
A statistical system for detecting *Salmonella* outbreaks in British livestock

R. KOSMIDER^{1*}, L. KELLY^{1,2}, S. EVANS¹ AND G. GETTINBY²

¹ Centre for Epidemiology and Risk Analysis, Veterinary Laboratories Agency, Woodham Lane, Addlestone, Surrey, UK

² Department of Statistics and Modelling Science, University of Strathclyde, Glasgow, Scotland, UK

(Accepted 23 November 2005, first published online 26 January 2006)

SUMMARY

Worldwide, early detection systems have been used in public health to aid the timely detection of increases in disease reporting that may be indicative of an outbreak. To date, their application to animal surveillance has been limited and statistical methods to analyse human health data have not been viewed as being applicable for animal health surveillance data. This issue was investigated by developing an early detection system for *Salmonella* disease in British livestock. We conclude that an early detection system, as for public health surveillance, can be an effective tool for enhanced surveillance. In order to implement this system in the future and extend it for other data types, we provide recommendations for improving the current data collection process. These recommendations will ensure that quality surveillance data are collected and used effectively to monitor disease in livestock populations.

INTRODUCTION

In the last few years, the United Kingdom has experienced economic difficulties within the livestock industry as a result of the emergence of novel diseases such as bovine spongiform encephalopathy (BSE) [1] and the increase, over time, in bacterial pathogens, such as *Salmonella* Typhimurium [2] and verocytotoxigenic *Escherichia coli* O157 [2]. In order to protect animal and public health and the economic viability of the industry, it is important that changes in the incidence of zoonotic pathogens within livestock are detected as early as possible. In doing so, control measures can be implemented in a timely fashion to prevent further spread of disease in the population.

Traditionally, in the United Kingdom, monitoring of veterinary diseases has been undertaken as part of a government health programme, which advocates the routine surveillance of animal pathogens [3]. Integral to this has been the analysis of data to determine seasonal trends, the frequency of outbreaks, and patterns in disease reporting over time (see [4], for example). However, as this analysis is carried out retrospectively, intervention measures often cannot be implemented until some time after an increase in reporting has been observed. This problem can be addressed through the development of early detection systems, which if implemented, enable prospective analysis of the data.

Early detection systems use data from on-going surveillance to identify potential outbreaks of disease soon after they have occurred in the field, by comparing the most recently recorded number of disease cases (the ‘current count’) with a threshold value derived from historical data. If the current count is

* Author for correspondence: Mrs R. Kosmider, Centre for Epidemiology and Risk Analysis, Veterinary Laboratories Agency, Woodham Lane, Addlestone, Surrey KT15 3NB, UK.
(Email: r.kosmider@vla.defra.gsi.gov.uk)

above the threshold value a warning is implemented indicating that the count is statistically aberrant and indicative of an outbreak situation, i.e. an increase in reports above expected. In this instance, an epidemiological investigation may be undertaken to determine whether there is evidence in the field to support the case for an outbreak. The warning process is typically undertaken using a statistical algorithm that automatically fits the historical data to an empirical model and provides an expected current count and threshold value. Early detection systems, therefore, can be used alongside traditional statistical analysis as an additional tool to provide enhanced surveillance.

To date, early detection systems have been predominantly used in the public health arena and their use within the animal health arena has been limited. In this paper we address the benefit of using an early detection system as a tool for enhanced animal surveillance by investigating the development of an early detection system for *Salmonella* isolated from domestic livestock species within Great Britain. We consider the available data, address inherent data issues, and for an appropriate subset, develop a detection system. Testing of the system leads to conclusions on feasibility and recommendations for improving the data collection process.

METHODS

Selection of surveillance data

In Great Britain, *Salmonella* infection in livestock is a reportable zoonoses under the Zoonoses Order 1989 and hence all laboratory isolations are reported to the Senior Veterinary Investigation Officer at one of the Veterinary Laboratories Agency's (VLA) Regional Laboratories or, in Scotland, to a Divisional Veterinary Manager. Through this reporting of *Salmonella* infection, data are available, centrally, on the isolation of numerous *Salmonella* serotypes in a number of animal species. *Salmonella* surveillance data from 1993 to December 2002 are considered in this study to ensure consistency in data quality as the reporting system was unchanged during this time period. Only isolations of *S. Typhimurium* from livestock displaying clinical illnesses and *S. Dublin* from clinically ill cattle were considered within the early detection system as this provided a clearly defined subset of data for analysis and reflected the nature of the surveillance system (most reports

in species other than poultry arise from diagnostic investigation of clinical illness).

