

This article was downloaded by: [Marianne Frisé]n

On: 01 December 2012, At: 23:01

Publisher: Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



IIE Transactions

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/uiie20>

Spatial outbreak detection based on inference principles for multivariate surveillance

Marianne Frisé^a

^a Department of Economics, University of Gothenburg, Gothenburg, SE, 40530, Sweden

Accepted author version posted online: 28 Nov 2012.

To cite this article: Marianne Frisé (2012): Spatial outbreak detection based on inference principles for multivariate surveillance, IIE Transactions, DOI:10.1080/0740817X.2012.748995

To link to this article: <http://dx.doi.org/10.1080/0740817X.2012.748995>



Disclaimer: This is a version of an unedited manuscript that has been accepted for publication. As a service to authors and researchers we are providing this version of the accepted manuscript (AM). Copyediting, typesetting, and review of the resulting proof will be undertaken on this manuscript before final publication of the Version of Record (VoR). During production and pre-press, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal relate to this version also.

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Spatial outbreak detection based on inference principles for multivariate surveillance

MARIANNE FRISÉN

Department of Economics, University of Gothenburg, Gothenburg, SE40530, Sweden

E-mail: marianne.frisen@statistics.gu.se

Spatial surveillance is a special case of multivariate surveillance. Thus, in this review of spatial outbreak methods, the relation to general multivariate surveillance approaches is discussed. Different outbreak models are useful for different aims. First, it makes a great difference which spreading pattern is of main interest to detect. We will discuss methods for the detection of: 1) Spatial clusters of increased incidence, 2) Increased incidence at only one (unknown) location, 3) Simultaneous increase at all locations, 4) Outbreaks with a time lag between the onsets in different regions. The sufficient reduction was used to find likelihood ratio methods for some of these spreading patterns. Secondly, an alternative to the common assumption of step change to an increased incidence level is suggested. The assumption is sometimes too restrictive and errors in the estimation of the baseline have great influence. Instead, a robust nonparametric model is suggested. The seasonal influenza in Sweden is used as an example. Here, the outbreak was characterized by a monotonic increase following the constant non-epidemic level. The semiparametric generalized likelihood ratio surveillance method used for this application is described. Thirdly, evaluation metrics are discussed. Evaluation in spatial and other multivariate surveillance requires special concern.

Keywords: Monitoring, Influenza, Sufficiency, Semiparametric, Generalized likelihood, Timeliness, Predicted value.

1. Introduction

The first versions of sequential surveillance were developed around 1930 by Walter A. Shewhart. He introduced control charts for industrial applications. The area of industrial quality control dominated the development of surveillance theory for a long time. Although industrial applications are still important, many new applications have come into focus. Emerging needs in other areas and the availability of powerful computing resources have encouraged the development of more advanced and efficient methods. The threats of bioterrorism and new contagious diseases have been important reasons behind the increased research activity in the theory of surveillance. Kaufmann et al. (1997) stated that the delay of one day in the detection of and response to an epidemic due to a bioterrorist attack could result in the loss of thousands of lives and millions of dollars. Even if such severe threats should become less pronounced in the future, there are now systems for the collection of public health information. Statistical surveillance systems will be expected to take care of such information for monitoring different kinds of health care services. Reviews on surveillance in public health are given by e.g. Sonesson and Bock (2003), Tsui *et al.* (2008), and Woodall *et al.* (2010).

Even though methods have been developed under different scientific cultures, inferential similarities can be identified. The aim of on-line surveillance is to give an alert signal as soon as possible after an important change has occurred. Overviews of the inferential issues in surveillance are given for example by Lai (1995), in the text book by Ryan (2000), and in the discussion in connection with the paper by Frisén (2009). Separate lines of research in theory, engineering, finance, and public health have come together in recent years, and progress in one area has been used in others. The theory of statistical surveillance will be further developed in

response to the demands of applications in various fields. A cross-fertilization back to the applications could then be expected.

The detection of outbreaks of epidemiological diseases is an important area of on-line surveillance. In recent years, there have been several events that highlight the importance of outbreak detection. The outbreaks of new kinds of influenza (SARS, avian flu, and H1N1) are such recent examples. By obvious reason there is a lack of knowledge on new diseases. Hopefully, methods designed and evaluated for known periodical diseases or symptoms, such as the yearly influenza epidemic, will be useful to detect such new and unexpected diseases. An early detection of the onset of an outbreak of a known disease is also important in itself. It is useful in order for health authorities to act timely and also for the planning of health care resources. The surveillance of seasonal influenza outbreaks in Sweden will be used as an example throughout the paper. Epidemics, such as influenza, are for several reasons very costly to society, and it is therefore of great value to monitor the epidemic period in order to properly allocate medical resources (Andersson *et al.* (2008b)). The incidences of different diseases and symptoms are monitored by international, national, and local authorities to detect outbreaks of infectious diseases. Today many countries have advanced routine systems to detect outbreaks of diseases or symptoms. Hulth *et al.* (2010) describe the practical usage of computer-supported outbreak detection in five European countries.

Several different definitions of an outbreak are used, explicitly or implicitly, in the literature. This is obvious in the broad overviews given by Shmueli and Burkom (2010) and Unkel *et al.* (2012). The present statistical review is based on different outbreak models and their correspondence to methods and evaluations. The choice of method and evaluation procedure

depends on which definition is used. Therefore, it is important to state the definition and the aim explicitly. Different methods are optimal under different conditions. The methods can be seen as complements to each other.

Seasonality, autocorrelation, and other complexities can be important for some applications. Often this can be handled by some regression technique to give standardized residuals. The data can often be explained by different models, and the aim is important for the choice of model. For example, Paul *et al.* (2008) and Held (2010) include outbreaks in the model and make outbreak detection by prediction. For influenza in Sweden Andersson *et al.* (2008a) concluded that simple models were useful for the outbreak detection purpose. Although complex stochastic models can be important in some situations, they will not be discussed in any detail in this review.

Sometimes information on the incidence at different locations is available. It is then desirable to generalize the univariate method to utilize this information. Two examples of reviews on spatial public health surveillance are Lawson and Kleinman (2005) and Tsui *et al.* (2011). Spatial surveillance is a special case of multivariate surveillance, as pointed out for example by Sonesson and Frisé (2005) and Jørgensen *et al.* (2008).

In Section 2, we will discuss different aims in outbreak detection and corresponding methods for univariate surveillance. In Section 3, we will discuss general approaches of how to construct multivariate surveillance from univariate surveillance. The discussion on outbreak models in Section 2 and the review on multivariate surveillance in Section 3 are brought together in Section 4. Here, we adapt the technique for multivariate surveillance to different aims of spatial outbreak detection. In Section 5, we discuss evaluation metrics. Concluding remarks are given in the final section.

2. Univariate outbreak detection

Different definitions of an outbreak are used, explicitly or implicitly, in the literature. To make the differences explicit, we start by describing univariate outbreak detection. This is a prerequisite for describing different aims of spatial outbreak detection in Section 4.

