Recent Research in Public Health Surveillance and Health Management

K.L. Tsui, D. Goldsman, W. Jiang, S.Y. Wong

Abstract - While the challenges of the next pandemic outbreak are overwhelming, either from swine flu, other infectious disease, bioterrorism, timely detection of disease outbreaks is most important for public health surveillance and society safety and stability. In public health surveillance, the objective is to systematically collect, analyze, and interpret public health data (chronic or infectious diseases) in order to understand trends, to detect changes in disease incidence and death rates, and to plan, implement, and evaluate public health practice. Recently much research has been conducted to develop methods and algorithms for health surveillance and disease detection. This paper presents an overview and reviews the recent research methods on temporal and spatiotemporal surveillance. Specific research challenges and future research directions are discussed. A real life example is used to compare the performance of three currently used surveillance methods, scan, EWMA, and CUSUM.

I. HEALTHCARE AND PUBLIC HEALTH SURVEILLANCE

The Institute of Medicine [1] reported that the number of deaths due to medical errors in U.S. hospitals may have exceeded 100,000 per year; and the numbers of unnecessary surgeries and hospital infections have topped 12,000 and 80,000, respectively. It is well recognized in the research community that many of these events can be avoided by the effective use of healthcare standardization, improvement and surveillance methods. On the other hand, in light of the latest outbreaks of H5N1 influenza and the continuing bioterrorism threat, there is an urgent need for specific research on the design and modelling of disease transmission and public health surveillance.

Many methods and algorithms for disease detection have been evaluated, but only on a rather small scale. For example, certain methods might be tested on just a few or even only one specific disease, in most cases influenza. This causes various problems. First, influenza is a periodically recurring disease resulting in numerous infections. Thus, it is relatively easy to detect, and many methods based on extensive historic data can be applied. In the case of bioterrorism defense, a possible attack would cause comparably few cases in comparison to an epidemic, and historical data on such events simply do not exist. When implementing approaches in disease surveillance, this paucity of data has to be taken into account.

The most common existing disease spread monitoring methods can be categorized into temporal, spatial, and spatiotemporal surveillance techniques. Most basic methods such as SPC, regression, time series, and forecast-based methods were originally developed as temporal approaches. On the other hand, popular health surveillance methods such as scan statistics were originally developed as spatial approaches [2] and later extended as temporal and spatiotemporal approaches [3]. Most spatial surveillance techniques rely on existing statistical clustering methods. Many techniques have been developed to expand those models to spatiotemporal methods that also search for clusters in time. For healthcare surveillance, sets-based methods [4] and risk adjustment methods [5] were developed for monitoring patient disease conditions. Woodall [6] provides detailed discussions of these methods for healthcare applications. Tsui et al. [7] provides detailed reviews on public health surveillance.

In public health applications, temporal surveillance refers to monitoring event occurrence at a single region or location and the objective is to detect accurately the time of change in occurrence rate as soon as possible. Spatiotemporal surveillance refers to monitoring event occurrences at multiple regions or locations simultaneously and the objective is to detect accurately the time and location(s) of change in occurrence rate as soon as possible. This paper focuses on the recent research in methodologies for temporal surveillance and spatiotemporal surveillance. Then, we will discuss the specific challenges and future research in temporal and spatiotemporal surveillance.

II. TEMPORAL SURVEILLANCE

- **Scan and CUSUM methods**

Three temporal cases in scan statistics are defined by Naus and Wallenstein [8]. Three different scenarios are addressed: (1) a window of pre-defined length scans for time intervals with an increased incidence rate where events are distributed over a continuous time frame, (2) time is split into intervals, so that a scan is performed for the maximum number of cases within a certain set of contiguous time intervals, and (3) Bernoulli case where a scan is performed on the occurrence of an event. Woodall et al. [9] point out that CUSUM statistics may be able to outperform scan statistics because of their optimality property. Joner et al. [10] evaluate the CUSUM's and scan statistic's performance under various rate increases of Bernoulli trials and show that the CUSUM outperforms the scan statistics in many situations. They considered both initial expected delay (out-of-control average run length (ARL)) and steady-state expected delay as performance measures. Kulldorff [2] proposes scan statistics for spatial surveillance, where the region diameter (window) of the scan statistic is chosen to maximize an unconditional likelihood function. As a result, Kulldorff's spatial scan statistic uses a variable window, unlike the traditional scan statistic which uses a fixed window [11]. Further, Kulldorff [3] extends the spatial scan statistics to spatiotemporal surveillance. Kulldorff's scan statistic for pure temporal surveillance has never been formally defined. However, Sonesson [12] points out that Kulldorff's scan statistics for spatiotemporal surveillance are similar to CUSUM statistics if conditional likelihood is considered. By removing the spatial component of Kulldorff's spatiotemporal scan statistics, one can interpret Kulldorff's scan statistics for temporal surveillance as CUSUM statistics. We shall continue our discussion regarding this topic in the section of spatiotemporal surveillance.

