Spatial and Spatio-temporal Epidemiology 2 (2011) 11–21

Contents lists available at ScienceDirect

Spatial and Spatio-temporal Epidemiology

journal homepage: www.elsevier.com/locate/sste

Incorporating retrospective clustering into a prospective cusum methodology for anthrax: Evaluating the effects of disease expectation

Ian Kracalik [a,*], Larissa Lukhnova [b], Alim Aikimbayev [c], Yerlan Pazilov [b], Gulnara Temiralyeva [b], Jason K. Blackburn [a]

[a] Spatial Epidemiology and Ecology Research Laboratory, Emerging Pathogens Institute and Department of Geography, University of Florida, Gainesville, Florida, USA
[b] Kazakh Science Center for Quarantine and Zoonotic Diseases, Almaty, Kazakhstan
[c] Republican Sanitary and Epidemiological Station, Almaty, Kazakhstan

ARTICLE INFO

Article history:
Received 6 January 2010
Revised 8 June 2010
Accepted 18 June 2010

Keywords:
Anthrax
Cumulative sum
Space–time clustering
Baseline expectation
GeoSurveillance

ABSTRACT

We analysed livestock anthrax in Kazakhstan from 1960–2006, using a prospective CUSUM to examine the affects of expectation on the detection of spatio-temporal clusters. Three methods for deriving baselines were used for CUSUM; a standard z-score, AVG, a spatially-weighted z-score derived from Local Moran's *I*, LISA, and a moving-window average, MWA. LISA and AVG elicited alarm signals in the second year that did not return below threshold during the 47-year period, while MWA signaled an alarm at year four and relented at year fifteen. The number of spatial clusters elicited varied: LISA $n = 16$, AVG $n = 11$, and MWA $n = 3$, although there were clusters present around Shymkent, in south-central Kazakhstan, in each method. The results illustrate that the selection of a baseline with an unknown background population has a significant effect on the ability to detect the onset of clusters in space and in time when employing a CUSUM methodology.

1. Introduction

The most common techniques for examining the presence of spatial or spatio-temporal clusters are retrospective analyses, which are often carried out as a onetime analysis of past data, after the onset, and often after the completion of an outbreak (Tango and Takashi, 2005). The prevalence of these techniques in the literature provides a valuable reference resource for analysing the spatial (Moran, 1950; Anselin 1995; Ord and Getis, 1995; Kulldorff, 1997) and/or temporal distribution (Knox, 1964; Mantel, 1967; Wallenstein, 1980) of past health events.

Yet, in human and veterinary epidemiology the objective is often to detect the onset of health events as quickly as possible. Statistical applications that allow for the continual evaluation of a disease status over time are advantageous since they may be able to identify the onset of clusters in a timelier manner. In this area of health analyses retrospective techniques may incur specific limitations due to issues of multiple hypotheses testing that occur when these methodologies are used to measure a disease status continuously over time (Tango, 2000). Alternatively, prospective statistical techniques such as the cumulative sum (CUSUM) approach (Page, 1954), originally developed for process control, can be used in a continuous detection system to monitor the status of disease over time in an attempt to detect the onset of clusters (Rogerson, 1997). A more comprehensive review of prospective techniques has been described elsewhere (Sonesson and Bock, 2003; Woodall et al., 2008).

For the purposes of this study we are particularly interested in the application of clustering techniques used in the monitoring of veterinary health. The application of space–time clustering using both retrospective techniques (Carpenter et al., 1996; Hoar et al., 2003; D'Orazi et al., 2007) and prospective techniques (Mostashari et al., 2003; Hohle et al., 2009) has been shown to be successful

* Corresponding author. Tel.: +1 760 481 8179.
E-mail address: itk@ufl.edu (I. Kracalik).

1877-5845/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.sste.2010.06.001

in research examining the spatial and temporal distribution of diseases in veterinary epidemiology. A study by Ward et al. (1996) found clustering of bluetongue virus serotypes among cattle herds in Queensland, Australia using the Cuzick and Edward's test. Research using retrospective space–time clustering techniques in the investigation of acute respiratory infections in cattle suggests that the identification of clusters link multiple illnesses to a single pathogen (Norstrom et al., 2000). Several studies have also utilized the SaTScan software to retrospectively identify the clustering of bovine spongiform encephalopathy in cattle (Doherr et al., 2002; Sheridan et al., 2005; Allepuz et al., 2007; Heres et al., 2008). Ward and Carpenter (2000) and Carpenter (2001) provide reviews of additional methods used to investigate the distribution of health events in veterinary epidemiology.

In the field of prospective surveillance Rogerson (1997) employed a one-sided CUSUM approach using a modified Tango's statistic to reanalyse data from Williams et al. (1978) in order to identify the presence of emerging space–time patterns of Burkitt's lymphoma in Uganda. This study found the CUSUM method detected the emergence of additional clusters previously unidentified by retrospective analyses. Research applying early detection methods in livestock surveillance incorporated a log-linear regression method to derive expected counts from a baseline dataset, which was constructed using historical case records, in order to elucidate anomalies in the distribution of *Salmonella spp.* infections (Kosmider et al. 2006). Additionally, Gilbert et al. (2005) illustrates the efficacy of the monitoring of livestock diseases by deriving model parameters from historical data, in conjunction with biotic and abiotic variables to predict a shifting geographic distribution of Bovine Tuberculosis (BTB) on a yearly basis. Like in the aforementioned case of BTB there is a crucial need to monitor other zoonotic livestock/wildlife pathogens (those transferrable from animals to humans) such as anthrax that threaten not only animal populations, but human populations as well.

Bacillus anthracis, the causative agent of anthrax, is a gram-positive spore-forming bacterium, that affects livestock and wildlife (primarily herbivorous ungulates), and secondarily humans (Van Ness, 1971). Outbreaks of the disease in Central Asia, including Kazakhstan (Woods et al., 2004), have increased in recent years due to inadequacies in public health and veterinary surveillance (Hugh-Jones, 1999).

Several studies have described the spatial and temporal distribution of anthrax infections in livestock (Dragon et al., 1999; Turner et al., 1999; Parkinson et al., 2003; Clegg et al., 2007; Himsworth and Argue, 2008; Mongoh et al., 2008). Van Ert et al. (2007) showed through mapping the phylogeography of *B. anthracis* that its global distribution may be influenced by its genetic variation. Research has also used GIS mapping in conjunction with ecological niche modeling to predict the potential geographic distribution of *B. anthracis* in the US (Blackburn et al., 2007) and in Kazakhstan (Joyner et al., 2010). However, few of these studies have applied spatio-temporal techniques to quantitatively describe the distribution of anthrax infections. This is also true in research looking at human infections of the disease, which have either focused on, the bioterrorist event in the US in 2001(Jernigan et al., 2002; Webb and Blaser, 2002), syndromic studies related to potential bioterrorism (Kleinman et al., 2005; Buckeridge et al., 2006), or the accidental release of weaponized anthrax in Sverdlovsk, Russia in 1979 (Meselson et al., 1994; Wilkening, 2006).

