



Research Report  
Statistical Research Unit  
Department of Economics  
Göteborg University  
Sweden

---

**Statistical Surveillance of Epidemics:  
Peak Detection of Influenza in Sweden.**

**David Bock, Eva Andersson &  
Marianne Frisé**

**Research Report 2007:5  
ISSN 0349-8034**

---

Mailing address:	Fax	Phone	Home Page:
Statistical Research Unit P.O. Box 640 SE 405 30 Göteborg Sweden	Nat: 031-786 12 74	Nat: 031-786 00 00  Int: +46 31 786 12 74	<a href="http://www.statistics.gu.se/">http://www.statistics.gu.se/</a>

# Statistical Surveillance of Epidemics: Peak Detection of Influenza in Sweden

David Bock, Eva Andersson and Marianne Frisén\*

Statistical Research Unit, Department of Economics, Göteborg University,  
SE 405 30 Göteborg, Sweden

## Summary

A statistical surveillance system gives a signal as soon as data give enough evidence of an important event. We consider on-line surveillance systems for detecting changes in influenza incidence. One important feature of the influenza cycle is the start of the influenza season, and another one is the change to a decline (the peak). In this report we discuss statistical methods for on-line peak detection. One motive for doing this is the need for health resource planning. Surveillance systems were adapted for Swedish data on laboratory verified diagnoses of influenza. In Sweden, the parameters of the influenza cycles vary too much from year to year for parametric methods to be useful. We suggest a non-parametric method based on the monotonicity properties of the increase and decline around a peak. A Monte Carlo study indicated that this method has useful stochastic properties. The method was applied to Swedish data on laboratory verified diagnoses of influenza for seven periods.

*Key words:* Disease surveillance; Monitoring; Non-parametric; Order restrictions.

## 1 Introduction

When adverse health events occur, quick and safe detection is essential. It is important that these events are detected very soon after they occur (a short delay) and that, when an alarm is called, it is probable that an event has occurred (a high predictive value). Continual observations are made and the theory of statistical surveillance can be used to evaluate the cumulated information in order to detect a change in the underlying process. For a review and discussion of prospective statistical surveillance in public health, see Sonesson and Bock (2003). Control charts first came in use 1930 for detecting changes in industrial production. However, the requirements for evaluations are more pronounced for public health and the need of simplicity is less. Gains in cross-fertilization between the areas are demonstrated by Woodall (2006).

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 allocate medical resources. A comparison between public health surveillance systems for monitoring influenza incidence in 24 European countries is made in Hannoun and Tumova (2000). Our report is part of the project "Statistical warning systems for emergency management" aiming to construct automated systems for monitoring epidemic diseases in Sweden. Even though most of the problems are the same for other diseases, such as e.g. tularemia and campylobacteriosis, we chose to describe the situation for influenza in order to be specific.

Most of the literature on public health surveillance is concerned with the detection of disease outbreaks. From a public health perspective, however, other features of the influenza cycle are of interest, such as the peak after which the influenza incidence starts to decline. One important reason for this is the need for health resource planning. Another reason is that there are other adverse health events that cause symptoms similar to influenza. One example is the avian flu and another one is anthrax exposure after a possible bioterrorism attack. Since these symptoms are similar to those of ordinary influenza, outbreaks of such adverse health events might be evident as an additional peak in the influenza cycle. Thus, in order to distinguish such a peak from one of the ordinary influenza, a method for peak detection of the influenza incidence is required. Peak detection was also the aim in surveillance for bioterrorism as described in Dafni et al. (2004).

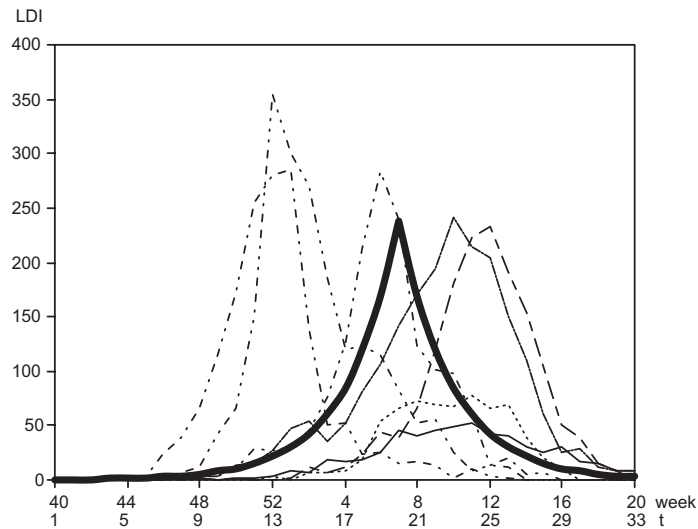
In Sweden, the Swedish Institute for Infectious Disease Control (SMI) carries out influenza surveillance and gives timely information to the society. A description of the Swedish system for collecting data on influenza can be found in e.g. Linde et al. (2004). Influenza incidence is measured by the weekly number of laboratory diagnosed influenza cases (LDI) and the weekly proportion of patients with influenza-like illnesses (ILI). Data on ILI are collected from a number of sentinel physicians. The LDI data consist of reports from five virus laboratories and a number of microbiological laboratories. Andersson et al. (2008) made the conclusion regarding the ILI variable that it should be interpreted with care since time-dependent effects, such as the physicians' inclination to send reports, seem to be present. For peak detection, the LDI variable seems to be more reliable. In this paper, the analyses and models are based on the LDI series from seven influenza periods.

---

\* Corresponding author: e-mail: marianne.frisen@statistics.gu.se, Phone: +46 31 773 1255, Fax: +46 31 773 12

Information about data for different geographical areas might be of value since the influenza spreads to different areas at different times. In an exploratory study of Swedish data (Bock and Pettersson (2006)), it was found that no spread between neighboring areas was evident but that the influenza first appeared in the big cities. In this paper, the spatial component is not considered. Instead, the analysis is based on aggregated data for the whole of Sweden.

Automatic systems for detecting influenza epidemics should have known properties concerning detection ability, risks of false alarms and predictive value as described in Section 3. Many surveillance systems are based on the deviations between the observed value and the expected value given that no change has occurred. Often, the expected value is assumed to be a known constant or a parametric function of time with known parameter values. A comparison could be made with an “average” curve as in e.g. Sebastiani et al. (2006). This will tell if this season is extraordinary compared with an average season. However, here the purpose is to detect the peak for the present season. As is seen in Figure 1 (which will be further described in the next section) the parameter values are not constant from one year to the next and using the “average” curve as the target value in the surveillance system, is not useful for peak detection. A non-parametric method is therefore suggested and evaluated.



