ISSN 0331 - 9504

Journal Of
The Nigerian Statistical Association

NAS

VOL. 19, 2007

JOURNAL OF THE NIGERIAN STATISTICAL ASSOCIATION VOL. 19, 2007

CONTENTS

ISSN 0331-9504

A REVIEW OF STATISTICAL MODELS FOR EARLY DETECTION OF OUTBREAK OF DISEASE EPIDEMICS AND APPLICATION

R.A ADEYEMI
Department of Crop Production,
School Of Agriculture and Agricultural Technology,
Federal University of Technology
MINNA – NIGERIA
email: adeyemira@yahoo.ca

Abstract

Statistical models provide a unique description to available data from public health surveillance systems which can provide meaningful measures of population risks for disease, disability, and death. Analysis and evaluation of these data help public health practitioners react to important health events in a timely manner both locally and nationally.

Different methods exist for monitoring Health Surveillance data, and no method is universally superior. This paper discusses some methods commonly used for detection of outbreaks of diseases epidemics and demonstrates the applicability of two of these models on hospital data. It compares the performance of EWMA and CUSUM models for detection of outbreaks of disease epidemics and finds that EWMA chart is slightly superior to CUSUM models for early detection of outbreaks of malaria and measles.

Key words: Dynamics Models, CUSUM, EWMA, Average Run Length, Acceptable Quality level, infectious disease

1. Introduction

Data from public health surveillance systems can provide meaningful measures of population risks for disease, disability, and death. Analysis and evaluation of these surveillance data help public health practitioners react to important health events in a timely manner both locally and nationally. Statistical methods such as outbreak detection methods are frequently used by epidemiologists and assist them in this task. These methods allow for the rapid assessment of changes in trends of different health outcomes. These detection methods are used to enhance the public health practitioner's ability to identify and characterize unusual trends or clusters in public health surveillance data (for detail, see Langmuir (1963); Thacker *et al* (1989); Teutsch *et al* (2000)).

For the purpose of this paper, Acceptable Quality level is defined as the outbreak frequency (level) at which the process is regarded as in control state while rejected quality level is regarded as the level in which a change in process mean level should be detected quickly (i.e. an out of control state). In this paper, it is important to note that hospital data were used for illustration of the methods and implement to detect outbreak of an infectious disease in Nigeria. Public health surveillance data could facilitate the timely detection of outbreaks which were previously difficult to recognize and allow the operators rapidly and promptly respond to outbreaks and take urgent action to savage any impending problem.

2. Review of Literature

The literature relating to process monitoring and detection methods can be broadly divided into two categories: case definition methods and pattern recognition methods. Pattern recognition methods, which are not discussed in detail in this paper are useful for identifying symptoms (or sets of symptoms) that deviate from the expected baseline and situation requiring attention. Pattern recognition methods may include advance statistical techniques such as Bayesian theory, nearest-neighbour techniques, linear discriminant functions or parameter estimation methods. Visually display of information using charts and graphs are likely methods to identify random variations that are not epidemiologically important. A retrospective analysis can be desirable to provide diagnostic information about the process. It is suggested that, prior to design and implementation of models for public health monitoring, such tests such as 'Span' test, ad-hoc test of significance, the distribution free- CUSUM test, and decision line approach should be run for preliminary diagnosis of the underscore process. (For details, see British Standards Institute1982, Part 4; Lucas, 1985; Osanaiye *et al*, 1989 and Hutwagner *et al,* 2000) ,.

In case definition methods, epidemiological expertise are employed to define an "event of interest" and thus track those syndromes considered of greatest importance, such as hepatitis A or influenza.

There are a number of different statistical methods for detection of aberration in surveillance data. A common group of methods based on regression models was originally proposed by Serfling (1963). In this approach, sinusoidal model is fitted using historical data, and alarm is raised if the observed number of events falls outside a limit in two consecutive sampling points. Related regression methods have been used in France, Hong Kong (Centre for Health Protection, CHP), USA (Centre for disease control and prevention, CDC), United Kingdom (Communicable Disease Surveillance Center CDSC)) and ISIS in the Netherlands.

Another approach that is growing in popularity for automated surveillance is based on CUSUM scheme, which was originally developed for industrial quality control. These methods usually incorporate historical information or are based on short – term data methods (Hutwaggner *et al,* 2005).

