Preventive Veterinary Medicine 119 (2015) 216–221

Contents lists available at ScienceDirect

Preventive Veterinary Medicine

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

Short communication

Syndromic surveillance system based on near real-time cattle mortality monitoring

G. Torres[a,*], V. Ciaravino[b,c], S. Ascaso[b], V. Flores[b], L. Romero[a], F. Simón[d]

[a] Subdirección General de Sanidad e Higiene Animal y Trazabilidad, Ministerio de Agricultura, Alimentación y Medio Ambiente, Madrid, Spain
[b] Centro Nacional de Epidemiologia, Instituto de Salud Carlos III, Madrid, Spain
[c] European Programme for Intervention Epidemiology Training (EPIET), ECDC, Stockholm, Sweden
[d] Centro de Coordinación de Alertas y Emergencias Sanitarias, Ministerio de Sanidad, Servicios Sociales e Igualdad, Madrid, Spain

ARTICLE INFO

Article history:
Received 19 July 2014
Received in revised form 4 March 2015
Accepted 7 March 2015

Keywords:
Syndromic surveillance
Cattle mortality
Time series
Cycling regression model
Cumulative sums

ABSTRACT

Early detection of an infectious disease incursion will minimize the impact of outbreaks in livestock. Syndromic surveillance based on the analysis of readily available data can enhance traditional surveillance systems and allow veterinary authorities to react in a timely manner.

This study was based on monitoring the number of cattle carcasses sent for rendering in the veterinary unit of Talavera de la Reina (Spain). The aim was to develop a system to detect deviations from expected values which would signal unexpected health events. Historical weekly collected dead cattle (WCDC) time series stabilized by the Box–Cox transformation and adjusted by the minimum least squares method were used to build the univariate cycling regression model based on a Fourier transformation. Three different models, according to type of production system, were built to estimate the baseline expected number of WCDC.

Two types of risk signals were generated: point risk signals when the observed value was greater than the upper 95% confidence interval of the expected baseline, and cumulative risk signals, generated by a modified cumulative sum algorithm, when the cumulative sums of reported deaths were above the cumulative sum of expected deaths.

Data from 2011 were used to prospectively validate the model generating seven risk signals. None of them were correlated to infectious disease events but some coincided, in time, with very high climatic temperatures recorded in the region. The harvest effect was also observed during the first week of the study year.

Establishing appropriate risk signal thresholds is a limiting factor of predictive models; it needs to be adjusted based on experience gained during the use of the models. To increase the sensitivity and specificity of the predictions epidemiological interpretation of non-specific risk signals should be complemented by other sources of information.

The methodology developed in this study can enhance other existing early detection surveillance systems. Syndromic surveillance based on mortality monitoring can reduce the detection time for certain disease outbreaks associated with mild mortality only detected at regional level. The methodology can be adapted to monitor other parameters routinely collected at farm level which can be influenced by communicable diseases.

* Corresponding author at: Subdirección General de Sanidad e Higiene Animal y Trazabilidad, Ministerio de Agricultura Alimentación y Medio Ambiente, C/ Almagro, 33, 28071 Madrid, Spain. Tel.: +34 91 3471639; fax: +34 91 3471639.
E-mail address: gjtorresp@magrama.es (G. Torres).

http://dx.doi.org/10.1016/j.prevetmed.2015.03.003
0167-5877/

1. Introduction

Veterinary disease surveillance methods have experienced an important evolution, moving from costly large labour intensive campaigns to more epidemiological targeted selective methods (Schwabe, 1982). Informatics and data mining techniques are now incorporated into veterinary epidemiology allowing the analysis of large amounts of data (Dorea et al., 2011). This provides an opportunity to explore data that are often generated automatically for purposes other than disease surveillance (Dupuy et al., 2013).

Syndromic surveillance based on cattle mortality monitoring at near real-time could aid in early detecting unexpected health events. Moreover, certain health events that result in a mortality rate too small to be detected at farm level might produce a signal at regional level using an appropriate algorithm (Perrin et al., 2012).

