# Use of Total and Differential Somatic Cell Counts from Composite Milk Samples to Detect Mastitis in Individual Cows

I.R. Dohoo, A.H. Meek, S.W. Martin and D.A. Barnum\*

# ABSTRACT

The objective of this study was to ascertain the value of variables measured on composite milk samples as predictors of mastitis in individual cows. The standard of comparison was the results obtained from the bacteriological examination of individual quarter foremilk samples. Cows were classified as negative or positive with regard to mastitis on the basis of one quarter sampling only and cows which were impossible to classify in this manner were omitted from subsequent analyses.

The variables that were examined were: the presence or absence of specific bacteria, demographic data, and logarithmically transformed total somatic cell counts and percentages of cell volume in channels 7 through 12 of a Coulter Counter. It was found that the inclusion of all variables resulted in correct classification of 95.9% of cows with regard to their mastitis status. Sequential elimination of individual variables or groups of variables in an attempt to simplify the procedure reduced the correct classification to 86.8% when only the log transformation of the total somatic cell count and the demographic data were included. The ability of a function which included the logarithm of the total somatic cell count, the logarithm of the percentage in channel 8 and demographic data, to classify cows was examined in detail and the sensitivity and specificity of the function also discussed. It is also shown that with increasing age the minimum total somatic cell count required to classify a cow as positive increased and possible explanations of this phenomenon are discussed.

# RÉSUMÉ

Cette étude consistait à vérifier la valeur des parametres variables d'échantillons composés de lait, comme detecteurs de la mammite, chez des vaches donnees. Les resultats de l'examen bacteriologique des premiers jets de lait de quartiers déterminés servirent de base de comparison. On classa les vaches comme atteintes ou non de mammite en se basant sur un seul echantillonnage; on elimina des analyses subsequentes celles qu'on ne reussit pas ainsi a classer.

On etudia les parametres suivants: la presence ou l'absence de bactéries spécifiques, les données démographiques, ainsi que la transformation logarithmique des comptages totaux des cellules somatiques et des pourcentages du volume cellulaire, du canal #7 au canal #12 d'un appareil Coulter. On r6alisa que <sup>l</sup>'inclusion de tous ces parametres permettait de classer correctement 95.9% des vaches, par rapport à la mammite. L'6limination s6quentielle de paramètres donnés ou de groupes de paramètres, dans le but de simplifier l'analyse, réduisit l'exactitude du classement à 86.8%, quand on ne tenait compte que de la transformation logarithmique du comptage total des cellules somatiques et des donn6es démographiques. On examina en detail la raison pour laquelle une fonction qui incluait la transformation logarithmique du comptage total des cellules somatiques et celle du pourcentage, dans le canal #8 de l'appareil Coulter, permettait de classer correctement les vaches. On commenta aussi la sensibilité et la spécificité de cette fonction. On démontra 6galement qu'avec le vieillissement, le comptage total minimal des cellules somatiques requis pour declarer une vache positive augmentait; on commenta enfin les raisons probables de ce phénomène.

# INTRODUCTION

The use of electronic particle counters to determine the total somatic cell count of milk is a well established and reliable procedure (3,4,8,10,12) and several authors (5,9,13) have suggested that the inclusion of differential counts (based on cell volume) may aid in the detection of bovine mastitis. The ability of Coulter Counters to classify cells by size into one of 16 channels makes this a fairly simple technique. At the Ontario Veterinary College (OVC), two counters are available to determine somatic cell counts in milk. A Coulter Milk Cell Counter with attached TA-II (MCC)1 automatically counts and

\*Departmentof Veterinary Microbiology and Immunology, Ontario Veterinary College, University of Guelph, Guelph, Ontario N1G 2W1. Submitted May 14, 1980.

'Coulter Electronics Ind., Hialeah, Florida.

records total somatic cell counts as well as the percentage of the total cell volume in channel 8 (particles with volumes of  $89.2 \mu^3$  to  $178.3 \mu^3$ ) and can process up to 200 samples per hour with minimal prior sample preparation. The other counter, <sup>a</sup> less automated Model TA (TA), can be used to determine total cell counts and full differential counts (channels 7 through 12). Newbould (9, 10) demonstrated that milk from infected quarters has a relative increase in the percentage of cells in channel 8 due primarily to increased numbers of neutrophils. Sheldroke et al (13) found that milk from mastitic quarters had a peak corresponding to a modal cell volume of 102  $\mu^3$ which is in the range covered by channel <sup>8</sup> of the MCC and TA.