At each time point, t , a new observation is made on a process \mathbf{Y} . We state the model for discrete time. Weekly data are available for influenza in Sweden. We want to detect the change from one state, non-epidemic, to another state, outbreak, as soon as possible after it has occurred, in order to give warnings and take corrective actions. Different univariate surveillance methods can be used depending on the optimality criteria. Frisé (2003) describes how different ways of summarizing the partial likelihoods (summation, maximum, etc) correspond to different optimality criteria, and optimal methods.

2.1 *Step change from an estimated baseline*

The most commonly used approach to outbreak detection aims to detect a step change from a parametric model for the non-epidemic periods. A signal is given as soon as one observation exceeds a threshold, usually a 95% prediction interval. A very successful method of this kind is the England-Wales method (see Farrington and Beale (1993), Farrington and Andrews (2004), Farrington *et al.* (1996), Farrington and Beale (1993), Farrington and Beale (1998), and Farrington and Andrews (2004). An alarm is triggered when the last observation exceeds a limit. With respect to how the information from different time points is summarized, this is a variant of the Shewhart surveillance method.

Likelihood based methods may give optimal weighting of the different observations. In Le Strat and Carrat (1999), a Hidden Markov model (HMM) is suggested, which allows for switching between states with different statistical properties: the non-epidemic state, with low incidence rates, and the epidemic state, characterized by an increased incidence. The cyclical regression suggested in Serfling (1963) was used to model the seasonal effect. It was found that the seasonality had a period of 52 weeks. Surveillance methods for the detection of a step increase in a Poisson distribution in public health were compared by Han *et al.* (2009).

Höhle and Paul (2008) use the CUSUM method to summarize the deviations from a parametric model and discuss the effect of estimating unknown parameters. In Andersson, *et al.* (2008a), it is concluded that parametric methods are not suitable when the parameters describing the incidence curve vary much from year to year, as is the case with influenza in Sweden.

2.2 *Gradually increasing incidence*

Outbreaks are always characterized by an increased incidence. However, the increase is seldom in one step. We will here consider an outbreak defined as a change from an unknown baseline to a monotonically increasing incidence, as suggested by Friséen and Andersson (2009). The suggested method aims at detecting a change from a constant level to an increasing regression, i.e. a change in monotonicity. In Andersson, *et al.* (2008a) on Swedish influenza data it was suggested that the incidence could be modeled using a Poisson distribution. The parameter $\lambda(t)$ of the distribution at time t has a constant value λ_0 before the outbreak but increases with time after the onset of the outbreak. We will use τ to denote the unknown time of the onset. For some applications, τ is naturally regarded as a parameter and for others as a stochastic variable.

We have

$$\lambda(t) = \begin{cases} \lambda_0, & t < \tau \\ \lambda_{t-\tau+1}, & t \geq \tau \end{cases},$$

with $\lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_s$. The aim at decision time s is to determine whether or not the outbreak has started yet, thus if $\tau \leq s$ or $\tau > s$. The state at the outbreak is characterized by a monotonically increasing expected incidence.

Generalized likelihood ratios (GLR) can be used by replacing the parameters with the maximum likelihood estimators under the monotonicity restrictions as derived in Frisé *et al.* (2010a) for the exponential family. The method is semiparametric since the distribution is parametric, but the regression is nonparametric. A user-friendly computer program can be downloaded at www.statistics.gu.se/surveillance. The method is also available in the R package *Surveillance*, described in Höhle (2010) and available on CRAN, and in the open JAVA package *CASE* described in Cakici *et al.* (2010).

The method was applied to the incidence of influenza in Sweden as a whole by Frisé *et al.* (2009) and compared favorably to other methods. In Section 4.4.2 we will adapt this method to spatial surveillance.

2.3 Varying sample size

Recently, monitoring of incidence rates of the Poisson distribution when sample size varies over time has drawn interest and this is of interest for outbreak detection. Most surveillance theory is developed for the case there we have a change from one distribution to another one at the change time τ . This is not the case here, but at time τ there is a change between states which each can

consist of several distributions due to the different sample sizes. Thus, not much of earlier optimality theory can be used. It adds to the complexity that the properties of methods depend heavily on how the sample sizes evolve with time and that different evaluation criteria put emphasis to changes which occur early or late (see Section 5 for evaluation criteria). An additional complexity is that the discreteness of the Poisson distribution makes the distribution of the relative incidence somewhat complex. Some of the most recent achievements will now be briefly described with respect to those issues.

Ryan and Woodall (2010) studied the case when the sample size varies randomly. They adjusted the incidences by the observed sample sizes and found that the CUSUM method has a better steady state average run length than the EWMA method. They also suggested a modified EWMA method by adding a lower reflecting barrier. Mei *et al.* (2011) examined several CUSUM variants. In one, an adaptive threshold is used. For the application of male thyroid cancer, which is used for illustration in this and many other papers, it was found that a logistic curve with an asymptote fits the sample sizes well. The estimated curve and variants with a more rapid increase or with a decrease are used in this and the following two papers. It was found that the expected delay depends on τ in opposite ways for increasing and decreasing sample sizes for the different methods. In standard settings, the CUSUM method is minimax optimal. Also here, the properties for the worst value of τ are studied. Asymptotic minimax optimality results are achieved under certain conditions. Shu *et al.* (2011) recommended extra weight to later observations by downgrading the previous evidence before combining it with the most recent evidence in the recursive formula of a CUSUM statistic. For the increasing logistic sample size model, it was demonstrated that the extra weight improved the ARL^1 . Zhou *et al.* (2012)

suggested a weighted likelihood-based EWMA method. They derived asymptotic bounds for ARL^0 and ARL^1 . A simulation study is made for eight methods and six different scenarios of evolving sample sizes. Several characteristics of the run length distributions for $\tau=1$, $\tau=20$ and $\tau=\infty$ (in-control) are presented. It is found that although some earlier suggested methods have good detection ability, they may have a high probability of early false alarm. The run length standard deviation is lower for the proposed method. The balance over time between false alarms and detection ability is illustrated by figures of the run length distributions. Another way to evaluate the balance is by the predictive value of an alarm (see Section 5.3).

The case of varying sample sizes illustrates the importance of the choice of evaluation criteria, which is the topic of Section 5 below. Simple summarizing criteria may not be enough in complex situations such as varying sample size.

3. Multivariate surveillance

Spatial surveillance is multivariate. After the section on univariate outbreak detection, we will now review general approaches to adapting univariate methods for multivariate problems. This is a prerequisite for Section 4 on spatial outbreak detection. General reviews on multivariate surveillance methods can be found for example in Basseville and Nikiforov (1993) and Bersimis *et al.* (2007). Jiang and Tsui (2008) base their framework on the relationship between control charts and statistical hypothesis testing. Frisén (2010) bases the overview on inference principles and emphasizes the times of the different changes. Tsui *et al.* (2012) propose a general framework for spatial and spatiotemporal surveillance based on likelihood ratio statistics.

Reviews on issues of multivariate surveillance in public health are given by e.g. Marshall *et al.* (2004), Burkom *et al.* (2005), and Rolka *et al.* (2007).

