A Citizen Observatory Approach for Developing a Disease Outbreak Early Warning System

Leah Mutanu Mwaura
United States International University-Africa
lmutanu@usiu.ac.ke

Dalton Ndirangu
United States International University-Africa
dndirangu@usiu.ac.ke

Patrick Wamuyu
United States International University-Africa
pwamuyu@usiu.ac.ke

Abstract

In health matters, early warning systems are timely surveillance systems that collect information on epidemic-prone diseases to trigger prompt public health interventions. However, these systems rarely apply statistical methods to detect changes in trends or sentinel events that would require intervention. Often, they rely on an in-depth review done by epidemiologists of the data coming in, which is rarely done systematically. This research introduced the use of ICT for collecting and analyzing citizen observations on disease trends and outbreaks. A citizen observatory ICT tool, which utilizes mobile and web features was developed. Data was collected on symptoms observed from diseases in four locations within Nairobi city. The system made use of mathematical models and outlier detection techniques to detect observations that deviated from the expected pattern in the dataset. New clusters were considered as outliers and the system flagged them as potential outbreaks. We clustered data using a K-Means algorithm and the Euclidean distance of each object from its corresponding cluster centre was obtained. From the results, the developed prototype was able to detect an outbreak of Flu and URTI diseases for the period of study. The proposed tool can therefore enhance the management of risks associated with disease outbreaks.

Keywords: *Early Warning Systems, citizen observatory, health surveillance, Outlier Detection, Modeling disease outbreak,*

1. Introduction

Early warning systems help stakeholders in better predicting and responding to natural disasters (Zschau, & Küppers, 2013). Recently the concept has found applications in predicting outbreaks (Semenza, J. C. 2015). Key innovations in automation of the process include solutions for collecting data and solutions for analyzing it (Liu, et, al. 2015). Citizen observatories are increasingly becoming a popular method for data collection in early warning systems (Brazzola, N., & Helander, S. E., 2018, Tapia, et al, 2014) with numerous applications in climate warning systems (Henriksen et al, 2018, Zommers & Singh 2014). A Citizen Observatory is defined as a software platform used to obtain volunteered information about a specific topic through different devices (Degrossi, Albuquerque, Fava, & Mendiondo, 2014). In health, Mathematical models are used for analytics and they range from simple compartmental transmission equations to complex equations that factor in environmental factors (Racloz, et, al. 2013). Experiments on these models are however confined to specific geographic areas or diseases such as dengue fever (Racloz, et, al. 2013, Hussain-Alkhateeb, et al, 2021) raising questions on their scalability. This study sets out to apply mathematical models in different geographic locations and on different diseases to evaluate the efficacy of the approach.

Kenya, like many developing countries, relies on health practitioners to provide information on disease outbreaks. These practitioners go out into the field to collect data from cases on the ground and analyse it. Invariably, they are unable to predict the outbreak before it occurs and have to deal with a reactive approach to handling disease outbreaks. Health problems are becoming an increasingly important issue, especially in developing countries. Therefore, we propose employing citizens' own digital devices to communicate environmental and human risk factors and risk mitigation strategies in the context of environmental and human health. Within a Citizen Observatory, citizens observe environmental conditions, receive a short-term (immediate) benefit and the Citizen Observatory itself creates a long-term positive impact on human health. There is need for a more proactive approach of involving citizens in the locality who can observe the spontaneous flare of diseases. The use of technology to develop a citizen observatory tool that acts as an early warning system would bridge this gap.

The rest of the paper is structured as follows: related work is presented in Section 2 while the methodology is presented in Section 3. In Section 4, the findings and discussion are provided. Section 5 has the conclusion drawn from the results of the research work.

2. Related Work

The term disease outbreak refers to an epidemic limited to localized increase in the incidence of disease (WHO, 2009). Epidemiologists classify outbreaks in different ways depending on how the disease spreads. Common source outbreaks begin from a single origin and last for a limited period of time and in a limited geographical location. They are characterized by dramatic single peak incidences and often go unreported due to the small numbers involved. Propagated outbreaks are normally caused by infections that can be passed from one person to the next. They continue over an extended period. Most outbreaks of respiratory diseases and some food or water borne outbreaks are propagated infections. The past decade alone has seen numerous propagated disease outbreaks ranging from Zika Virus, Ebola Virus, and recently Corona Virus. These outbreaks resulted in numerous losses of life and devastating economic impacts. A proactive approach to addressing the outbreaks would mitigate these challenges. By collecting data from citizen on the ground and employing predictive analytics, an early warning solution developed would enrich this proactive approach.

2.1 Citizen Observatories

Palacin-Silva et al. (2016) perceived that citizen science has rapidly spread in the last decades around the world as a genuinely interactive and inclusive opportunity for engaging citizens in the continuous collection of data relevant for science, governance, businesses, communal living, and individual concerns. Grossberndt, and Liu (2016), carried a study on citizen participation approaches in environmental health. They concurred that Environmental health is a very complex topic since it involves a whole range of factors which in many instances, has many high risks and uncertainty. In another related study, Zaman et al. (2014), advocate the use of participatory sensing, which appropriates wearable devices such as mobile phones to enable ad-hoc, person-centric mobile sensing networks. They conclude by emphasizing the need for campaign support, which empowers communities by providing them with a tool to answer local concerns and set up grassroots sensing actions. Lanfranchi et al. (2014) reaffirmed that Citizens' observatories are emerging as a means to establish interaction and co-participation between citizens and authorities during both emergencies and the day-to-day management of fundamental resources. Liu et al. (2014), in their study on environmental health, reiterated that there has been a trend to view the Citizens' Observatory as an increasingly essential tool that provides an approach for better observing, understanding, protecting, and enhancing our environment. They concluded by underscoring the need for the development of novel monitoring technologies and advanced data management strategies to survey, capture and analyze the data, thus facilitating their exploitation of policy for society. Crowd sourcing applications make use of a two-tier client-server architecture comprising of the citizen client side and the cloud server side. The client side collects data from key stakeholders using devices such as mobile phones and transmits it to servers on the cloud where analytics are conducted using mathematical algorithms informed by experts such as the ministry of health (MOH). An illustration of the crowd sourcing architecture described by Kurtah, et, al. (2019) is shown in Figure 1. This architecture inspired the design of the solution in this study. The challenge, however, remains with ICT out of reach of the poor especially those in rural areas. As recommended in (POST, 2011), new ICTs such as the internet and mobile phones can help combat this challenge. Statistics from CCK[1] show that the adoption of mobile phones has been rapid even among low-income earners with over 25 million mobile phone subscribers. The observations in (Arunga, 2007) attribute this rapid adoption to low cost and technical simplicity. Therefore, mobile solutions were adopted for this study.

