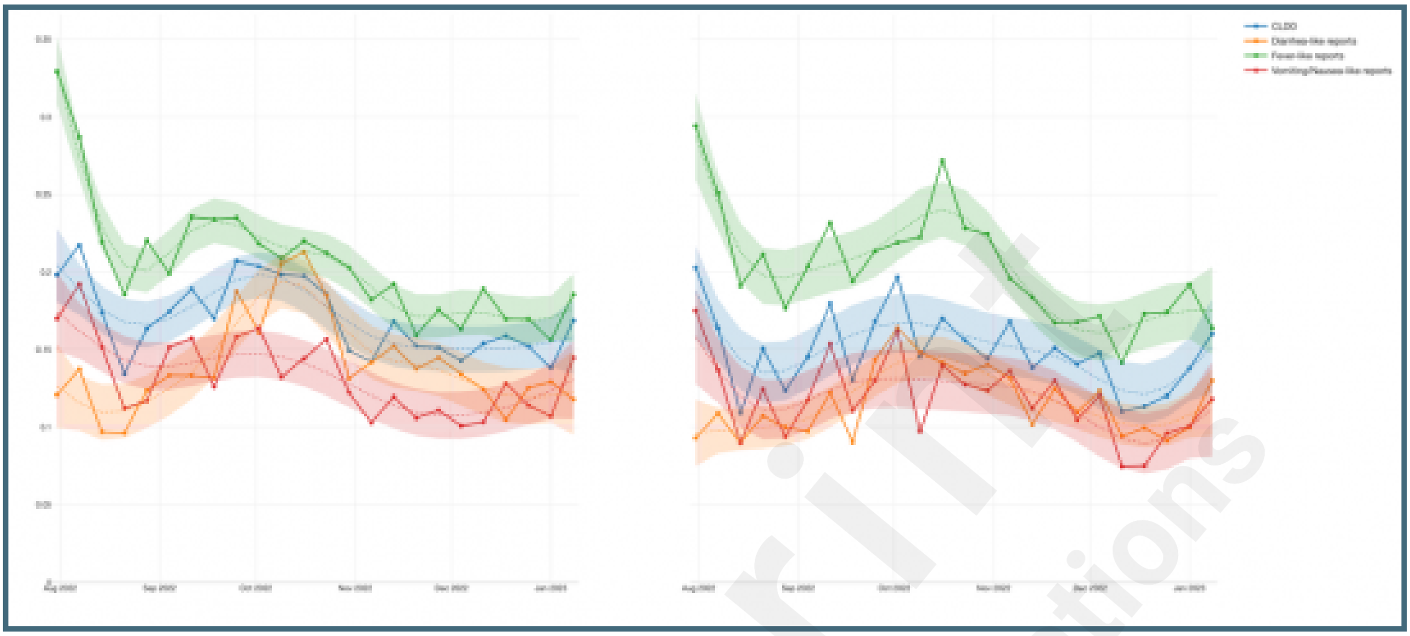
Participation profile rate according to the status of the IVR call aggregated.



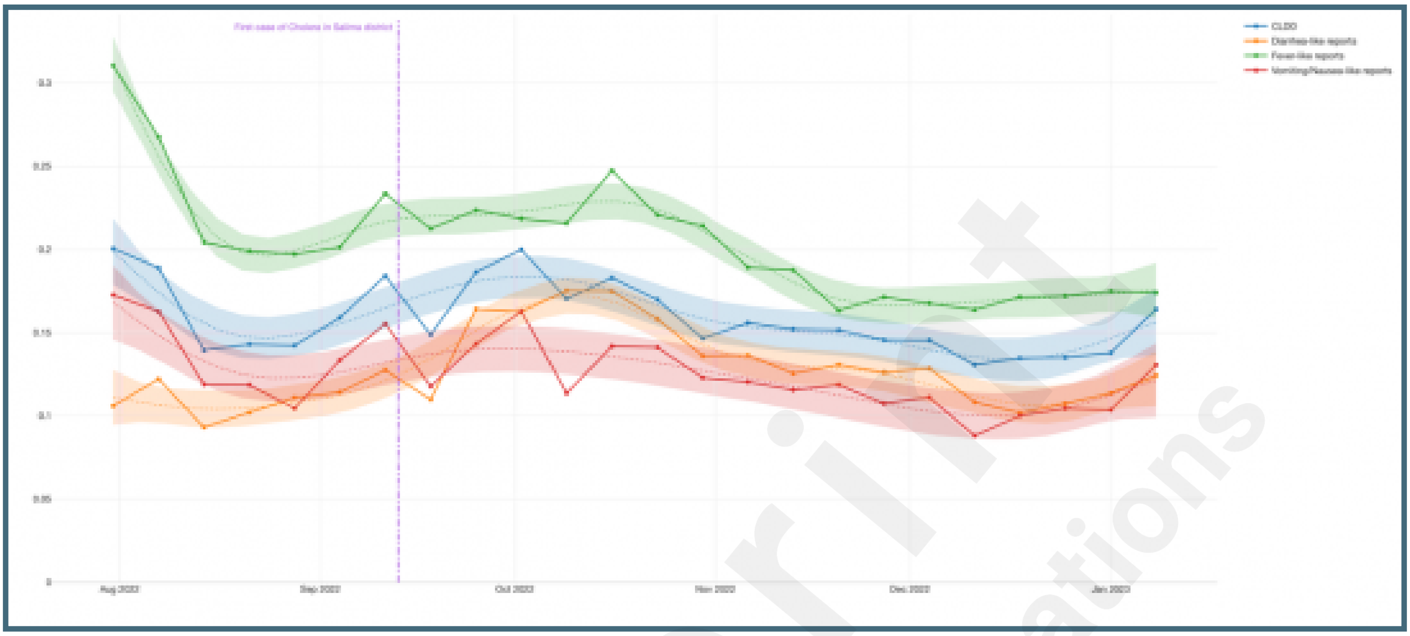
Participation profile rate according to the status of the IVR call disaggregated by TA areas.