The few studies that have utilized spatio-temporal statistical techniques to analyse the distribution of anthrax infections have illustrated the potential usefulness of these tools. Initial research by Smith et al. (1999) identified three anthrax isolates responsible for wildlife epidemics in Kruger National Park (KNP), South Africa and found using the Mantel's test they were clustered in both space and time. A subsequent study by Smith et al. (2000) indicated through the use of SaTScan that there was distinct spatio-temporal clustering of two major anthrax strain types, A and B within KNP, due to possible differences in soil composition that may have exerted an influence on the location of each strain. Current research on the distribution of anthrax outbreaks is limited to retrospective analyses allowing for the implementation of prospective methodologies to add to the current body of anthrax literature.

The purpose of this current study was to conduct an exploratory analysis of the spatial and temporal distribution of historical anthrax outbreaks among livestock in Kazakhstan utilizing a prospective CUSUM approach. Specifically this study had two objectives: (1) examine the methods for deriving a baseline rate of disease for use in a CUSUM methodology when no population data are available and (2) to evaluate the influence that various derived expectations of disease have on the detection of clusters in space and time in an annual CUSUM methodology. This study represents one of the first prospective statistical examinations of anthrax in livestock.

2. Methodology

As part of a larger effort to map and model the geographic distribution of anthrax and its control in Kazakhstan, the Kazakh Science Center for Quarantine and Zoonotic Disease developed a spatial database of database totaling 3963 outbreaks that were reported over a 74-year period from 1933–2006 (Aikembayev et al., 2010). This current study employs a selection of that historical record. A subset of the data representing livestock outbreaks (a combination of small and large ruminant outbreaks) between 1960 and 2006 comprised of 2920 outbreaks was selected for analyses in order to analyse the distribution of anthrax positive outbreaks from the post-vaccination time period. For the purposes of this study an outbreak was defined as a location that reported one or more positive confirmations of an infection of anthrax in livestock.

In order to derive baseline expectations for analyses using CUSUM an additional subset of the data representing outbreaks from 1950 through 1959 were selected. This subset of data was employed to calculate baseline values using a moving-window average and a standard *z*-score technique (Fig. 1). A subset of data from this period was necessary since the calculation of baseline rates from, for

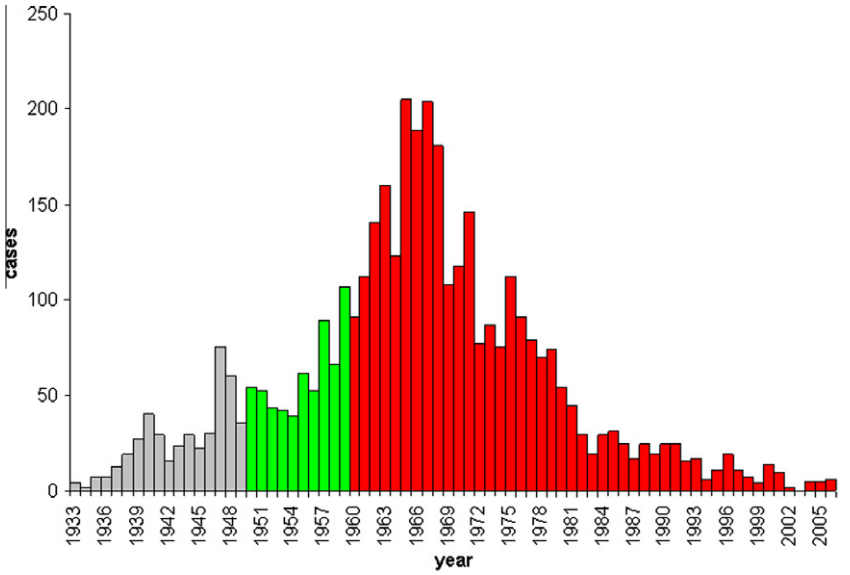

Fig. 1. Total number of anthrax outbreaks by year, among livestock, in Kazakhstan. Grey coloured bars indicate a subset of data that was not used in this study 1933–1949 and the green coloured bars represent subset of data that were used in part to calculate expected values 1950–1959 but were not incorporated spatially into the analyses. Red coloured bars portray the subset of the data that was used in the analyses representing outbreaks from the post-vaccination time frame 1960–2006. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

example, the year 1960 would require the inclusion of case numbers from previous years. The locations and dates of the livestock outbreaks were provided as latitude/longitude coordinate pairs georeferenced to the nearest community (Fig. 2). Outbreak locations were then aggregated to 174 administrative rayons (region or county equivalent) by year using Hawth's Tool point-in-polygon count (Beyer, 2004). This provided forty-seven fields of data representing the total number of outbreaks per rayon for each year from 1960 to 2006.

2.1. Prospective analysis

Examining the emergence of disease clusters on a continuous basis in both space and time may allow for a more rapid evaluation of disease status. The CUSUM technique measures sequential deviations from an expected mean of a variable over time, in an attempt to represent events of the variable that might require immediate attention. In order to implement an annual CUSUM approach the freeware software package GeoSurveillance 1.1 (Rogerson and Yamada, 2004) was used. The one sided univariate cumulative sum (CUSUM) statistic that was implemented in GeoSurveillance requires a standardized z-score as the input variable and is written following (Rogerson, 1997):

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

where S_t is the summation of all events from the z-score Z_t that deviate more than k standard units from a mean and are accumulated over time. When the accumulation of S_t over time crosses a pre-designated threshold h an alarm signal is triggered indicating an emerging cluster of outbreaks. Alarm events in a CUSUM methodology in this instance are monitored using a control chart that analyses a process compared to a critical h. Events that are summated in excess of h are deemed to be an out of control signal (exceeding a threshold) while events that fall below h are deemed to be an in control signal (below a threshold). The selection of h is based on the average run length (ARL_o) or the amount of time that is desired between false alarms (spurious signals above baseline). For example, in an annual study an average run length of 100 would signify that on average a false alarm is triggered every 100 years (Rogerson, 1997). A value of $k = 0.5$, as in this case, is often chosen to minimize the amount of time required to achieve an alarm signal and represents one-half of the total deviation from the mean that is to be registered (Rogerson, 1997).

In this study two different average run lengths were used to evaluate the signaling sensitivity of emerging anthrax outbreaks on an annual basis during the time period 1960–2006 (total of 47 years) and to also assess the effect that varying target values provided in the form of z-scores would have on the performance of the control chart. The first CUSUM evaluation used an arbitrary ARL of 100 years and was adjusted for multiple testing using a Bonferroni correction (# of Regions × ARL) resulting in a value of $h = 7.9$. The second CUSUM evaluation set the probability of a false alarm occurring in the study period of n observations at 0.1 and employed an ARL that was derived using the equations set forth in (Rogerson 2001):

$$1 - \exp(-(\#\ \text{of outbreaks})/\text{ARL}) = 0.1$$

The equation produced an average run length of 27,705. The ARL of 27,705 was adjusted for multiple testing using a Bonferroni correction (ARL × number of regions) resulting in a value $h = 17.4$.