[1]CCK - Communications Commission of Kenya

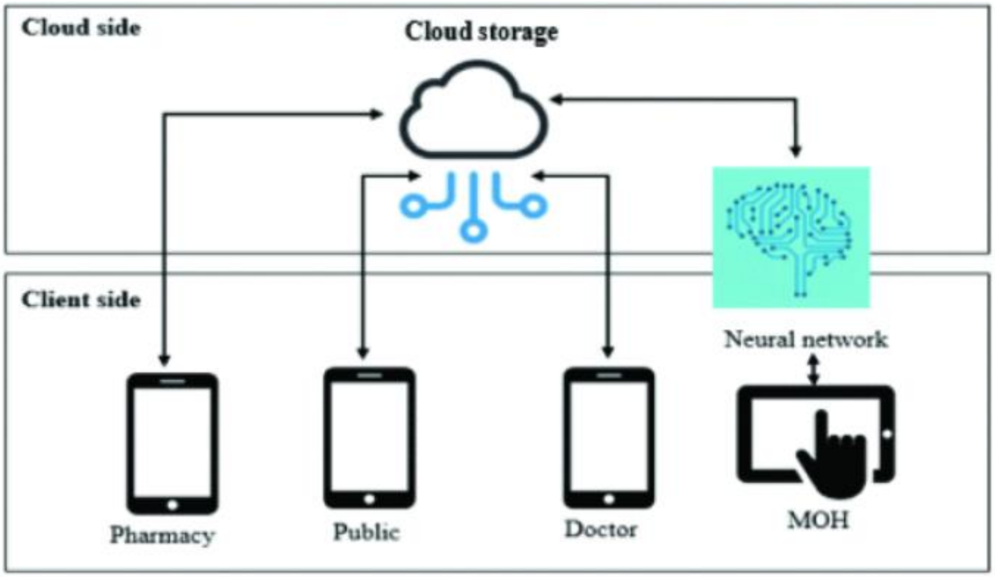


Fig 1. Crowd sourcing architecture (Kurtah, et, al. 2019)

2.2 Early Warning Systems

One of the Kenyan vision 2030 objectives for the health sector is the provision of "Preventive Health Care" services for all Kenyans (Government of Kenya, 2007); it aims specifically at enhancing the promotion of individual health and lifestyle (Ndung'u, Thugge, & Otieno, 2009). ICT can play a role in embracing Sustainable Development Goals such as combating serious disease by improving access to information (POSTNOTE, 2011). For example, the use mobile phone text messaging (Zurovac et al, 2012) can improve the delivery of health services and health outcomes in terms of an EWS in the context of malaria control in Africa. Another example is the EWS-development in the context of the severe acute respiratory syndrome (SARS), (Eysenbach, 2003). The public health and infectious disease research community widely praised the role of ICT in early detection as well as in fostering global collaboration and information exchange during the SARS epidemic, which broke out in 2002. Disease early warning systems have been developed in the health sector and implemented in some countries in the world such as Yugoslavia (WHO, 2017), Serbia (WHO, 2004), and Afghanistan (Noormal, 2011) among others. In Kenya there are no obvious pointers to the literature done on human disease EWS other than those for animal diseases compared to other countries. Additionally, in these countries, EWS depends on GIS information collection, relayed to health centres for dissemination to the citizen, and therefore they are not online systems. Improvements call for ensuring that the information is also accessible to the citizen directly whenever they need it. A citizen participatory early warning system should incorporate four key elements, namely: risk knowledge; monitoring and warning services; dissemination and communications; and response capability (Dutta, et al, 2015) as shown in Figure 2. This study starts by identifying risks based on common diseases found in the area and their definitive symptoms. A citizen observatory portal developed monitors the symptoms to forecast outbreaks. Output from the system is communicated to relevant stakeholders through reports to develop a response plan. The Early Warning Systems should be able to mine the data and perform cluster analysis to detect disease outbreaks. Given that disease outbreaks change with time, the system should be able to evolve through an adaptive process.

Fig. 2. Elements of an Early warning System (Luther, et al, 2017)

2.3 Mathematical Modelling

A study that identifies Epidemiologic Determinants for Modelling Pneumonic Plague Outbreaks observes that mathematical models based on historical data can quantitatively assess the transmissibility and potential health effects of plague outbreaks under a range of assumptions (Gani, R., & Leach, S., 2004). Mathematical models are fitted to data and their parameters (Susceptible, Infected, and Recovered) estimated (Martcheva, M., 2015, Keeling, & Rohani, 2011, Fraser et, al. 2004). A study that identified Epidemiologic Determinants for Modelling Pneumonic Plague Outbreaks observes that mathematical models based on historic data can quantitatively assess the transmissibility and potential health effects of plague outbreaks under a range of assumptions (Gani, R., & Leach, S., 2004). Mathematical models can aid in anomaly detection using statistical and machine learning-based techniques to identify disease outbreaks within a specific area (Muruti, et al., 2018). Statistical techniques include non-parametric models such as Histograms and parametric models such as regression models. Machine-learning based models include classification models using neural networks, clustering models, and nearest neighbor models such as distance-based models. This study explores both statistical and machine-learning based models using histograms and nearest neighbour outlier detection approaches. To describe how often a disease or another health event occurs in a population, the prevalence and incident rates serve as measures of disease frequency. The prevalence reflects the number of existing cases of disease while the incidence reflects the number of new cases of disease (Noordzij, et al, 2010). It is worth noting that outbreak forecasting requires an integrative approach to modelling because outbreaks vary by disease and social network heterogeneity therefore dynamic modelling approaches become useful for prediction (Scarpino, & Petri 2019).

