



## Somatic cell count assessment at the quarter or cow milking level

H. Mollenhorst,<sup>\*1</sup> P. P. J. van der Tol,<sup>†</sup> and H. Hogeveen<sup>\*‡</sup>

<sup>\*</sup>Department of Farm Animal Health, Faculty of Veterinary Medicine, Utrecht University, PO Box 80151, 3508 TD Utrecht, the Netherlands

<sup>†</sup>Research Department, Software Engineering, Lely Industries NV, Weverskade 110, 3147 PA Maassluis, the Netherlands

<sup>‡</sup>Business Economics Group, Wageningen University, PO Box 8130, 6700 EW Wageningen, the Netherlands

### ABSTRACT

The aim was to investigate whether on-line somatic cell count (SCC) assessment, when combined with electrical conductivity (EC), should be implemented at the udder quarter or at the cow level. Data were collected from 3 farms with automatic milking systems, resulting in 3,191 quarter milkings used in the analyses. Visual observations of foremilk and quarter milk samples for laboratory SCC analysis were used to define 2 gold standards. One was based on visual observation only and the other was based on a combination of visual observation and SCC (using a reference value of 500,000 cells/mL), which means that a quarter milking must have visually abnormal milk as well as an increased SCC to be categorized positive. On-line SCC assessment took place at the quarter level during the first part of the milking. Composite cow level samples were used for laboratory SCC analysis and to compare the performance of SCC assessment at quarter and cow levels. The EC at the quarter level was measured by in-line sensors of the automatic milking system. Alerts for SCC indicators were calculated based on straightforward reference values. Alerts for EC were based on straightforward reference values, or on interquarter ratios. The latter was calculated by dividing the value of a given quarter by the average value of the 2 lowest quarters of that milking. The EC and SCC indicators were combined with either a Boolean “and” or “or” function. Receiver operating characteristic curves were used to visually present results using different threshold values. Sensitivity, specificity, and success rate at the quarter level and false alert rate per 1,000 cow milkings were used to compare indicators at given sensitivity or specificity levels. Quarter level SCC assessment was superior to cow level assessment (transformed partial area under the curve = 0.70 vs. 0.62) when combined with EC measurement at quarter level. When aiming for the same sensitivity level (e.g., 50%) with all visual abnormal milk as the gold standard, more false alerts were

generated with cow level assessment (137 per 1,000 cow milkings) compared with quarter level SCC assessment (75 per 1,000 cow milkings). As a comparison, using EC alone resulted in 292 false alerts per 1,000 cow milkings in the same situation. Therefore, it is concluded that quarter level SCC assessment was superior to cow level assessment when combined with EC measurement at quarter level.

**Key words:** abnormal milk, detection, on-line assessment, somatic cell count

### INTRODUCTION

Every farmer needs to monitor udder health. Several indicators are available for the control of mastitis. In conventional milking parlors, farmers assess the milk and udder mostly visually during milking. Additionally, sensors can be installed to assist with diagnosis of mastitis. Much research was done on electrical conductivity (EC; Davis, 1947; Lansbergen et al., 1994) and development of algorithms based on EC combined with milk temperature and yield (Maatje et al., 1992; de Mol et al., 1997). Furthermore, Caverio et al. (2006) and de Mol and Woldt (2001) used additional data and techniques, such as the fuzzy set theory, to improve detection performance.

When using an automatic milking system (AMS), the farmer relies mainly on the alerts for monitoring udder health. Demands for performance are different than when used in conventional milking systems. Most AMS generate alerts based on EC and milk yield and color, sometimes using indicators based on milk flow or temperature. A new development is the use of on-line (real-time) SCC assessment (Kamphuis et al., 2008). In general, SCC data are available for test days only, which are monthly or less. Furthermore, these data are at the cow level, so it is not known which quarter is causing the increased SCC. During the past decade, methods for on-line SCC assessment were developed (Whyte et al., 2005). At the cow level, log-transformed on-line SCC correlated ( $r = 0.76$ ) with log-transformed, laboratory determined SCC that was regarded as the gold standard (GS) for SCC (Kamphuis et al., 2008). On-line SCC at the cow level improved detection of

Received October 21, 2009.

Accepted February 15, 2010.

<sup>1</sup>Corresponding author: H.Mollenhorst@uu.nl

**Table 1.** Descriptive statistics per farm of the data used for analysis

Farm	Cows (n)	Cow milkings (n)	Hours observed (n)	Positive cows/quarter milkings (n/n)
1	61	396	64	7/13
2	82	187	42	3/4
3	48	233	42	2/2
Total	191	816	148	12/19

abnormal milk compared with using EC alone. That research focused on detection of clinical mastitis cases only and on-line SCC was determined at the cow level.