Tsui et al., [7] highlight the importance of investigating and comparing the performance measures of surveillance methods under various scenarios. Recent research has been undertaken to fill in the existing research gap. Mei, Han, and Tsui [13] compare the effect of detecting a change in the mean of Poisson distribution from three classes of detection schemes. Taking into account the effect of population size, the study investigates the performance of three methods: (1) generalized likelihood ratio statistics, (2) weighted likelihood ratios, and (3) adaptive threshold method by simulation studies and asymptotic analysis. The results point out that the weighted likelihood ratio based detection schemes are the best when the population sizes increase. Contrary, when the population sizes decrease, the generalized likelihood ratio based detection schemes are the best. In the sense of small detection delays under Lorden's worst-case detection delay criterion, the detection schemes based on the adaptive threshold method seem to be robust regardless the change in population size. Han et al. [14] compare the performance of temporal scan statistics, CUSUM, and exponential weighted moving average (EWMA) under Poisson distribution. The study shows that the Poisson CUSUM and EWMA charts generally outperformed the Poisson scan statistic methods and the EWMA charts outperformed the CUSUM charts in situations with a small shift and an early change in time.

- **Forecast-based surveillance methods**

Due to the non-stationary nature of disease data, it is often appealing to model the baseline pattern before applying monitoring methods. Regression and time series models are the most popular methods for modelling baseline patterns. Alternatively, adaptive forecasting methods and nonparametric regression can be effective tools for capturing complex seasonal patterns. Forecast-based methods have been developed in manufacturing SPC applications for a long time. Alwan and Roberts [15] propose monitoring residuals of fitted time-series models to the original process. Since then, many residual charts (e.g., Jiang et al., [16]) have been proposed and investigated.

Since count data are usually monitored in health applications, generalized linear models (GLMs) are popular tools for modelling in disease monitoring. The regression-based approach has been widely implemented in ESSENCE for the Greater Washington DC area [17]. Daily ICD-9 code data were collected and compared to baseline values for the occurrence of certain diseases, for instance, the cases of diarrhea in three different regions of the U.S. were identified using this method in 2002. Burkom et al. [18] have applied non-adaptive and adaptive regression models for quickly picking up short-term trends and compared them to a Holt-Winters model. They took into account an 8-week sliding baseline and a holiday indicator to enhance forecast accuracy on holidays and prevent overly high forecasts on those days. Due to the fact that disease spread data is subject to many influence factors ranging from interventions through the healthcare system to changes in the weather, non-adaptive regression methods do not perform well in Burkom's evaluation on actual disease data. Although the adaptive regression methods have better performance, they still cannot entertain holiday problems for accurate forecasting.

Box-Jenkins autoregressive integrated moving average (ARIMA) formulation is one of the most popular and well-established methods for time series analysis. In disease surveillance, it is often not feasible to assume stationarity of data. Lai [19] suggests stabilizing the variation by performing a log transformation of the time series that it might be useful to undertake curve fitting to improve the performance of time series approaches on quickly changing rates of disease cases. Instead of evaluating an ARIMA approach of automated time series forecasting methods, Burkom et al. [17] use pre-selected ARIMA models or a pre-selected set of smoothing variables in general.