The process of modelling disease outbreaks starts by collecting historical data and using it to develop models. Finally, predictions on new data conducted using the models as illustrated in Figure 3.

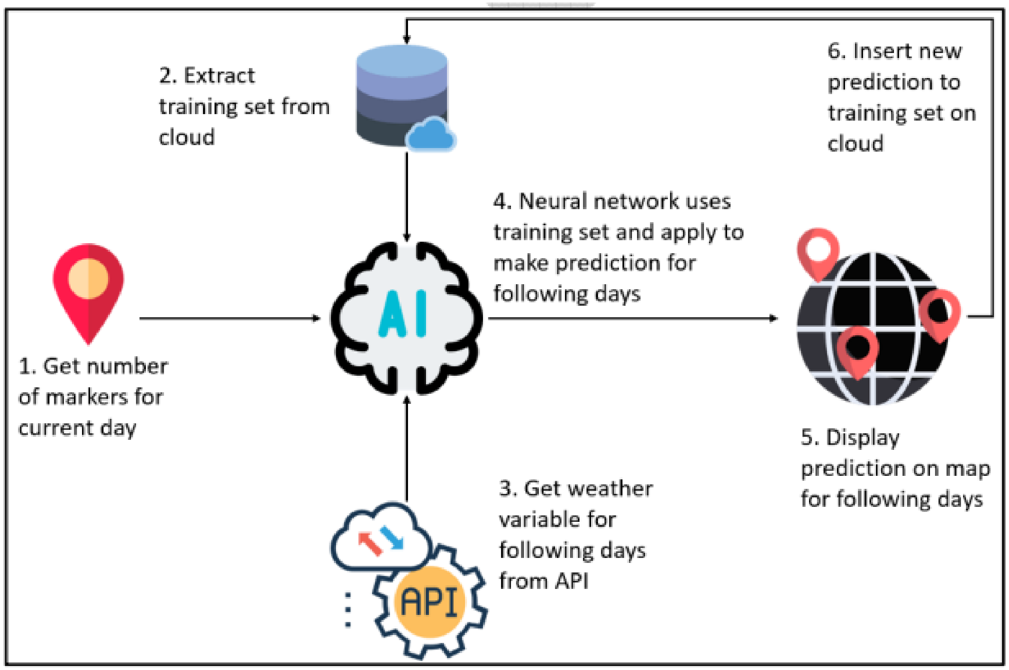


Fig. 3. Disease Propagation Prediction (Kurtah, et, al. 2019)

Existing studies employ different models for disease early warning systems. Each model is applicable to specific diseases and data types. A comparative study on early warning system for infectious diseases (Hu, et, al. 2021) highlights the need for solutions that can adapt to a wide range of diseases by investigating various diseases, technologies, geographic locations and socio-cultural information.

3. Methodology

The general objective of the project was to develop a citizen observatory tool used to gather data for predicting the outbreak of infectious diseases. Specifically, the researchers set out (i) to investigate the properties of infectious diseases within the Nairobi city, (ii) to develop a citizen Observatory tool to gather the information used to predict disease outbreaks, and (iii) to test and evaluate the efficacy of the proposed tool. The research focused on designing a tool that aids in proactively managing disease outbreaks. It targeted infectious diseases that are common in the tropics. The research used four different sites within the city that represent upper class, middle, and low-income areas. To maximize data collection, the study used health facilities where visiting patients provided symptoms. For this study, the researchers aimed at 50 respondents for the pilot and 1000 respondents for the actual study. The study sample was based on the proportion of mobile users and derived using equation (i), while the pilot sample was based on 5% of the population (equation (ii)) as recommended by Saunders et, al. (2009)

$$Study\ Sample = \frac{\boldsymbol{z^2\ p(1-p)}}{\boldsymbol{m^2}} \qquad \text{(i)}$$

$$Pilot\ Sample = 5\%\ of\ Study\ Sample \qquad \text{(ii)}$$

*Where: **z** = confidence level at 95% (standard value of 1.96), **p** = Kenyan mobile penetration according to CAK, **m** = margin of error at 3% (standard value of 0.03)*

Study Sample $= \dfrac{1.96^2 \text{ x } .0.479(1-.0.479)}{.03^2}$

$= 1067.1 \approx \mathbf{1067}$

Pilot Sample =0.05*1067

=**53**

The number of pilot respondents received was 60, while the study sample was 1027. We used purposeful sampling to identify the data collection sites due to time limitations and budget constraints. Data collection from these sites targeted every patient (100%) who visited the facility with symptoms of infectious diseases. A preliminary survey of common infectious diseases and the symptoms conducted during the feasibility study informed the design of the citizen observatory tool. The data collected was analysed with Data Analytic Tools to identify disease outbreaks. We tested the developed prototype in selected locations and monitored it for half a year to evaluate its efficacy. The results of this study are discussed in the next section.

4. Experimental Results and Discussions

This project took a Citizen Science approach for data collection whereby members of the public were involved in data collection. An abundance of mobile apps makes participation in Citizen Science projects easier by reducing the effort in data collection and entry. This is especially useful in longitudinal type of studies where data collection is conducted over a period of time rather than once.