Surveillance data naturally form a time series and a variety of methods have been developed for analyzing this form of data. These include ARIMA models from the Box Jenkin's family (Chatfield C., 2001) and dynamic linear models (West *et al,* 1997). Reis *et al* (2003) proposed a hybrid method which uses CUSUM on top of an ARIMA model.

3. Outbreak Detection Models

3.1 Infectious Disease Methods

Infectious disease methods are categorized according to the length of time over which data is used for the surveillance system. Long-term implementation methods expect surveillance to last longer than 30 days. For most long-term systems, it is recommended that only the most recent 5 years of data should be used for the baseline because significant changes occur in populations and in surveillance systems over time. For long-term systems with at least 3 years of baseline data, specific methods have to be implemented for various types of diseases (Hutwagner *et al,* 1997; Stroup *et al,* 1989; Farrington *et al* ,1996; and Simonsen *et al,*1997) schemes such as Shewhart scheme (Lucas, 1985). CDC has routinely applied this method to analyze laboratory-based Salmonella serotype data using the Salmonella Outbreak Detection Algorithm.

The method uses the formula

$$S_t = \max(0, S_{t-1} + \frac{(X_t - \mu_0 + k\sigma_{xt})}{\sigma_{xt}} \qquad (4)$$

for counts of 5 or more with a decision limit, where X_t is the count for the current week, μ_0 is the 5-year weekly mean, σ_{xt} is the standard deviation of the 5-year weekly counts for a given week, k is the detectable shift from the mean, S_t is the current CUSUM value and S_{t-1} is the previous CUSUM calculation.

However, a simpler version of (4) called decision limit standard CUSUM charts are often used to detect an upward or downward shift in process quality (one –sided CUSUM chart) or shift in both directions (two-sided CUSUM chart). To monitor a positive shift from the goal value k, the CUSUM Statistic

$$S_t = \max(0, X_t - k + S_{t-1}) \qquad (5)$$

is used to detect positive shift, while the statistic

$$S_t = \min(0, X_t - k + S_{t-1}) \qquad (6)$$

is used to detect a negative shift. The process is taken to be out of control if $S_t \geq h$ for an upward shift or $S_t \leq -h$ for a downward shift. In this work the starting values $S_0 = 0$. The procedures for determination of the CUSUM k and h are discussed extensively in the literature (Lucas, 1985; Osanaiye *et al* 1989).

3.5 Quality Control Compound Smoothing Method

A quality control method called compound smoothing was developed by Stern and Lightfoot (1999). This method was applied to serotype Salmonella and Shigella data in Australia at national, state, and specified geographic levels. Weekly data are collapsed into monthly data over 5 years for a total of 60 observations, and the compound smoothing technique is applied to the median for consecutive monthly values using 4-month intervals, 2-month intervals, then 5-month intervals, and finally 3-month intervals across the entire time considered.

The current count X_o is then compared with a threshold using the formula

$$X_0 > \beta + 2\sigma_x \qquad (7)$$

where β is the smoothed baseline, and σ_x is the standard deviation calculated as the differences between the smoothed value and the raw value for each data point.

3.6 Exponential Weighted Moving Average (EWMA) Model

It was first introduced by Robert (1959) and later by Wortham and Ringer (1971), who proposed it for applications in the process industries as well as for applications in financial and management control systems for which single observations, x_i, per time

If fewer than 3 years of baseline data are available, one needs to consider alternative outbreak detection methods.

3.2 Historical Limits Method

Centre for Disease Control (CDC) in United States has consistently used historical limits method on infectious disease surveillance data in USA. The method was applied to incidence data for nine diseases (hepatitis A, hepatitis B, hepatitis C/non-A/non-B, legionellosis, meningococcal infections, measles, mumps, pertussis, and rubella) reported to the National Notifiable Diseases Surveillance System (NNDSS). The method was used to compare the number of reported cases in the current 4-week period for a given health outcome with historical incidence data on the same outcome from the preceding 5 years and is based on a comparison of the ratio of current reports with the historical mean. The method uses the formula

$$\frac{x_0}{\mu} > 1 + \frac{2\sigma_x}{\mu} \qquad (1)$$

where x_0 is the current total for a 4-week interval, μ is the mean of 15 totals of 4-week intervals (including the same 4-week period, the preceding 4-week period, and the subsequent 4-week period over the preceding 5 years of historical data), and σ_x is the standard deviation of these 15 historical incidence data values.

