Journal of Dairy Research (2005) **72** 406–414. © *Proprietors of Journal of Dairy Research 2005*
doi:10.1017/S0022029905000993 *Printed in the United Kingdom*

Visual scoring of clots in foremilk

Morten D Rasmussen*

Danish Institute of Agricultural Sciences, Research Centre Foulum, DK-8830 Tjele, Denmark

Received 19 November 2004 and accepted for publication 2 February 2005

The necessary unequivocal and generally accepted definitions of normal and abnormal milk are not available. A precise definition is needed in order for companies to develop sensors to detect and sort abnormal milk at the time of milking. Experts at a workshop defined abnormal milk to be that from cows whose foremilk had changed in homogeneity or was coloured by blood. The objectives of this paper were: firstly, to explore how different groups of people scored the appearance of foremilk; and secondly, to develop a method suitable as an objective reference for testing of manual and automatic detection systems. Consumers, farmers and advisors did not agree on the visual appearance of normal, watery, clotty milk, or milk with blood, and experience is needed to score the visual appearance of foremilk correctly. It seems reasonable to expect a sensitivity of at least 70% for detection of abnormal milk during foremilking. Filter sizes 0·05, 0·07, 0·1, 0·2, 0·5, 1·0, and 2·0 mm were used to filter milk from cows with visually abnormal foremilk. If clots appeared in the foremilk, clots appeared on all size filters, but the filter with pore size 0·1 mm was the easiest to read and work with. The filter method is not reliable in identifying quarters with watery, yellowish, or bloody milk, whereas the method seems consistent, and at least as good as scoring of visual appearance in finding clots in the milk. Clots should show clearly on the filter to be counted as abnormal milk. All clinical cases with clots in the foremilk can be found on the filter and such cases have high somatic cell count (SCC). Both trained and untrained persons using the filter method can score normal and abnormal foremilk with a high specificity (>90%) and a high sensitivity (>80%). The filter method is recommended as a reference for scoring the homogeneity of foremilk.

Keywords: Normal and abnormal milk, definition, automatic milking systems.

Consumers expect milk to be healthy and wholesome and to be produced by healthy animals. The hygienic quality of the bulk milk is monitored by random sampling and analysis for somatic cell count (SCC) and total bacterial count. Additionally, it is expected that the milker diverts milk from cows with clinical mastitis. The general conditions for hygienic milk production in the EU are defined in Commission Directive 89/362/EEC (1989), and Chapter III-4 states that the milker must inspect the appearance of the milk before milking the individual cow and withhold milk from delivery if abnormalities are detected. Fulfilment of this directive is presently a problem with automatic milking systems (AMS) because normally a person is not present to inspect the foremilk. The proposal for the coming Hygiene Regulation part B, *Hygiene during milking, collection and transport* (Official Journal of the EU 2004/C 48 E/23) opens up the possibility of having other methods in place which produce similar results to the human checking of foremilk for abnormalities. This means that technical solutions may replace visual inspection for detection of abnormal milk, either before or during milking, and subsequently separate abnormal milk.

However, unequivocal and generally accepted definitions of normal and abnormal milk are not available. Before AMS companies can develop sensors to detect abnormal milk, a precise definition of abnormal or unacceptable milk is needed. A workshop on the definition of normal and abnormal milk at time of milking was held at the Danish Institute of Agricultural Sciences in November 2002 (Rasmussen, 2002). Participants in the workshop were scientists, legislators, veterinarians and people from the milking machine and dairy factory industries. There was a consensus at the workshop that there should be no double standards and that the reference method must apply to conventional as well as automatic milking (Rasmussen, 2004). This means that fairly simple and robust methods are needed. The participants in the workshop were in favour of basing the definition of abnormal milk caused by clinical mastitis on the

*For correspondence; e-mail: MortenD.Rasmussen@agrsci.dk

homogeneity of the milk and not on colour since the colour of milk changes with breed, stage of lactation and feeding. Moreover, there was consensus at the workshop that the cell count of milk should not be included in the definition of abnormal milk at the time of milking. A high cell count is a clear indicator of inflammation in the udder (Kitchen, 1981), but cannot be required to be measured at every milking for determination of abnormal milk. Consequently, a definition of abnormal milk including a cell count limit as proposed by Smith et al. (2001) cannot be used universally. It is still recommended that cell count is part of the milk quality survey of bulk milk (Rasmussen, 2004).

Very few scientists have dealt with an exact definition of the visual appearance of milk in relation to hygienic quality. Many papers deal with treatment of clinical mastitis, prevalence and incidence rates, and effect on milk yield and milk quality. However, the actual appearance of clots or flakes in the milk is seldom presented. Rasmussen & Larsen (2003) proposed that agglutination, changes in the casein fraction and proteolytic activity, could be reasons for milk to clot. Clinical mastitis covers a whole range of conditions from watery milk to milk being strongly abnormal in terms of consistency, colour, smell and taste, and the general condition of the quarter and cow may or may not be affected. Consequently, the milker may use different senses to detect clinical mastitis and sort milk (Hillerton, 2000). In principle, AMS have the same opportunities to test the milk but sorting of milk requires an exact definition of normal and abnormal milk. Changes in milk composition due to mastitis are numerous (Kitchen, 1981; Harmon, 1994; Korhonen & Kaartinen, 1995) and it is obvious that picking one method and definition will compromise some of the others. Using the appearance of foremilk as a reference will, however, make the method universal and applicable to all kinds of milking.