The objective of this study was to identify variables determined from a composite milk sample, that are useful in predicting whether or not a cow has subelinical mastitis. First, indirect indicators of mastitis (total and differential cell counts) were compared with direct indicators (bacteriological data), using discriminant analysis, in order to determine if the more readily ascertained indirect indicators were reliable predictors of mastitis. Second, a comparison of a complete differential somatic cell count (total cell count and channels 7 through 12) was made with a partial differential count (total cell count and channel <sup>8</sup> only). The MCC automatically provides the latter data and therefore offers the possibility of a rapid inexpensive screening test for mastitis. Finally, the contributions of various variables in a classification function were examined in more detail. Cows were classified as positive or negative based on the examination of individual quarter foremilk samples.

# MATERIALS AND METHODS

# COLLECTION OF MILK SAMPLES

A group of 39 dairy herds in the Belleville area of Ontario were used for the study. Individual foremilk quarter samples and composite samples were collected using sterile techniques in screw top vials and the samples were shipped to the Ontario Veterinary College in refrigerated coolers. The date of birth, last calving date and daily milk production as recorded on the latest Dairy Herd Improvement Association test prior to sampling were recorded for each cow.

#### EXAMINATION OF MILK SAMPLES

Upon arrival at the OVC, each quarter sample was cultured using routine procedures and a California Mastitis Test (CMT) was performed. On the basis of these two tests quarters were classified with respect to subclinical mastitis as follows: a negative quarter was defined as one having no major pathogens isolated and <sup>a</sup> CMT reaction of trace or less, whereas a positive quarter was one from which a major pathogen was isolated and the CMT was one or higher. Corynebacterium bovis and nonhaemolytic staphylococci were considered minor pathogens and a quarter containing only these organisms was considered negative if the CMT score was trace or less. Quarters with <sup>a</sup> low CMT reaction but from which a major pathogen was isolated were classified as "group 3" while those with an elevated CMT but no major pathogens were classified as "group 4".

The composite samples were also cultured, following which all samples were fixed with formalin, incubated and a somatic cell count performed using the MCC which automatically records the total somatic cell count (SCC) in thousands of cells per mL and the percentage of the total cell volume in channel 8 (PCH8).

The same sample was then counted on the TA counter using the method described by Newbould (8). In this case the total (SCC-B) and the percentage of total cell volume in channels 7 through 12 (P7, P8, P9, P10, P11, P12) were recorded.

#### ANALYTIC METHODS

Initially descriptive statistics of the variables under study and transformations of some of the variables were examined to determine if certain transformations were appropriate in order to make the variable's distribution more nearly normal. Data from cows with unusual histories, such as clinical mastitis present at the time of sampling or treatment for clinical mastitis within the past week, were removed from the file.

Cows were then classified with respect to subelinical mastitis as negative if all milking quarters were negative or positive if they had one or more positive quarters. Cows that had a combination of negative, "group 3" and/or "group 4" quarters were excluded from subsequent discriminant analyses as it was impossible to classify the cow as positive or negative with regard to mastitis, based on one sample only.

Of 1020 cows originally sampled, were classified as negative, were classified as positive and were unclassified.

Common logarithmic transformations were made on all total somatic cell counts, resulting in the new variables, LSCC and LSCC-B (1). Logarithmic, square root and logistic transformations were applied to the differential percentages. It was found that the skewness of the distribution was minimized by the common logarithmic transformation. Consequently, it was applied to the raw data resulting in the following new variables (LPCH8, LP7, LP8, LP9, LP10, LP11, LP12). Values of LPCH8 that corresponded to values of  $LSCC \geqslant 3.301$  (equivalent to 2,000,000 cells/mL) were adjusted upwards so that a regression of LPCH8 on LSCC had a zero slope for  $LSCC \geqslant 3.301$ . The formula for this adjustment was:

#### $LPCH8$  (adjusted) = LPCH + 0.54858 (LSCC - 3.301),

where 3.301 is the common log of 2,000 (equivalent to 2,000,000 cells/mL) and -0.54858 was the slope of the initial regression of

LPCH8 on LSCC for samples with more than 2,000,000 cells/mL. Hereafter, LPCH8 will refer to the adjusted LPCH8.

