Monitoring trends and developments in cattle health using routinely collected census data

G VAN SCHAIK[1,2]*, H BROUWER-MIDDELESCH[1], L VAN WUYCKHUISE[1],
AJG DE BONT-SMOLENAARS, IMGA SANTMAN-BERENDS[1]

[1]GD Animal Health Service, Deventer, The Netherlands; [2]Department of Farm Animal Health,
Faculty of Veterinary Medicine, Utrecht University

*g.v.schaik@gdanimalhealth.com

Abstract

In the Netherlands, a national Cattle Health Surveillance System (CHSS) is in place that consists of several surveillance components including a quarterly data analysis component to monitor trends and developments in cattle health using routine census data called the Trend Analysis Surveillance Component (TASC). The challenges that were faced during the development of the TASC and the merits of this surveillance component are discussed to provide guidance to those who want to develop such a system. Challenges that had to be solved were subdivided into process-oriented challenges, e.g. data availability, confidentiality, etc. and analytical issues including data validation, aggregation and modeling. The results of TASC had to provide information on cattle health that was intuitive to the stakeholders and could support decision making. Key monitoring indicators that relate to cattle health were developed including parameters on mortality, fertility, udder health, milk production and antimicrobial usage. Population-Averaged Generalised Estimating Equations, with the appropriate distribution were used for analysis. Both trends in time and associations between cattle health indicators and potential confounders are monitored, discussed and reported to the stakeholders on a quarterly level. The TASC provides insight in cattle health parameters, it visualises trends in time, can be used to support or nuance signals that are detected in one of the other surveillance components, is sustainable and provide warnings or initiate changes in policy when unfavorable trends occur.

Keywords: *Cattle health, monitoring and surveillance systems, trend analysis, census data*

Introduction

Monitoring and surveillance systems in livestock are important to ensure high standards of animal health, food safety and public health (1). There are no specific requirements for national monitoring and surveillance programs of cattle health in general. Therefore, a large variety in monitoring and surveillance activities exist (2). Traditionally, samples were collected and tested for diseases. Currently, quality of routinely collected data is improving, providing new possibilities for monitoring and surveillance (3).

In the Netherlands, a national surveillance system, the Cattle Health Surveillance System (CHSS) is in place to monitor trends and developments in animal health and for early detection of exotic diseases or new disorders. The CHSS consists of several surveillance components including enhanced passive reporting, diagnostic and post-mortem examinations, random surveys for prevalence estimations and trend analysis on animal health data.

The routine trend analysis surveillance component (TASC) in which routinely collected Dutch cattle census data is analysed on a quarterly basis as part of the CHSS is described in this paper. The aim of this paper is to provide guidance to others that want to develop a monitoring and surveillance system that incorporates a component in which census data are analysed on a routine basis.

Material and methods

The quarterly TASC aims to monitor trends and developments in cattle health. The results provide insight in cattle health parameters, the so-called key monitoring indicators (KMIs). When developing TASC, there were several challenges that had to be solved.

Data availability

Every quarter of the year, six nationally operating data collecting organisations covering all Dutch cattle herds, provide census data on mortality, slaughter, animal trade, herd health status, antimicrobial usage and bulk milk somatic cell count (BMSCC). In addition, two milk control organisations that together cover about 80% of the dairy herds, provide routinely collected cattle data about milk production, udder health, fertility and metabolic parameters on a quarterly level.

Confidentiality

At the start of the CHSS, all farmers were asked for permission to use their routinely collected herd data for the surveillance system. The farmers were assured that all data is anonymised and that the results of TASC are not traceable to individual farms or animals. Data from herds of which the farmers did not give their consent are excluded from the analysis (1.4% of the farmers).

Agreements with the data collecting organisations

Contracts that covered the moment of data delivery, the format of the data, the information included and the criteria that had to be met in order to be allowed to use the data, were

signed by all collaborating parties. To assure confidentiality, all data collecting organisations provide their raw data directly to an external firm (IntoFocus Data Transformation Services (IDTS)) using a secured FTP server. IDTS encrypts all variables in the data that might link the data back to the original source and the same encryption code is used for all datasets, to ensure that the data of different sources can be combined for analysis (Figure 1).