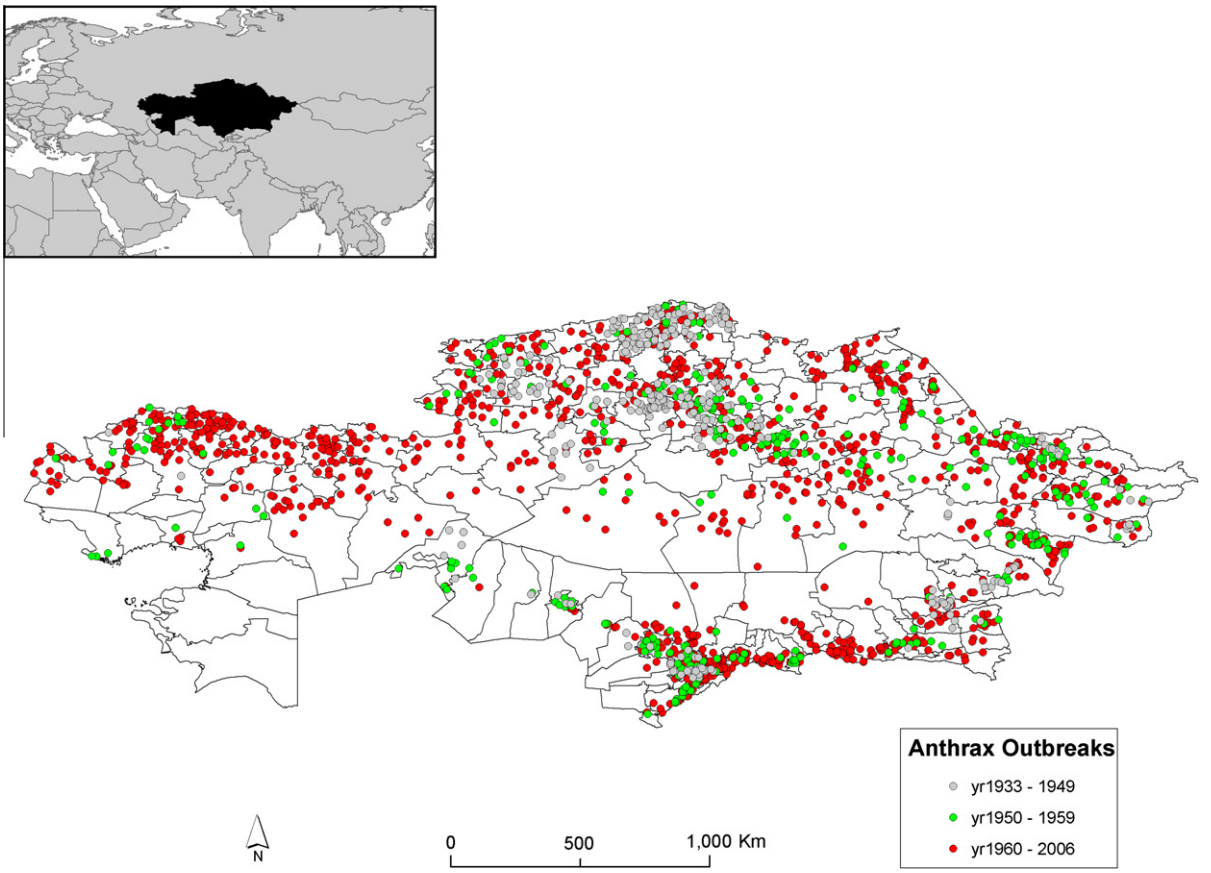

Fig. 2. Location of anthrax positive infections in livestock across Kazakhstan from 1930–2006 consisting of 3963 outbreaks. Grey points indicate anthrax positive outbreaks during the time period 1930–1949, green dots represent outbreaks during the time period 1950–1959, and red dots portray outbreaks from the post-vaccination time period 1960–2006. Map symbol colours match the colours used in the histogram of outbreak numbers in Fig. 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

2.2. Evaluating the effects of expectation on CUSUM

Choosing a baseline rate of disease in livestock epidemiology is a crucial step in locating areas of high case numbers. This is especially true in endemic regions with little knowledge of the underlying population at risk. Since selecting a level of disease expectation will strongly influence the response of public health and veterinary management to a health event, multiple scenarios should be examined. Due to a lack of livestock population data that corresponds to the village-level outbreak data, there was no inherent expectation associated with the livestock contracting anthrax. Therefore, using models that implemented a standard observed and expected calculation of *z*-scores was less suitable for this data. In order to gain an understanding of how disease expectation would affect the CUSUM control charts varying degrees of expectation were derived. Standardized *z*-scores used as input variables in GeoSurveillance 1.1 were obtained using three different methods. Each different methodology employed here was used to substitute for a baseline level of disease presence for each year during the 47-year period.

2.2.1. LISA-based calculation of expectation (LISA)

Instituting a baseline expectation from administrative aggregations may, in some instances, be inappropriate since the occurrence of disease often does not stop at arbitrary administrative boundaries. Therefore, incorporating analyses that take into account the spatial relationship among neighboring regions in the calculation of a baseline disease rate may provide a more accurate assessment of expectation. The first methodology, referred to as LISA from here on, employed the use of the retrospective Local Moran's *I* statistic used for measuring local spatial autocorrelation (Anselin 1995). This test evaluates the existence of spatial autocorrelation or local clusters by examining the contribution of a rayon to the global autocorrelation Moran's *I* statistics for the entire country. The statistic was implemented in GeoDa 0.9.5-I (Anselin et al., 2006) using a Queen contiguity matrix, and 999 permutations at an $\alpha < 0.05$.

LISA calculations were performed a total of 47 times, once for each year. The resulting output provided standardized z-scores for all rayon across the time period. Standardized *z*-scores were computed using the Local Moran's *I* Statistic following Anselin (1995):

$$(E); E[I_i] = -w_i/(n-1)$$

$$(S); S = w_{i(2)}(n-b_2)/(n-1) + 2w_{i(kh)}(2b_2-n)/(n-1)(n-2) - w_i^2/(n-1)^2$$

for each statistic I_i the expected value (E) is divided by the square root of the variance (S). Computing z-scores in this manner explored the use of consecutive retrospective analyses in a CUSUM methodology. The application of this statistic did not include any controls for population or herd density.

2.2.2. Expectation based on the average number of outbreaks (AVG)

Often times epidemiological investigations into a disease outbreak will uncover a temporal relationship. Deriving the expectation of disease from a temporal component may therefore be a useful technique in establishing a baseline rate. In this case z-scores were derived using the second methodology, referred to as AVG from here on, that incorporated a calculation based on the average number of outbreaks for all rayons in a specific year. Calculations were made comparing the outbreaks within a given rayon to the average and the standard deviation of all outbreaks for each time period. This provided time specific z-scores for each rayon based on the average number outbreaks for each year and its contribution to the average. The methodology presented here allows for the baseline rate of disease to be based on a temporal component and a comparison of the global average of the study area.