The process under surveillance is a p -variate vector, denoted by $\mathbf{Y} = \{\mathbf{Y}(t), t = 1, 2, \dots\}$, where $\mathbf{Y}(t) = \{Y_1(t), Y_2(t), \dots, Y_p(t)\}$. The components of the vector here represent the incidence of a disease at p different locations. At decision time s , we base the decision on the available information, $\mathbf{Y}^s = \{\mathbf{Y}(1), \mathbf{Y}(2) \dots \mathbf{Y}(s)\}$.

There are several approaches to multivariate surveillance. The most commonly used approach is the reduction to one scalar statistic for each time. This will be described in Section 3.1. Another common approach, described in Section 3.2, is to use several univariate systems in parallel. There are also other approaches but these two are used in many of the spatial outbreak methods described in Section 4.

3.1 *Reduction to one scalar statistic for each time*

It is always sensible to reduce dimension in multivariate problems provided that this does not reduce the important information. The most far-reaching reduction is the reduction to a scalar for each time. This is a common way to handle multivariate surveillance. In multivariate surveillance the observations at each time point consist of a vector. This vector can be transformed to a scalar statistic for that time point. The accumulation over time can then be made by a univariate surveillance method.

One natural reduction when dealing with multivariate normal variables is to use the Hotelling T^2 statistic suggested by Hotelling (1947). This statistic is defined as

$$T^2(t) = (\mathbf{Y}(t) - \boldsymbol{\mu}_0(t))^T \mathbf{S}_{\mathbf{Y}(t)}^{-1} (\mathbf{Y}(t) - \boldsymbol{\mu}_0(t)),$$

where $\mathbf{S}_{Y(t)}$ is the sample covariance matrix. Originally, the Hotelling T^2 statistic was used in a Shewhart approach, and this is sometimes referred to as the Hotelling T^2 control chart.

One example of scalar accumulation is when, for each time point, a statistic representing the important aspects of the spatial pattern is constructed from a purely spatial analysis. This statistic can then be used in a surveillance method. The reduction to a univariate variable can be followed by univariate monitoring of any kind. For examples of how this is used in cluster detection, see Section 4.1.

For the influenza incidence in a country, like Sweden, it is natural to use the incidence for the country as a whole for each time point, as in Frisé and Andersson (2009). Using the sum over locations means that no regional information is utilized. Other reductions than the sum are more efficient, as is seen in Section 4.4.

3.2 *Parallel outbreak detection*

Another stepwise way to handle multivariate surveillance is to first reduce the data to a surveillance statistic for each variable. This is, in a sense, an opposite technique to that in Section 3.1. Both methods are very common. They have different optimality properties. By the parallel approach, each process is monitored separately and an overall alarm is called if some condition is fulfilled. The most common overall alarm condition is that one of the surveillance statistics indicates an alarm. References to methods using this approach for spatial surveillance will be given in Sections 4.1 and 4.2.

4. Methods for spatial outbreak detection

Different applications require different models and methods. Some structure is necessary in order to get detection power. Otherwise, the power will be diluted. On the other hand, false assumptions may lead to misleading results. We will discuss methods for the detection of the following kinds of outbreaks: 1) Spatial clusters of increased incidence, 2) Increased incidence at only one (unknown) location, 3) Simultaneous increase at all locations, 4) Outbreaks with a time lag between the onsets in different regions.

4.1 *Clusters*

Unusual aggregations and spreading patterns in space can give information on health hazards. Cluster detection methods aim to detect increased incidences in regions that are close in space. Thus, an outbreak is characterized by a non-homogenous distribution in space, and the aim is to detect a departure from spatial randomness. Different assumptions are made on the spatial process and different kinds of clusters are of interest.

There is much literature on cluster detection. Older literature considers retrospective hypothesis testing. Simple retrospective tests of the space-time interaction can be based on the Knox statistic described by Knox (1964) where the proportion of observations close in time and space is recorded. Many other statistics for retrospective tests of clustering have also been suggested based on different statistical models.

Prospective surveillance will now be discussed with emphasis on the inference approaches. Most methods are stepwise. Either the spatial pattern is summarized first and then the time-

development, as described in Sections 3.1, or vice versa, as in Section 3.2. We will now give examples of how specific methods use the different approaches.

By the summarizing technique described in Section 3.1 we first compute a statistic measuring the clustering tendency for each time point and then summarize over time by some univariate surveillance method. Rogerson (2001) suggested that the local Knox statistic could be used for each time, and the information was accumulated by the univariate CUSUM method. The method was criticized by Marshall *et al.* (2007) by distributional arguments and by the ARL performance. Zhou and Lawson (2008) characterized the spatial pattern by a Bayesian model for each time, and the statistic was then monitored by the EWMA method.

The general multivariate technique of parallel surveillance (described in Section 3.2) has also been used. The timely detection of clusters in space within the framework of Bayesian hierarchical Poisson models is treated by Corberan and Lawson (2011). They summarize increasing tendencies for each region in parallel by an advanced estimation technique and comparisons between successive values. Then, they summarize over regions by noting for how many of the regions a threshold is exceeded. Raubertas (1989) uses a first step by smoothing the spatial data and then makes a parallel CUSUM for each region and makes a general alarm as soon as there is an alarm for any of the regions.

The many papers by Kulldorff (for example Kulldorff (2001)) and the computer programs (see Kulldorff (2010)) on scans for areas with high incidence of a disease have been important for the development of spatial surveillance. Shmueli and Burkom (2010) state that it is the most widely used method for detecting the emergence of localized disease clusters. By the SaTScan method (Kulldorff (2001)) possible circular regions are scanned for increased incidence. The

method is based on likelihood ratios and developed within the framework of statistical tests. It has later been developed in different directions (for example elliptical regions).

Sonesson (2007) showed that the scan method by Kulldorff can be expressed as a CUSUM method since the partial likelihood ratios are summarized by maximization over possible outbreak times. The important difference to CUSUM, as it is used in surveillance, is that Kulldorff conditions on the total number of events for the time periods (between the possible change point and the decision time). This condition makes the method fit within the hypothesis testing framework and p-values can be calculated. It also gives conditional levels of incidence.

Sonesson (2007) suggested a new method based on the parallel CUSUM by Raubertas (1989) but with circular clusters, as in the method by Kulldorff. He also suggested that other surveillance methods than CUSUM should be used for summarizing the partial likelihoods over time. Sonesson (2007) examined the surveillance properties of the methods by Raubertas, Kulldorff, and himself by simulations for some simple situations and by application to the detection of tularemia clusters in Sweden. This seems to be the first evaluation of the widely used method by Kulldorff in a surveillance context.

Assunção and Correab (2009) make advanced use of the spatial and temporal likelihood expressions. The statistic for the spatial interaction is a likelihood ratio based on a fixed shift within the cluster. A fixed radius for the cluster and estimated purely spatial and purely temporal functions are used. The summarizing over time is made by the Shiryaev Roberts approach.

Spatial surveillance is often seen as identical with cluster detection and there is plenty of literature on cluster detection. However, there are also other spatial abnormalities which demand detection methods and some will now be described.

