Genetic Analyses of Mastitis Data Using Animal Threshold Models and Genetic Correlation with Production Traits

D. Hinrichs, E. Stamer, W. Junge, and E. Kalm

Institute of Animal Breeding and Husbandry, Christian-Albrechts-University, D-24098 Kiel, Germany

ABSTRACT

In the present study, 6 different mastitis data sets of 3 dairy herds with an overall herd size of 3200 German Holstein cows were analyzed. Data collection periods included the first 50, 100, or 300 d of lactation. The 3 data collection periods were analyzed with a lactation model and a test-day model. All models were animal threshold models. Mastitis frequencies in the lactation model data sets varied between 29 and 45%, and varied between 3 and 6% in the test-day model data sets. Depending on the period of data collection, heritabilities of liability to mastitis in the lactation models were 0.05 (50 d), 0.06 (100 d), and 0.07 (300 d). In the test-day models, heritabilities were slightly higher with values of 0.09 (50 and 100 d), and 0.06 (300 d). Between lactation models, the rank correlations between the relative breeding values were high and varied between 0.86 and 0.94. Rank correlations between the relative breeding values of the test-day models ranged from 0.68 to 0.87. The rank correlations between the relative breeding values of lactation models and test-day models varied from 0.51 and 0.80. Genetic correlations between mastitis and milk production traits were estimated with a linear animal test-day model. The correlations with mastitis were 0.29 (milk yield), 0.30 (fat yield), 0.20 (fat content), 0.34 (protein yield), and 0.20 (protein content). The estimated genetic correlation between mastitis and somatic cell score was 0.84.

(Key words: mastitis, lactation model, test-day model, animal threshold model)

INTRODUCTION

Mastitis is the most frequent and expensive disease in dairy cattle. Costs due to mastitis include veterinary and treatment costs, reduced milk production during the remaining part of lactation, the loss of milk that has to be discarded due to contamination with antibiotics,

early culling, decreased milk quality, and increased disease risk in the future (Heringstad et al., 2000).

The heritability of mastitis has been estimated in several studies. The most common approach when utilizing mastitis data in genetic evaluation is to consider mastitis as a binominal trait and to apply linear models, which assume a normal distribution of the data (Heringstad et al., 2000). Heritabilities from these models are low, with most values appearing in the interval 0.02 to 0.03 (Lund et al., 1994; Sander Nielsen et al., 1996; Heringstad et al., 2001a). Threshold models are an alternative to the traditional linear models, as they take the binary nature of the data into account. The estimates of heritabilities increase when threshold models are used and range from 0.06 to 0.12 (Simianer et al., 1991; Heringstad et al., 2001b; 2003a). It should be noted that the definition of mastitis as a binominal trait does not fully use all information provided by the data, because some cows have more than one case of mastitis. Therefore, alternative modeling is an important area of research. Rekaya et al. (1998, 2003) suggested the development of test-day models for longitudinal binary response for the analysis of mastitis field data. The estimates for mastitis heritabilities of those models ranged from 0.05 to 0.11 (Rekaya et al., 1998; Kadarmideen et al., 2000; Heringstad et al., 2003b; Chang et al., 2004).

The unfavorable genetic correlation between mastitis and milk yield is generally accepted. The average genetic correlation between mastitis and milk yield was 0.43 (Sander Nielsen et al., 1996; Luttinen and Juga, 1997; Lund et al., 1999).

In several countries, SCC is used as an indirect indicator for udder health. Heritabilities of SCC vary from 0.08 to 0.19 (Lund et al., 1994; Schutz et al., 1994; Reents et al., 1995). The efficiency of SCC as a selection criterion to reduce the frequency of mastitis depends on the genetic correlation between these 2 traits (Heringstad et al., 2000). Mrode and Swanson (1996) concluded that the average genetic correlation between mastitis and SCC is roughly 0.70. If milk samples for SCC are collected monthly, mastitis cases with rapid recovery are not detected by SCC because only 10 to 20% of environmental mastitis infections are detected by monthly collected SCC (Shook and Schutz, 1994).

Received February 3, 2004.

Accepted February 13, 2005.

Corresponding author: D. Hinrichs; e-mail: dhinrichs@tierzucht.uni-kiel.de.

