

Published in final edited form as:

Stat Med. 2011 November 20; 30(26): 3095–3116. doi:10.1002/sim.4340.

Conditional predictive inference for online surveillance of spatial disease incidence

Ana Corberán-Vallet^{*†} and Andrew B. Lawson

Division of Biostatistics and Epidemiology, Medical University of South Carolina, 135 Cannon St, Suite 303, Charleston, SC 29425, USA

Abstract

This paper deals with the development of statistical methodology for timely detection of incident disease clusters in space and time. The increasing availability of data on both the time and the location of events enables the construction of multivariate surveillance techniques, which may enhance the ability to detect localized clusters of disease relative to the surveillance of the overall count of disease cases across the entire study region. We introduce the surveillance conditional predictive ordinate as a general Bayesian model-based surveillance technique that allows us to detect small areas of increased disease incidence when spatial data are available. To address the problem of multiple comparisons, we incorporate a common probability that each small area signals an alarm when no change in the risk pattern of disease takes place into the analysis. We investigate the performance of the proposed surveillance technique within the framework of Bayesian hierarchical Poisson models using a simulation study. Finally, we present a case study of salmonellosis in South Carolina.

Keywords

public health surveillance; spatial data; Bayesian hierarchical models; lagged loss function; conditional predictive ordinate; multiple comparisons

1. Introduction

On the past decade, following the intentional release of anthrax in the USA in 2001, there has been a growing interest in the development of statistical methodology for early detection of disease outbreaks. The ability to rapidly detect unusual aggregations of disease is of critical importance to facilitate timely public health response and, consequently, to reduce morbidity and mortality. A wide range of statistical methods including process control charts, time-series methods, and regression techniques have been proposed for surveillance of univariate time series of counts of disease [1, 2]. However, the increased need for outbreak detection has coincided with an increasing availability of data, especially data on the location of events. As a consequence, practical statistical surveillance usually implies analyzing multiple time series simultaneously. The use of spatial information promises to increase the power to detect small localized outbreaks of disease relative to the surveillance of the total count across a larger region, where sharp increases in a relatively small number of regional counts may be diluted by the natural variation associated with overall counts. In

addition, spatiotemporal surveillance facilitates public health interventions once an incident disease cluster is identified.

Unlike testing methods [3–5], modeling for spatiotemporal disease surveillance is relatively recent, and this is a very active area of statistical research [6]. Models describing the behavior of diseases in space and time allow covariate effects to be estimated and provide better insight into etiology, spread, prediction, and control of diseases. Most spatiotemporal models have been developed for retrospective analyses of complete data sets often with the objective of estimating the true relative risk of a disease of interest across a geographic study region. However, surveillance data in public health accumulate over time, and sequential estimation of the model using all the data collected so far is essential to early detection of disease outbreaks.

When small area disease data in the form of counts are available, Poisson models with a mean which is a function of the expected counts of disease and the unknown area-specific relative risks are commonly used to describe the behavior of diseases. Recently, these models have been applied in a surveillance context. Lawson *et al.* [7] proposed to monitor the model-based residuals of the fitted model in sequential time periods to detect changes in the risk pattern of a disease. More recently, Zhou and Lawson [8] have presented an approximate procedure where a spatial-only model is fitted at each time period to the current spatial data and outbreak detection is carried out by comparing the estimated relative risks for the small areas with a baseline level, which is calculated as an exponentially weighted moving average (EWMA) of historical estimates. Other statistical models proposed in a surveillance context include generalized linear mixed models [9], hidden Markov models [10], and log-Gaussian Cox processes [11] for point processes.

In this paper, we use Bayesian hierarchical Poisson models to describe the space–time behavior of disease. We show how the conditional predictive ordinate (CPO) [12, 13], which is a Bayesian diagnostic to detect observations discrepant from a given model, can be adapted in a surveillance context to detect small areas of unusual disease aggregation. For each small area and time period, we define the surveillance CPO (SCPO) as the conditional predictive density of the new observation given the data collected so far. We then carry out a parallel approach across the different areas under surveillance, where in each area alarms are sounded if the corresponding SCPO is below a specified critical level. When many areas are monitored simultaneously, it is important to deal with the problem of multiple comparisons [14]. Otherwise, because we are performing separate evaluations for each small area to determine whether the areas are significant at some fixed critical level, an inappropriate number of false alarms could be obtained. From the Bayesian viewpoint, there is no need to introduce a penalty term for performing hundreds of simultaneous tests [15]. A natural approach would be to assume that there is a common probability that a given area signals an alarm when no change in the relative risk pattern takes place. Here, we assume that the number of alarms sounded at each time point under the null follows a binomial distribution with parameters the number of small areas under surveillance and the probability of false alarm. With the aim of reducing the probability of false alarms while keeping a high outbreak detection capability, an alarm for an out-of-control system will be generated at a particular time period if the probability of observing at least the same number of small areas signaling an alarm is below a specified critical level. All the areas showing an increased disease incidence will then be reported. Finally, we show how the SCPO can be easily extended to incorporate information from spatial neighbors, which facilitates outbreak detection when changes in disease risk affect neighboring areas simultaneously.

We organize this paper as follows. In the next section, we describe the modeling framework. In Section 3, we introduce the SCPO. Section 4 deals with the multiple comparisons

problem. Section 5 describes the measures used for the evaluation of the proposed surveillance technique and shows the results obtained in a simulation study. Section 6 provides an application to reported salmonellosis cases in South Carolina from January 1995 to December 2003.

2. Spatiotemporal modeling

In what follows, we assume that we will monitor a map of m small areas over T time periods, y_{it} being the count of disease in area i , $i = 1, 2, \dots, m$, and temporal period t , $t = 1, 2, \dots, T$. When small area count data are available, it is common to consider hierarchical Poisson models. At the first level of the model, the Poisson distribution is considered for modeling the within-area variability of the counts

$$y_{it} \sim Po(e_{it}\theta_{it}) \quad (1)$$

where the expected count of disease e_{it} represents the background population effect and θ_{it} is the unknown area-specific relative risk. At the second level of the model, the logarithm of the relative risk is usually decomposed in additive components representing spatial, temporal, space–time interaction, and possible covariate effects [16]. A model could be assumed for the log relative risk of the form [17]

$$\log(\theta_{it}) = \rho + u_i + v_i + \tau_t + \delta_{it} \quad (2)$$

where ρ is the overall level of the relative risk; u_i and v_i represent respectively spatially correlated and uncorrelated extra variation; τ_t is a component describing the temporal variation; and δ_{it} represents the interaction between spatial and temporal effects in the maps. To capture the extent of variation present in count data, the prior distributions for these parameters are usually defined as follows. The prior distribution for the intercept is assumed to be zero mean Gaussian with variance σ_ρ^2 . The conditional autoregressive model proposed by Besag *et al.* [18] is used as a prior for the spatially correlated effect, that is,

$$u_i | u_{(i)} \sim N\left(\bar{u}_{n_i}, \frac{\sigma_u^2}{m_i}\right)$$

where $u_{(i)} = (u_1, u_2, \dots, u_{i-1}, u_{i+1}, \dots, u_m)$, n_i is the set of spatial neighbors of the i th region, \bar{u}_{n_i} is the mean of the neighborhood, m_i is the cardinality of n_i , and σ_u^2 is the correlated spatial component variance. Here, the neighborhood is assumed to consist of first-order spatial neighbors defined by a common boundary. Different definitions could also be considered. For instance, a region could be a neighbor if the centroid is within a certain distance of the region in question. Note that this conditional autoregressive model is useful in defining a correlated heterogeneity prior distribution. The uncorrelated spatial effect defined as

$$v_i \sim N(0, \sigma_v^2)$$

allows for uncorrelated extra variation. Because the unobserved effects within a study area could take on a variety of forms, it is always prudent to include both spatial random effects. An autoregressive prior distribution is usually defined for the temporal effect

$$\tau_t | \tau_{t-1} \sim N(\nu \tau_{t-1}, \sigma_\tau^2)$$

where ν is the autoregressive parameter and σ_τ^2 is the variance. This allows for a type of nonparametric temporal effect (random walk when $\nu = 1$). Finally, the prior distribution for the space–time interaction term can be specified as

$$\delta_{it} \sim N(0, \sigma_\delta^2)$$

where σ_δ^2 is the spatiotemporal component variance. More complex prior distributions for this interaction component can be found in [19]. However, because identifiability of highly structured interaction from main effects may be doubtful, we consider this simpler description.

Lawson *et al.* [7] and Vidal and Lawson [20] applied this spatiotemporal model in a surveillance context, where they monitored model-based residuals of the fitted model in sequential time periods to detect changes in risk. It is important to emphasize that the sequential refitting of the model presents a problem because, as time progresses, the size of the data and the parameter space will expand. Within the Markov chain Monte Carlo (MCMC) framework [21], this implies restarting the simulation process from scratch at each time period, which can be very time consuming and hampers application in real-time contexts.

In the surveillance context, as in this study, the model must be chosen that it can describe the overall behavior of disease in space and time and also that it is sensitive to temporal changes in the spatiotemporal structure. A relatively simple model capturing the normal historical variation without absorbing the changes in the model fit is advisable [22]. In this paper, we investigate the adequacy of two models for spatiotemporal disease surveillance: the convolution model originally proposed by Besag *et al.* (Besag–York–Mollie (BYM) model, [18])

$$\log(\theta_{it}) = \rho + u_i + v_i \quad (3)$$

and the spatiotemporal (ST) model defined by

$$\log(\theta_{it}) = \rho + u_i + v_i + \delta_{it} \quad (4)$$

Convolution models with both uncorrelated and correlated heterogeneity are simple models where the relative risk, which is described solely in terms of spatial effects, is assumed to be constant over time. Given their robustness and good performance in recovering the true disease relative risk surface [23,24], these models are widely used in disease mapping. The spatiotemporal model (4), while still providing a parsimonious representation, allows for a more flexible modeling of the spatiotemporal variation in disease incidence, where unusual departures from predictable patterns based on the overall spatial risk surface can be captured by the space–time interaction component. Consequently, this model is more likely to absorb changes in risk in the model fit. If time series data are available over a relatively long period and an overall time trend is present, we may include a temporal effect (model (2)) in the aforementioned equations so that we can effectively detect changes in disease incidence beyond those expected on the basis of the main spatial and temporal effects. Otherwise,

because we are interested in monitoring changes in disease risk over time, modeling the temporal variation is not advisable.

Prior distributions for the model parameters are specified as for those in model (2). Here, a fully Bayesian analysis is carried out and, consequently, prior distributions must be specified for all the hyper-parameters. We are interested in noninformative prior distributions so that inference for parameters is based only on the assumed model and the observed data. Various noninformative prior distributions have been suggested in Bayesian literature for variance parameters in hierarchical models. A common practice is to put a vague inverse Gamma(0.5,0.005) prior on the variance parameters [25]. A Gamma(0.01,0.01) prior for precision parameters is an alternative choice, which is less likely to cause identifiability problems for the two types of heterogeneity in convolution models [26]. Recently, the use of a noninformative uniform distribution has been recommended as a robust prior for standard deviation parameters [27]. In what follows, we assume a uniform distribution in the interval (0,10) for σ_ρ , whereas we use a Gamma(2,0.5) prior for the precision parameters τ_u , τ_v , and τ_δ , which provides reasonable noninformativeness.

