ANALYSIS OF CLINICAL MASTITIS IN NORWEGIAN CATTLE WITH A LONGITUDINAL THRESHOLD MODEL

B. Heringstad[1], Y. M. Chang[2], D. Gianola[1,2] and G. Klemetsdal[1]

[1] Department of Animal Science, Agricultural University of Norway, Ås, Norway
[2] Department of Animal Sciences, University of Wisconsin, Madison 53706, USA

INTRODUCTION

In genetic evaluation of mastitis in first-lactation Norwegian Cattle (NRF), the trait is treated as binary. The score is based on whether or not a cow has had at least one case of veterinary treated clinical mastitis (CM) in the period going from 15 days before to 120 days after calving (Heringstad *et al*., 2000). This definition does not utilize fully the information available from the Norwegian health recording system, where all veterinary treatments are recorded. Some cows have several mastitis treatments during lactation and, also, the date of treatment is known, so it is possible to ascertain at what stage of lactation the disease takes place.

In recent years, there has been an increasing use of test-day models for genetic analysis of dairy production traits, as shown, e.g. in reviews of Swalve (2000) and Jensen (2001). Applications have mainly focused on continuous traits, such as milk yield. On the other hand, Rekaya *et al*. (1998) suggested a Bayesian approach for analyzing longitudinal binary traits. They presented an application to mastitis in Holsteins, with data consisting of sequences of binary responses.

The objectives of this study were: 1) to analyze longitudinal binary CM data for first lactation NRF using a longitudinal threshold model; 2) to examine relevant selection criteria from this model that can be used for ranking and selection purposes, and 3) to compare sire evaluations from the longitudinal model with those from cross-sectional models.

MATERIAL AND METHODS

Data. The data were from first-lactation daughters of 245 NRF sires having their first progeny test in 1991 or 1992. Only records of daughters with a first calving in 1990 through 1992, and from herds having at least 5 daughters of any of these test-bulls were kept. The resulting dataset included 36,178 first-lactation cows from 5286 herds. The pedigree file represented 437 males (via relationships through sires and maternal grandsires of the 245 sires above). All cases of veterinary treated CM in first-lactation, from 30 days before calving to culling, second calving or 300 days after calving, whichever occurred earlier, were included. About 77% of the cows had no CM during first lactation, while 16%, 5%, and 2 % had one, two, and three cases, respectively. Only 315 cows had more than four recorded CM treatments. Within cows with mastitis, the mean number of CM treatments during first lactation was 1.5. For each cow, the period going from 30 days before calving to 300 days after first calving was divided into 11 intervals of 30 days length each. Within each such interval, presence or absence of mastitis was scored based on whether or not the cow had at least one veterinary treatment of CM recorded. Mastitis frequencies were 4.7 and 10.1% in the first two intervals, respectively, and between

1.4 and 2.2 % in the following 9 intervals. If the cow had mastitis within the interval, the day of treatment was used in the time dependent function, otherwise the midpoint of the interval was used.

Statistical analyses. The binary CM data were analyzed with a threshold-liability model (e.g. Gianola, 1982). Based on preliminary studies, a longitudinal model with regressions on a Legendre polynomial function of order four was used for the underlying liability to CM. The model included effects of year of calving (3 classes), herd (5,286 classes), permanent environment (36,178 cows), regressions within age×season level, sire specific regressions, and error. The permanent environmental effects account for covariances between liabilities in different periods, assuming a constant correlation. Age×season of first calving effects were in 12 classes, from 3 age (<24, 24-27, and >27 months) and 4 season (March-May, June-August, September-November, December-February) levels. The variance of the residual distribution was assumed constant from period to period, and set equal to 1. A Bayesian approach was used for inferences. Proper uniform priors were assumed for the effects of year and for the "fixed" regressions by age×season classes. The effects of herd and of permanent environment were assigned independent normal priors with means zero and unknown variances. Independent scaled inverse chi-square prior distributions were assumed for the variances of herd and permanent environment. A multivariate-normal prior distribution was used for the sire effects, and the 5×5 (co)variance matrix of the sire regression coefficients, including the zero-order term, was assigned an inverse Wishart prior distribution. Draws from the posterior distributions of the parameters were obtained using a Gibbs sampler with data augmentation (Sorensen *et al.*, 1995). Inferences were based on 160,000 samples, with the first 40,000 iterations discarded as burn-in.