Table 1. Length of data collection, mastitis frequency (%), number of observations, cows, and sires for different data sets.¹

Date set	Model	Data recording period	Mastitis frequency	Observations	Cows	Sires
1	Lactation model	50 d	29.1	11,317	6492	181
2	Lactation model	100 d	34.4	11,317	6492	181
3	Lactation model	300 d	45.4	11,317	6492	181
4	Test-day model	50 d	5.9	503,829	6564	193
5	Test-day model	100 d	4.1	963,055	6653	193
6	Test-day model	300 d	2.6	2,539,621	6804	193

¹In data sets 1 to 3, mastitis frequency is the percentage of cows with mastitis. In data sets 4 to 6, mastitis frequency is the percentage of test-days on which the cow had mastitis.

The aim of the present study was the estimation of heritabilities of liability to mastitis. Three different data collection periods (50, 100, or 300 d of lactation) and 2 models (lactation animal threshold models and test-day animal threshold models) were compared. Finally, genetic correlations between mastitis and different traits of milk yield and somatic cell score are presented. The estimated genetic correlations are based on a linear animal model.

MATERIALS AND METHODS

The data were collected on 3 large commercial dairy farms with a total herd size of 3200 German Holstein cows in the years 1998 to 2001. All medical treatments by veterinary or farm staff were recorded and allocated to a disease class. Available disease classes were udder diseases, fertility diseases, metabolic diseases, claw and leg diseases, and other. In the disease class "udder diseases," most treatments (98%) were associated with mastitis. This resulted in 32,107 treatments for mastitis. Six data sets were compared. The first 3 data sets contained only one observation per cow per lactation, and the data collecting period ended at 50, 100, or 300 d in lactation, resulting in data sets 1, 2, and 3, respectively. Each observation was allocated a mastitis code, "1" if a cow showed mastitis during the observed period and "0" otherwise. The remaining 3 data sets contained one observation for every test-day. Data collecting periods were the same as in the first 3 data sets, resulting in data sets 4, 5, and 6, respectively. If there was a recorded treatment, the day of treatment and the following 5 d were coded with "1". All other days were coded with "0". The 5-d period was chosen because it was the average time in which milk has to be discarded. Only records from sires with at least 10 daughters were considered. This was done to reduce computation time. Table 1 shows the number of observations, cows, and sires for the different data sets. The number of cows in the data sets for test-day models was higher because it was not necessary that all animals began with lacta-

tion d 1 for this analysis. Therefore, animals that were between lactation d 2 and 300 at the beginning of the sampling period could be kept in the analysis. This resulted in data sets with different number of animals (see Table 1). In the data sets 1, 2, and 3 culled cows are represented by the information collected before the cows were culled and the liability values were only sampled for this time. In the data sets 4, 5, and 6, all cows with at least one test-day were included.

In a first step, all data sets were checked for extreme categories. Classes with all responses equal to "0" or "1" were excluded from the analysis. They were not combined with other classes. Thereafter, the significance of systematic environmental effects was investigated in a fixed-effect analysis, ignoring random permanent environmental and genetic effects. In the lactation models, the lactation number and the herd-year-season of calving were considered as systematic environmental effects. Test-day models considered herd-week of observation, herd-year-season of calving, and age of calving as systematic environmental effects. In the lactation models, lactation numbers above 3 were summarized in one class. The age of calving in the test-day models includes the lactation number, with lactation number 1 and 2 each divided into 3 different classes: 1 (20 to 24 mo), 2 (25 to 29 mo), and 3 (30 to 49 mo) for lactation 1, and 4 (30 to 39 mo), 5 (40 to 45 mo), and 6 (46 to 56 mo) for lactation 2, respectively. Age of calving class 7 (45 to 75 mo) was equal to lactation 3; and lactation 4 and higher lactations were summarized to age of calving class 8 (≥ 75 mo). Herd-year-season of calving classes were equal for the lactation models and the test-day models. Every year was divided into 3 seasons of calving. A multicode was then created, resulting in 36 different herd-year-seasons of calving in the lactation models, and 42 different herd-year-seasons of calving in the test-day models. In the lactation models, 6 herd-year-seasons of calving had to be excluded because they were extreme categories. The test-day models included 612 different herd-week of observation classes. The fixed

effect of herd-week of observation had no extreme categories. This was caused by the average large herd size.

The analysis of the fixed effects was done with a GLM procedure with probit link function using the GENMOD procedure of the SAS package (SAS Institute, 1999). Statistical testing for each factor was done by removing this particular factor from the model and comparing the likelihood of the reduced model with the likelihood of the full model. A threshold liability model (Wright, 1934; Dempster and Lerner, 1950; Gianola and Foulley, 1983) was used for the estimation of random permanent environmental variance, additive genetic variance, and breeding values.