**Figure 1:** Seven influenza periods and an “average” curve (bold). Both the time as calendar week number and the time as observation week number are given.

In Section 2 methodology of statistical surveillance of a peak is described. The problem is described in Section 2.1. Peak detection methods using parametric and non-parametric models for the incidence are described in Sections 2.3 and 2.4, respectively. In Section 3, the properties of a method based on order restrictions are investigated both by a Monte Carlo study and by application to Swedish data. Conclusions are given in Section 4.

## 2 Statistical Surveillance to Detect a Peak

### 2.1 The Surveillance Problem

In statistical surveillance, a process is monitored on-line with the aim of quick detection of an important underlying change at an unknown time point. For a review of optimality and methods, see Frisén (2003). Public health surveillance is described by e.g. Sonesson and Bock (2003), Frisén and Sonesson (2005) and Woodall (2006). The importance of using statistical surveillance in routine communicable disease surveillance is described by O'Brien and Christie (1997). An annotated bibliography for syndromic surveillance, including influenza surveillance, is available at the web site of Centers for Disease Control and Prevention (CDC).

The literature on detecting a decline after a peak in the incidence of influenza is very sparse. Most work has concerned the detection of the start of an epidemic. Some authors (for example Brookmeyer and You (2006)) assess the end of a period of adverse events. We address the problem where the expected influenza incidence, at an unknown time, reaches a peak which we want to detect as soon as possible after it has occurred. The peak represents a change in the expected influenza incidence, from an increasing to a decreasing incidence. Detecting

a peak can also be seen as detecting a turning point. This is considered for applications in economics by e.g. Koskinen and Öller (2003), Andersson et al. (2004), Andersson et al. (2005) and Andersson et al. (2006b).

Figure 1 shows Swedish data for seven influenza periods. Data were collected from week 40 in the autumn until week 20 in the spring (33 weeks). With consecutive ordering of the observations times this represents  $t = 1, \dots, 33$ . In Figure 1 we see that the peak time varied between  $t=13$  and  $t=25$  and it is also evident that the start time of the increase and the slopes of the curves vary from one year to the next. The “average” curve inserted for comparison. This average curve is achieved by first calculating the parameters (intercept and slope) of the exponential regression for each season. After that the median intercept and median slope among the seven periods are used as rough estimates of the average intercept and slope. The exponential curve with these parameters (the “average curve”) is plotted in Figure 1.

In a surveillance system based on weekly data, we get a new observation on an indicator of the influenza incidence  $X(t)$  each week  $t$ . Thus, at decision week  $s$  we have  $s$  available observations,  $X_s = \{X(1), X(2), \dots, X(s)\}$ . The time,  $\tau$ , at which the expected influenza incidence changes from increasing to decreasing, is regarded as a discrete-valued random variable with probability function  $P(\tau=i)$ , for  $i=1, 2, \dots$ . Based on  $X_s$ , the aim is to discriminate between  $D(s) = \{\tau > s\}$  and  $C(s) = \{\tau \leq s\} = \{\bigcup_{i=1}^s C_i\}$ , where  $C_i = \{\tau=i\}$ . The index  $s$  in  $D(s)$  and  $C(s)$  is suppressed when obvious.  $D$  implies that the change has not yet occurred (the decline has not yet started) while  $C$  implies that the process has changed (the peak has occurred and the incidence has started to decline). At each time point  $s$ , we calculate the value of an alarm statistic (a function of all available observations) and compare it to an alarm limit  $k_s$ . As soon as the value of the alarm statistic exceeds the limit, we infer that the critical event  $C(s)$  has occurred.

## 2.2 Some General Surveillance Methods

It has been shown that methods based on the likelihood ratio have optimality properties (see Shiryaev (1963) and Frisén and de Maré (1991)). The likelihood ratio, at time  $s$ , between event  $C = \{\tau \leq s\}$  and  $D = \{\tau > s\}$  is based on all  $s$  partial likelihood ratios, i.e.

$$LR(s) = \frac{f(x_s | C)}{f(x_s | D)} = \sum_{i=1}^s w_i \cdot \frac{f(x_s | \tau = i)}{f(x_s | \tau > s)},$$

where  $w_i = P(\tau=i)/P(\tau \leq s)$  and  $f$  is the density function. An alarm is given as soon as the LR statistic exceeds a time-dependent limit,  $k_s$ . This is further exemplified in Section 2.3.

The Shiryaev-Roberts approach (SR) by Shiryaev (1963) and Roberts (1966) is the LR method for the limiting case when  $\tau$  follows a geometric distribution with an intensity  $v$  that tends to zero i.e.  $v = P(\tau=t | \tau \geq t) \rightarrow 0$ . The resulting alarm statistic has equal weights,  $w_i$ , for the  $s$  partial likelihood ratios and a constant alarm limit. Thus, the SR method can be seen as using a non-informative distribution for  $\tau$ . The SR method is based on the alarm statistic

$$SR(s) = \sum_{i=1}^s \frac{f(x_s | \tau = i)}{f(x_s | \tau > s)}.$$

An early and still dominating surveillance method is the Shewhart method (Shewhart (1931)) which gives an alarm as soon as the last observation exceeds a limit. The limit is chosen so that the probability of exceeding the alarm limit would be constant for each time point if we disregard the fact that we have a surveillance system with repeated decisions. The history and optimality of the Shewhart method is described by Frisén (2007).

Since the Shewhart method does not utilize all available information, it is optimal only in very special situations.

## 2.3 Parametric Peak Detection

Parametric models, both for the curve that describes how the expected value depends on time and for the stochastic distribution, would be useful as a basis for a surveillance method. Andersson et al. (2008) discuss different parametric models such as a piecewise linear curve, a model with constant levels within the phases for the successive differences, the trigonometric model of Serfling (1963) and a piecewise exponential model,

$$\mu(t | \tau) = \begin{cases} \beta_0 \cdot \exp(\beta_1 \cdot t), & t < \tau \\ \beta_0 \cdot \exp(\beta_1 \cdot (\tau-1) + \beta_2 \cdot (t-\tau+1)), & t \geq \tau \end{cases}$$

which have all been suggested in an influenza context. They conclude that the trigonometric model resulted in a poor fit. The other models, and especially the exponential one, fitted the data well. Analyses by Andersson et al. (2008) of the residuals from the exponential curve demonstrated that the simple model

$$X(t|\tau) = \mu(t|\tau) + \varepsilon(t),$$

where  $\varepsilon \sim \text{iid } N(0, \sigma^2)$ , was satisfactory in a neighborhood of the peak where no small counts appear. Since here the aim is to detect a change in  $\mu$  the LR statistic at decision time  $s$  would be

$$\text{LR}(s) = \sum_{i=1}^s w_i \cdot \frac{f(x_s | \mu = \mu^{\text{Ci}})}{f(x_s | \mu = \mu^{\text{D}})}.$$