Tsui et al. [20] implement a forecasting approach to three data streams simulated from different types of real-life outbreaks: E. Coli, cryptosporidium, and influenza. It is found that the forecasting approach with EWMA residual charts can be extremely effective for detecting outbreaks if the baseline can be modelled accurately. A two-step approach is proposed for a surveillance system to estimate baseline pattern and detect outbreaks. To improve the model fit, a nonparametric regression approach is chosen to model the baseline pattern. For outbreak detection, a modified EWMA method with simulation-based fine-tuning approach is proposed. Three outbreak types of diseases with three variables ED (visits to the emergency department), OTC (over-the-counter drug sales), and TH (number of nurse hotline telephone calls) are being considered. The authors apply different algorithm scoring methods as the outbreak patterns for the three diseases are different. The *E. Coli* outbreak, which is evaluated based on the average length of the delay until detection, shows a sudden increasing profile and is easy to detect from monitoring, whereas the cryptosporidium and influenza outbreaks are hard to detect and are evaluated by the correct number of outbreaks detected. The study shows that the average delay time is around six days and the correct rates of outbreak detection are around 95 to 99%.

- **Research Challenge in Temporal Surveillance**

The main advantage of the residual charts is that many existing monitoring charts can be applied to the residuals (or forecast errors) of the fitted model. The main disadvantage is that the charting performance critically depends on how well the forecast model fits the process. In addition, even if the forecast model fits well with the historical data, there is no guarantee that the selected model will remain unchanged in the future. As a result, adaptive forecasting methods, such as exponential smoothing [21, 22] may be good alternatives for monitoring dynamic processes. Jiang et al. [16] propose a robust class of forecasting methods (PID charts) for SPC applications. Jiang, Au, and Tsui [23] propose a robust adaptive forecasting approach for monitoring the forecast errors of multiple customer time series in business activity monitoring. As pointed out in these papers, while it is possible to develop robust forecasting methods for the baseline data, it is challenging to distinguish the anticipated change patterns in the baseline (e.g., growth or seasonal patterns) from the unanticipated changes (e.g., customer fraud, outbreaks, etc.). It will be of great interest to investigate the two conflicting objectives of (i) being robust to various patterns in fitting baseline processes and at the same time (ii) being sensitive in detecting process shifts (outbreaks). We discuss various specific future research ideas below.

Traditional SPC monitoring methods often focus on detection of step shifts in mean performance. Tsung and Tsui [24] show that standard control charts may not be efficient in detecting special types of mean shifts, especially when the shift only lasts for a short period of time (i.e., the window of opportunity). Shu, Jiang, and Tsui [25] develop a weighted CUSUM chart for detecting patterned mean shifts resulting from forecasting or feedback control. The new approach is efficient when the stream of data to be monitored experiences a dynamic pattern of mean shift, i.e., the mean shift in the sequence of data is not constant but varies over time. In disease surveillance, we often encounter a dynamic patterned mean shift. Therefore, the weighted CUSUM approach may be useful for detecting disease outbreak patterns.

To detect changes in time series data, Alwan and Roberts [15] propose to monitor the forecasting residuals of fitted ARIMA models. A major challenge of the approach is the difficulty of finding the correct model and the impact of model misspecification. Due to their simplicity, adaptability, and computational convenience, EWMA models have been widely used and considered a strong competitor to ARIMA models. They are particularly useful when ARIMA models cannot be easily characterized. There are three popular exponential smoothing techniques for different forecasting problems: single exponential smoothing for stationary time series, double exponential smoothing [21] for time series with a trend, and triple exponential smoothing [22] for seasonal time series. In addition, Montgomery and Mastrangelo [26] propose monitoring the residuals of EWMA forecasting. While EWMA has been recognized as a powerful forecasting tool, very little research has been conducted to investigate the performance of residual charts from EWMA forecasting. In disease surveillance, the background patient count trends can be very different for different diseases. As shown in Jiang et al. [16], the EWMA statistic may be quite useful for forecasting highly varying time series. Burkom et al. [18] apply EWMA smoothing to a temporal surveillance problem. While the forecast-based monitoring approach has great potential in disease surveillance applications, much research is needed to identify specific forecasting strategies and to determine optimal parameters in various applications.

To better facilitate the detection of an increase in occurrence rates, research can be conducted in one-sided Poisson EWMA charts. Some foundational works have been already done to discuss one-sided EWMA charts for monitoring continuous data, and one-sided multivariate EWMA charts [27]. To study other types of outbreak patterns, especially transient outbreaks, conventional detection methods have focused on detection of step shifts in mean. In health surveillance, many types of the change may occur. The research direction could be pointed to more sophisticate methods to investigate the baseline parameter under an in-control process that shifts and drifts over time because of changes in population or seasonal effects, for instance, the application of a risk adjustment method. Also, more efficient detection methods that can handle unexplained noises spikes should be developed [14].