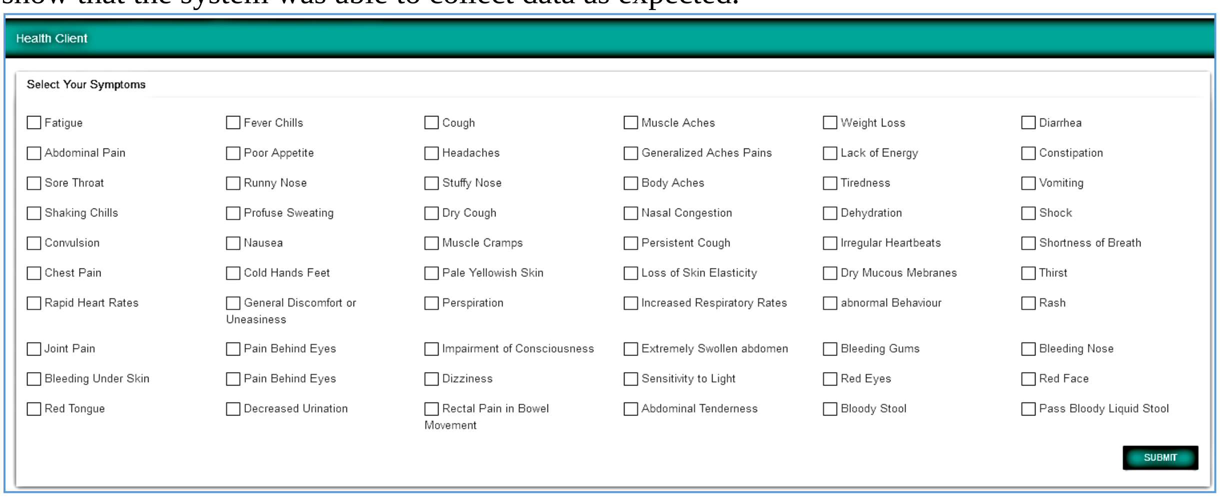
4.1 System Development

The development process started with the identification of methods, data, and metadata required across different tasks, platforms, and sites. Each observation site/data point had a universally unique identifier denoted by its geo-coordinates as seen in the sample data set provided Table 1

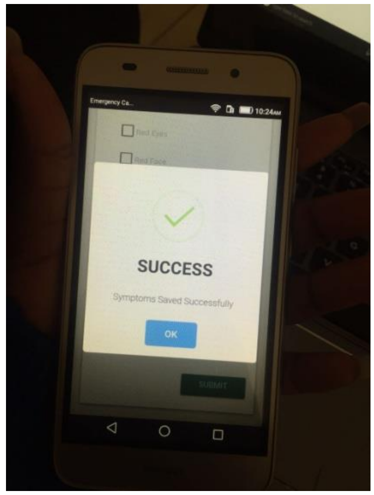
Table 1: Data collected uniquely identified by their geo-coordinates

Date	Device	Symptoms	Lt	Ln
9/25/2017 8:44	Browser	Bloody Stool,Abdominal Tenderness	-1.26295	36.80131
9/26/2017 5:24	ANDROID	Poor Appetite,Weight Loss,Cough	-1.263	36.80132
9/26/2017 5:24	ANDROID	Poor Appetite,Weight Loss,Cough	-1.263	36.80132
9/26/2017 4:13	ANDROID	Fever Chills,Weight Loss,Abdominal Pain	-1.26301	36.80132
9/26/2017 4:13	ANDROID	Cough,Abdominal Pain,Generalized Aches Pains	-1.26301	36.80132

The design of the mobile app developed, aimed at reducing the effort of the participant by ensuring all the queries were on one page. It is accessible via both a mobile phone (Figure 4 (a)) and a web browser as shown in Figure 4 (b). The presentation layer provides the client user interface. A Geolocation API determines the users location hence reducing on the user input requirements. To ensure privacy of the participants, no personal data was collected. Additionally, the data was removed from online repositories and research devices were returned to the researchers as soon as the data collection exercise was completed. Participation in the research was voluntary with permission obtained from the health facilities. The results presented in Table 1 show that the system was able to collect data as expected.



(a) App Questionaire



(b) App Interface confirming submission

Fig. 4. Prototype Interface

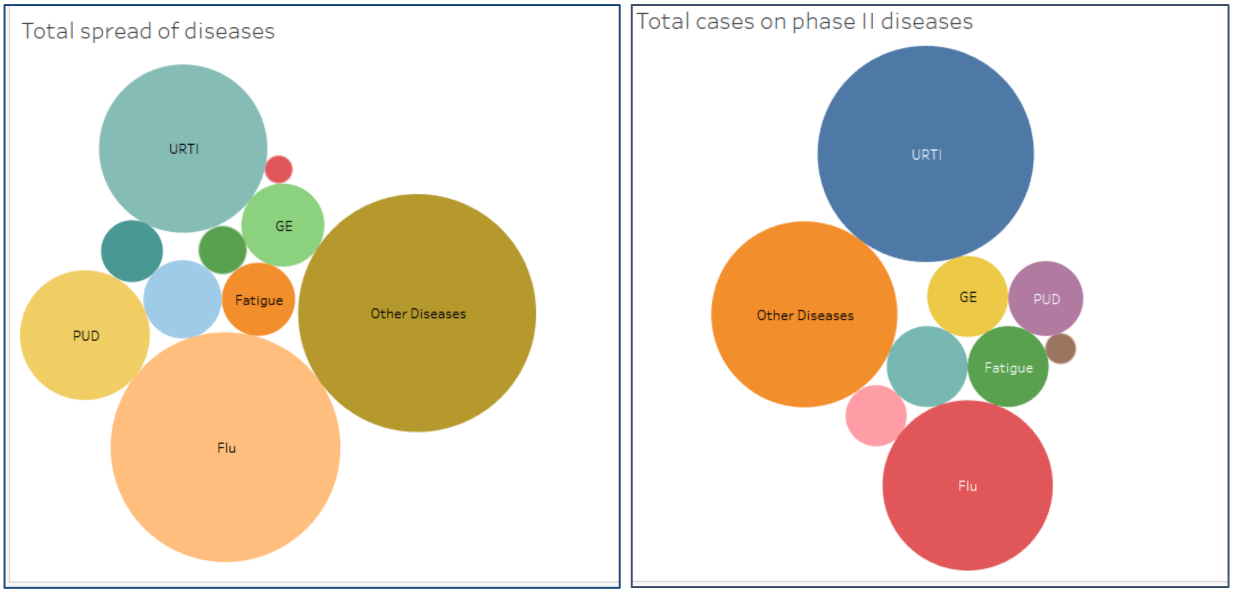
4.2 Data Analysis

The next step involved analysis of the data and for patterns that could indicate outbreaks. The sum total of all diseases across all sites was obtained and depicted using bubble charts. Bubble charts were used for comparing the relationships between data objects in three numeric-data dimensions: the X-axis data (Longitude), the Y-axis data (latitude), and data represented by the bubble size (Frequency of infections). The frequency of infections was obtained by the summation formula provided in equation (iii).