4.2 *Outbreak at only one location out of many*

Sometimes it can be assumed that the change appears in only one out of many locations. Tartakovsky and Veeravalli (2004) described the parallel surveillance of a multichannel system where one (and only one) channel may have a change. A parallel approach for each location may be assumed to be the best solution and is usually used for such problems. However, the properties of the method depend on how the results from different parallel surveillance methods are combined. The most commonly used method is the union intersection method, where an alarm is given as soon as there is an alarm for any of the locations. As regards the combination of p-values in hypothesis testing, it was demonstrated by Frisé (1974) that the choice of method for combination depends on the size of the change. This indicates that no truly simple optimality result can be expected.

Sonesson (2007) compared the detection ability of cluster detection methods for some different cluster sizes and found that the SaTScan method by Kulldorff (2001) was the best choice when the cluster was constituted by only one position. Jiang *et al.* (2011) examined the effect of radius in cluster detection and concluded that a correspondence between actual radius and that of the method has a large effect.

If only a few locations out of many are affected, similar properties as for only one location will occur.

4.3 *Simultaneous change at all locations*

The case where the onsets are simultaneous, that is $\tau_i = \tau$ for $i=1, \dots, p$, is of special interest. In most papers on multivariate surveillance it is implicitly assumed that the changes occur

simultaneously. Multivariate surveillance is often evaluated by ARL¹. The consequences of this will be further discussed in Section 5.2.

Wessman (1998) and Frisé *et al.* (2011) demonstrated that if all processes have the same change points, then the univariate vector of partial likelihood ratios $\{L(s,t), t=1,\dots,s\}$, for $\tau < s$ against $\tau = s$, is sufficient for the sequence of distributional families. Thus, in order to monitor a simultaneous and fully specified change, it is possible to construct a univariate surveillance procedure based on the sufficient sequence of likelihood ratios. Zhou *et al.* (2010) used this result for the simultaneous shifts of mean and variance in a normal distribution.

4.4 *Outbreaks with time lags*

The relation between different locations is important in the monitoring of outbreaks. In some situations the time lag in the onset of the outbreak in different regions is rather well known. The spreading pattern of influenza in Sweden is described in Section 4.4.1. Another example where a time lag is relevant is when there is an early but rough indicator which may be combined with a later and more accurate one. In Hulth *et al.* (2009) and Ginsberg *et al.* (2009) it was shown that data of search patterns on the Internet could be used as an early proxy for influenza incidence. In Ginsberg *et al.* (2009) it was found that the lag between information by Google and traditional CDC-data was about two weeks. The method described in Section 4.4.2 may possibly be useful also for such situations, where the lag is in the reporting rather than in the onset of the outbreak at the various locations. However, the motive for the suggested method was the spreading pattern.

4.4.1 *The spatial spreading pattern of influenza in Sweden*

Data on epidemic diseases are collected by national and international institutes, for example the European Centre for Disease Prevention and Control in Europe and the Centers for Disease Control and Prevention in the US. In Sweden, data on infectious diseases are collected by the Swedish Institute for Infectious Disease Control, SMI. The monitoring of influenza in Sweden is mostly based on the sum of reports of all Swedish laboratory diagnoses of influenza (LDI). Andersson, *et al.* (2008a) and Andersson, *et al.* (2008b) give descriptions of the collection and properties of these data for the whole country.

The spatial pattern of how a disease spreads between regions is important. Spatial clustering of adverse health events is often discussed. However, in some situations, such as in the case of influenza in Sweden, the outbreak pattern is not characterized by clustering.

Schiöler (2011) searched for geographical patterns in the spread of influenza in Sweden. A spread from south to north or from west to east had been discussed. No such pattern was found. Instead it was found that influenza epidemics tend to start in the larger cities and then spread to the smaller ones. Data from areas classified as Metropolitan areas generally showed an earlier outbreak than those from the Locality areas. The Metropolitan areas have major international airports nearby, and commuting to other countries is common. This is a possible explanation for the early start of the influenza season in these areas. This explanation is also in accordance with the results of Crepey and Barthelemy (2007), who investigated the relation between traveling and influenza in the US as well as in France and found a stable impact.

The time difference in the onset of the influenza outbreak between the group of Metropolitan areas and the group of Locality areas was about one week. This information was used to increase the efficiency of the surveillance system, as described in the next section.

4.4.2 *Spatial outbreak surveillance based on sufficient reduction*

Knowledge about the relation between the times of the onsets of the outbreaks is essential. In Frisé *et al.* (2010b) it was demonstrated that the relation between the change points of the different processes is very important, since it affects the properties of different surveillance methods in different ways. Thus, any knowledge on the change points should be utilized. Different methods are suitable for different relations. Without any information about this relation it is not possible to derive a good method. The time τ_i of the onset of the outbreak of process Y_i may not be the same for all $i=1,\dots,p$. For notational convenience we order the processes according to the order of onset, so that $\tau_1 \leq \dots \leq \tau_p$. For some purposes, it may be of interest to monitor each location separately. However, the aim here, as often, is to detect an outbreak in any of the processes, which means that we aim at detecting the first one.

Due to the complexity of multivariate problems we will now examine the possibilities, in the case of outbreak detection, to minimize the complexity without any loss of information. A sufficient reduction will not reduce the information and still allows a joint solution to the full surveillance problem.

We assume that the different processes are identically distributed except for the time of the onset. and that the distributions of the processes all belong to the one-parameter exponential family. For the influenza in Sweden, the Poisson distribution was demonstrated to be relevant by Andersson, *et al.* (2008a).

Frisén, *et al.* (2011) studied the case of step changes in the general one-parameter exponential family (including the Poisson distribution) and identified the sufficient reduction for some cases. Modification to gradual change is of interest for influenza outbreaks. In Section 2.2 we studied the univariate case where a process Y_i increases monotonically from the onset of the outbreak τ_i and onwards. In Schiöler and Frisé (2012) it was demonstrated that a sufficient reduction to a univariate statistic exists for the situation with known time lags between the outbreaks of several processes.

The sufficient statistic can be used in a univariate surveillance method. Schiöler and Frisé (2012) derive a generalized likelihood ratio method based on the sufficient reduction. If the outbreaks appear simultaneously for the different processes, then we have a univariate sufficient statistic with one change point, and optimal methods can be derived. However, when the outbreaks appear at different times, the sufficient statistic has more than one change point in the distribution. Even though each component has one change point, the distribution of the sufficient statistic is not constant either for $t < \tau_i$ or for $t \geq \tau_i$. The proofs commonly used for minimax or expected delay optimality require that there is only one change between two distributions. Since no method with exact optimality can be expected, the properties of the methods have to be described. This can be done for example by a simulation study based on models relevant for the application. A simulation study was made to examine robustness and other properties of the method, as discussed in Section 5.

The method was applied to nine seasons of influenza in Sweden and the multivariate method proved more efficient than the earlier method based on the total incidence in Sweden.