III. SPATIOTEMPORAL SURVEILLANCE

- **Scan and CUSUM statistics**

Kulldorff et al. [2] and Kulldorff [3] propose retrospective and prospective scan statistics, respectively, based on the basic theory for the spatial case explained above. Woodall et al. [9] review and examine the various scan statistics used for spatiotemporal surveillance. They point out that this method can only work if a cluster contains at least two incidence counts because otherwise the likelihood function could be maximized by decreasing the size of the cycle while centered on a single data point. By extending the scan statistics in Kulldorff [3], Sonesson [12] proposes a CUSUM method for timely detection of emerging clusters of diseases. Based on a general likelihood function, he defines a general CUSUM statistic as the maximum over all individual CUSUM likelihoods over all possible subsets of regions (variables).

By considering the CUSUM and other LR-based approaches in temporal surveillance for industrial quality control, Tsui, Han, Jiang, Woodall [28] develop a generic framework based on likelihood ratio statistics over windows of test for both the spatial surveillance and spatiotemporal surveillance problems. The LR-based framework includes many existing methods as special cases by taking different operators over the LR statistics in variable windows. The study demonstrates the use of the summation of likelihood ratios over all possible windows is often more powerful than the use of the maximum of the same set of likelihood ratios. Taking summation for spatial and spatiotemporal surveillance, the framework outperforms the common procedure that takes the maximum over all windows of the LR statistics. It is also found that when the outbreak coverage is known, scan statistics with an appropriate radius that matches the actual outbreak coverage often perform better than the statistics under- or over-scanned.

Most existing health surveillance research is based on the assumption that observations from different regions are independent. Recently, Jiang, Han, Tsui, Woodall [29] propose a set of multivariate surveillance schemes generalized from well-known detection methods in multivariate statistical process control based on likelihood ratio tests. A multivariate CUSUM method using regression-adjusted clusters for spatiotemporal surveillance in the presence of spatial correlation is proposed. It shows that the proposed schemes outperform the existing surveillance methods and provide faster and more accurate detection of outbreaks. It also points out that estimating the outbreak magnitude may help the detection of outbreaks when there are medium-sized spatial correlations.

- **Multivariate Surveillance Methods**

In disease surveillance, the SPC-based spatiotemporal surveillance problem has been studied by just a few researchers. Rogerson [30] presents some control chart-based work on the spatiotemporal regional case. He extends the retrospective method of Tango [31] to a prospective application using a CUSUM method. Rogerson and Yamada [32] consider the spatiotemporal problem for which the counts in the sub-regions are correlated at each particular time, with the correlation decreasing as the distance between the sub-regions increases. They compare the performance of the use of multiple CUSUM charts [33] for each region against a multivariate CUSUM method [34].

Joner et al. [35] show that the use of a one-sided version of the multivariate EWMA chart of Lowry et al. [36] is a better approach to use in this case.

- **Research Challenge in Spatiotemporal Surveillance**

We discuss specific challenges and future research in spatiotemporal surveillance methods in public health applications. As mentioned earlier, there are close relationships among the scan/CUSUM statistics, the multivariate methods for disease surveillance, as well as the traditional methods proposed in the multivariate SPC literature for normal data.

Under various applications, different CUSUM statistics have been proposed for monitoring multi-streams (regions) data. Tartakovsky et al. [37] study a simpler CUSUM statistic similar to the one proposed in Sonesson [12]. Their idea is to first compute the temporal CUSUM statistic for each region, then to define the overall test statistic as the maximum over the individual CUSUM statistics over the multiple streams. Mei [38] considers a similar idea to Tartakovsky et al. [37] but defines the test statistic as the sum of the individual CUSUM statistics over multiple streams. Raubertas [39] develops another CUSUM statistic: Instead of computing the temporal CUSUM statistic for each region, he proposes to compute the simultaneous CUSUM statistic for the center and four (or fewer) adjacent regions for each region center, and then define the overall testing statistic as the maximum over these simultaneous CUSUM statistics.

