International Journal of Scientific & Engineering Research Volume 9, Issue 7, July-2018 1417
ISSN 2229-5518

Real-Time Intelligent Framework for Detecting Malaria Outbreak

Babagana Modu, M. A. Bilkisu

Abstract— Occurrence of malaria out break in developing countries have pose social and economic problem to the populace. Despite effort for preventive measures, by the World Health Organisation (WHO) and other health organisation agencies, there are a lot of uncertainty regarding where and when the outbreak of malaria will strike remains a key challenges. Several works were presented applying wide range of techniques towards reducing the incidence of malaria out break. The out break of malaria parasite is still on the high rate due to inadequate mechanism for tracking, detecting and preventing the outbreak in advance. This research seeks to propose a framework using real-time incidence of malaria out break to monitor the pattern of incidences of the out break. This proposed framework using real-time statistical control charts such as cumulative sum (CUSUM) and exoponential weighted moving average (EWMA) will be of great importance to hospitals, public health officials and policy makers with prior information to better prepare in anticipation of malaria outbreak for better management and planning.

Index Terms— Malaria, Outbreak, Detection, Hospitals Protocols, CUSUM, EWMA, Prevention

1 Introduction

THE occurrence of malaria outbreaks in developing countries have pose social and economic problem to the populace. Besides this, the uncertainty surrounding when and where malaria outbreak will occur is believed to have been the major challenges facing preventive measures. Despite concerted efforts by World Health Organization (WHO) and other donor agencies in providing interventions and strategizing policy towards malaria prevention and control. Yet, the malaria infection is still persisting, despite the interventions, policies and researches. Globally, it exposed about 3.2 billion people into risk of being infected with estimated deaths of nearly a million every year [1]. But, the mechanism for detecting likelihood of the malaria outbreak prior to the incidence remained key challenges. However, the availability of electronics healthcare data would enhanced development of a system for monitoring and detecting outbreak [2].

The prevention and control of diseases faces serious challenges particularly in the realm of detection of new cases and sometimes identification of aberration. For example, when an information reached health workers about the emergence of a new epidemic in an area. Surveillance and emergency team will be deployed for response. In applying this approach, the situation will sometimes go out of hand before it could be brought back to normalcy. To alleviate the delay between outbreak notification and intervention, it is pertinent to deploy a system that will keep-track of real-time health information on various diseases and conditions reported and recorded in the Electronic Database of hospitals. This will enable online processing of information and visualizing the pattern of disease episode as it is happening at any point in time for better monitoring aberration. In the past, several techniques have been utilized such as mathematical modelling [3-5], computational modelling [6] [7] and data science approach [8-10] towards malaria control. They suggested some findings for using these approaches such as identification of hidden ecological factor, large scale driver, an intelligent system for predicting malaria outbreak among other. However, malaria transmission is still persisting, and in this study, we aim to diversify the approaches. Thus, technology brings innovation and development, and also is a panacea to all sort problems. The scourge of malaria not only affects human life alone but also adversely imposes greater burden to the economy of a nation. Due to correlation of malaria incidence with rainy season [11] [12], work absconds by civil servants, school absenteeism by students and inability of farming activities by farmers as a results of sickness all these affect phase of economic development. Moreover, malaria infection have afflicted huge burden on education and productivity of developing countries as a result of those at risk [13].

This paper was proposed to address two key issues: Firstly, to design a framework for detecting malaria outbreak based on reported cases of malaria incidence. Secondly, to design the application platform coded with embedded control charts algorithm to visualize real-time pattern of the malaria incidence. In this study, we aim to use hospital records of reported cases of malaria to design a framework for detecting outbreak. For example, in the hospital settings, people are being diagnosed and were subjected for a test. Then, treatment will be given based on the outcome of the examination. Consequently, a number of cases on several disease conditions will be obtained and kept for retrospective purposes. The rate at which cases are being reporting to the hospital is empirically expected to give a pattern showing up or down. Fundamentally, the use of appropriate statistical methods such as Descriptive Statistics [14], Statistical Control Charts [15] among other are suitable for visualizing and identifying aberrations. However, deploying these techniques