5. Evaluation of methods for spatial outbreak detection

In order to interpret an alarm or a lack of alarm one needs to know the properties of the method used. If frequent false alarms can be expected, an alarm may not give cause for serious concern.

In order to choose an appropriate method in practice one needs to know which factors influence the detection ability and which methods are preferable in a given situation. The urgent need for further discussion on metrics for evaluation of surveillance was stressed in Fricker (2011) as well as in the discussion and rejoinder.

The choice of the alarm limit can be made by examining the results of some limits when applied to a model, historical data, or simulated data. If a simulation study is used, it should be based on models relevant for the application. Even if a nonparametric method is suggested, a parametric model is needed to generate data for the simulation study.

Some evaluation criteria will now be discussed. We will discuss measures for false alarms and detection delay as well as a measure for the balance between these measures.

5.1 *False alarms*

Surveillance, where we make successive decision, differs from hypothesis testing in the view of false alarms. In hypotheses testing a fixed significance level is important. In contrast, the commonly used methods in surveillance have a probability of any false alarm that tends to one as time tends to infinity. It has been suggested that this can be avoided by letting the alarm limit tend to infinity. However, this will seriously harm the ability to detect late changes, as demonstrated by Bock (2008). Thus, false alarms are not regarded in the same way in surveillance as in hypothesis testing. The False Discovery Rate (FDR), was suggested by

Benjamini and Hochberg (1995) for hypothesis testing with multiple comparisons. FDR is the proportion of rejections which are false. It has been suggested also for surveillance problems but the FDR is not as important in surveillance as in hypothesis testing because of the different views on false alarms.

Since false alarms are unavoidable in surveillance, the distribution of t_A , the time to false alarm is important. The false alarm measure most commonly used in surveillance is $ARL^0 = E[t_A | \tau = \infty]$. This measure is often useful, but for some situations it is not enough. Zhou, *et al.* (2012) demonstrated the importance of studying several characteristics of the distribution, such as percentiles and the standard deviation of the run length at the complex situation of varying sample sizes (see Section 2.3). Mei (2008) demonstrated that (for a complex null state) ARL^0 might not be finite even when the expected value at a change is finite. He suggests alternative measures, some of them related to percentiles.

ARL^0 is naturally generalized for multivariate surveillance as $E[t_A | \tau_{\min} = \infty] = E[t_A | \tau_1 = \infty, \dots, \tau_p = \infty]$ where t_A is the time of the general alarm for the multivariate situation. ARL^0 of the multivariate procedure, as suggested above, can be interpreted as the expected time until a false alarm. When methods are compared it is common to use the same false alarm tendency for all methods. It requires a very large number of replicates to ensure that there is not an error in the false alarm value that makes the comparison biased. The technique suggested by Frisén and Sonesson (2006) can be used to ensure that the alarm limit was determined with enough accuracy to make the error in the important properties negligible. Alternatively, the median run length, MRL^0 , can be used instead of the expected value with the same generalization for the multivariate situation as for ARL^0 . MRL^0 is more convenient to use in simulation studies since it

requires considerably fewer replicates to give the same accuracy of the alarm limit. For both the spatial and the univariate influenza in Sweden MRL^0 was set to 780 for all methods.

In univariate theoretical work the false alarm probability, $PFA=P(t_A<\tau)$, is sometimes used. This is naturally generalized for multivariate surveillance as

$$PFA = P(t_A < \tau_{\min}) = \sum_{i=1}^{\infty} \left(P(t_A < \tau_{\min} | \tau_{\min} = i) P(\tau_{\min} = i) \right).$$

Note that the distribution of τ_{\min} influences the suggested multivariate PFA expression.

5.2 Delay

Timeliness in detection is of extreme interest in surveillance. Unfortunately, the expected delay can be quite different for different change points. The expected delay is a function of τ . In Section 5.2.1 we will describe this dependency of the delay on τ . Often, a summarizing measure which gives only one value of the delay is used, and such measures are described in Section 5.2.2.

5.2.1 Delay as a function of the time of the change

We start by recapturing the univariate case where the expected delay for a specific value of τ is $ED(\tau) = E\{\max(0, t_A - \tau)\}$. Since $ED(\tau)$ for most methods tends to zero, it is useful to study the delay conditional on no alarms before τ . For a specific value of τ , the Conditional Expected Delay, CED, is

$$CED(\tau) = E[t_A - \tau | t_A \geq \tau].$$

The first use of the term CED and calculation for a specific value of τ , different from 1 and ∞ , seems to be by Zacks and Kenett (1994). Andersson (2002, Frisén and Wessman (1998) reported

CED as a function of τ , and in Friséen (2003, Friséen and Sonesson (2006) it was strongly advocated that the whole CED curve should be studied.

CED can be generalized for multivariate surveillance to $CED(\tau_1, \tau_2, \dots, \tau_p) = E[t_A - \tau_{\min} | \tau_{\min} \leq t_A]$, see Friséen, *et al.* (2010b). This delay measure depends on all the change points. However, there is often some relation between the change times which simplifies the picture. The CED curves differ considerably for different relations between the values of the change times. For simultaneous outbreaks, the multivariate CED reduces to the univariate one. For the case of influenza in Sweden, the assumption of a time lag considerably reduced the complexity. The CED had its worst value (three weeks) when the outbreak occurred at the start of the surveillance but was reduced to about two weeks when the outbreak occurred late. This is natural since we used a nonparametric method which needed time to gather information. Assunção and Correia (2009) suggested a modified measure, $CED^*(\tau)$, for cluster detection. Here the delay is not the time between the outbreak and the alarm. Instead, only those events which belong to the space time cluster are regarded.

Sometimes the time available for action is limited. In such situations it is important to use a surveillance system with high detection ability within the existing time frame. This property can be measured by the Probability of Successful Detection, PSD, which was suggested by Friséen (1992). It measures the probability that an alarm is called within d time units. In the multivariate case it was defined by Friséen (2009b) as

$$PSD(d, \tau_1, \dots, \tau_p) = P(t_A - \tau_{\min} \leq d | t_A \geq \tau_{\min}).$$

PSD can be calculated and compared for different values of d , as is done in Marshall, *et al.* (2004). If we expect sudden and major changes, we may want a method with high immediate

detection ability (a high PSD for a small d). In a situation where we expect small changes, the long term detection ability (a high PSD for a large d) may be more important.

5.2.2 Summarized delay

The most commonly used measure of the detection ability is the average run length, given that the change occurs immediately ($\tau=1$). This is widely used in univariate surveillance and often named zero-state ARL or ARL^1 . In univariate surveillance the ARL^1 has a simple relation to CED, namely $ARL^1=CED(1)+1$. To consider only $\tau=1$ in the univariate case is a limitation, and the univariate ARL^1 is criticized as a formal optimality criterion for example by Frisé (2003). ARL^1 is the most commonly used evaluation measure also in the multivariate case. However, it is seldom explicitly defined. One possibility is to define the multivariate zero-state ARL as $E[t_A|\tau_{\min}=1]$. However, the values of CED for $\tau_{\min}=1$ vary a lot for different relations between the values of τ_{\min} and the change times of the other processes. Thus, there is no unique ARL^1 with the definition $E[t_A|\tau_{\min}=1]$. Another possibility is to define the multivariate zero-state ARL as $E[t_A|\tau_1=\tau_2=\dots=\tau_p=1]$. This is the definition implicit in most publications. Here, it is assumed that all processes change at the same time. As was seen in Section 4.3 a sufficient reduction to a univariate problem exists when all processes change at the same time. Thus, a reduction to a univariate surveillance statistic is the proper procedure by the sufficiency principle. Thus, we have a univariate inference. ARL^1 is thus questionable as a formal measure for comparing methods for genuinely multivariate problems.