Note that the difference between Tartakovsky et al. [37] and Mei [38] is similar to the difference between the M-statistic (the maximum of standardized $\overline{X}$ statistics of individual responses) defined in Hayter and Tsui [40] and the T^2 statistic defined in Hotelling [41], where the Shewhart $\overline{X}$ statistics from individual variables are considered instead of the CUSUM statistics. As pointed out in Hayter and Tsui, the advantage of the M-statistic (or equivalently the maximum statistic in Tartakovsky et al. [37]) is that, once the test statistic triggers the alarm, it immediately provides information on which variables are responsible for the alarm. In disease surveillance, this implies that, once the test statistic signals, it immediately provides information on which regions are the locations where the outbreak occurs. The problem of identifying which variables (or regions) are responsible after the alarm is triggered is well-known — the identification problem — in the multivariate SPC literature (see Mason et al. [42]; Mason and Young [43]).

While the identification problem has been investigated extensively for multivariate Shewhart charts with normal data, very little has been done for non-normal data and/or for multivariate CUSUM or EWMA methods. In addition, most research has been focused on how to perform diagnosis to identify the variables that are responsible for the alarm; however, a more important and fundamental problem in variables (regions) identification has been overlooked — that is the lack of performance measures for identification correctness. Note that after the alarm is signaled, certain variables (regions) based on the diagnosis procedure will be identified as the variables where changes have taken place. In reality, it is possible that only some (but not all) of these regions have changed. When none of the identified regions have actually changed, the performance is measured by the false alarm rate (or in-control ARL). When all of the identified regions have changed, the performance is measured by the out-of-control ARL (or expected delay). However, when only some (but not all) of the identified regions have changed, the performance needs to be measured by simultaneous measures of the out-of-control ARL as well as the correct rate of identification. This creates a research challenge on the choice of appropriate performance measures as well as on understanding the conflicting behavior of the expected delay and the correct identification rate.

For CUSUM charting, Woodall and Ncube [33] define a multiple CUSUM method under multivariate normal distributions. The idea is similar to considering the maximum over the CUSUM statistics of individual responses. Similar to the T^2 statistics, Crosier [34] proposes a multivariate CUSUM method as charting the T^2 statistic over the CUSUM statistics of individual responses. An alternative multivariate CUSUM statistic is defined as the CUSUM statistic of the univariate T^2 statistic from the multivariate vector of responses (see Pignatiello and Runger, [44]). While some simulations have been conducted for evaluation purposes, it is not clear which CUSUM statistic has better performance in general. Jiang and Tsui [45] relate the T^2 chart, M-chart, and regression-adjusted chart [46] and show that the performance of a multivariate control chart may depend on correlation of the variables as well as the direction and magnitude of the process shift. They propose a hybrid method that combines the T^2 chart and regression-adjusted chart for better performance.

For EWMA charting, Lowry et al. [36] propose a multivariate EWMA method as charting the T^2 statistic over the EWMA statistics of individual responses. Tsui and Wooodall [47] propose a multivariate EWMA method as charting the EWMA statistics of the loss function from the multivariate vector of responses. Similar to the research of MSPC charts for multivariate normal distributions, it is of great interest to investigate the various multivariate CUSUM and EWMA charting methods for public health surveillance of multivariate Bernoulli and Poisson data.

Currently most spatiotemporal surveillance methods are based on the assumption of a known/given outbreak pattern. One of the critical research directions in spatiotemporal surveillance would be to estimate both outbreak coverage and magnitude while recent studies only have little knowledge about them and substitute the estimates into the corresponding scan statistic in practice [29].

IV. A TEMPORAL SURVEILLANCE EXAMPLE

Below we illustrate some standard methods described above (scan, EWMA, CUSUM) with an example from Han et al. [14]. Our dataset contains the incidence of male thyroid cancer in New Mexico, 1973-2005, which is available through data from the Surveillance, Epidemiology, and End Results (SEER) Program at the National Cancer Institute (www.seer.cancer.gov/data/). The SEER program collects cancer incidence and mortality from the cancer registries in the United States. Figure 4 plots the annual incidence of thyroid cancer per 100,000 men. The main goal of this application is to detect the change in rates as early as possible. It can be seen from Figure 1 that the rate increases after 1989 or so, and therefore, it is assumed that there are no shifts between 1973 and 1988. We used this steady-state period to estimate the baseline incident rate, which is around 2. We targeted to detect a 25 % increase of the incident rate, which is equivalent to the targeted shift size of 2.5.