The joint posterior distribution for the ST model, proportional to the product of the likelihood function and the prior distributions, is given by

$$\begin{aligned} \pi(\rho, u_i, v_i, \delta_{it}, \sigma_\rho^2, \sigma_u^2, \sigma_v^2, \sigma_\delta^2 | y) &\propto \prod_{t=1}^T \prod_{i=1}^m \frac{(e_{it} \theta_{it})^{y_{it}} \exp\{-e_{it} \theta_{it}\}}{y_{it}!} \\ &\times \sigma_u^{-m} \exp\left\{-\frac{1}{2\sigma_u^2} \sum_i \sum_{j \in n_i} (u_i - u_j)^2\right\} \\ &\times \sigma_v^{-m} \exp\left\{-\frac{1}{2\sigma_v^2} \sum_{i=1}^m v_i^2\right\} \\ &\times \sigma_\delta^{-mT} \exp\left\{-\frac{1}{2\sigma_\delta^2} \sum_{t=1}^T \sum_{i=1}^m \delta_{it}^2\right\} \\ &\times \sigma_\rho^{-1} \exp\left\{-\frac{\rho^2}{2\sigma_\rho^2}\right\} \\ &\times \pi(\sigma_u^2) \\ &\times \pi(\sigma_v^2) \\ &\times \pi(\sigma_\delta^2) \times \pi(\sigma_\rho^2) \end{aligned}$$

This posterior distribution is analytically intractable but can be sampled using MCMC simulation techniques. We can obtain the posterior distribution for the BYM model by removing in the previous equation the terms associated with the space–time interaction effects.

2.1. Lagged loss function

An important issue relating to the estimation of spatiotemporal models is that the size of the data and possibly the parameter space will expand as time progresses. Under conventional MCMC sampling, the available sample of parameter values at each time period must be discarded and a new sample has to be drawn by restarting the simulation process from scratch, which can be time consuming or even unfeasible. With the use of a spatial-only model, a fixed parameter space is assured. However, as new data arrive, the estimation of

the model in a surveillance context can still be relatively slow. For model (4), both the data set and the parameter space enlarge. One approach to this problem is the use of particle filtration, which resamples the past realizations from the posterior distribution to provide reweighted estimates to take account of incoming data and parameters [28]. However, these methods crucially depend on having sampled enough of the process variation at the beginning of the surveillance exercise to allow subsequent resample. Otherwise, as data become available, the ability of the samples to cover the posterior region of interest may deteriorate.

An alternative approach adopted in the estimation of spatiotemporal models in a surveillance context, which prevents refitting the whole data set and allows for a static data size, is the use of a lagged loss function or ‘sliding window’ with fixed time units within which the effects are estimated [10, 11, 17]. This mechanism, also known as batch processing [29], has been widely used in public health applications. For instance, in the univariate case, the log-linear regression model of Farrington *et al.* [30] is applied to the early detection of aberrations in laboratory-based surveillance data received at the Communicable Disease Surveillance Centre in the UK. At each time period, the estimated model using a set of recent values with similar conditions is used to predict a threshold above which an observed count is declared to be unusual.

In our particular setting, this procedure is equivalent to working with the generalized likelihood function given by

$$GL(l, t) = \prod_{k=1}^t \prod_{i=1}^m f(y_{ik}|\theta_{ik})^{\delta_k}$$

where δ_k takes the value 1 if $t - l + 1 \leq k \leq t$, l being the length of the sliding window, and the value 0 otherwise. That is,

$$GL(l, t) = \prod_{k=t-l+1}^t \prod_{i=1}^m f(y_{ik}|\theta_{ik}),$$

which corresponds to the full likelihood function when $l = t$. Let us assume that the BYM model is used to describe the relative risk surface.

Let us assume that the BYM model is used to describe the relative risk surface. Let $\Phi = \{\rho, u_1, u_2, \dots, u_m, v_1, v_2, \dots, v_m\}$ be the model parameter vector and $y_t = (y_{1t}, y_{2t}, \dots, y_{mt})'$ the $m \times 1$ data vector at time t . When the lagged loss function approach is used to estimate the model, only the most recent data are allowed to update the model through the posterior distribution

$$\pi^G(\Phi|y_{t-l+1}, y_{t-l+2}, \dots, y_t) \propto GL(l, t) \pi(\Phi). \tag{5}$$

For the ST model in Equation (4), the parameter vector at each time period is divided into two sets: $\Phi = \{\rho, u_1, u_2, \dots, u_m, v_1, v_2, \dots, v_m, \delta_{t-l+1}, \delta_{t-l+2}, \dots, \delta_t\}$ and $\Delta = \{\delta_1, \delta_2, \dots, \delta_{t-l}\}$ being the vector $(\delta_{1t}, \delta_{2t}, \dots, \delta_{mt})'$. Inferences about Φ , the parameter vector of interest, are based on the posterior distribution $\pi^G(\Phi|y_{t-l+1}, y_{t-l+2}, \dots, y_t)$. Hence, parameter estimates based on the posterior distribution (5) and those derived from the conventional posterior distribution, $\pi(\Phi|y_1, y_2, \dots, y_t) = \int \pi(y_1, y_2, \dots, y_{t-l}) \pi^G(\Phi|y_{t-l+1}, y_{t-l+2}, \dots, y_t)$, will be similar if the data within the sliding window provide the

same information as the whole history of the process. When no change in the relative risk pattern takes place, the longer the window length, the higher the estimation accuracy. However, in an online surveillance context where changes in risk are expected, we want the spatial effects to adapt to those changes quickly so that the surveillance technique can detect further changes as new data become available. As Jiang and Zhang [31] stressed, the use of a shorter window length is fundamental for discarding the out-of-date information effectively and achieving fast tracking of the changed parameters. Otherwise, windows that are too wide will smooth over possible shifts in the process. We consider different window lengths in this paper. This allows us to examine their ability to estimate and detect changes in the relative risk pattern of disease.

In an effort to overcome the estimation problem arising when hierarchical Poisson models are used in a surveillance context, Zhou and Lawson [8] presented an approximate procedure where the BYM model is fitted at each time period to the current spatial data. Changes in the risk pattern are then detected by comparing the estimated relative risks for each of the small areas with a baseline level calculated as an EWMA of historical estimates. There is, however, only heuristic justification for the proposed procedure. First, a sliding window of length 1 cannot guarantee accurate model estimates, specially for large Poisson mean values. In addition, an EWMA approach is only justified under Gaussian model assumptions.

3. Monitoring changes in relative risk patterns

As time progresses, a decision concerning whether the disease incidence has increased has to be made sequentially on the basis of the data collected so far. That is, at each time point, we are interested in evaluating how unusual the observations for the new time period are. A commonly used approach when addressing a problem involving both spatial and temporal components is to perform a parallel surveillance for each location and sound a general alarm when there is an alarm for any of the small areas. For instance, Höhle *et al.* [32] applied the Farrington algorithm [30] for the monitoring and surveillance of multivariate count time series in veterinary public health. An alarm for the whole system is generated if there is at least one observation above the threshold predicted by the corresponding univariate algorithm. A reduction of the multivariate surveillance problem is an alternative technique. In this case, either the observations at each time period or the alarm statistics of the parallel surveillance methods are transformed into a scalar statistic, which will be monitored over time. For a global increase in disease incidence, Wessman [33] showed that there exists a sufficient univariate reduction of the multivariate observation process. Hence, without loss of information, univariate surveillance methods can be applied to monitor the resultant scalar statistic. Likelihood-ratio-based surveillance procedures [34] have also been proposed to jointly handle correlated time series count data (see [35] and the references therein for a detailed overview of multivariate surveillance). Determining what type of change we aim to detect becomes then a crucial issue. For instance, a method derived to detect a global change may not be capable of detecting a small cluster of increased disease incidence.

In our surveillance context, we believe that a global increase in the incidence in all the areas occurring at the same time point is unlikely, and so our primary concern is to design a surveillance technique that, at each time period, alerts us to any small area of increased disease incidence in need of further investigation. Therefore, we propose to carry out a parallel approach across the different areas under surveillance, where different signals will be generated for those areas of unusual disease aggregation. By making the univariate surveillance tests dependent on the fitted spatiotemporal model, we overcome the information loss problem associated with the parallel approach, which leads to insufficient

and suboptimal surveillance methods [36]. In addition, we focus here on detection of sudden jumps in risk, and so no accumulation of information over time is made (for detection of an increasing incidence, see for instance, [37]).

A possible approach within this framework is the analysis of residuals. Lawson *et al.* [7] introduced surveillance residuals as the difference between the observed data for the new time period and the data predicted under the model when it is fitted using the data from previous periods. The absolute value of the surveillance residuals (summed over space) is monitored over time to detect changes in risk. Once a change is detected, the p -value surface is analyzed to determine the location and extension of possible outbreaks. For each small area and time point, the Bayesian p -value is defined as the probability that replicated data, based on the posterior predictive distribution, could be more extreme than the observed data, that is,

$$p_{it} = \Pr(y_{it}^* > y_{it} | \text{data}),$$

extremely small p -values indicating that the observation for the new time period is not representative of the data expected under the fitted model. Note that the equivalent of a parametric bootstrap [38] in the Bayesian setting is necessary to generate the set of simulated counts $\{y_{it}^*\}$, which can be cumbersome on some occasions. In addition, fitting the model with data in sliding windows leads to misleading results because, due to censoring effects, bigger residuals at the start of the surveillance window can be confounded with changes in risk [20].

We do not pursue this approach here. Instead, we show how the conditional predictive ordinate [12] can be adapted for a surveillance context to detect small areas of unusual disease aggregation.

3.1. Surveillance conditional predictive ordinate

Box [39] emphasized the complementary roles of the posterior and predictive distributions arguing that the posterior distribution provides a basis for estimation of parameters conditional on the adequacy of the entertained model, whereas the predictive distribution enables criticism of the entertained model in light of current data. Geisser [12] introduced the CPO as a Bayesian diagnostic to detect observations discrepant from a given model. Let $y = (y_1, y_2, \dots, y_n)'$ be the vector of conditionally independent observations with density $f(y_i | \theta)$ under model \mathcal{M} and $y_{(i)} = (y_1, y_2, \dots, y_{i-1}, y_{i+1}, \dots, y_n)'$ the data vector with y_i deleted. For the i th observation, the CPO is defined as

$$\text{CPO}_i = f(y_i | y_{(i)}) = \int f(y_i | \theta, y_{(i)}) \pi(\theta | y_{(i)}) d\theta,$$

small CPO values indicating a poor fit by the model. Gelfand *et al.* [13] further discussed the CPO, proposing a cross-validation approach which is based on conditional predictive distributions arising from single observation deletion to address model determination. A

surrogate for the marginal likelihood $f(y)$ can be defined as $\text{PML} = \prod_{i=1}^n f(y_i | y_{(i)})$ [40], which leads to the pseudo-Bayes factor that can be used as an alternative strategy to overall model selection. For the purpose of model assessment, the presence of many small CPOs criticizes the model.