2.2.3. Moving-window average calculation of expectation (MVA)

Comparing the historical persistence of outbreaks to the current disease status on a region-by-region basis may allow for the more accurate determination of a disease presence. Regions, or rayons in this case, that show a high level of temporal endemicity may require a higher expectation of disease compared to rayons with a sporadic or non-existent presence of the disease. The third calculation of z-scores, referred to as MWA from here on, used a moving-window average methodology that compared the number of outbreaks in a rayon back to itself, rather than other rayons (as in AVG calculations), in a single time period. Adjusting for the lack of population homogeneity in this case can be achieved by conditioning the total number of outbreaks observed to calculate the expected number of outbreaks for each location (i) a form of indirect adjustment (Ward and Carpenter, 2000). Therefore, in this instance a moving average was applied to the data to construct an expected value. The expected value for a rayon at a specific time period was based on the average number of outbreaks from the previous × number of years in that same rayon. The calculation used was (e.g. for 3 years):

$$\mathrm{Exp}_{yi} = \frac{yi-1 + yi-2 + yi-3 \ldots yi-N}{3 \ldots N}$$

where Exp_{iy} is the expected value for a rayon i with x number of outbreaks at time y. Calculations of expected values using the MWA method were performed for the previous 1, 2, 3, 4, 5, 6, and 10 years. The expected values obtained from the previous calculation were then used along with the observed number of outbreaks to compute z-scores using the following standard formula (Lee et al., 2007):

$$Z_i = \sqrt{obs} + \sqrt{obs+1} + \sqrt{4\exp+1}$$

The corresponding expected values were used to calculate z-scores for each MWA and then incorporated into a CUSUM model. Selection of a MWA time range of ten years was selected based in part on the fact that it produced higher max CUSUM values.

2.3. Z-score distribution

In order to visualize potential differences in the distribution of values between methodologies, z-scores from a single year, 1960, were displayed graphically. A graph was created that displayed the range of z-scores for the MWA, LISA, and AVG methodologies using outbreak data for the year 1960. In this approach the direction and magnitude of the deviation from the mean is visible for each methodology and each rayon during a single year.

3. Results

3.1. Prospective analysis

The CUSUM analysis using an ARL of 100 ($h = 7.9$), shows the spatial relationship of each of the three different methods for calculating z-scores MWA, AVG, and LISA during the period 1960–2006 (Fig. 3). During the 47-year period the MWA methodology had the lowest number of rayons eliciting alarm signals $n = 3$, while the LISA methodology showed the highest number of rayons with alarm events $n = 16$ and the AVG methodology had $n = 11$ rayons signal an alarm. The presence of alarm events was consistent across all three methodologies in the rayons of Southern Kazakhstan near the city of Shymkent, where the same two rayons signaled alarms events. Additionally, an emerging cluster was identified in a single rayon to the west of Almaty, which signaled an alarm event using all three methods. Furthermore, there was also a similarity in the spatial distribution of the AVG and LISA methodology. The spatial distribution of rayons signaling alarm events in those two methods were around Shymkent, to the west of Almaty, and the south of Semipalantisk with a few signals west of Aqtobe.

Temporal results portraying the control chart for the initial CUSUM analysis ($h = 7.9$) showed that all three methods for calculating z-scores triggered alarm events during the time period 1960–2006 (Fig. 4). The control chart revealed that AVG and LISA methods triggered alarm events at year two and persisted as an alarm signal for the duration of the study period. An alarm signal was also triggered using the MWA method, but at a later time than AVG and LISA, at year four. The alarm signal in the MWA method did not persist for the length of the time period like that of the AVG and LISA methods, instead it initiated at year four and relented at year 15.

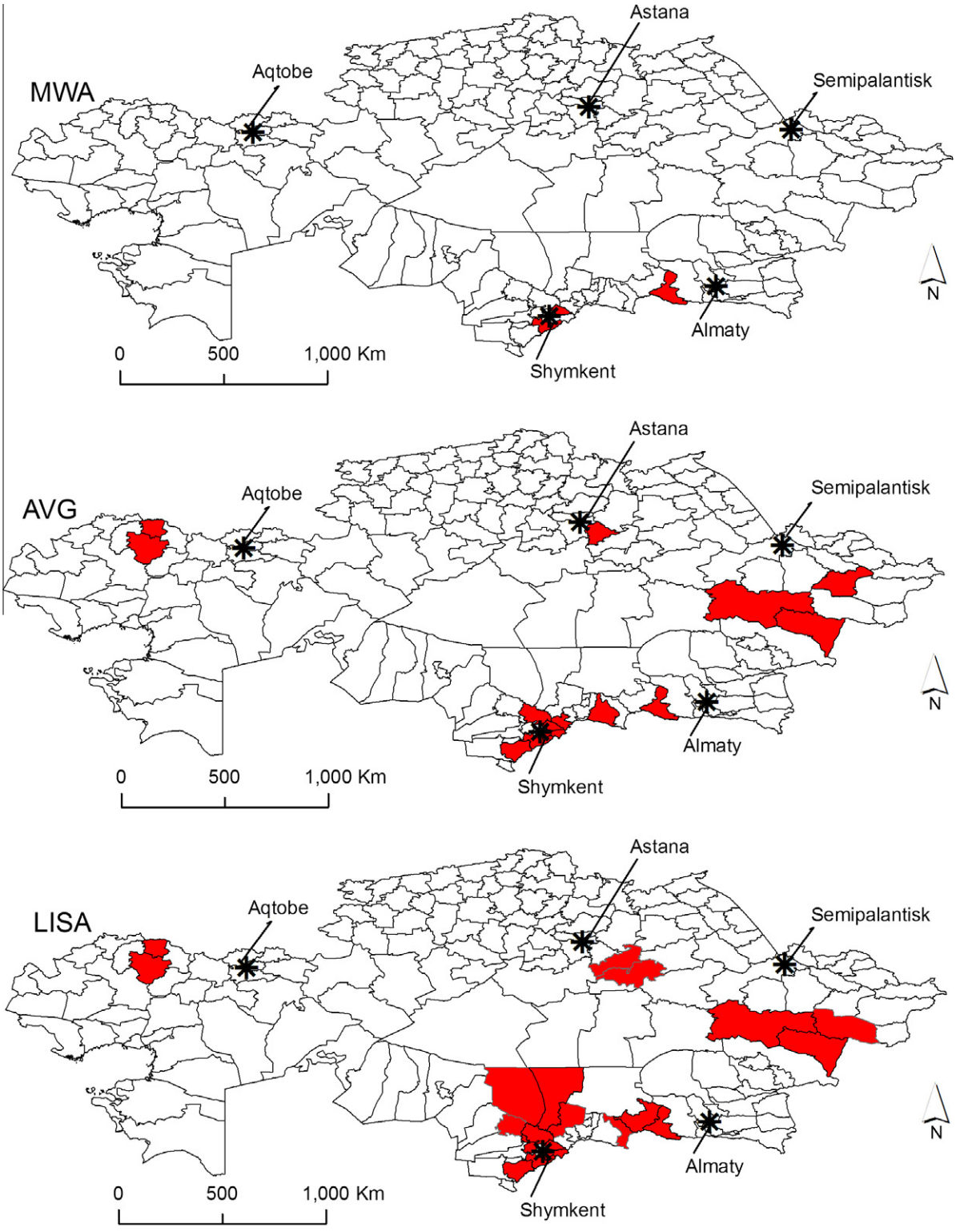

Fig. 3. Spatial distribution of alarm signal events in CUSUM (h = 7.9) shown in red, over a 47-year period using the three different methods for calculating z-scores MWA, AVG, and LISA. Results show that LISA method for calculating z-scores had the highest number of rayons eliciting an alarm. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

Temporal results from the CUSUM analysis using a critical threshold of h = 17.4 revealed that not all of the methods triggered alarm signals (Fig. 5). The control chart shows that the LISA method and AVG method both initi-

