Journal of Biostatistics and Epidemiology

*J Biostat Epidemiol. 2019;5(3): 204-209*

**Original Article**

# Evaluation the Performance of Exponentially Weighted Moving Average in the Detection of Cholera Outbreaks: Using the Reported Cholera Outbreaks in Literature

Yousef Alimohamadi [1,2], Firooz Esmaeilzadeh[3], Parisa Jalali [2], Mohsen Mohammadi [2] Mojtaba Sepandi [4,5*]*

[1]Pars Advanced and Minimally Invasive Medical Manners Research Center, Pars Hospital, Iran University of Medical Sciences, Tehran, Iran.
[2] Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran.
[3] School of Nursing and Midwifery, Maragheh University of Medical Sciences, Maragheh, Islamic Republic of Iran.
[4] Health Research Center, Life Style Institute, Baqiyatallah University of Medical Sciences, Tehran, Iran.
[5] Department of Epidemiology and Biostatistics, Faculty of Health, Baqiyatallah University of Medical Sciences, Tehran, Iran.

ARTICLE INFO

Received 07.02.2019
Revised 24.06.2019
Accepted 08.07.2019
Published 01.08.2019

**Key words:**
EWMA;
Sensitivity;
Specificity;
Outbreak

ABSTRACT

**Introduction:** Timely Detection of outbreaks of infectious diseases can have a very important role in surveillance systems. the presence of appropriate methods can have a very important role for this purpose, the aim of the current study was to Evaluation The Performance of Exponentially Weighted Moving Average in the detection of cholera outbreaks using the reported cholera outbreaks in literature

**Methods:** In the current study the EWMA method was evaluated. To assess the performance of the mentioned methods the six real outbreaks algorithm reported in the literature were used. These reported outbreaks were the daily counts of cholera cases in different countries. After insertion of each outbreak, 7 days inserted as non-outbreaks days. All analyses performed by MedCalc18.11, Stata version15 and excel 2010.

**Results:** the sensitivity of EWMA was 56.4% (95% CI: 54.3%- 58.5%). The highest sensitivity for outbreak detection was seen in EWMA1 79.18(73.56-84.09) and the lowest was seen in EWMA4 12.2(8.4-17.0). EWMA2 with $\lambda$= 0.2 had the best performance with sensitivity 69.8 (63.6-75.5) and specificity 91.4(76.9-98.2) and AUC= 0.80.

**Conclusion:** The EWMA method can be very useful in the detection of outbreaks, but the use of this method along the other models may increase the sensitivity of outbreaks detection.

## Introduction

Even with the development of preventative measures, cholera as a life-threatening water-borne infectious disease which can be characterized by diarrhea, accompanied by numerous voluminous watery stools and vomiting, remains a public health burden in developing countries. Globally, an estimated 1.3 billion people are at risk of cholera (1). The causative agent of cholera is Vibrio cholera O1, a Gram-negative pathogen(2). An outbreak is defined as more cases of a disease than expected in a specific location over a specific period. Suspicion often arises when health care workers report an unusual cluster or a single, unexpected presentation. An increasing number of methods are being developed to detect outbreaks of infectious diseases using routinely collected data(3). One of the most known methods and

* Corresponding Author: msepandi @gmail.com

Please cite this article in press as Alimohamadi Y, Esmaeilzadeh F, Jalali P, Mohammadi M, Sepandi M. Evaluation the Performance of Exponentially Weighted Moving Average in the Detection of Cholera Outbreaks: Using the Reported Cholera Outbreaks in Literature. J Biostat Epidemiol. *2019; 5(3): 204-209*

algorithms used by the surveillance systems to detect outbreaks is Exponentially Weighted Moving Average (EWMA). This algorithm is a group-based method of statistical process control and efficiency in detecting small changes(4, 5). Early response to health events, especially public health emergencies with international concern is a major public health priority. Outbreak detection methods and algorithms as the main tools for public health surveillance systems are under the umbrella of temporal and spatial methods(6). Timely detection of outbreaks and bioterrorism is necessary by public health surveillance systems. Moreover, few published studies addressed the performance of statistical methods in cholera surveillance. This study aimed to address the performance of EWMA in the detection of outbreaks detection by using the reported cholera outbreaks in literature.