Each component is weighted by the probability,  $w_i$  of a turn at that time. The probability of a turn does depend on the assumed distribution for  $\tau$ . A common assumption is a geometric distribution, thus  $w_i = (v(1-v)^{i-1})/(1-(1-v)^s)$ . The alarm statistic of the LR method at decision time  $s$  would be

$$\sum_{i=1}^s \frac{v(1-v)^{i-1}}{1-(1-v)^s} \cdot \frac{\exp\left(-\left(\frac{1}{2\sigma^2}\right) \sum_{t=i}^s (x(t) - \beta_0 \exp(\beta_2 \cdot t))^2\right)}{\exp\left(-\left(\frac{1}{2\sigma^2}\right) \sum_{t=i}^s (x(t) - \beta_0 \exp(\beta_1 \cdot t))^2\right)}.$$

The alarm limit of the LR method is  $k_s = (P(\tau > s)/P(\tau \leq s)) \cdot k$ , where  $k$  is a constant. It is thus not constant over time.

In the Shiryaev-Roberts approach for detecting a change in  $\mu$ , the components are not weighted, thus the SR statistic at time  $s$  would be

$$\text{SR}(s) = \sum_{i=1}^s \frac{f(x_s | \mu = \mu^{\text{Ci}})}{f(x_s | \mu = \mu^{\text{D}})}.$$

The alarm limit of the Shiryaev-Roberts method is constant. The reason for choosing the SR method instead of the LR method is that an informative prior will have a rather strong influence on the properties of the surveillance. In Andersson (2004) it is shown that an informative prior works well only for the case when the observed peak time is the same as the most likely peak time in the prior. For cases with other peak times, the inclusion of an informative prior results in poor properties (long delay and low predictive value). It might in fact be extra important to detect a peak which does not come as expected.

General approaches for surveillance can be applied to the problem of peak detection. The full likelihood ratio method is optimal in the sense that the expected delay is minimized for a fixed false alarm probability. For further details see Shiryaev (1963) and Frisén and de Maré (1991). A related surveillance approach is the use of Hidden Markov Models (HMM). Le Strat and Carrat (1999) and Rath et al. (2003) used HMM for the ILI incidence rates in France. In an HMM, the process is assumed to be switching between different states (which differ with respect to their statistical properties) and the switching is governed by an unobservable Markov chain. The periods are usually classified by means of the posterior probability. In Le Strat and Carrat (1999) and Rath et al. (2003), the Markov chain has two states (non-epidemic and epidemic). The model is of the same type as that described in the beginning of this section. The two states are for  $t < \tau$  and  $t \geq \tau$  respectively. Whereas Le Strat and Carrat (1999) used the cyclical regression model of Serfling for  $\mu(t|\tau)$  with a Gaussian distribution of  $\varepsilon(t)$  for each state, Rath et al. (2003) used an exponential and a Gaussian distribution for non-epidemic and epidemic periods, respectively, both without seasonal components.

In Le Strat and Carrat (1999) and Rath et al. (2003) estimation and classification were made on a fixed historical series, but HMM might also be used for surveillance. In the article by Andersson et al. (2005), a theoretical comparison was made between an HMM-based surveillance method and likelihood ratio methods for surveillance. One important difference is that in an HMM, knowledge about the type of the next change to be detected is not utilized. Except for this, the method is equivalent to the LR method described in Section 2.2. When data for both epidemic and non-epidemic periods are available for many years HMM might be convenient for classification purpose. In practice, however, it is clear whether to expect an increase or a decline, and therefore it is better to use the LR method than the HMM for surveillance.

The full LR is optimal, but it requires information on the intensity of peak occurrence,  $v$ , and to avoid risky assumptions, we here use the SR approach (see Section 2.2).

The SRLin method for the detection of a peak was suggested by Andersson (2002). It is based on the SR approach, and the expected value is assumed to be a piecewise linear curve. Other parametric curves can easily be used in the SR approach, by specifying  $\mu$  according to the desired parametric function. As long as the assumptions made regarding the parametric model are sound, the surveillance system will be powerful.

In a parametric surveillance system the parameters have to be specified for example by estimation from historical data. The possibility of basing a parametric surveillance system on past data for future use is hampered by lack of consistency between periods, since the risk of misspecification is then substantial. In our material there were large differences between the different influenza periods (i.e. between years) in that the parameters varied considerably between them. The curves differed much in height as well as growth. The periods with high peaks often had a larger growth and vice versa. Also the residual variances differed between the periods. The residual variance was often smaller for periods with low peaks. If we used parametric models, the risk of misspecification would consequently be high. In a surveillance method based on a parametric model for  $\mu$  (e.g. an exponential model) the parameters would have to be estimated, e.g. from past peaks. Each peak can be adequately estimated by an exponential curve, but the parameters of the exponential curve are different for each period. A comparison with an average curve (Figure 1), would only tell if the present data differ from the average curve but not be adequate to detect the peak of the present period. Misspecifications can have an adverse effect on the performance of the surveillance system, see e.g. Andersson et al. (2005). Consider the seven seasons and the average curve in Figure 1. Say that we are at the beginning of one of the seasons with a high early peak (around week 52), and that we, in our surveillance system, use the average curve as the expected value. An alarm will be called when the difference between the observed value and the expected value is a large enough negative value. Because the average curve has a much later peak time, this difference will stay positive almost until the end of that season, thus producing no alarm. The opposite problem can also occur. Thus a misspecification can result in a much longer time to alarm and a low predictive value of alarms. A surveillance method based on a parametric model might hence not be suitable in the current setting. Due to this, the parametric methods will not be addressed further.

## 2.4 Non-Parametric Peak Detection

We do not know of any earlier methods for surveillance of influenza incidences which do not rely on any estimates of the expected value of the incidences. We will use a surveillance method that does not rely on a parametric curve for  $\mu$ . At time  $s$ , we want to discriminate between  $D=\{\tau>s\}$ , under which  $\mu$  is monotonically increasing with time

$$\mu(1)\leq\mu(2)\leq\dots\leq\mu(s),$$

and  $C=\{\tau\leq s\}=\{\tau=1\},\{\tau=2\},\dots,\{\tau=s\}=\{C_1, C_2, \dots, C_s\}$ , under which  $\mu$  is first monotonically increasing and then monotonically decreasing with time

$$\mu(1)\leq\mu(2)\leq\dots\leq\mu(i-1)\leq\mu_{\max}\text{ and } \mu_{\max}\geq\mu(i)\geq\mu(i+1)\geq\dots\geq\mu(t),$$

where  $i=1, 2, \dots, s$  and where  $\mu_{\max}$  is not necessarily observable and where at least one inequality is strict in the second part. For some values of  $i$ , the left hand side set can be empty.