The objectives of this paper were to explore how different consumer groups, farmers, and advisors score the appearance of foremilk in order to set a limit or sensitivity for what they think is acceptable; and secondly, to develop a method that can be used as an objective reference when testing different sensing systems.

Material and methods

Visual scoring of dishes with milk from cows with clinical mastitis

A test panel of 15 comprising five milk quality advisors, five milkers, and five consumers not dealing professionally with raw milk scored the visual appearance of normal milk and milk from five cows with clinical mastitis and high SCC. Three containers were consecutively filled with about 50 ml of foremilk from each of the four quarters and two samples of 10 ml were taken from each container and then poured into separate dishes. The test panel scored a total of 120 Petri dishes with milk four-at-a-time to simulate scoring of milk from the four quarters. The four dishes were placed on a dark brown board which could be tilted to observe the viscosity and homogeneity of the sample. The milk samples were milked out and presented to the panel for scoring within 15 min. Samples were scored as normal, watery, containing clots, blood or colostrum. The panel had 30 s to score each series of four samples, write their scoring on the sheet provided along with their initials, drop the sheet in a closed box and move to the next booth. Colour was measured by use of a Chroma meter (CR-300, Minolta Co., Osaka 541-8556, Japan) where 2 ml of a mixture was transferred to a black capsule and colour measured immediately. The colour was expressed on three scales within the visible spectrum: 'L' – a light/dark scale that runs from 0 (black) to 100 (white), 'a' – a red/green scale (– is green, and + is red), and 'b' – a blue/yellow scale (– is blue, and + is yellow). The remaining milk in each container was analysed for fat, protein and SCC.

Visual scoring of slides with abnormal milk

It was difficult to keep milk in a homogeneous state in the small dishes and have many people score the milk samples one at a time. Consequently, a new series of pictures was taken of foremilk that was normal (visually normal and low SCC), watery (appeared thin and shiny), contained blood, and was clotty (quarters with clinical mastitis). This 'true' status of the milk samples was set following consensus on visual scores between the author and two technicians. Each individual milk sample was placed in a laboratory syringe from where it was sprayed onto the plate of a black strip cup. Four pictures were taken during this simulation of pre-milking from the beginning with a lot of milk on the plate and until the plate had almost drained. Sixty slides of four such pictures of milk from the same 'pre-milking' were shown. Twelve slides were repeated. The first 40 slides did not contain blood and the test panel was told that. The last 20 pictures included 8 pictures with a blood content of 0·03% to 1·0%. Out of the 60 slides, 26 were from quarters with the foremilk appearing normal, 8 were watery, 18 were from quarters with clinical mastitis where clots appeared in the foremilk, and 8 were normal milk intentionally mixed with blood. These scorings were regarded as the true status. The slides were scored by 20 milk quality advisors, 24 farmers, 25 veterinarians from the Danish Food and Veterinary Administration, and by a consumer group representing 12 students, 6 technicians, and 15 housewives with no direct relation to dairy farming.

Sorting of milk based on a filter method

Experiments were set up to find an objective method to score the appearance of milk. Filter sizes 0·05, 0·07, 0·1, 0·2, 0·5, 1·0, and 2·0 mm were used to filter milk from

https://doi.org/10.1017/S0022029905000993 Published online by Cambridge University Press

cows with visually abnormal foremilk. The idea was that clots on the filter could then be scored instead of the visual appearance in a strip cup. A strip cup was formed out of black plastic so that the foremilk could drain through the filter and into a container. In this manner, filters could be changed with the container and the sampled milk analysed. About 10 ml of milk from each quarter was foremilked into the modified strip cup. Foam turned out to disturb the scoring but pouring about 10 ml of water through the filter solved this problem. Milk of different breeds and fat percentages was filtered. Milk did not pass the two smaller size filters as easily as the larger ones. The larger filter sizes were more difficult to read than the smaller ones and were also more difficult to mount in the modified strip cup. The 0·1-mm filter was the most convenient filter to work with and, additionally, this pore size is about the particle size that is visible to the human eye. Multiple persons scored several foremilk samples on the 0·1-mm filter and full agreement was obtained. Unfortunately, these results were not filed and cannot be reported here.

Milk from quarters with clinical mastitis was filtered through the largest down to the smallest filter and clots were visible on all filters. Filters were easily blocked by milk from clinical mastitis. Milk from a few cases of clinical mastitis was then filtered from the smallest to the largest size filter and again clots were visible on all filters. Obviously, there were factors in the milk that made it clot again when some clots were removed, and the pore size of the filter seemed to be of less importance.

A herd with three automatic milking units and about 130 cows was foremilked once weekly during a year and the milk scored for visual appearance, CMT-score and appearance of the filter. Two trained technicians performed the scoring. In all, 24 167 foremilk samples were scored. Foremilk was scored as normal, watery, clots, yellowish, or containing blood. CMT was scored on a scale of 1–5 and the expected cell counts of the CMT-scores were: (1) <150 000, (2) 150 000–300 000, (3) 300 000–800 000, (4) >800 000 and (5) $>3\times10^6$ cells/ml. The filter was scored as normal, very few (<4) and small flakes (<2 mm), or clots.