3.3 Log-Linear Regression Model

A log-linear regression model was developed by Farrington *et al* (1997). The method has been used to analyze national-level infectious disease data reported to the Communicable Disease Surveillance Center (CDSC) in the United Kingdom. The log-linear regression model is usually used to assess different types of variation in communicable disease incidence data and is represented by

$$\log \mu_i = \alpha + \beta t_i \qquad (2)$$

where μ_i is the mean of the baseline, α is the threshold value, βt_i is the systematic component of the model, and it is sometimes the week indicator. The exceedance score X can be estimated by

$$X = \frac{y_0 - \mu_0}{Y - \mu_0} > 1 \qquad (3)$$

for counts considered to be epidemiologically significant (e.g., counts of 5 or more in the preceding 4 weeks), where y_0 is the current weekly count, μ_0 is the corresponding historical mean, and U is the expected acceptable shift based on a 2/3 power transformation.

3.4 Quality Control Cumulative Sums Methods

Cumulative Sum (CUSUM) chart was originally developed by Page (1954) for process control of variables, which are widely used in the manufacturing industry. The CUSUM schemes usually give tighter process control than classical quality control schemes such as Shewhart scheme (Lucas, 1985). CDC has routinely applied this method to analyze laboratory-based Salmonella serotype data using the Salmonella Outbreak Detection Algorithm.

The method uses the formula

$$S_t = \max(0, S_{t-1} + \frac{(X_t - \mu_0 + k\sigma_{xt})}{\sigma_{xt}} \quad (4)$$

for counts of 5 or more with a decision limit, where X_t is the count for the current week, μ_0 is the 5-year weekly mean, σ_{xt} is the standard deviation of the 5-year weekly counts for a given week, k is the detectable shift from the mean, S_t is the current CUSUM value and S_{t-1} is the previous CUSUM calculation.

However, a simpler version of (4) called decision limit standard CUSUM charts are often used to detect an upward or downward shift in process quality (one –sided CUSUM chart) or shift in both directions (two-sided CUSUM chart). To monitor a positive shift from the goal value k, the CUSUM Statistic

$$S_t = \max(0, X_t - k + S_{t-1}) \quad (5)$$

is used to detect positive shift, while the statistic

$$S_t = \min(0, X_t - k + S_{t-1}) \quad (6)$$

is used to detect a negative shift. The process is taken to be out of control if $S_t \geq h$ for an upward shift or $S_t \leq -h$ for a downward shift. In this work the starting values $S_0 = 0$. The procedures for determination of the CUSUM k and h are discussed extensively in the literature (Lucas, 1985; Osanaiye *et al* 1989).

3.5 Quality Control Compound Smoothing Method

A quality control method called compound smoothing was developed by Stern and Lightfoot (1999). This method was applied to serotype Salmonella and Shigella data in Australia at national, state, and specified geographic levels. Weekly data are collapsed into monthly data over 5 years for a total of 60 observations, and the compound smoothing technique is applied to the median for consecutive monthly values using 4-month intervals, 2-month intervals, then 5-month intervals, and finally 3-month intervals across the entire time considered.

The current count X_0 is then compared with a threshold using the formula

$$X_0 > \beta + 2\sigma_x \quad (7)$$

where β is the smoothed baseline, and σ_x is the standard deviation calculated as the differences between the smoothed value and the raw value for each data point.

3.6 Exponential Weighted Moving Average (EWMA) Model

It was first introduced by Robert (1959) and later by Wortham and Ringer (1971), who proposed it for applications in the process industries as well as for applications in financial and management control systems for which single observations, x_i, per time period may be available (i.e. where sub-grouping may not be practical). The EWMA charting procedures are sometimes used to monitor the rate of occurrences of rare events where the time between two consecutive occurrences is exponentially distributed.

To use EWMA chart detecting an increase in the exponential mean β, Z_t is potted against t where

$$Z_t = \max\{A, (1-\lambda_R)Z_{t-1} + \lambda_R X_t\} \tag{8}$$

for t = 1, 2,... where λ_R is a smoothing constant such that $0 < \lambda_R < 1$, A is a nonnegative boundary and $Z_0 = u$, $A \le u \le h_R$. An out of control signal is given at first t for which $Z_t \ge h_R$.