Let $y_t = (y_{1t}, y_{2t}, \dots, y_{mt})'$ be the $m \times 1$ data vector at time point t , $y_{1:t} = (y'_1, y'_2, \dots, y'_t)$ the vector of all the observations up to time t , and $\theta_t = (\theta_{1t}, \theta_{2t}, \dots, \theta_{mt})'$ the relative risk vector at time t . In our surveillance context, the CPO at each time period

$$CPO_t = f(y_t | y_{1:t-1}) = \int f(y_t | \theta_{t-1}, y_{1:t-1}) \pi(\theta_{t-1} | y_{1:t-1}) d\theta_{t-1}$$

can be used to detect an overall process change. However, as a global measure, localized clusters of insufficient size may be missed because of dilution effects.

In this paper, we propose to modify the CPO to obtain what we call the surveillance CPO (SCPO), which is computed for each new observation using only the data from previous time periods. Otherwise, the inclusion of observations from the new time period may lead to a different model for the relative risk pattern. In particular, we define the SCPO for the i th small area and time point t as

$$SCPO_{it} = f(y_{it} | y_{1:t-1}) = \int f(y_{it} | \theta_{i,t-1}, y_{1:t-1}) \pi(\theta_{i,t-1} | y_{1:t-1}) d\theta_{i,t-1} \tag{6}$$

In general, $f(y_{it} | \theta_{i,t-1}, y_{1:t-1})$ can be computed as

$$f(y_{it} | \theta_{i,t-1}, y_{1:t-1}) = \int f(y_{it} | \theta_{it}) \pi(\theta_{it} | \theta_{i,t-1}, y_{1:t-1}) d\theta_{it}$$

where $f(y_{it} | \theta_{it})$ is Poisson and $\pi(\theta_{it} | \theta_{i,t-1}, y_{1:t-1})$ can be derived from the model describing the relative risk surface. For the spatiotemporal model (4), $\log(\theta_{it})$ can be stated as $\log(\theta_{i,t-1}) + \Delta \delta_{it}$, where $\Delta \delta_{it} = \delta_{it} - \delta_{i,t-1}$ represents the variation in risk at time t . In our surveillance context, we can assume that $\Delta \delta_{it} = 0$, that is $\theta_{it} = \theta_{i,t-1}$, if no change in risk takes place at time t . On the basis of this approximation, a Monte Carlo estimate for the SCPO, which does not have a close form, can be obtained from a posterior sampling algorithm as

$$SCPO_{it} \approx \frac{1}{J} \sum_{j=1}^J Po(y_{it} | e_{it} \theta_{i,t-1}^{(j)}) \tag{7}$$

where $\{\theta_{i,t-1}^{(j)}\}_{j=1}^J$ is a set of relative risks sampled from the posterior distribution that corresponds to the previous time period. Hence, if there is no change in risk y_{it} will be representative of the data expected under the previously fitted model. Otherwise, SCPO values close to zero will be obtained. When the BYM model is assumed to describe the relative risk surface, the constant risk θ_i will be considered instead of the relative risk at the previous time point. If the lagged loss function approach described in Section 2.1 is used, the model describing the structure of the risks will be estimated using only the data in the interval $[t-l, t-1]$, l being the length of the sliding window.

As we have mentioned before, one advantage of this surveillance technique is that it depends on the joint model for the multivariate time series, which accounts for the spatial correlation in the data. Yet, one limitation of the proposed approach is that it may fail to detect both emerging outbreaks of disease occurring in some neighboring areas which are small in magnitude to be detected by individual tests or outbreaks occurring on the border between areas, which may be missed because they do not conform to the predefined geographical areas. On those occasions, a different technique using information from spatial neighbors may facilitate the outbreak detection work. For the i th small area, let $n_i = \{i_1, i_2, \dots, i_{m_i}\}$ be

the set of first-order spatial neighbors, m_j being the cardinality of n_j . We propose to consider a multi-variate extension of the SCPO incorporating information from the spatial neighborhood as an alternative surveillance technique to the univariate SCPO. This multivariate extension, which is denoted here as clustered SCPO (CSCPO), is defined as

$$\begin{aligned} \text{CSCPO}_{it} &= f(y_{it}, y_{i_1t}, y_{i_2t}, \dots, y_{i_{m_i}t} | y_{1:t-1}) \\ &= \int \int \dots \int f(y_{it}, y_{i_1t}, y_{i_2t}, \dots, y_{i_{m_i}t} | \theta_{i,t-1}, \theta_{i_1,t-1}, \theta_{i_2,t-1}, \dots, \theta_{i_{m_i},t-1}, y_{1:t-1}) \\ &\quad \times \pi(\theta_{i,t-1}, \theta_{i_1,t-1}, \theta_{i_2,t-1}, \dots, \theta_{i_{m_i},t-1} | y_{1:t-1}) d\theta_{i,t-1} d\theta_{i_1,t-1} d\theta_{i_2,t-1} \dots d\theta_{i_{m_i},t-1} \quad (8) \\ &= \int \int \dots \int f(y_{it} | \theta_{i,t-1}) \prod_{k=1}^{m_i} f(y_{i_kt} | \theta_{i_k,t-1}) \\ &\quad \times \pi(\theta_{i,t-1}, \theta_{i_1,t-1}, \theta_{i_2,t-1}, \dots, \theta_{i_{m_i},t-1} | y_{1:t-1}) d\theta_{i,t-1} d\theta_{i_1,t-1} d\theta_{i_2,t-1} \dots d\theta_{i_{m_i},t-1}, \end{aligned}$$

values of the CSCPO close to 0 indicating clusters of neighboring areas of unusual disease incidence. Once a sample of relative risk values has been drawn from the posterior distribution that corresponds to the previous time period, $\pi(\theta_{1,t-1}, \theta_{2,t-1}, \dots, \theta_{i,t-1}, \dots, \theta_{m,t-1} | y_{1:t-1})$, the CSCPO can be approximated as

$$\text{CSCPO}_{it} \approx \frac{1}{J} \sum_{j=1}^J Po(y_{it} | e_{it} \theta_{i,t-1}^{(j)}) Po(y_{i_1t} | e_{i_1t} \theta_{i_1,t-1}^{(j)}) \dots Po(y_{i_{m_i}t} | e_{i_{m_i}t} \theta_{i_{m_i},t-1}^{(j)})$$

3.2. Monitoring measures

Effective measures based on the SCPO values have to be constructed to assess the system state each time new observations become available. For the univariate SCPO, a possible approach is to trigger an alarm for the i th small area if the corresponding SCPO is below a specified critical level. However, setting the same cutoff value for all the areas may lead to misleading results because the value of the SCPO depends on the mean of the Poisson distribution. For small values of the mean, the probability mass of the Poisson distribution is tightly concentrated near zero. As the mean and, consequently, the variance increase, the distribution becomes less concentrated around the mean and the probability associated with each value decreases. Here, we propose to scale the SCPO by dividing by the value of the posterior predictive density at the posterior mean of the Poisson distribution, that is,

$$\text{sSCPO}_{it} = \frac{\text{SCPO}_{it}}{f(e_{it} \widehat{\theta}_{i,t-1} | y_{1:t-1})} \quad (9)$$

where $\widehat{\theta}_{i,t-1}$ is the posterior mean of the relative risk at the previous time period. Hence, the scaled SCPO will take values close to 1 if the observation at time t is close to the data expected under the fitted model and values close to 0 otherwise. An alarm for the i th small area will be triggered at time t if $\text{sSCPO}_{it} < \alpha$ and $y_{it} > e_{it} \widehat{\theta}_{i,t-1}$. Notice that the first condition detects discrepant observations, and the second one prevents us from signaling an alarm for those areas of decreased disease incidence.

If the CSCPO is used instead, an alarm will be sounded for the i th small area at time t if the corresponding scaled CSCPO $_{it}$ given by

$$\text{sCSCPO}_{it} = \frac{\text{CSCPO}_{it}}{f(e_{it} \widehat{\theta}_{i,t-1}, e_{i_1t} \widehat{\theta}_{i_1,t-1}, \dots, e_{i_{m_i}t} \widehat{\theta}_{i_{m_i},t-1} | y_{1:t-1})} \quad (10)$$

is below a specified critical level, $y_{it} + \sum_{k=1}^{m_i} y_{ikt} > e_{it} \widehat{\theta}_{i,t-1} + \sum_{k=1}^{m_i} e_{ikt} \widehat{\theta}_{i,t-1}$ and $y_{it} > e_{it} \widehat{\theta}_{i,t-1}$. In this case, the first two conditions determine if the total number of counts of disease in the neighborhood is significantly higher than expected. The last condition is necessary to properly identify those areas with more counts of disease than expected in a neighborhood of increased disease incidence.

4. Multiple comparisons problem

In the spatiotemporal disease surveillance setting, it is important to consider the problem of multiple comparisons. Otherwise, because we are performing separate evaluations for each small area to determine whether the areas are significant at some fixed level α , the probability of false alarm will be close to 1 for large values of m . Benjamini and Hochberg [41] proposed to control the false discovery rate, which is defined as the expected proportion of false rejections among all rejected hypothesis. This procedure, which has proved to be a more powerful approach than the Bonferroni correction to address the problem of multiplicity, has been the object of numerous studies in recent years [42,43].

From a Bayesian viewpoint, even though there is no need to introduce a penalty term for performing multiple comparisons simultaneously [15], the multiple comparisons problem has to be carefully addressed to assure a good performance of the surveillance procedure. Yet, little work has been done that deals with multiple testing in spatial and spatiotemporal surveillance. Kleinman *et al.* [9] suggested not reporting the individual p -values directly but the number of tests that would be required to see a p -value as small as the one observed so that the user focuses on the more alarming values. However, this approach does not prevent the signaling of an alarm at every time period. Zhou and Lawson [8] proposed a parallel approach which signals an alarm for an out-of-control system if any small area is significantly anomalous. For a critical level equal to 0.05, the probability of signaling a false alarm when no change in risk takes place can be higher than 0.5. By reducing the value of the critical level for the individual comparisons, the probability of a false alarm is reduced but also the power of the method for outbreak detection. Besides, on those occasions, although the surveillance procedure may be able to detect a small area of increased disease incidence and trigger an alarm, it would hardly recover the true outbreak behavior.

A natural Bayesian approach is to assume that there is a common probability that a given area signals an alarm when no change in risk takes place, that is,

$$\alpha^* = \Pr(\text{sSCPO}_{it} < \alpha \quad \& \quad y_{it} > e_{it} \widehat{\theta}_{i,t-1} | \text{no change})$$

α being the specified critical level for the individual comparisons. We can then model the number of alarms sounded at each time period under the null by the binomial distribution with parameters m , the number of small areas under surveillance, and α^* , the probability of each area signaling a false alarm. Let n_t be the number of small areas signaling an alarm at time t . In an effort to reduce the probability of false alarm while keeping a high outbreak detection capability, we propose to evaluate the probability of observing n_t or more alarms, that is,

$$p_t = 1 - \sum_{n=0}^{n_t-1} \binom{m}{n} (\alpha^*)^n (1 - \alpha^*)^{m-n},$$

and generate an alarm for an out-of-control system if this probability is below a critical level β , which controls the probability of false alarm at each time period. It is important to emphasize here that we consider two different alarm strategies and, consequently, two different critical levels. First, on the basis of the sSCPO and α , alarms are sounded for those areas of increased incidence. If the number of unusual observations is significant according to β , an alarm for an out-of-control system will be triggered and all the alarms associated with small areas of unusual aggregation of disease will then be reported. By setting $\beta = 1$, an alarm for an out-of-control system will be triggered whenever there is, at least, one small area significantly anomalous. So, this approach can be seen as a generalization of previous approaches. Because we do not reduce the critical level of the individual comparisons, an important advantage of this procedure is that it allows us to detect those small areas of increased disease incidence, and so it is likely to recover the true outbreak behavior.