Assessment of the data

An initial crucial stage of developing an early detection system is an assessment of the data to determine if the disease burden can be estimated at a specific point in time. This estimation is critical because it forms the basis of the derivation of the expected and threshold values and hence affects the detection of outbreaks; a system with a threshold value set too low will raise false alarms whereas a system with a threshold value set too high will not detect outbreaks. Due to inherent attributes of the surveillance data (i.e. presence of past outbreaks, inaccurate representation of geographical areas, seasonality, under-reporting, reporting delays), there are difficulties in estimating this true disease situation. Some of these issues have been addressed previously particularly in relation to human surveillance data (see [5], for example). For *Salmonella* animal surveillance additional issues arise which affect the estimation of the baseline incidence of disease.

In animal surveillance the epidemiological unit is often the herd or flock. In these circumstances, the presence of an infected individual in the group may strongly influence the probability of another individual within the group becoming infected. This phenomenon, known as the 'herd effect', means that animals in a herd/flock are dependent, that is each case of infection is influenced by another case, whilst individual herds (depending on the degree of mixing) are considered independent. This is important when considering two very different approaches used for describing the disease burden: isolations and incidents. An isolation is a single report of a pathogen within an individual animal or groups of animals (epidemiological unit) whereas an incident comprises the first isolation and all subsequent isolations of the same serotype and phage type from the epidemiological unit of animals within a standard time period, usually 30 days; allowance is made for local epidemiological knowledge [4]. Modelling the number of incidents over time is, therefore, an estimation of the number of livestock units that are infected whereas the number of isolations represents the number of individual reports of infection [6]. Within the *Salmonella* database, isolations are clearly identified whereas incidents are calculated using an algorithm applied to the database.

Animals are sampled for numerous reasons as dictated by various animal health policies. *Salmonella*, specifically, may be isolated from samples taken for: diagnosis of clinical illness, routine surveillance, monitoring requirements under the Poultry Breeding Flocks and Hatcheries Order (PBFHO) 1993, investigation under the Zoonoses Order 1989 and unspecified reasons. All of these reasons impact on what the data informs us about the burden of *Salmonella* infection in the domestic livestock population. For instance, without the availability of denominator data (i.e. the number of livestock units and tests taken) samples for surveillance reasons may result in a biased measurement of *Salmonella* infection in the population. This, however, is not the case for samples taken under the PBFHO 1993 as the Order dictates a specific sampling scheme of which and when samples should be taken. Assuming that the population does not alter significantly, positive samples can be considered to be representative of the population. Given this, the number of positive samples (i.e. the numerator data) is representative of the true *Salmonella* burden in the population over time. Samples collected for unknown reasons unfortunately do not provide any information about the disease burden and cannot be analysed within an early detection system, thus emphasizing the importance of the collection of accurate epidemiological data with the reports to enable interpretation of the data. For these reasons, not all the data within the *Salmonella* dataset can be combined within one early detection system. Rather, the data needs to be partitioned by reason for sampling and analysed separately, using appropriate statistical methods.

An influential factor in the collation of animal surveillance data over time is the changing circumstances affecting the livestock industry. For example, epidemics of major contagious pathogens, as occurred in 2001 with the foot-and-mouth disease (FMD) outbreak, impact on the normal surveillance of animal disease. In addition, the varying economic viability of livestock enterprise often means it is too costly to request a veterinary surgeon to take a sample. Therefore, the submission rate of samples for the clinical signs of *Salmonella* disease varies over time, a factor that is driven primarily by costs. This is in contrast to human surveillance where the isolation of disease is reliant on the patient going to their general practitioner and submitting a sample, a factor driven by awareness.

A further unique facet of animal health data is that it incorporates all the species that are affected by the

pathogen of interest. For *Salmonella* data, this means a wide range of livestock species are represented within the data. Hence, there is potentially a greater amount of surveillance data for animal health than public health and, further, a greater variation in the frequency in which different serotypes for each affected livestock species are reported. The statistical method chosen, therefore, should be robust enough to cope with these differences.

