Preventive Veterinary Medicine 121 (2015) 1–7

Contents lists available at ScienceDirect

Preventive Veterinary Medicine

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

Investigating the potential of reported cattle mortality data in Switzerland for syndromic surveillance

Rahel Struchen[a,*], Martin Reist[b], Jakob Zinsstag[c], Flavie Vial[a]

[a] Veterinary Public Health Institute, Vetsuisse Faculty, University of Bern, Schwarzenburgstrasse 155, 3003 Bern, Switzerland
[b] Swiss Federal Food Safety and Veterinary Office, Schwarzenburgstrasse 155, 3003 Bern, Switzerland
[c] Swiss Tropical and Public Health Institute, University of Basel, Socinstrasse 57, 4051 Basel, Switzerland

ARTICLE INFO

Article history:
Received 13 October 2014
Received in revised form 5 April 2015
Accepted 19 April 2015

Keywords:
Syndromic surveillance
Cattle mortality
Early detection
Time-series analysis
Animal health surveillance

ABSTRACT

Systems for the identification and registration of cattle have gradually been receiving attention for use in syndromic surveillance, a relatively recent approach for the early detection of infectious disease outbreaks. Real or near real-time monitoring of deaths or stillbirths reported to these systems offer an opportunity to detect temporal or spatial clusters of increased mortality that could be caused by an infectious disease epidemic. In Switzerland, such data are recorded in the "Tierverkehrsdatenbank" (TVD). To investigate the potential of the Swiss TVD for syndromic surveillance, 3 years of data (2009–2011) were assessed in terms of data quality, including timeliness of reporting and completeness of geographic data. Two time-series consisting of reported on-farm deaths and stillbirths were retrospectively analysed to define and quantify the temporal patterns that result from non-health related factors.

Geographic data were almost always present in the TVD data; often at different spatial scales. On-farm deaths were reported to the database by farmers in a timely fashion; stillbirths were less timely. Timeliness and geographic coverage are two important features of disease surveillance systems, highlighting the suitability of the TVD for use in a syndromic surveillance system. Both time series exhibited different temporal patterns that were associated with non-health related factors. To avoid false positive signals, these patterns need to be removed from the data or accounted for in some way before applying aberration detection algorithms in real-time. Evaluating mortality data reported to systems for the identification and registration of cattle is of value for comparing national data systems and as a first step towards a European-wide early detection system for emerging and re-emerging cattle diseases.

1. Introduction

Systems for the individual identification and registration of cattle were implemented in all European Union (EU) states in the aftermath of the bovine spongiform encephalopathy (BSE) crisis in 1996 (Council Regulation (EC) No. 1760/2000 of 17 July 2000). These computerised, databased systems were designed to restore consumer faith in food safety, by enabling the tracing of cattle suspected of having BSE from the slaughterhouse back to their various holdings of origin. Since then, they have proven valuable for other types of epidemiological investigations, for example tracing the movements of animals potentially infected with other agents such as bovine viral diarrhoea virus (Presi et al., 2011), estimating population dynamics for modelling disease transmission (Green et al., 2006) and designing cost-effective disease control and monitoring programs (Blickenstorfer et al., 2011; Schärrer et al., 2014).

These identification systems have gradually been receiving attention for use in syndromic surveillance (Dupuy et al., 2013). Syndromic surveillance is a recent surveillance approach, based on the continuous monitoring of unspecific health related data in (near) real-time. Its primary purpose is the early detection of potential health threats, to inform timely and effective control measures (Triple-S Project, 2011). Assessing the impact of identified events on population health (Elliot et al., 2010) is another reported benefit. Data most accessible for syndromic surveillance are those that are stored electronically, collected and analysed in a timely fashion and that have extensive geographic, demographic and temporal coverage (Mandl et al., 2004). In veterinary public health, syndromic data are typically clinical observations collected from veterinary practitioners (Vourc'h et al., 2006) or diagnostic laboratory test requests (Dórea et al., 2013). Such pre-diagnostic data are assumed to contain earlier, but weaker signatures of a disease outbreak (Yahav and Shmueli, 2007). Various statistical algorithms exist (Unkel et al.,

* Corresponding author. Tel.: +41 31 631 5741; fax: +41 31 631 5749.
E-mail address: rahel.struchen@vetsuisse.unibe.ch (R. Struchen).

http://dx.doi.org/10.1016/j.prevetmed.2015.04.012
0167-5877/

2012) for identifying unexpected patterns in these data that may result from infectious disease outbreaks.

Mortality (or fertility) data from cattle identification systems may be useful for syndromic surveillance. In a retrospective analysis of mortality data from the French National Cattle Register, Perrin et al. (2012) reported a positive association between the spatiotemporal distribution of weekly excess mortality and notifications of infected herds during the Bluetongue outbreak in 2007 and 2008. Such systems contain large amounts of data routinely collected on a daily basis over several years. Historical data are needed for constructing a baseline model defining expected normal behaviour in time series from these databases. Reporting is compulsory, ensuring reasonably good coverage of the population. Existing data collection, transmission and storage infrastructures can be used, making surveillance convenient and reducing surveillance costs. However, these data were originally collected for purposes other than surveillance, and for this reason may be of insufficient quality, have limited timeliness or may contain biases. The temporal patterns observed in the data, such as seasonality, day-of-week effects or global trends (Nöremark et al., 2009; Robinson and Christley, 2006), are to a large degree caused by factors which are not health-related. These patterns need to be removed from the data or accounted for before applying aberration detection algorithms in a prospective fashion (Shmueli and Burkom, 2010).

