Preventive Veterinary Medicine 127 (2016) 15–20

Contents lists available at ScienceDirect

Preventive Veterinary Medicine

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

Monitoring endemic livestock diseases using laboratory diagnostic data: A simulation study to evaluate the performance of univariate process monitoring control algorithms

Ana Carolina Lopes Antunes[a,*], Fernanda Dórea[b], Tariq Halasa[a], Nils Toft[a]

[a] Section for Epidemiology, National Veterinary Institute, Technical University of Denmark, Denmark
[b] Department of Disease Control and Epidemiology, National Veterinary Institute, Sweden

ARTICLE INFO

Article history:
Received 29 January 2016
Received in revised form 1 March 2016
Accepted 7 March 2016

Keywords:
Endemic disease
Laboratory results
Univariate process monitoring control algorithms

ABSTRACT

Surveillance systems are critical for accurate, timely monitoring and effective disease control. In this study, we investigated the performance of univariate process monitoring control algorithms in detecting changes in seroprevalence for endemic diseases. We also assessed the effect of sample size (number of sentinel herds tested in the surveillance system) on the performance of the algorithms.

Three univariate process monitoring control algorithms were compared: Shewart *p* Chart[1] (PSHEW), Cumulative Sum[2] (CUSUM) and Exponentially Weighted Moving Average[3] (EWMA). Increases in seroprevalence were simulated from 0.10 to 0.15 and 0.20 over 4, 8, 24, 52 and 104 weeks. Each epidemic scenario was run with 2000 iterations. The cumulative sensitivity[4] (CumSe) and timeliness were used to evaluate the algorithms' performance with a 1% false alarm rate. Using these performance evaluation criteria, it was possible to assess the accuracy and timeliness of the surveillance system working in real-time.

The results showed that EWMA and PSHEW had higher CumSe (when compared with the CUSUM) from week 1 until the end of the period for all simulated scenarios. Changes in seroprevalence from 0.10 to 0.20 were more easily detected (higher CumSe) than changes from 0.10 to 0.15 for all three algorithms. Similar results were found with EWMA and PSHEW, based on the median time to detection. Changes in the seroprevalence were detected later with CUSUM, compared to EWMA and PSHEW for the different scenarios. Increasing the sample size 10 fold halved the time to detection (CumSe = 1), whereas increasing the sample size 100 fold reduced the time to detection by a factor of 6.

This study investigated the performance of three univariate process monitoring control algorithms in monitoring endemic diseases. It was shown that automated systems based on these detection methods identified changes in seroprevalence at different times. Increasing the number of tested herds would lead to faster detection. However, the practical implications of increasing the sample size (such as the costs associated with the disease) should also be taken into account.

1. Introduction

Surveillance systems are critical for the effective and timely control of infectious diseases. Surveillance based on the continuous monitoring of secondary animal health data sources is a growing field, but the ability of automated systems to detect changes in the disease burden depends upon the choice of data source, their representativeness and sampling strategy (Buckeridge, 2007).

Sentinel surveillance systems are used when the health status of a population is periodically assessed based on a limited number of herds. These systems can be used for monitoring trends in diseases, in order to identify outbreaks and monitor the burden of disease in a population, providing a more cost-effective

Abbreviations: PRRSV, Porcine Reproductive and Respiratory Syndrome Virus; SPF System, Specific Pathogen Free System; EWMA, Exponentially Weighted Moving Average; CUSUM, Cumulative Sums; PSHEW, Shewart *p* Chart; UCL, Upper control limit; CumSe, Cumulative sensitivity.

* Corresponding author at: Section for Epidemiology, National Veterinary Institute, Technical University of Denmark, Bülowsvej 27, 1870 Frederiksberg C, Denmark.
E-mail address: aclan@vet.dtu.dk (A.C. Lopes Antunes).

[1] Shewart *p* Chart (PSHEW).
[2] Cumulative Sum (CUSUM).
[3] Exponentially Weighted Moving Average (EWMA).
[4] Cumulative sensitivity (CumSe).

http://dx.doi.org/10.1016/j.prevetmed.2016.03.005
0167-5877/

alternative to other disease surveillance methods. The testing frequency and sample size required for sentinel surveillance are dependent upon several factors, such as the goal of the surveillance, the etiology of the infectious agent, and the diagnostic test sensitivity and specificity (McCluskey, 2003).

One example of sentinel surveillance is the Danish monitoring program for Porcine Reproductive and Respiratory Syndrome Virus[5] (PRRSV), targeting swine breeding herds. Despite disease control efforts, PRRSV has continued (since its first diagnosis in 1992) to contribute to economic losses due to mortality in piglets, respiratory problems in growers and finishers, and reproductive problems in sows (Kvisgaard et al., 2011). The surveillance is primarily based on serological testing in order to maintain the Specific Pathogen Free System[6] (SPF System) certificate (Specific Pathogen Free System (SPF-SuS), 2015). The frequency of testing is dependent upon the health status defined as "red", "blue" and "green". The majority of Danish breeding herds have the "red" health status and therefore are tested on a monthly basis. In order to gain or maintain the SPF status, farmers must participate in a voluntary control program, for which they must provide health declarations and information on their herd health status for seven diseases, including PRRSV.