From a fully Bayesian viewpoint, a tri-level hierarchical model allowing for joint estimation of the model parameters and the probability of a false alarm α^* would be advisable [44]. However, because the probability of interest depends on the SCPO, this is not a feasible procedure. Note that for each area and time period, a set of relative risks sampled from the posterior distribution that corresponds to the previous time period is required to compute the SCPO (Equation (7)). A possible solution is to apply the surveillance procedure to a set of historical data under the null, that is, when there is no change in risk, to estimate the number of alarms at each time period and, therefore, the probability of false alarm.

This approach can also be applied when working with the CSCPO, which is computed for each small area taking into account the information from the spatial neighborhood. In that case, α^* represents the probability that a given area signals a false alarm based on the CSCPO.

5. Simulation study

To assess the effectiveness of the proposed surveillance procedure for outbreak detection, we conducted a simulation study. In this section, we show a summary of the results obtained in four different surveillance scenarios mimicking possible disease outbreak patterns.

5.1. Data

We used the US state of Georgia, which consists of $m = 159$ counties, as the base map to generate the observed disease count data at county level for $T = 20$ time periods. Having a study region with a large number of subregions is ideal because it allows a wide range of different outbreak sizes to be simulated. We used the number of inpatients from asthma in 2007, obtained from the Georgia Department of Community Health, to calculate the expected counts for the mapped area. Note that we assumed a constant expected count rate over time, which is reasonable if we assume the population does not change during the surveillance exercise. The importance of evaluating surveillance procedures using epidemiologically and spatially appropriate data has been stressed in literature. Asthma is a convenient choice because, as one of the most common chronic diseases in the USA, it has become a major public health challenge.

The model considered for the true relative risks under the null was

$$\theta_{it} = \exp\{\rho + v_i + \delta_{it}\}; \quad i=1, 2, \dots, m; t=1, 2, \dots, T$$

where ρ is zero mean Gaussian with precision τ_ρ ; (v_1, v_2, \dots, v_m) is assumed to be a realization of a multivariate Gaussian distribution with zero mean vector and precision

matrix $\tau_v I_m$; and each $\delta_i = (\delta_{i1}, \delta_{i2}, \dots, \delta_{iT})$ is assumed to follow a random walk independently of all other counties, that is, $\delta_{it} \sim \mathcal{N}(\delta_{i,t-1}, \tau_\delta)$; we set the precision parameters τ_ρ , τ_v , and τ_δ equal to 10,000, 100, and 2500 respectively, which produced a small variation in disease risk over both space and time.

We considered four different scenarios representing possible changes in the relative risk pattern over space and time: isolated local jumps, clustered local jumps, and a global jump (Figure 1). As we have mentioned before, a global change affecting all the counties at the same time point is unlikely. Yet, it is important to assess the behavior of our surveillance technique in that scenario. We assumed that irreversible changes occurred at time period $t = 18$ so that enough historical data were available to evaluate the effect of the window length on the performance of the procedure.

Scenarios 1 and 2. Local jumps in risk were generated for nine counties at time $t = 18$. In scenario 1, a jump of 60% is defined as

$$\theta'_{it} = 1.6 \theta_{it} \quad i \in R_0; t \in T_c$$

where $T_c = \{18, 19, 20\}$. Note that this jump is proportional to the relative risk. Consequently, those counties with high values of the relative risk will undergo more noticeable jumps. However, the contrary may occur on some occasions. In scenario 2, the simulated jumps took the form

$$\theta'_{it} = \theta_{it} + 0.5 \quad i \in R_0; t \in T_c$$

so local jumps of different intensities, varying between 28% and 75%, were generated, the counties with the smallest relative risks being the ones undergoing the biggest increases.

Scenario 3. Clustered jumps in risk of different intensities were generated for three different regions

$$\begin{aligned} \theta'_{it} &= 1.4 \theta_{it} & i \in R_1; t \in T_c \\ \theta'_{it} &= 1.5 \theta_{it} & i \in R_2; t \in T_c \\ \theta'_{it} &= 1.3 \theta_{it} & i \in R_3; t \in T_c \end{aligned}$$

Scenario 4. An irreversible global jump in risk of 30% was generated at time period $t = 18$

$$\theta'_{it} = 1.3 \theta_{it} \quad i = 1, 2, \dots, m; t \in T_c$$

Once the values for the expected counts and the relative risks were specified, the observed counts in the mapped area were generated using the Poisson model

$$y_{it} \sim \text{Po}(e_{it} \theta'_{it}) \quad i = 1, 2, \dots, m; t = 1, 2, \dots, T$$

To allow for sampling variability, we simulated 500 data sets for each of the relative risk patterns. After experimenting with a range of different number of replications, we chose this number as a balance choice between accuracy and computational efficiency. The results presented are averaged over these 500 realizations.

5.2. Evaluation of the surveillance method

It is not evident how to measure the performance of a multivariate surveillance method where repeated decisions are made over time. A trade-off between false alarms and short delay times or detection probability must be struck, so a single optimality criterion is not always enough and evaluation by several measures might be necessary to characterize the behavior when the process is in and out of control.

The surveillance method developed in this paper is envisaged as a screening instrument that allows us to identify small areas with a sudden jump in risk and, consequently, prompting further investigation. Hence, sensitivity and specificity are advisable measures. Let us divide the results obtained for each scenario during the outbreak period T_c into true alarms (TA), false alarms (FA), false no alarms (FNA), and true no alarms (TNA) (Table I). In particular, we define the sensitivity as the proportion of all the areas undergoing a change in risk which signal an alarm at any time during the outbreak period, that is, $TA / (TA + FNA)$. Similarly, the specificity is given by the proportion of in-control regions which are correctly identified as such, $TNA / (FA + TNA)$. Note that our definition of sensitivity and specificity may differ from others proposed in a surveillance context (see, for example, [8, 10]). The reason is that our primary concern is in detecting all the areas of increased disease incidence. Hence, we are more concerned with recovering the true outbreak behavior, rather than just detecting a small area of unusual disease aggregation when a change in risk takes place. It is important to mention that when a sliding window approach is used to estimate the model describing the behavior of disease, the proposed sensitivity and specificity are time independent. For each time point $t \in T_c$, the spatiotemporal model is fitted using only the data observed between $t - l$ and $t - 1$, l being the length of the sliding window. Unusual observations at time t are then detected by the corresponding SCPO.

As Sonesson and Bock [36] pointed out, timeliness is of extreme interest in public health surveillance, and so additional measures accounting for timeliness in detection should be considered together with the common test characteristics of sensitivity and specificity. In this study, we consider the median time to outbreak detection (MTD), which is defined as follows. For each small area undergoing an outbreak, we define the time to detection as the number of time periods from the beginning of the outbreak until the first alarm is triggered, the outbreak taking place at any time during the surveillance period. If no alarm is sounded, that is, if the outbreak is missing, an infinite time to detection is assigned. The MTD is then given by the median of the times to detection corresponding to those areas of increased incidence. Consequently, an MTD equal to infinite means that the surveillance technique has not detected, at least, half of the areas of increased incidence. This measure avoids some of the problems arising when using the mean time to detection: if missed outbreaks are ignored when computing the mean, a method that alerts late can have a larger mean than a method that does not alert at all.

Finally, to characterize the behavior of the surveillance mechanism when the process is in control, we consider the probability of a false alarm for an out-of-control system (PrFA) when the proposed procedure for addressing the multiple comparisons problem is applied. This probability can be estimated by the total number of time points associated with a false alarm divided by the total number of non-outbreak time periods. The commonly used average in-control run length (ARL^0) can be estimated from the PrFA taking into account their relationship under the null when evidence against the in-control situation is not aggregated over time. We do not present ARL^0 but favor PrFA here.

5.3. Results

We sequentially fitted the two models described in Section 2, the BYM and the ST models, as new data became available. We considered different window lengths to assess their influence on model estimation and outbreak detection capability. We carried out the posterior sampling using mixed Gibbs–Metropolis steps. We fixed a burn-in period of 25,000 iterations to assess the convergence of MCMC chains. To reduce the correlation for the samples, we kept one posterior sample in five iterations after the burn-in until a set of 5000 iterations was obtained. We performed a sensitivity analysis to check whether the prior assumptions had an undue effect on the results. We experimented with a range of different hyperprior specifications (Section 2), all of them providing similar results. The results presented here correspond to the case of a $\text{Gamma}(2,0.5)$ prior for the precision parameters. Because of the huge amount of information, we present a selection of the most representative results in this section.

For a range of different window lengths, Table II displays the sensitivity, specificity, and MTD of the proposed surveillance technique in scenarios 1 and 3 when we use the BYM model to describe the behavior of disease. Note that we obtain one measure value for each simulation scenario and data set. The results presented here are averaged over the 500 realizations simulated for each scenario. For the MTD we do not present the mean value but favor the median here. If the mean is considered, a MTD equal to infinite is obtained if there is just one data set whose corresponding MTD equals infinite. To assess the variability of the results, we present the 95% credible intervals (CI) together with posterior average estimates. It should be pointed out that, although the specificity remains the same, the sensitivity of the surveillance technique improves as the window length increases. This is due to the fact that unusual counts of disease can be better detected as we increase the amount of historical observations under non-outbreak periods. However, the use of historical data beyond 6 time periods back does not have a significant impact on the performance of the surveillance technique. Although we do not show it here, we obtain similar results for scenarios 2 and 4 and for the ST model. Hence, we recommend the use of a sliding window, which prevents refitting the whole data set and reduces computational requirements. This is specially important for Bayesian hierarchical model-based surveillance methods because sequential estimation of the model using MCMC simulation techniques can be time consuming. In addition, once an increased disease incidence has been detected, we want the surveillance mechanism to adjust to the new risk pattern so that it can look for further anomalies. Otherwise, the use of the whole history will lead to multiple sequential alarms for the detected outbreak, and so it may hinder identification of additional changes in the disease incidence. In this study, we can recommend a length of 6 time units as an appropriate choice. However, the optimal cutoff value will be application specific and different lengths should be considered depending on the amount of data available and the relative risk pattern.

Table III shows the sensitivity, specificity, and MTD of the proposed surveillance technique by simulation scenario when both the BYM and the ST models are fitted to the data within a sliding window of length 6 time periods. Because we can only compute specificity for those situations where changes in risk do not affect all the areas under surveillance, we present only the sensitivity and MTD for scenario 4. An important result of our simulation study is that the SCPO based on the BYM model provides more satisfactory results. This is due to the fact that the constant area-specific relative risks in the BYM model are estimated so that they can accommodate all the observed counts within the sliding window. Changes in risk will then be detected taking into account all the previously observed data. On the contrary, when the ST model is used to describe the relative risk surface, detection of changes in risk will depend mainly on the observed data for the last time period. Observations bigger than the ones observed in previous time points will lead to higher estimates of the corresponding

relative risks, which may hinder the outbreak detection capability of the procedure. Similarly, counts of disease smaller than the ones previously observed may result in false alarms. These results support the idea that simple models describing the overall behavior of disease in space and time are advisable in surveillance contexts. In what follows, we will use the BYM model to describe the relative risk pattern.

