Retrospective analysis of 11 years of livestock necropsy data: evaluation for animal health surveillance

C FAVERJON[1*], S KÜKER[1], L FURRER[2], J BEREZOWSKI[1], H POSTHAUS[3], F RINALDI[2], F VIAL[1†]

[1] Veterinary Public Health Institute, Vetsuisse Faculty, University of Bern, Liebefeld, Switzerland; [2] Institute of Computational Linguistics, University of Zürich, Zürich, Switzerland; [3] Institute of Animal Pathology, University of Bern, Bern, Switzerland; [†] Current affiliation: Epi-Connect, Djupdalsvägen 7, 14251 Skogås, Sweden

*celine.faverjon@vetsuisse.unibe.ch

Abstract

Livestock necropsy reports from diagnostic laboratories may be of interest for disease surveillance. However, they are usually created using natural language making the extraction of relevant data complicated. To evaluate necropsy reports for animal health surveillance, we first developed a text mining tool to automatically classify necropsy reports for cattle and pigs into 13 syndromic categories primarily based on topography of organ systems, before retrospectively describing 11 years of necropsy submissions to one of the two largest Swiss veterinary diagnostic laboratories using time series analysis. The main syndromic categories identified were gastrointestinal system, serous membranes and respiratory system. The proportion of submissions represented by different syndromes and their seasonal patterns differed between age classes, in particular for cattle. Thus, we recommend that the different age classes should be monitored separately should these data be integrated in a prospective surveillance system.

Keywords: *Necropsy data, syndromic surveillance, text mining, post mortem*

Introduction

Syndromic surveillance (SyS) is increasingly used to enhance traditional passive surveillance systems in veterinary medicine or public health. SyS seeks to identify the early and often weak signal of an outbreak, through the routine analysis of pre-diagnostic and unspecific health-related data.

Different types of data have already been explored in veterinary SyS, but an increasing amount of investigations are performed to evaluate potential new data sources. Data related to dead animals offer rare information about diseases of unknown etiology. Post-mortem meat inspection data have been used for routinely monitoring animal health, but meat inspection provides limited information about current diseases in a population because only "clinically-healthy" animals are sent to slaughter (1). The monitoring of on-farm mortality, also referred to as fallen stock, provides information about a very specific and relevant subset of the population: "clinically-ill" animals that have died on the farm and/or were euthanized.

Necropsies performed by diagnostic laboratories on fallen stock could be used to provide more detailed information on the cause of death. In wildlife, necropsy data are already an important source of information about infectious disease (2,3) and are well integrated into national wildlife surveillance systems (e.g. in Sweden (4) and Canada (5)). However, necropsy data coming from livestock have not yet been integrated into national disease surveillance systems in Europe.

Systematic evaluation of livestock necropsy data could be of great value for early detection of new diseases, or of changing disease patterns in a population (6). These data are already being used in wildlife (3) to identify pathological patterns. They could also be used to produce time series of submissions, and/or of syndromic groups (i.e. all animals presenting with the same pathological pattern). These time series could then be easily monitored in an operationalized SyS system using aberration detection algorithms (7). However, necropsy reports tend to be created in natural language, making it difficult to obtain the structured data required for aberration detection algorithms. Advanced analytical techniques such as text mining are becoming increasingly important and may help defining syndromic groups in free text necropsy reports.

The objective of our study was to evaluate free text livestock necropsy reports that were automatically classified into syndromes using a tool developed in a previous study (8). As a proof of concept, we used data provided by the Institute of Animal Pathology (ITPA) in Bern, Switzerland focusing on cattle and pigs necropsies as they are of high economic relevance to veterinary public health.

Material and methods

Data used

A total of 6031 pigs and 2911 cattle were necropsied at the ITPA between 2000 and 2011. Post mortem findings were entered manually as free text by veterinary pathologists. Final necropsy reports contained macroscopic findings, morphological diagnoses as well as results from additional diagnostic tests.