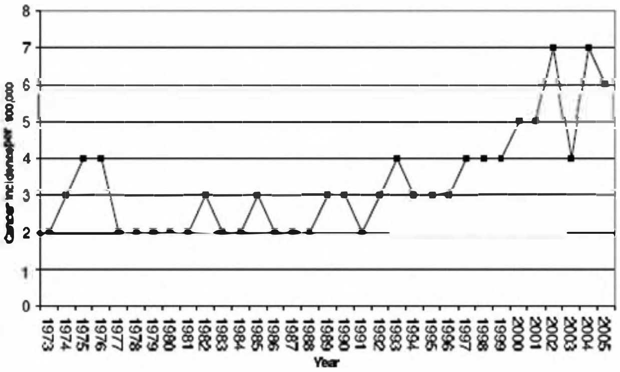

Figure 1: The trend of male thyroid cancer incidence between 1973 and 2005.

The parameters of the three detection methods were determined according to Section 4.1 of Han et al. [14]. The target in-control average run length (ARL0) was set to 1,000. The parameters were λ_0=2 and λ_1=2.5 for CUSUM, m=37 for scan statistics, and α=0.02 for EWMA. Consequently, the thresholds of the three methods for target ARL0 are 97 (scan statistics), 16.6 (CUSUM), and 2.33 (EWMA). Figure 2 shows the statistics over time from the three detection methods. In order to ensure comparability of the different methods and use the same threshold, we adjusted the values of the statistics of the scan statistic method and CUSUM by dividing them by 41.63 and 7.124, respectively. It can be observed that the scan statistic, EWMA, and CUSUM methods trigger an alarm in 2004, 1999, and 2000, respectively. Assuming that early detection is desirable, EWMA and CUSUM triggered an alarm faster than the scan statistic method, which is consistent with the simulation results reported in Han et al. [14]. Nevertheless, it is difficult to argue that EWMA and CUSUM were better methods in this particular example as the true outbreak time is unknown for this specific example. More importantly, the message here is that different surveillance methods based on the same data set can result into different outbreak alarm time. Hence, one should be careful about selecting surveillance methods in real life applications. More investigations and simulation studies in comparing the three methods can be found in [14] and [28].

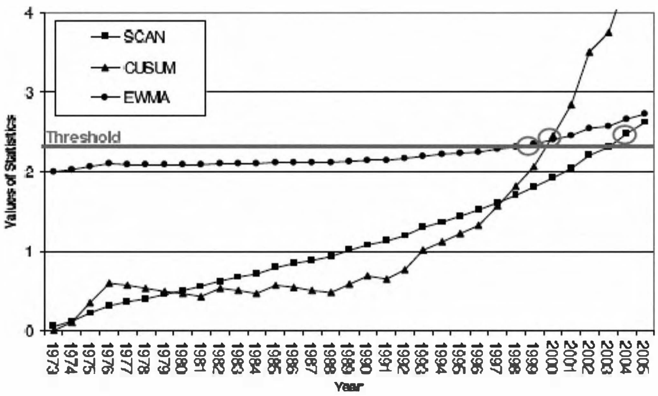

Figure 2: Plots of scaled scan statistic, CUSUM statistic, and EWMA statistic. The circle indicates the first time point when each method triggers an alarm.

Conclusions

With the urgent needs of healthcare performance improvement, the latest outbreaks of avian influenza, and the continuing bioterrorism threat, timely detection of increases in the rate of unusual events is an important objective in public health surveillance and health management [6, 48]. Also, due to the advancement of health information technology, medical information systems, and public health and syndromic data collection systems, there are great opportunities as well as challenges for research in public health and disease surveillance. The quality of the baseline data can be readily contaminated by unexplained noise spikes. These noises may adversely affect the performance of detection methods. Further, the baseline (phase I) data are generally non-stationary and correlated. There are many types of outbreak patterns related to diseases to be detected. Research effort should be devoted to develop robust surveillance methods detecting these various outbreak patterns. Sophisticate surveillance methods aiming to investigate the baseline parameter under an in-control process that shifts and drifts over time because of changes in population or seasonal effects may be required [14].

Finally, the current performance measures for temporal surveillance are not satisfactory for comparing spatiotemporal surveillance methods [7]. Although some recent research has been undertaken to fill in the existing research gap, we believe more research is needed to comprehensively consider the underlying assumptions and scenarios under real life healthcare and public health surveillance applications.