Given the above considerations, it is apparent there are numerous data issues that infringe on the ability to develop an early detection system for all *Salmonella* animal health surveillance data. We, therefore, selected a subset of the data for the system: isolations of *S. Typhimurium* from livestock sampled for clinical illness and of *S. Dublin* detected from cattle sampled for clinical illness. These subsets spanning 1993 to December 2002 were selected because it was considered comparable to human surveillance data and, therefore, statistical methods applied to human health were explored for their applicability to this data subset.

Early detection system

Several statistical methods have been applied previously for the early detection of outbreaks or clusters of disease: time-series analysis [7, 8], regression analysis [9], scan statistics [10] and cumulative sums [11]. Time series is an intuitive approach to adopt given the fact that surveillance data naturally conforms to a series of data points over time. However, when applied to a dataset with numerous serotypes, each with a specific pattern as in the case of the *Salmonella* dataset, a separate time-series model is required for each serotype. Another commonly applied approach is regression analysis, a statistical method for fitting a model to observed data to make predictions and place error bounds on these predictions [12]. A primary advantage of this technique is that trends and seasonality observed in the dataset can be easily incorporated as has been demonstrated within a log-linear regression model developed for the Communicable Disease Surveillance Centre (CDSC) public health surveillance data [5]. In particular, this robust model captures the common attributes of both animal and human surveillance data and, thus, was selected for the outbreak detection system for *Salmonella*.

The regression model, in its application to human surveillance data, has been previously described [5, 13]. In this paper, the model is described for the subset

of the *Salmonella* animal health data. To accomplish this several assumptions are made. First, samples submitted for clinical illness are regarded as arriving at a constant rate for each comparable time period in the year. Under this assumption, the denominator data can be considered to remain constant and the observed counts representative of the *Salmonella* infection burden in the livestock population. Second, counts are assumed to follow a Poisson distribution. It is further assumed that each sample (i.e. isolation) is independent and as such the system aims to detect aberrations in reports of infection rather than in the number of livestock units infected. This assumption is made on the basis that as a result of the number of animal movements during the period under study there are not well-defined distinct closed livestock populations but rather a mixed livestock population. It is acknowledged that violation of these assumptions will affect the validity of the regression model. However, upon the identification of an aberration it can be determined, after examining the data, whether the reports are from the same holding or multiple holdings thereby indicating either a local or disperse *Salmonella* infection.

Before applying the regression, the marked seasonal trends in *Salmonella* reporting [4], which can affect the derivation of the expected counts, were addressed. First, as the daily counts of *Salmonella* reports were considered too low to analyse efficiently, these counts were aggregated into months. Therefore, for each calendar month in the historical dataset, the number of counts, or isolations, of *Salmonella* was tallied, using the date the sample was entered onto the database as a reference. This date was chosen as the point of reference rather than the date the sample was taken as it is always known. Next, the historical data were automatically segmented into small windows of time, centred on the current observed month. More specifically, a segment of data incorporated the current month and one month of data either side of the current month. This process was repeated for each year included in the database resulting in three data points per year. The resulting subset of data is referred to as the baseline dataset.

Using the baseline dataset a regression analysis is applied. The model assumes that the current count, y_i , follows a Poisson distribution with mean, μ_i , and variance, $\phi\mu_i$, where ϕ is a dispersion parameter [5]. The dispersion parameter is incorporated to account for the fact the surveillance data may not adequately fit the Poisson distributional assumption of equal

mean and variance due to over- or under-dispersion present in the data. Allowance is also made for a linear trend in the frequency of *Salmonella* reports over time, an assumption that is later tested and validated. To incorporate this linear component within the nonlinear Poisson regression model, a logarithmic link function is used [14]. Hence, the initial log-linear regression model is described by [5]

$$\log(\mu_i) = \beta_0 + \beta_1 t_i, \tag{1}$$

where μ_i is the mean number of cases at time i , β_0 is the regression constant, β_1 is the coefficient for the linear time trend and t_i is the time measured in months. The parameters, β_0 and β_1 are estimated using quasi-likelihood methods rather than maximum-likelihood methods, as is commonly used for estimating parameters within generalized linear models (GLM) due to the presence of the dispersion parameter.