Data classification

We built a set of 13 syndromic categories primarily based on topography of organ systems: "gastrointestinal", "respiratory",

"heart", "lymphatic", "nervous", "musculoskeletal", "urinary", "reproduction" and "serous membranes". For the monitoring of congenital malformations and neoplasia, two specific categories were created: "malformations" and "neoplasia". For diagnoses referring to pathologic patterns that could not be assigned to one of the categories above (for example intoxications, macroscopic findings of the skin or septicemia), the category "other" was added. We constructed a veterinary terminology resource in order to classify most of the necropsy reports into at least one of the above syndromic categories. The construction of the terminology and its validation are thoroughly described in (8). The proportion of reports misclassified by the tool has been estimated to be low. However, the reports that did not fit into one of the categories previously defined were grouped into a last category named "non-classified". This category included thus animals with unclear cause of death and animals that were misclassified.

Data description

After a brief assessment of data quality, animals necropsied at ITPA were described in term of species, sex, age, and geographical origin. We defined three age classes for cows (i.e. calves (<6 months), young (6 months – 2 years), and adults (>2 years)) and three age classes for pigs (i.e. piglets (0-5weeks), weaner pigs (6-12 weeks), and fattening (or adult) pigs (>12 weeks)). The main syndromic categories were identified for each species and age classes. Associations between categories were also considered.

Data modelling

First the number of animals necropsied was modelled using generalized linear models (GLM) in R (9). Then, the two of the most prevalent syndromic categories were further investigated. Alternative GLMs were evaluated on the training data from 2000 to 2006 and validated using data from 2007 to 2011. The models were then evaluated and compared using AIC, the auto-correlation and partial auto-correlation functions of the residuals, and the root-mean-squared error.

Results

Data description

Necropsied animals originated mostly from west-central Switzerland. Most of the cattle necropsied were female more than two years old (49%) and calves (40%). In swine; piglets, weaner and fattening pigs represented respectively 43%, 31% and 25% of the animals necropsied.

Ten percent of the submissions were not classified into a specific syndrome category. The four most frequent pathological syndrome classifications were the same in both species: gastrointestinal system (pigs 55%, cows 41%), serous membranes (pigs 16%, cows 28%), respiratory system (pigs 18%, cows 22%) and "other" (pigs 15%, cows 14%). In cows, the reproduction system also represented a large proportion of animals (19%) but this syndrome was less common in pigs (4,4%). A relatively large proportion of animals were classified into more than one syndromic category (36% of pigs and 45% of cows). The most frequently associated

categories were the same for pigs and cows and were the categories identified as the most frequent for each species (see above). Differences were observed between age classes. As an example, reproduction system syndromes were mainly found in adult cows and respiratory syndromes mainly in young animals.

Data modelling

For cattle, the total number of submissions was fairly stable over the study period. However, models tested were unable to properly fit the data when all age classes were merged into one time series. Three different types of GLMs were thus fitted for each age class. Important differences were observed among age classes (Figure 1): adults presented two annual peaks, one in spring and one autumn; whereas younger animals only had one peak, in winter or spring.

Figure 1. Mean of the fitted values for each age class in cows (data after 2007 not shown).
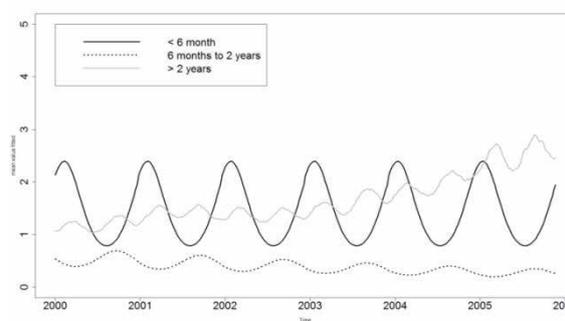
The total number of pig submissions decreased over time. Similar to cows, data for the different age classes were analyzed separately. The same seasonal pattern was found across all age classes, with fewer pigs being necropsied over the summer (Figure 2).

We then focused on reports classified as gastrointestinal and/ or respiratory syndromes, the two most common syndromes. The best models were obtained when age classes where considered separately. The models fitted were close to those obtained when the total number of necropsied animal was considered. As an example, respiratory and gastrointestinal syndrome occurred at the same time in calves (during winter). In adult cows two peaks where observed in respiratory and gastrointestinal syndromes: one in spring and another in autumn.