Visual scoring of abnormal milk on filters

Pictures were taken during the foremilking of 20 cows of which 10 had one or two quarters with clinical mastitis. In total, 14 quarters had clinically abnormal milk and 66 were normal. The foremilk was run through filters of the sizes 0·1, 0·2 and 0·5 mm and pictures were taken of these filters. A panel comprising six veterinarians and six consumers not dealing professionally with raw milk scored the visual appearance of the foremilk and filters. In total, 80 slides of four pictures each were shown. The panel had 10 s to evaluate each slide and write their score on the sheet of paper provided. Pictures were scored as normal, small flakes, or clots. There were only two samples with small flakes and because of the difficulty of judging the size of these flakes on the screen, these samples were included as samples with clots.

Statistical methods

The statistical procedure PROC GENMOD (SAS Institute, 1999) was used to test scoring of samples. Model 1 was used for experiment 1:

$$Y_{ik}=\mu+\text{Subsample}_i+\text{Person(Group)}_k+\text{Group}_l+\text{Sample}_j*\text{Group}_{jl}+\varepsilon_{ik} \quad (1)$$

where
Y_{ik}=Visual score,
μ=overall mean,
Subsample_i=the fixed effect of Subsample (i=1, …, 120),
Sample_j=the fixed effect of Sample included as a repeated subject (j=1, …, 60),
Person(Group)_k=the fixed effect of person within Group (k=1, …, 15),
Group_l=the fixed effect of Group (l=milk quality inspector, milker, consumer),
$\text{Sample}*\text{Group}_{jl}$=the interaction between Sample and Group,
ε_{ik}=residual error.

A mean deviation of the score was calculated within sample and person and set at 0 if the person had the same score and at 1 if scores differed. This binary outcome was tested with the following model where the explaining variables are as described above:

$$Y_{ik}=\mu+\text{Sample}_i+\text{Person(Group)}_k+\text{Group}_l+\varepsilon_{ik}/\text{link=logit dist=bin} \quad (2)$$

The correct results of the scores were known for the second experiment and the binary outcome of correct or incorrect was tested with the following model:

$$Y_{ik}=\mu+\text{Sample}_i+\text{Visual}_j+\text{Person(Group)}_k+\text{Group}_l+\text{Visual}*\text{Group}_{jl}+\varepsilon_{ik}/\text{link=logit dist=bin} \quad (3)$$

where
Y_{ik}=Binary outcome (correct, incorrect),
μ=overall mean,
Sample_i=the effect of sample included as a repeated subject (i=1, …, 60),
Person(Group)_k=the fixed effect of Person within Group (k=1, …, 102),
Visual_j=the fixed effect of Visual appearance (j=normal, watery, clots, blood),
Group_l=the fixed effect of Group (l=milk quality inspector, veterinarian, farmer, others),
$\text{Visual}*\text{Group}_{jl}$=the interaction between Visual appearance and Group
ε_{ik}=residual error.

https://doi.org/10.1017/S0022029905000993 Published online by Cambridge University Press

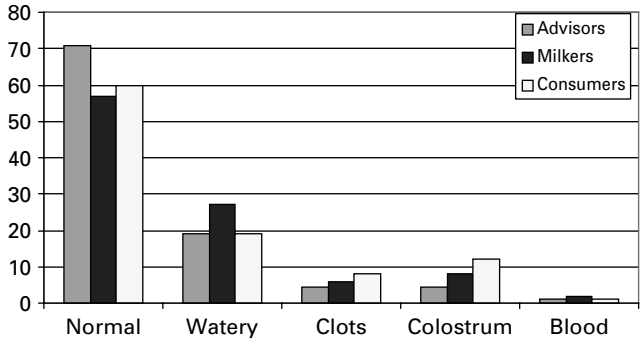

Fig. 1. Percentages of dishes with normal milk and milk from quarters with clinical mastitis scored visually as normal, watery, containing clots or blood. Milk quality advisors, milkers (operators), and consumers comprised the test panel.

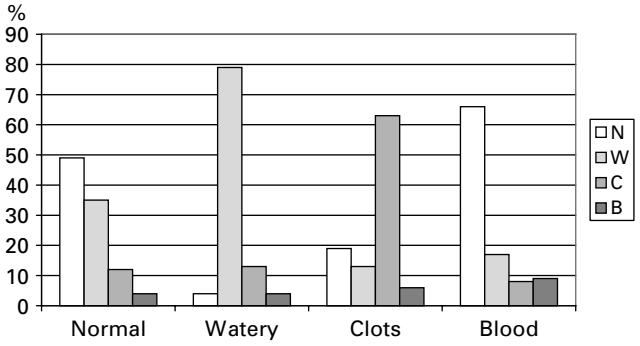

Fig. 2. Slides with normal and abnormal milk scored as being Normal (N), Watery (W), containing Clots (C), or Blood (B) in percentage of the true status shown on the X-axis.

The model used for experiment 4 was:

$$Y_{ijk} = \mu + \text{Sample}_i + \text{Filter}_j + \text{Person(Group)}_k + \text{Visual}_l + \text{Group}_m + \text{Visual} * \text{Group}_{lm} + \varepsilon_{ijk}/\text{link=logit dist=bin} \quad (4)$$

where
Y_{ijk}=Binary outcome (correct, incorrect),
μ=overall mean,
Sample_i=the effect of sample included as a repeated subject (i=1, …, 80),
Filter_j=the fixed effect of filter size (j=100, 200, 500 µm),
Person(Group)_k=the fixed effect of Person within Group (k=1, …, 12),
Visual_l=the fixed effect of the true Visual appearance on the filter (l=normal, flakes and clots),
Group_m=the fixed effect of Group (m=consumer, veterinarian),
$\text{Visual} * \text{Group}_{lm}$=the interaction between visual appearance on the Filter and Group
ε_{ik}=residual error.