To date, no data were available for on-line SCC at quarter level and visual assessment of abnormal milk for a large quantity of milkings. Therefore, the question remained whether sampling at the quarter level has additional value. Nielsen et al. (2005) stated that mastitis indicators should be sampled at the quarter level because the dilution effect makes the indicator less accurate. When on-line SCC assessment is performed at the cow level only, it may be possible to identify the affected quarter by other sensor data, such as EC.

The aim was to investigate whether on-line SCC assessment should be implemented at the udder quarter or at the cow level when combined with EC. Two different GS were used: all visually abnormal milk or visually abnormal milk that also has an increased quarter SCC (QSCC).

## MATERIALS AND METHODS

### *On-Farm Data Collection*

Data were collected from 3 farms that were selected because they had 2 AMS (Lely Astronaut A3 with MQC-C software, Lely Industries NV, Maassluis, the Netherlands) and 100 to 120 animals (Table 1). One AMS was equipped for sampling purposes and the second functioned as a back-up. Farms 1 and 3 had 2 separate groups of cows. On these farms, cows that were anticipated by the farmer to have abnormal milk were transferred from the group that was not involved in the study to the study group to increase the number of positive milkings.

Each farm was sampled for at least 36 h over 3 to 4 d. Initially, each farm was involved once. The first farm was visited again because the number of positive milkings was considered too low after visiting the 3 farms for 1 period each. All cows visiting the AMS were sampled, resulting in 3,191 quarter milkings being used. Observations were performed by 9 students (studying or finished Veterinary Medicine at Utrecht University, Utrecht, the Netherlands) who were selected for their experience with milking and affinity for dairy farm-

ing. Furthermore, 5 students were instructed on the first farm by practicing the protocol together to attain consistent standards. Two students who were involved later were instructed individually by the coordinator and 2 by an experienced student who also attended the first instructional meeting. In general, 1 pair worked from 0700 until 1400 h and a second pair from 1400 until 2100 h.

### *Data Collection Protocol*

The following steps were carried out for each cow that was milked by the AMS:

1. The first 2 squirts of milk from each quarter were discarded to exclude cases of teat canal infection.
2. The next 2 to 3 squirts of milk were scored visually (normal, abnormal) at first sight on a black quarter assessment paddle (commonly used for California mastitis test scoring).
3. After cleaning the udder and automatic cluster attachment, samples (20 mL/quarter) were taken by the AMS from the first milk from the long milk tube for on-line SCC assessment.
4. During milking, EC, color, and milk yield were measured by in-line sensors.
5. Mechanical milk meters (WB Mini-Test, Tru-Test Ltd, Auckland, New Zealand) were used to collect proportional samples from individual quarters throughout the entire milking. At the end of milking, samples were taken from the 4 Tru-Test bottles and analyzed for QSCC.
6. Composite cow milking samples were collected (Shuttle, Lely Industries NV, Maassluis, the Netherlands) and analyzed for cow SCC (CSCC).

The on-line SCC sensor operated using the following procedure (adapted from Whyte et al., 2004):

1. A start-of-milking signal from the AMS initiated collection of a 20-mL milk sample per quarter.
2. A 4-mL sample was precisely mixed with reagent in the mixing chamber. Accuracy of the 4-mL

sample volume was important because it affected the accuracy of the result.

3. The mixture was transferred to the measurement chamber.
4. The viscosity was determined by the time taken for the gel to drain out of the measurement chamber.
5. If the outflow passage was blocked, vacuum was applied to clear the blockage.
6. At the end of the measurement cycle, the mixing and measurement chambers were flushed to minimize cross-contamination between samples.

All laboratory analyses were performed using Combi-Foss 6000 (Foss Electric, Hillerød, Denmark) at Qlip NV (Zutphen, the Netherlands)

### Definition of Abnormal Milk

In this study, 2 different GS were used, both at the quarter level. The first was based on visual observation only. The second was based on a combination of visual observation and QSCC (using a reference value of 500,000 cells/mL), which meant that a quarter milking must have visually abnormal milk as well as an increased SCC to be marked positive. The first GS supported a strict interpretation of EU legislation (Regulation EC 853/2004; EC, 2004). The second GS was in agreement with Rasmussen et al. (2005), who proposed a clearer classification when evaluating the performance of AMS.