Instead of a parametric model for  $\mu$ , we use only the monotonicity definition of a peak. A surveillance method based on non-parametric estimation under such order restrictions as above was suggested by Frisén (2000) and described and evaluated by Andersson (2002). It was further evaluated and used for different applications by Andersson et al. (2006b), Andersson (2004), Andersson et al. (2004), Andersson (2005), Andersson et al. (2005) and Andersson et al. (2006b). This method is based on the ratio between the maximum likelihoods, where the likelihoods are maximized under the order restrictions above. The maximum likelihood ratio is

$$\frac{\max f(x_s | C)}{\max f(x_s | D)}$$

We consider, as before, the model

$$X(t|\tau) = \mu(t|\tau) + \varepsilon(t),$$

where  $\varepsilon \sim \text{iid } N(0, \sigma^2)$ . The aim is to detect a peak in  $\mu$ , i.e. a change from a monotonical increase to a monotonical decrease. Thus, we have the non-parametric Likelihood-Ratio method, LRnp:

$$\frac{\max f(x_s | C)}{\max f(x_s | D)} = \sum_{i=1}^s w_i \cdot \frac{f(x_s | \mu = \hat{\mu}^{Ci})}{f(x_s | \mu = \hat{\mu}^D)},$$

where  $\hat{\mu}^{Ci}$  and  $\hat{\mu}^D$  are estimated by way of non-parametric regression under the order restrictions above. The solution technique is the pool-adjacent-violators algorithm (see e.g. Robertson et al. (1988)). For the monotonic case ( $\hat{\mu}^D$ ), also called isotonic regression, we estimate  $\mu(t)=E[X(t)]$  as a function of time, under an isotonicity restriction. This is a least square estimate where the sum of squares is minimized under the monotonicity restriction:  $\{\mu(1) \leq \mu(2) \leq \dots \leq \mu(s)\}$ . For an iid Gaussian process, it is also the maximum likelihood estimate. For the unimodal case ( $\hat{\mu}^{Ci}$ ) we also have a least square estimate, but this time under another restriction: the unimodal case is described by Frisén (1986) as consisting of an up-phase, where  $\mu(t)$  is monotonically increasing with  $t$ , and a down-phase, where the regression is monotonically decreasing with  $t$ . Under the assumption that  $\tau=i$ , we have a partition into monotonous phases and hence we can use the least square isotonic regression for each phase. This means that the sum of square should be minimized under the restriction  $\{\mu(1) \leq \mu(2) \leq \dots \leq \mu(i-1) \leq \mu_{\max}$  and  $\mu_{\max} \geq \mu(i) \geq \mu(i+1) \geq \dots \geq \mu(t)\}$ . Just as  $\hat{\mu}^D$ , it is a least square estimate, and for an iid Gaussian process it is also the maximum likelihood estimate.

Here we suggest the SR approach regarding  $\tau$ , which implies that all weights in the LRnp method are equal. An alarm is given as soon as

$$\text{SRnpPeak}(s) = \sum_{i=1}^s \frac{f(x_s | \mu = \hat{\mu}^{Ci})}{f(x_s | \mu = \hat{\mu}^D)} > k,$$

where  $k$  is a constant. The time of alarm,  $t_A$ , for SRnpPeak is

$$t_A = \min[t: \text{SRnpPeak}(t) > k].$$

The properties of the SRnpPeak method are investigated in Section 3.1 by simulations and the method is also applied to seven periods of influenza epidemics in Section 3.2. A user-friendly freeware computer program which performs the calculations is available from the authors. This program is run as a macro in Excel and the interface is very similar to those of the standard procedures of Excel.

The method can be adjusted for different applications by the choice of the variance  $\sigma^2$  and the alarm limit  $k$ . In the evaluation in Section 3.1 and in the application to Swedish data in Section 3.2, we have chosen  $\sigma=16$  as this is the median standard deviation among those resulting from the exponential curves for the seven available periods. We want an estimate of the variance near the peak, and thus, when estimating the residual variance, the observations near the peak were used ( $\text{LDI} \geq 30$ ). The effect of the unavoidable misspecification for some periods is analyzed in the next section. In the following, the alarm limit was chosen to be  $k=10$ . The value of  $k$  was not determined by any formal optimization criteria but chosen by the simulation results to give a reasonable balance between false alarms and delay for applications like that described in Section 4. A larger value of  $k$  would give fewer false alarms and a better predictive value of an alarm. A smaller value would give a quicker detection but worse predictive value. In order to address the behaviour near the peak the surveillance starts with the first observation for which  $\text{LDI} \geq 30$ .

### 3 Properties of the Non-Parametric Peak Detection Method

In Section 3.1 we will examine the stochastic properties by a Monte Carlo study of the SRnpPeak method specified at the end of Section 2.4. In Section 3.2, we will apply the method to the actual data for seven influenza periods.

The SRnpPeak method does not require any parametric assumptions about the curve shape. However, it is easier to detect a peak for a steep curve with a small variance than for a flat curve with a large variance. To demonstrate this we use specific parametric models in our simulation study. The chosen models have relevance for the application. In Andersson et al. (2008) it was concluded that piecewise exponential models (with different start times and slopes) describe the expected influenza incidence rather well. Exponential curves are therefore used in the simulations in Section 3.1. We will examine the properties of the SRnpPeak method for two curves: one curve, “the average curve”, (see Figure 1) that mimics the average influenza peak and one curve with a lower average growth (“the low curve”). The false alarm probability and other properties will depend on the specific curve. We will examine the robustness of the method for other parameters. The lack of symmetry will be addressed by two different non-symmetric peaks which mimic a rapid and a slow decline, respectively. Moreover, the effect of misspecifying the variance will be studied. The robustness study gives information about the behaviour for periods where the situation is extreme, which is useful since the parameters vary much between the different influenza periods.

