GENETIC ANALYSIS OF MASTITIS AND NUMBER OF SERVICES TO CONCEPTION IN NORWEGIAN RED USING A CENSORED THRESHOLD MODEL

B. Heringstad1,2, I.M. Andersen-Ranberg2, Y.M. Chang3, and D. Gianola1,3

1Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, P.O. Box 5003, N-1432 Ås, Norway.
2Geno Breeding and A. I. Association, P.O. Box 5003, N-1432 Ås, Norway.
3Department of Dairy Science, University of Wisconsin, Madison 53706, USA.

INTRODUCTION

In genetic evaluation, both clinical mastitis (CM) and female fertility are often treated as binary traits. For mastitis, the binary variable takes value 1 if a cow has at least one CM case within a given time interval, and 0 otherwise. Likewise, the fertility trait non-return rate takes value 0 if a cow has a second insemination within, e.g., 56 d after a first one, and 1 otherwise. With these definitions, cows with one or more cases of mastitis are treated as equally liable to the disease, whereas cows with one or multiple returns after insemination are scored as equally fertile. An alternative is to use number of CM cases (NCM) and number of services to conception (STC) as response variables. There are few studies on genetic associations between health and fertility in dairy cattle, and estimates of the genetic correlation between NCM and STC are not available in the literature. A reason why these traits are not used more widely is the problem of censoring, as cows with mastitis and/or fertility problems are more likely to be culled than healthy, fertile, cows. The risk of culling increases as NCM or STC increase.

Chang et al. (2006) developed an ordinal threshold model that takes censoring into account. Our objective was to extend this model to a bivariate setting for two ordered categorical traits, and to infer genetic parameters for NCM and STC for Norwegian Red (NRF).

MATERIAL AND METHODS

Data. The dataset had records from 620,492 first-lactations daughters of 3,064 NRF sires. Only first batch daughters (i.e., birth year of daughter - birth year of sire < 6 years) were included. First calving was from 1980 through 2004, and age at first calving was between 20 and 40 months. The dataset was restricted further to include only cows from herd-5-year classes with at least 5 uncensored first lactation cows, and from sires with at least 20 uncensored daughters. For each cow, CM cases from 30 days before first calving to either culling, 300 days after first calving, or second calving, whichever occurred first, were included. Cases were considered to be distinct only if the interval between treatments was longer than 5 days, to avoid multiple counting of the same mastitis episode. NCM was defined as the number of CM treatments during 1st lactation, and scored into 3 categories: 0, 1, and 2 or more cases of CM. During 1st lactation 81 % of the cows had no CM, 16 % had 1 case, and 3 % had 2 or more cases. If a cow was culled before 300 d after first calving, her record was considered censored; overall censoring rate for NCM was 30 %. Mean NCM per cow was higher for censored cows (0.27 vs. 0.21). A total of 460,445 (74%) of the cows had at least 1 insemination. Double insemination was defined as a new insemination 0–5 days after a first one. For each cow, all services in first lactation (other than double inseminations) were counted. STC had 5 categories, with a mean of 1.6; 61 % of the cows had 1 service, and 26%, 9% and 3% had 2, 3, and 4 services, respectively, and only 1% of the cows had 5 or more services. Cows without record of a second calving were considered censored for STC (16%). Mean STC was 1.5 for uncensored and 1.8 for censored cows. The pedigree file had 3756 sires, including the 3,064 with daughters in the dataset.
Model. A bivariate ordinal threshold-liability model (e.g., Gianola, 1982) was used for analysis of NCM and STC. For each trait the threshold model postulates an underlying continuous variable, liability (λ), such that the observed category is j if $T_{j-1} \leq \lambda < T_j$, where $T_{j-1}$ and $T_j$ are thresholds, and $j = 1, 2, \ldots, J$ indexes the category to which the observation belongs. The thresholds satisfy $\infty = T_0 \leq T_1 \leq T_2 \leq \cdots \leq T_J = \infty$. The first threshold $T_1$ is set to zero, because the parameter cannot be identified in a probit analysis. If an observation in category j is censored, its corresponding liability must be larger than $T_{j-1}$. The model fitted allowed for censoring of either trait. In matrix notation the model is: $\lambda = X\beta + Z_h h + Z_s s + e$, where $\lambda$ is a vector of unobserved liabilities for NCM and STC; $\beta$ is a vector including age at first calving (21 levels) and month-year of first calving (288 levels) effects; $h$ is a vector of herd-5-year period of calving effects (51,808 levels); $s$ is a vector of sire transmitting abilities; $e$ is a vector of residuals for the 2 traits; and $X$, $Z_h$, and $Z_s$ are appropriate incidence matrices. Both residual variances were set equal to 1. Residuals were assumed to be independent between cows and correlated between records of the same cow.