$$\sum_{i=n}^{m} a_i = a_n + a_{n+1} + a_{n+2} + \ldots + a_{m-2} + a_{m-1} + a_m \qquad \text{(iii)}$$

Where a_i represents a recorded incident and m is the number of incidences

The total of all diseases across all sites for different phases, illustrated in Figure 5, shows that Flu was the most common disease generally found in the city during the first phase (Figure 5(a)) while *Upper Respiratory Tract Infections (*URTI) was the most common disease when data was collected four months later (Figure 5 (b)).



(a) Phase I *(b) Phase II*

Fig. 5. Total Spread of Diseases across both Phases

Further analysis based on the different sites used revealed locations with higher incidences of diseases. The Map in Figure 6 shows the sites that data was collected from and the spread of diseases across the different sites during the first phase of data collection. The results show that some sites had higher incidences than others. Based on this the researchers analysed the incidences for different sites

Fig. 6. Spatial distribution of Diseases for Phase I

The results in Figure 7 show that different sites used for data collection encountered different disease types. Although Flu was the most common ailment generally experienced during the first phase, only two sites reported this as the most common ailment. A closer examination reveals that URTI, Allergenic Skin reactions, and Pulmonary Disease (PUD) were common in different sites. With this information, health stakeholders can proactively put checks in place to manage the diseases. This shows that citizens' observatories have the potential to proactively manage outbreaks by informing the efficient distribution of resources as highlighted by Lanfranchi et al. (2014).

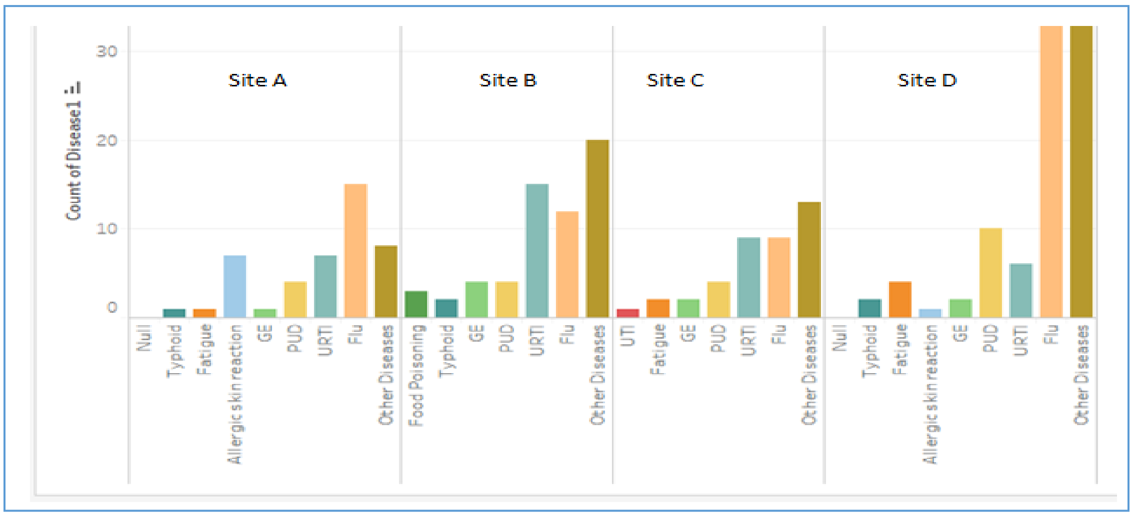


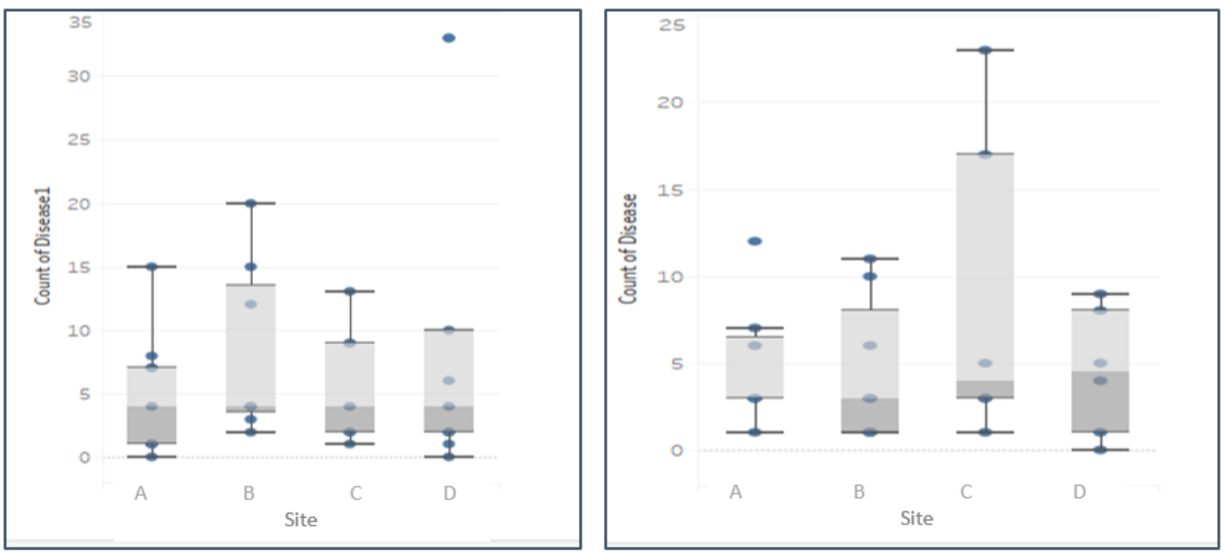
Fig. 7. Phase I Spread of Diseases Across different sites