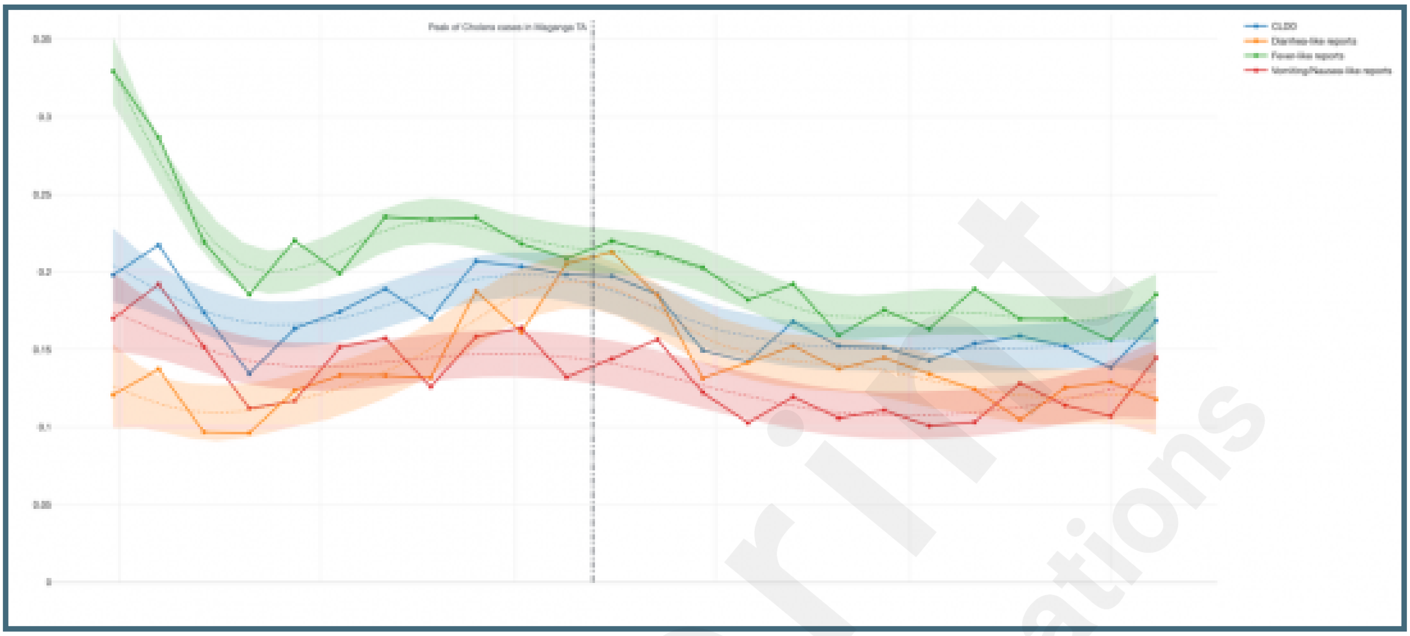
Time-series visualization of the signals including the trend line (GAM) of CLDD, Diarrhea-like reports, Fever-like reports and Vomiting/Nausea-like reports for Maganga TA and Ndindi TA. On y-axis the percentage of reports which answered positive for the symptoms or syndrome.