### Detection Algorithms

Detection of abnormal milk by the AMS could be based on different alerts associated with different sensors or combinations of sensors. In this study, alerts for SCC indicators were calculated based on straightforward reference values, so alerts were generated when the value was larger than a given threshold. Alerts for EC were based on straightforward reference values, or on interquarter ratios (**IQR**). The IQR was calculated by dividing the value of a given quarter by the average value of the 2 lowest quarters of that milking. To illustrate the effect of using on-line SCC at the cow or quarter level, performance characteristics of 4 single indicators (EC, EC-IQR, on-line QSCC, and laboratory determined CSCC) and of different combinations of EC and SCC indicators (combined with either a Boolean “and” or “or” function) were calculated. A combination with a Boolean “and” meant that both indicators were above its specified reference value, whereas a combination with a Boolean “or” meant that only 1 of the 2 were above its specified reference value.

### Data Analysis

To assess the performance of on-line QSCC, it was compared with the laboratory determined QSCC. The SCC data were log-transformed before Spearman rank correlations were calculated.

According to the classification based on the GS and detection algorithms, samples were divided into 4 groups: true positive (**TP**), GS-positive quarter milking that is alerted; false negative (**FN**), GS-positive quarter milking that is not alerted; true negative (**TN**), GS-negative quarter milking that is not alerted; and false positive (**FP**), GS-negative quarter milking that is alerted. Because all milkings were checked visually, time windows as proposed by Sherlock et al. (2008) were not used to extend the validity of observations or alerts.

Sensitivity (**SN**) was calculated as the number of TP quarter samples over the total number of GS-positive quarter samples:

$$SN = TP / (TP + FN) \times 100\% \quad [1]$$

Specificity (**SP**) was calculated as the number of TN quarter samples over the total number of GS-negative quarter samples:

$$SP = TN / (TN + FP) \times 100\% \quad [2]$$

For evaluation and comparison of different tests with continuous or ordinal scales, receiver operating characteristic (**ROC**) curves can be very helpful (Gardner and Greiner, 2006). An ROC curve is a graph in which the true positive rate (SN) is plotted against the false positive rate ( $100 - SP$ ) over the range of possible threshold values (Dettileux et al., 1999).

The upper left point of a complete ROC graph (0, 100) represents a perfect classifier with an SN and SP of 100%. The origin of the graph (0, 0) represents a classifier with a very high threshold, classifying all samples as negative. The upper right point of a complete ROC graph (100, 100) represents a classifier with a very low threshold, classifying all samples as positive. The line  $y = x$  represents a random classifier with different levels of positive classified samples (Provost and Fawcett, 2001). Any improvement over this random classifier results in an ROC curve at least partially above this straight line (Cortes and Mohri, 2005). The area under the curve (**AUC**) can be used as a summary statistic to assess performance of an indicator. The AUC is 0.50 for a random classifier and 1 for the perfect classifier.

When only certain parts of the ROC curves are of interest (e.g., when an SP of more than 99% is required), a partial AUC can be calculated for that part of the

ROC curve only. The partial AUC can be transformed (**TPAUC**) by the following algorithm to be interpreted in the same way as the total AUC (McClish, 1989):

$$\text{TPAUC} = 1/2 [1 + (A - \text{min})/(\text{max} - \text{min})], \quad [3]$$

where A is the partial AUC, min is the AUC of the random classifier for that part of the curve, and max is the AUC of the perfect classifier for that part of the curve.

In this study, ROC curves were obtained by plotting SN and (100 - SP) values for every point of the EC index (Lely Industries NV), for every 0.01 increase of IQR starting from IQR = 1.00, and for every 10,000 cells/mL for all SCC indicators. Only the points with the highest SN for a given SP were depicted in the figures.

Based on observations of Hillerton (2000), Rasmussen (2004) proposed 80% SN and 99% SP as target values. Dependent on the number of GS-positive cases (around 20), this guaranteed a minimum SN of approximately 70% for a 95% confidence interval. For focus on relevant SP ranges, the ROC curves were plotted for SP of 97% and higher, and TPAUC were calculated for SP of 99% and higher.