Means are presented as Least Squares Means.

Results and Discussion

Visual scoring of dishes with milk from cows with clinical mastitis

On average, samples were scored as normal, watery, clots, colostrum or containing blood in 63%, 22%, 6%, 8% and 1% of the samples (Fig. 1). Milk quality advisors scored more of the samples as normal than milkers and consumers did ($P<0{\cdot}001$). Consumers scored more of the samples as containing colostrum than the other groups did ($P<0{\cdot}001$), and mainly when the fat percentage was high. An explanation for this could be that consumers are only accustomed to looking at milk taken directly from the refrigerator and expect the milk to be white. In general, the test panel did not agree on the scoring and only 10% of the samples were given exactly the same score by everybody. Milk quality advisors agreed the most with 37% of the samples having the same score, and milkers and consumers agreed the least with only 12%and 18% of the samples being scored equally within the group ($P<0{\cdot}01$). Milk quality advisors were the most consistent in their scoring with 82% of the subsamples having the same score compared with 65% and 67% for milkers and consumers ($P<0{\cdot}001$).

There was no conclusive score of the milk samples, but the scoring could be compared with the SCC and a colour scanning. Milk samples that most of the test panel scored as having clots had high SCC. However, 25% of the samples scored as normal milk had SCC $>10^6$/ml and some were even above 10^7/ml. The main conclusion is that it is not possible to differentiate between milk samples with high and low SCC just by looking at the visual appearance. The R^2 value of logSCC regressed on the outcome of the colour scanner was 0·19, but 0·69 for the visual mean score of each sample regressed on the colour. There seems to be a possibility of using colour scanning as an aid in the differentiation between normal and abnormal appearance of the milk as confirmed by Ouweltjes & Hogeveen (2001) and Espada & Vijverberg (2002). However, if a high SCC is included in the definition of abnormal milk at time of milking, this property has to be measured more directly.

Visual scoring of slides with abnormal milk

Only about 50% of the normal samples were scored as normal milk (Fig. 2), but about 35% were scored as being watery. The change from having much to almost no milk on the strip cup could be interpreted as if the milk was watery when the layer of milk became thin (see Fig. 3). Droplets of milk on the plate of the strip cup may have been interpreted as clots since 12% of the normal samples were scored as having clots. Close to 80% of the watery samples were scored correctly as watery and this high score was probably given because these samples were very thin and watery. Milk from cows with clinical mastitis changed from having few flakes to being abnormal in both

https://doi.org/10.1017/S0022029905000993 Published online by Cambridge University Press