- *Babagana Modu is currently pursuing PhD in Data Science at University of Bradford, United Kindgom. E-mail: b.modu@bradford.ac.uk*
- *M. A. Bilkisu is currently pursuing PhD in Survival Analysis at UMM, Malaysia. E-mail: bilkisuahmad971@gmail.com*

to retrospective cases for studying disease outbreak or other ramification are prone to limitation. Due to static nature of the analysis, it is less effective in providing early warning signal or prior information. Warning signal would assist hospitals, public health officials and policy makers with tip-off for better preparedness. But, the real-time monitoring of disease cases is the most efficient and effective paradigm for malaria outbreak prevention.

The subsequent sections of this work are as follows. In Section II, present the theoretical framework comprises of malaria transmission, related works using statistical control chat for disease detection such CUSUM (Cumulative Sum) and EWMA (Exponential Weighted Moving Average) and other charts. In Section III, we describe the outbreak detection framework using the protocols in the hospital settings. The model formulation describing the control chart for monitoring and detecting outbreak are vividly illustrated and their limitations presented in Section IV. In Section V, we wrap-up the work by summarizing our findings and indicate the future direction.

2 Theoretical Framework

2.1 Review Stage Malaria Transmission

Malaria is a vector-borne disease caused by genus species called plasmodium, and it spread into human population by female mosquitoes (Anopheles species) [16].There are over 100 hundreds of malaria parasites exists, but only four are known to infect human these includes: plasmodium falciparum; plasmodium malarae; plasmodium ovale and plasmodium vivax [17]. The most virulent species is plasmodium falciparum [17], which predominantly widespread in the World with high degree of infection burden on human population over all other species. Another species of plasmodium was later discovered in South-Eastern Asian called plasmodium knowlesi [18], this one has zoonotic characteristics in nature.

2.2 Related Work

A lot of studies are carrying out day-by-day in developing a system for detecting aberration in diseases incidence pattern for better planning and prevention of outbreak ahead. Recent study by [19] uses EWMA for detecting outbreak of Salmonella based on weekly reported cases in German Federal states. A spatio-temporal EWMA statistic were used and identified the disease clusters in multiple regions pooling it together for increasing the speed of detection compared to the independent approach. Three different control charts were compared to monitor and study their detection performance in nosocomial infection outbreak and the results showed that the traditional shewhart chart is most appropriate for monitoring periods with large deviations [20]. However, in the work of [21], which analyzes influenza cases using EWMA control chart to detect the start and end of influenza outbreaks in relation to its seasonality dependence usually winter and early spring.

3 Outbreak Detection Framework

The system architecture and protocols depicted in Figure 1, portrays the bureaucracy for reporting and recording diseases cases in hospital settings through electronic database. This is being done for effective data management and real-time monitoring of aberration to detects and avert outbreak ahead of time. The architecture presented in Figure 1, comprises of two mutually independent units connected according to their operational function, these are: Health centers and Central data processing hub.