For evaluating the practical relevance of a detection system, Sherlock et al. (2008) proposed relative measures in addition to the SN because SP was not very informative in practice when prevalence was low. The success rate (**SR**; also known as predictive value positive) indicated the proportion of alerts that are true alerts and was calculated as the number of TP quarter samples over the total number of alerted quarter samples:

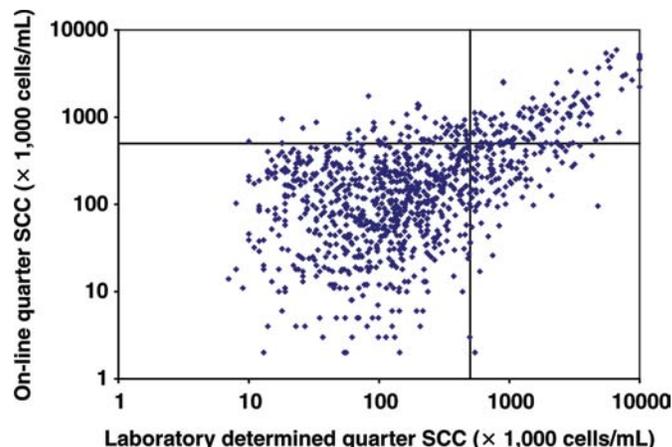
$$\text{SR} = \text{TP}/(\text{TP} + \text{FP}) \times 100\%. \quad [4]$$

The false alert rate (**FAR**) indicated the number of false alerts that a farmer may have to deal with per 1,000 cow milkings:

$$\text{FAR} = 1,000 \times \text{FP}/\text{total number of cow milkings}. \quad [5]$$

For the calculation of FAR, a maximum of 1 FP/cow milking was used. In contrast with SN and SP, the SR and FAR were dependent on the prevalence and varied by farm.

Significant differences between SN, SP, FAR, and SR values were assessed by logistic regression (Pepe, 2003). The likelihood ratio statistics for type 3 analysis were used to correct for adverse farm influences, including the interaction effects of farm practices and test methods. All data preparation steps and statistical analyses were performed using SAS 9.1 (SAS Institute Inc., 2002).



**Figure 1.** On-line quarter SCC assessments against laboratory-determined composite quarter SCC (reference lines at 500,000 cells/mL). Color version available in online PDF.

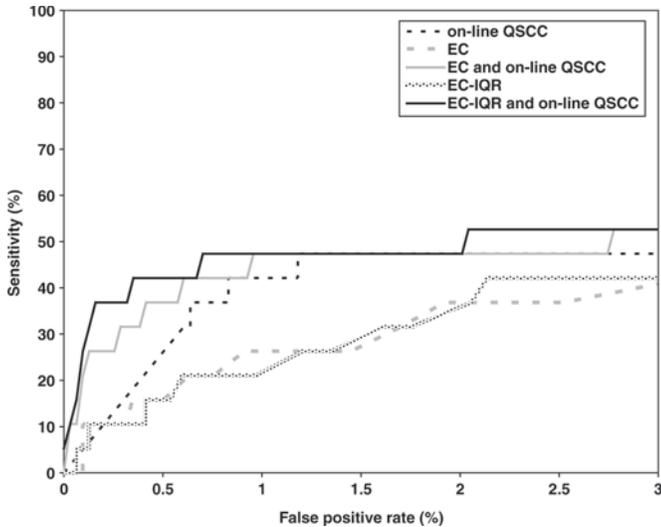
## RESULTS

Figure 1 shows the relation between log-transformed laboratory determined QSCC and log-transformed on-line QSCC [Spearman rank correlation coefficient ( $r_s$ ) = 0.47,  $P < 0.001$ ].

Within the SP range 97 to 100%, straightforward reference values and IQR performed similarly when EC was used alone (Figure 2). Both had a TPAUC for the SP range 99 to 100% of around 0.57. Combining EC with on-line QSCC could improve performance considerably. Because the combination of on-line QSCC with EC-IQR performed descriptively better (TPAUC = 0.70) than with the straightforward EC reference values (TPAUC = 0.67), only EC-IQR was used in further analyses.

The combination of EC with CSCC (TPAUC = 0.62) descriptively outperformed EC alone when SP was set high (>98.5%). On-line QSCC outperformed EC alone. A combination of on-line QSCC and EC gave the best results (Figure 3). At high SP values (>99%), performance was improved by the use of a Boolean “and” function (TPAUC = 0.70) rather than a Boolean “or” function (TPAUC = 0.62).

