Health monitoring in Austria – statistical models based on somatic cell count at cow level for early detection of udder health problems developed

P. Winter[1], J. Hofrichter[2], W. Obritzhauser[4], K. Zottl[5] & C. Egger-Danner[3]

[1]*Department for Farm Animals and Veterinary Public Health, University of Veterinary Medicine Vienna, Austria*
[2]*Data, Statistics and Risk Assessment, Austrian Agency for Health and Food Safety, Austria*
[3]*ZuchtData EDV-Services, Vienna, Austria*
[4]*Chamber of Veterinaries, Vienna, Austria*
[5]*Performance Recording for Lower Austria, Zwettl, Austria*

Abstract

An Austrian wide health monitoring system for cattle has been established since 2006. A monitoring tool based on somatic cell count (SCC) is a further step for early and accurate awareness of subclinical mastitis cases. A logistic regression model based on the Dairy Herd Improvement (DHI) data for the Simmental Fleckvieh breed was developed. For validation cows marked suspicious of suffering from intramammary infections (IMI) were sampled for bacteriological examination twice in a weekly interval. The positive predictive value of the logistic regression model was 96.5 % (CI 90.0-99.3). The positive predictive value of cows with two consecutive warnings based on the logistic regression model was 99.0 %. Farmers will be able to select cows for further investigations and following treatment protocols more accurately than using the 200.000 cells/ml benchmark. The risk assessment will be implemented in the DHI reports which are sent in a monthly interval to the farmers. Potentially subclinical infected cows will be summarized in the category "cows with high somatic cell counts and mastitis – further investigations are recommended".

Keywords: Health monitoring, cattle, logistic regression model, somatic cell count.

1.0 Introduction

A project to establish an Austrian wide health monitoring system for cattle has been established since 2006 (Egger-Danner *et al.*, 2009). Within the project diagnostic data, which have to be documented by law (law of animal drug control) are standardised and recorded into a central database for performance recording and breeding. All farms under performance recording are free to join the project. Presently about 13.100 farms with about 220.000 cows are participating.

One project aim is the provision of breeding values for health traits. To increase the health status of the animals by management measures health reports are provided. One of the most frequent diagnoses are acute and chronical mastitis; 28% of all diagnoses recorded within the Austrian health monitoring project are acute or chronical mastitis. A monitoring tool based on somatic cell count (SCC) is a further step for early and accurate awareness of subclinical mastitis cases.

Somatic cell count has been extensively used as a tool for monitoring mastitis in dairy herds. Monthly SCC values that are routinely measured at cow level are useful figures of prevalence and incidence and are often used to monitor dynamics of IMI within herds. Farm managers use SCC to identify cows requiring interventions such as bacteriological examination, treatment, or removal from the herd (Ruegg 2003). Interpretation of SCC is sometimes difficult, because it is a variable parameter that is influenced by many factors, such as diurnal variation, stage of lactation, parity and intramammary infections (Lam *et al*. 2009). Because of the variability of SCC, for monitoring infections dynamics of longitudinal data are necessary. An operational threshold of practical value with minimal diagnostic error, 200.000 cells/ml, was proposed by Schukken *et al*. (2003). The choice of SCC thresholds depends on the purpose of the test. Lowering the threshold increases the sensitivity and consequently provides minimal false negative results whereas raising the threshold increases the specificity, providing minimal false positive results (Pantoja *et al*. 2009).

The objective of the study was to develop a statistical model assessing the individual dynamics of SCC for identifying uninfected and potentially infected cows instead of using the fixed threshold of 200. 000 cells/ml. Therefore, two statistical models based on monthly DHI data were developed. These models are a logistic regression model and a model based on Classification and Regression Trees (CART). Based on the validation results the model, which performs clearly better, will be implemented in the DHI reports.

2.0 Materials and methods

2.1 Data

For developing and validating the statistical models data sets from the routine bacteriological laboratory at the Clinic for Ruminants from 2000 to 2007 were used. Each data set included the result of a bacteriological examination and the monthly DHI data from the previous 6 months regarding SCC, milk yield, milk contents, breed, age and days in lactation. Bacterial examination was conducted according to NMC (1999) recommendations. A cow was considered infected if a mastitis pathogen was isolated at least in one quarter.

2.2 Statistical Analysis

Two different statistical models were used to analyse these data. These models are a logistic regression and a CART model. The logistic regression model yields a prediction for the probability of the presence of IMI based on DHI data. The aim of the CART model was to split the data in several subgroups. Then for each subgroup a specific threshold value for the SCC can be specified.