Although frequency analysis enabled the researcher to identify the most common disease in a given location, they do not indicate an outbreak. To identify an outbreak further analysis was required to measure outliers i.e., deviation from other observations. In statistics, an outlier is an observation point that is distant from other observations. An outlier is an observation that lies an abnormal distance from other values in a random sample from a population. Traditionally outliers were data items considered anomalous due to several causes such as erroneous measurements or anomalous process conditions. However, recently they have found many applications, such as network intrusion, medical diagnosis, or fraud detection. In this study, outliers helped in the detection of emergent behaviors in disease patterns as an indicator of an outbreak. For example during the first phase the flu figures ranged between 2 to 10 cases, however on a specific day 34 cases were observed. The analytics algorithm flagged this as an outlier, and a possible indicator of an outbreak. In phase I Flu was an outlier detected in Site D (Figure 8(a)) while in Phase II Flu was an outlier detected in Site A as shown in Figure 8(b). We can therefore conclude that these two sites had a potential Flu outbreak during those two periods of data collection, subject to the accuracy of the data collection and diagnosis process.

For outlier detection we used the Nearest-Neighbor outlier detection technique which requires a distance (or similarity measure) defined between two data instances. It assumes that outliers are distant from their neighbors or that their neighborhood is sparse. According to Dang et. al., (2015), for a point *p(x, y),* the distance between *p* and its kth nearest neighbor $P_k(X_k, Y_k)$ is given by equation (iv).

$$D^k(p) = \sqrt{(y - y_k)^2 + (x - x_k)^2} \quad \text{(iv)}$$

To visualize the results, a box plot was used in this study. It is a useful graphical display for describing the behavior of the data in the middle as well as at the ends of the distributions. The box plot uses the median and the lower and upper quartiles (defined as the 25th and 75th percentiles). If the lower quartile is Q1 and the upper quartile is Q3, then the difference (Q3 - Q1) is called the interquartile range or IQ. Box plots have fences that show the outliers. A point beyond an inner fence on either side is considered a mild outlier. A point beyond an outer fence is considered an extreme outlier.



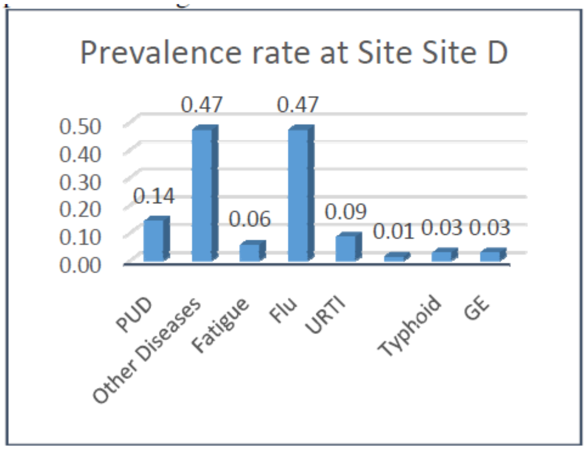
(a) Phase I *(b) Phase II*

Fig. 8. Outlier Detection of Diseases

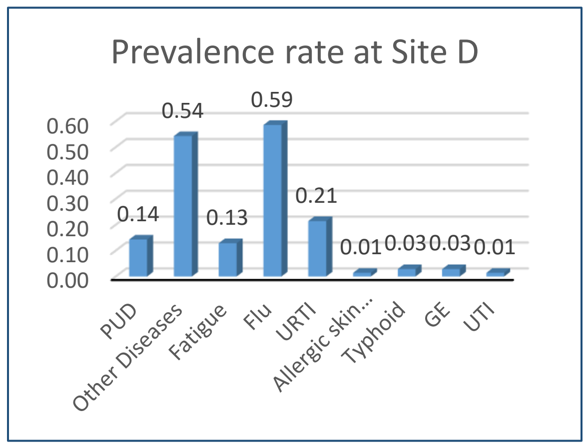
Although outliers are useful in detecting outbreaks, on their own they are not adequate as they do not factor in the size of the population. A more accurate picture is obtained when the number of cases is examined against the population. To achieve this, mathematical modelling was performed based on the work of Noordzij, et al, (2010). The epidemiology of infectious diseases involves the study of the prevalence, incidence, and determinants of infections in populations. Our study focused on the prevalence and incident rates. The prevalence of a disease is the proportion of a population that is infected during a period (period prevalence) or at a particular date in time (point prevalence). Our study focused on point prevalence due to the challenge of tracking individual cases over a period. Point prevalence is given by equation (v), where n is the population.

$$\frac{\textit{All new and pre-existing cases during a given time period}}{\textit{Population during the same time period}} \times 10^n \qquad \text{(v)}$$

From the results, we note that Flu was highly prevalent during both phases of data collection. There was also an increase in this prevalence rate during both periods. Food-related infections however remained low with relatively constant prevalence rates. A significant increase in the prevalence rate was noted for both URTI and Fatigue, an indicator of outbreaks. These results are presented in Figure 9(a & b) for both Phases I and II for Site D.



(a) Phase I



(b) Phase II

Fig. 9. Point Prevalence rates for Site D

A limitation of point prevalence is that it is based on a single examination. A better measure is the incident rate, which looks at a stated period. Prevalence differs from incidence in that; prevalence rate gave us an idea of how widespread the disease was. Incidence rate on the other hand gave us an idea of the risk/probability of contracting the disease. The incidence of a disease is therefore the rate at which new cases occur in a population during a specified period. A person-time rate is generally calculated from a long-term cohort follow-up study, wherein enrollees are followed over time and the occurrence of new cases of the disease is documented. Typically, each person is observed from an established starting time until one of four "endpoints" is reached: onset of disease, death, migration out of the study ("lost to follow-up"), or the end of the study. Similar to the incidence proportion, the numerator of the incidence rate is the number of new cases identified during the period of observation. Due to the time available for this study, the researchers were not able to track specific individuals over a long period but rather for four months only. We however show how this analysis would occur using data from different individuals and show the incident rate per location rather than per individual. The incidence rate is the ratio of the number of cases to the total time the population is at risk of disease.