References

1. Institute of Medicine (2000). To Err is Human, National Academy Press, Washington, DC.
2. Kulldorff, M. (1997). "A spatial scan statistic", Communications in Statistics-Theory and Methods 26(6): 1481-1496.
3. Kulldorff, M. (2001). "Prospective time periodic geographical disease surveillance using a scan statistic", Journal of the Royal Statistical Society Series A-Statistics in Society 164: 61-72.
4. Chen, C. (1978). "A surveillance system for congenital malformations", Journal of the American Statistical Association 73: 323-327.
5. Steiner, S. H., Cook R. J., Farewell V. T., and Treasure T. (2000). "Monitoring surgical performance using risk-adjusted cumulative sum charts", Biostatistics 1(4): 441-452.
6. Woodall, W. H. (2006). "The use of control charts in health-care and public-health surveillance", Journal of Quality Technology 38(2): 89-104.
7. Tsui K.L., Chiu W., Gierlich P., Goldsman D., Liu X., Maschek T. (2008). "A review of healthcare, public health, and syndromic surveillance", Quality Engineering 20(4): 435-450.

8. Naus, J. and Wallenstein S. (2006). "Temporal surveillance using scan statistics", Statistics in Medicine 25(2): 311-324.
9. Woodall, W. H., Marshall J. B., Joner M. D., Jr., Fraker J. E., and Abdel-Salam A. G. (2008). "On the use and evaluation of prospective scan methods for health-related surveillance", Journal of the Royal Statistical Society: Series A 171: 223-237.
10. Joner, M. D. Jr., Woodall W. H., Reynolds M. R. Jr. (2007). "Detecting a rate increase using a Bernoulli scan statistic", Statistics in Medicine, 27, 2555-2575.
11. Glaz, J., Naus J., and Wallenstein S. (2001). Scan Statistics. Springer, New York.
12. Sonesson, C. (2007). "A CUSUM framework for detection of space-time disease clusters using scan statistics", Statistics in Medicine 26(26): 4770-89.
13. Mei Y., Han S.W., Tsui K. L. (2009). "Early detection of a change in Poisson rate after accounting for population size effects", Submitted to Statistica Sinica.
14. Han S. W., Tsui K. L., Ariyajunya B., Kim S. B. (2009). "A comparison of CUSUM, EWMA, and temporal scan statistics for detection of increases in Poisson rates", Quality and Reliability Engineering International. Forthcoming.
15. Alwan, L. C. and Roberts H. V. (1988). "Time series modeling for statistical process-control", Journal of Business and Economic Statistics 6(1): 87-95.
16. Jiang, W., Wu H. Q., Tsung F. G., Nair V. N., and Tsui K. L. (2002). "Proportional integral derivative charts for process monitoring", Technometrics 44(3): 205-214.
17. Lewis, M. D., Pavlin J. A., Mansfield J. L., O'Brien S., Boomsma L. G., Elbert Y., and Kelley P. W. (2002). "Disease outbreak detection system using syndromic data in the greater Washington DC area", American Journal of Preventive Medicine 23(3): 180-186.
18. Burkom, H. S., Murphy S. P., and Shmueli G. (2007). "Automated time series forecasting for biosurveillance", Statistics in Medicine 26(22): 4202-4218.
19. Lai, D. (2005). "Monitoring the SARS Epidemic in China: A time series analysis", Journal of Data Science 3(3): 279-293.
20. Tsui, K. L., Chiu, W. Goldsman, D., Han, S. W., Kim, M. Liu, X. Lo, L. and Woodall, W. H. (2009). submitted to International socity for disease surveillance (ISDS)
21. Holt, C. C. (1957). "Forecasting seasonals and trends by exponentially weighted moving averages", ONR Research Memorandum, Carnegie Institute 52.
22. Winters, P. R. (1960). "Forecasting sales by exponentially weighted moving averages", Management Science 6(3): 324-342.
23. Jiang, W., Au T., and Tsui K. L. (2007). "A statistical process control approach to business activity monitoring", IIE Transactions 39(3): 235-249.
24. Tsung, F. G. and Tsui K. L. (2003). "A mean-shift pattern study on integration of SPC and APC for process monitoring", IIE Transactions 35(3): 231-242.