After fitting the regression model, the expected count for each time t_i can be estimated by solving equation (1) for μ_i . This expected count is, however, influenced by the presence of past outbreaks, i.e. historic increases in reporting, affecting the sensitivity of the system for detecting future outbreaks. To address this, the standardized Anscombe residuals are calculated and any value with a large residual is weighted such that points with a large residual are given a lower weight than those points with a small residual. This weighting procedure for each point in the baseline data set is defined by Farrington *et al.* [5].

$$w_i = \begin{cases} S_i^{-2}\gamma & S_i < 1 \\ \gamma & \text{otherwise} \end{cases}, \tag{2}$$

where γ is a constant such that the sum of the weights are equal to the number of baseline data points (n) and S_i^{-2} are the standardized Anscombe residuals. The value of γ is estimated empirically each time the model is initiated. Given that $\sum_{i=1}^n w_i = n$ and using equation (2) yields,

$$n = \sum_{i=1}^{n-x} \gamma + \sum_{i=1}^x S_i^{-2}\gamma = (n-x)\gamma + \gamma \sum_{i=1}^x S_i^{-2}, \tag{3}$$

where x is the number of times $S_i < 1$. Solving equation (3) for γ gives

$$\gamma = \frac{n}{(n-x) + \sum_{i=1}^x S_i^{-2}}, \tag{4}$$

The w_i 's are calculated using equation (2) and are multiplied by each value in the baseline dataset to provide a more indicative observed value in the absence of an outbreak. The log-linear regression

model is refitted using this adjusted baseline dataset and a more realistic expected count is derived in the absence of past outbreaks.

After seasonality and the presence of past outbreaks have been accounted for within the regression model, the assumption of linearity is tested using a 5% significance level and a *t* test. If the linear trend is not significant, the model is re-estimated to exclude the trend. In summary, therefore, the model is estimated twice to adjust for past outbreaks and in the absence of any linear time trend. Using this revised model, the expected count is estimated.

A confidence limit for the expected count is calculated. The upper and lower confidence limits define the interval which contains the expected count with 95% probability [13]. Based on this, a value above the upper confidence limit can be considered to be statistically aberrant. Therefore, the upper confidence limit is defined as the threshold value. These limits are derived using a 2/3 power law transformation to accommodate the fact that the Poisson distribution is often highly skewed [5]. The derived threshold value, *U*, can then be compared to the current observed count.

In order to provide an indication of the degree to which the current observed count deviates from the threshold value, either above or below [5] an exceedance score is derived using the following equation

$$X = \frac{y_0 - \hat{\mu}_0}{U - \hat{\mu}_0},$$

where y_0 is the observed current count, $\hat{\mu}_0$ is the expected count for the current time period, and *U* is the threshold value. An exceedance score >1 indicates that the current count exceeds the threshold value and may require further epidemiological investigation. The main benefit of deriving the exceedance score is it enables the different species and serotypes to be ranked and compared with ease, a factor that is important in communicating the results to the relevant stakeholders.

Testing the early detection system

The early detection system described was implemented using R, a freely available language environment for statistical programming and graphics (see www.r-project.org for details). To test the system, historical data spanning January 1993 to December 2002 were used. The data for 2001 were removed due to the FMD outbreak which affected the collection of samples for animal surveillance. Assuming August 2002 is the

current month, the current counts, expected counts, threshold value and exceedance scores were derived using the full dataset and output from the system.

In addition, outputs for the expected count and threshold counts were derived for each month of the historical dataset spanning from January 1996 to December 2002 for *S. Typhimurium* DT104 and *S. Dublin* in cattle. A minimum of 3 years' data is considered in order to have at least 10 data points in the baseline for which to fit the regression model. Using this time series, the efficiency of the system over time can be determined rather than exploring a single time point as per above. Further, it can be assessed if the algorithm can be applied, successfully, to more than one *Salmonella* spp. serotype.

