

Vetsyn: Veterinary Syndromic Surveillance Streamlined into one R Package

Fernanda C. Dórea*, Stefan Widgrén and Ann Lindberg

Department of Epidemiology and Disease Control, National Veterinary Institute, Uppsala, Sweden

Objective

To describe an R package that was designed to provide ready implementation of veterinary syndromic surveillance systems, from classified data to the generation of alerts and an html interface.

Introduction

The field of veterinary syndromic surveillance (VSS) is developing fast, with countries exploring a great variety of data sources [1]. After implementing two VSS systems we have demonstrated that the steps from classified data to full system implementation can be streamlined, and published a guideline for implementation [2]. All the steps described have been made available in an R package (<https://github.com/nandadorea/vetsyn>). We aim to demonstrate the utility and potential of this streamlined approach.

Methods

The *vetsyn* package takes advantage of object-oriented programming in R. Using S4 classes, we have defined a *syndromic* class, which stores in a single object all the data needed and produced during syndromic monitoring. The object contains the following slots:

- `@observed`: a matrix storing the observed event counts of one syndrome per column, and one time point (day or week) per row ($t \times s$). This is constructed with a single function starting from a dataset of health events, where the events have been classified into syndromes.
- `@dates`: a data frame storing date related information for each time point in the dataset (rows in all other slots).
- `@baseline`: a $t \times s$ matrix which serves as a training baseline for outbreak-signal detection algorithms. During the implementation of the system a function can be used to construct this baseline from historical data, removing outliers based on parametric (generalized linear models) or non-parametric methods (moving percentiles). After the system is implemented the baseline continues to grow by automatically removing detected outbreak-signals from observed data [3].

- `@alarms`: a $t \times s \times x$ array ($t \times s \times x$ algorithms for detection), in which a third dimension is added to accommodate for the parallel use of multiple aberration detection algorithms. Users can choose to use control charts (EWMA, CUSUM and Shewhart) or the Holt-Winters exponential smoothing to detect outbreak-signals, and those can be coupled with parametric (generalized linear models) or non-parametric (differencing) pre-processing to remove temporal effects. The user can also set up multiple detection limits, in which case for each syndrome, time-point, and algorithm used, the stored result is a detection score, rather than a binomial indicator of whether an alarm happened.

- `@UCL`: a $t \times s \times x$ array, which stores, for each syndrome, time-point and algorithm used, the upper confidence limit of the detection method chosen, that is, the minimum value that would have generated an alarm.

One function helps the user set up email alerts to be generated in case of an alarm. Another function allows the user to generate an html interface where all syndromes monitored are tabulated, an alarm score is plotted for each syndrome (Figure), and time series graphs show any alarms detected.

Conclusions

We have combined several statistical analyses methods already tested for use in animal data streams, and R functions to support data management and output visualization. A full tutorial is available, which further helps streamlining the process, facilitating the implementation of syndromic surveillance systems by veterinary epidemiologists.

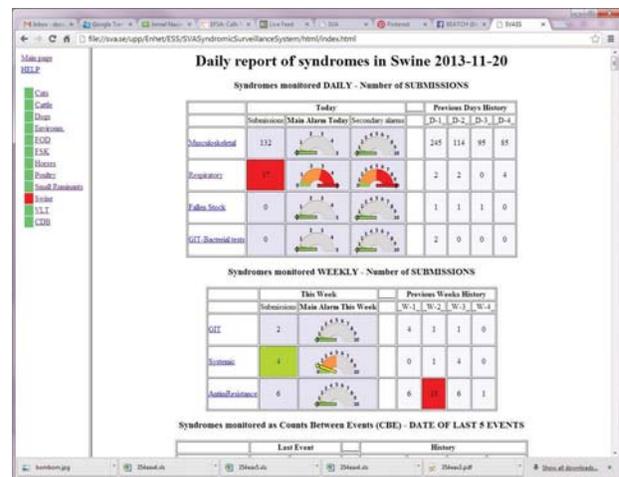


Figure: Example of the html interface generated to display the information stored in a *syndromic* object created with the *vetsyn* package.

Keywords

Animal health; syndromic surveillance; R programming

References

- [1] Dupuy C, Bronner A, Watson E, et al. Inventory of veterinary syndromic surveillance initiatives in Europe (Triple-S project): Current situation and perspectives. *Prev Vet Med* 2013 (111) 3-4 220-9
- [2] Dórea FC, Lindberg A, McEwen BJ, et al. Syndromic surveillance using veterinary laboratory requests: a practical guide informed by experience with two systems. *Prev Vet Med* 2014 .DOI: 10.1016/j.prevetmed.2014.04.001
- [3] Dórea FC, McEwen BJ, McNab BW, et al. Syndromic surveillance using veterinary laboratory data: data pre-processing and algorithm performance evaluation. *J Royal Soc Interface*. 2013 DOI:10.1098/rsif.2013.0114

*Fernanda C. Dórea

E-mail: fernanda.dorea@sva.se