## Methods

### Outbreaks data:

To assess the understudy method, six real outbreaks algorithm reported in the literature were used (7-12). These reported outbreaks were the daily counts of cholera cases in different countries. After insertion of each outbreak, 7 days inserted as non-outbreaks days. Overall the used time-series data set includes 280 days (245 outbreak days and 35 non-outbreak days). The EWMA was applied to daily reported counts of cholera data to detect inserted outbreaks.

Outbreak detection method EWMA statistics are defined by the following recursive equation(13):

$$EWMA_t=Yt+(1-\lambda)EWMA_{t-1}. \qquad (1)$$

Where Yt equals the number of suspected cases of cholera in day t, $\lambda$ is the weighting parameter that has been considered as 0.1 for EWMA1, 0.2 for EWMA2 and so on (Table 1). The upper control limit for outbreak detection is as follow:

$$\text{Upper Control Limit}=EWMA_0+ k\times \sigma_{EWMA}$$

Where k is a constant parameter, $\sigma_{EWMA}$ and $EWMA_0$ are the standard deviation ($\sigma$) and the mean ($\mu$) of data in the absence of the outbreak. In the current study, the amount of K determined 2(K=2) and the $\mu+2\sigma$ considered as an upper limit for outbreak detection

Measures of the algorithm's performance the performance of EWMA algorithms in the detection of cholera outbreaks was measured using sensitivity, specificity, false alarm rate, likelihood ratios and area under the receiver operating characteristics (ROC) curve (AUC) and accuracy. The total number of outbreak-days was considered as the gold standard to calculate appropriate measures to evaluate the performance of algorithms. Accordingly, the denominator for sensitivity and specificity formulas was 245 outbreak days and 35 non-outbreak days, respectively. AUC with 95% confidence intervals (95% CI) was used to compare different algorithms and greater values indicate better performance. Briefly, greater values of AUC indicate better performance of a specific EWMA algorithm in comparison to other algorithms. AUC values have been reported by percentage throughout the text and displayed in Figure. All analyses performed by MedCalc18.11, Stata version15 and excel 2010.

## Results

Overall Sensitivity of the EWMA for all occurred outbreaks was 56.4% (95% CI: 54.3%- 58.5%). the false and negative alarm rate for EWMA was 8 %( 95% CI: 5%- 11%), 43(40-45) respectively. Among the different algorithms, EWMA1 with $\lambda$= 0.1 had the highest sensitivity with 79.18(73.56-84.09) in detecting inserted outbreaks and EWMA4 with $\lambda$= 0.4 had the lowest sensitivity= 12.2 (8.4-17.0) in detecting inserted outbreaks. Also EWMA2 with $\lambda$= 0.2 had the best performance with sensitivity 69.8 (63.6-75.5) and specificity 91.4(76.9-98.2) and AUC= 0.80. Tables 2 and 3 show disaggregated measures of EWMA performance by different values of parameter entitled EWMA1 to EWMA9 including sensitivity, specificity, false alarm rate, false-negative rate, positive and negative likelihood ratios, and Roc Area. In total, the AUC of the EWMA for all of the occurred outbreaks was 0.74. The same values according to the different parameters for EWMA1 to EWMA9 are shown in Fig. 1 and table3.

Table 1-Characteristics of the used algorithms in the study for outbreak detection.

| Algorithm no | λ |
|---|---|
| EWMA1 | 0.1 |
| EWMA2 | 0.2 |
| EWMA3 | 0.3 |
| EWMA4 | 0.4 |
| EWMA5 | 0.5 |
| EWMA6 | 0.6 |
| EWMA7 | 0.7 |
| EWMA8 | 0.8 |
| EWMA9 | 0.9 |

Table 2- Sensitivity, specificity, false alarm rate, false-negative rate, positive and negative likelihood ratios of the used EWMA algorithms.