Similarly, a lower- sided EWMA chart intended for detecting a decrease in the exponential mean β is obtained by plotting

$$Z_t = \min\{B, (1-\lambda_L)Z_{t-1} + \lambda_L X_t\} \tag{9}$$

against t, for t = 1, 2,... where λ_L is a smoothing constant such that $0 < \lambda_L < 1$, B is a nonnegative boundary and $Z_0 = u$, $h_L \le u \le B$. A signal is given at first t for which $Z_t \le h_L$. The detail procedure for designing EWMA chart is given by Gan (1998)

3.7 Cyclical Regression Models

To assess the relationship between influenza virus circulation and cause-specific morbidity and mortality, cyclical regression models have been applied to influenza mortality data and to administrative data on hospitalizations and outpatient visits associated with influenza like illness. In the United States, pneumonia and influenza deaths are monitored using excess death models. Time series autoregressive integrated moving average (ARIMA) models have been applied to pneumonia and influenza deaths to detect outbreaks or increases in the number of reported cases (Semonen *et al*, 1997). CDC implemented this model to evaluate pneumonia and influenza mortality trends. It used 5 years of weekly historical mortality data with epidemic periods removed to estimate a baseline curve of expected weekly deaths for the subsequent season.

The method uses the formula

$$y = \beta_0 + \beta_1{*}T + \beta_2{*}T^2 + \beta_3\mathrm{Cos}(2\pi T) + \beta_4\mathrm{Sin}(2\pi T) \tag{10}$$

where y represents number of deaths from pneumonia and influenza in a particular week, β_0 is the intercept, β_1 and β_2 represent terms associated with the secular trends, and β_3 and β_4 represent cyclical terms associated with seasonal trends.

3.8 Modified CUSUM Methods

Long-term methods are usually implemented with limited baseline data. It entails CUSUM-V1, CUSUM-V2, and CUSUM-V3. These three versions are based on the CUSUM formula (5) with S_{t-1}, the previous CUSUM value equated to zero for CUSUM-V1 and CUSUM-V2. For CUSUM-V1and CUSUM-V2, a flag is signalled if the current count is greater than the baseline mean plus three standard deviations.

The difference among these methods results from computation of the baseline values used for their mean and standard deviation. CUSUM-V1 uses data 1 to 7 days in the past for calculating the mean and standard deviation, and CUSUM-V2 and CUSUM-V3 use data from 3 to 9 days in the past for calculating the mean and standard deviation. In addition, CUSUM-V3 bases its calculation of an average run length time on 3 days (i.e. flag up time).

3.9 Short-Term Methods

The short-term implementation methods, or drop-in surveillance, are represented primarily by traditional quality control methods such as the P-Chart, the moving average, and CUSUM. For these methods, the average run length is usually 30 days. The length of time for the surveillance periods tends to be very short (approximately 21 days); therefore, seasonality factors are less important in the assessment of daily aberrations.

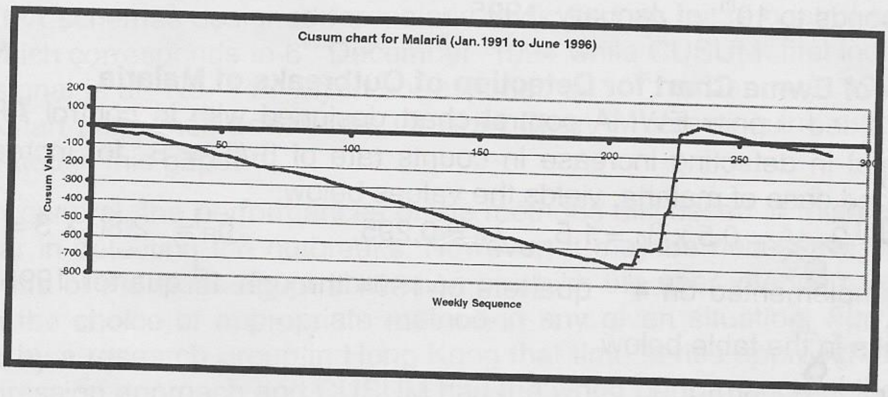
4 Design of Cusum and Ewma Methods

The data used for illustration of the methods in this paper was obtained from Research and Records Unit of Ahmadu Bello Teaching Hospital (ABUTH), Kaduna, Nigeria. It consists of 285 weekly sequences of reported cases of five infectious diseases (Malaria, Measles, Meningitis, Pneumonia and Tuberculosis) between 1st January 1991 and 11th June 1996. These are extracted from a form containing 34 infectious diseases (including Food Poison and HIV/AID) tabulated by age and sex of which these five diseases are prominent (Adeyemi, 2006). Data on malaria was used to illustrate the applicability of two of these methods (CUSUM and EWMA) in this paper.