25. Shu, L. J., Jiang, W., and Tsui, K. L. (2008). "A weighted CUSUM chart for detecting patterned mean shifts", Journal of Quality Technology, 40:194-213.
26. Montgomery, D. C. and Mastrangelo C. M. (1991). "Some statistical process-control methods for autocorrelated data", Journal of Quality Technology 23(3): 179-193.
27. Shu L, Jiang W, Wu S. (2007). "A One-Sided EWMA Control Chart for Monitoring Process Means. Communications in Statistics", Simulation and Computation. 36:901-920.
28. Tsui, K. L., Han, S. W., Jiang W., Woodall, W. H. (2009). "Likelihood ratio methods for spatial and spatiotemporal surveillance", Submitted to Technometrics.
29. Jiang, W., Han, S. W. Tsui, K. L., Woodall, W. H. (2009). "Spatiotemporla surveillance methods in the presence of spatial correlation", Submitted to Statistics in Medicine.
30. Rogerson, P. A. (1997). "Surveillance systems for monitoring the development of spatial patterns", Statistics in Medicine 16(18): 2081-93.
31. Tango, T. (1995). "A class of tests for detecting general and focused clustering of rare diseases", Statistics in Medicine 14(21-22): 2323-2334.
32. Rogerson, P. A. and Yamada I. (2004). "Monitoring change in spatial patterns of disease: Comparing univariate and multivariate cumulative sum approaches", Statistics in Medicine 23(14): 2195-2214.
33. Woodall, W. H. and Ncube M. M. (1985). "Multivariate CUSUM quality control procedures", Technometrics 30(3): 291-303.
34. Crosier, R. B. (1988). "Multivariate generalizations of cumulative sum quality-control schemes", Technometrics 30(3): 291-303.
35. Joner, M. D. Woodall Jr., W. H., Reynolds M. R., Jr., and Fricker R. D. (2008). "A One-Sided MEWMA Chart for Health-Related Surveillance", to appear in Quality and Reliability Engineering International.
36. Lowry, C. A., Woodall W. H., Champ C. W., and Rigdon S. E. (1992). "A multivariate exponentially weighted moving average control chart", Technometrics 34(1): 46-53.
37. Tarakovsky, A. G. and Veeravalli V. V. (2004). "Change-point detection in multichannel and distributed systems with applications", in: Applications of Sequential Methodologies (N. Mukhopadhyay, S. Datta and S. Chattopadhyay, Eds), Marcel Dekker, Inc., New York, 331-363.
38. Mei, Y. (2008) "Robust scalable schemes for monitoring multiple data streams", submitted for publication.
39. Raubertas, R. F. (1989). "An analysis of disease surveillance data that uses the geographic locations of the reporting units", Statistics in Medicine 8:267-271.
40. Hayter, A. J. and Tsui K. L. (1994). "Identification and qualification in multivariate quality control problems", Journal of Quality Technology 26(3): 197-208.
41. Hotelling, H. (1947). Multivariate Quality Control. Techniques of Statistical Analysis, Eisenhart, Hastay, and Wallis (eds.), McGraw-Hill, New York.
42. Mason, R. L., Tracy N. D., and Young J. C. (1995). "Decomposition of T^2 for multivariate control chart interpretation", Journal of Quality Technology 27:99-108.
43. Mason, R. L. and Young J. C. (2002). Multivariate Statistical Process Control with Industrial Applications

(ASA-SIAM Series on Statistics and Applied Probability). American Statistical Association, VA.

44. Pignatiello, J. J. Jr. and Runger G. C. (1990). "Comparisons of multivariate CUSUM charts", Journal of Quality Technology 22(3):173-186.
45. Jiang, W. and Tsui K. L. (2008). "A theoretical framework and efficiency study of multivariate statistical process", IIE Transactions, 40(7): 650-663.
46. Hawkins, D. M. (1991). "Multivariate Quality Control Based on Regression-Adjusted Variables", Technometrics 33(1): 61-75.
47. Tsui, K.-L. and Woodall W. H. (1993). "Multivariate control charts based on loss functions", Sequential Analysis 12(1): 79-92.
48. Rolka, H., Burkom H., Cooper G. F., Kulldorff M., Madigan D., and Wong W. K. (2007). "Issues in applied statistics for public health bioterrorism surveillance using multiple data streams: Research needs", Statistics in Medicine 26(8): 1834-1856.