| Algorithm | sensitivity | specificity | False Alarm rate | False negative rate | Positive likelihood ratio | Negative likelihood ratio | Accuracy |
|---|---|---|---|---|---|---|---|
| EWMA1 | 79.18(73.56-84.09) | 40(23.87-57.89) | 60(40-80) | 21(16-26) | 1.32 | 0.52 | 74.29(68.75-79.30) |
| EWMA2 | 69.8 (63.6-75.5) | 91.4(76.9-98.2) | 5(-2-9) | 30(24-36) | 8.14 | 0.3 | 72.5(66.8-77.6) |
| EWMA3 | 74.7(68.1-80.5) | 40.0(23.8-57.9) | 3(0-8) | 26(20-31) | 1.2 | 0.6 | 69.6(63.3-75.4) |
| EWMA4 | 12.2(8.4-17.0) | 100.0 (90.0 -100.0) | 0 | 88(84-92) | - | 0.9 | 23.2(18.4-28.6) |
| EWMA5 | 57.5(51.1-63.8) | 100.0 (90.0-100.0) | 0 | 42(36-49) | - | 0.4 | 62.8(56.9-68.5) |
| EWMA6 | 56.3(49.8-62.6) | 100.0 (90.0-100.0) | 0 | 44(37-50) | - | 0.4 | 61.8(55.8-67.5) |
| EWMA7 | 57.5(51.1-63.8) | 100.0 (90.0-100.0) | 0 | 42(36-49) | - | 0.4 | 62.8(56.9-68.5) |
| EWMA8 | 56.7(50.3-63.0 | 100.0 (90.0-100.0) | 0 | 43(37-49) | - | 0.4 | 62.1(56.1-67.8) |
| EWMA9 | 56.7(50.3-63.0 | 100.0 (90.0-100.0) | 0 | 43(37-49) | - | 0.4 | 62.1(56.1-67.8) |
| Overall EWMA | 56.4(54.3-58.5) | 92(88.5-94.8 | 8(5-11) | 43(40-45) | 7.11 | 0.4 | 60.8(58.9-62.8) |

Table3-the ROC curve area according to different parameter for EWMA1 to EWMA9

| Column1 | Total days | ROC Area | Std. Err. | [95% Conf. Interval] |
|---|---|---|---|---|
| EWMA1 | 280 | 0.59 | 0.04 | 0.50 -0.68 |
| EWMA 2 | 280 | 0.80 | 0.02 | 0.75 - 0.86 |
| EWMA 3 | 280 | 0.79 | 0.02 | 0.75 - 0.83 |
| EWMA 4 | 280 | 0.56 | 0.01 | 0.54 - 0.58 |
| EWMA 5 | 280 | 0.78 | 0.01 | 0.75 - 0.81 |
| EWMA 6 | 280 | 0.78 | 0.01 | 0.75 - 0.81 |
| EWMA 7 | 280 | 0.78 | 0.01 | 0.75 - 0.81 |
| EWMA 8 | 280 | 0.78 | 0.01 | 0.75 - 0.81 |
| EWMA 9 | 280 | 0.78 | 0.01 | 0.75 - 0.81 |

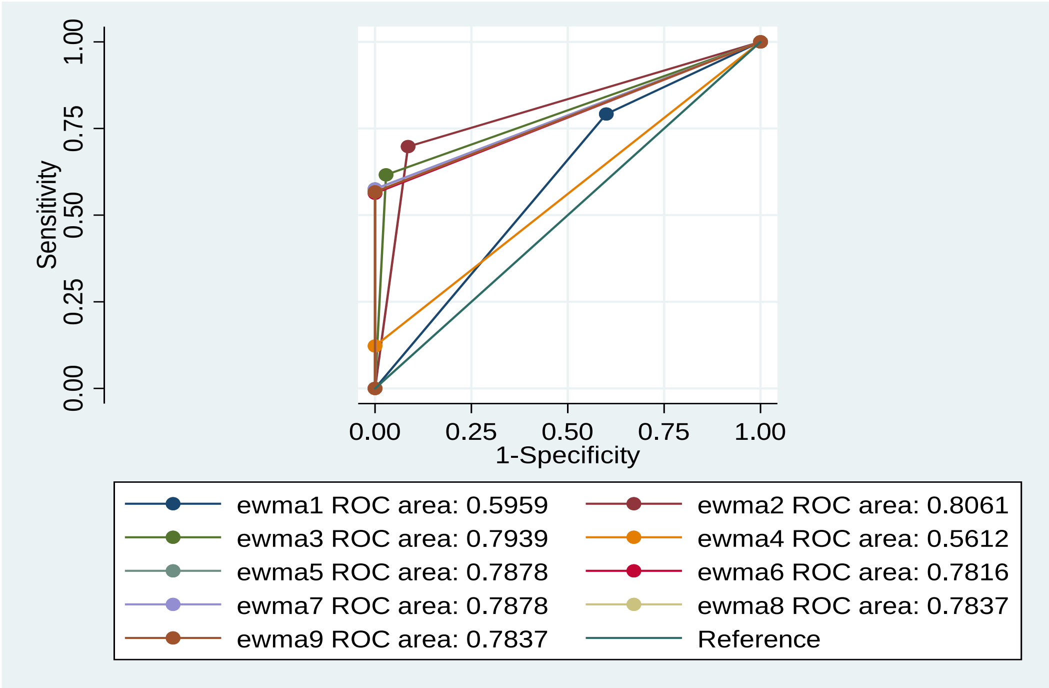


Fig 1. The area under the ROC curve for different used EWMA algorithms.