When the 3 most important indicators (EC-IQR and its combinations with either on-line quarter or cow level SCC) were compared at an SP level of 99.0% (at the quarter level), SN and SR of EC-IQR combined with on-line QSCC were not significantly different from the other 2 (Table 2). The FAR (per 1,000 cow milkings) was lowest for the combination with cow level SCC because in that case most false alerts at the quarter level were from the same milking as the TP alerts. These false alerts were counted only once when FAR was expressed per 1,000 cow milkings. When compared at an SN level



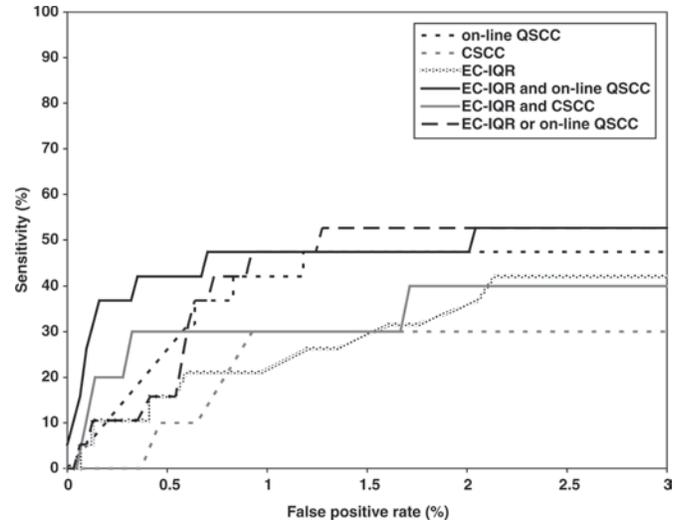
**Figure 2.** Receiver operating characteristic curves for detection of all visual abnormal milk to compare the use of straightforward reference values with the use of interquarter ratios for electrical conductivity (EC-IQR), alone or in combination with on-line quarter SCC assessment (QSCC).

of 50%, the combination of EC-IQR with on-line QSCC performed best on SP, FAR, and SR (Table 2).

Figure 4 shows the 3 most important indicators, but with a different GS. This time, quarter milk had to be visually abnormal and have a laboratory determined QSCC of  $>500,000$  cells/mL. The general pattern stayed the same; however, the SN level increased. When using this GS, a combination of 80% SN and 99% SP could be reached. The TPAUC of EC-IQR combined with on-line QSCC was 0.92. At an SP level of 99%, the combination with on-line QSCC had a significantly higher SN than EC-IQR alone or combined with cow level SCC. At an SN level of 80%, the combination with on-line QSCC outperformed both other indicators significantly on SP, FAR, and SR (Table 3).

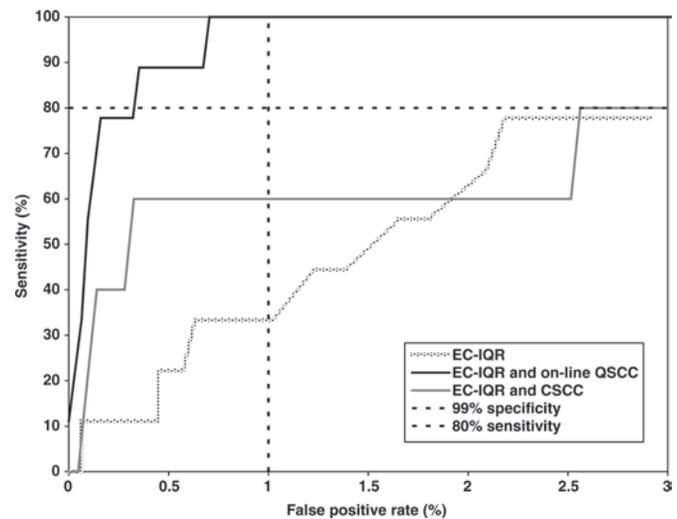
## DISCUSSION

The correlation between on-line QSCC and laboratory determined QSCC ( $r_s = 0.47$ ) was not very high and was lower than that found by Kamphuis et al. (2008) and Leslie et al. (2007). However, the sample taken by the on-line SCC sensor was not from the same part of the milking as the QSCC sample. The on-line SCC sensor analyzed milk from the first part of the milking, which was more in line with general practices such as visual observation and California mastitis test (Schalm and Noorlander, 1957) scoring, whereas the samples for laboratory analysis were taken from the composite quarter milk. The sampled milk fraction influences these results (Sarıkaya and Bruckmaier, 2006).



**Figure 3.** Receiver operating characteristic curves for detection of all visual abnormal milk to compare the use of on-line quarter SCC (QSCC) with the use of composite cow SCC (CSCC), alone or in combination with the interquarter ratios for electrical conductivity (EC-IQR).