Figure 2 shows the spatial distribution of the sSCPO at time 18 when the BYM model is used to describe the behavior of disease. It can be seen that counties with low sSCPO values correspond to those where changes in risk were generated (compare Figure 2 with Figure 1). Counties with no simulated increase in risk have much larger sSCPO, all of them greater than 0.5. It should be emphasized here that, rather than adding more counts directly, we simulated jumps in the relative risks and then we generated the counts of disease using the Poisson model and the simulated risks. As a consequence, generated counts corresponding to those areas with a jump in risk may not present an increased disease incidence from the beginning of the outbreak.

In Table IV, we compare the detection power of both the univariate and clustered SCPO (Equations (6) and (8), respectively) by simulation scenario. For comparative purposes, we also present the results obtained with two alternative surveillance procedures: the Bayesian p -value surface [7, 20] and the widely used log-linear regression model developed by Farrington *et al.* [30], which is implemented in the R software package *surveillance* [45]. We set the corresponding critical levels for these techniques equal to 0.10 for the scaled SCPO, 0.0025 for the scaled CSCPO, 0.01 for the Bayesian p -value, and 0.04 for the Farrington algorithm. Thresholds for detecting areas of unusual risk have been proposed in a nonsurveillance context [46, 47]. The thresholds here were chosen so that a specificity around 0.93 was assured in all the scenarios, that is, 93% of the small areas without a jump in risk during the outbreak period were properly identified as so by the surveillance techniques. For all the techniques and simulation scenarios, the 95%CI for the specificity was within the interval [0.87, 0.97]. Let us comment the results. The analysis of the Bayesian p -value surface, as our surveillance approach, requires the use of a spatiotemporal model describing the behavior of disease over space and time under ‘normal’ conditions. Although a spatiotemporal model as in Equation (2) was originally proposed [7], it can be seen that the use of the BYM model provides more satisfactory results. For each small area and time period, the Bayesian p -value is given by the probability that replicated data from the posterior predictive distribution could be more extreme than the new observation, that is, the Bayesian p -value looks at the area under the posterior predictive distribution from the observed value. So, the Bayesian p -value and the SCPO, which corresponds to the value of the posterior predictive density at the new observation, are based on the same predictive distribution and, consequently, similar results are expected with both techniques although at different critical levels. For those observations above the mean, the smaller the value of the posterior predictive density, the smaller the area of the upper tail. The main advantage of the SCPO is that it can be easily extended to incorporate information from the spatial neighborhood. The use of the CSCPO improves considerably the outbreak detection capability when changes in risk affect neighboring areas simultaneously. This is an important result that should be kept in mind because surveillance procedures usually focus on either individual tests for each of the small areas in the map or a scalar statistic which is a function of all the individual tests. Finally, as we have mentioned before, a common approach for the surveillance of multivariate time series of disease counts is to apply a univariate method to each time series independently and report all the alarms. The Farrington algorithm is routinely applied in public health applications to detect aberrations in laboratory-based surveillance data. Recently, this method has been applied in the analysis of multivariate surveillance data arising in veterinary public health [32]. However, because correlation between the series is ignored, more satisfactory results can be obtained by using

a multivariate surveillance technique. It is important to note here that we are just comparing the detection capability of the surveillance techniques during the outbreak period for a given value of the specificity. In practice, when the time series are monitored separately by univariate techniques, higher thresholds would be set as the number of small areas under surveillance increases so as to keep in-control alarm rates comparable with those in univariate surveillance contexts. For instance, one would set the critical level of the Farrington algorithm at $\alpha = 0.01/m$ to control the PrFA at 0.01. In this case, the sensitivity of the surveillance procedure will decrease considerably. As shown next, the procedure proposed in this paper to control the probability of false alarm provides an alternative satisfactory solution to the problem of multiple comparisons.

Let us assume that an alarm for the i th small area is sounded at time t if the observed count of disease is higher than expected and the corresponding sSCPO is below the critical level $\alpha = 0.10$. Table V displays, for each scenario, the probability of observing at least the same number of small areas signaling an alarm at each time point during the outbreak period. Note that this probability depends on the number of small areas undergoing an outbreak. In addition, the number of small areas signaling an alarm decreases as time progresses. This is due to the fact that higher estimates of the area-specific relative risks are obtained as data from outbreak time periods are included in the sliding window. Let us assume that an alarm for an out-of-control system is generated at a particular time point if this probability is below the critical level $\beta = 0.01$, which allows us to control the probability of a false alarm for the whole study region when no change in risk takes place at 0.01. As the results in Table V demonstrate, an alarm for the whole study region will be generated as soon as a jump in risk takes place with probability 1 for all the scenarios. By setting the critical level $\alpha = 0.10$, 93% of the areas with a jump in risk are then reported as areas of increased disease incidence in scenario 1. For scenarios 2–4, the percentage of correctly detected areas are respectively 0.90, 0.77, and 0.53 (Table III). The probability of false alarm (PrFA) can be estimated using the 500 data sets simulated for each scenario under normal conditions. Let $N_t = 22000$ be the total number of non-outbreaks time periods (4 scenarios \times 500 data sets \times 11 time periods) and N_{fa} the total number of points time associated with a false alarm. The estimated PrFA = N_{fa}/N_t is equal to 0.006, smaller than 0.01 as expected.

6. Case study

In this section we study the number of reported salmonellosis cases in South Carolina during 1995–2003, which has also been studied in [8]. The salmonellosis data set contains information on the date of diagnosis for each case and the patient's county of residence at the time of diagnosis. Even though these data have a fixed past time, in this section we show the results obtained in the prospective analysis of the data using our surveillance procedure.

Salmonellosis is an infection with *Salmonella* bacteria, which is usually transmitted to humans by eating foods contaminated with animal feces. Most people infected with *Salmonella* develop diarrhea, fever, vomiting, and abdominal cramps. In most cases, the illness lasts 5 to 7 days, and most people recover without treatment. In severe cases, the *Salmonella* infection may spread from the intestines to the blood stream and then to other body sites causing death unless the person is treated promptly with antibiotics.

Figure 3 depicts the monthly counts of salmonellosis cases in South Carolina during this 9-year period, with a clear seasonal cycle of length 12. A total of 6958 cases were diagnosed from January 1, 1995 to December 31, 2003. As shown, the number of salmonellosis cases diagnosed in October 1996 is extremely large, approximately three times of those numbers in October of other years, which indicates that something unusual happened at that time period. Although not as remarkable, unexpected high counts can also be observed at

different time points during the surveillance window, for instance, in September 1996, June 2000, February–March 2001, or the period from August to November 2002.

To apply the proposed surveillance procedure to the Salmonella data, we first need to define the expected counts in each county and time period. Because of the lack of a standard population with age-specific rates for the disease, expected counts are calculated by internal standardization [48]. That is, the total number of cases diagnosed at a particular time period is used to estimate a disease rate, and expected counts in each county are then calculated by assuming they are proportional to the counties' population during that time period. This approach was also used in [8]. The focus of their study was to analyze the salmonellosis data for October only during the period from 1995–2003. Here, however, we are interested in analyzing the whole data set. Because more cases are diagnosed in summer, we need to deal with seasonality. This can be done by either incorporating the seasonal effects in the expected counts or in the relative risks. In the first case, we use each monthly count to estimate the disease rate for that month and assume a constant relative risk during the surveillance exercise. In the second approach, we calculate the rates of disease using the 12-month cumulative counts, where months looking highly unusual are excluded when using the internal standardization method. We can then incorporate seasonal effects into the model through the relative risk decomposition as follows

$$\log(\theta_{it}) = \rho + u_i + v_i + \sum_{s=1}^{12} \alpha_s I_s(t) \quad (11)$$

where u_i and v_i represent the spatial variation, $\{\alpha_s\}_{s=1}^{12}$ are the seasonal effects, and $I_s(t)$ is the indicator function which takes the value 1 if time t corresponds to month s , and 0 otherwise. It is important to note that we define different risks to account for the seasonality of the data, but the risks so defined are constant over time. Alternatively, sine and cosine terms can be used instead of dummy variables to model the seasonal variation in the data. Yet, the model given by Equation (11) fits easily, and it can be seen as a generalization of the BYM model allowing for inclusion of covariates which, in general, do not necessarily have to correspond to seasonal effects.

The spatiotemporal model in Equation (11) is sequentially estimated as new data arrive every month, and the SCPO values associated with the new data are analyzed to detect changes in the relative risk pattern. Because of the abnormally high values in September and October 1996 and the lack of long-term historical data, we initially fit the model using the data for 1995 and start the surveillance exercise in January 1996. As new data become available, the length of the sliding window is progressively increased from 12 to 36, which allows the seasonal effects to be estimated using the data from 3-month sets. An alarm for the i th county is sounded at time t if the observed count of disease is higher than expected and the corresponding sSCPO is below 0.08. The total number of alarms for the first 7 months is used to estimate the distribution of the number of counties signaling an alarm at each time period when no change in risk takes place, which is given by the binomial distribution $Bi(46, 0.02)$. From August 1996, in an effort to control the probability of false alarm at 0.01, a general alarm for an out-of-control system is sounded at time t if there are at least five counties showing an increased disease incidence. Nevertheless, our surveillance procedure allows us to keep track of those areas of unusual incidence at each time point regardless of whether we generate an alarm for an out-of-control system. Table VI shows the number of counties signaling an alarm at each month from January 1996 to December 2003. As it can be seen, possible outbreaks of salmonellosis took place in October 1996, July 1997, June 2000, February 2001, and the time period from August to November 2002. These results are consistent with our expectations (Figure 3).

Values of the sSCPO equal to 0 at different times during the surveillance exercise strongly suggest that something unusual happened, even though the total number of counties signaling an alarm is not significant enough to sound a general alarm for an out-of-control system. This is the case of the Greenville and Spartanburg counties in September 1996. A total of 92 and 13 cases were diagnosed, respectively, in these counties during September 1996, the counts of disease observed up to that time point being always lower than 6. An alarm for an out-of-control system is sounded in October 1996, when the outbreak of salmonellosis spread to other counties. Similarly, the Richland county presents a sSCPO equal to 0 in March 2001, which indicates that the outbreak detected in February 2001 is still 'alive'. The 69 cases of salmonellosis reported during March 2001 in this county are the cause of the peak we observe in the time plot of the series. The surveillance procedure alerts us to this high count, which is about 23 times the number of cases diagnosed in March of previous years. However, because only this county presents an extremely increased disease incidence, the surveillance procedure does not sound a general alarm for the whole system. Figure 4 displays the spatial distribution of the sSCPO for the salmonellosis data in October 1996, February 2001, and October–November 2002. For illustrative purposes, we also include maps for September 1996 and March 2001 which, as we have mentioned before, include areas of increased disease incidence.

These results are consistent with what Zhou and Lawson [8] observed in their analysis of the salmonellosis data for October during the period from 1995–2003. In that paper, the authors detect an outbreak of salmonellosis in Greenville, Pickens, and Spartanburg during October 1996. In October 2002, the authors also detected an increased incidence in the Charleston county.