The response variable of the logistic regression was the bacteriological examination. The explanatory variables are monthly DHI data from the previous 6 months as described above. The considered model was

logit( )= X , which is $\pi = \dfrac{\exp(X\beta)}{1+\exp(X\beta)}$,

where is the probability of the presence of IMI, X is the design matrix of the explanatory variables and is the vector of unknown parameters (Hofrichter *et al*. 2010, Winter *et al*. 2009). The model selection was done by a backward selection using the Analysis of Deviance table. Given an appropriate model, a cutoff value for must be determined to specify the level at which probability a cow should be considered as infected. This was done using the Receiver Operating Characteristic (ROC) curve. The point with minimal quadratic distance to the point of a sensitivity and specificity of 100% is recommended (Dohoo *et al*. 2003).

For the CART model the DHI data of the previous two months were used. The SCC values were used as response variable to split the data into several subgroups based on to the DHI data. An appropriate set of subgroups was found by pruning the tree with respect to a minimal cross validated error. Then for each subgroup a threshold value for the SCC was specified. This was again done by using the ROC curve within each subgroup in the same way as described above.

2.3 Implementation - DHI monthly reports

To support management decisions of cattle breeders and herd health management of their veterinarians, health reports are provided for the first time within a health monitoring project in Austria. The health report is available approx. 10 times a year. It combines performance recording data and data from health monitoring to enable early detection of health problems and their therapy. The report is providing relevant information for individual cows in the fields Fertility, Udder health, Metabolism and Digestive tract, Feet and Legs and Miscellaneous (e.g. Disposals, Calving difficulties). At these reports the SCC information based on the logistic regression model is implemented (Figure 1).

3. Results

Initially 7,437 Simmental Fleckvieh (SI) cows were enrolled, with a total of 33.0 % positive diagnosed cows. Combining these data with the DHI data, complete data were available for 5,115 cows. A logistic regression model was fitted. The relevant explanatory variables are SCC of the two previous months, the age in lactation, the days in lactation and the amount of urea in the milk. Sensitivity and specificity of this model are listed in table 1. For the second kind of model based on the CART analysis, the data sets were splitted into two subgroups with respect to the age in lactation. The performance of this model is listed in Table 1, too.

Table 1 shows that the logistic regression performs better in terms of sensitivity and specificity than the model based on CART. Additionally, both models yield to a much better sensitivity, than using a fixed threshold of 200.000 cells/ml.

4. Validation of the models

After the two final models (logistic regression and CART) for each breed had been found, these models were validated with additional data from the year 2007. Complete data were available for 206 SI cows. The sensitivity and specificity for these models applied on these validation data sets are listed in table 1.

A comparison of the performance of the CART model applied on the two different data sets shows, that sensitivity and specificity differ more than in the logistic regression approach. This indicates that the CART model has less performance than the logistic regression, and is additional confirmed by lower values of sensitivity and specificity.

Table 1. Sensitivity (Se) and specificity (Sp) for a fixed threshold at SCC 200.000, the regression model and the CART model, applied on the model data set (model) and the validation data set (val), for Simmental.

	SCC >200.000		Logistic regression				CART			
	Se	Sp	Se		Sp		Se		Sp	
Breed	model		model	val	model	val	model	val	model	val
SI	46.6%	88.7%	72.4%	84.4%	72.6%	71.3%	68.6%	77.5%	70.5%	68.3%

Application of the model

The logistic regression model was implemented in the DHI database. For the application a dataset from Simmental Fleckvieh cows in Lower Austria from 2009 was used. Remarkable is the group between 100 and 150,000 cells, in which 43% of the cows were considered as potentially infected after two consecutive warnings using the regression model (Table 2).

Table 2. Distribution of suspicious Fleckvieh cows based on the logistic regression formula grouped by SCC categories of SCC of cows of Lower Austria in 2009.

SCC on day of milk recording	Nr MR	% suspicious with one warning	% suspicious with two consecutive warnings
<50	166,012	2.2	1.0
50 < 100	107,859	27.1	14.7
100 -< 150	56,127	73.4	43.0
150 -< 200	33,473	86.8	54.9
200 -< 250	21,537	89.1	59.6
250 -< 300	14,585	89.4	60.9
300 -< 350	10,293	89.0	61.6
350 -< 400	7,656	88.7	60.5
>= 400	45,242	85.8	57.5
Total	462,784	41.1	25.7

Figure 1. The health reports are expanded by the warning cows suspicious for SCC based on the logistic regression formula.

Cows with high somatic cell counts and mastitis – further investigations recommended

Name	Animal-ID	L.	Days	BE*	21.08.2009 SCC		21.07.2009 SCC	12.06.2009 SCC
LISA	AT 999.444.972	2	283	BE*	568		205	132
SUMSI	AT 999.136.847	4	121		40	D	268	174
BIENE	AT 999.326.745	5	215	BE*	182		108	48
STRAUSSA	AT 999.327.845	4	28		31		T	T
LOLITA	AT 999.857.145	5	11		16		T	T

BE: Bacterial examination recommended*

As shown in figure 1 for cow „BIENE" a bacteriological examination is recommended even the SCC is below the threshold of 200.000 cells.