Using IQR instead of straightforward reference values improved performance. When more sophisticated algorithms are used, results can probably be improved further. Also, adding information from more sensors (e.g., color or yield) could probably improve results. The main focus here was on whether SCC at the quarter level could improve detection performance more than SCC at the cow level. The combination of SCC and EC with Boolean functions was suitable for this.



**Figure 4.** Receiver operating characteristic curves for detection of milk that was visually abnormal and had an increased SCC (laboratory-determined quarter SCC above 500,000 cells/mL) to compare the use of the interquarter ratios for electrical conductivity (EC-IQR) alone or in combination with either on-line quarter SCC (QSCC) or composite cow SCC (CSCC).

**Table 2.** Performance characteristics for different combinations of SCC and electrical conductivity (EC) indicators compared with visual observations only as gold standard (GS) at specificity (SP) level of 99.0% or sensitivity (SN) level of 50% (or nearest level above)

Indicator <sup>1</sup>	SP = 99.0%				SN = 50%		
	GS-positive quarters (n)	SN (%)	False alert rate <sup>2</sup>	Success rate (%)	SP (%)	False alert rate <sup>2</sup>	Success rate (%)
EC-IQR	19	21.1	36.8 <sup>b</sup>	11.4	90.4 <sup>c</sup>	291.7 <sup>c</sup>	3.2 <sup>b</sup>
EC-IQR and on-line QSCC	19	47.4	35.6 <sup>b</sup>	22.5	98.0 <sup>a</sup>	74.8 <sup>a</sup>	13.5 <sup>a</sup>
EC-IQR and CSCC	10	30.0	12.7 <sup>a</sup>	12.5	94.6 <sup>b</sup>	137.4 <sup>b</sup>	4.1 <sup>b</sup>

<sup>a-c</sup>Values within a column with different superscripts differ significantly ( $P < 0.05$ ) after correction for interaction with the farm.

<sup>1</sup>EC-IQR = interquarter ratio of EC; QSCC = quarter level SCC; CSCC = cow level SCC (laboratory).

<sup>2</sup>Expressed per 1,000 cow milkings.

When 2 indicators were used, there were 2 straightforward ways to combine them using Boolean functions. Using “or” resulted in worse performance than using “and.” Because detection levels of both indicators must be adapted to the Boolean function used, it was better to implement them in an algorithm or to adapt the detection levels to the way a farmer uses the separate alerts lists.

In general, performance indicators are expressed per measurement unit. The FAR was introduced as a performance indicator that could be more informative for a farmer or consultant (Sherlock et al., 2008). Then the question arises as to whether the total number of false alerts at the quarter level is informative because the costs of a false alert are mostly at the cow level. Therefore, it is better to calculate the FAR per 1,000 cow milkings. Data in Tables 2 and 3 show that when comparing indicators at a specified SP (at the quarter level), cow level indicators were at a disadvantage for achieving a high SN because most TP alerts were accompanied by 1 or more false alerts on the same cow milking. As a consequence, the FAR per 1,000 cow milkings was lower than for the quarter level indicators at the same SP level. Therefore, it was difficult to compare the cow and quarter level indicators at a specified SP because both SN and FAR were better for the

cow level indicator, whereas normally there would be a trade-off between both when indicators are at the same level. When indicators were compared at a specified SN, the worse performance of the cow level indicators was more pronounced in the SP than in the FAR.

Two GS were used, namely all visually abnormal milk and a subgroup containing only those visually abnormal milkings that had an increased laboratory determined QSCC as well. The first GS was based on EU legislation (Regulation EC 853/2004; EC, 2004), which corresponds with the United States Grade A Pasteurized Milk Ordinance, stating that all visually abnormal milk must be discarded. Rasmussen et al. (2005) state that “evaluation and testing of the ability of AMS to detect and sort milk must be done with respect to a clear classification into normal and abnormal milk” and combine visual observations with SCC data to define abnormal milk. Therefore, the second GS was used also.