Stepwise discriminant analyses were then performed using various sets of variables in order to determine how well different variables were able to distinguish between negative and positive cows. Within each analysis, stepwise inclusion of variables was based on their ability to maximize Rao's V statistic, resulting in the greatest overall separation of the groups. At each step inclusion of an additional variable and retention of any variable already incorporated in the function depended on the partial F ratio for the variable being greater than unity. Finally, correlations between the logarithmically transformed results from all cows determined on the MCC and TA counter were examined.

#### RESULTS

Table <sup>I</sup> provides a description of the variables used in the analyses. Table II contains means for several of the variables in the negative and positive groups of cows, determined either from the composite sample or the cow's history. Positive cows were generally older, slightly later in lactation, were producing slightly less milk and had higher total and channel 8 counts than negative cows.

Seven separate discriminant analyses were performed. Table III provides a summary of the variables made available in each analysis and the percent of cows correctly classified. In each succeeding analysis, either the amount of data that was made available was reduced or the data that were used came from the more automated source (i.e. MCC as opposed to TA counter). The fourth discriminant analysis was examined in more detail and Table IV lists the variables that were selected by the stepwise procedure (not in order of selection) along with their standardization discriminant coefficients (s.d.c.).

TABLE I. Variables Used in Discriminant Analyses to Classify Cows with Regard to Mastitis Status

Variable Name	Variable Definition
Age Dm Prodn	Age in years Days in milk Daily milk production (kg) at previous DHIA test
Sh	Staphylococcus aureus
Sa Sna	Streptococcus agalactia Streptococcus nonagalactia
Coli	Coliforms
Nm Chovis	Nonhaemolytic micrococci Corynebacterium bovis
<b>Op</b>	Other pathogen
Onp <b>LSCC</b>	Other nonpathogen
LPCH <sub>8</sub>	Log of SCC ('000/mL) determined on MCC Log of % in channel 8 determined on MCC
LSCC-B LP7 to LP12	Log of SCC('000/mL) determined on TA Log of % for channels 7 to 12 determined on TA

TABLE II. Means of Some Variables for Cows With and Without Mastitis. Date from 39 Dairy Herds in Eastern Ontario, 1979



'Geometric means

<sup>b</sup>Difference between positive and negative significantly different at  $p < 0.05$ 

Confirence between positive and negative significantly different at  $p<0.01$ 



TABLE III. Variables Made Available for Discriminant Analyses and the Percentage of Cows Correctly Classified with Regard to Mastitis

All positive and negative cows, (subsequently referred to as the set of cases), were randomly divided into two subsets containing 379 and 423 cases respectively and the fourth disciminant analysis (using LSCC, LPCH8, Age, Dm and Prodn) was rerun on the first subset to obtain a classification function. This function was subseTABLE IV. The Standardized Discriminant Coefficients of Variables Selected in the Fourth Discriminant Analysis<sup>a</sup>



'See Table III

<sup>b</sup>See text for explanation





'The discriminant function from analysis 4 (Table III) was used in this series  $P$ Represents the actual probabilities in this set of cases

TABLE VI. Mean Age and Percentage of Cows Positive and Negative for Several Age Groups with the Minimum Total Somatic Cell Count Required to Classify a Cow as Positive<sup>®</sup>

Age Range (yrs.)	Number of cows	Mean Age (yrs.)	$\%$ Neg	% Pos	Minimum SCC (2000/mL)
2-3.99	282	3.1	84.8	15.2	183
4-5.99	247	4.9	68.8	31.2	217
6-8.99	207	7.3	49.8	50.2	266
9	66	10.9	29.2	70.8	269
all ages	802	5.4	66.2	33.8	228

'Based on the sixth discriminant analysis (Table III)

TABLE VII. The Logarithmic Means, the Standard Deviations and the Untransformed Means of the SCC and Percent Channel <sup>8</sup> on the TA and MCC Coulter **Counters** 

		Mean Log value	S.D. Log value	Untransformed Means
Total SCC	MCC	$2.28^*$	0.43	191,000/mL
	TA	$2.33^*$	0.43	212,000/mL
% channel 8	MCC	$0.91^*$	0.42	8.13%
	TA	$1.11^*$	0.36	12.89%

'Difference between mean on MCC and mean on TA significantly different at p< 0.01

quently used to classify both the first and second subsets with 86.02% of the first subset and 88.12% of the second subset being correctly classified for a pooled average of 87.13%.