The following single trait repeatability lactation model was used for analyses of data sets 1 to 3:

$$E[\pi_{ijkl}] = \Phi(K_i + L_j + \text{DIM} + (\text{DIM})^2 + m_k + t_l)$$

where $E[\pi_{ijkl}]$ = expected percentage of animals with mastitis; Φ = cumulative probability function of the standard normal distribution; K_i = effect of the i th herd-year-season of calving ($i = 1, \dots, 36$); L_j = effect of the j th lactation ($j = 1, \dots, 4$); DIM = linear regression on DIM ; $(\text{DIM})^2$ = quadratic regression on DIM ; m_k = permanent environmental effects ($k = 1, \dots, 6492$); and t_l = additive genetic effect of the l th animal ($l = 1, \dots, 15,612$).

In the lactation model, DIM are the days from calving to the end of sampling period or culling. Cows that were culled before the end of the sampling period were included in the data sets. This was possible because the model contained 2 regression coefficients on the length of the sampling period. Furthermore, it should be noted that the single trait repeatability lactation model assumes that mastitis is the same trait in all lactations.

The following test-day repeatability model was applied to the data sets 4 to 6:

$$E[\pi_{ijklm}] = \Phi(B_i + K_j A_k + f_{k(\text{DIM})} m_l + t_m)$$

where $E[\pi_{ijklm}]$ = expected percentage of animals with mastitis; Φ = cumulative probability function of the standard normal distribution; B_i = effect of the i th herd-week of observation ($i = 1, \dots, 612$); K_j = effect of the j th herd-year-season of calving ($j = 1, \dots, 42$); A_k = effect of the k th age of calving ($k = 1, \dots, 8$); $f_{k(\text{DIM})}$ = lactation curve nested within age of calving k ; m_l = permanent environmental effects ($l = 1, \dots, 6804$); and t_m = effect of the m th animal ($m = 1, \dots, 15,612$).

The number of permanent environmental effects was 6564, 6653, and 6804 in data sets 4, 5, and 6, respectively. It should be noted that the test-day repeatability model assumes that mastitis is the same trait each day of lactation in all lactations. A pedigree file was created

by tracing the pedigree of cows with data 3 generations back. The resulting file had 15,612 animals.

In the test-day model, $f_{k(\text{DIM})}$ is a fixed effect for the first 7 d of lactation. The decision to treat the first 7 d of lactation differently was done according to the study of Schomaker et al. (2002). After the seventh day of lactation, $f_{k(\text{DIM})}$ is identical to the function used by Ali and Schaeffer (1987). One of the main reasons for the use of the Ali and Schaeffer function was the flexibility of this function. Therefore, decreasing and increasing parts of a trajectory could be described by one function. This function is also used for the official SCS breeding value estimation in Germany. Furthermore, Kadarmideen et al. (2000) examined the effect of using the Ali-Schaeffer (Ali and Schaeffer, 1987), Wood (Wood, 1967), and Morant and Gnanasakthy (Morant and Gnanasakthy, 1989) curves on variance component estimates for mastitis, and they found very similar results for all 3 functions. More recently, Chang et al. (2004) compared different random regression functions for the analysis of clinical mastitis using longitudinal threshold models. These authors concluded that the Legendre polynomial was best supported by their data.

The posterior distribution of the permanent environmental variance and the additive genetic variance for liability to mastitis were determined with the Gibbs sampling algorithm implemented in the LMMG_TH program, a threshold derivative of LMMG (Reinsch, 1996). The LMMG_TH program is based on data augmentation (Sorensen et al., 1995). For random effects (permanent environmental, animal effect), a multivariate normal prior distribution with zero mean and appropriate variance-covariance matrices were used. Improper flat priors were assumed for "fixed effects" (herd-year-season of calving, lactation, age of calving, herd-week of observation) as well as for variance components.

For all data sets, 100,000 iterations were generated and the results of each iteration were retained. The first 25,000 iterations were discarded as burn-in, and inferences were based on the remaining 75,000 iterations. The convergence was determined by visual inspection of plots of realized parameter values against iteration number.