## Discussion

Studying the use of EWMA algorithms reflects the benefits of this method in the early detection of outbreaks. But since this method has its limitations, proper use and proper positioning can give the best returns to the early detection of outbreaks. Based on a true comparison between the results of this study and similar studies, the sensitivity of the EWMA model to the identification of all outbreaks was compared with the sensitivity of this model in detecting influenza-induced outbreaks, which, concerning values of 54% versus 70% sensitivity This model is less likely to detect cholera outbreaks than to detect flu outbreaks. This comparison compares the false alarms of this model with the best performance of the model in detecting influenza outbreaks, and this difference is significant with 8% vs. 2% (5) In our study, Parameter 0.2 had the best performance in detecting an outbreak, and in the same study, Parameter 0.9 was the best performance. This difference, which is also seen in other studies, indicates a lack of a single algorithm with an alarm and an appropriate feature to identify all outbreaks. (5) However, in identifying some outbreaks, including meningitis, there is no definition of a specific threshold level algorithm for use in all outbreaks of this disease that can cover different conditions and do not have a proper function in all circumstances, so it is sometimes necessary attention to the epidemiology of the disease in question, the importance of detecting the time and cost of examining false alarms from different algorithms, even in detecting outbreaks of a disease(14) The results of this study on the suitability of using this model for detecting cholera outbreaks are consistent with the results of similar studies that look at the performance of this model in identifying health outcomes. The study by Moscarelli et al., which explores the use of this model in identifying a pattern When death is due to cardiac surgery, it has been shown that using this model in monitoring reduces mortality following heart surgeries. (15)Other strengths of this approach can be easily understood by the clinical staff. This theme helps with quick diagnosis and prompt and bad interventions the goal is to reduce infections in infants and reduce the use of antibiotics in them (16) Also, due to the high sensitivity and high profile of this model and the help that early detection of infectious diseases in domestic animals, it can be much successful in reducing economic losses. (17) In general, the use of this model in the early diagnosis of many diseases suggests to its proper function in preventing and controlling the outbreak in time. (18)However, along with all the benefits mentioned, some studies have shown the limitations of this model by using It offers it in combination with other models. Regarding the weaknesses of this model, it is possible to reduce the ability to detect an outbreak with precision in control measures to reduce the false alarms of this model. On the other hand, if we want to pay more attention to the early detection of an outbreak by this model, the chance of false warnings and the subsequent costs increase (14) Also, the use of the EWMA model for detecting small outbreaks, such as the local outbreak of measles with a low incidence, is not a good method, and it is believed to detect these outbreaks 2 to 7 days (19) And in cases where the EWMA chart is used as an alert for increasing death or other health outcomes, the weaknesses of this model in such situations can be seen in the slow detection of adverse health outcomes (16). Therefore, the use of this method alone works It's not right, and it's best to use the combination methods that can best be used to detect outbreaks. (14, 20, 21)

## Conclusions

The results of this study showed that the EWMA method in the detection of an outbreak has a good function, but the use of this model along with other predictive models may increase the sensitivity and general characteristics of the combined use and increase the accuracy of the diagnosis that is necessary for Future studies will address this issue.

Ethics approval and consent to participate

This study was approved by the Ethics Committee of Research at Baqiyatallah University of Medical Sciences, Iran (IR.BMSU.REC.1398.071).

## Acknowledgments

Many thanks to all the authors who helped in the preparation of this article.

## References

1.Ali M, Nelson AR, Lopez AL, Sack DA. Updated global burden of cholera in endemic countries. PLoS neglected tropical diseases 2015;9(6):e0003832.

2.Islam MT, Khan AI, Sayeed MA, Amin J, Islam K, Alam N, et al. Field evaluation of a locally produced rapid diagnostic test for early detection of cholera in Bangladesh. PLoS neglected tropical diseases 2019;13(1):e0007124.