The fourth discriminant analysis was also repeated several times with various prior probabilities of group (i.e. positive or negative) membership. The results of sequentially increasing the prior probability of positive group membership are summarized in Table V.

The set of cases was divided into four separate subsets on the basis of age and the sixth discriminant analysis (using LSCC only) was repeated on each subset. Table VI contains the average age for each group, the percentage of cows in the group that are positive or negative and the minimum total somatic cell count that would be required in order for a cow to be classified as positive.

The means and standard deviations for the total and percent channel 8 counts for each machine are presented in Table VII. The correlation coefficients for the logarithm of the total count was 0.91 and for the logarithm of the percentage in channel 8 was 0.88.

# DISCUSSION

The objective of this study was to

determine how well various variables, measured on a composite milk sample, were able to differentiate between cows that had mastitis in at least one quarter and cows that did not have mastitis. Discriminant analysis was the analytical method of choice for reaching this objective and some of the factors which might influence the results of the analysis will be elaborated on throughout the discussion.

The transformation of a variable in order to make its distribution more nearly normal makes its use as a discriminating variable more compatible with the statistical theory of discriminant analysis. For total somatic cell counts, Ali and Shooke determined that a logarithmic transformation is optimal (1). Three different transformations of differential percentages were examined and based on the logarithmic transformation's ability to minimize skewness it was selected for use in the analysis. In addition, experience with the MCC indicates that for total cell counts greater than 2,000,000 cells/mL the percent of total cell volume recorded in channel 8 is erroneously low, possibly due to more than one cell passing through the orifice of the counting tube at a time. Therefore, using the formula previously described, adjustments were made to the 26 values with cell counts of 2,000,000 cells/mL or more.

The first analysis, which included all possible variables (Table III) provided a base line classification level against which other procedures were compared. The high level of success (95.9% correctly classified) was expected because bacteriological data were included and the classification of the cows as negative or positive had been based, to a large extent, on the bacteriological examination of individual quarter samples. The fact that 100% of the cows were not correctly classified was possibly due to failure of organisms isolated from quarter samples to appear in the composite or the appearance in the composite of a pathogenic organism (usually Staphylococcus aureus) when no pathogenic organism was isolated from the quarter samples.

Since the routine culturing of milk samples requires sterile collection of samples and is a moderately time-consuming procedure one of the objectives was to determine how well cows could be classified on the basis of somatic cell counts (total and differential) only. Other workers have reported the benefits of recording cell volumes (5, 9, 13) and an objective in this study was to determine if this benefit could be obtained by recording channel 8 values only or if other channels contributed a significant amount of extra information. Therefore, in the second analysis the values for all channels were made available and, for comparison purposes, in the third analysis only the channel 8 and total SCC values were used. There was a drop of 5.8 percentage points in overall classification between analyses 1 and 2 and the subsequent exclusion of channels 7 and 9 through 12 resulted in a further drop of less than one percentage point. The decrease accompanying the exclusion of the bacteriological data may be more a function of the definition of mastitis used (i.e. requiring the isolation of a pathogenic bacterium) than the ability to predict cows that have an inflammatory process in the udder. The minimal decrease in classification that occurred between analyses 2 and 3 indicates that if differential counts are to be used it is only necessary to record percent channel 8 values.

In order to determine if the automatic MCC was comparable to the TA counter, analysis <sup>4</sup> was conducted using data from the MCC. A drop of 2.4 percentage points in the "percent correctly classified" was found but this loss may be offset by the advantage of the MCC being fully automated and able to process up to 200 samples per hour with minimal prior preparation.

A comparison of results from analyses 4 and 5 revealed that overall there was no advantage to be gained by including channel 8

values if the total somatic cell count was available. However, it was decided to use the function derived in analysis 4, which selected LSCC, LPCH8, Age, and Dm as important variables for further analysis. This was because the MCC used at the OVC requires no additional effort to derive percent channel 8 values and in specific cases such as samples taken early in lactation, knowledge of the percent channel 8 value may help determine a cow's status with regard to mastitis.