RESULTS

The patterns of *S. Typhimurium* disease reporting from January 1993 (month 0) to December 2002 (month 108) vary among animal species (Fig. 1). For example, it can be seen that for cattle and to some extent for pigs, that the number of isolations of *S. Typhimurium* has decreased over time with a peak of cases occurring in the mid-1990s. Indeed, the maximum numbers of reports are for cattle (DT104) in 1996 (month 37). In contrast, for ducks, pheasants, and partridges a more constant pattern emerges with, on average, <5 isolations per month which may be an artefact of population size and sampling. For all species considered the most recent counts in the dataset are relatively low. The maximum current count is for pigs, of which 14 reports were recorded in August 2002.

A different pattern in the number of isolations, over time, is illustrated for *S. Dublin* isolated from clinically ill cattle during the period January 1993 to December 2002 (Fig. 2). Here a seasonal pattern with a recent marked incline in the number of reports is observed; 197 isolations were reported in October 2002. Prior to this in August 2002, there were 52 reports of *S. Dublin* from clinically ill cattle.

Counts predicted by the system for *S. Typhimurium*, assuming August 2002 is the current month, are shown in the Table. For the majority of species, particularly turkeys, cattle (non-DT104) and ducks, the expected count is in agreement with the trends observed in Figure 1. A similar observation is made for *S. Dublin* in cattle, for which the expected count (50.0) is in close agreement with the observed count (52). In contrast, however, for DT104 in cattle,