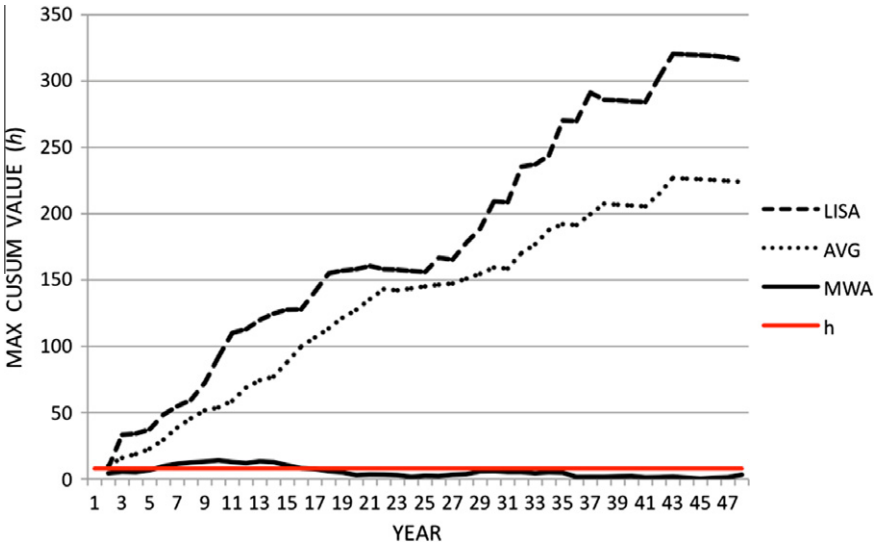

Fig. 4. Max CUSUM of anthrax outbreaks during the time period 1960–2006 with an $h = 7.9$, portraying the control chart for the LISA, AVG, and MWA methods. The three different methodologies each signal an alarm event represented by the crossing of the critical h threshold (red line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

ated alarm signals at year two and persisted for the duration of the study period. The MWA method, however, failed to elicit an alarm signal during the 47-year period.

The evaluation of multiple MWA time frames showed differences in the ability to trigger an alarm signal based on the size of the moving window (Fig. 6). The graph shows that as the time range for each MWA increases so does the max CUSUM value. Additionally, the graph indicates that the max CUSUM values from the averages for years one through four would not trigger an alarm signal during the 47-year period using the chosen values of $h = 7.9$ or 17.4. There is a noticeable decrease in the inter-annual variability of the larger moving windows, as the length of time in the moving window increases the amount of variability in the signal decreases. In the case of the 10 year window the initial max CUSUM is greater than the other window time frames, but decreases at year thirty below the level of the window time frames 1–5.

3.2. Z-score evaluation

The range of z-scores from each of the three methodologies is illustrated for all rayons for the year 1960 (Fig. 7). The distribution of z-scores illustrates the differences between the magnitude and direction of z-scores for each methodology used in the calculation of an expected disease rate. The range of z-scores when compared between methodologies shows that within certain rayons there is an inverse relationship among values. Rayons numbered 17–32 on the graph in Fig. 7 revealed that the LISA meth-

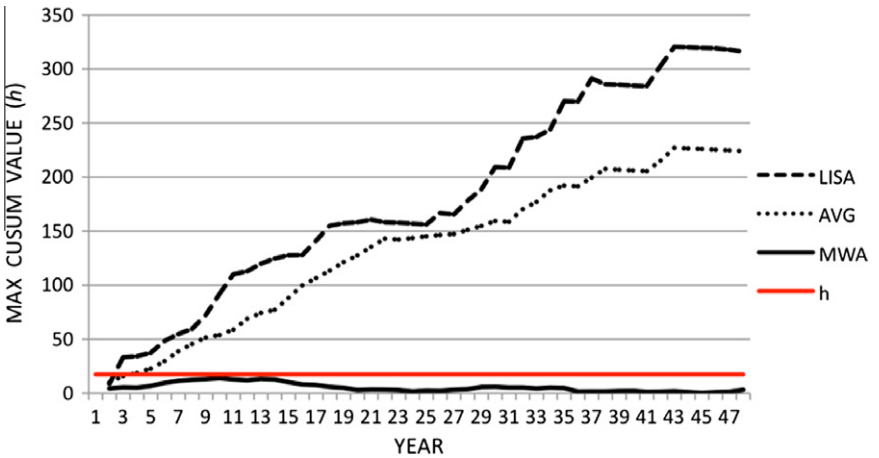

Fig. 5. Max CUSUM of anthrax outbreaks in livestock during the time period 1960–2006 with an $h = 17.4$, showing the control chart for the LISA, AVG, and MWA methods. Alarm events are signaled for the LISA and AVG methods, represented by the crossing of the critical h threshold (red line), but not the MWA. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

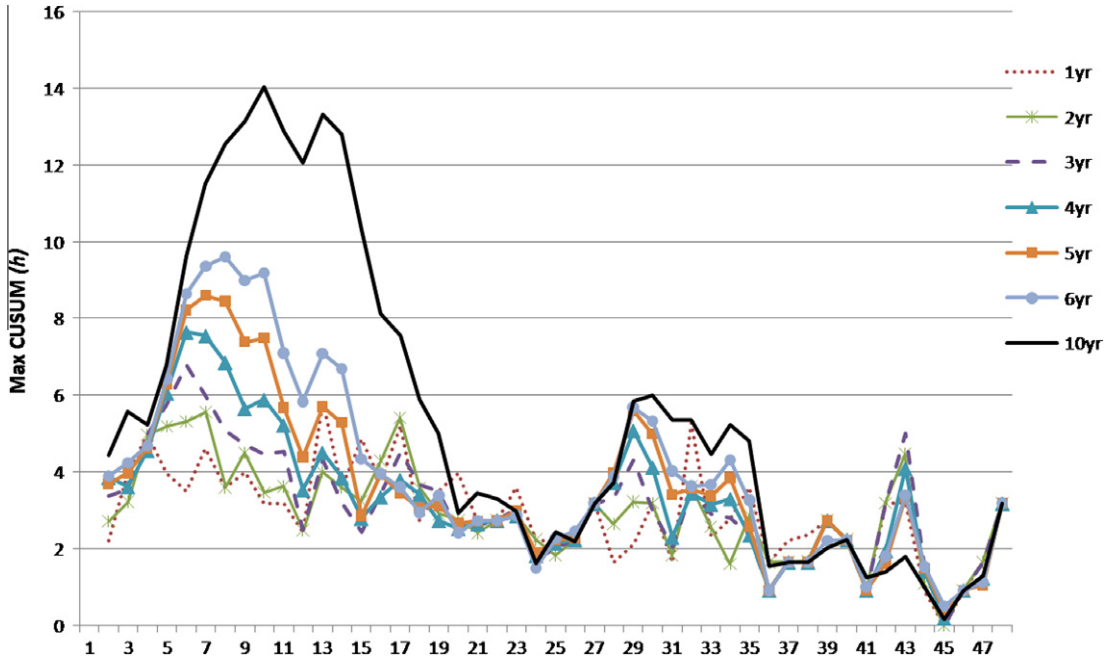

Fig. 6. Control chart for anthrax outbreaks in Kazakhstan using varying expectations. Chart portrays signal levels derived from various *z*-score calculations using expected values from a range of moving-window averages. Moving-window averages were calculated for the previous 1-, 2-, 3-, 4-, 5-, 6-, and 10-year periods for a rayon.