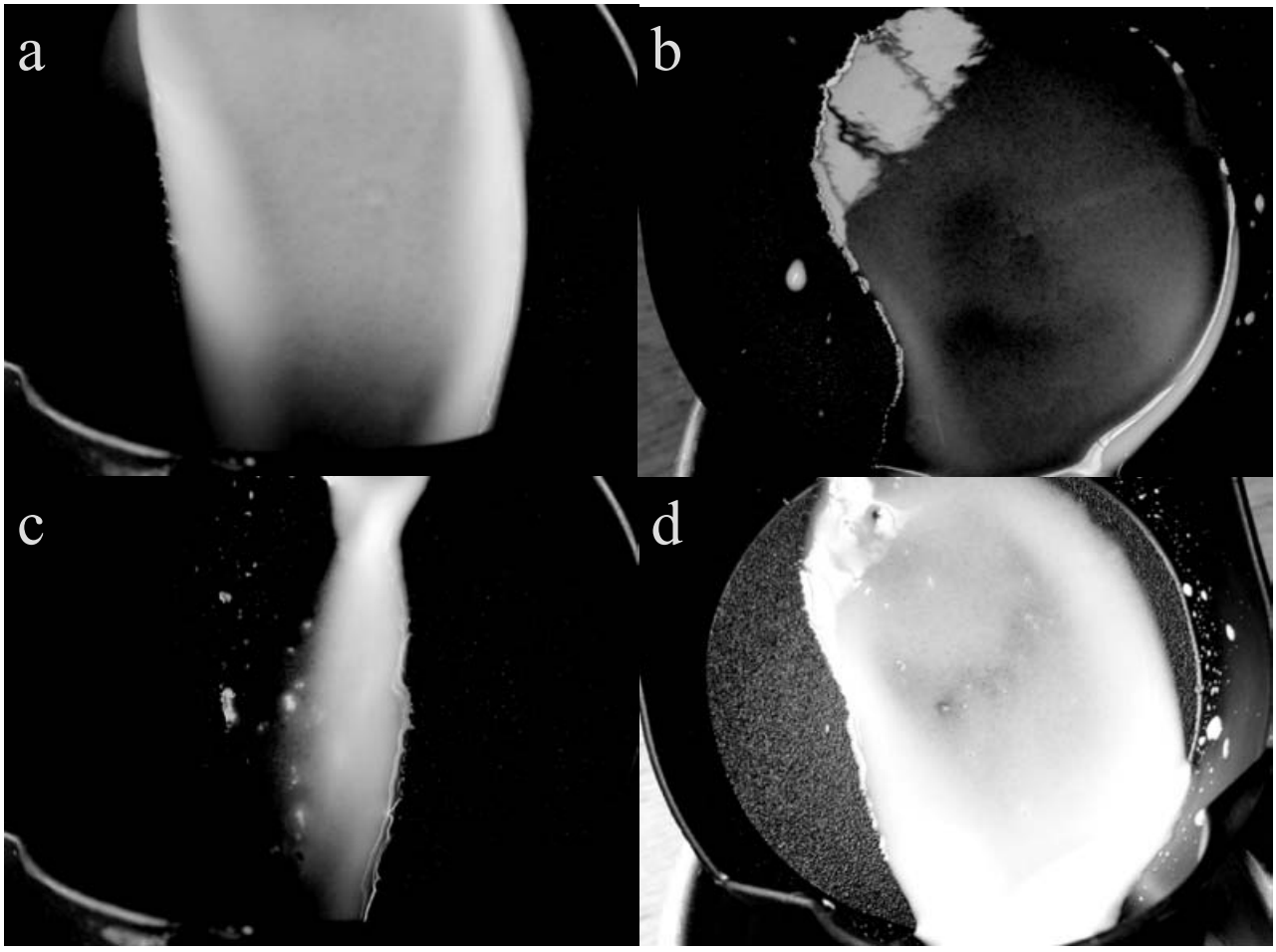

Fig. 3. Examples of slides with normal milk (a and b) and milk with clots (c and d) being scored correctly (a and c) or mainly incorrectly (b and d).

homogeneity and colour. The panel scored 63% of the abnormal samples correctly and about 20% were scored as being normal. It proved to be very difficult to identify samples with blood, less than 10% of which were placed correctly. Blood in milk does not show well on a black plate but is easily detected when comparing a container with a small percentage of blood being next to a container with white milk (Rasmussen & Bjerring, 2005).

The different groups scored slides differently in relation to the true status of the milk sample (Fig. 4). Milk quality advisors had the highest percentage of normal milk samples scored correctly and farmers the least ($P<0{\cdot}001$), which was mainly because farmers scored 37% of the normal samples as watery. Milk quality advisors and veterinarians scored 88% and 86% of the true watery samples correctly, where in fact all did well. For samples with clots in the milk, veterinarians scored best ($P<0{\cdot}01$) having 76% of the samples scored correctly, closely followed by the farmers. Milk quality advisors recognized 63% of the samples with clots as having clots and the group of others settled on 57%. Hillerton (2000) gives a sensitivity of 80% for manual detection of clots in the milk, which is comparable to the figures found here for people trained to look at raw milk. It seems reasonable to expect a sensitivity of at least 70% for detection of abnormal milk during foremilking.

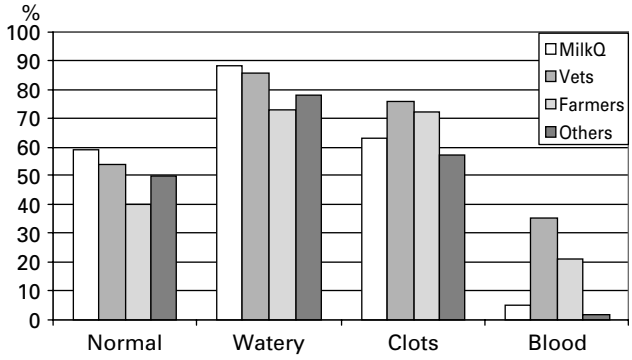

Fig. 4. The percentage of samples (slides) with normal, watery, clots, or blood in the milk scored correctly by groups of milk quality advisors, veterinarians, farmers and other consumer groups.

https://doi.org/10.1017/S0022029905000993 Published online by Cambridge University Press

Table 1. The visual appearance of foremilk in a strip cup and on a filter of size 0·1 mm: numbers of samples in each category (% of the row in brackets)

	Appearance on the filter		
Visual appearance	Normal	Few, small flakes	Clots
Normal	24 942 (97·6)	550 (2·2)	69 (0·3)
Watery	117 (79·0)	22 (14·9)	9 (6·1)
Clots	4 (2·0)	6 (3·1)	185 (94·9)
Yellowish	124 (53·2)	94 (40·3)	15 (6·4)
Blood	26 (59·1)	14 (31·8)	4 (9·1)

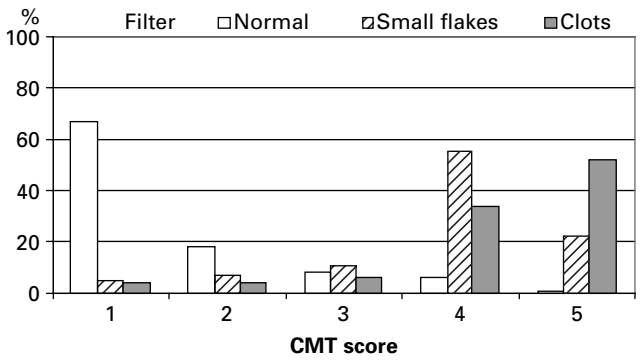

Fig. 5. The distribution of CMT-scores of foremilk samples that appeared normal, had small flakes, or clots on a filter with a pore size of 0·1 mm.

Although not impressively so, veterinarians and farmers were better at detecting samples with blood than the other two groups ($P<0·001$). However, these slides were difficult to score correctly even knowing the true result, and the technical conditions of the presentation almost certainly influenced the results. It was estimated that the technical conditions played a minor role in identifying quarters with normal, watery, and abnormal milk. Milk quality advisors and veterinarians were the most consistent in reproducing their score of the 12 repeated slides and they also had the highest percentage of correct answers for these slides.

