

EVALUATION OF BOVINE TEAT CONDITION IN COMMERCIAL DAIRY HERDS:

3. GETTING THE NUMBERS RIGHT

D.J.Reinemann⁵, M.D.Rasmussen², S. LeMire⁵, F.Neijsen³, G.A.Mein¹, J.E.Hillerton⁴,
W.F.Morgan¹, L.Timms⁵, N. Cook⁵, R.Farnsworth⁵, J.R.Baines⁴, and T. Hemling⁵
“Teat Club International”, c/o F. Neijshuis, Research Institute for Animal Husbandry
PO Box 2176, 8203 AD Lelystad, The Netherlands. email: f.neijshuis@pv.agro.nl
Co-authors from: Australia¹, Denmark², The Netherlands³, UK⁴, USA⁵

Paper Presented at the Proceedings, AABP-NMC International Symposium on Mastitis and Milk Quality,
Vancouver, BC, Canada. September 2001.

Classification of bovine teat condition can be used to assess the effects of milking machines, milking management or environment on teat tissue and the risk of new intra-mammary infections. Veterinarians and others require a simple and reliable method for evaluating teat health in dairy herds. Teat condition can be classified using continuous measures (e.g., teat thickness measured in mm) or categorical scores (e.g., smooth or rough). Categorical data are often presented using a numerical scale (e.g., teat end condition = 1, 2, 3, or 4). This may lead to false statistical conclusions if the relative risk factors are not distributed according to their numerical value. Other common statistical problems of evaluating teat condition in the field are insufficient sample size, using improper statistical tests for skewed data, and samples not representing the population in question. For more information on sample size, statistical inference, and guidelines for the collection of statistically defensible data, see LeMire et al. (1998, 1999). This paper provides simple guidelines for statistical evaluation of the teat conditions (Table 1) proposed in the companion paper by Mein et al. (2001).

Sampling all cows: In small herds (up to 100 cows) it may be practical to score the entire herd. In this case the entire population of cows has been measured and the sample mean is equal to the population mean. No additional statistical analysis is required to estimate this true population mean. Statistical analysis is still required, however, to determine if changes in this population mean are significant.

Random Sampling: In larger herds, random sampling is desirable in order to reduce the amount of time and resources for estimation of the true mean of the herd. The validity of all statistical analysis is based on a random sample of independent subjects from the entire population of subjects. Random samples should be taken from either all the cows or the production groups of interest. The outcome of the scoring is only valid for the population represented by the sample. One of the most common errors in estimating the condition of a herd based on a sample is lack of randomization. As the percentage of the population sampled decreases, the importance of random sampling increases. Most common statistical tests assume that the measure of interest in the population of subjects is normally distributed about some average value. Since the outcome of teat scoring is a yes/no response the following tables are based on a binomial distribution.

False Positive Error: Mein et al. (2001) propose describing the teat condition status of a herd of cows by the proportion of cows that have a particular teat condition classed as abnormal, e.g., more than 20% of cows with rough teat-ends. A statistical test will result in claims that the percentage of cows or quarters in a population differs from some critical value based on measurements of a given sample of the population. A false positive (Type I) error is the claim of difference (more than 20% of cows are affected) when the difference does not exist (20% or

fewer cows are affected). The probability of a false positive error is denoted by α . The commonly accepted level of error in scientific literature is a false positive probability of $\alpha < 0.05$. If statistical tests indicate that $\alpha < 0.05$ we can say with 95% confidence that the true population mean does indeed differ from the critical value. The accuracy of the estimate will increase (confidence interval decrease) with increasing sample size. While the 95% confidence level is normally the minimum acceptable for scientific publications, the risk of a false positive error is generally of less concern when diagnosing conditions on a farm. If we conclude that more than 20% of cows are affected at the 90% confidence level, there is a 1 in 10 probability of a wrong result (fewer than 20% are affected). According to the recommendations of Mein et al. (2001) we would undertake further investigation needlessly 1 out of 10 times. At the 80% confidence level the probability of a wrong result is 1 in 5.

Correlation between quarters: If one teat of a cow is affected, and another teat on that cow is more likely to be affected than a teat on a different cow, then the teat scores are correlated within cow. There are likely situations in which observations of quarters will be correlated within cow. In other situations, they may not be correlated. The statistical analysis presented here will be invalid if based on the number of teats affected whenever correlation exists between teat condition scores within cow. Furthermore, if a cow has one teat with poor condition, the associated risk factors are present for that cow. We recommend, therefore, that statistical analysis be done at the cow level.

How many cows?

Tables 2-4 present the minimum and maximum number of cows required to conclude that the problem with the poor teat condition is less than or greater than the expected value for confidence intervals of 80% (- +) and 90% (-- ++). These tables are based two tailed tests assuming that the true probability of finding a teat condition problem is 20, 10, and 5% respectively. These tables can be used for any test with a yes/no result.