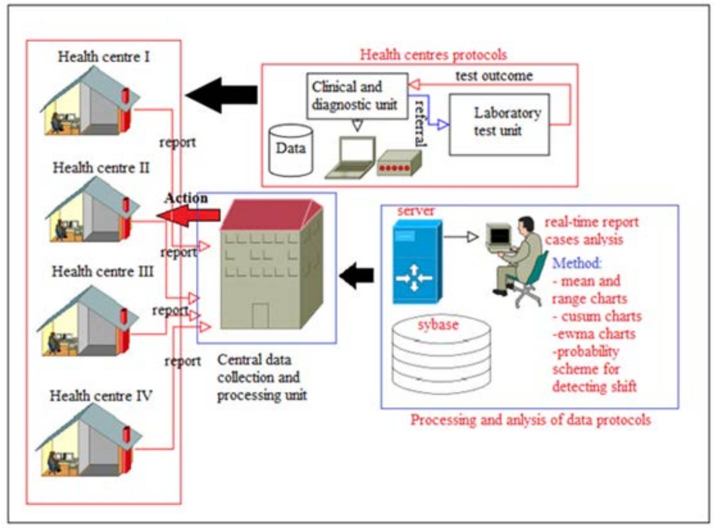

Figure1: Shows the architecture of the protocols followed in hospital settings

3.1 Health Centers

The health centers are basically the designated hospitals clustered around a certain location of: city, province, state or country at large. In this study, we are considering four health centers shown in Fig.1 (I, II, III, IV) to illustrates how the health information are being flow from one unit to another within hospital settings. Health centers are mainly the primary source of gathering data on diseases and other conditions related to health. Shown in Figure 1, that for each of the health center there are units such as: clinical and diagnostic, laboratory and data processing unit. The topology provide a layout of gaining information systematically without skipping a step until discharge. This is a regular practice in the hospitals, as patient arrive hospital will be issued a card to see a general practitioner (GP) for diagnosis and after that if there is need for test then will be refereed to laboratory. Results obtained in the laboratory after the test would be taken back to the GP for further action and at the same the outcome of the test would also forwards to data processing units for documentation. For each of the health centers according to the topology in Figure 1 is responsible for collecting individual health information on every subjects into their computer systems on real-time scenario for further action. The information stored in the data processing units of individual's health centers would be pass unto hub (the central data processing unit).

3.2 Data Processing Hub

This unit is responsible for receiving and visualizing information that transmitted electronically from the health centers through online platforms. For instance, a person is diagnosed and was subjected for a test in the laboratory. If the outcome shows malaria is positive, the result will then supplied into online platform and the information will be transmists to the progressing hub. Doing this, would enable to track, detect and prevent outbreak based on the pattern of the reported malaria cases in hospital. In the hospitals, medical records is the unit solely responsible for keeping up–to–date health information. Due to laxity of medical record personnel, health information at the time of enquiry are often not available. This is perhaps attributed to lack of qualified personnel with prerequisite skills and commitments toward work. Collecting the consortium of health information and kept it in a booklet is a traditional approach which cannot be accommodated again in this modern age of computer system and internet compliance. The methods of keeping records in a booklets and files suffers a great challenges especially in an events of fire, flood, thief and other natural and artificial calamity that might cause loss of records. However, even though the computer based approach cannot be hundred percent reliable but in comparison to the modern way would makes it satisfactorily more precise.

4 MODEL FORMULATION

This section describes the statistical charts like cumulative sum (CUSUM) and exponential weighted moving average (EWMA) for better understanding of the concept.

4.1 Cummulative Sum (CUSUM)

The Cumulative Sum (CUSUM) chart is an efficient charting approach compare to the traditional Shewhart due to its ability to detect small shift in the mean of a process. In this study, we consider the sensitivity of CUSUM chart and proposed using it to monitor the pattern of real-time cases of malaria incidence reported. Real-time monitoring will enable to detect when the cases are going up, this would help in tracking and averting outbreak in advance. In relation to computation of Average Run Lengths (ARL's), the CUSUM chart outperform Shewhart control charts when it is desired to detect shifts in the mean that are 2 sigma or less.

In mathematical sense, the CUSUM chart is formed by plotting: $S'_m = 1/\sigma_{\bar{x}} \sum_{i=1}^{n} (\bar{x}_i - \widehat{\mu_o})$ where m is the samples, and n is size of each samples, $\widehat{\mu_o}$ is the estimate of the in-control mean and $\sigma_{\bar{x}}$ is the standard deviation of the sample means. However, the choice of $\widehat{\mu_o}$ and $\sigma_{\bar{x}}$ for plotting CUSUM is usually determined by the statistical package used for the analysis. For the process to remain in control, it should be centered at $\widehat{\mu_o}$, the CUSUM plot will show a pattern showing the variation in a random centered about zero. If however, the process mean shifts upward, the charted CUSUM points will eventually drift upwards, and vice versa if the process mean decreases.

4.2 Exponential Weighted Moving Average (EWMA)

Let us consider $y_1, y_2, y_3, \cdots$ to represents the time series of the recorded cases of malaria incidence reported to hospital. We also assume that $y_1, y_2, y_3, \cdots$ are identically independent distributed (iid) Poisson random variable with mean μ . When the pattern of the reported cases is within the tolerable bound if $\mu = \mu_o$ or otherwise. To monitor the pattern of cases of malaria incidence, we propose the use of an Exponential Weighted Moving Average (EWMA) which was first introduced by [22].