When the CSCPO is used to detect emerging clusters of disease, the surveillance procedure generates the first alarm for an out-of-control system in September 1996, the following counties in the northwest region of South Carolina prompting further investigation: Abbeville, Greenville, Greenwood, Lexington, Newberry, Pickens, and Spartanburg. Note that only three of these counties undergo a significant increase in risk at that time period to be detected by the univariate SCPO (Figure 4). In October 1996, alarms based on the CSCPO were sounded for Abbeville, Anderson, Pickens, Greenville, and Spartanburg. Hence, when a set of neighboring areas show an increased disease incidence, the CSCPO facilitates detection of those regions undergoing an incipient outbreak of disease at the very moment of their onset.

In this case study, we are dealing with sparse data. Except for some of the most populated counties, counts of disease observed at each time period are generally equal to zero. It is worth pointing out that, in this case, the use of the sSCPO to detect areas of increased disease incidence provides more satisfactory results than the analysis of the Bayesian p -value surface. This is due to the fact that Bayesian p -values tend to signal an alarm for those counties which at a particular time present one or two cases of disease after a period of time with no reported cases. Note that the estimated relative risk for a particular county will be practically zero if no cases have been reported up to that time point. Consequently, most of the replicated data from the corresponding conditional predictive distribution will be zero, and the Bayesian p -value associated with a new count equal to one will be close to zero.

Finally, we present the results obtained when the Farrington algorithm is applied to monitor each one of the $m = 46$ time series independently, beginning from February 1996. In this case, seasonality is handled by estimating the log-linear regression model using past observations within a window of size 3 around the period of the observations. We assume that an alarm is generated for county i at time t if the observation y_{it} is above the upper limit of its $(1 - \alpha) \times 100\%$ prediction interval. The two-third-power transformation is applied to

normalize the distribution before computing the predictive interval. To control the probability of false alarm at each time period at 0.01, we set α equal to $0.01/m$. Figure 5 displays those time periods for which there is at least one county causing an alarm. For comparative purposes, we also present the results corresponding to the critical level $\alpha = 0.01$. Because it is common to observe counts of disease higher than expected for just one county at a particular time period, an excessive number of alarms are generated when the problem of multiple comparisons is not dealt with. However, areas of unusual aggregation of disease in need of further investigation are missed when the critical level α is reduced. For instance, only the counties of Greenville and Spartanburg and the counties of Greenville and Pickens are detected respectively in September and October 1996. During the period from August to November 2002, just one county causes an alarm in November 2002. The reason is twofold. First, only those extremely unusual observations are detected when the critical level is reduced. Second, because the time series are analyzed separately, correlation between the series is ignored.

7. Discussion

This paper has focused on the development of statistical methodology for prospective spatiotemporal public health surveillance. In particular, we have shown how the CPO can be adapted in a surveillance context to detect areas of unusual aggregation of disease. The resulting technique, called the SCPO, has the potential to be applied in any surveillance context where a statistical model is used to describe the count of disease occurrences. Here, we have shown its application in the monitoring of noninfectious diseases within the framework of Bayesian hierarchical Poisson models. A key finding from our simulation study is that, in terms of outbreak detection, the SCPO based on the convolution model provides more satisfactory results. On the basis of the goodness of fit, previous studies using Poisson models in a surveillance context have used a spatiotemporal model including both spatial and temporal effects as well as a space–time interaction term [7, 20]. Recently, it has been pointed out the importance of maintaining a relatively simple model capable of describing the overall behavior of disease over space and time without absorbing changes in the model fit [8, 22]. However, no such comparative evaluation can be found in the literature. For the monitoring of infectious disease data, the convolution model provides a suitable framework for describing the behavior of disease during nonepidemic stages. The SCPO can then be used to detect the onset of epidemics. Because evidence against the in-control situation is not accumulated over time, the SCPO may fail to detect a gradual change in risk. On those occasions, a weighted sum of previous SCPO or a multivariate extension incorporating information from previous counts of disease may help to detect a continuing change in disease incidence over time.

The SCPO and the previously proposed Bayesian p -value are both model-based surveillance techniques, which use the posterior predictive distribution to detect unusual observations, and so both techniques have a similar detection power. Yet, some features of the SCPO make our procedure a more convenient one. First, unlike surveillance residuals, the SCPO is not affected by censoring effects. Hence, the use of a lagged loss function or ‘sliding window’ is a feasible solution to the estimation problem, which allows the computation time to be reduced substantially. It is important to emphasize that, from a theoretical viewpoint, the use of a lagged loss function is preferable to the surveillance mechanism proposed in [8], where the spatial model is fitted to the data from a single time point. In our surveillance procedure, we calculate accurate estimates of the spatial effects using the most recent observations. Second, changes in risk can be detected without having to implement the equivalent of a parametric bootstrap in the Bayesian setting, which can be cumbersome on some occasions. In addition, when dealing with sparse data such as the salmonellosis data analyzed in this paper, the use of the SCPO provides more satisfactory results than the

analysis of the Bayesian p -value surface. Finally, the SCPO can be easily extended to the multivariate case. In this paper, a CSCPO incorporating information from the first-order spatial neighborhood has been proposed. As shown in the simulation and case studies, this technique improves considerably outbreak detection capability when changes in risk affect neighboring areas simultaneously. Very often in practice, it is of interest to simultaneously monitor multiple diseases. On those occasions, a multivariate SCPO incorporating information from the different diseases could be defined in a similar way.

An important feature of our surveillance procedure is that different alarms are sounded for those areas of increased disease incidence. We have shown how the multiple comparisons problem can be successfully addressed by assuming that there is a common prior probability that a given area signals an alarm when no change in risk takes place. An alarm for an out-of-control system is then generated if there is a significant number of small areas showing an unusual aggregation of disease. This number will be set according to the desired probability of a false alarm. Because we do not reduce the critical level of the individual comparisons, this approach allows us to recover the true outbreak behavior and facilitates the carrying out of further investigations and preventive actions.

Finally, we have shown how the BYM model can be extended to incorporate covariates in the relative risk decomposition. In our example, these covariates represent seasonal effects. It would be valuable to investigate the effect of incorporating covariates related to possible disease risk factors, such as air pollution levels or poverty rates. This would allow us to detect changes in risk by monitoring changes in the covariates, providing a better understanding of emerging trends or outbreaks of disease. A very fruitful area for further research is the incorporation of covariates associated with syndromes for diseases such as job or school absenteeism, physician visits, or over-the-counter pharmacy sales. This would allow for detection of adverse health effects at the earliest possible time, possibly even before disease diagnoses can be confirmed.

Recently, integrated nested Laplace approximations have been used to approximate posterior marginals in latent Gaussian models [49]. Schrödle *et al.* [50] applied these techniques to obtain parameter estimates for different spatiotemporal models for retrospective analyses of veterinary surveillance data. Although this approach may be useful for computing the SCPO, computation of the CSCPO is not straightforward. Yet, its application in prospective public health surveillance may be a fruitful area for further research.

Acknowledgments

This work was supported by Grant Number R03CA162029 from the National Cancer Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

References

1. Le Strat, Y. Overview of temporal surveillance.. In: Lawson, A.; Kleinman, K., editors. *Spatial and Syndromic Surveillance for Public Health*. Wiley; Chichester: 2005. p. 13-29.
2. Unkel S, Farrington CP, Garthwaite PH, Robertson C, Andrews N. Statistical methods for the detection of infectious disease outbreaks: a review. *Journal of the Royal Statistical Society, series A*. 2012; 175:1–34.
3. Lowry CA, Woodall WH, Champ CW, Rigdon SE. A multivariate exponentially weighted moving average control chart. *Technometrics*. 1992; 34:46–53.
4. Kulldorff M. Prospective time periodic geographical disease surveillance using a scan statistic. *Journal of the Royal Statistical Society, series A*. 2001; 164:61–72.

5. Rogerson, PA. Spatial surveillance and cumulative sum methods.. In: Lawson, AB.; Kleinman, K., editors. *Spatial and Syndromic Surveillance for Public Health*. Wiley; Chichester: 2005. p. 95-114.
6. Robertson C, Nelson TA, MacNab YC, Lawson AB. Review of methods for space-time disease surveillance. *Spatial and Spatio-temporal Epidemiology*. 2010; 1:105–116. [PubMed: 22749467]
7. Lawson AB, Clark A, Vidal C. Developments in general and syndromic surveillance for small area health data. *Journal of Applied Statistics*. 2004; 31:951–966.
8. Zhou H, Lawson AB. EWMA smoothing and Bayesian spatial modeling for health surveillance. *Statistics in Medicine*. 2008; 27:5907–5928. [PubMed: 18759375]
9. Kleinman K, Lazarus R, Platt R. A generalized linear mixed models approach for detecting incident clusters of disease in small areas, with an application to biological terrorism. *American Journal of Epidemiology*. 2004; 159:217–224. [PubMed: 14742279]
10. Watkins RE, Eagleson S, Veenendaal B, Wright G, Plant AJ. Disease surveillance using a hidden Markov model. *BMC Medical Informatics and Decision Making*. 2009; 9:39. [PubMed: 19664256]
11. Diggle P, Rowlingson B, Su T. Point process methodology for on-line spatio-temporal disease surveillance. *Environ-metrics*. 2005; 16:423–434.
12. Geisser S. Discussion on sampling and Bayes' inference in scientific modelling and robustness (by G.E.P. Box). *Journal of the Royal Statistical Society, series A*. 1980; 143:416–417.
13. Gelfand, AE.; Dey, DK.; Chang, H. Model determination using predictive distributions with implementation via sampling-based methods (with discussion).. In: Bernardo, JM.; Berger, JO.; Dawid, AP.; Smith, AFM., editors. *Bayesian Statistics 4*. Oxford University Press; Oxford: 1992. p. 147-167.
14. Berry DA, Hochberg Y. Bayesian perspectives on multiple comparisons. *Journal of Statistical Planning and Inference*. 1999; 82:215–227.
15. Scott JG, Berger JO. An exploration of aspects of Bayesian multiple testing. *Journal of Statistical Planning and Inference*. 2006; 136:2144–2162.
16. Lawson, AB. *Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology*. Chapman & Hall; Boca Raton: 2009.
17. Lawson, AB. Some considerations in spatial-temporal analysis of public health surveillance data.. In: Brookmeyer, R.; Stroup, DF., editors. *Monitoring the Health of Populations: Statistical Principles and Methods for Public Health Surveillance*. Oxford University Press; Oxford: 2004. p. 289-314.
18. Besag J, York J, Mollié A. Bayesian image restoration with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*. 1991; 43:1–59.
19. Knorr-Held L. Bayesian modelling of inseparable space-time variation in disease risk. *Statistics in Medicine*. 2000; 19:2555–2567. [PubMed: 10960871]
20. Vidal CL, Lawson AB. Monitoring changes in spatio-temporal maps of disease. *Biometrical Journal*. 2006; 48:463–480. [PubMed: 16845909]
21. Gamerman, D.; Lopes, HF. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. 2nd ed.. Chapman & Hall; Boca Raton: 2006.
22. Lawson, AB. Spatial and spatio-temporal disease analysis.. In: Lawson, AB.; Kleinman, K., editors. *Spatial and Syndromic Surveillance for Public Health*. Wiley; Chichester: 2005. p. 55-76.
23. Lawson AB, Biggeri AB, Boehning D, Lesaffre E, Viel JF, Clark A, Schlattmann P, Divino F. Disease mapping models: an empirical evaluation. *Statistics in Medicine*. 2000; 19:2217–2241. [PubMed: 10960849]
24. Best N, Richardson S, Thomson A. A comparison of Bayesian spatial models for disease mapping. *Statistical Methods in Medical Research*. 2005; 14:35–59. [PubMed: 15690999]
25. Bernardinelli L, Clayton D, Montomoli C. Bayesian estimates of disease maps: how important are priors?. *Statistics in Medicine*. 1995; 14:2411–2431. [PubMed: 8711278]
26. Banerjee S, Carlin BP. Semiparametric spatio-temporal frailty modeling. *Environmetrics*. 2003; 14:523–535.
27. Gelman A. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*. 2006; 1:515–534.