Examples could easily have been chosen of normal and abnormal milk that would have a high probability of being scored correctly. However, many of the samples with abnormal milk were selected from quarters where foremilk (probably cisternal milk) was clearly abnormal, and ejected milk (probably alveolar milk) instantly appeared normal until milk drained off the strip cup and left clots behind. The focus of the panel test was abnormal milk and this may have moved the panel to score more samples as being abnormal than normal in order to be sure to find all abnormal milk samples. The percentage of abnormal milk samples (57%) was much higher than during normal milking, which gives a better evaluation but probably overestimates the percentage of correct answers for abnormal milk. The large differences between percentages of correctly scored samples within groups of observers highlight the need for a more objective method of categorizing foremilk.

Sorting of milk based on a filter method

Visually abnormal milk was found by the technicians in 2·4% of the samples, CMT-score 5 in 1·5%, and clots on the filter in 3·7% of the samples. About 0·3% of the samples appearing visually normal showed as clots on the filter (Table 1) and a further 2·2% were found with small flakes. Watery milk was not detected on the filter in 79% of the cases. Only four of the samples scored visually as milk with clots did not show clots on the filter. Three of these samples were from one cow at one milking with no special remarks and the last one was marked as being yellowish and creamy. Consequently, virtually all milk samples with clots will show on the filter as well. About 47% of the samples being yellowish and 41% of the samples with blood in the milk could be detected on the filter, and the filter method is consequently not useful to detect these abnormalities. CMT-score 5 was given to 65% of the samples with visual clots *v.* 54% of the samples showing clots on the filter (Fig. 5). About 55% of the foremilk samples with small flakes on the filter had CMT-score 4. Only 14% of the samples with clots on the filter had CMT-scores of 1 to 3. Out of the group with CMT-score 5, 60% of the samples were visually abnormal, but 75% showed as clots or small flakes on the filter.

The visual appearance of the milk from an infected quarter will depend on the physiological status of the cow and may or may not show as distinct clots. Out of 184 quarters with clots in the milk, 28% had clots in the milk at the next weekly scoring. Out of 285 quarters with clots on the filter, 34% kept this status at the next weekly scoring. This difference was not significant (Chi-square test). However, the immediate conclusion from this is that the filter method is at least as stable (if not better) in categorizing the visual appearance of the foremilk as the visual method.

In conclusion, the filter method is not reliable in identifying quarters with watery, yellowish, or bloody milk whereas the method seems consistent and as good as the visual method of finding clots in the milk. Clots should show clearly on the filter to be counted as abnormal milk.

Visual scoring of abnormal milk on filters

Overall, the panel scored 96% of the pictures of filters correctly. The lowest percentage of correctly scored normal milk samples (specificity) was 91 and four of the persons scored 100% (Fig. 6). The percentage of correctly scored samples with flakes or clots on the filter (sensitivity) was 83% at the lowest and 98% for the best scoring

https://doi.org/10.1017/S0022029905000993 Published online by Cambridge University Press

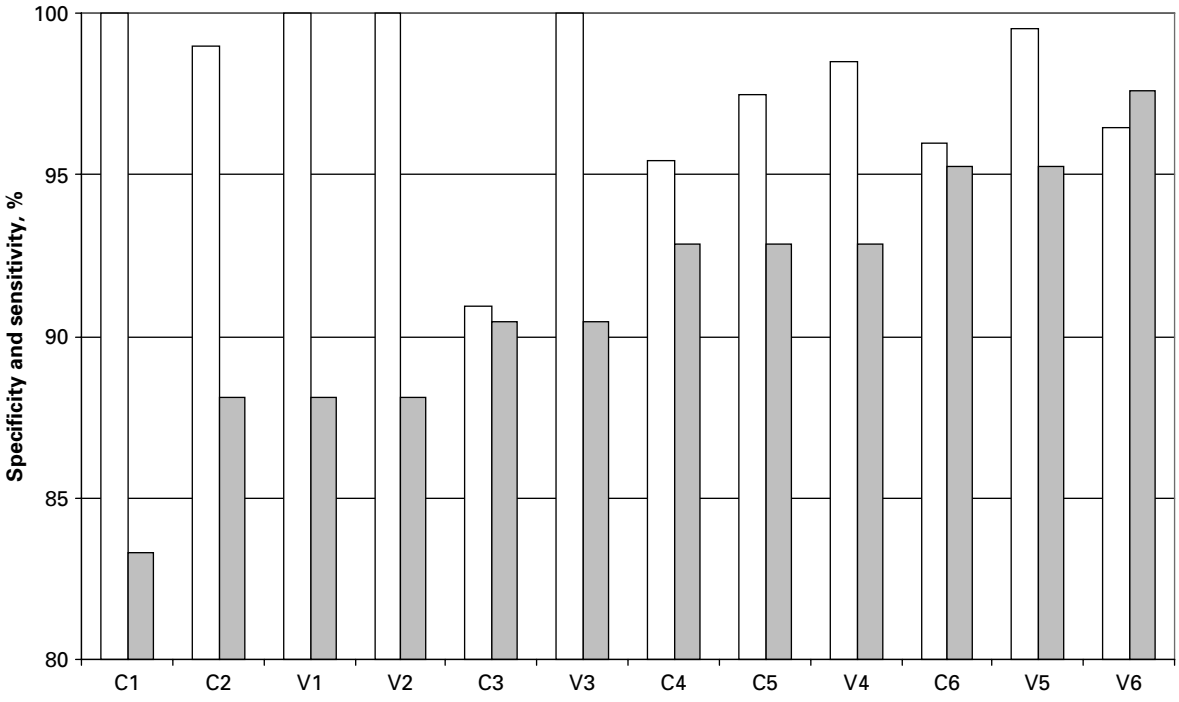

Fig. 6. Specificity (□) and sensitivity (■) for scoring of foremilk sampled through filters with a pore size of 0·1 mm and scored by consumers (C1 to C6) and veterinarians (V1 to V6).