The incidence rate is given by equation (vi).

$$\frac{\textit{Number of new cases of disease or injury during specified period}}{\textit{Time each person was observed, totaled for all persons}} \qquad \text{(vi)}$$

The results obtained when the incident rate was calculated are presented in Figure 10. From the results, we note that there was a risk of contracting URTI at Site D while there were food-poisoning related risks in Site B. This is more meaningful for an early warning system, which should facilitate a proactive approach for disease risk management. The data obtained from the prototype developed therefore provides useful information for disease risk management.

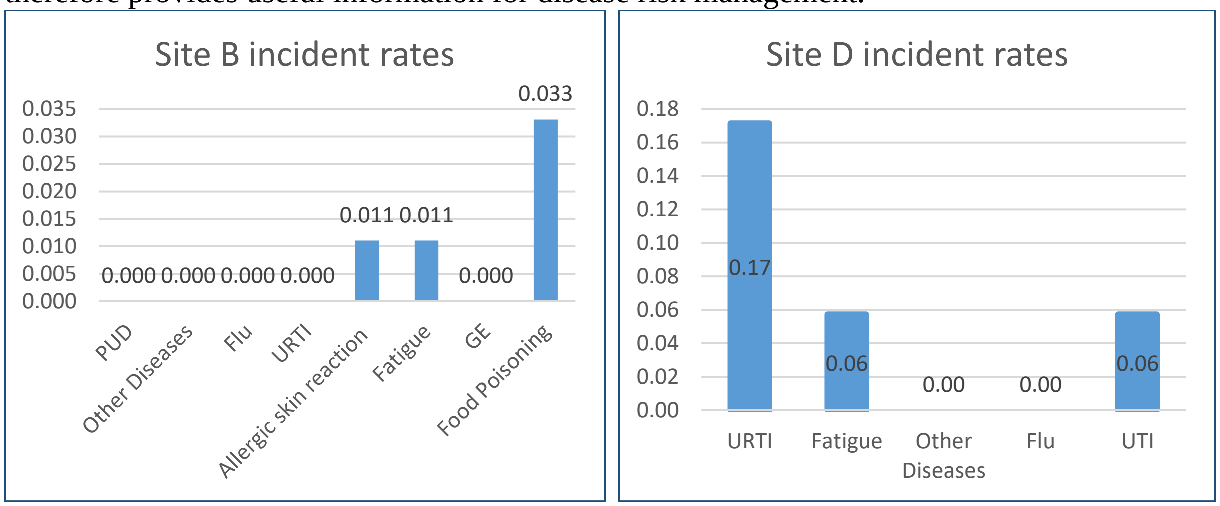


Fig. 10. Incidence rates of diseases

The results show that Mathematical models based on historic data, can quantitatively assess the transmissibility and potential health effects of plague outbreaks under a range of assumptions as discussed by Gani, & Leach (2004). Additionally, outbreak forecasting requires an integrative approach to modelling because outbreaks vary by disease, geographic location and socio-cultural practices. Therefore, dynamic modelling approaches should be adopted for prediction as pointed out by Scarpino, & Petri (2019).

5. Conclusion

The research set out to design a Citizen Observatory approach that can be used for developing an early warning system for disease outbreaks. The study designed an observatory where citizens were able to send in their observations. This study only took place for four months. It is envisaged that if a longitudinal study was conducted then it would be possible to measure the success of the system and compare it with other contemporary systems.

Mathematical modelling algorithms present great opportunities for predicting outbreaks using both supervised and unsupervised approaches. Our future work will focus on exploring unsupervised approaches such as Expectation-Maximization (EM) clustering which provides better optimization than distance-based or hard membership algorithms, such as k-Means. EM easily accommodates categorical and continuous data fields making it the most effective technique available for proper probabilistic clustering. Experiments with EM clustering however show some gaps that are typical of natural data. To address these alternative algorithms that improve the machine learning process such as supervised learning approaches will also be examined used. These include neural network classification algorithms. Findings from the study show that the Early Warning System can be used for disease outbreaks by having citizens themselves report their health status. This is the basis of Citizen Observatory portals, which encourage citizens to participate by reporting changes in their environment and reporting to health officers. When these changes reach predefined thresholds that would make them a menace, timely interventions would be taken.

The prototype designed used only the English language. The researchers however noted the need for localizing the application to incorporate local languages; for example, Kiswahili is widely used in East and Central Africa. Further, the application was tested at only four clinics using patients visiting the clinics. Tests should be scaled up to other locations for more conclusive results. Further, including feature phones that utilize USSD technology would increase the number of responses.

References

Angloher, G., Bento, A., Bucci, C., Canonica, L., Erb, A., von Feilitzsch, F & Huff, P. (2014). Results on low mass WIMPs using an upgraded CRESST-II detector. *The European Physical Journal C*, *74*(12), 3184.

Brazzola, N., & Helander, S. E. (2018). Five approaches to build functional early warning systems. *United Nations Development Programme*.

Dutta, R., Basnayake, S., & Ahmed, A. K. (2015). Assessing gaps and strengthening early warning system to manage disasters in Cambodia. IDRiM Journal, 5(2), 167-175.

Eysenbach, G., & Street, E. (2003). SARS and Population Health Technology Corresponding Author : *JOURNAL OF MEDICAL INTERNET RESEARCH*, *5*(2), 1–6. http://doi.org/10.2196/jmir.5.2.e14

Fraser, C., Riley, S., Anderson, R. M., & Ferguson, N. M. (2004). Factors that make an infectious disease outbreak controllable. Proceedings of the National Academy of Sciences, 101(16), 6146-6151.