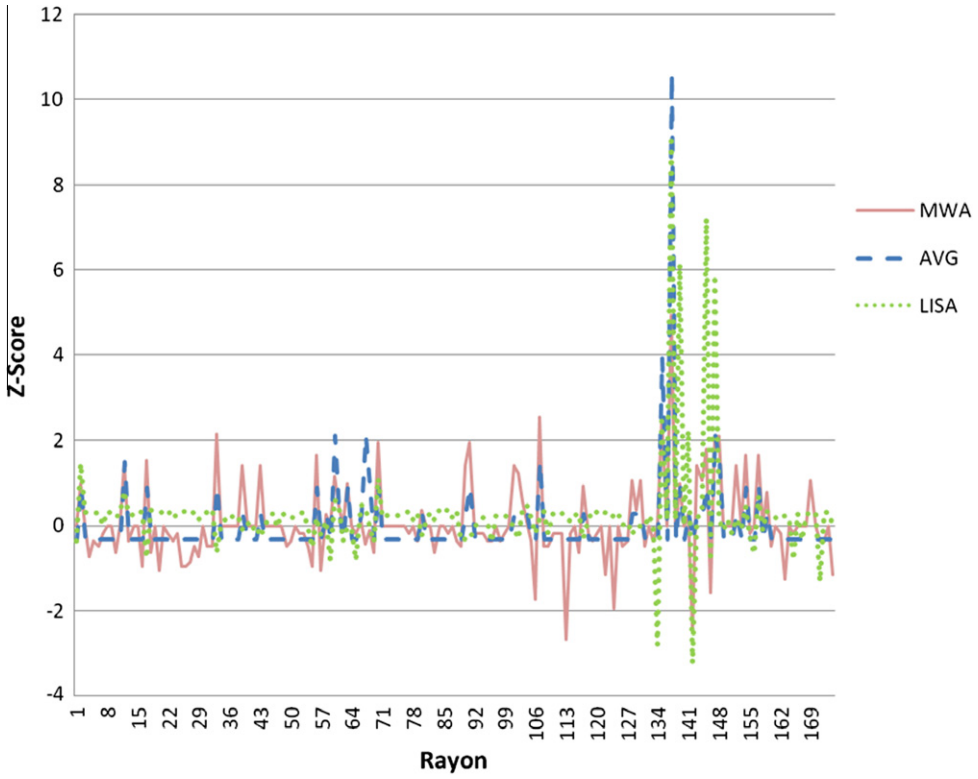

Fig. 7. Distribution of *z*-scores from a single year (1960) across all 174 rayons in Kazakhstan using three different *z*-score calculation methods. The graph shows differences in the *z*-scores for a given rayon using the MWA, AVG, and LISA methodology.

odology has, at times, an inverse relationship compared to the MWA and AVG methodologies. Although the inverse relationship between the LISA method and the other two methodologies does not hold true across all rayons, the graph does depict a general inverse trend among *z*-scores.

4. Discussion

The methods set forth in this paper introduce techniques for selecting a baseline rate of disease for use in a prospective CUSUM. While prospective methodologies have been widely used in the analysis of human disease data, their application in veterinary health is far less extensive (Kosmider et al. 2006). This is probably due to the difficulties encountered when attempting to analyse the spatial and/or temporal patterns of livestock diseases when population figures are not available, or when they are provided at too coarse a resolution. A number of techniques that have been proposed in the literature can be used to aid in the analysis of outbreak data both retrospectively and prospectively (Ord and Getis 1995; Ward and Carpenter 2000; Carpenter 2001; Kulldorff et al. 2001; Muscatello et al., 2005; Jefferson et al., 2008). However, in statistical applications such as CUSUM there is often uncertainty in deciding how to approximate a baseline rate of disease. This issue is often further exacerbated by a lack of guidance in the literature (Watkins et al., 2008).

It has been suggested that incorporating historical data from epidemic, non-epidemic, or spatial associations be used in order to construct a baseline rate of disease (Carpenter, 2002; Hutwagner et al., 2005; Sonesson, 2007; Watkins et al., 2008). In our prospective analyses of anthrax in livestock we applied three methodologies for calculating baseline in a CUSUM approach including incorporating a retrospective spatial statistic: LISA, AVG, and MWA. While not unexpected, we found that the selection of a methodology to derive a disease baseline strongly influences the identification of clusters in space and time. Furthermore, the results in this study point to the importance of exploring data from multiple analytical perspectives. This study shows that the CUSUM analysis elicited additional cluster signals not found in the original application of the Local Moran's *I* statistic on anthrax outbreaks in Kazakhstan from 1960 through 2006 performed in Kracalik (2009). That is, the prospective methodology identified emerging clusters not identified by a single retrospective cluster analysis of the entire 47-year period. These findings are similar to those presented by Rogerson (1997), for human cancer data, supporting that prospective methods can potentially identify clusters undetected by retrospective techniques.

Each of methodologies used incorporated a different baseline calculation derived from a temporal component (AVG), a spatial component (LISA), or a combination of the two (MWA). Therefore, as expected each of the methodologies produced differences in the spatial and/or temporal distribution of clusters. While the total number of spatial clusters differed between methodologies there were some consistencies in the distribution of the clusters. Emerging clusters were present in each of three methodologies in southern Kazakhstan around the city of Shymkent (Fig. 3). Persistence of clusters in this area, despite a lack of population data, suggest that future work should focus on possible ecological conditions that might promote anthrax persistence in this part of the country.

Spatial and/or temporal variations in the signaling of emerging alarm events between methods may have been a result of the differences in the magnitude and direction of their *z*-scores (Fig. 7). However, it is interesting to note in the control charts and the *z*-score distributions that despite the fact that the MWA method and AVG method share a similar max CUSUM chart their *z*-scores are a times inversely related.

In CUSUM large deviations from baseline would result in larger *z*-scores and a greater contribution to the max CUSUM, translating into fluctuations in the corresponding control chart. The LISA and AVG methodology produced a max CUSUM signal in the control chart that was higher than the MWA. A persistently higher max CUSUM in this case may have also resulted in a greater number of emerging spatial clusters. This was possibly due to the fact that for the LISA and AVG methods the *z*-scores were derived from a calculation based on the number of outbreaks in surrounding rayons. The LISA methodology was based on a Local Moran's *I* calculation that compared the local spatial clustering in a particular rayon to its contribution to the global autocorrelation, whereas the AVG values were obtained by comparing a single rayon to the outbreaks in specific time period. On the other hand, the *z*-score calculation in the MWA method is only comparing outbreak numbers in a rayon back to itself using a predefined moving-window average time frame.