The opposite way to avoid the dependence on τ in the univariate case is to use the steady-state conditional expected delay, $CED_{SS} = \lim_{\tau \rightarrow \infty} CED(\tau)$. The asymptotic measure is often used also in

multivariate situations, but this is not without problems. It is often implicit that the situation $\tau_1=\tau_2=\dots=\tau_p=t$, $t\rightarrow\infty$ is considered. For equal change points we have a unique delay value for each method. However, this is another example of the situation where univariate surveillance can be used instead of multivariate surveillance since there is a sufficient reduction to univariate surveillance. For other situations than simultaneous changes there is no simple asymptotic CED. Even though all the τ values tend to infinity, it also matters how they do this. Specification of how the times of the change points are related is necessary.

A commonly used optimality criterion is the minimax criterion by Lorden (1971). This avoids the dependency on τ by using the least favorable value of τ . This and other advanced measures can be generalized to the multivariate situation, but some structure on the problem is necessary to make the measure comprehensive. The sufficiency principle can be a useful at derivation of relevant measures.

5.3 *Predictive value*

When a method calls an alarm, we need to know whether this alarm is a strong or weak indication of a change. Methods designed to have a fast initial response often has a high probability that an early alarm is false. This is so even for the Shewhart method, while the balance between false alarms and detection ability is more stable over time for other methods. Zhou, *et al.* (2012) observed that for methods designed for varying sample size over time (see Section 2.3) it is an important difference in this balance over time between methods. The predictive value is a well-established measure in epidemiology. In surveillance, however, we need a variant that also incorporates time. The difference in surveillance, as compared to situations involving only one decision, is that we can get an alarm at any time point, and

therefore the predictive value is a function of the time of the alarm. In order to interpret an alarm at time t_A , it is necessary to consider the balance between the risk of false alarms, the detection ability, and the probability of a change. If we have one change point τ and this is regarded as a random variable, this can be done by the probability of an outbreak at an alarm, as suggested by Frisé (1992):

$$PV(t) = P(\tau \leq t | t_A = t) = \frac{\sum_{i=1}^t P(t_A = t | \tau = i) P(\tau = i)}{\sum_{i=1}^t P(t_A = t | \tau = i) P(\tau = i) + P(t_A = t | \tau > t) P(\tau > t)}.$$

In a multivariate setting this was generalized by Frisé, *et al.* (2010b) to

$$PV(t) = P(\tau_{\min} \leq t | t_A = t) = \frac{\sum_{i=1}^t (P(t_A = t | \tau_{\min} = i) P(\tau_{\min} = i))}{\sum_{i=1}^t (P(t_A = t | \tau_{\min} = i) P(\tau_{\min} = i)) + P(t_A = t | \tau_{\min} > t) P(\tau_{\min} > t)}.$$

The components in the formula generally depend on the relation between the change points. The predictive value also depends on whether outbreaks appear frequently or rarely. Knowledge of the exact distribution of τ_{\min} is seldom available, but it can still be useful to have a rough indicator. For the influenza in Sweden τ_{\min} was assumed to be geometrically distributed with $P(\tau_{\min} = i) = (1 - \nu)^{i-1} \nu$. This may not give the closest fit to the onset times in Sweden, but in order to detect outbreaks which occur at unexpected times we did not want to include information on which week is the most common for the onset. From all available historical data on seasonal influenza the level of intensity was roughly estimated to be $\nu = 0.1$. With this intensity the PV is above 0.99 with the method used, and for a lower intensity, $\nu = 0.01$, which

weakens the PV, it is above 0.95. The method and alarm limit used were considered potentially useful for the application since the predictive value was high.

6. Discussion

Statistical surveillance, where timely decisions are made, has become an important part of statistics. Simple applications for industrial quality control dominated the development of surveillance theory for a long time. Today emerging needs in public health and other areas have triggered the development of more advanced methods. The new methods adjust to the complexity of the applications by advanced inference. Although methods have been developed under different scientific cultures, inferential similarities can be identified. Time is essential in surveillance in contrast to the situation in the well developed areas of hypotheses testing and estimation. We have no fixed hypotheses but the problem changes with time. It is not only necessary to make the right decision it is also essential when it is made. The theory of surveillance tells us when we have enough information for the decision.

The detection of outbreaks of epidemiological diseases is an important area of on-line surveillance. Several different definitions of an outbreak are used, explicitly or implicitly, in the literature. The definitions correspond to different aims and have consequences for which methods are suitable.

A spatial setting is often present in public health surveillance. Spatial surveillance is a special case of multivariate surveillance. For multivariate problems it is always hard to derive optimal

methods and sometimes even hard to define optimality. The description of general multivariate approaches was used to identify inferential approaches to spatial outbreak methods.

Different kinds of spatial outbreaks are characterized by the relation between the times of the outbreaks at different locations. The most commonly used definition of a spatial outbreak is an emerging cluster. Besides that, there are also other spatial outbreak situations. The contrast between an outbreak at only one location among many and simultaneous outbreaks at all locations is interesting. A case between these two extremes is the one of a time lag between the outbreaks at different locations. The sufficient reduction plays an important role in finding a likelihood ratio method. An outbreak can have a step change, but an outbreak can also be characterized by a monotonic increase following the non-epidemic constant level. A semiparametric maximum likelihood ratio surveillance method was used for the monitoring of the seasonal influenza in Sweden.

Evaluation metrics appropriate for spatial outbreak detection is a topic of great concern. Simple measures may not always agree with the aim of the application. A prospective method should be evaluated with respect to timeliness. The timeliness concerns both false alarms and detection delay. The predictive value of an alarm can be used to describe the balance, and it can also be useful for deciding how to act at an alarm. Different methods are optimal under different conditions, and the methods can be seen as complements to each other.

The future for research on surveillance is bright. The number of publications on surveillance has increased considerably during the last ten years compared with some competing subjects. The theory of statistical surveillance will be further developed in response to the demands of

applications in various fields. A cross-fertilization back to the applications could then be expected.

Acknowledgement

The author is grateful to the referees for their constructive comments.