We propose that all teats on each sample cow are scored with a minimum of 80 cows scored or 20% of the herd whichever is larger. Sampling more cows will increase the accuracy of the diagnosis. For each category of teat condition, the number of cows with at least one unacceptable teat condition score as well as the total number of unacceptable teat condition scores should be calculated. Use Tables 2-4 to determine the probability that the number of cows affected is below or above the threshold criteria for each teat condition measure. Record the diagnoses for each condition (--, -, o, +, ++) in Table 1.

The ratio of affected teats to affected cows will provide useful information on the distribution of problem teat conditions among cows. The relationship between problem and non-problem conditions also provides additional diagnostic information.

The first five conditions of Table 1 can be the result of short-term responses to machine milking. If any of these conditions score ++, further investigation into milking-induced problems should be conducted. Condition number 6, Teat End Roughness can be influenced by machine milking influences that have persisted over a period of several weeks. Teat End Roughness is also influenced by environmental conditions. Skin lesions also have primary causes other than machine milking (e.g., freezing and chemical exposure). If some of the short-term responses have a positive result then the probability of a machine milking related cause of the abnormal teat condition increases.

The examples below present statistical methods to determine the accuracy of estimating a population mean (e.g. does this population of cows have more than 20% of teats with poor teat condition score).

Table 1. Frequencies of teat condition problems that indicate need for further investigation.

| Teat Condition Measure | Criteria | Diagnosis | | | | |
|--------------------------|-----------------------------------|-----------|---|---|---|----|
| | | -- | - | o | + | ++ |
| 1. Color | > 20% visibly reddened or blue | | | | | |
| 2. Swelling at Teat Base | > 20% swelling or palpable rings | | | | | |
| 3. Swelling at Teat End | > 20% firm, hard or swollen | | | | | |
| 4. Openness | > 20% classified as open | | | | | |
| 5. Vascular Damage | > 10% petechiations | | | | | |
| 6. Teat End Roughness | > 20% Rough and Very Rough | | | | | |
| 7. Open Lesions | > 5% open lesions or cracked skin | | | | | |

Table 2. Binomial test results for 20% of cows expected to have a problem.

| Number of Cows Scored | Number of Cows with Problem Condition | | | | |
|-----------------------|---------------------------------------|----|--------------------|----|----|
| | -- | - | 20% Expected Value | + | ++ |
| 80 | 8 | 9 | 16 | 22 | 23 |
| 120 | 15 | 16 | 24 | 31 | 33 |
| 160 | 21 | 23 | 32 | 40 | 42 |
| 200 | 28 | 30 | 40 | 49 | 51 |

Table 3. Binomial test results for 10% of cows expected to have a problem.

| Number of Cows Scored | Number of Cows with Problem Condition | | | | |
|-----------------------|---------------------------------------|----|--------------------|----|----|
| | -- | - | 10% Expected Value | + | ++ |
| 80 | 2 | 3 | 8 | 13 | 14 |
| 120 | 5 | 6 | 12 | 18 | 19 |
| 160 | 8 | 9 | 16 | 22 | 24 |
| 200 | 11 | 12 | 20 | 27 | 28 |

Table 4. Binomial test results for 5% of the samples expected to have a problem.

| Number of Cows Scored | Number of Cows with Problem Condition | | | | |
|-----------------------|---------------------------------------|---|-------------------|----|----|
| | -- | - | 5% Expected Value | + | ++ |
| 80 | 0 | 0 | 4 | 7 | 8 |
| 120 | 1 | 1 | 6 | 10 | 11 |
| 160 | 2 | 3 | 8 | 13 | 14 |
| 200 | 3 | 4 | 10 | 15 | 16 |

(Tables 2 – 4 note: The values in the tables are the critical values to conclude that less or more than the expected value is affected at the 80% (- or +) and 90% (- - or ++) confidence levels. * implies that even with zero responses, we cannot conclude that fewer than the expected percentage of cows are affected.

Example I:

- Teat-end roughness scores are recorded for 240 quarters of 80 randomly selected cows

- Further investigation is required according to Mein et al. (2001) if more than 20% of cows have teat condition scores of R and VR.
- Consult Table 2 (20% expected value) with a sample size of 80
- If 8 or fewer cows have at least one teat with R and VR score we may conclude with 90% confidence that fewer than 20% of the population is affected (--).
- If 9 cows have at least one teat with R and VR score we may conclude with 80% confidence that fewer than 20% of the population is affected (-). More cows must be scored to increase the confidence of a negative diagnosis.
- If from 10 to 21 cows have at least one teat with R and VR score we may not conclude that the percentage of cows affected is different from 20% (either more or less). More cows must be scored to make a diagnosis.
- If 22 cows have at least one teat with R and VR score we may conclude with 80% confidence that more than 20% of the herd is affected (+). More cows can be scored to increase the confidence of the diagnosis.
- If 23 or more cows have at least one teat with R and VR score we can conclude with 90% confidence that more than 20% of the population is affected (++).