4.1. Cusum Plot for Malaria

Figure 4.1 shows a graph of the CUSUM value versus time sequence for 285 weeks (January 1991 to June 1996) for malaria. The CUSUM value is obtained by subtracting the mean from each observation and the differences cumulated over time.

Figure 4.1: CUSUM plot for Malaria



From the figure, it is clear that from the starting point of the chart up to and including 204th observation, there was downward trend in the CUSUM indicating that these counts are less than the average process value. From the 205th observation to the 234th observation, there is an upward trend on the CUSUM graph indicating an increase

The difference among these methods results from computation of the baseline values used for their mean and standard deviation. CUSUM-V1 uses data 1 to 7 days in the past for calculating the mean and standard deviation, and CUSUM-V2 and CUSUM-V3 use data from 3 to 9 days in the past for calculating the mean and standard deviation. In addition, CUSUM-V3 bases its calculation of an average run length time on 3 days (i.e. flag up time).

3.9 Short-Term Methods

The short-term implementation methods, or drop-in surveillance, are represented primarily by traditional quality control methods such as the P-Chart, the moving average, and CUSUM. For these methods, the average run length is usually 30 days. The length of time for the surveillance periods tends to be very short (approximately 21 days); therefore, seasonality factors are less important in the assessment of daily aberrations.

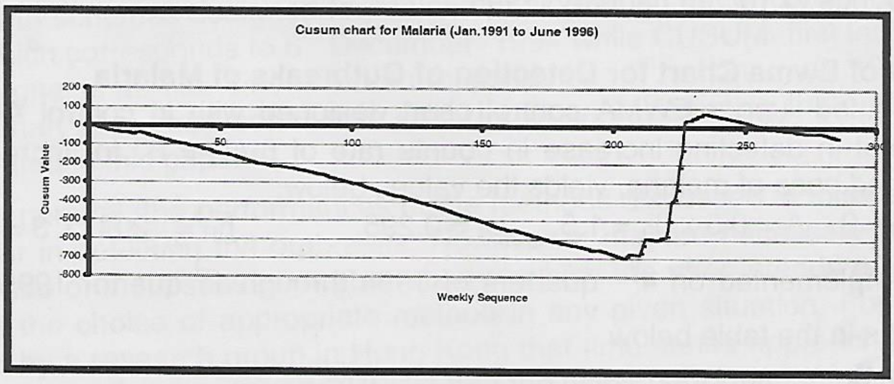
4 Design of Cusum and Ewma Methods

The data used for illustration of the methods in this paper was obtained from Research and Records Unit of Ahmadu Bello Teaching Hospital (ABUTH), Kaduna, Nigeria. It consists of 285 weekly sequences of reported cases of five infectious diseases (Malaria, Measles, Meningitis, Pneumonia and Tuberculosis) between 1st January 1991 and 11th June 1996. These are extracted from a form containing 34 infectious diseases (including Food Poison and HIV/AID) tabulated by age and sex of which these five diseases are prominent (Adeyemi, 2006). Data on malaria was used to illustrate the applicability of two of these methods (CUSUM and EWMA) in this paper.

4.1. Cusum Plot for Malaria

Figure 4.1 shows a graph of the CUSUM value versus time sequence for 285 weeks (January 1991 to June 1996) for malaria. The CUSUM value is obtained by subtracting the mean from each observation and the differences cumulated over time.

Figure 4.1: CUSUM plot for Malaria



From the figure, it is clear that from the starting point of the chart up to and including 204th observation, there was downward trend in the CUSUM indicating that these counts are less than the average process value. From the 205th observation to the 234th observation, there is an upward trend on the CUSUM graph indicating an increase in value. That is, these counts are greater than process mean level . After 235th observation, the CUSUM dropped again.

4.2 Design of Cusum Chart for Detection of Outbreaks of Malaria

A CUSUM scheme designed for incidence of malaria reported between 1991- 1996 at ABTU obtained parameters k = 7 and h = 7, with ARL at AQL, L_a = ARL (μ_a) = 95.5 and the corresponding ARL at RQL, ARL (μ_d) =3.14 . This implies that an out of control signal will be indicated whenever $S_i = \max(0, X_i - 7 + S_{i-1}) \geq 7$ for a case of Malaria.