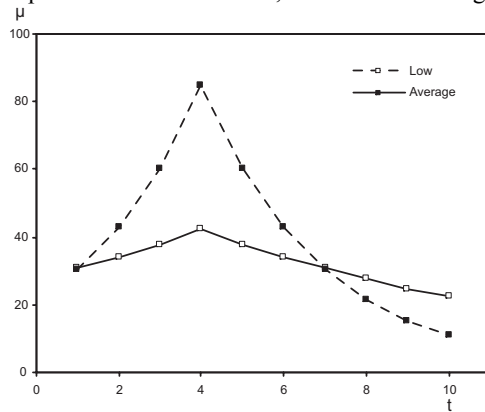
Though timeliness is mentioned as important in public health surveillance (see for example Buehler et al. (2004)), specific evaluation measures reflecting timeliness are seldom used, as pointed out by Sonesson and Bock (2003). Instead measures like sensitivity and predictive value positive (PVP) are often applied. These measures focus on the ability to classify cases correctly and do not explicitly reflect timeliness and the aspect of repeated decisions.

#### 3.1 Monte Carlo Results

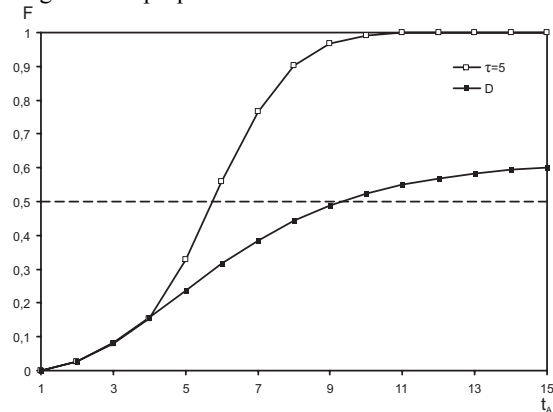
The symmetric average curve with parameters  $\beta_0=3.92$  and  $\beta_1=-\beta_2=0.34$  represents the main alternative as regards what can be expected for the influenza. This is a rough estimate of an average behaviour as illustrated in Figure 1. However, the incidence will sometimes be much higher or much lower. Low curves are expected to cause more problems, given that the variance is the same, since it is more difficult to detect the peak in a low curve. Therefore, we will concentrate on a low curve with parameters  $\beta_0=4.09$  and  $\beta_1=-\beta_2=0.11$ . This is a rough estimate of a low but not unlikely behaviour as seen in Figure 1. The two curves are given in Figure 2. In the method and the generation of data, the standard deviation  $\sigma=16$  is used. Thus, in the simulation study, observations were generated using the following model:

$$X(t) = \exp(\beta_0 + \beta_1 * \tau + \beta_2*(t - \tau)) + \varepsilon(t),$$

where  $\beta_0$ ,  $\beta_1$  and  $\beta_2$  took the values given above,  $\varepsilon \sim \text{iid } N(0, \sigma=16)$ . Simulations were made for  $\tau = \{1, 2, \dots, 10\}$  and also for state D where  $\tau = \infty$  (no peak). The number of replicates for each value of  $\tau$  is 1 million. The replicates were run for  $t=1, \dots, T$  where  $T$  is large enough for the properties to have stabilized.



**Figure 2:** Examples of curves used in the simulation study. Here  $\tau=5$ .



**Figure 3:** The distribution function  $F$  for the alarm time  $t_A$  for the low curve. The run length distributions are given for states D and  $\tau=5$ . The 50th percentile is marked by a dashed line.



When the expected value is increasing (i.e. when the process is in state D), the peak has not yet come and all alarms are false. The false alarms can be controlled by the average run length until a false alarm,  $ARL^0 = E[t_A|D]$ . Hawkins (1992), Gan (1993) and Andersson (2002) suggest that the median run length,  $MRL^0$ , should be used for control as this makes interpretations for skewed distributions easier and the computer time for calculations much shorter. Due to the exponential growth, the curve will soon be very steep and very few or no false alarms will be given. This will influence measures that characterize the false alarms, for example  $MRL^0$ . For the low curve, the  $MRL^0$  equals 9.32 while the median delay, when  $\tau=5$ , is 4 (see Figure 3). In general we want a large  $MRL^0$  and a short delay for motivated alarms. For the average curve described above, the  $MRL^0$  is not finite, but there is a probability of no alarm, due to the large value of  $\beta_1$ . A consequence is that it is not possible to make the two situations comparable by a fixed  $MRL^0$ . However, we can conclude that the risk of a false alarm is very small for the average curve (with the alarm limit we have chosen) but higher for a low curve.

If a change occurs, various actions could be taken, for instance the re-allocation of medical resources. When an alarm is triggered, it is probably a signal of a change, but we need to know how much trust to put in an alarm that comes at time  $t$ . The trust is reflected by the predictive value of an alarm so this measure gives important information for decision-making. Frisén (1992) suggested the following time-dependent predictive value:

$$PV(t)=P(\tau \leq t|t_A=t).$$

The false alarms influence the predictive value of an alarm. The predictive value requires a distribution of  $\tau$ , which is often assumed geometric, an assumption which is also made here. The intensity  $v$  is estimated from the seven periods of the LDI data available. The maximum likelihood estimator of  $v$  is  $\hat{v}=1/\bar{\tau}$  where  $\bar{\tau}$  is the average time until the turn. This was determined as the time between the start of the surveillance and the values of  $\tau$ , as determined by unimodal regression (see Section 2.4). This analysis resulted in the estimate  $\hat{v}=0.2$ . A constant predictive value with respect to time is desirable since this allows the same action to be taken whenever an alarm occurs. In the case of the LR method, Frisén and Wessman (1999) found the predictive value to be relatively constant.

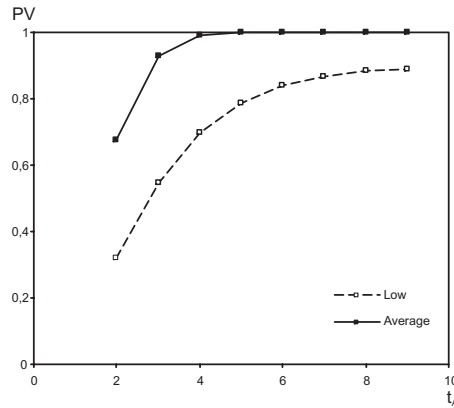


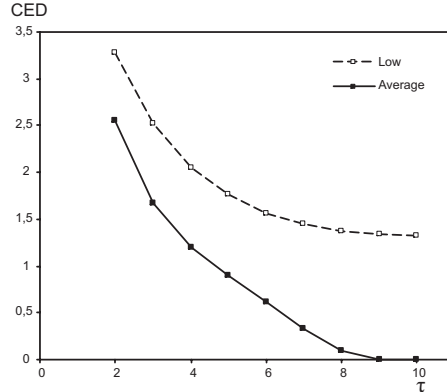
Figure 4: Predictive value,  $PV(t)$ , as a function of the time of alarm,  $t_A=t$ .