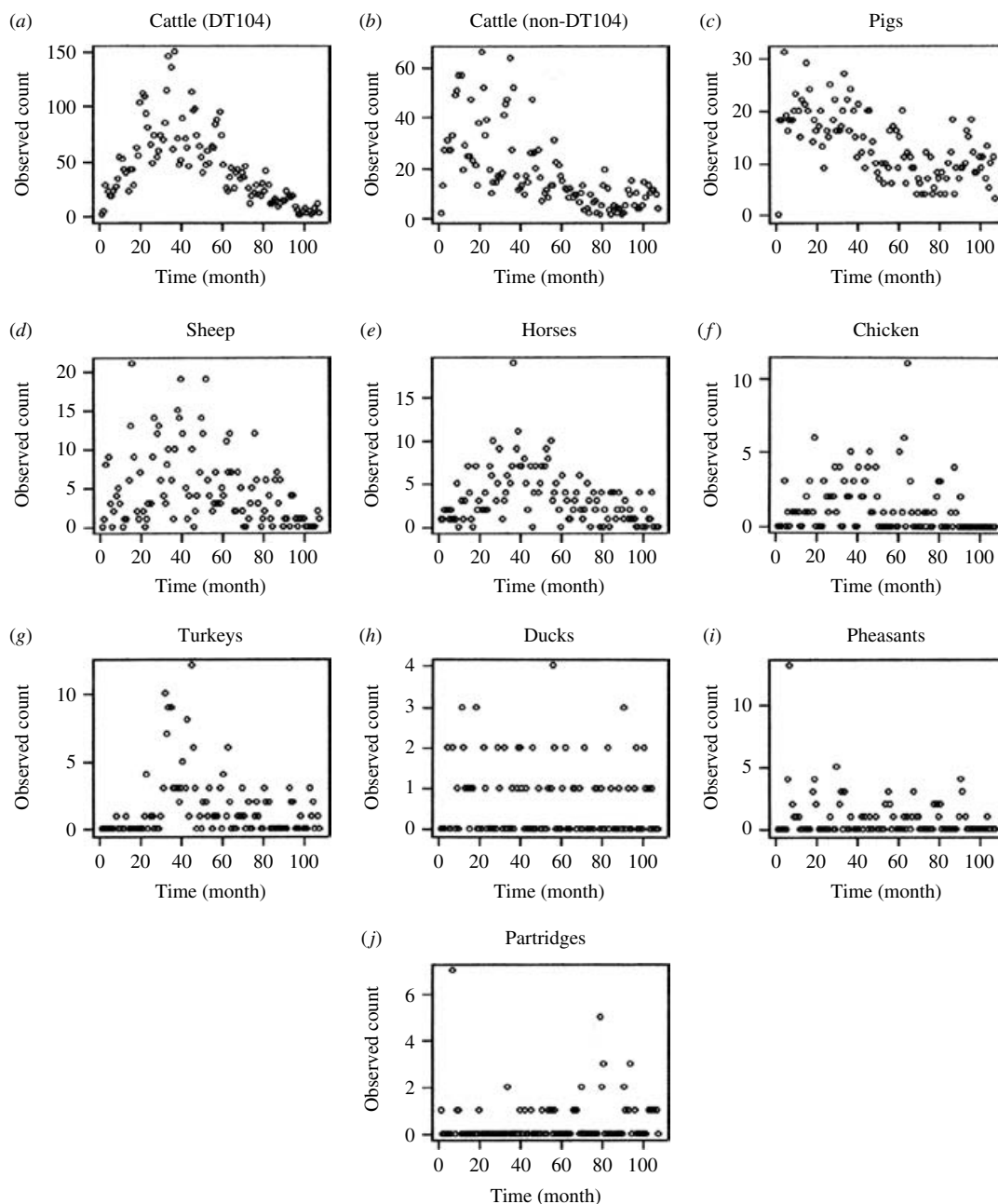


Fig. 1. Number of *S. Typhimurium* isolations from January 1993 (month 0) to December 2002 (month 108) minus data for 2001, in each reported species.

the expected count is substantially greater than is observed based on the trends in Figure 1 which is attributable to the presence of the large epidemic in the 1990s. Thus, it is apparent that the system does not sufficiently adjust for this large-scale epidemic within the historical data. All of the exceedance scores are below 1, except for pigs, and thus a potential outbreak(s) was only detected in one species in this month.

Assessment of the current count compared to the expected count and the threshold value for the period

January 1996 (month 50) to December 2002 (month 108) is depicted in Figure 3 for both *S. Typhimurium* DT104 and *S. Dublin* in cattle. It is apparent that the large epidemic of DT104 in the 1990s is influencing the ability of the algorithm to produce reasonable estimates for the expected count. In contrast, for *S. Dublin*, it can be seen that the observed and expected trends are closely aligned, indicating that the algorithm is able to produce more accurate predictions of potential outbreaks for this serotype. Indeed,

Table. Output from the early detection system for *Salmonella Typhimurium* in various livestock species (current month = August 2002)

Species	Current count	Expected count	Threshold value	Exceedance score
Cattle (DT104 only)	5	18.2	31.3	-1.0
Cattle (non-DT104)	4	3.4	9.6	-0.09
Pigs	14	5.7	13.4	1.0
Sheep	0	1.3	5.6	-0.3
Horses	0	2.7	8.3	-0.5
Chicken	0	1.5	6.0	-0.3
Turkeys	2	2.4	0.1	-0.06
Ducks	0	0.8	1.6	-1.0
Pheasant	0	1.9	6.9	-0.4
Partridge	0	3.4	9.7	-0.6

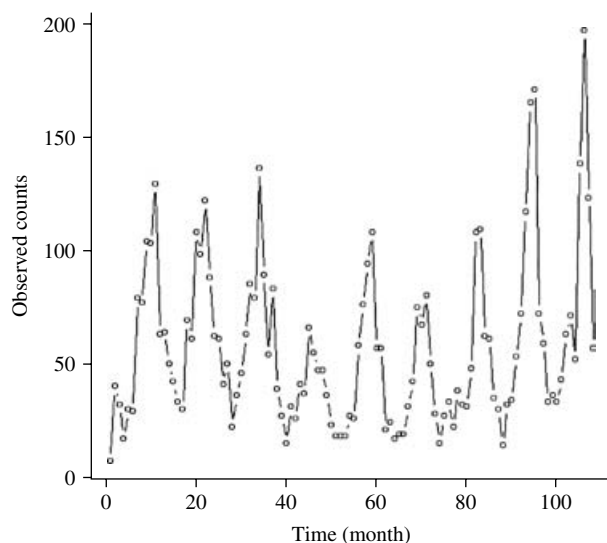


Fig. 2. Number of *S. Dublin* isolations from January 1993 (month 0) to December 2002 (month 108) minus 2001, in clinically ill cattle.

for the later months (months 106 and 107), in particular, the current count is greater than the threshold value suggesting that potential outbreaks have occurred during this period.