4.1 Dissemination of the reports

The reports are provided electronically to farmers and veterinarians. Most of the information is also available at an internetplattform. A future perspective is the implementation of the logistic regression formula within the internet platform.

5. Discussion

According to Pantoja *et al.* (2008) cows with SCC > 200. 000 cells/ml are 2. 7 times more likely to experience a first case of mastitis. Cows with intramammary infections at dry-off can be identified adequately combining information from SCC and clinical mastitis records, but decisions regarding the optimal selection criteria depend on herd characteristics such as prevalence of IMI and type of microorganisms present in the herd (Torres *et al*. 2008).

Using the risk assessment model these cows are detected on a very early stage of infection and no further herd information is necessary. It has to be considered that the sensitivity is 72,4%, which means that 27,6% of the cows are false positive highlighted. Therefore the application of the model should never be the base of a treatment decision or any other decisions like culling. Highlighted cows require further investigations like performing a California Mastitis Test or a bacteriological examination of quarter milk samples. Normally about 30% of the Austrian dairy cows are subclinically infected, at least in one quarter (Hofrichter *et al*. 2010).

6. Conclusion

The design of the DHI reports summarizing the cows in special categories gives the farmers and the veterinarians a precise overview of the heath status of the herd. The regularly interpretation of the DHI data using the logistic regression model is a potential tool for monitoring udder health in Austria. Potentially subclinically infected cows will be determined as soon as somatic cell counts are exceeding their individual base line.

This risk assessment allows an early detection of IMI and supports the reduction of infection reservoirs and of economic losses. For the correct interpretation and use, continuous information and training for farmers is needed. Support by the veterinarians is recommended.

7. Acknowledgements

The study was funded by the Austrian Federal Ministry of Health (BMGF 74020/0034-IV/6/2006; BMGF-70420/0109-IV/6/2006) and the Ministry of Agriculture, Forestry, Environment and Water Management (BMLFUW) of Austria. The authors thank the Austrian Breeding Association (ZuchtData, Mr. Mayerhofer) for the cooperation and assistance.

References

Egger-Danner, C., B. Fuerst-Waltl, W. Holzhacker, R. Janacek, J. Lederer, C. Litzllachner, C.Mader, M. Mayerhofer, J. Miesenberger, W. Obritzhauser, G. Schoder, and A. Wagner, 2007. Establishing a health monitoring for cattle in Austria: first experiences. In: EAAP-Book of Abstracts No 13: 58th Annual Meeting of the EAAP, 26-29.8.2007 Dublin, Ireland, 363.

Dohoo I, Martin W, Stryhn H (2003) : Veterinary Epidemiologic Research. Atlantic Veterinary College, Charlottetown, PE, Canada.

Hofrichter, J., Winter, P., Fuchs, K. (2010): Comparison of two statsitical models based on Somatic cell count for diagnosis of intramammary infections at cow level. 5 th IDF Mastitis Conference, 21.-24. 03. 2010, New Zealand, p. 558-561.

Lam TJGM, Olde Riekerink RGM, Sampimon OC, Smith H (2009): Mastitis diagnostics and performance monitoring: a practical approach. Irish Vet. J. 62, Suppl, 34-39.

National Mastitis Council (1999): Laboratory Handbook on Bovine Mastitis. Madison WI, USA: NMC.

Pantoja JCF, Hulland C, Ruegg P. (2009): Dynamics of somatic cell counts and intramammary infections across dry period. Prev. Vet. Med. 90, 43-54.

Pantoja JCF, Hulland C, Ruegg P. (2009): Somatic cell count status across the dry period as a risk factor for the development of clinical mastitis in the subsequent lactation. J. Dairy Sci. 92, 139-148.

Schukken YH, Wilson DJ, Welcome F, Garrison-Tikofsky L, Gonzalez RN. (2003): Monitoring udder health and milk quality using somatic cell counts. Vet. Res. 34, 579-596.

Torres AH, Rajala-Schultz PJ, De Graves FJ, Hoblet KH. (2008) : Using dairy herd improvement records and clinical mastitis history to identify subclinical mastitis infections at dry-off. J. Dairy Res. 75, 240-247.

Ruegg PL. (2003): Investigations of mastitis problems on farms. Vet. Clin. North Am. Food Anim. Pract. 19 (1), 47-74.

Winter, P., Hofrichter, J. (2009): Risk assessment of monthly determinded somatic cell counts and cow individual parameters for detecting subclinically intramammary infections at cow level. European Buiatrics Forum, 1.-4. 12. 2009, Marseille, Frankreich, p 85.