Analysis 6 demonstrated that once the cell count data are included, the addition of historical data (Age, Dm and Prodn) is only of marginal benefit. The fact that the historical information alone was able to correctly classify 70.6% of cows (analysis 7) does not imply that by themselves they are valuable predictors. In this set of cases, 66% of the cows were negative and therefore any function could correctly classify 66% by predicting that all cows were negative. Therefore, historical information alone resulted in a gain of less than five percentage points in the overall classification.

Many packaged programs (including SPSS) for discriminant analysis use all cases in the data set to develop the discriminant function and then classify the same set of cases. This may lead to an overly optimistic estimate of the ability of the function to classify new cases (6). To obviate this problem the cases were randomly divided into two subsets and a discriminant function was derived on the first subset and subsequently used to classify new cases from the second subset. The discriminant function that was derived from the first subset was virtually identical to the one developed from the complete set of cases and was equally as successful at correctly classifying cows. This provides additional evidence of the potential value of this function in the field as a predictor of an individual cow's status with regard to mastitis from a composite milk sample.

The estimates of the ability of the various discriminant functions to

correctly classify cows may be overly optimistic for another reason. The necessity of classifying cows as positive or negative resulted in the exclusion of 218 cows from subsequent analyses and this group may in fact be more difficult to correctly classify using information from a composite sample. The possibility also exists that some variables which do not appear important in classifying the majority of cows are important in this group. Subsequent studies, using repeated quarter sampling, may be able to reduce or eliminate the number of cows that have to be excluded.

In analyses 1 through 5 the prior probabilities that a cow had, or did not have, mastitis were equal. Changing the prior probability of group membership (see Table V) is analogous to altering the critical level of a screening test. As the prior probability of a cow being positive was increased from 0.15 to 0.65 the likelihood of correctly classifying a positive cow was increased, while the number of negative cows that were incorrectly classified as positive also increased. This is equivalent to increasing the sensitivity (ability to correctly detect positives) of the test and decreasing the specificity (ability to correctly classify negatives). Except at the extremes, the overall percentage correctly classified remained fairly constant. The level at which to set the prior probabilities then becomes a decision of the user, depending on whether it is important to detect all cases of mastitis (and risk having more false positives) or to be certain that all positives actually are positive (and risk more cases going undetected).

The relative importance of any variable in a discriminant function is given by its s.d.c. The greater the absolute magnitude of the s.d.c. the greater is the importance of that variable. However, in any multivariate analysis the standardized coefficients of highly correlated variables are unstable and may be difficult to interpret. This occurs because once one of the correlated variables has been entered into the

discriminant function the amount of additional information that the other correlated variable(s) can provide is greatly diminished resulting in a relatively small standardized coefficient. In this case LSCC and LPCH8 are highly correlated (pooled within group correlation coefficient = 0.83) and therefore their joint effect should be considered, not their importance relative to each other. Consequently it can be seen that in the function developed in analysis 4, the cell count and channel 8 together play a very important role, while age provides a moderate amount of extra information and days in milk is a relatively unimportant variable.

The fact that production was not selected as a discriminating variable should not be interpreted as indicating no association between mastitis and production but instead it indicates that once the other information was known, the milk yield did not help predict the cow's mastitis status. Again, this is partially due to the correlation between production and age and days in milk (pooled within group correlation coefficients of 0.21 and -0.45 respectively).

The inclusion of age as a moderately important variable after the cell count had been entered demonstrates that in general older cows are more likely to be positive for mastitis. In order to determine the minimum somatic cell count required to classify a cow of a specific age as positive, the set of cases was divided into four subsets based on the age of the cow and analysis number 6 (using LSCC only) was repeated on each subset. The minimum cell count required for a positive classification rose with age as is shown in Fig. 1.

There are several possible explanations for this rise. Previous work has shown that there is a slight increase with age in somatic cell counts in bacteriologically negative cows (2, 7). Natzke et al suggest that part of this rise may be due to the greater prevalence of resolved infections in older cows (7). In addition several authors have demonstrated that infections



Fig. 1. Relationship between age and the minimum somatic cell count required to classify a cow as positive.

with minor pathogens increase somatic cell counts and consequently any rise in the prevalence of these minor infections with age would result in a corresponding increase in the average somatic cell count (2, 7, 14). Also, Eberhart demonstrated that older cows produce higher somatic cell counts in response to minor pathogens than young cows do and thus even if the prevalence of minor infections remains constant with age the average somatic cell counts for older cows would increase (2). This study did not subdivide negative cows into bacteriologically negative cows and cows with minor infections so evaluation of these explanations was not possible.