Recommendations

It is apparent that, given specific assumptions, an early detection system can be applied to a specific subset of the *Salmonella* surveillance data. However, in order to make full use of all the available data, it is important that measures in the data collection process

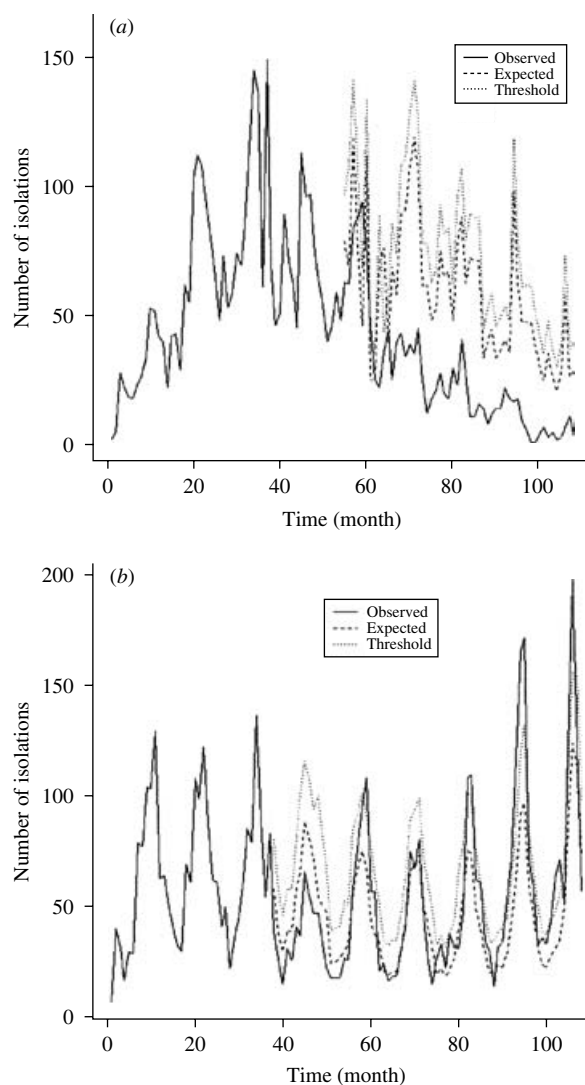


Fig. 3. Illustration of the observed, estimated expected counts and threshold values for (a) *S. Typhimurium* DT104 and (b) *S. Dublin* isolated from clinically ill cattle during the period January 1996 (month 50) to December 2002 (month 108) minus 2001.

are taken to achieve this important goal. These measures include incorporating qualitative information regarding the sampling plan for samples submitted for surveillance activities, encourage samples for clinical illness to be submitted routinely in order to gauge as accurately as possible the disease burden in the population, reduce the number of samples submitted for unknown reasons, collect denominator data (e.g. the number of animals and livestock units tested), and collate within a central database denominator data from private and public laboratories. This list is not exhaustive but highlights the complexities involved with gathering and using surveillance data within an early detection system.

DISCUSSION

Recently, Defra has launched its UK veterinary surveillance strategy the aim of which is to enhance current veterinary surveillance within an improved and comprehensive network of surveillance partners [15]. Part of achieving this goal is developing and applying surveillance methods used in other fields as well as enhancing approaches that are currently used. In the case of *Salmonella* surveillance retrospective data analysis methods are currently applied to the surveillance data to observe trends and seasonal patterns. This analysis is published annually, providing epidemiologists, veterinarians, policy officials and other interested parties with valuable insight into the *Salmonella* infection burden in the population [4]. However, to fulfil the primary aims of the enhanced veterinary surveillance strategy, it is imperative that detection of increases in reports is undertaken soon after they occur in the field. To achieve this, early detection systems can be used.

The work presented here suggests that it is possible to apply an early detection system to animal health data and provide policy-makers and interested parties with informative results regarding the possible presence of outbreaks, particularly where reports have been relatively stable in recent years. However, for *S. Typhimurium*, specifically, due to the large-scale historic epidemic observed for this serotype, further work is required to ensure that predictions produced from such a system are more sensitive to the current stable reporting. The system developed here for detecting *Salmonella* outbreaks is based on a subset of the full database; various issues associated with the data meant that all data could not be included within one system. A full system may be more beneficial to users and, thus, it is important that the data issues are addressed by improving the data collection process and by exploring further statistical methods. For instance, denominator data would be particularly useful, as would the recording of the sampling schemes and maintaining of records. A reduction in the time delay from acquisition of samples to the recording of the results into the central database would also be beneficial. Several of these issues are currently being addressed by one of the main strategic goals of the UK veterinary surveillance strategy which requires better value from surveillance information and activities.