(Fig. 6). The group of veterinarians had a higher mean percentage of correct scores (96·8%), than the consumer group with 94·9% ($P<0·05$). This difference was mainly due to the fact that veterinarians had a higher sensitivity (95·3%) than the consumer group (91·6%). In any case, the percentage of correct scores was much higher with this filter method than in the other experiments. The percentage of correct scores was 97·6% for the filter size of 500 µm *v.* about 95% for the two smaller filters ($P<0·05$). This difference was due to a sensitivity of 98% for the larger size filter *v.* about 91% for the smaller filters. The pictures of the large size filter was generally easier to read than the pictures of the smaller filters, which differs from the direct observations made in the previous experiment where the smaller size filters were judged as the best ones. In conclusion, trained and untrained persons using the filter method can score normal and abnormal foremilk with a high specificity and sensitivity.

The purpose of detecting abnormal milk

A full session at the symposium 'Automatic milking – a better understanding' (Meijering et al. 2004) focused on detection of abnormal milk and the 19 papers reflect that a lot of work is focusing on this subject and that there is a need for this information. Most of the papers reported on indirect measurements of abnormal milk by the use of electrical conductivity, but reports were also presented on the use of spectrophotometry (Wiedemann & Wendl, 2004), near infra-red spectroscopy (Tsenkova et al. 2004) and gel formation during CMT-testing (Whyte et al. 2004). Maassen-Francke et al. (2004) developed an optical sensor composed of a digital camera to detect flakes or clots $\geqslant$0·1 mm. The equipment was tested for its ability to distinguish between clots and other particles like straw, sawdust, sand, foam or spots caused by reflection. About 90% of the objects were classified correctly, which looks promising in relation to automatic diversion of milk changed in homogeneity. Such a method for detection of abnormal milk would probably correlate well with the proposed reference method in this paper. Most of the current indirect detection methods have a broader objective than just finding clots in the milk and are not designed for automatic diversion, but more for the detection of quarters or cows developing subclinical or clinical mastitis. It is very important with regard to production economy and animal welfare that the udder health is monitored at every milking. There may be very different purposes in identifying infected quarters: subclinical for observation, diversion of abnormal milk, diversion of milk with high SCC, treatment of clinical mastitis, culling of cows or quarters, or re-acceptance for delivery of milk. All such information is needed for management decisions in well-managed herds. However, inclusion of all these purposes into one list of cows whose milk is diverted would cause an economic loss for the farmer (Pietersma & Hogeveen, 2004).

Cows and quarters with visually abnormal foremilk have high CMT-scores (and SCC) leading to a poorer milk quality for dairy processing because the activity of

https://doi.org/10.1017/S0022029905000993 Published online by Cambridge University Press

degradative enzymes increases with the SCC. A detection system based on measurement of SCC and diversion of the milk with the highest cell counts to bring the bulk tank SCC within the quality limits, would offer the best price for the produced milk (Nielsen et al. 2002). However, not all abnormal milk will be withheld from delivery and especially not if SCC is only measured on composite milk where milk from diseased quarters are diluted with milk from healthy quarters. Having quality limits for bulk milk SCC regulates the proportion of cows delivering milk with high SCC, but does not necessarily exclude milk from cows with abnormal milk. Diversion of milk from cows where the foremilk appears abnormal would for one thing detect cows and quarters with clinical mastitis, but would also ensure an aesthetic milk production. Consumers expect the milk to be homogeneous, white and produced by healthy animals. Consequently, milk that differs from normal in colour and homogeneity should not be delivered for consumption, as stated at the workshop on definition of abnormal milk (Rasmussen, 2002) and in accordance with the EU Commission Directive. Consequently, the detection system for automatic diversion has to be specifically targeted at abnormal milk, which further requires that the definition of abnormal milk and the reference method are objectively described.

According to the first experiment in this paper, only 10% of the evaluators fully agreed on the visual scoring of foremilk, which illustrates the need for more objective methods for evaluating the performance of different manual or automatic detection systems. Much greater agreement between evaluators was found using the filter method. The filter method may not be the best at identifying cows with udder health problems, but it can detect virtually all cases where milk has changed in homogeneity. Discarding milk from such cows and quarters will ensure that visually abnormal milk is withheld from delivery.

Conclusions

Experience is needed to score the visual appearance of the foremilking correctly and the method is subjective. It seems reasonable to expect a sensitivity of at least 70% for detection of abnormal milk during foremilking. Scoring of foremilk run through a filter with a pore size of 0·1 mm is less subjective but does not give the same result as visual scoring in a strip cup. The filter method is not reliable in identifying quarters with watery, yellowish, or bloody milk, whereas the method seems consistent and as good as the visual scoring in finding clots in the milk. According to the outcome of a workshop on definition of abnormal milk, milk from cows with clinical mastitis should only be scored on homogeneity and not on colour. Clots should show clearly on the filter for the milk to be counted as abnormal. All clinical cases with clots in the foremilk will be found on the filter and such cases have high SCC. Both trained and untrained persons using the filter method can score normal and abnormal foremilk with a high specificity (>90%) and a high sensitivity (>80%). The filter method is recommended as a reference for scoring the homogeneity of foremilk and thereby classifies milk from the quarter and cow in question as being normal or abnormal. Different manual and automatic systems for detection of clinical mastitis may then be tested against this reference.