References

- Andersson, E. (2002) Monitoring cyclical processes - A nonparametric approach. *Journal of Applied Statistics*, **29** (7) 973-990.
- Andersson, E., Bock, D. and Friséén, M. (2008a) Modeling influenza incidence for the purpose of On-Line monitoring. *Statistical Methods in Medical Research*, **17** 421-438.
- Andersson, E., Kuhlmann-Berenzon, S., Linde, A., Schiöler, L., Rubinova, S. and Friséén, M. (2008b) Predictions by early indicators of the time and height of yearly influenza outbreaks in Sweden. *Scand. J. Public Health*, **36** 475-482.
- Assunçãoa, R. and Correab, T. (2009) Surveillance to detect emerging space–time clusters. *Computational Statistics & Data Analysis*, **53** (8) 2817-2830.
- Basseville, M. and Nikiforov, I. (1993) *Detection of abrupt changes- Theory and application*, Englewood Cliffs: Prentice Hall.
- Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, **57** (1) 289-300.

- Bersimis, S., Psarakis, S. and Panaretos, J. (2007) Multivariate Statistical Process Control Charts: An Overview. *Qual. Reliab. Eng. Int.*, **23** 517-543.
- Bock, D. (2008) Aspects on the control of false alarms in statistical surveillance and the impact on the return of financial decision systems. *Journal of Applied Statistics*, **35** (2) 213-227.
- Burkom, H. S., Murphy, S., Coberly, J. and Hurt-Mullen, K. (2005) Public Health Monitoring Tools for Multiple Data Streams. *Morbidity and Mortality Weekly Report (Supplement)*, **54** 55-62.
- Cakici, B., Hebing, K., Grünewald, M., Saretok, P. and Hulth, A. (2010) CASE –a framework for computer supported outbreak detection. *BMC Medical Informatics and Decision Making*, **10** (14).
- Corberan, A. and Lawson, A. (2011) Conditional Predictive Inference for On-line Surveillance of Spatial Disease Incidence *Statistics in Medicine*, **30** (26) 3095-3116.
- Crepey, P. and Barthelemy, M. (2007) Detecting Robust Patterns in the Spread of Epidemics: A Case Study of Influenza in the United States and France. *American Journal of Epidemiology*, **166** (11) 1244-1251.
- Farrington, C. P. and Andrews, N. J. (2004) Outbreak detection: application to infectious disease surveillance. In *Monitoring the Health of Populations* (eds R. Brookmeyer and D. F. Stroup). Oxford: Oxford University Press.
- Farrington, C. P., Andrews, N. J., Beal, A. D. and Catchpole, M. A. (1996) A Statistical Algorithm for the Early Detection of Outbreaks of Infectious Disease. *Journal of the Royal Statistical Society A*, **159** 547-563.

- Farrington, C. P. and Beale, A. D. (1993) Computer-aided detection of temporal clusters of organisms reported to the Communicable Disease Surveillance Centre. 1993:3 R78-R82.
- Farrington, C. P. and Beale, A. D. (1998) The Detection of Outbreaks of Infectious Disease. In *Geomed 97* (eds L. Gierl, A. Cliff, A. J. Valleron, C. P. Farrington and M. Bull), pp. 97-117. Leipzig: Teubner verlag
- Fricker, J. R. D. (2011) Some methodological issues in biosurveillance. *Statistics in Medicine*, **30** (5) 403-415.
- Frisén, M. (1974) Stochastic deviation from elliptical shape. Thesis. *Department of Statistics, University of Gothenburg, Gothenburg.*
- Frisén, M. (1992) Evaluations of Methods for Statistical Surveillance. *Statistics in Medicine*, **11** (11) 1489-1502.
- Frisén, M. (2003) Statistical surveillance. Optimality and methods. *International Statistical Review*, **71** (2) 403-434.
- Frisén, M. (2009) Optimal sequential surveillance for finance, public health and other areas. Editor's special invited paper. *Sequential Analysis*, **28** (3) 310-337, discussion 338-393.
- Frisén, M. (2010) Principles for Multivariate Surveillance In *Frontiers in Statistical Quality Control 9* (eds H.-J. Lenz, P.-T. Wilrich and W. Schmid), pp. 133-144. Heidelberg: Physica-Verlag
- Frisén, M. and Andersson, E. (2009) Semiparametric surveillance of outbreaks. *Sequential Analysis*, **28** (4) 434-454.
- Frisén, M., Andersson, E. and Pettersson, K. (2010a) Semiparametric estimation of outbreak regression. *Statistics*, **44** (2) 107 - 117.

- Frisén, M., Andersson, E. and Schiöler, L. (2009) Robust outbreak surveillance of epidemics in Sweden *Statistics in Medicine*, **28** (3) 476-493.
- Frisén, M., Andersson, E. and Schiöler, L. (2010) Evaluation of Multivariate Surveillance. *Journal of Applied Statistics*, **37** (12) 2089-2100.
- Frisén, M., Andersson, E. and Schiöler, L. (2011) Sufficient reduction in multivariate surveillance. *Communications in Statistics -Theory and Methods*, **40** (10) 1821-1838.
- Frisén, M. and Sonesson, C. (2006) Optimal surveillance based on exponentially weighted moving averages. *Sequential Analysis*, **25** (4) 379-403.
- Frisén, M. and Wessman, P. (1998) Quality improvement by likelihood ratio methods for surveillance. In *Quality Improvement Through Statistical Methods* (ed B. Abraham), pp. 187-193. Boston: Birkhauser.
- Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S. and Brilliant, L. (2009) Detecting influenza epidemics using search engine query data. *Nature*, **457** (7232) 1012-1014.
- Han, S., Tsui, K.-L., Ariyajunyab, B. and 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*, **26** 279-289.
- Held, L. (2010) A model-based view on outbreak detection. *Conference of Statistical Methods for Outbreak Detection*, Milton Keynes, UK.
- Hotelling, H. (1947) Multivariate Quality Control In *Techniques of statistical analysis* (eds C. Eisenhart, M. W. Hastay and W. A. Wallis), pp. 111-184. New York: McGraw-Hill.

- Hulth, A., Andrews, N., Ethelberg, S., Dreesman, J., Faensen, D., van Pelt, W. and Schnitzler, J. (2010) Practical usage of computer-supported outbreak detection in five European countries. *Eurosurveillance*, **15** (36) 1-6.
- Hulth, A., Rydevik, G. and Linde, A. (2009) Web Queries as a Source for Syndromic Surveillance. *PLoS ONE*, **4** (2) e4378.
- Höhle, M. (2010) Aberration Detection in R Illustrated by Danish Mortality Monitoring In *Biosurveillance* (eds T. Kass-Hout and X. Zhang): CRC Press.
- Höhle, M. and Paul, M. (2008) Count data regression charts for the monitoring of surveillance time series. *Computational Statistics & Data Analysis*, **52** (9) 4357-4368.
- Jiang, W., Han, S. W., Tsiu, K.-L. and Woodall, W. H. (2011) Spatiotemporal surveillance methods in the presence of spatial correlation *Statistics in Medicine*, **30** (5) 569-583.
- Jiang, W. and Tsui, K.-L. (2008) A theoretical framework and efficiency study of multivariate statistical process control charts. *IIE Transactions*, **40** (7) 650 - 663.
- Joner, J. M. D., Woodall, W. H., Reynolds Jr, M. R. and Fricker, R. D. (2008) A One-sided MEWMA Chart for Health Surveillance. *Quality and Reliability Engineering International*, **24** 503-518.
- Kaufmann, A. F., Meltzer, M. I. and Schmid, G. P. (1997) The economic impact of a bioterrorist attack: Are prevention and postattack intervention programs justifiable? *Emerging Infectious Diseases*, **3** (2) 83-94.
- Knox, E. G. (1964) The detection of space-time interactions. *Applied Statistics*, **13** 25-29.
- Kulldorff, M. (2001) Prospective time periodic geographical disease surveillance using a scan statistic. *Journal of the Royal Statistical Society A*, **164** (1) 61-72.