This is defined mathematically as $Z_t = \mu_o$, and $Z_t = \gamma y_t + (1-\gamma) Z_{(t-1)}$ by taking the average of Z_t then $E(Z_t) = \mu$, for pattern of malaria cases to be in control when $E(Z_t) = \mu = \mu_o$. The variance $V(Z_t) = \gamma/(\gamma - 1)\,[1-(1-\gamma)^{2t}]\mu_o$, the exact variance when t is large, then variance is approximately $V(Z_t) \approx \gamma\mu_o/(2-\gamma) = V(Z_\infty)$. The EWMA control chart for Poisson random variable with mean μ, to detect outbreak when out-of-control signal such as $Z_t < \delta_{ll}$ or $Z_t > \delta_{ul}$, where $\delta_{ll} = \mu_o + A_{ll}(V(Z_t))^{1/2}$ and $\delta_{ul} = \mu_o + A_{ul}(V(Z_t))^{1/2}$. Therefore, $A = A_{ll} = A_{ul}$ are often chosen to be equal. Although it is sometimes may be advantage to construct asymmetric control charts.

6 CONCLUSION

Most often in developing countries, a large number of casualties are being recorded from outbreaks. This is due to inadequate of malaria outbreak detection systems. A lot of lives lost and some were severely afflicted with illness. The intervention in terms of emergency rescue, quarantine and provision of health facilities are mostly comes after outbreak stroke. However, the level of damage caused by this delay cannot be easily measured because of deadly nature of malaria. Globally, malaria alone accounts for about half a million deaths annually according to World Health Organization, 2015. Despite, continues efforts for mitigating malaria scourge, still it remained the challenges today. In the previous works presented in this paper, retrospective analysis of malaria cases have been used and gave suggestions on the preventive measures. Using this approach provide no information about early warning signal for malaria outbreak, but rather identify the likelihood outbreak. This paper was proposed to address to key issues: Firstly, to design a framework for detecting malaria outbreak based on reported cases of malaria incidence. Secondly, to design the application platform coded with embedded control charts algorithm to visualize real-time pattern of the malaria incidence. However, we are considering the former in this paper, while the latter is reserved for the future work. In this study, we used protocols in hospital settings and designed a framework for transmitting malaria cases electronically from respective health centers to data processing hub. The data transmitted to the hub, will then instantaneously analyze and visualize a pattern showing real-time scenario indicating the behavior of malaria occurring. Doing so, will enable to track, detect and efficiently prepared to avert outbreak. This framework will be facilitated by the use of some statistical control charts (CUSUM and EWMA) algorithm which would be embedded in the platform. These control chats were chosen because of their potential ability and sensitivity to a small shift unlike the Shewhart charts. This framework, also have widerange of operational coverage that cut-across other diseases. The implementation of this framework will curtail malaria outbreak and also assist hospitals, public health workers and policy makers for better management and planning

DECLARATION

The authors have unanimously agreed to publish in IJSER with no comflict of interest.

REFERENCES

[1] World Health Organization. Malaria rapid diagnostic test performance: results of WHO product testing of malaria RDTs: round 6 (2014-2015). (2015).

[2] Buendia, R. J. M., & Solano, G. A. (2015, July). A disease outbreak detection system using autoregressive moving average in time series analysis. In *Information, Intelligence, Systems and Applications (IISA), 2015 6th International Conference on* (pp. 1-5). IEEE.

[3] Okuneye, K., & Gumel, A. B. (2017). Analysis of a temperature-and rainfall-dependent model for malaria transmission dynamics. *Mathematical biosciences*, *287*, 72-92.

[4] Nah, K., Nakata, Y., & Röst, G. (2014). Malaria dynamics with long incubation period in hosts. *Computers & Mathematics with Applications*, *68*(9), 915-930.