Figure 2. Mean of the fitted values for each age class in pigs (data after 2007 not shown).
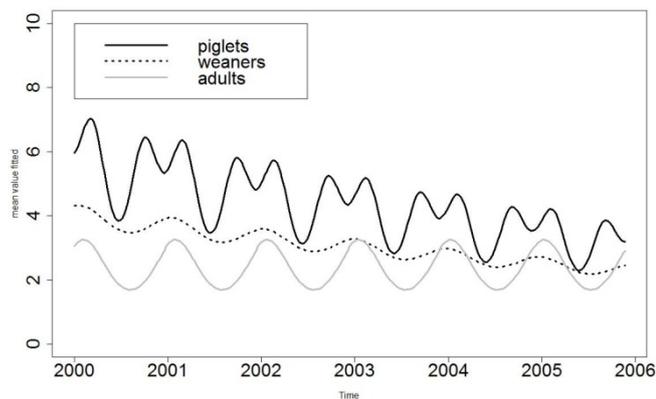
Discussion

Our results showed that important differences exist between species and age classes regarding the pathological patterns and the seasonality of submissions. The main pathological findings were gastrointestinal related in calves, piglets and weaners, gastrointestinal and respiratory in young cattle, and reproduction in adult cows and fattening pigs. These results are consistent with previous studies for both species in Switzerland (10-13). The seasonality of submissions in cattle was consistent with previously reported on-farm cattle mortality in Switzerland (14). More work is needed to investigate the differences between age classes and the seasonal pattern observed in pigs. These differences suggest that age classes in necropsy report data in Switzerland should be monitored as separate time series for SyS.

The data used in our study was of acceptable quality even though some information such as the age of animals was missing in more than 20% of reports (data not shown). The data collected were consistent with the sex and age ratio of the Swiss cattle and pig population. However, most of animals necropsied came from western-central Switzerland and do not represent the entire Swiss populations. Further work is thus needed to improve the quality and the representativeness of the data. The two main recommendations from this study are: 1) to include data from other veterinary laboratories in order to obtain more complete coverage of the Swiss populations, and 2) to improve data reporting especially regarding the age of necropsied animals, in order to allow for monitoring of age classes as separate time series. The addition of the animal's or herd history and comparison between the reason for submission and pathological findings should be also considered for future work.

The agreement between classification of necropsy reports done by the text mining tool and the classification done by experts of veterinary pathology has been estimated to be good or acceptable (8). However, it would be interesting to test different categories in order to perform more detailed analysis and also to better identify specific shifts in the background noise when working with aberration detection algorithms in SyS. This especially applies to the category "other", as it covers a broad range of lesions. An increase of conjunctivitis or mastitis for example may be masked by the general noise in this category. Due to the flexibility of the classification tool (8), testing different syndromic categories can be done easily. It is also possible to expand the tool to include clinical data and other languages which offers promising opportunities.

References

1. **Dórea FC *et al*.** *Prev Vet Med* 101(1-2), 1-17, 2011
2. **Cooper JE.** *OIE Rev Sci Tech* 21(1), 77-89, 2002
3. **Warns-Petit E *et al*.** *BMC Vet Res* 6(1), 56, 2010
4. **National Veterinary Institute SVA.** *Wildlife disease monitoring in Sweden*. Uppsala, Sweden, 2014
5. **Canadian Wildlife Health Cooperative** (CWHC). Available from: http://www.cwhc-rcsf.ca/
6. **O'Toole D.** *Vet Pathol* 47(1), 40-4, 2010
7. **Vial F, Berezowski J.** *Prev Vet Med* 120(1), 27-38, 2015
8. **Furrer L *et al*.** In: *Proceedings of the 11th International Conference on Terminology and Artificial Intelligence*, 2015
9. **R Development Core Team**. Vienna, Austria: R Foundation for Statistical Computing, 2008
10. 10.**Busato A *et al*.** *Prev Vet Med* 30(1), 9-22, 1997
11. **Luginbühl A *et al*.** 154(7), 277-85, 2012
12. **de Vries V.** University of Zurich, 2010
13. **Pill K.** University of Veterinary Medicine Hanover, 2014
14. **Struchen R *et al*.** *Prev Vet Med* 121(1), 1-7, 2015

Acknowledgment