Previous work (3) has shown the comparability of results from Coulter Electronics semiautomatic electronic cell counter (Model TA) and the automatic milk cell counter (MCC). Similar results were obtained in this study with high correlation coefficients for both total and channel 8 values, except that the values for percentage in channel <sup>8</sup> on the MCC were, on average, lower than the corresponding values on the TA.

# ACKNOWLEDGMENTS

The excellent technical work of

R. Johnston, R. Johnson, E. Stewart and N. Smart is gratefully acknowledged. This work was financed by grants from the Ontario Milk Marketing Board and the Ontario Ministry of Agriculture and Food. The senior author is supported on a fellowship from the Medical Research Council of Canada.

#### REFERENCES

- 1. ALI, A.K.A. and G.E. SHOOK. An optimum transformation for somatic cell concentration in milk. J. Dairy Sci. 63: 487-490. 1980.
- 2. EBERHART, R.J., H.C. GILL-MORE, L.J. HUTCHINSON and S.B. SPENCER. Somatic cell counts in DHI samples. Proc. Natn. Mastitis Council. pp. 32-40. 1979.
- 3. GINN, R.E., V.S. PACKARD and D.R. THOMPSON. Comparison of the automated with the semi-automatic coulter counter method and the direct microscopic somatic cell count (dmsec) on raw milk samples. J. Fd Protect. 42: 567-568. 1979.
- 4. HEESCHEN, W. Determination of somatic cells in milk (technical aspects of counting). Proc. I.D.F. Seminar on Mastitis Control. Doc. 85: 79-92 Int. Dairy Fed., Brussels, Belgium. 1975.
- 5. MEEK, A.H., D.A. BARNUM and F.H.S. NEWBOULD. Use of total and differential somatic cell counts to differentiate potentially infected from potentially non-infected quarters and cows and between herds of various levelsof infection. J. Fd Protect. 43: 10-14. 1980.
- 6. MORRISON, D.G. On the interpretation of discriminant analysis. J. Market. Res. 6: 156-163. 1969.
- 7. NATZKE, R.P., R.W. EVERETT and D.S. POSTLE. Normal milk somatic cell counts. J. Milk Fd Technol. 35: 261-263. 1972.
- 8. NEWBOULD, F.H.S. Electronic counting of somatic cells in farm bulk tank milk, J. Milk Fd Technol. 37: 504-509. 1974.
- 9. NEWBOULD, F.H.S. A technique for differential somatic cell counts in milk. Proc. I.D.F. Seminar on Mastitis Control, Doc. 85: 136-141 Int. Dairy Fed.,

Brussels, Belgium. 1975.

- 10. NEWBOULD, F.H.S. Effect of several factors on total and differential counting of somatic cells in farm bulk tank milk. J. Fd Protect. 41: 96-99. 1978.
- 11. NIE, N.H., C.H. HULL, J.G. JEN-KINS, K. STEINBRENNER and D.H. BENT. Statistical Package for the Social Sciences. 2nd Ed. New York: McGraw-Hill Inc. 1975.
- 12. PEARSON, J.K.L., C.L. WRIGHT and D.O. GREER. A study of methods for estimating the cell content of bulk milk. J. Dairy Res. 37: 467-480. 1970.
- 13. SHELDRAKE, R.F., R.J.T. HOARE, V.E. WOODHOUSE and G.D. McGREGOR. Cell volume to aid analysis and technique of somatic cell counts in milk. J. Dairy Sci. 60: 882- 888. 1977.
- 14. SHULTZ, L.H. Somatic cell counting of milk in production testing programs as a mastitis control technique. J. Am. vet. med. Ass. 170: 1244-1246. 1977.
- 15. SYRSTAD, O., I. RON and J. WIG-GEN. Factors affecting cell counts in milk from individual cows. Nordisk Veterinaermedicin. Denmark 31: 114- 121. 1979.