A Bayesian approach using MCMC methods (Sorensen and Gianola, 2002), as applied by Chang et al. (2006), was used. Independent proper uniform priors were assigned to each of the elements of $\beta$. Multivariate normal prior distributions were assigned to herd-5-year effects: $h \sim N(0, H \otimes I)$, and to sire transmitting abilities: $s \sim N(0, G \otimes A)$. Inverse Wishart prior distributions were used for the (co)variance matrices of herd-5-year ($H$) and sire ($G$) effects, while the residual covariance was assigned an uniform prior bounded between –1 and 1. Thresholds were assumed to be distributed as ordered statistics from a uniform distribution, where $0 \leq T_1 \leq \cdots \leq T_J$. Draws from posterior distributions, except those of the thresholds, were obtained using a Gibbs sampler (Sorensen and Gianola, 2002). A Metropolis algorithm was used for sampling thresholds, as described by Chang et al. (2006). Inferences were based on 90,000 samples, after the first 10,000 were discarded as burn-in.

To evaluate the effect of taking censoring into account, the model was also run ignoring censoring, i.e., treating censored observations as uncensored. For comparison purposes, the data were analyzed also with a linear sire model (same explanatory structure) and standard REML-BLUP procedures, using the DMU-package (Madsen and Jensen, 2005), ignoring censoring and the categorical nature of the traits.

RESULTS AND DISCUSSION

Posterior mean and standard deviation (SD) of heritability of liability to NCM and to STC, and correlations between the two traits are given in Table 1. All posterior distributions were sharp and symmetric, as illustrated for heritability and genetic correlation in Figure 1. The posterior means (SD) of heritability of liability to NCM and to STC were 0.08 (0.004) and 0.029 (0.002), respectively (Table 1). The heritability of liability to NCM is within the range of estimates from cross-sectional threshold model analyses of first lactation mastitis treated as a single binary trait, which range from 0.06 to 0.12 (Lund et al., 1999; Heringstad et al., 2003; Zwald et al., 2004). Heritability of liability to binary CM in the same dataset (unpublished) was slightly lower (0.07 vs. 0.08); the small difference is not surprising, since 97% of the records were in the 0 and 1 classes. Lund et al. (1999) and Sørensen et al. (2000) used linear model analyses, and found that the heritability of counts of clinical mastitis cases was not higher than that of clinical mastitis treated as a binary trait.
Table 1. Posterior mean (standard deviation) of heritabilities ($h^2 = 4\sigma^2_s/(\sigma^2_s + \sigma^2_h + 1)$), and genetic, herd-5-year, and residual correlations between NCM and STC, from censored threshold model (TM) and TM ignoring censoring, together with linear model estimates.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Censored TM</th>
<th>TM ignoring censoring</th>
<th>Linear model estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heritability NCM</td>
<td>0.08 (0.004)</td>
<td>0.07 (0.003)</td>
<td>0.03</td>
</tr>
<tr>
<td>Heritability STC</td>
<td>0.03 (0.002)</td>
<td>0.03 (0.002)</td>
<td>0.02</td>
</tr>
<tr>
<td>Genetic correlation</td>
<td>0.21 (0.039)</td>
<td>0.09 (0.040)</td>
<td>0.10</td>
</tr>
<tr>
<td>Herd-5-year correlation</td>
<td>0.17 (0.014)</td>
<td>0.15 (0.014)</td>
<td>0.15</td>
</tr>
<tr>
<td>Residual correlation</td>
<td>0.05 (0.002)</td>
<td>-0.01 (0.002)</td>
<td>-0.01</td>
</tr>
</tbody>
</table>

The point estimate of heritability of liability to STC was in agreement with other threshold model estimates (Chang et al., 2006; González-Recio, et al., 2005), which range between 0.03 and 0.05.

Figure 1. Posterior distributions of heritability of liability to STC, heritability of liability to NCM, and of the genetic correlation between NCM and STC from threshold models taking censoring into account (CTM) or ignoring censoring (TMic)

The posterior distribution of the genetic correlation between NCM and STC from the censored threshold model was sharp and symmetric (Figure 1), with posterior mean (SD) 0.21 (0.039). The correlation between herd-5-year effects was positive (0.17) while the residual correlation between the two traits was close to zero (Table 1). This suggests that herds with high NCM tend to have higher STC.

Table 1 shows that heritability estimates from the threshold models (liability scale) were higher than corresponding linear model estimates. This is expected, and in agreement with previous studies (e.g., Heringstad et al., 2003). Relative to a threshold model ignoring censoring, in which censored records were treated as “complete” (Table 1 and Figure 1), the censored threshold model produced higher heritability estimates for both traits, and a larger genetic correlation (0.21 vs. 0.09). Residual correlation was close to zero, and the herd-5-year correlation was similar in all 3 models. The correlation (rank correlation) between sire evaluations from the censored threshold model and from the linear model was 0.90 (0.90) for NCM and 0.87 (0.86) for STC.

The results indicate that censoring affects parameter estimation as well as genetic evaluation and ranking of sires. Genetic (co)variance between NCM and STC seems to be understated in models ignoring censoring. With a model that takes censoring into account such as ours, NCM and STC may become more appealing candidate traits for genetic evaluation of mastitis and fertility.
ACKNOWLEDGMENTS
Access to the data was given by the Norwegian Dairy Herd Recording System and the Norwegian Cattle Health Service in agreement number 004.2005. This work is part of project no. 167893/110 financed by the Research Council of Norway. Support was also received from the Babcock Institute for International Dairy Research and Development, University of Wisconsin, Madison.

REFERENCES