During the past decade, several studies have applied statistical quality control methods for syndromic surveillance in human and veterinary medicine (Dupuy et al., 2015; Dórea et al., 2013a,b; Jackson et al., 2007; Mandl et al., 2004). These studies mainly focused on detecting simulated outbreaks representing different scenarios of disease spread sometimes associated with emerging or re-emerging diseases. However, it may not be possible to generalize the performance of these algorithms in detecting disease outbreaks when monitoring changes in the burden of endemic diseases. Due to the availability of control measures such as vaccination or health management programs, the dynamic of disease spread is expected to be different for endemic diseases, resulting in a lower incidence rate when compared to exotic diseases. Moreover, the natural immunity developed from previous exposure to the agent also reduces an animal's susceptibility to endemic diseases. Therefore, it is unlikely that "extreme" changes in incidence and prevalence would be observed for diseases already present (and controlled) in the population. The dynamics of within-herd and between-herd endemic disease transmission also depends on the nature of the pathogen (Carslake et al., 2011), and can contribute to different temporal patterns of endemic disease spread.

For endemic diseases, it might be beneficial to monitor changes in disease prevalence rather than incidence. In these cases, the data differ from that obtained from traditional biosurveillance (generally related to incidence monitoring), as it is focused on the endemic scenario with less frequently sampled data. Moreover, this reflects the added complexity of monitoring endemic diseases, as disease burden is affected not only by the incidence but also by the disease duration and recovery rate.

In this study, we investigated the performance of three univariate process monitoring control algorithms commonly used in biosurveillance (Wagner and Moore, 2006) when applied to endemic disease monitoring. The algorithms were chosen for this study based on the simulated scenarios and on the type of simulated data (proportion data). The aim was to demonstrate that monitoring based on the weekly seroprevalence of a subset of the population for an endemic disease could be used to detect changes in disease occurrence in an accurate and timely manner. In addition, the impact of sample size (i.e. the number of sentinels) was explored. The design of our study was based on the Danish PRRSV monitoring program.

2. Methods

2.1. The Danish PRRSV monitoring program

Compulsory serological testing is performed on a monthly basis for all herds certified as SPF, which includes almost all Danish breeding herds (Specific Pathogen Free System (SPF-SuS), 2015).

Laboratory submission data stored in the National Veterinary Institute—Technical University of Denmark (DTU Vet) information management system and in the Laboratory for Swine Diseases-SEGES Pig Research Centre (VSP-SEGES) were used to determine the weekly number of Danish breeding herds tested for PRRSV and the corresponding between-herd seroprevalence from January 2007 to December 2014. Each laboratory submission consisted of individual blood samples collected from different animals in the same herd on the same day. Only submissions where at least two individual blood samples were subject to serological tests, including Blocking Enzyme-Linked Immunosorbent Assay (ELISA) (Sørensen et al., 1997; IDEXX, Ludwigsburg, Germany) and/or Immunoperoxidase monolayer assay (IPMA) (Bøtner et al., 1994), for one or both PRRSV strains were included in the analysis. Results from experimental studies were not included in the analysis.

Herds were classified as PRRSV seropositive when at least two individual blood samples in each submission tested PRRSV positive, independently of the PRRSV strain. The between-herd PRRSV seroprevalence was calculated weekly as the proportion of PRRSV positive herds within the total number of herds tested. The average between-herd PRRSV seroprevalence was 0.10 and the median weekly number of herds tested for PRRSV was 54 (minimum = 4, maximum = 85, standard deviation = 12.7).

2.2. Simulation experiment

As no additional knowledge of the true PRRSV seroprevalence was available, a simulation experiment was devised to derive the number of seropositive herds over a week, in order to control the development of changes. A baseline scenario of PRRSV seroprevalence of 0.1 was defined based on the data. In this scenario, the number of positive herds (X) per week from 2007 to 2014 were drawn from a binomial distribution ($X \sim bin(n, p)$) with a probability (p) of 0.1 and a sample size (n) equal to the number of Danish breeding herds tested for PRRSV in a given week. The weekly seroprevalence was calculated as the simulated number of seropositive herds, divided by the total number of herds tested that week. This simulation produced a stationary process representing an endemic disease under control.

The seroprevalence was increased from $p = 0.1$ to $p = 0.15$ and $p = 0.20$, over 4, 8, 24, 52 and 104 weeks. These 10 scenarios were designed to represent possible seroprevalence increases for the disease and population under study, considering the control measures in place. The final week of the simulated increase corresponded to the maximum increase. Following this, seroprevalence was maintained at the increased level (0.15 and 0.20). Two of these scenarios are illustrated in Fig. 1. The simulated increases in seroprevalence were started in random weeks between 2009 and 2012, and the weeks preceding this increase were used to train the algorithms.

2.3. Univariate process monitoring control algorithms

Three univariate process monitoring control algorithms used in previous studies in veterinary science (Dórea et al., 2013a,b; Dupuy et al., 2015) were investigated: Exponentially Weighted Moving Average (EWMA), Cumulative Sums (CUSUM) and Shewart p Chart

[5] Porcine Reproductive and Respiratory Syndrome Virus (PRRSV).
[6] Specific Pathogen Free System (SPF System).