The scheme (k, h) = (7, 7) implemented on 4th quarters of 1994 through 1st quarter 1995 gives the results in the table below

Sample No.	X	X-k	∑(x-k)	Si =max {0, X_i – k +S_{i-1})
1	1	-6	0	0
2	0	-7	-7	0
3	2	-5	-12	0
4	1	-6	-18	0
5	1	-6	-24	0
6	1	-6	-30	0
7	2	-5	-35	0
8	18	11	-24	0
9	17	10	-14	0
10	5	-2	-16	0
11	4	-3	-19	0
12	4	-3	-22	0
13	15	8	-14	0
14	47	40	26	26*
15	43	36	52	52*
16	6	-1	51	51*

This shows that an out of control signal is first indicated at No. 14 , which corresponds to 10th of January 1995.

4.2 Design of Ewma Chart for Detection of Outbreaks of Malaria

A one sided- upper EWMA control chart designed with in control ARL of 800 which is optimal in detecting increase in counts rate of $\beta_1/\beta_0 = 6$ for increase in the count of reported case of malaria, yields the values below:

$Q_0 = \beta_0 = 3$ $A = 0.5 \times \beta_0 = 1.5$, $\lambda_R = 0.295$, $h_R = 2.80 \times 3 = 8.40$

The scheme implemented on 4th quarters of 1994 through 1st quarter 1995 gives the following results in the table below

Sample No.	X_i	$Z_t = \text{Max}\{A, X\lambda_R + (1-\lambda_R)Z_{t-1}$	$Z \geq 8.4$
1	1	3	--
2	0	2.115	---
3	2	2.081075	--
4	1	1.762158	--
5	1	1.537321	--
6	1	1.378812	--
7	2	1.562062	--
8	18	6.411254	--
9	17	9.534934	* * *
10	5	8.197128	"
11	4	6.958976	"
12	4	6.086078	"
13	15	8.715685	"
14	47	20.00956	"
15	43	26.79174	"

***An out of control signal is first indicated on sample No. 9, which corresponds to 6th December 1994.

5. Summary and Discussion

The outbreaks of two diseases were used for illustration and monitored over 285 weeks (between Jan 1991 and June 1996) by using two of the methods(CUSUM and EWMA models). It was observed that both counted data CUSUM and EWMA charts designed had performed creditable as the were sensitive to detect the outbreaks of epidemics. In both all cases of the designed schemes for monitoring disease outbreaks, the EWMA schemes designed for malaria indicated an out of control signal on sample No. 9, which corresponds to 6th December 1994 while CUSUM first indicated an out of control signal is at No. 14 , which corresponds to 10th of January 1995. In summary, EWMA chart performed better than CUSUM Chart in detecting the outbreak of malaria as illustrated in this paper.

In general, the performances of the methods discussed in this paper are likely to be similar in detecting the outbreaks. However, it may be necessary to investigate the background of the diseases or aberration of which the user intends to study in order to facilitate the choice of appropriate method in any given situation. For instance, it was reported by a research group in Hong Kong that time series approach performed better than regression approach and CUSUM had the worst performance of the three methods for detection of outbreaks of influenza. It was also shown that the peak of the outbreaks occurred in January, February or March each year.

Thus, monitoring models are powerful tools, which can be used to obtain first hand information about an aberration in events of disease attack, disability or death. The EWMA and CUSUM schemes illustrated in this paper further confirms the vital role Quality Control techniques are capable of playing in a non-manufacturing sector as well.

6. Conclusion

The main goal in the development and implementation of these methods is to provide local and state health departments with a tool to assist in the best application of often-limited resources during epidemiological investigations of important public health events. The methods also allow users to select validated aberration detection methods and modify sensitivity and specificity estimates to values considered to be of public health importance by local and state health departments.

More so, these methods can be used with a variety of data sources to produce outputs that enable the users to determine how many resources and personnel they can invest in investigating specific aberrations. Several different public health officials at different locations can implement this output simultaneously.

Acknowledgement

I wish to special acknowledge the assistance and contribution of Prof. Osainaye, University of Ilorin, Nigeria . I also appreciate the invaluable comment and relevant contribution from references below.