For the average curve, PV is very high except possibly for an alarm which occurs immediately after the start of the surveillance, as is seen in Figure 4. At the start there is a high risk of false alarms, which gives a lower predictive value for the low curve.

Even though the false alarm tendency will rapidly diminish because of the exponential growth, this will not have any dramatic effect on the detection ability. The reason is that when the peak value of  $\mu$  is high (due to a fast growth), the decline immediately after the peak will also be large, resulting in an easily detected decline. The evaluation of the detection ability can be made in various ways. One measure is the conditional expected delay

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

As is seen in Figure 5 the conditional expected delay, CED, is very short for the average curve (except for a decline which occurs very early). For the low curve the detection has a longer delay. CED(t) decreases monotonically with the time of the peak. Thus, it takes less time to detect a late change.

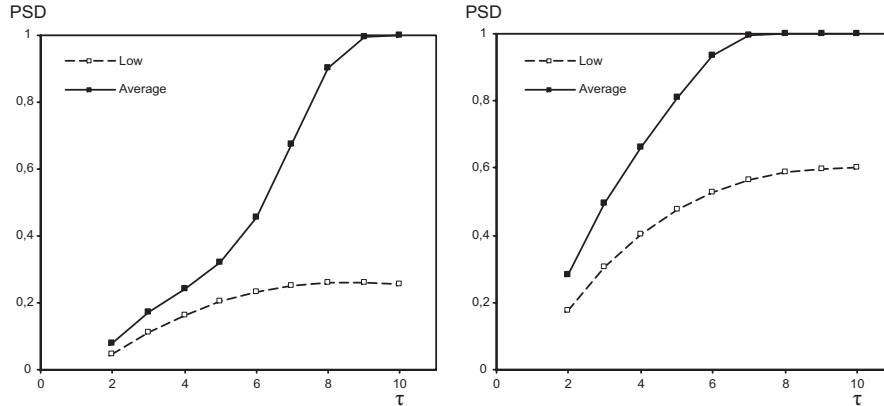


**Figure 5:** Conditional expected delay,  $CED(t)$ , as a function of the time of the peak,  $\tau=t$ .

In some situations, there is limited time available for successful action after a change. In order to evaluate a method from this point of view, the probability of successful detection (see Frisén (1992)) is useful;

$$PSD(d, t) = P(t_A - \tau \leq d \mid t_A \geq \tau, \tau = t)$$

Goldenberg et al. (2002) used the proportion of changes detected within a certain number of days. This measure is similar to PSD but does not consider the dependency on the time of the change. Here we analyze  $PSD(d, t)$  for  $d=0$  and  $1$ , which measure the probability that the detection is made at the same time as the change occurs and within one time unit after the change, respectively. The relation between the values of PSD in Figure 6 shows the same tendency as the CED: also PSD improves with a late change.



**Figure 6:** Probability of successful detection,  $PSD(d=0, t)$  (left) and  $PSD(d=1, t)$  (right) as a function of the time of the peak,  $\tau=t$ .

The conclusion is that for the average case, the properties are very good. For a lower peak, they are not as good but still acceptable.

The observed LDI curves are not all symmetric. The slope before the peak will influence the false alarms. We will here study the post-peak growth since it influences the detection probability. We compare three exponential peaks with the same pre-peak growth ( $\beta_0=3.92$ ,  $\beta_1=0.34$ ) but with different values of  $\beta_2$  ( $-0.34$ ,  $-0.23$  or  $-0.86$ ).

Since the pre-peak growth is the same for all three situations, so is the false alarm behavior. The difference in the timeliness of the motivated alarms has only a small impact on PV. The difference in PV, due to the slope of the decline, is not large, except at  $t=2$  where the slow decline ( $\beta_2 = -0.23$ ) has the smallest PV.

When the peak is non-symmetric, a slow decline after the peak gives a longer expected delay, compared to the symmetric case. A steep decline yields a shorter delay.

The conclusion is that the method has a better detection ability when the peak has a steeper decline.

The study of the effect of the variance will give information about how serious a misspecification in the method is. It is difficult to get a good variance estimate when there are no replicates (only one observation each week) and not much information from earlier studies of the process. If we use the wrong model regarding the growth, the variance will be overestimated, and if we use the same data from which the expected value is estimated, the variance will be underestimated.

The  $\sigma$  used in the method is set to 16. The value  $\sigma=16$  was also used when generating data in the simulation studies above. To address the effect of a misspecification of  $\sigma$ , we here generate observations with  $\sigma=10$ , 16 and

50 and investigate the behavior of the method in the case of the average curve. A process with a small standard deviation yields few alarms as opposed to the larger value of  $\sigma$ . Therefore the PV is higher for small values of  $\sigma$ . A large variance in the process does affect the results of the method in different ways: We will get more alarms, both false and motivated ones. Thus, the delay of an alarm will be shorter. The CED(2) is 3.22, 2.56 and 1.30 weeks for  $\sigma$  equal to 10, 16 and 50, respectively. If we wanted to control the false alarms for different values of  $\sigma$  (for example having a fixed MRL<sup>0</sup>), the situation with a small  $\sigma$  would correspond to a lower alarm limit. Then, the CED values for a small  $\sigma$  would become much better. Here we evaluate a specified method (with  $\sigma=16$ ), and the conclusion is that in a period with a smaller variance than the one specified in the SRnpPeak method, it yields a very good predictive value but a longer delay until alarm.

### 3.2 Application to Swedish Data

In order to demonstrate what can be achieved by a statistical surveillance method, we apply SRnpPeak to the LDI data for seven periods. The specification of the method is the same as in the previous section; the standard deviation in the likelihood expression is set to 16 and the alarm limit is set to 10. The surveillance starts as soon as  $LDI \geq 30$ . The results are shown in Figure 7.

Ideally, the alarm should come at the first time after the peak. This is the case for some periods, but for others there is a short delay. Retrospectively it is easy to determine the peak even “by eye”. However, the advantage of the automatic system is that it address the much harder problem of evaluating the information sequentially by using only the observations up to the decision time. In general, the signals of a good automatic sequential system should appear close to the retrospectively determined peak. A disadvantage with subjective judgements is that the predictive value and other characteristics would not be known. As was seen in the previous section, the surveillance system will yield few false alarms. If more false alarms are accepted, then the alarm signals will appear with less delay after the peak, but the predictive value will be affected negatively.

## 4 Conclusions

The theory of statistical surveillance gives us the tools for controlling the properties of systems for automatic surveillance of influenza. In this report, the aim was to suggest and evaluate systems for automatic detection of the influenza peak. The surveillance system gives a signal when there is enough information to conclude that the peak has been reached and that the decline has started. The system was designed for data on laboratory verified diagnoses of influenza.