28. Doucet, A.; de Freitas, N.; Gordon, N. *Sequential Monte Carlo Methods in Practice*. Springer; New York: 2001.
29. Berzuini C, Best NG, Gilks WR, Larizza C. Dynamic conditional independence models and Markov Chain Monte Carlo methods. *Journal of the American Statistical Association*. 1997; 92:1403–1412.
30. Farrington CP, Andrews NJ, Beale AD, Catchpole MA. A statistical algorithm for the early detection of outbreaks of infectious disease. *Journal of the Royal Statistical Society, Series A*. 1996; 159:547–563.
31. Jiang J, Zhang Y. A revisit to block and recursive least squares for parameter estimation. *Computers and Electrical Engineering*. 2004; 30:403–416.
32. Höhle M, Paul M, Held L. Statistical approaches to the monitoring and surveillance of infectious diseases for veterinary public health. *Preventive Veterinary Medicine*. 2009; 91:2–10. [PubMed: 19539387]
33. Wessman P. Some principles for surveillance adopted for multivariate processes with a common change point. *Communications in Statistics - Theory and Methods*. 1998; 27:1143–1161.
34. Frisén M, de Maré J. Optimal surveillance. *Biometrika*. 1991; 78:271–280.
35. Sonesson, C.; Frisén, M. Multivariate surveillance.. In: Lawson, AB.; Kleinman, K., editors. *Spatial and Syndromic Surveillance for Public Health*. Wiley; Chichester: 2005. p. 153-166.
36. Sonesson C, Bock D. A review and discussion of prospective statistical surveillance in public health. *Journal of the Royal Statistical Society, Series A*. 2003; 166:5–21.
37. Frisén M, Andersson E, Schiöler L. Robust outbreak surveillance of epidemics in Sweden. *Statistics in Medicine*. 2009; 28:476–493. [PubMed: 19012277]
38. Efron, B.; Tibshirani, RJ. *An Introduction to the Bootstrap*. Chapman & Hall; New York: 1993.
39. Box GEP. Sampling and Bayes' inference in scientific modelling and robustness (with discussion). *Journal of the Royal Statistical Society, series A*. 1980; 143:383–430.
40. Geisser S, Eddy W. A predictive approach to model selection. *Journal of the American Statistical Association*. 1979; 74:153–160.
41. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*. 1995; 57:289–300.
42. Storey JD. A direct approach to false discovery rates. *Journal of the Royal Statistical Society, Series B*. 2002; 64:479–498.
43. Catelan D, Biggeri A. Multiple testing in disease mapping and descriptive epidemiology. *Geospatial Health*. 2010; 4:219–229. [PubMed: 20503190]
44. Catelan D, Lagazio C, Biggeri A. A hierarchical Bayesian approach to multiple testing in disease mapping. *Biometrical Journal*. 2010; 52:784–797. [PubMed: 20809523]
45. Höhle M. Surveillance: an R package for the monitoring of infectious diseases. *Computational Statistics*. 2007; 22:571–582.
46. Richardson S, Thomson A, Best N, Elliott P. Interpreting posterior relative risk estimates in disease-mapping studies. *Environmental Health Perspectives*. 2004; 112:1016–1025. [PubMed: 15198922]
47. Hossain MM, Lawson AB. Space-time Bayesian small area disease risk models: development and evaluation with a focus on cluster detection. *Environmental and Ecological Statistics*. 2010; 17:73–95. [PubMed: 20953293]
48. Banerjee, S.; Carlin, BP.; Gelfand, AE. *Hierarchical Modeling and Analysis for Spatial Data*. Chapman & Hall; Boca Raton: 2003.
49. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations (with discussion). *Journal of the Royal Statistical Society, Series B*. 2009; 71:319–392.
50. Schrödle B, Held L, Riebler A, Danuser J. Using integrated nested Laplace approximations for the evaluation of veterinary surveillance data from Switzerland: a case-study. *Journal of the Royal Statistical Society, Series C*. 2011; 60:261–279.

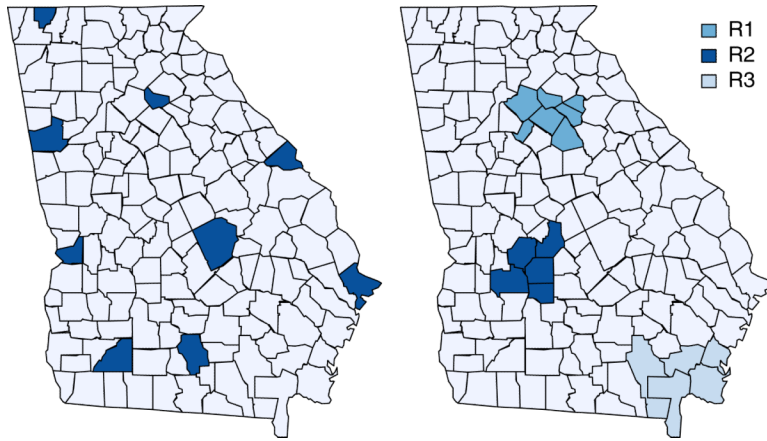


Figure 1. Simulation study. Regions where changes in risk are considered. Left, isolated local jumps (scenarios 1 and 2); right, clustered local jumps (scenario 3).

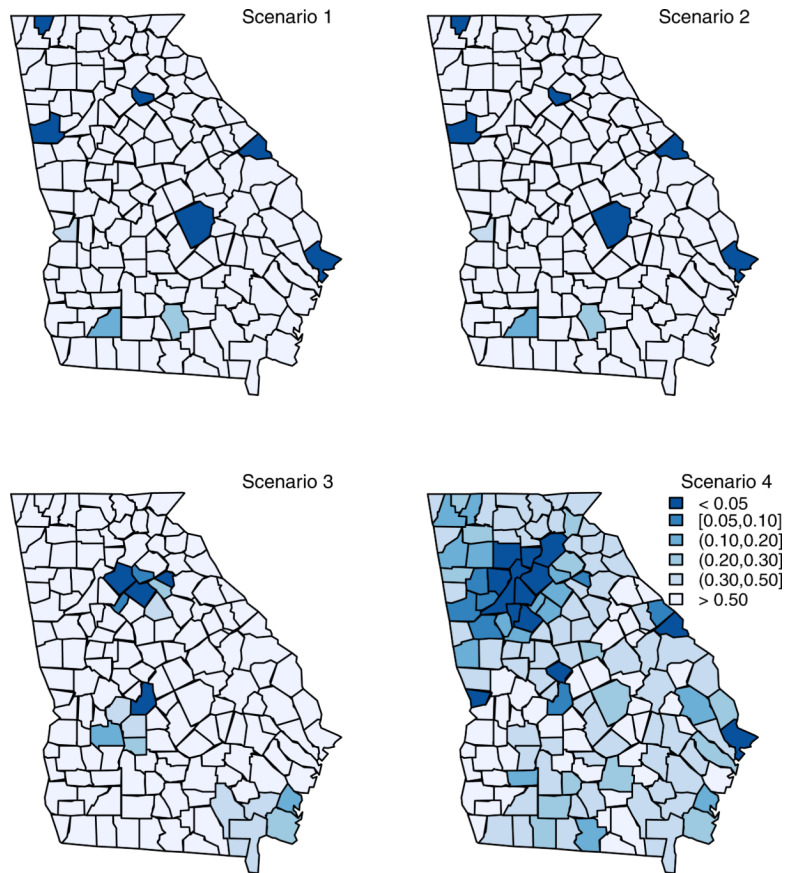


Figure 2. Simulation study. Spatial distribution of the scaled surveillance conditional predictive ordinate from the Besag–York–Mollie model at time period 18 under different scenarios.

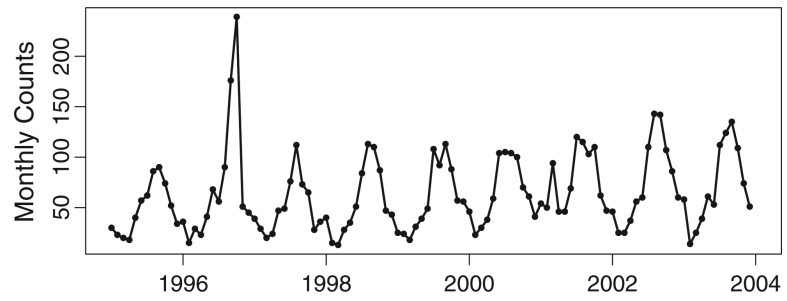


Figure 3. Monthly counts of reported salmonellosis cases in South Carolina during 1995–2003.

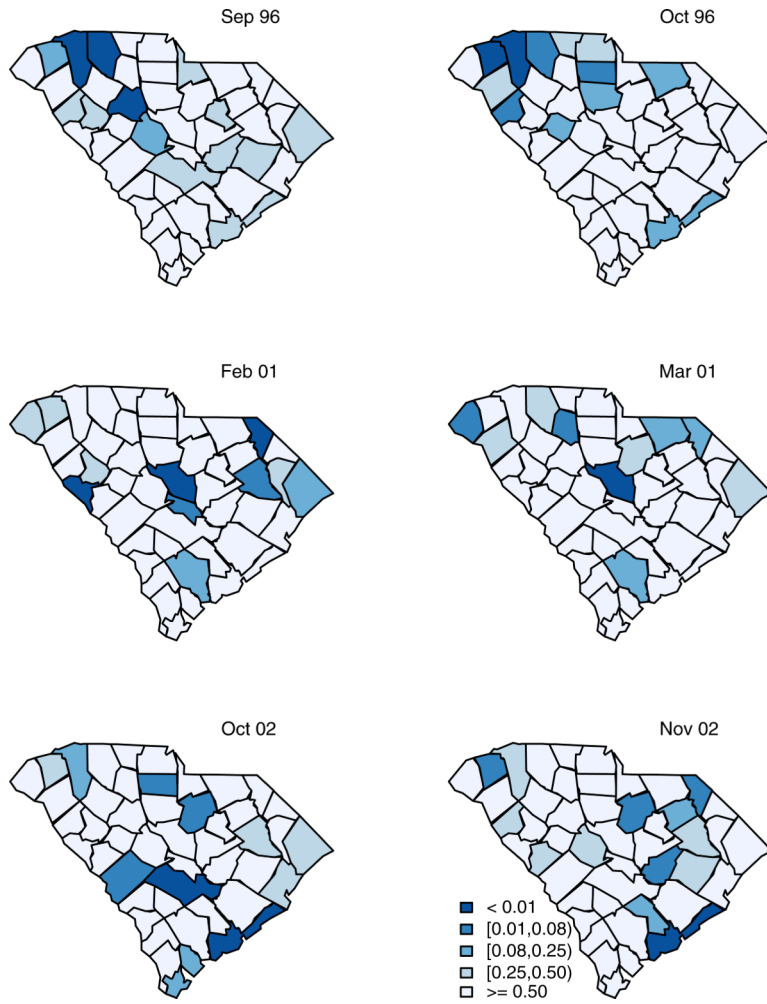


Figure 4. Spatial distribution of the scaled surveillance conditional predictive ordinate for the salmonellosis data at those months undergoing a possible outbreak of disease.