The technical help of Martin Bjerring, Erik S Frimer, Mette L Bech and Birgitte Madsen in performing foremilk scoring, setting up panel tests and slide shows is gratefully acknowledged. The European Commission (QLK5-2000-31006) and the Danish Directorate for Food, Fisheries and Agri Business financed this work. The contents of this publication are the sole responsibility of its authors, and do not necessarily represent the views of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use, which may be made of the information.

References

Commission Directive 89/362/EEC 1989 General conditions of hygiene in milk production holdings

Espada E & Vijverberg H 2002 Milk colour analysis as a tool for the detection of abnormal milk. In *Proceedings of the First North American Conference on Robotic Milking*, March 20–22, Toronto, Canada, IV-28–IV-38

Harmon RJ 1994 Symposium – Mastitis and genetic evaluation for somatic-cell count – Physiology of mastitis and factors affecting somatic cell counts. *Journal of Dairy Science* **77** 2103–2112

Hillerton JE 2000 Detecting mastitis cow-side. *National Mastitis Council Annual Meeting* **39** 48–53

Kitchen BJ 1981 Review of the progress of dairy science: Bovine mastitis: milk compositional changes and related diagnostic tests. *Journal Dairy Science* **48** 167–188

Korhonen H & Kaartinen L 1995 Changes in the composition of milk induced by mastitis. In *The Bovine Udder and Mastitis*, pp. 76–82 (Eds M Sandholm, T Honkanen-Buzalski, L Kaartinen & S Pyörälä). Helsinki, Finland: University of Helsinki, Faculty of Veterinary Medicine

Maassen-Francke B, Wiethoff M, Suhr O, Clemens C & Knoll A 2004 A method to detect flakes and clots in milk in automatic milking systems. In *Automatic Milking – A Better Understanding*, p. 251 (Eds A Meijering, H Hogeveen & CJAM de Koning). The Netherlands: Wageningen Academic Publishers

Meijering A, Hogeveen H, de Koning CJAM (Eds) 2004 *Automatic Milking – A Better Understanding*, 525 pp. The Netherlands: Wageningen Academic Publishers

Nielsen AHN, Skjøth F & Rasmussen MD 2002 Economic consequences of sorting of milk based on visual appearance, CMT-score and somatic cell count on both quarter level and composite milk level. *Internal report no. 169 for Workshop of the EU-project (QLK5-2000-31006): Implications of the introduction of automatic milking on dairy farms*. November 27, 2002, 91–100

Ouweltjes W & Hogeveen H 2001 Detecting abnormal milk through colour measuring. *National Mastitis Council Annual Meeting* **40** 217–219

Pietersma D & Hogeveen H 2004 Cost of discarding milk with automatic separation of abnormal milk. In *Automatic Milking – A Better Understanding*, pp. 221–227 (Eds A Meijering, H Hogeveen

& CJAM de Koning). The Netherlands: Wageningen Academic Publishers

Rasmussen MD & Bjerring M 2005 Visual scoring of milk mixed with blood. *Journal of Dairy Research* (In press)

Rasmussen MD (Ed.) 2002 Definition of normal and abnormal milk at time of milking. *Internal report no. 169 for Workshop of the EU-project (QLK5-2000-31006): Implications of the introduction of automatic milking on dairy farms*. November 27, 2002. 102 pp

Rasmussen MD 2004 Detection and separation of abnormal milk in automatic milking systems. In *Automatic Milking – A Better Understanding*, pp. 189–197 (Eds A Meijering, H Hogeveen & CJAM de Koning). The Netherlands: Wageningen Academic Publishers

Rasmussen MD & Larsen LB 2003 Milking hygiene: new issues and opportunities from automatic milking. *Italian Journal of Animal Science* **2** 283–289

SAS Institute Inc 1999 SAS OnlineDoc®, Version 8.2, Cary, NC: SAS Institute Inc

Smith KL, Hillerton JE & Harmon RJ 2001 NMC guidelines on normal and abnormal milk based on SCC and signs of clinical mastitis. National. Mastitis Council. Madison WI, USA

Tsenkova R, Morita H, Shinzawa H & Hillerton JE 2004 Detection of abnormal udder tissue and milk by near infra-red spectroscopy (cow side). In *Automatic Milking – A Better Understanding*, p. 259 (Eds A Meijering, H Hogeveen & CJAM de Koning). The Netherlands: Wageningen Academic Publishers

Whyte DS, Orchard RG, Cross P, Frietsch T, Claycomb RW & Mein GA 2004 An on-line somatic cell count sensor. In *Automatic Milking – A Better Understanding*, pp. 235–240 (Eds A Meijering, H Hogeveen & CJAM de Koning). The Netherlands: Wageningen Academic Publishers

Wiedemann M & Wendl G 2004 The use of spectral photometry for detection of mastitis milk. In *Automatic Milking – A Better Understanding*, pp. 228–234 (Eds A Meijering, H Hogeveen & CJAM de Koning). The Netherlands: Wageningen Academic Publishers