If the method is used in an automatic warning system, it is important to remember that it only produces warnings. Medical expertise must decide on the appropriate action. The predictive value, reported here, gives information about how likely it is that there has been a peak, given an alarm at the specific time. If the summarized information indicates that the warning was a false alarm, then the computation of the alarm statistic can be continued by just ignoring the signal. On the other hand, if there is an indication of more than one peak during the influenza season this is very unusual and will motivate an epidemiological investigation to ensure that this is not caused by some influenza-like dangerous disease. In other applications, several peaks might be natural. In that case, the classification of a process in different regimes might be of interest. If so, the alarm statistic presented here can be used in a Hidden Markov Model to classify the process in parts which are increasing and decreasing, respectively.

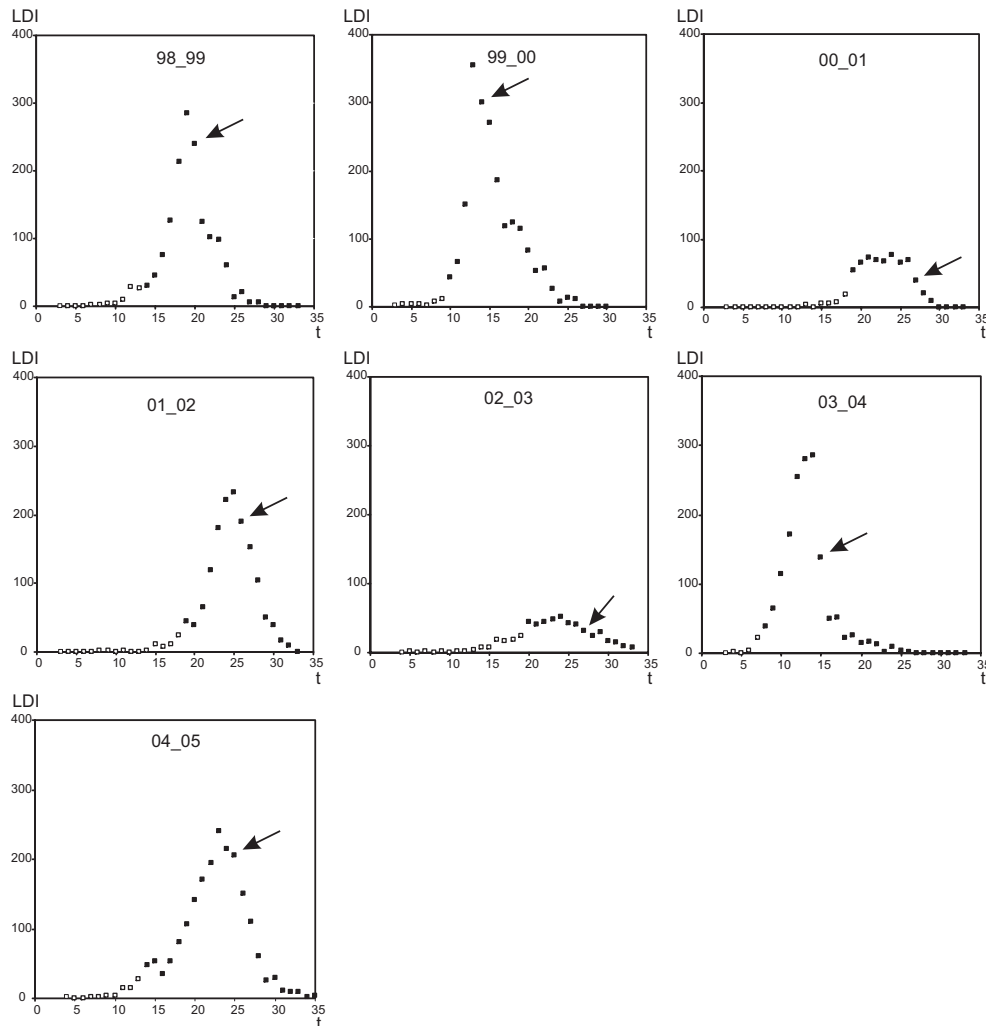
Both parametric and non-parametric surveillance methods have been discussed. Parametric methods are powerful if the assumptions are correct but might give misleading results in case of misspecification. Since the stochastic properties of the influenza incidence varied greatly between the different influenza periods, the risk of misspecification was substantial. Therefore a surveillance system based on a parametric model might not be suitable for influenza. Instead a non-parametric method, based only on the monotonicity properties of a peak, has been suggested.

The methods are based on the maximum likelihood expressions for unimodal regression and the normal distribution with a constant variance. This was the model which Andersson et al. (2008) found most appropriate near the peak of the influenza. For other situations the variance might not be constant. If the variance is known, the maximum likelihood estimator is a weighted version of the one used here (Frisén (1986)). The alarm statistic can be adjusted to the situation with a varying variance by weighting the observations with the reciprocal of the variance. In other situations, such as the onset of the influenza (Andersson et al. (2008)) it would be more appropriate to use the Poisson distribution. This can be approximated by a normal distribution with a varying variance but better still is to use the exact distribution when the maximum likelihood is derived.

The non-parametric surveillance method was evaluated for different situations relevant to influenza (an average and a low peak, non-symmetric peaks and different standard deviations). The SRnpPeak method had very good stochastic properties for the situation which mimics an average influenza period. For a lower peak the

detection ability was not as good but still acceptable. The predictive value, reflecting the trust in alarms, was higher for a small variance than for a large one. A slow decline after the peak gave a larger delay than a steep one. From the simulation studies we conclude that the method has a satisfactory ability to detect a peak and that the predictive value of a signal is good.

The non-parametric method was applied on seven periods of Swedish influenza data. The detections of the peaks were timely. However, the method can be further improved by the use of information from more data of better quality. Regional information could also be of value.



**Figure 7:** LDI data for seven periods. The observation weeks  $t=1, \dots, 33$  corresponds to calendar weeks 40 in the autumn to 20 in the spring. The transparent marks indicate the observations before the surveillance period and the black marks are the observations of the surveillance period. The alarm times are marked with arrows.

**Acknowledgements** The data were made available to us by the Swedish Institute for Infectious Disease Control, and we are grateful for discussions about the aims and the data quality. The research was supported by the Swedish Emergency Management Agency and the Bank of Sweden Tercentenary Foundation.