[5] Beck-Johnson, L. M., Nelson, W. A., Paaijmans, K. P., Read, A. F., Thomas, M. B., & Bjørnstad, O. N. (2013). The effect of temperature on Anopheles mosquito population dynamics and the potential for malaria transmission. *PLOS one*, *8*(11), e79276.

[6] Zhou, Y., Arifin, S. M., Gentile, J., Kurtz, S. J., Davis, G. J., & Wendelberger, B. A. (2010, July). An agent-based model of the Anopheles gambiae mosquito life cycle. In *Proceedings of the 2010 Summer Computer Simulation Conference* (pp. 201-208). Society for Computer Simulation International.

[7] Illangakoon, C., McLeod, R. D., & Friesen, M. R. (2014, June). Agent based modeling of malaria. In *Humanitarian Technology Conference-(IHTC), 2014 IEEE Canada International* (pp. 1-6). IEEE.

[8] Modu, B., Asyhari, A. T., & Peng, Y. (2016, December). Data Analytics of climatic factor influence on the impact of malaria incidence. In *Computational Intelligence (SSCI), 2016 IEEE Symposium Series on* (pp. 1-8). IEEE.

[9] Modu, B., Asyhari, A. T., Konur, S., & Peng, Y. (2017, June). An Assessment on the Hidden Ecological Factors of the Incidence of Malaria. In *Multidisciplinary Digital Publishing Institute Proceedings* (Vol. 1, No. 3, p. 131).

[10] Modu, B., Polovina, N., Lan, Y., Konur, S., Asyhari, A. T., & Peng, Y. (2017). Towards a Predictive Analytics-Based Intelligent Malaria Outbreak Warning System. *Applied Sciences*, *7*(8), 836.

[11] MacLeod, D. A., and Morse, A. P. (2014). Visualizing the uncertainty in the relationship between seasonal average climate and malaria risk. Scientific reports, 4, 7264.

[12] Krefis, A. C., Schwarz, N. G., Krger, A., Fobil, J., Nkrumah, B., Acquah, S. and May, J. (2011). Modelling the relationship between precipitation and malaria incidence in children from a holoendemic area in Ghana. The American journal of tropical medicine and hygiene, 84(2), 285-291.

[13] D. J. Gubler, Resurgent vector-borne diseases as a global health problem, Emerging Infectious Diseases 4 (3) (1998), p. 442-45

[14] MacDonald, P. D., Cressman, G., McKay, M., Loo, S., McClure, E., & MacGuire, E. (2017). Real-time surveillance and response system for Ebola and other emerging infections. *Online journal of public health informatics*, *9*(1).

[15] Neuburger, J., Walker, K., Sherlaw-Johnson, C., van der Meulen, J., & Cromwell, D. A. (2017). Comparison of control charts for monitoring clinical performance using binary data. *BMJ Qual Saf*, bmjqs-2016.

[16] K. Mendis, B. J. Sina, P. Marchensini, and R. Carter, The neglected burden of Plasmodium vivax malaria, A. J. Trop. Med. Hyg. 64 (1, 2) S (2001), p. 97-106.

[17] Martcheva, M., and Prosper, O. (2013). Unstable dynamics of vector borne diseases: Modeling through delay-differential equations. In Dynamic models of infectious diseases (pp. 43-75). Springer New York.

[18] Singh, B., and Daneshvar, C. (2013). Human infections and detection of Plasmodium knowlesi. Clinical microbiology reviews, 26(2), 165-184.

[19] Salmon, M., Schumacher, D., Burmann, H., Frank, C., Claus, H., & Höhle, M. (2016). A system for automated outbreak detection of communicable diseases in Germany. *Eurosurveillance*, *21*(13), 30180.

[20] Dassanayaka, S. K., & French, J. (2016). An Improved EWMA-Based Method for Outbreak Detection in Multiple Regions. *Online Journal of Public Health Informatics*, *8*(1).

[21] Steiner, S. H., Grant, K., Coory, M., and Kelly, H. A. (2010). Detecting the start of an influenza outbreak using exponentially weighted moving average charts. BMC medical informatics and decision making, 10(1), 37.

[22] Roberts, S. W. (1966). A comparison of some control chart procedures. *Technometrics*, *8*(3), 411-430.