- Kulldorff, M. (2010) SaTScan version 8.2.1: software for the spatial, temporal, and space-time scan statistics. (Available from <http://www.satscan.org>).
- Lawson, A. B. and Kleinman, K., eds (2005) *Spatial and Syndromic Surveillance for Public Health*, Chichester: Wiley.
- Le Strat, Y. and Carrat, F. (1999) Monitoring epidemiologic surveillance data using hidden Markov models. *Statistics in Medicine*, **18** (24) 3463-3478.
- Lorden, G. (1971) Procedures for reacting to a change in distribution. *Annals of Mathematical Statistics*, **42** (6) 1897-1908.
- Marshall, C., Best, N., Bottle, A. and Aylin, P. (2004) Statistical issues in the prospective monitoring of health outcomes across multiple units. *Journal of the Royal Statistical Society A*, **167** (3) 541-559.
- Marshall, J. B., Spitzner, D. J. and Woodall, W. H. (2007) Use of the local Knox statistic for the prospective monitoring of disease occurrences in space and time. *Statistics in Medicine*, **26** (7) 1579-1593.
- Mei, Y. (2008) Is average run length to false alarm always an informative criterion? *Sequential Analysis*, **27** (4) 354-376.
- Mei, Y., Han, S. W. and Tsui, K.-L. (2011) Early detection of a change in poisson rate after accounting for population size effects. *Statistica Sinica*, **21** 597-624.
- Paul, M., Held, L. and Toschke, A. M. (2008) Multivariate modelling of infectious disease surveillance data. *Statistics in Medicine*, **27** 6250-6267.
- Raubertas, R. F. (1989) An Analysis of Disease Surveillance Data That Uses the Geographic Locations of the Reporting Units. *Statistics in Medicine*, **8** (3) 267-271.

- Rogerson, P. A. (2001) Monitoring point patterns for the development of space-time clusters. *Journal of the Royal Statistical Society A*, **164** (1) 87-96.
- 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.
- Ryan, A. G. and Woodall, W. H. (2010) Control charts for Poisson count data with varying sample sizes. *Journal of Quality Technology*, **42** 260--274.
- Ryan, T. P. (2000) *Statistical methods for quality improvement*, New York: John Wiley & Sons.
- Schiöler, L. (2011) Characterization of influenza outbreaks in Sweden. *Scandinavian Journal of Public Health*, **39** 427-436
- Schiöler, L. and Friséén, M. (2012) Multivariate outbreak detection. *Journal of Applied Statistics*, **39** (2) 223-242.
- Serfling, R. (1963) Methods for current statistical analysis of excess pneumonia-influenza deaths. *Public Health Reports*, **78** (6) 494-506.
- Shmueli, G. and Burkom, H. S. (2010) Statistical Challenges Facing Early Outbreak Detection in Biosurveillance. *Technometrics*, **52** (1) 39-51.
- Shu, L., Jiang, W. and Tsui, K. L. (2011) A comparison of weighted CUSUM procedures that account for monotone changes in population size. *Statistics in Medicine*, **30** (7) 725-741.
- Sonesson, C. (2007) A CUSUM framework for detection of space-time disease clusters using scan statistics. *Statistics in Medicine*, **26** 4770-4789.
- Sonesson, C. and Bock, D. (2003) A review and discussion of prospective statistical surveillance in public health. *Journal of the Royal Statistical Society A*, **166** (1) 5-21.

- Sonesson, C. and Frisé, M. (2005) Multivariate surveillance. In *Spatial surveillance for public health* (eds A. Lawson and K. Kleinman), pp. 169-186. New York: Wiley.
- Tartakovsky, A. G. and Veeravalli, V. (2004) Change-point detection in multichannel and distributed systems In *Applications of Sequential Methodologies* (eds N. Mukhopadhyay, S. Datta and S. Chattopadhyay). New York: Marcel Dekker, Inc.
- Tsui, K.-L., Chiu, W., Gierlich, P., Liu, X. and Maschek, T. (2008) A Review of Healthcare, Public Health and Syndromic Surveillance. *Quality Engineering*, **20** 435-450.
- Tsui, K.-L., Han, S. W., Jiang, W. and Woodall, W. H. (2012) A Review and Comparison of Likelihood Based Charting Methods. *IIE Transactions*(in press).
- Tsui, K.-L., Wong, S. Y., Jiang, W. and Lin, C.-F. (2011) Recent Research and Developments in Temporal and Spatiotemporal Surveillance for Public Health *IEEE Transaction on Reliability*, **60** (1) 49-58.
- Unkel, S., Farrington, C. P., Garthwaite, P. H., Robertson, C. and Andrews, N. (2012) Statistical methods for the prospective detection of infectious disease outbreaks: a review. *Journal of the Royal Statistical Society A*, **175** (1).
- Wessman, P. (1998) Some Principles for surveillance adopted for multivariate processes with a common change point. *Comm. Stat. Theor. Meth.*, **27** (5) 1143-1161.
- Woodall, W. H., Grigg, O. A. and Burkom, H. S. (2010) Research issues and ideas on health-related surveillance In *Frontiers in Statistical Quality Control 9* (eds H.-J. Lenz, P.-T. Wilrich and W. Schmid), pp. 145-155. Heidelberg: Physica-Verlag.

- Zacks, S. and Kenett, R. S. (1994) Process tracking of time series with change points In *Recent Advances in Statistics and Probability* (eds J. P. Vilaplana and M. L. Puri), pp. 155-171. Zeist, The Netherlands: International Science Publishers.
- Zhou, H. and Lawson, A. B. (2008) EWMA smoothing and Bayesian spatial modeling for health surveillance. *Statistics in Medicine*, **27** (28) 5907-5928.
- Zhou, Q., Luo, Y. and Wang, Z. (2010) A control chart based on likelihood ratio test for detecting patterned mean and variance shifts *Computational Statistics & Data Analysis*, **54** (6) 1634-1645.
- Zhou, Q., Zou, C., Wang, Z. and Jiang, W. (2012) Likelihood-based EWMA charts for monitoring poisson count data with time-varying sample sizes. *Journal of the American Statistical Association*, in print.

Marianne Frisén is professor emerita in Statistics at Statistical Research Unit, Department of Economics, University of Gothenburg, Gothenburg, Sweden. She received a PhD in Statistics from this university. She is an elected member of the International Statistical Institute. Her main interests are in statistical surveillance, order restricted inference, robust methods, applied work, and the foundations of statistical inference.