Early detection of animal diseases is an important component of the '*Swiss Animal Health Strategy 2010+*[1]' that aims to maintain and improve the high standard of animal health in Switzerland. The Swiss Federal Food Safety and Veterinary Office (FSVO) plans to build a national early detection system by 2016, using information from various sources. Many syndromic surveillance systems rely on clinical data collected from veterinarians (Dórea et al., 2011). In Switzerland, there was no computerised system for recording clinical data in place yet by the time this study started and therefore, alternative data were being evaluated. For example, the (near) real-time monitoring of cattle deaths could be used to identify temporal or spatial clusters of increased mortality, which may be indicative of a disease outbreak. Outbreaks of emerging or re-emerging diseases such as Rift Valley fever or brucellosis, and changes in endemic diseases such as botulism or leptospirosis, may produce clusters of excess cases in reported mortality data. In Switzerland, cattle mortalities are reported by farmers to the system for the identification and registration of cattle, the “Tierverkehrsdatenbank” (TVD). The objectives of our study were to assess the quality of the TVD data, and to define the temporal patterns caused by non-health related factors in two non-slaughter mortality time-series. Understanding non-health related patterns is a prerequisite step before choosing appropriate prospective methods for detecting temporal aberrations in the data that might be linked to disease outbreaks.

2. Materials and methods

2.1. Data source

In Switzerland, it has been compulsory, since 2000, for cattle farmers to report all births and deaths of animals on their holding and all movements to and from their holding to the TVD. Births are required to be reported within 30 days, whereas all movements (on- and off-farm) and non-slaughter deaths are required to be reported within 3 days (Animal Health Ordinance (AHO), SR 916.401). Reporting is either electronic via the internet or by paper forms. Detailed information is captured for animals (e.g. sex, breed) and farms (e.g. location, farm type). Monetary incentives to report do exist. For example, farmers receive a carcass disposal fee for each dead calf whose birth had been previously reported to the TVD, and slaughterhouses receive a similar fee for each slaughtered animal whose movement history is complete. Missing or incomplete reports are penalised by reducing the incentive. Two possible syndromic indicators were identified in the TVD: on-farm deaths and stillbirths. Both events are recorded as separate entries. On-farm deaths include deaths occurring either naturally or by euthanasia. Stillbirth reporting is not compulsory and therefore an explicit definition of stillbirths could not be made.

2.2. Data extraction

All reported on-farm deaths and stillbirths for the period from January 1st 2009 to December 31st 2011 were extracted from the TVD. Data prior to 2009 were available, but were excluded because analyses had shown that the quality of the data improved notably due to incentives that were implemented in 2009. The date that each event (on-farm death or stillbirth) was reported to have occurred and the date it was reported were extracted. Additional data about the animal affected (e.g. breed) and the farm that reported the event (e.g. geographical coordinates) were also extracted. Data were stored in a PostgreSQL database (The PostgreSQL Global Development Group) and data handling using SQL was performed with Squirrel (Bell et al.).

2.3. Data quality and descriptive statistics

Data quality was assessed by estimating the timeliness of report submission and the completeness of records. Timeliness was defined as the time between the reported occurrence of an event and its reporting to the system and is further referred to as reporting timeliness. Completeness was estimated by calculating the proportion of reports with missing geographic data. Geographic data were chosen because there are critically important for geo-locating disease and outbreak occurrences.

For the descriptive analyses, mortality data were stratified by sex and production type. Production type was subset into dairy, beef, mixed and other according to the breed of cattle reported. On-farm deaths were also divided into five categories by the age at death: (1) up to 7 days old, (2) 8–120 days, (3) 121 days to 1 year, (4) 1–2 years, and (5) more than 2 years. Classification was based on livestock units used to calculate direct payments offered to farmers for their services for the common good. For the time series analyses, daily time series for the total numbers of on-farm deaths and stillbirths were generated. All statistical analyses were performed in R (R Core Team, 2013) using the packages timeDate (Wuertz et al., 2013), TSA (Chan and Ripley, 2012) and gcmr (Masarotto and Varin, 2012).

2.4. Model building and comparison

Regression models were fitted to the data to determine the effects of trend, seasonality, day of the week and bank holidays. Poisson and negative-binomial regression models were applied to the daily counts of on-farm deaths or stillbirths. Alternative models tested included different variants for some of the predictors: trend was defined either as a continuous time variable or as categorical variables for the years. Seasonality was modelled using either categorical variables for each month or months grouped into two or four seasons, or a sinusoidal function using the term: $\sin(2\pi t/365)+\cos(2\pi t/365)$ where t is the day number from 1 to 1095. Days of the week were included using categorical variables either for each day or grouped into Mondays, other weekdays and weekends. Bank holidays were included in the model using a variable for common Swiss holidays and a second variable for the

[1] See http://www.blv.admin.ch/gesundheit_tiere/03007/index.html?lang=en.