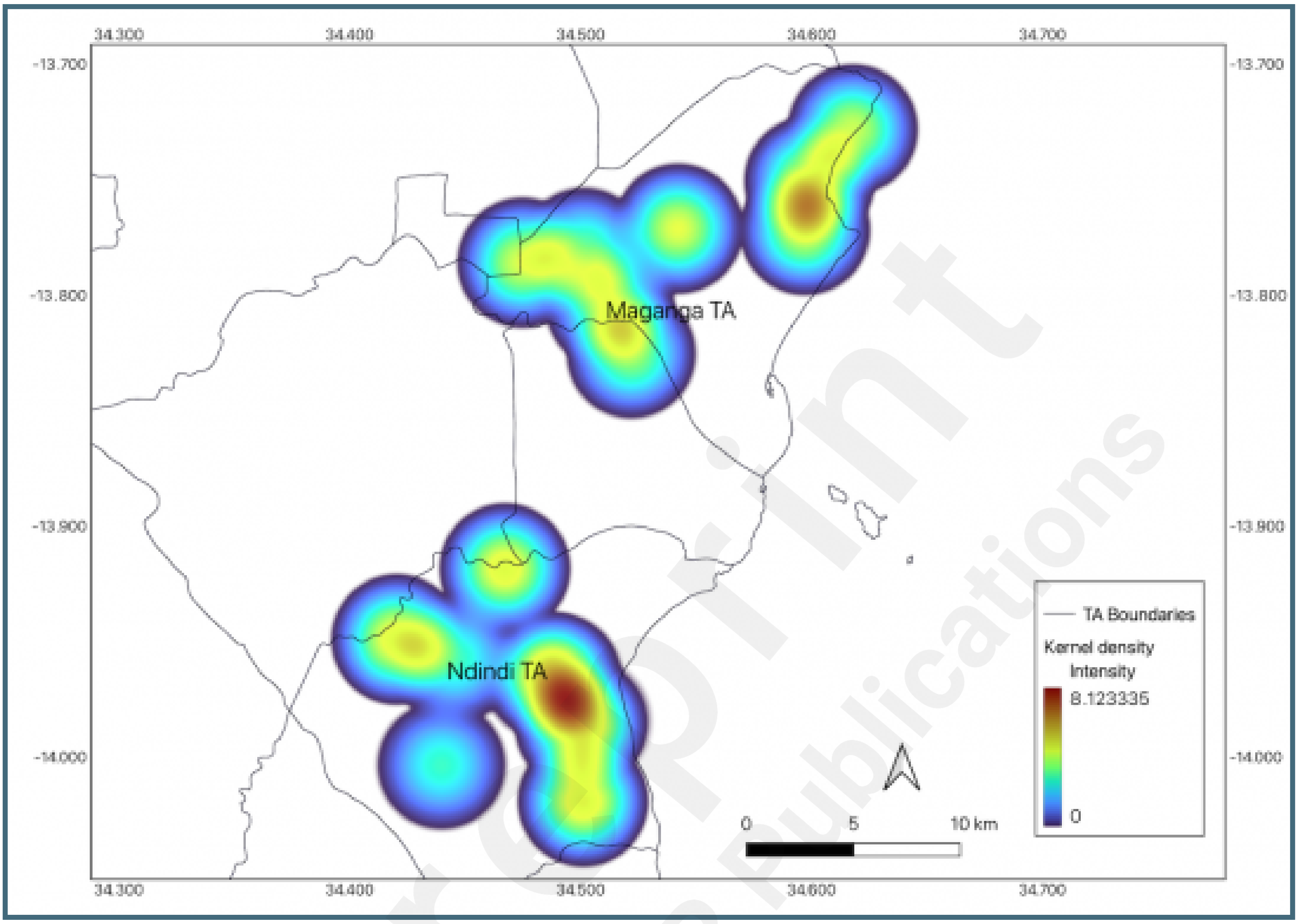
Time-series visualization of the signals including the trend line (GAM) of CLDD, Diarrhea-like reports, Fever-like reports and Vomiting/Nausea-like reports aggregated by Salima district. It includes also the date of the first Cholera case reported in the district. On y-axis the percentage of reports which answered positive for the symptoms or syndrome.



Time-series visualization of the signals including the trend line (GAM) of CLDD, Diarrhea-like reports, Fever-like reports and Vomiting/Nausea-like reports aggregated for Maganga TA. It includes also the peak of cholera cases in that TA area. On y-axis the percentage of reports which answered positive for the symptoms or syndrome.



Kernel density estimator for the CLDD reports in both Maganga TA and Ndindi TA areas aggregating the whole period of the study.



Related publication(s) - for reviewers eyes onlies

Text file with tracking changes.
URL: http://asset.jmir.pub/assets/ee048b8054ed221e405351ae9118178d.pdf

Text with track changes.
URL: http://asset.jmir.pub/assets/5aa41ee4202a39c96121344bfb6d1e14.pdf