Figure 1. Overview of the process of the quarterly trend analysis surveillance component on cattle census data.
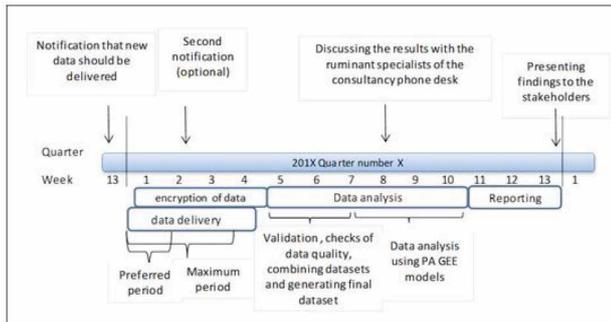
Data quality and uniformity
For TASC, standard software scripts in SAS 9.3® (4) were developed to automate the validation process and merging of the datasets. Standard quality checks are executed on the number of observations, duplicates in the data and extreme values.

Validation and aggregation
Every quarter, the data collecting organisations provide in total eleven datasets of which the number of observations vary between a minimum of 143,000 and a maximum of almost 92 million records. SAS® is used because of its ability to deal with big data, the data validation possibilities and because the program is user friendly. During the whole data preparation process, routine checks and preliminary descriptive statistics are conducted to assure that errors that evolve during the validation process are detected as soon as possible. All software scripts are standardised and optimised and the same scripts are used every quarter of the year.

Within TASC, cattle health parameters were either selected or created. Five different cattle herd types were distinguished based on the provided data:
1. Dairy: milk sold for processing
2. Suckler: >80% female cattle and births
3. Young stock rearing: >95% female cattle <2 years of age
4. Beef: >80% male cattle, mainly off-farm movements to slaughter.
5. Small scale: <20 cattle

For each of the herd types, the trends and developments in cattle health are described using a large number of key indicators (KIs) e.g. mortality, somatic cell count parameters,

antibiotic use (AMU) in cows (>2 years), etc. These key indicators are grouped within six KMIs: durability, herd health, udder health, fertility, metabolic disorders and AMU. For each of the key indicators, a value per herd per quarter of the year for a period of five years is calculated. Eventually the datasets are combined into the final dataset that is transferred to Stata® 14 (5) for analysis.

Analysis
Multivariable Population-Averaged Generalised Estimating Equations (PA-GEE) models, with the appropriate distribution and link function are used. In each of the models, a number of potential confounders are included such as herd size, growth in herd size, replacement rate, location, milk production level (dairy only), season, milk price (dairy only), beef price, purchase, status for endemic diseases, presence of an automatic milking system (AMS), AMU and a variable representing the trend in time. Effects of the independent variables are presented by either odds ratios (OR), incidence rate ratios (IRR) or linear estimates. Dependent variables are presented as observed values, predicted values and the trend in time for the KIs that are displayed. Every quarter, the results of TASC are discussed with the cattle health specialists that are involved in the CHSS. The short term trend of the last six months, the long-term trend during the whole analysed period of five years and whether the direction of the trend in time is favorable or unfavorable is discussed and presented to the stakeholders.
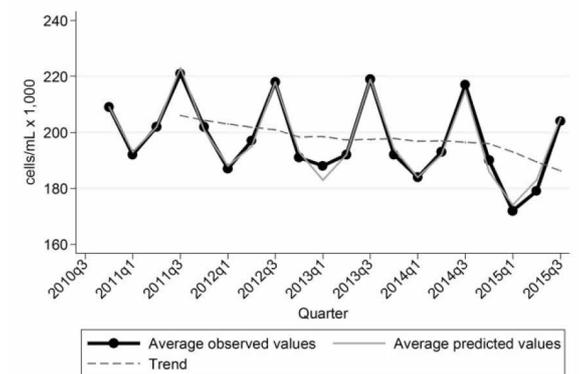
Results
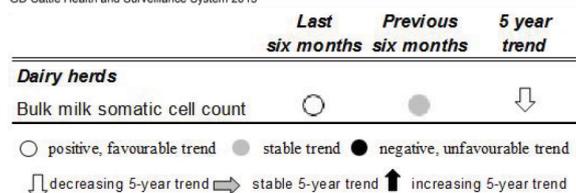*An example: Bulk Milk Somatic Cell Count (BMSCC)*
In the Netherlands, the quarterly BMSCC is normally distributed and a PA GEE analysis with a Gaussian distribution, identity link function and independent correlation structure was used for analysis. In the presented model, data from the period between 1 October 2010 until 30 September 2015 were analysed. In total, 308,021 quarterly herd level observations based on the average BMSCC in the specific quarter from 17,818 dairy herds were included.