Genetic evaluation. Sire evaluations obtained with the longitudinal model may be presented as curves depicting the probability of mastitis along lactation. However, this information can be summarized into a single number that can be used for ranking and selection of sires. For example, the expected fraction of days without mastitis (EFD) in an interval (t_1, t_2) was calculated for each sire as: $EFD(t_1,t_2)_i = \sum_{j=t_1}^{t_2}[1-\Phi(\mu_{ij})]/(t_2 - t_1) = 1 - MD_i/(t_2 - t_1)$, where t_1 and t_2 are the first and the last day of the interval, respectively, $\Phi(\mu_{ij})$ is the probability that a daughter of sire *i* will have mastitis at day *j*, and $MD_i = \sum_{j=t_1}^{t_2}\Phi(\mu_{ij})$ gives the expected number of days with mastitis in the interval from t_1 to t_2. For EFD's, $\Phi(\mu_{ij})$ were of the form $\Phi(\mu_j + \boldsymbol{\varphi}_4'(j)\bar{\mathbf{s}}_i)$, where μ_j is the probit for interval j, $\bar{\mathbf{s}}_i$ is the 5×1 posterior mean vector of sire-specific regression coefficients, and $\boldsymbol{\varphi}_4(j)$ is a vector of the fourth-order Legendre polynomial evaluated at day *j*. EFD's were calculated for: 1) the total 330 day period EFD(-30,300); 2) from 30 days before to 30 days after calving, EFD(-30,30); 3) from 30 days before to 120 days after calving, EFD(-30,120); and 4) from 120 to 300 days after calving, EFD(120,300). The EFD's were compared to sire evaluations from two threshold model analyses of mastitis treated as a single binary trait. One evaluation was for the interval from 30 days before to 300 days after calving (P300), and the second was from 15 days before to 120 days after calving (P120). These evaluations had the form $\Phi(\mu + \bar{\mathbf{s}}_i)$, where $\bar{\mathbf{s}}_i$ is the posterior mean of sire *i* and μ is the probit corresponding to the overall mean incidence of mastitis.

Assuming that "transmitting ability at day *j*" is $\boldsymbol{\varphi}_4'(j)\mathbf{s}_i$, the between-sire variance of liability to CM at day *j* is $\sigma_s^2(j) = \boldsymbol{\varphi}_4'(j)\mathbf{G}\boldsymbol{\varphi}_4(j)$, where $\boldsymbol{\varphi}_4'(j)$ is a 5×1 vector with elements of the fourth-order Legendre polynomial evaluated at day *j*, and **G** is the matrix of the estimated sire (co)variances between the regression coefficients. The estimated heritability of liability to CM at any day *j* of lactation is: $h^2(j) = 4\sigma_s^2(j) / [1 + \sigma_{pe}^2 + \sigma_s^2(j)]$.

RESULTS AND DISCUSSION

The posterior means (SD) of the herd and permanent environmental variances were 0.06 (0.004) and 0.11 (0.008), respectively. Heritability of liability to CM ranged between 0.07 and 0.13 before calving, from 0.04 to 0.15 the first 270 days after calving, and increased sharply thereafter. The trajectory of heritability along lactation was in agreement with Chang et al. (2001) who analyzed the same data set with an 11-variate threshold model, assuming that mastitis was a different trait in each period. The estimated probability of CM along lactation, for the five highest ranking and the five lowest ranking sires based on EFD(-30,300), is shown in Figure 1. The "best" sires have a much lower peak around calving and a lower probability of CM throughout lactation than the "worst" sires.