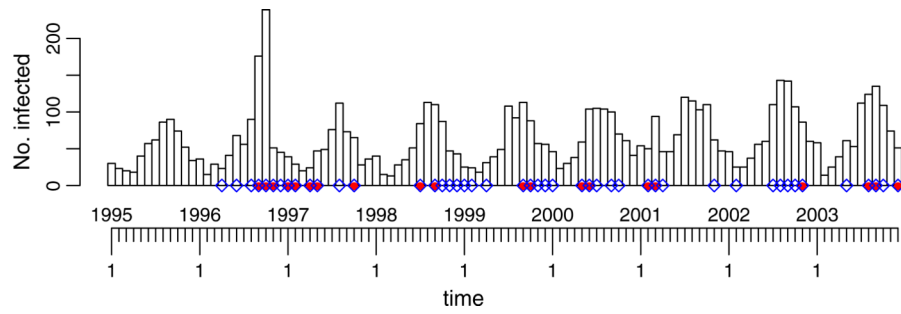


Figure 5. Salmonellosis case study. Resulting alarms from the surveillance of the $m = 46$ time series using the Farrington algorithm. Diamonds and solid points represent alarms corresponding to the critical levels $\alpha = 0.01$ and $\alpha = 0.01/46$ respectively.

Table IOutcomes of the surveillance procedure during the outbreak period T_c .

	Areas that signal an alarm at any time during T_c	Areas that do not signal an alarm during T_c	Total
Areas with a change in risk	TA	FNA	m_0
Areas without a change in risk	FA	TNA	$m - m_0$
Total	N_a	$m - N_a$	m

TA, true alarms; FA, false alarms; FNA, false no alarms; TNA, true no alarms.

Table II

Simulation study. Sensitivity, specificity, and median time to outbreak detection (both posterior average estimates and 95% credible intervals) of the proposed surveillance technique when the Besag–York–Mollie model is fitted to the data within a sliding window of length l . Decision rule: An alarm for the i th county is sounded at time t if the observed count of disease is higher than expected and the corresponding scaled surveillance conditional predictive ordinate is below the critical level 0.10.

l	Scenario 1		Scenario 3	
	Sensitivity	Specificity	MTD	Specificity
3	0.91 [0.78,1]	0.93 [0.89,0.97]	0 [0,0]	0.93 [0.89,0.97]
6	0.93 [0.78,1]	0.93 [0.89,0.97]	0 [0,0]	0.93 [0.89,0.97]
10	0.94 [0.78,1]	0.93 [0.89,0.97]	0 [0,0]	0.93 [0.89,0.97]
h	0.95 [0.78,1]	0.93 [0.88,0.96]	0 [0,0]	0.92 [0.88,0.97]

MTD, median time to outbreak detection.

Table III

Simulation study. Sensitivity, specificity, and median time to outbreak detection (both posterior average estimates and 95% credible intervals) of the proposed surveillance technique when it is based on the Besag–York–Mollie model (upper half) and the spatiotemporal model (lower half). Decision rule: An alarm for the i th county is sounded at time t if the observed count of disease is higher than expected, and the corresponding scaled surveillance conditional predictive ordinate is below the critical level α .

α	Scenario 1				Scenario 2				Scenario 3				Scenario 4			
	Sensitivity	Specificity	MTD	MTD	Sensitivity	Specificity	MTD	MTD	Sensitivity	Specificity	MTD	MTD	Sensitivity	Specificity	MTD	MTD
0.20	0.96 [0.89,1]	0.86 [0.80,0.91]	0 [0,0]	0 [0,0]	0.95 [0.78,1]	0.86 [0.80,0.91]	0 [0,0]	0 [0,0]	0.86 [0.71,1]	0.86 [0.80,0.92]	0 [0,1]	0 [0,1]	0.67 [0.61,0.74]	0.61 [0.54,0.67]	1 [0,1]	1 [0,1]
0.15	0.95 [0.78,1]	0.90 [0.85,0.94]	0 [0,0]	0 [0,0]	0.93 [0.78,1]	0.90 [0.85,0.94]	0 [0,0]	0 [0,0]	0.83 [0.65,1]	0.90 [0.85,0.94]	0 [0,1]	0 [0,1]	0.61 [0.54,0.67]	0.61 [0.54,0.67]	1 [1,2]	1 [1,2]
0.10	0.93 [0.78,1]	0.93 [0.89,0.97]	0 [0,0]	0 [0,0]	0.90 [0.67,1]	0.93 [0.89,0.97]	0 [0,0]	0 [0,0]	0.77 [0.59,0.94]	0.93 [0.89,0.97]	0 [0,1]	0 [0,1]	0.53 [0.47,0.60]	0.53 [0.47,0.60]	2 [1,Inf]	2 [1,Inf]
0.05	0.89 [0.67,1]	0.97 [0.93,0.99]	0 [0,0]	0 [0,0]	0.84 [0.67,1]	0.97 [0.94,0.99]	0 [0,1]	0 [0,1]	0.68 [0.50,0.88]	0.97 [0.94,0.99]	0 [0,Inf]	0 [0,Inf]	0.42 [0.35,0.48]	0.42 [0.35,0.48]	Inf (Inf,Inf)	Inf (Inf,Inf)
0.20	0.94 [0.78,1]	0.86 [0.79,0.91]	0 [0,0]	0 [0,0]	0.91 [0.67,1]	0.86 [0.80,0.91]	0 [0,0]	0 [0,0]	0.79 [0.59,0.94]	0.86 [0.80,0.91]	0 [0,1]	0 [0,1]	0.58 [0.51,0.65]	0.58 [0.51,0.65]	1 [1,2]	1 [1,2]
0.15	0.91 [0.78,1]	0.90 [0.84,0.94]	0 [0,0]	0 [0,0]	0.87 [0.67,1]	0.90 [0.85,0.94]	0 [0,0]	0 [0,0]	0.73 [0.53,0.94]	0.90 [0.85,0.94]	0 [0,2]	0 [0,2]	0.50 [0.43,0.57]	0.50 [0.43,0.57]	2 [1,Inf]	2 [1,Inf]
0.10	0.87 [0.67,1]	0.94 [0.90,0.97]	0 [0,0]	0 [0,0]	0.82 [0.56,1]	0.94 [0.90,0.97]	0 [0,1]	0 [0,1]	0.64 [0.41,0.82]	0.94 [0.90,0.97]	0 [0,Inf]	0 [0,Inf]	0.41 [0.35,0.48]	0.41 [0.35,0.48]	Inf (Inf,Inf)	Inf (Inf,Inf)
0.05	0.81 [0.61,1]	0.97 [0.95,1]	0 [0,0]	0 [0,0]	0.72 [0.44,1]	0.97 [0.94,0.99]	0 [0,Inf]	0 [0,Inf]	0.52 [0.29,0.71]	0.97 [0.94,0.99]	1 [0,Inf]	1 [0,Inf]	0.29 [0.23,0.35]	0.29 [0.23,0.35]	Inf (Inf,Inf)	Inf (Inf,Inf)

Table IV

Simulation study. Sensitivity and median time to outbreak detection (both posterior average estimates and 95% credible intervals) of the surveillance conditional predictive ordinate, clustered surveillance conditional predictive ordinate, Bayesian p -value, and the Farrington algorithm by simulation scenario.

Procedure	Scenario 1		Scenario 2		Scenario 3		Scenario 4	
	Sensitivity	MTD	Sensitivity	MTD	Sensitivity	MTD	Sensitivity	MTD
SCPO	0.93 [0.78,1]	0 [0,0]	0.90 [0.67,1]	0 [0,0]	0.77 [0.59,0.94]	0 [0,1]	0.53 [0.47,0.60]	2 [1,Inf]
CSCPO	0.77 [0.56,1]	0 [0,1]	0.67 [0.44,0.89]	0 [0,Inf]	0.85 [0.65,1]	0 [0,1]	0.69 [0.63,0.77]	0 [0,1]
Bayesian p -value from ST model	0.88 [0.67,1]	0 [0,0]	0.83 [0.67,1]	0 [0,1]	0.68 [0.47,0.88]	0 [0,Inf]	0.44 [0.37,0.50]	Inf [2,Inf]
Bayesian p -value from BYM model	0.94 [0.78,1]	0 [0,0]	0.90 [0.67,1]	0 [0,0]	0.78 [0.59,0.94]	0 [0,1]	0.53 [0.47,0.60]	2 [1,Inf]
Farrington	0.91 [0.78,1]	0 [0,0]	0.87 [0.67,1]	0 [0,0]	0.73 [0.53,0.88]	0 [0,2]	0.50 [0.43,0.58]	2 [1,Inf]

SCPO, surveillance conditional predictive ordinate; CSCPO, clustered surveillance conditional predictive ordinate; ST, spatiotemporal; BYM, Besag-York-Mollie; MTD, median time to outbreak detection.

Table V

Simulation study. Probability of observing at least the same number of counties signaling an alarm at each time point during the outbreak period. Decision rule: An alarm for a particular county is sounded at time t if the observed count of disease is higher than expected and the corresponding scaled surveillance conditional predictive ordinate is below 0.10.

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
$P(n = n_{18})$	0.005 [0,0.042]	0.007 [0,0.042]	0.001 [0,0.005]	0 [0,0]
$P(n = n_{19})$	0.007 [0,0.073]	0.016 [0,0.101]	0.005 [0,0.042]	0 [0,0]
$P(n = n_{20})$	0.019 [0,0.101]	0.037 [0,0.214]	0.022 [0,0.214]	0 [0,0]

Table VI

Salmonellosis case study. Number of counties signaling an alarm (n_t) at each time point t during the surveillance period (January 1996–December 2003). Decision rule: An alarm for a particular county is sounded at time t if the observed count of disease is higher than expected and the corresponding scaled surveillance conditional predictive ordinate is below 0.08. Months with five or more counties signaling an alarm are highlighted in bold.

t	n_t	t	n_t	t	n_t	t	n_t	t	n_t	t	n_t	t	n_t
1	1	17	4	33	3	49	2	65	1	81	4		
2	0	18	2	34	2	50	1	66	1	82	5		
3	2	19	5	35	2	51	3	67	1	83	5		
4	2	20	2	36	3	52	0	68	2	84	2		
5	1	21	0	37	0	53	1	69	1	85	1		
6	1	22	1	38	1	54	5	70	1	86	0		
7	0	23	0	39	2	55	2	71	0	87	0		
8	3	24	1	40	2	56	1	72	1	88	0		
9	3	25	2	41	1	57	1	73	1	89	2		
10	5	26	0	42	1	58	1	74	1	90	2		
11	2	27	0	43	3	59	3	75	1	91	1		
12	3	28	3	44	2	60	1	76	2	92	2		
13	0	29	1	45	0	61	2	77	0	93	4		
14	3	30	1	46	0	62	5	78	1	94	3		
15	2	31	3	47	2	63	3	79	3	95	0		
16	2	32	3	48	0	64	2	80	6	96	2		