The MWA method, however, is more prone to inter-annual variation in the control chart signal depending on the time frame average selected (Fig. 6). The CUSUM signal from each moving window time frame apparently undergoes a smoothing effect as the time frame of the moving window increases from one year to ten years. There is a noticeable decrease in the inter-annual variability when the 10-year average is compared to the one-year average. Shorter moving-window average time frames appear to have a higher sensitivity to small changes in outbreak numbers while a large moving average is less sensitive to the number of outbreaks in single years. Essentially a larger moving-window average will not be as likely to trigger a distinct alarm signal for a given year, but rather may have an alarm signal present due to an artifact of high cases from previous years. However, longer windows, such as a 10-year, show an elevated alarm signal initially and then decreases at year forty-two below that of the other time frames selected. Shorter moving window time frames seem to elicit lower max CUSUM values requiring the need for a smaller ARL. This leads us to the conclusion that there is a trade-off in the amount of time selected in a moving-window average; with signal sensitivity decreasing as the moving average time frame increases and the ability to trigger an alarm event increases.

The validity of the statistic is greatly dependent on knowledge of the population at risk. Rogerson (2001) suggests that long-term evaluations of disease status using a CUSUM approach may be impacted by population move-

ments. In this case there were no known estimates of population at a meaningful resolution to calculate prevalence or incidence. This is a major caveat of the data and the interpretation of the results must be looked at from an exploratory perspective. However, for the purposes of this study, the focus was on the comparison of methodologies for calculating baseline rather than extracting information on surveillance measures from exact cluster locations.

The methods introduced in this study only represent a portion of the requirements for applying a CUSUM methodology and additional parameters need to be taken into consideration. Baseline in this study differs from the selection of a critical threshold of h used as an indicator for emerging alarm signals or cluster events. Over long time periods the selection of a single critical h may become inappropriate for the data being analysed. In situations such as this it may be necessary to reset the critical h to a new more suitable threshold that takes into account the level of sensitivity required between timeliness of the statistic and the false alarm rate. Furthermore, the methods presented in this study represent only a few of the many methods for calculating a baseline rate of disease. When at all possible, estimates of the population should be used to construct appropriate determinations of a disease status in order to gain a more appropriate representation of the disease status. Yet, in instances when there are no data available on the population being studied one or more of these methods presented here may provide an exploratory tool for analysing the spatial and or temporal distribution of health events. In this case consistent clusters were detected in southern Kazakhstan from each of the three methods for estimating baseline, suggesting further efforts may be warranted for understanding the ecology of the disease in this region. In the future more testing needs to be done to look at the distribution of z-scores derived from different baseline methodologies and to also examine the data at a finer temporal resolution.

Acknowledgements

This project was funded by the United States Defense Threat Reduction Agency (DTRA) as part of the Biological Threat Reduction Program's Cooperative Biological Research Program in Kazakhstan. Y. Sansyzbayev, M.E. Hugh-Jones, A. Curtis, and T.A. Joyner assisted with the original database development. We would like to thank two anonymous reviewers for strengthening this manuscript.

References

Aikembayev Alim M, Lukhnova Larissa, Temiraliyeva Gulnara, Meka-Mechenko Tatyana, Pazylov Yerlan, Zakaryan Sarkis, et al. Historical distribution and molecular diversity of *Bacillus anthracis,* Kazakhstan. Emerg Infect Dis 2010;16(5):789–96.

Allepuz A, Lopez-Quilez A, Forte A, Fernandez G, Casal J. Spatial analysis of bovine spongiform encephalopathy in Galicia, Spain (2000–2005). Prev Vet Med 2007;79:174–85.

Anselin Luc. Local indicators of spatial autocorrelation. Geographic Analalysis 1995;27:93–115.

Anselin Luc, Syabri I, Kho Y. GeoDa: an introduction to spatial data analysis. Geographic Analalysis 2006;38:5–22.

Beyer H.L. Hawth's analysis tools for ArcGIS. Available at <http://www.spatialecology.com/htools>; 2004.

Blackburn Jason K, McNyset Kristina M, Curtis Andrew, Hugh-Jones Martin E. Modeling the geographic distribution of *Bacillus anthracis,* the causative agent of anthrax disease, for the contiguous United States using predictive ecologic niche modeling. Am J Trop Med Hyg 2007;77(6):1103–10.

Buckeridge David L, Owens Douglas K, Switzer Paul, Frank John, Musen Mark A. Evaluating detection of an inhalation anthrax outbreak. Emerg Infect Dis 2006;12(12):1942–9.

Carpenter TE, Hird David W, Snipes Kurt P. A space-time investigation of the epidemiology of fowl cholera. Prev Vet Med 1996;28:159–63.

Carpenter Tim E. Methods to investigate spatial and temporal clustering in veterinary epidemiology. Prev Vet Med 2001;48:303–20.

Carpenter TE. Evaluation and extension of the cusum technique with an application to *Salmonella* surveillance. J Vet Diagn Invest 2002;14:211–8.

Clegg SB, Turnbull PCB, Foggin CM, Lindeque PM. Massive outbreak of anthrax in wildlife in the Malilangwe wildlife reserve, Zimbabwe. Vet Rec 2007;160:113–8.

Doherr MG, Hett AR, Rufenacht J, Zurbriggen A, Heim D. Geographical clustering of cases of bovine spongiform encephalopathy (BSE) born in Switzerland after the feed ban. Vet Rec 2002;151:456–72.

Augusta D' Orazi, Mignemi Manuela, Geraci Francesco, Vullo Angela, Gesaro Michele Di, Vullo Stefano, et al. Spatial distribution of brucellosis in sheep and goats in Sicily from 2001 to 2005. Vet Ital 2007;43(3):541–8.

Dragon DC, Elkin BT, Nishi JS, Ellsworth TR. A review of anthrax in Canada and implications for research on the disease in northern bison. J Appl Microbiol 1999;87:208–13.

Gilbert M, Mitchell A, Bourn D, Mawdsley J, Clifton-hadley R, Wint WC. Attle movements and bovine tuberculosis in Great Britian. Nature 2005;435:491–6.

Heres, Lourens, Brus Dick J, Hagenaars Thomas J. Spatial analysis of BSE cases in The Netherlands. BMC Vet Res 2008;4(21).

Himsworth Chelsea G, Argue Connie K. Anthrax in Saskatchewan 2006: an outbreak overview. Can Vet J 2008;49(3):235–7.

Hoar BR, Chomel BB, Rolfe DL, Chang CC, Fritz CL, Sacks BN, et al. Spatial analysis of *Yersinia pestis* and *Bartonella vinsonii* subsp. *Berkhoffii* seroprevalence in California coyotes (*Canis latrans*). Prev Vet Med 2003;56:299–311.

Hohle Michael, Paul Michaela, Held Leonhard. Statistical approaches to the monitoring and surveillance of infectious diseases fro veterinary public health. Prev Vet Med 2009;91:2–10.

Hugh-Jones M. 1996–97 global anthrax report. J Appl Microbiol 1999;87:189–91.

Hutwagner LC, Thompson WW, Seeman GM, Treadwell T. A simulation model for assessing aberration detection methods used in public health surveillance for systems with limited baselines. Statistic in Med 2005;24(4):543–50.

Jefferson Henry, Dupuy Bruce, Chaudet Herve, Texier Gaetan, Green Andrew, Barnish Guy, et al. Evaluation of a syndromic surveillance for the early detection of outbreaks among military personnel in a tropical country. J Pub Health. 2008;30(4):375–83.

Jernigan Daniel B, Raghunathan Pratima L, Bell Beth P, Brechner Ross, Bresnitz Eddy A, Butler Jay C, et alThe National Anthrax Investigation Team. Investigation of bioterrorism-related anthrax, United States, 2001: epidemiologic findings. Emerg Infect Dis 2002;8(10):1019–28.