## References

- Andersson, E. (2002) Monitoring cyclical processes - a nonparametric approach. *Journal of Applied Statistics*, **29**, 973-990.
- Andersson, E. (2004) The impact of intensity in surveillance of cyclical processes. *Communications in Statistics-Simulation and Computation*, **33**, 889-913.
- Andersson, E. (2005) On-line detection of turning points using non-parametric surveillance: The effect of the growth after the turn. *Statistics and Probability Letters*, **74**, 433-439.
- Andersson, E., Bock, D. and Friséén, M. (2004) Detection of turning points in business cycles. *Journal of Business Cycle Measurement and Analysis*, **1**, 93-108.
- Andersson, E., Bock, D. and Friséén, M. (2005) Statistical surveillance of cyclical processes. Detection of turning points in business cycles. *Journal of Forecasting*, **24**, 465-490.
- Andersson, E., Bock, D. and Friséén, M. (2006a) Exploratory analysis of Swedish influenza data. 1:2006, Report from the Swedish Institute for Infectious Disease Control.
- Andersson, E., Bock, D. and Friséén, M. (2006b) Some statistical aspects on methods for detection of turning points in business cycles. *Journal of Applied Statistics*, **33**, 257-278.
- Andersson, E., Bock, D. and Friséén, M. (2008) Modeling influenza incidence for the purpose of on-line monitoring. *Statistical Methods in Medical Research*, to appear.
- Bock, D. and Pettersson, K. (2006) Exploratory analysis of spatial aspects on the Swedish influenza data. 3:2006, Report from the Swedish Institute for Infectious Disease Control.
- Brookmeyer, R. and You, X. (2006) A Hypothesis Test for the End of a Common Source Outbreak. *Biometrics*, **62**, 61-65.
- Buehler, J. W., Hopkins, R. S., Overhage, J. M., Sosin, D. M. and Tong, V. (2004) Framework for Evaluating Public Health Surveillance Systems for Early Detection of Outbreaks. *Morbidity and Mortality Weekly Report Recommendations and Reports*.
- Dafni, U. G., Tsiodras, S., Panagiotakos, D., Gkolfinopoulou, K., Kouvatseas, G., Tsourti, Z. and Saroglou, G. (2004) Algorithm for statistical detection of peaks--syndromic surveillance system for the Athens 2004 Olympic Games. *Morbidity and Mortality Weekly Report*, **24**, 86-94.
- Friséén, M. (1986) Unimodal regression. *The Statistician*, **35**, 479-485.
- Friséén, M. (1992) Evaluations of Methods for Statistical Surveillance. *Statistics in Medicine*, **11**, 1489-1502.
- Friséén, M. (2000) Statistical Surveillance of Business Cycles. Research Report, Department of Statistics, Göteborg University.
- Friséén, M. (2003) Statistical surveillance. Optimality and methods. *International Statistical Review*, **71**, 403-434.
- Friséén, M. (2007) Properties and Use of the Shewhart Method and Followers. *Sequential Analysis*, **26**.
- Friséén, M. and de Maré, J. (1991) Optimal Surveillance. *Biometrika*, **78**, 271-80.
- Friséén, M. and Sonesson, C. (2005) Optimal surveillance. In *Spatial surveillance for public health* (Eds, Lawson, A. and Kleinman, K.) Wiley, New York.
- Friséén, M. and Wessman, P. (1999) Evaluations of likelihood ratio methods for surveillance. Differences and robustness. *Communications in Statistics. Simulations and Computations*, **28**, 597-622.
- Gan, F. F. (1993) An optimal design of EWMA control charts based on median run-length. *Journal of Statistical Computation and Simulation*, **45**, 169-184.
- Goldenberg, A., Shmueli, G., Caruana, R. A. and Fienberg, S. E. (2002) Early statistical detection of anthrax outbreaks by tracking over-the-counter medication sales. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 5237-5240.
- Hannoun, C. and Tumova, B. (2000) Survey on influenza laboratory diagnostic and surveillance methods in Europe. *European Journal of Epidemiology*, **16**, 217-222.
- Hawkins, D. L. (1992) Detecting shifts in functions of multivariate location and covariance parameters. *Journal of Statistical Planning and Inference*, **33**, 233-244.
- Koskinen, L. and Öller, L.-E. (2003) A classifying procedure for signalling turning points. *Journal of Forecasting*, **23**, 197-214.
- Le Strat, Y. and Carrat, F. (1999) Monitoring epidemiologic surveillance data using hidden Markov models. *Statistics in Medicine*, **18**, 3463-3478.
- Linde, A., Brytting, M., Johansson, M., Wiman, Å., Högberg, L. and Ek Dahl, K. (2004) Annual Report September 2003 - August 2004: The National Influenza Reference Center. Swedish Institute for Infectious Disease Control.
- O'Brien, S. J. and Christie, P. (1997) Do CuSums have a role in routine communicable disease surveillance? *Public Health*, **111**, 255-258.
- Rath, T. M., Carreras, M. and Sebastiani, P. (2003) Automated Detection of Influenza Epidemics with Hidden Markov Models. In *Advances in Intelligent Data Analysis V* Springer-Verlag, Berlin, Germany, pp. 521-531.
- Roberts, S. W. (1966) A Comparison of some Control Chart Procedures. *Technometrics*, **8**, 411-430.
- Robertson, T., Wright, F. T. and Dykstra, R. L. (1988) *Order Restricted Statistical Inference*, John Wiley & Sons Ltd.
- Sebastiani, P., Mandl, K. D., Szolovits, P., Kohane, I. S. and Raponi, M. F. (2006) A Bayesian dynamic model for influenza surveillance. *Statistics in Medicine*, **25**, 1803-1816.
- Serfling, R. (1963) Methods for current statistical analysis of excess pneumonia-influenza deaths. *Public Health Reports*, 494-506.
- Shewhart, W. A. (1931) *Economic Control of Quality of Manufactured Product*, MacMillan and Co., London.
- Shiryayev, A. N. (1963) On optimum methods in quickest detection problems. *Theory of Probability and its Applications*, **8**, 22-46.

- 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**, 5-21.
- Woodall, W. H. (2006) The Use of Control Charts in Health-Care Monitoring and Public-Health Surveillance. *Journal of Quality Technology*, **38**, 89-134.



## Research Report

- |        |                                      |  |
|--------|--------------------------------------|--|
| 2007:1 | Andersson, E.:                       | Effect of dependency in systems for multivariate surveillance.                     |
| 2007:2 | Frisén, M.:                          | Optimal Sequential Surveillance for Finance, Public Health and other areas.        |
| 2007:3 | Bock, D.:                            | Consequences of using the probability of a false alarm as the false alarm measure. |
| 2007:4 | Frisén, M.:                          | Principles for Multivariate Surveillance.  |
| 2007:5 | Andersson, E., Bock, D. & Frisé, M.: | Modeling influenza incidence for the purpose of on-line monitoring.                |