When a farmer wants to detect cows with clinical mastitis that have to be treated with antibiotics, a quite narrow definition of abnormal milk is appropriate, such as the second GS used in this study. When EU legislation (Regulation EC 853/2004; EC, 2004) is interpreted quite strictly, the first GS used (visually abnormal milk) is more appropriate. When separation of abnormal milk is involved, it is better to express

**Table 3.** Performance characteristics for different combinations of SCC and electrical conductivity (EC) indicators compared with visual observations combined with SCC (laboratory determined quarter SCC above 500,00 cells/mL) as gold standard (GS) at specificity (SP) level of 99.0% or sensitivity (SN) level of 80% (or nearest level above)

Indicator <sup>1</sup>	SP = 99.0%				SN = 80%		
	GS-positive quarters (n)	SN (%)	False alert rate <sup>2</sup>	Success rate (%)	SP (%)	False alert rate <sup>2</sup>	Success rate (%)
EC-IQR	9	33.3 <sup>b</sup>	35.8 <sup>b</sup>	9.1	92.1 <sup>c</sup>	245.4 <sup>c</sup>	3.1 <sup>b</sup>
EC-IQR and on-line QSCC	9	100 <sup>a</sup>	35.8 <sup>b</sup>	22.5	99.6 <sup>a</sup>	13.6 <sup>a</sup>	42.1 <sup>a</sup>
EC-IQR and CSCC	5	60.0 <sup>b</sup>	12.7 <sup>a</sup>	12.5	97.4 <sup>b</sup>	56.4 <sup>b</sup>	6.8 <sup>b</sup>

<sup>a-c</sup>Values within a column with different superscripts differ significantly ( $P < 0.05$ ) after correction for interaction with the farm. These results must be viewed with caution because numbers of GS-positive quarters were low and SCC was used to define the GS and is part of some of the indicators.

<sup>1</sup>EC-IQR = interquarter ratio of EC; QSCC = quarter level SCC; CSCC = cow level SCC (laboratory).

<sup>2</sup>Expressed per 1,000 cow milkings.

all performance indicators at the cow level because discarding milk from only one quarter is not allowed (Regulation EC 853/2004; EC, 2004).

Results of this paper should not be seen as detection performances with practical relevance. Because the number of GS-positive cases was too small to split the data set into a test and a training data set, no further validation could be performed. Results from the second GS (visually abnormal and increased laboratory determined QSCC) must be considered with extra care because the number of GS-positive cases was even lower and SCC was used to define the GS and was part of some of the indicators. The results could be used to illustrate the effect of using SCC assessment at the cow or quarter level, as was done here.

Kamphuis et al. (2008) showed that in New Zealand, the cow level SCC combined with quarter level EC improved detection of abnormal milk compared with EC or the cow level SCC alone. The current study showed that a combination with the quarter level SCC can further reduce the number of false alerts.

## CONCLUSIONS

To detect abnormal milk, quarter level SCC assessment was superior to cow level assessment, even when combined with EC measurement at quarter level. When aiming for the same SN level, far more false alerts were generated with the cow level assessment.

## ACKNOWLEDGMENTS

We gratefully acknowledge the participating farmers, the students, and the technicians of Lely Industries NV for their contribution to the acquisition of data used. This research was supported by the Dutch Technology Foundation STW (Utrecht, the Netherlands), applied science division of the Netherlands Organization for Scientific Research (NWO), and the Technology Program of the Ministry of Economic Affairs (The Hague, the Netherlands).