Breeding values were estimated during each iteration and the given breeding values are the mean breeding value over all iterations. They were mainly used for model comparison. Another possibility of model comparison is the calculation of the Bayes factor. This method was used by Rekaya et al. (2003). Rank correlations among the breeding values of the different data sets were estimated as Spearman correlation coefficients, using the CORR procedure of SAS (SAS Institute, 1999). For the estimation of rank correlations only,

Table 2. Number of observations (N), mean, standard deviation (SD), minimum, and maximum of milk production traits.¹

Trait	N	Mean	SD	Minimum	Maximum
Somatic cell score	74,056	3.00	1.73	-2.60	9.60
Milk yield (kg)	74,206	28.00	8.45	0.60	82.40
Fat content (%)	74,051	4.22	0.80	0.60	10.00
Fat yield (kg)	74,051	1.15	0.33	0.10	4.80
Protein content (%)	74,053	3.46	0.37	1.54	9.30
Protein yield (kg)	74,053	0.95	0.25	0.10	2.80

¹Based on data set 6.

the 70 sires with at least 30 daughters with mastitis information were used. Furthermore, breeding values were multiplied by -1 so that bulls with daughters with less mastitis than average received higher breeding values. Relative breeding values were calculated as:

$$\text{RBV} = \frac{[(\text{breeding value/additive genetic standard deviation}) \times 12] + 100}{100}$$

For estimation of the genetic correlations, results of milk performance tests were combined with data set 6. On all 3 farms, milk performance was recorded monthly. This resulted in 74,206 available milk yield records and 74,051, 74,053, and 74,056 records for fat (content and yield), protein (content and yield), and SCS, respectively (Table 2). The average milk yield in the performance test was 28.00 kg per cow with 4.22% fat and 3.46% protein. This resulted in an average fat and protein yield of 1.15 and 0.95 kg per cow, respectively. In addition, the mean SCS was 3.00 (Table 2).

Bivariate animal repeatability test-day models were used for estimation of genetic correlations between mastitis and milk production traits. The following linear model was used for all traits:

$$y_{ijklmn} = \mu + B_i + K_j + A_k + f_{k(\text{DIM})} + m_l + t_m + e_{ijklmn}$$

where y_{ijklmn} = production observation of cow m , in herd week i , with herd-year-season of calving j , and age of calving k ; μ = overall mean; B_i = fixed effect of the i th herd-week of observation ($i = 1, \dots, 612$); K_j = fixed effect of the j th herd-year-season of calving ($j = 1, \dots, 42$); A_k = fixed effect of the k th age of calving ($k = 1, \dots, 8$); $f_{k(\text{DIM})}$ = lactation curve nested within age of calving k ; m_l = random permanent environmental effects ($l = 1, \dots, 6804$); t_m = random effect of the m th animal ($m = 1, \dots, 15,612$); and e_{ijklmn} = random residual effect.

Genetic correlations between mastitis and milk production traits were estimated with the program VCE 4.0, version 4.2.5 (Groeneveld, 1998).

RESULTS

The mastitis frequencies of the different data sets are given in Table 1. During the first DIM, mastitis frequency increased rapidly. Maximum mastitis frequency was 19.7% after 7 DIM. Mastitis frequency depends on the data-recording period (Table 1). In the lactation model (data sets 1 to 3), mastitis frequency varied between 29.1 and 45.4%, and in the test-day model (data sets 4 to 6), from 2.6 to 5.9%, respectively. The frequencies from data set 1 to 3 are equal to the percentage of cows with at least one case of mastitis during the data collecting period. In data sets 4 to 6,

Figure 1. Trajectory of mastitis during the first 300 d of lactation.

Table 3. Number of observations, lactation length, standard deviation (SD), and mastitis frequency in different lactations for the lactation model and the test-day model.¹

Lactation number	Observations	Mean lactation length (d)	SD	Mastitis frequency (%)
Lactation model				
1	4885	142	87	42.5
2	2889	141	85	45.6
3	1891	139	85	49.2
>3	1652	135	85	48.9
Test-day model				
1	1,077,685	142	87	2.1
2	712,812	141	85	2.5
3	426,586	139	85	3.1
>3	322,538	135	85	3.3

¹In data sets 1 to 3, mastitis frequency is the percentage of cows with mastitis. In data sets 4 to 6, mastitis frequency is the percentage of test-days on which the cow had mastitis.

the mastitis frequency is equal to the percentage of test days coded as mastitis days. Figure 1 suggests that most cases of mastitis occurred during the first 30 d of lactation. Table 3 gives a further description of the data, and contains the distribution of records over lactation, the mean number of days in each lactation, and its standard deviation, and the disease frequencies in different lactations.