Joyner, Andre Timothy, Lukhnova Larissa, Pazilov Yerlan, Temiralyeva Gulnara, Hugh-Jones Martin E, et al. Modeling the potential distribution of Bacillus anthracis under multiple climate change scenarios for Kazalhstan. PLoS ONE 2010;5(3).

Kleinman KP, Abrams AM, Kulldorff M, Platt R. A model-adjusted space-time scan statistic with an application to syndromic surveillance. Epidemiol Infect 2005;133:409–19.

Kosmider R, Kelly L, Evans S, Gettinby G. A statistical system for detecting salmonella outbreaks in British livestock. Epidemiol Infect 2006;134:952–60.

Knox G. The detection of space-time interactions. Applied Statistics 1964;13:25–9.

Kracalik, Ian T. Spatial and temporal analyses of anthrax: an exploratory retrospective and prospective examination of outbreaks in Kazakhstan. M.A. thesis. California State University, Fullerton, 2009. Proquest dissertations and thesis. Web. April 2; 2010.

Kulldorff Martin. A spatial scan statistic. Commun Stat-theor Meth 1997;26(6):1481–96.

Kulldorff Martin. Prospective time periodic geographical disease surveillance using a scan statistic. J R Statist Soc A 2001;164:61–72.

Lee, Gyoungju, Ikuho Yamada, Peter Rogerson. GeoSurveillance 1.1 User's Manual; 2007.

Mantel N. The detection of disease clustering and a generalized regression approach. Cancer Res 1967;27:209–20.
Meselson Matthew, Guillemin Jeanne, Hugh-Jones Martin, Langmuir Alexander, Popova Ilona, Shelokov Alexis, et al. The Sverdlovsk anthrax outbreak of 1979. Science 1994;266:1202–8.
Mongoh Ndiva M, Dyer NW, Stoltenow CL, Hearne R, Khaitsa ML. A review of management practices for the control of anthrax in animals: the 2005 anthrax epizootic in North Dakota-case study. Zoonoses Pub Health 2008;55:279–90.
Moran PA. Notes on continuous stochastic phenomena. Biometrika 1950;37:17–23.
Mostashari Farzad, Kulldorff Martin, Hartman Jessica J, Miller James R, Kulasekera Varuni. Dead bird clusters as an early warning system for West Nile virus activity. Emerg Infect Dis 2003;6(9): 641–6.
Muscatello, David J, Churches Tim, Kaldor Jill, Zheng Wei, Chiu Clayton, et al. An automated, broad-based, near real-time public health surveillance system using presentations to hospital emergency departments in New South Wales, Australia. BMC Pub Health 2005; 5:141.
Norstrom, Madelain, Pfeiffer Dirk U, Jarp Jorun. A space-time cluster investigation of an outbreak of acute respiratory disease in Norwegian cattle herds. Prev Vet Med 2000;47:107–19.
Ord Jk, Getis Arthur. Local spatial autocorrelation statistics: distributional issues and an application. Geographic Analalysis 1995;4(27): 286–306.
Page ES. Continuous inspection schemes. Biometrika 1954;41:100–15.
Parkinson Robert, Rajic Andrijana, Jenson Chris. Investigation of an anthrax outbreak in Alberta in 1999 using a geographic information system. Can Vet J. 2003;44:315–8.
Rogerson Peter. Surveillance systems for monitoring the development of spatial patterns. Stat Med 1997;16:2081–93.
Rogerson Peter. Monitoring point patterns for the development of space-time clusters. J R Statist Soc A 2001;164:87–96.
Rogerson Peter, Yamada Ikuho. GeoSurveillance. Software for Monitoring Change in Geographic Patterns. Lect Notes Comput Sci 2004;3044: 1096–106.
Sheridan Hazel A, McGrath Guy, White Paul, Fallon Richard, Shoukri Mohamed M, Wayne Martin S. A temporal-spatial analysis of bovine spongiform encephalopathy in Irish cattle herds, from 1960 to 2000. Can J Vet Res 2005;69:19–25.
Smith KL, De Vos V, Bryden HB, Hugh-Jones ME, Klevytska A, Price LB, et al. Meso-scale ecology of anthrax in southern Africa: a pilot study of diversity and clustering. J Appl Microbiol 1999;87:204–7.
Smith KL, De Vos V, Bryden HB, Hugh-Jones ME, Keim P. *Bacillus anthracis* diversity in Kruger national park. J Clin Microbiol 2000;38(10): 3780–4.
Sonesson Christian, Bock David. A review and discussion of prospective statistical surveillance in public health. J R Statist Soc A 2003;166:5–21.
Sonesson Christian. A CUSUM framework for detection of space-time disease clusters using scan statistics. Stat Med 2007;26:4770–89.
Tango Toshiro. A test for spatial disease clustering adjusted for multiple testing. Stat Med 2000;19(2):191–204.
Tango, Toshiro, Kunihiko Takashi. A flexibility shaped scan for detecting clusters. Int J Health Geographics 2005;11(4).
Turner AJ, Galvin JW, Rubira RJ, Miller GT. Anthrax explodes in an Australian summer. J Appl Microbiol 1999;87:196–9.
Van Ert MN, Easterday WR, Huynh LY, Okinaka RT, Hugh-Jones ME, Ravel J. Global genetic population structure of *Bacillus anthracis*. PLoS ONE 2007;2(5).
Van Ness Glenn B. Ecology of anthrax. Science 1971;172:1303–6.
Wallenstein S. A test for detection of clustering over time. Am J Epidemiol 1980;104:576–84.
Ward MP, Carpenter TE, Johnson SJ. Spatial analysis of seroconversion of sentinel cattle to bluetounge viruses in Queensland. Aust Vet J 1996;2(74):128–31.
Ward Michael P, Carpenter Tim E. Techniques for analysis of disease clustering in space and in time in veterinary epidemiology. Prev Vet Med 2000;45:257–84.
Watkins S, Rochelle E., Eagleson S, Veenendaal B, Wright G, Plant A. Applying cusum-based methods for the detection of outbreaks of Ross River virus disease in Western Australia. BMC Med Inform Decision Making 2008;8(37).
Webb Glenn F, Blaser Martin J. Mailborne transmission of anthrax: modeling and implications. PNAS 2002;99(10):7027–32.
Wilkening, Dean A. Sverdlovsk revisited: modeling human inhalation anthrax. PNAS 2006;103(20):7589–94.
Williams EH, Smith PG, Day NE, Gesser A, Ellice J, Tukei P. Space time clustering of Burkitt's lymphoma in the West Nile district of Uganda. Br J Cancer 1978;37:109–22.
Woodall William H, Brooke Marshall J, Joner Michael D, Fraker Shannon E, Abdel-Salam Abdel-Salam G. On the use and evaluation of prospective scan methods for health-related surveillance. J Royal Stat Soc A 2008;171:223–37.
Woods Christopher W, Ospanov Kenes, Myrzabekov Akylbek, Favorov Michael, Pliykaytis Brain, Ashford David A. Risk factors for human anthrax among contacts of anthrax-infected livestock in Kazakhstan. Am J Trop Med Hyg 2004;71:48–52.