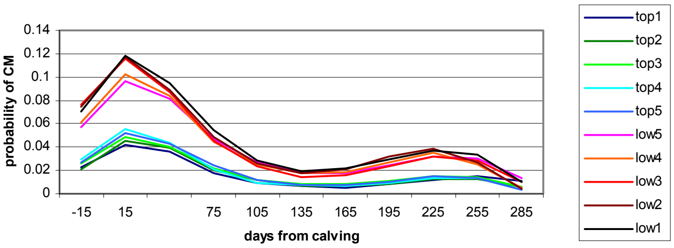

Figure 1. Probability of clinical mastitis along first lactation for the five highest ranking and the five lowest ranking sires based on the expected fraction of days without mastitis from 30 days before to 300 days after calving, EFD(-30,300)

Rank correlations between EFD's from the longitudinal model and sire evaluations from the two cross-sectional models (P120 and P300) are shown in Table 1. In general, rank correlations between all sire evaluations that addressed the first part of lactation were high (≥0.89). Rank correlations between EFD(120,300) and other evaluations were lower (0.64-0.87), as expected, since EFD(120,300) makes use of the regression curve for the last part of lactation only. Although, the rank correlations were high, this does not imply that the same sires would be selected, especially if selection intensity is high. For example, if 10 of the 245 sires were to be selected, only 4, 5 or 7 of them would be in common if selection were based on EFD(-30,300) versus EFD(-30,30), P120 or P300, respectively.

Table 1. Rank correlations between sire evaluations from the longitudinal model, expected fraction of days without mastitis (EFD) for four time intervals, and the two cross-sectional models (P120 and P300)

	EFD(-30,30)	EFD(-30,120)	EFD(120,300)	P120	P300
EFD(-30,300)	0.94	0.97	0.87	0.89	0.96
EFD(-30,30)		0.98	0.68	0.94	0.93
EFD(-30,120)			0.74	0.93	0.95
EFD(120,300)				0.64	0.82
P120					0.93

An advantage of using a longitudinal model for CM is the ability to take multiple treatments and time aspects into account, as well as to account for environmental effects peculiar to each test-interval. Cows that are culled before the end of lactation do not cause data sampling biases in a longitudinal model, since both "incomplete lactations" and records in progress can be included. In this study, and for illustration, we used the expected fraction of days without mastitis for sire ranking. However, other criteria may be of interest. For example, more weight could be placed on mastitis in early lactation (since these cases may result in higher costs) by computing the probability of no mastitis for given intervals for each sire, and then weighting the information in some manner.

ACKNOWLEDGMENTS

The Norwegian Dairy Herd Recording (Husdyrkontrollen) is acknowledged for providing data, and GENO Breeding and A. I. Association for furnishing pedigree information on sires. This work is part of the "Healthy Cow" project financed by the Research Council of Norway. Support has also been received from the Babcock Institute for International Dairy Research and Development, University of Wisconsin-Madison, by NFS grant DEB-00.89742, and by grant NRICGP/USDA 99-35205-8162.

REFERENCES

Chang, Y.M., Rekaya, R., Gianola, D., Heringstad, B. and Klemetsdal, G. (2001) *J. Dairy Sci.* **84** (Suppl.1) : 110.

Gianola, D. (1982) *J. Anim. Sci.* **54** : 1079-1096.

Heringstad, B., Klemetsdal, G. and Ruane, J. (2000) *Livest. Prod. Sci.* **64** : 95-106.

Jensen, J. (2001) *J. Dairy Sci.* **84** : 2803-2812.

Rekaya, R., Rodriguez-Zas, S.L., Gianola, D. and Shook G.E. (1998) *Book of Abstr.* 49th EAAP **4** : 44.

Sorensen, D.A., Andersen, S., Gianola, D. and Korsgaard, I. (1995) *Genet. Sel. Evol.* **27** : 229-249.

Swalve, H.H. (2000) *J. Dairy Sci.* **83** : 1115-1124.