The results outlined for cattle (DT104 only) demonstrate the impact that the presence of a large

epidemic has on the sensitivity of the system. It is, therefore, of vital importance that the early detection system is continually revised and monitored for its ability to correctly identify outbreaks. To address the epidemic of *S. Typhimurium* DT104 in cattle in the 1990s, it is recommended to use the recent 'stable' years of historic data and to acquire more data before implementing a system that will be able to detect any sudden or minor increases in disease reporting. Work has been undertaken to examine the effect on the sensitivity of estimating the expected count by using only data from January 1997 to December 2002. It has been observed that this does provide more accurate predictions but further testing of the system, in the future, when more data becomes available needs to be undertaken to determine the correct balance between the number of years to be included in order to provide an accurate expected count whilst maintaining a sensitive system.

CONCLUSION

Based on the observations made here, we conclude that the use of automated early detection systems for animal health is a feasible aim of any surveillance strategy. We plan to communicate the data issues and extend the system to development of a working system in which all stakeholders will have confidence.

ACKNOWLEDGEMENTS

The authors thank the Department for Environment, Food and Rural Affairs for funding this study.

DECLARATION OF INTEREST

None.

REFERENCES

1. **Lord Phillips of Worth Matravers.** The BSE Inquiry: inquiry into the emergence and identification of Bovine Spongiform Encephalopathy (BSE) and variant Creutzfeldt–Jakob Disease (vJCD) and the action taken in response to it up to 20 March 1996. London: The Stationery Office, 2000.
2. **Thorns C.** Bacterial food-borne zoonoses. *Revue Scientifique et Technique de l'Office International des Epizooties* 2000; **19**: 226–239.
3. **Veterinary Investigation Surveillance Report 2001 and 1994–2001.** Weybridge, Surrey: Veterinary Laboratories Agency Publication, 2001.

4. **Evans S, Kidd S.** *Salmonella in Livestock Production in GB 2002*. Weybridge, Surrey: Veterinary Laboratories Agency Publication, 2002.
5. **Farrington CP, et al.** A statistical algorithm for the early detection of outbreaks of infectious disease. *Journal of the Royal Statistical Society* 1996; **159**: 547–563.
6. **Smith-Palmer A, et al.** Epidemiology of *Salmonella enterica* serovars Enteritidis and Typhimurium in animals and people in Scotland between 1990 and 2001. *Veterinary Record* 2003; **153**: 517–520.
7. **Watier L, Richardson S.** A time series construction of an alert threshold with application to *S. Bovismorbificans* in France. *Statistics in Medicine* 1991; **10**: 1493–1509.
8. **Dessau RB, Steenberg P.** Computerized surveillance in clinical microbiology with time series analysis. *Journal of Clinical Microbiology* 1993; **31**: 857–860.
9. **Stroup DF, Williamson DD, Herndon JL.** Detection of aberrations in the occurrence of notifiable disease surveillance data. *Statistics in Medicine* 1989; **8**: 323–329.
10. **Wallenstein S.** A test for detection of clustering over time. *American Journal of Epidemiology* 1980; **11**: 367–372.
11. **Carpenter TE.** Evaluation and extension of the cusum technique with an application to *Salmonella* surveillance. *Journal of Veterinary Diagnostic Investigation* 2002; **14**: 211–218.
12. **Ryan TP.** *Modern Regression Methods*. Chichester: John Wiley and Sons, 1997.
13. **Farrington CP, Beale AD.** Computer-aided detection of temporal clusters of organisms reported to the Communicable Disease Surveillance Centre. *Communicable Disease Report* 1993; **3**: R78–R82.
14. **Dobson AJ.** *An Introduction to Generalized Linear Models*, 2nd edn. London: Chapman & Hall/CRC, 2002.
15. **Defra.** Partnership, priorities, and professionalism: a strategy for enhancing veterinary surveillance in the UK. London: Department for Environment, Food and Rural Affairs, 2003.