The aim of this study was to model already available historical rendered cattle data to prospectively detect departures from expected weekly values. The algorithm will generate automatic risk signal intending to be part of a near-real time prospective syndromic surveillance.

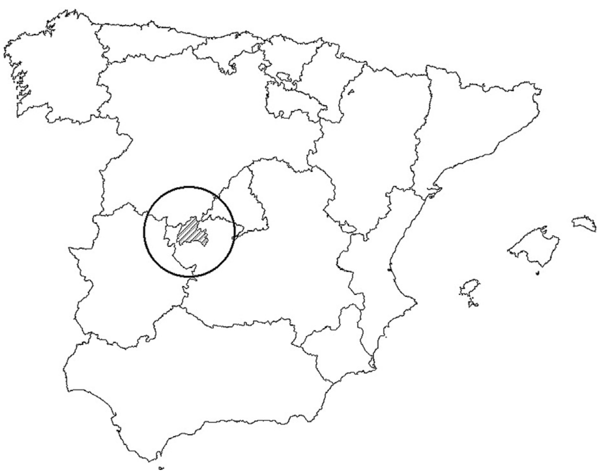
2. Material and methods

2.1. Study area

The study was carried out in Spain in the area defined by the veterinary unit of Talavera de la Reina (Autonomous Community of Castilla-La Mancha, Spain). It comprises 34 municipalities (Map 1).

2.2. Data source

The main dataset was obtained by compiling two databases. The Spanish National Cattle Register (*REGA*) which contains cattle holdings data as requested by European Union regulation. Farmers must report to REGA soon after there are any changes to the census data of the farm. This database is regularly supervised by official veterinarians. The second data source was the national fallen stock database, which is updated daily and it is managed by the National Agriculture Insurance Agency (ENESA) which is a government agency under the authority of the Ministry of Agriculture, Food and Environment.



Map 1.

The datasets were merged using a common and unique farm identifier and cleaned to create a dataset stored in Microsoft Excel®.

The study dataset included the following information attributed to each collected carcass: date of collection; geographical coordinates of the farm, type of production system (fatten, suckler or dairy) and unique farm identifiers. A historical dataset from 2005 until 2010 was used to build the models. Collected data from 2011 were used to validate the model's predictions.

2.3. Case definition and exclusion criteria

Dead cattle collected by authorized rendering companies in the study area during the study period and included in the national fallen stock database.

We excluded records that had missing values for date of carcass collection, farm identifiers, municipality or type of production.

2.4. Statistical analysis

A descriptive analysis of the historical mortality data and its temporality was performed, stratified by type of production system: dairy, suckler or fatten cattle.

Characterization of the specific time series was carried out using summary statistics by week, month and year of reporting.

Data variance was first stabilized by applying the Box–Cox transformation which chooses an optimal transformation to remediate deviations from the assumptions of the linear regression model (Box and Cox, 1964). The Shapiro–Wilk test confirmed data were normally distributed after applying the Box–Cox transformation.

Historical weekly collected dead cattle (WCDC) time series transformed and adjusted by the minimum least squares method were used to build a cycling regression model based on a Fourier transformation with the purpose of controlling the trend and seasonality (Serfling, 1963). The cyclical regression model is a linear regression model based on weekly counts that introduce sine and cosine terms to account for seasonal waves.

The trend of the time series was controlled before applying the Fourier transformation. $Y = \alpha + \beta t + \frac{r \cdot \cos(2\pi)(t+\text{phase})}{\text{period}}$ where Y = transformed number of weekly collected dead cattle.

Models were run with weekly data (from Monday to Sunday). When appropriate the 53rd week was excluded in order to maintain the same number of weeks in each of the study years.

Three different models were built according to the type of production systems to estimate specific mortality baseline for dairy, suckler and fatten cattle.

The historical time series used as a reference to build the models were selected based on the stability of data. The Kruskal–Wallis test was performed to analyse variance and