Example II :

- 480 quarters on 120 randomly selected cows are scored for vascular damage.
- 47 quarters on 20 cows have at least one teat with petechial haemorrhages
- Consulting Table 3 we may conclude with 90% confidence that more than 10% of cows have experienced vascular damage during milking (++).
- The ratio of quarters to cows is 47:20 or 2.35:1. On average each affected cow has 2 affected teats indicating that correlation between teats within cow is likely.

Example III:

- 480 quarters on 120 randomly selected cows are scored for open lesions.
- 1 cow is found to have at least one teat with open lesions or cracked skin.
- From Table 4 we can conclude with 90% confidence that fewer than 5% of cows in this herd have at least one teat with open lesions (--).

Cause and Effect

Another reason for scoring teat condition might be to make claims about cause and effect relationships (e.g., did a change in management practice result in a change in teat condition scores). It requires considerably more planning and effort to substantiate claims of cause and effect than to estimate a population mean.

It is important to have a clear idea of the desired comparison. The causal variable (e. g., change in milking vacuum) and the effect (e.g., change in teat condition score) must be clearly identified. The size of a meaningful effect must also be determined (e.g., change of 5% in the number of cows with poor teat condition). This information is used to frame a question such as: *Did a change of 2 kPa in milking vacuum result in more than 5% change in the number of cows with poor teat-end condition?* It is important to randomize cows to treatments to avoid unknown or unintended differences between treatment groups. Randomization of cows to treatment can be difficult because of the logistical problems of efficient cow movement.

Variability in the population, variability introduced by measurement error, and the number of measurements will determine the precision of measuring that outcome. Reducing variability or increasing sample size can improve predictive ability. The number of lactating cows fixes the maximum sample size on a farm. Other practical considerations (e.g., the way cows are grouped during milking) may reduce the maximum sample size further. The short-term effects of machine milking can be tested over a period of several days. The main advantage of short-term tests is that there is less chance for other factors like weather or changes in herd composition to influence the factors of interest. This time restriction reduces the variability from other sources and increases the ability to detect differences correlated with changes in the milking equipment or procedures. Paired tests use the same cows for measurement before and after the treatment is applied and generally reduce the variability in the response. Variability can be reduced by blocking on known sources of variability and randomize on unknown or uncontrollable sources of variability. The paper by LeMire et al, (1999) presents a more detailed discussion of experimental design and blocking strategies

Conclusions

- Score teat condition on all teats of all cows in the herd if time and herd size allow.
- If time is limited and/or for large herds, **randomly** select at least 80 cows or 20% of the herd (whichever is largest number of cows) for teat condition scoring.
- Keep a record of both the number of teats and the number of cows with problem conditions. If a cow has one or more teats with a problem condition then that cow is considered to have a problem condition.
- Use Tables 2-4 to record diagnosis (--, -, o, +, ++) for each condition on Table 1. Diagnoses - - and – indicate that a problem condition is not likely to exist in that herd for that teat condition score. Diagnoses of + or ++, indicate that it is likely that a problem condition exists for that teat condition and further investigation into possible causes is warranted. A diagnosis of o indicates that we cannot rule out a problem but we likewise cannot say with confidence that a problem exists. A larger sample size or further investigation may be warranted.
- Use the combination of problem and non-problem conditions as a guide to identify possible causes (Mein et al. 2001).
- Use the ratio of affected teats to affected cows as a guide to identifying whether the problem condition is evenly distributed among cows (low correlation between teats within cow) or if certain cows are more likely affected (high correlation between teats within cow).

References

- LeMire, S.D., D.J. Reinemann, G.A. Mein, M.D. Rasmussen, 1998. Statistical Considerations for Milking Time Tests. Paper No. 983128, Written for presentation at the 1998 ASAE annual International Meeting, Orlando Florida July 12-15, 1998
- LeMire, S.D., D.J. Reinemann, G.A. Mein, M.D. Rasmussen, 1999. Recommendations for Field Tests of Milking Machine Performance. ASAE Paper No. 993020 Written for Presentation at the 1999 ASAE/CSAE International Meeting, Toronto, Ontario, Canada, July 18-21, 1999

Mein, G.A., F. Neijenhuis, W.F. Morgan, D.J. Reinemann, J.E. Hillerton, J.R. Baines, I. Ohnstad, M.D. Rasmussen, L. Timms, J.S. Britt, R. Farnsworth, N. Cook, and T. Hemling. 2001. Evaluation of bovine teat condition in commercial dairy herds: 1. Non-infectious factors. Proceeding NMC and AABP, Vancouver, Canada.