3.Watkins RE, Eagleson S, Hall RG, Dailey L, Plant AJ. Approaches to the evaluation of outbreak detection methods. BMC public health 2006;6(1):263.

4.Shu L, Su Y, Jiang W, Tsui K-L. A comparison of exponentially weighted moving average-based methods for monitoring increases in incidence rate with varying population size. IIE Transactions 2014;46(8):798-812.

5.Solgi M, Karami M, Poorolajal J. Timely detection of influenza outbreaks in Iran: Evaluating the performance of the exponentially weighted moving average. Journal of infection and public health 2018;11(3):389-392.

6.Karami M. Validity of evaluation approaches for outbreak detection methods in syndromic surveillance systems. Iranian journal of public health 2012;41(11):102-103.

7.Dan-Nwafor CC, Ogbonna U, Onyiah P, Gidado S, Adebobola B, Nguku P, et al. A cholera outbreak in a rural north central Nigerian community: an unmatched case-control study. BMC public health 2019;19(1):112.

8.Oguttu DW, Okullo A, Bwire G, Nsubuga P, Ario A. Cholera outbreak caused by drinking lake water contaminated with human faeces in Kaiso Village, Hoima District, Western Uganda, October 2015. Infectious diseases of poverty 2017;6(1):146.

9.Gidado S, Awosanya E, Haladu S, Ayanleke HB, Idris S, Mamuda I, et al. Cholera outbreak in a naïve rural community in Northern Nigeria: the importance of hand washing with soap, September 2010. The Pan African Medical Journal 2018;30.

10.Pande G, Kwesiga B, Bwire G, Kalyebi P, Riolexus A, Matovu JK, et al. Cholera outbreak caused by drinking contaminated water from a lakeshore water-collection site, Kasese District, south-western Uganda, June-July 2015. PloS one 2018;13(6):e0198431.

11.Moradi G, Rasouli MA, Mohammadi P, Elahi E, Barati H. A cholera outbreak in Alborz Province, Iran: a matched case-control study. Epidemiology and health 2016;38.

12.Karami M. Maosumi Asl H, Mohammadin M, Rae-ofi H, Saghafipour A, Noroozi M, Khedmati E. Qom cholera outbreak in 2011: influential and determinant factors. IJE 2012;8(3):84-92.

13.Lucas JM, Saccucci MS. Exponentially weighted moving average control schemes: properties and enhancements. Technometrics 1990;32(1):1-12.

14.Faryadres M, Karami M, Moghimbeigi A, Esmailnasab N, Pazhouhi K. Levels of alarm thresholds of meningitis outbreaks in Hamadan Province, west of Iran. J Res Health Sci 2015;15(1):62-5.

15.Moscarelli M, Athanasiou T, Sevdalis N, Vescovi F, Fattouch K, Nasso G, et al. Controlled Exponentially Weighted Moving Average Chart in Cardiac Surgery: A Simulation Study Across 9 Italian Cardiac Centers. Semin Thorac Cardiovasc Surg 2016;28(2):253-258.

16.Bowen JR, Callander I, Richards R, Lindrea KB. Decreasing infection in neonatal intensive care units through quality improvement. Arch Dis Child Fetal Neonatal Ed 2017;102(1):F51-F57.

17.Silva GS, Schwartz M, Morrison RB, Linhares DCL. Monitoring breeding herd production data to detect PRRSV outbreaks. Prev Vet Med 2017;148:89-93.

18.Steiner SH, Grant K, Coory M, Kelly HA. Detecting the start of an influenza outbreak using exponentially weighted moving average charts. BMC Med Inform Decis Mak 2010;10:37.

19.Karami M, Soori H, Mehrabi Y, Haghdoost AA, Gouya MM. Real time detection of a measles outbreak using the exponentially weighted moving average: does it work? J Res Health Sci 2012;12(1):25-30.

20.Texier G, Alldoji RS, Diop L, Meynard J-B, Pellegrin L, Chaudet H. Using decision fusion methods to improve outbreak detection in disease surveillance. BMC medical informatics and decision making 2019;19(1):38.

21.Dawson P, Gailis R, Meehan A. Detecting disease outbreaks using a combined Bayesian network and particle filter approach. J Theor Biol 2015;370:171-83.