Posterior distributions of the heritability of liability to mastitis are given in Figures 2 and 3. The posterior distributions of the heritabilities of liability to mastitis estimated with the lactation model (data sets 1 to 3) were similar for the 3 data sets (Figure 2). The posterior distributions of the heritability for liability to mastitis estimated with data sets 4 and 5 (test-day model) were also similar, whereas the posterior distribution of the heritability of data set 6 had a different shape (more peaked) and a lower mean, but it was similar to the

posterior distributions of data sets 1 to 3 (Figures 2 and 3).

Posterior means and the standard deviations of the variance components, heritabilities, and repeatabilities are given in Table 4. In data sets 1 to 3 (lactation model), the posterior mean of the permanent environmental variance was equal for all data sets ($\sigma_{pe}^2 = 0.09$), and the posterior mean of the additive genetic variance increased from $\sigma_a^2 = 0.05$ to $\sigma_a^2 = 0.08$ with increasing sampling period. In data sets 4 to 6 (test-day model), the posterior mean of the permanent environmental variance decreased from $\sigma_{pe}^2 = 2.51$ to $\sigma_{pe}^2 = 1.12$, and the posterior mean of the additive genetic variance decreased from $\sigma_a^2 = 0.35$ to $\sigma_a^2 = 0.15$ with the increasing data collection period. This led to slightly higher posterior means for the heritability of liability to mastitis and significantly higher posterior means for the repeat-

Figure 2. Posterior distributions of heritability of liability to mastitis in data sets 1 to 3 (data recording period in parentheses) for the lactation model.

Figure 3. Posterior distributions of heritability of liability to mastitis in data sets 4 to 6 (data recording period in parentheses) for the test-day model.

ability in the data sets 4 to 5 (test-day model), compared with data sets 1 to 3 (lactation model). This could also be seen in Table 4. The posterior standard deviations of the heritabilities for liability to mastitis and of the repeatabilities are low (0.01 to 0.02).

Spearman correlation coefficients between the breeding values of different data sets are given in Table 5 and vary between $r_s = 0.51$ and $r_s = 0.94$. All correlation coefficients were significant ($P < 0.0001$). Within-model correlations, i.e., the correlations between data sets 1 to 3 (lactation model) and between data sets 4 to 6 (test-day model) were higher for the lactation models ($r_s = 0.86$ to 0.94), compared with the test-day models ($r_s = 0.68$ to 0.87). Between lactation models and test-day models, e.g., data set 1 and data set 4, rank correlations varied between $r_s = 0.51$ and $r_s = 0.80$.

The results of the relative breeding value calculations are given in Table 6. The mean of the relative breeding value is 100 and the standard deviation is 12. As expected, the average relative breeding value for all data sets varied between 100 and 102. Minimum relative breeding values were similar for all data sets and varied between 86 and 95. Maximum relative breeding values

ranged from 106 to 125. The difference between maximum and minimum breeding values was higher in data sets 4 to 6 (test-day model) and ranged from 28 to 39, compared with data sets 1 to 3 (lactation model), where the differences varied from 17 and 21 (Table 6).

The heritability for liability to mastitis estimated with a linear model was 0.01. Genetic correlations between mastitis and milk production traits and heritabilities of milk production traits are in Table 7. Highest genetic correlation was estimated between mastitis and SCS ($r_g = 0.84$). Furthermore, all milk production traits have an unfavorable genetic correlation with mastitis, varying between $r_g = 0.20$ and $r_g = 0.34$. Heritabilities for the milk production traits ranged from $h^2 = 0.16$ to $h^2 = 0.45$, and heritability for SCS was $h^2 = 0.16$. The standard errors of the estimated heritabilities are low ($\pm sh^2 = 0.01$). Standard errors for the genetic correlations were larger and ranged from $\pm srg = 0.01$ to $srg = 0.08$ (Table 7).