Because census data is modeled, only associations that are both statistically significant and biologically relevant are presented. In consultation with udder health specialists, the threshold value for biological relevance was set at and in- or decrease of 10,000 cells/ml compared to the reference category. The results showed that the average BMSCC in the third quarter of 2015 was lower compared to the BMSCC in the same quarter of previous years (Figure 2). Therefore, the short-term trend was determined to be favorable. The long-term trend during the five year period showed a decrease in BMSCC in time, which was also favorable (Figure 2). The 10% dairy herds with the lowest milk production had a higher (+51 x 10³ cells/ml) BMSCC compared to the average Dutch dairy herd. Herds with a conventional milking parlor and dairy herds with the highest average milk production had a lower BMSCC (-24 and -38 x 10³ cells/ml, respectively).
Figure 2. Results of bulk milk SCC for 17,818 Dutch dairy herds between Oct 2010 and Sep 2015.

GD Cattle Health and Surveillance System 2015

	Last six months	Previous six months	5 year trend
Dairy herds			
Bulk milk somatic cell count	○	●	⇩

○ positive, favourable trend ● stable trend ● negative, unfavourable trend

⇩ decreasing 5-year trend ⇨ stable 5-year trend ⬆ increasing 5-year trend

Communication of the result

For presentation purposes, the results of the key indicators are summarised and presented per KMI group. Changing trends, are visualised in more detail using figures such as Figure 2. Possible causes associated with the changing trends are discussed in the text and relevant associations with the independent variables are described. Every quarter, stakeholders are informed about the findings of TASC through a meeting with the surveillance steering committee in which each of the stakeholder organisations are represented.

Discussion

The CHSS aims at early detection of (re) emerging exotic diseases, detection of new phenomena and monitoring trends and developments in Dutch cattle health. The specific aim of TASC is to monitor trends and developments in cattle health to support or nuance signals that are derived from the other components. A drawback from TASC is that it cannot be used for early warning because the census datasets are very large and the process of unlocking, encrypting, validating and analysing the data takes time.

The CHSS system is not static; in all surveillance components, including TASC, on-going efforts are taken to adapt the system to altering circumstances in the cattle industry and to retain high quality standards. For TASC, this meant that new techniques for analysis were implemented and that additional KMIs were added.

The TASC, does not aim to provide information for individual herds given that the system aims at monitoring trends and developments in the complete Dutch cattle industry. The availability of routinely collected census data facilitated the possibility to explore whether census data on cattle health could be used to monitor cattle health on herd level (6) or to serve as proxy for key indicators which are usually collected in large scale field studies (7,8).

Conclusion

As presented, the development and implementation of TASC within the CHSS was challenging and difficulties had to be solved before evolving to the quality level the system has today. Although the system was developed for the cattle population in the Netherlands, similar challenges can arise when such a TASC is developed for other species and/or production systems in other countries. The CHSS as conducted by GD Animal Health in the Netherlands, performs according to expectations of stakeholders. The use of automatically collected and routinely available census data to monitor trends in time for a wide range of cattle health indicators offers insight in general cattle health in the Netherlands. In addition, TASC provides information to support or nuance signals of changes in cattle health. The information in this paper that described both the challenges, but also the solutions that were found and the merit of the complete system, might be valuable for those who want to develop a data analysis system for monitoring and surveillance purposes.

References

1. **OIE.** 2014. ISBN 978-92-9044-842-6,
2. **Geraghty** *et al. Prev. Vet. Med.* 116, 1-11, 2014
3. **Dupuy** *et al. BMC Vet. Res.* doi: 10.1186/1746-6148-9-88, 2013
4. **SAS 9.3 Institute Inc.** Cary, NC, USA, 2010
5. **Stata.** StataCorp LP, TX, USA, 2014
6. **Brouwer** *et al. Prev. Vet. Med.* 122, 76-82, 2015
7. **Santman-Berends** *et al. J. Dairy Sci.* 98, 6965-6977, 2015
8. **De Vries** *et al. J. Dairy Sci.* 97, 715-73, 2014