## REFERENCES

- Cavero, D., K. H. Tolle, C. Buxade, and J. Krieter. 2006. Mastitis detection in dairy cows by application of fuzzy logic. *Livest. Sci.* 105:207–213.
- Cortes, C., and M. Mohri. 2005. Confidence intervals for the area under the ROC curve. Pages 305–312 in *Advances in Neural Information Processing Systems 17*. Proc. 2004 Conf. L. K. Saul, Y. Weiss, and L. Bottou, ed. MIT Press, Cambridge, MA.
- Davis, J. G. 1947. The rapid abnormality indicator. A simple electrical apparatus for the rapid detection of abnormal (mastitis) milk. *Dairy Industries* 12:35–40.
- de Mol, R. M., G. H. Kroeze, J. M. F. H. Achten, K. Maatje, and W. Rossing. 1997. Results of a multivariate approach to automated oestrus and mastitis detection. *Livest. Prod. Sci.* 48:219–227.
- de Mol, R. M., and W. E. Woldt. 2001. Application of fuzzy logic in automated cow status monitoring. *J. Dairy Sci.* 84:400–410.
- Detilleux, J., J. Arendt, F. Lomba, and P. Leroy. 1999. Methods for estimating areas under receiver-operating characteristic curves: Illustration with somatic-cell scores in subclinical intramammary infections. *Prev. Vet. Med.* 41:75–88.
- EC (European Commission). 2004. Regulation (EC) No. 853/2004 of the European Parliament and of the Council of 29 April 2004 laying down specific hygiene rules for the hygiene of foodstuffs. *Off. J. L* 139:55–205.
- Gardner, I. A., and M. Greiner. 2006. Receiver-operating characteristic curves and likelihood ratios: Improvements over traditional methods for the evaluation and application of veterinary clinical pathology tests. *Vet. Clin. Pathol.* 35:8–17.
- Hillerton, J. E. 2000. Detecting mastitis cow-side. Pages 48–53 in Proc. National Mastitis Council 39th Annual Meeting. Atlanta, GA. NMC, Madison, WI.
- Kamphuis, C., R. Sherlock, J. Jago, G. Mein, and H. Hogeveen. 2008. Automatic detection of clinical mastitis is improved by in-line monitoring of somatic cell count. *J. Dairy Sci.* 91:4560–4570.
- Lansbergen, L. M. T. E., M. Nielsen, T. J. G. M. Lam, A. Pengov, Y. H. Schukken, and K. Maatje. 1994. Evaluation of a prototype online electrical-conductivity system for detection of subclinical mastitis. *J. Dairy Sci.* 77:1132–1140.
- Leslie, K. E., R. Dingwell, L. Yan, A. Bashiri, and P. Johnstone. 2007. An evaluation of Sentsortec Cellsense® for determining udder health status in lactating dairy cattle. Pages 232–233 in Proc. National Mastitis Council 46th Annual Meeting. San Antonio, Texas. NMC, Verona, WI.
- Maatje, K., P. J. M. Huijsmans, W. Rossing, and P. H. Hogewerf. 1992. The efficacy of in-line measurement of quarter milk electrical conductivity, milk yield and milk temperature for the detection of clinical and subclinical mastitis. *Livest. Prod. Sci.* 30:239–249.
- McClish, D. K. 1989. Analyzing a portion of the ROC curve. *Med. Decis. Making* 9:190–195.
- Nielsen, N. I., T. Larsen, M. Bjerring, and K. L. Ingvarsten. 2005. Quarter health, milking interval, and sampling time during milking affect the concentration of milk constituents. *J. Dairy Sci.* 88:3186–3200.
- Pepe, M. S. 2003. *The Statistical Evaluation of Medical Tests for Classification and Prediction*. Oxford University Press, Oxford, UK.
- Provost, F., and T. Fawcett. 2001. Robust classification for imprecise environments. *Mach. Learn.* 42:203–231.
- Rasmussen, M. D. 2004. Detection and separation of abnormal milk in automatic milking systems. Pages 189–197 in *Automatic Milking: A Better Understanding*. A. Meijering, H. Hogeveen, and C. J. A. M. de Koning, ed. Wageningen Academic Publishers, Wageningen, the Netherlands.
- Rasmussen, M. D., M. Bjerring, and F. Skjøth. 2005. Visual appearance and CMT score of foremilk of individual quarters in relation to cell count of cows milked automatically. *J. Dairy Res.* 72:49–56.
- Sarikaya, H., and R. M. Bruckmaier. 2006. Importance of the sampled milk fraction for the prediction of total quarter somatic cell count. *J. Dairy Sci.* 89:4246–4250.
- SAS Institute Inc. 2002. *SAS Online Doc*. SAS Institute Inc, Cary, NC.
- Schalm, O. W., and D. O. Noorlander. 1957. Experiments and observations leading to development of the California mastitis test. *J. Am. Vet. Med. Assoc.* 130:199–204.
- Sherlock, R., H. Hogeveen, G. Mein, and M. D. Rasmussen. 2008. Performance evaluation of systems for automated monitoring of udder health: Analytical issues and guidelines. Pages 275–282 in *Mastitis Control—From Science to Practice*. T. J. G. M. Lam, ed. Wageningen Academic Publishers, Wageningen, the Netherlands.
- Whyte, D. S., R. G. Orchard, P. S. Cross, T. Frietsch, R. W. Claycomb, and G. A. Mein. 2004. An on-line somatic cell count sensor. Pages 235–240 in *Automatic Milking: A Better Understanding*. A. Meijering, H. Hogeveen, and C. J. A. M. de Koning, ed. Wageningen Academic Publishers, Wageningen, the Netherlands.
- Whyte, D., M. Walmsley, A. Liew, R. W. Claycomb, and G. A. Mein. 2005. Chemical and rheological aspects of gel formation in the California Mastitis Test. *J. Dairy Res.* 72:115–121.