References

1. **Adeyemi R.A (2006)**. "Retrospective analysis and Design of CUSUM and EWMA charts for detection of outbreak of Disease epidemics" unpublished M.Sc. Project., University of Ilorin, Nigeria.
2. **British Standard 5730, Part 2 (1990); Part 4 (1982)**
3. **Centers for Disease Control and Prevention (1990)**. Guidelines for investigating clusters of health events. Appendix. Summary of methods for statistically assessing cluster of health events; MMWR Morb Mortal Wkly Rep.;39 (RR-11):17–23.
4. **Centre for Health Protection CHP (March, 2005)**, Department of Health, Hong Kong Special Administrative Region.
5. **Chatfield, C. (2001)**. Time-series forecasting. Boca Raton: Chapman & Hall.
6. **Farrington, C.P., N.J. Andrews, A.D. Beale, and M.A. Catchpole (1996)**, A statistical algorithm for the early detection of outbreaks of infectious disease. Journal of the Royal Statistical Society, Series A, **159**(3): p.547-63.
7. **Hutwagner L.C, Maloney E.K, Bean N.H, Slutsker L, Martin S.M (1997)**. Using laboratory-based surveillance data for prevention: An algorithm for detecting Salmonella outbreaks. *Emerging Infectious t Disease*.;3:395–400.
8. **Hutwagner L, Thompson W, Groseclose S, Williamson G.D (2000)**. An evaluation of alternative methods for detecting aberrations in public health surveillance data. *American Statistical Association 2000 Proceedings of the Biometrics Section*. 82–85.
9. **Hutwagner, L.C., T. Browne, G.M. Seeman, and A.T. Fleischauer, (2005)** Comparison of aberration detection methods using simulated data. *Emerging Infectious Diseases*,. **11**(2): p. 314-6.
10. **Osanaiye P.A. and Talabi O. (1989)**. On some non-manufacturing applications of counted data cumulative sum (CUSUM) control schemes, *The Statistician*.
11. **Page, E.S. (1954)** "Continuous Inspector Schemes, "*Biometrika*, **41**, 100-114.
12. **Langmuir A.D. (1963)**. The surveillance of communicable diseases of national importance. *N Engl J Med.*; 268:182–192.

13. **Lucas J.M (1985)**. Counted data CUSUM's, *Technometrics* ; 27:129-44.
14. **Gan, F.F. (1991)**. An optimal design of CUSUM quality control charts," *Journal of Quality Technology*, 23, 279- 286.
15. **Stroup D.F, Williamson G.D, Herndon J.L, Karon J. (1989)**. Detection of aberrations in the occurrence of notifiable diseases surveillance data. *Statistics in Medicine*.; 8:323–329.
16. **Farrington C.P, Andrews N.J, Beale A.D, Catchpole M.A. (1996)**. A Statistical algorithm for the early detection of outbreaks of infectious disease. J Roy Stat Soc A.; 159:547–563.
17. **Reis, B.Y., Pagano, M. and Mandl, K.D. (2003)**. Using temporal context to improve Biosurveillance. Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(4): p. 1961-5.
18. **Scientific Committee on Advanced Data Analysis and Disease Modeling (SCADADM, 2003)**. Working Group on Influenza-like Illness Statistical Detection of Aberrations (WGILI) Submitted to Working Group Report on A Comparison of Methods for Early Detection of the Influenza Peak Season in Hong Kong.
19. **Serfling, R.E.(1963)**. Methods for current statistical analysis of excess pneumonia-influenza deaths. Public Health Reports. **78**: p. 494-506.
20. **Simonsen L, Clarke J.M, Stroup D.F, Williamson G.D, Arden N.H, Cox NJ. (1997)**. A method for timely assessment of influenza- associated mortality in the United States. *Epidemiology*.; 8:390–395.
21. **Stern L, Lightfoot D. (1999)**. Automated outbreak detection: a quantitative retrospective Analysis. *Epidemiology Infection*;122:103–110.
22. **Stoumbos Z.G. (2000)**.The state of statistical process control as we proceed into the 21st century. J Am Stat Assoc.;95:992–998.
23. **Thacker S.B. & Berkelman R.L. (1988)**. Public health surveillance in the United States; *Epidemiology* Review.;10:164–190.
24. **Thacker SB, Berkelman RL, Stroup D.F (1989)**. The Science of public health surveillance, *Journal Public Health Policy*.;10:187–203.
25. **Teutsch S.M & Churchill RE,(2000)**. Principles and Practice of Public Health Surveillance. New York, NY: Oxford University Press.
26. **West, M. and J. Harrison (1997)**. Bayesian forecasting and dynamic models. 2nd ed., New York: Springer.