Gani, R., & Leach, S. (2004). Epidemiologic determinants for modeling pneumonic plague outbreaks. *Emerging infectious diseases*, 10(4), 608.

Grossberndt, S., & Liu, H. Y. (2016). Citizen participation approaches in environmental health. In *Environmental Determinants of Human Health* (pp. 225-248). Springer, Cham.

Henriksen, H. J., Roberts, M. J., van der Keur, P., Harjanne, A., Egilson, D., & Alfonso, L. (2018). Participatory early warning and monitoring systems: A Nordic framework for web-based flood risk management. International journal of disaster risk reduction, 31, 1295-1306.

Hu, S. N., Cheng, X., & Chen, D. (2021). Comparative study on early warning methods of infectious diseases. In E3S Web of Conferences (Vol. 251). EDP Sciences.

Hussain-Alkhateeb, L., Rivera Ramírez, T., Kroeger, A., Gozzer, E., & Runge-Ranzinger, S. (2021). Early warning systems (EWSs) for chikungunya, dengue, malaria, yellow fever, and Zika outbreaks: What is the evidence? A scoping review. PLoS neglected tropical diseases, 15(9), e0009686.

Keeling, M. J., & Rohani, P. (2011). Introduction to simple epidemic models. *In modeling infectious diseases in humans and animals* (pp. 15-53). Princeton University Press.

Kurtah, P., Takun, Y., & Nagowah, L. (2019, July). Disease propagation prediction using machine learning for crowdsourcing mobile applications. In 2019 7th International Conference on Information and Communication Technology (ICoICT) (pp. 1-6). IEEE.

Lanfranchi, V., Wrigley, S. N., Ireson, N., Ciravegna, F., & Wehn, U. (2014). Citizens ' Observatories for Situation Awareness in Flooding. In Hiltz, S.R., M. S. Pfaff, L. Plotnick, & P. C. Shih (Eds.), *11th International ISCRAM Conference* (pp. 145–154). University Park.

Liu, H., Kobernus, M., Broday, D., & Bartonova, A. (2014). A conceptual approach to a citizens ' observatory - Supporting community-based environmental governance. *Environmental Health*.

Liu, Y., Hu, J., Snell-Feikema, I., VanBemmel, M. S., Lamsal, A., & Wimberly, M. C. (2015). Software to facilitate remote sensing data access for disease early warning systems. Environmental Modelling & Software, 74, 247-257.

Luther, J., Hainsworth, A., Tang, X., Harding, J., Torres, J., & Fanchiotti, M. (2017, May). World Meteorological Organization (WMO)—concerted international efforts for advancing multi-hazard early warning systems. In Workshop on World Landslide Forum (pp. 129-141). Springer, Cham.

Martcheva, M. (2015). Introduction to epidemic modeling. *In An introduction to mathematical epidemiology* (pp. 9-31). Springer, Boston, MA.

Muruti, G., Rahim, F. A., & bin Ibrahim, Z. A. (2018, November). A survey on anomalies detection techniques and measurement methods. In 2018 IEEE Conference on Application, *Information and Network Security (AINS)* (pp. 81-86). IEEE.

Noordzij, M., Dekker, F. W., Zoccali, C., & Jager, K. J. (2010). Measures of disease frequency: prevalence and incidence. *Nephron Clinical Practice,* 115(1), c17-c20.

Noormal, B. (2011). *Disease Early Warning System (DEWS) Surveillance , Early Detection and Response to Communicable Diseases ANNUAL REPORT 2011*.

Palacin-Silva, M., Seffah, A., Heikkinen, K., Porras, J., Pyhälahti, T., Sucksdorff, Y & Junttila, S. (2016). State-of-the Art Study in Citizen Observatories: Technological Trends, Development Challenges and Research Avenues.

POSTNOTE 2011, Cyber Security in the UK, Houses of Parliament, Parliament Office of Science and Technology, Number 389 September 2011United Nations. (2001). *A /56/326* (Vol. 52607).

Racloz, V., Ramsey, R., Tong, S., & Hu, W. (2012). Surveillance of dengue fever virus: a review of epidemiological models and early warning systems. PLoS neglected tropical diseases, 6(5), e1648.

Saunders, M., Lewis, P., & Thornhill, A. (2009). Research methods for business students. Pearson education.

Scarpino, S. V., & Petri, G. (2019). On the predictability of infectious disease outbreaks. *Nature communications*, 10(1), 1-8.

Semenza, J. C. (2015). Prototype early warning systems for vector-borne diseases in Europe. International journal of environmental research and public health, 12(6), 6333-6351.

Tapia, A. H., LaLone, N. J., MacDonald, E., Priedhorsky, R., & Hall, M. (2014, May). Crowdsourcing rare events: Using curiosity to draw participants into science and early warning systems. In ISCRAM.

WHO, (2004). *Weekly Epidemiological Record.* Geneva, World Health Organization. https://www.who.int/publications/journals/weekly-epidemiological-record, Accessed 20th March 2018

WHO, (2009). Outbreak Investigation and Response, https://www.who.int/diseasecontrol_emergencies/publications/idhe_2009_london_outbreaks.pdf, Accessed 10th November 2021

WHO, (2011), Early warning systems, http://www.who.int/csr/labepidemiology/projects/earlywarnsystem/en/, Accessed 20th March 2018

Zaman, J., D'Hondt, E., Boix, E. G., Philips, E., Kambona, K., & De Meuter, W. (2014, March). Citizen-Friendly participatory campaign support. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2014 IEEE International Conference on* (pp. 232-235). IEEE.

Zschau, J., & Küppers, A. N. (Eds.). (2013). Early warning systems for natural disaster reduction. Springer Science & Business Media.

Zommers, Z., & Singh, A. (Eds.). (2014). Reducing disaster: early warning systems for climate change. Springer Science+ Business Media.

Zurovac, D., Talisuna, A. O., & Snow, R. W. (2012). Mobile phone text messaging: tool for malaria control in Africa. PLoS medicine, 9(2), e1001176.