DISCUSSION

Mastitis is one of the greatest problems affecting commercial milk production. In the present study, 29.1 to

Table 4. Posterior means and standard deviation (SD) of additive genetic (σ_a^2) and permanent environmental (σ_{pe}^2) variance, heritabilities (h^2), and repeatabilities (r) for different data sets.¹

Date set	Model	Data collection period	σ_a^2	σ_{pe}^2	h^2	r
1	Lactation model	50 d	0.05 (0.02)	0.09 (0.03)	0.05 (0.02)	0.13 (0.02)
2	Lactation model	100 d	0.07 (0.02)	0.09 (0.03)	0.06 (0.02)	0.14 (0.02)
3	Lactation model	300 d	0.08 (0.02)	0.09 (0.03)	0.07 (0.02)	0.14 (0.02)
4	Test-day model	50 d	0.35 (0.07)	2.51 (0.11)	0.09 (0.02)	0.74 (0.01)
5	Test-day model	100 d	0.26 (0.04)	1.80 (0.07)	0.09 (0.01)	0.67 (0.01)
6	Test-day model	300 d	0.15 (0.02)	1.12 (0.03)	0.06 (0.01)	0.56 (0.01)

¹In data sets 1 to 3, the mastitis frequency is equal to the percentage of cows with mastitis; in data sets 4 to 6, the mastitis frequency is equal to the percentage of mastitis days.

Table 5. Rank correlations (r_s) between sire evaluations from different data sets and models (includes only sires with at least 30 daughters).^{1,2}

	Data set 2	Data set 3	Data set 4	Data set 5	Data set 6
Data set 1	0.93	0.86	0.73	0.79	0.80
Data set 2		0.94	0.57	0.72	0.78
Data set 3			0.51	0.68	0.79
Data set 4				0.86	0.68
Data set 5					0.87

¹Data sets 1 to 3 = lactation model; data sets 4 to 6 = test-day model.

²Data sets 1 and 4 = 50 d; data sets 2 and 5 = 100 d; data sets 3 and 6 = 300 d.

45.4% of the cows were affected by at least one case of mastitis (lactation model, data sets 1 to 3), and 2.6 to 5.9% of the test days in data sets 4 to 6 (test-day model) were mastitis days (Table 1). Mastitis frequencies in this study were high compared with literature reports by Heringstad et al. (1999, 2001b, 2003a) (17.0 to 30.0%), but they are in agreement with the mastitis frequency given by Fournichon et al. (2001) (29.0 to 51.0%). It should be noted that the results of Heringstad et al. (1999, 2001a, 2003a,b) were based only on observations from first lactation. However, mastitis frequency in the first lactation in this study was higher than that given by Heringstad et al. (1999, 2001a, 2003a,b). Mastitis frequency increased with increasing lactation number (Table 3). That the mastitis frequency did not increase more in higher lactations could be related to the large herds. In these herds, it may be more likely that cows with mastitis are culled in higher lactations.

Trajectory of mastitis frequency during lactation (Figure 1) was similar to that reported by Heringstad et al. (2003b). Figure 1 shows that most cases of mastitis occur during the first 30 d of lactation when the physiological demands of the cow are high.

Posterior distributions of heritabilities for liability to mastitis for the data sets that were analyzed with the lactation model (Figure 2) were similar to those given

Table 7. Genetic correlations (r_g) and their standard errors (SE) between mastitis and milk production traits and heritabilities (h^2) of milk production traits.¹

Trait	h^2	r_g
Somatic cell score	0.16 (0.01)	0.84 (0.02)
Milk yield (kg)	0.36 (0.01)	0.29 (0.08)
Fat content (%)	0.38 (0.01)	0.20 (0.08)
Fat yield (kg)	0.34 (0.01)	0.30 (0.06)
Protein content (%)	0.45 (0.01)	0.20 (0.01)
Protein yield (kg)	0.33 (0.01)	0.34 (0.08)

¹Based on data set 6.

by Heringstad et al. (2003a), who found that the distributions were fairly constant with increasing data collection period.

The heritabilities for data sets 1 to 3 (lactation model) were between $h^2 = 0.05$ and $h^2 = 0.07$ and in agreement with literature reports given by Heringstad et al. (2001b, 2003a), but they were slightly lower than the heritabilities given by Lund et al. (1999).

Additive genetic variance and permanent environmental variance in the data sets 4 to 6 (test-day model) were higher compared with the data sets 1 to 3 (lactation model; Table 4), because data sets 4 to 6 used all available mastitis information. Therefore, repeated cases of mastitis were distinct observations in these data sets. It should be noted that additive genetic variance and permanent environmental variance decreased with an increasing period of data collecting. The decreased additive genetic variance suggests that mastitis in earlier phases of lactation (when the animals are facing high physiological demands) was more related to the cows' genetic character, whereas mastitis in the later phases of lactation was more affected by random, genetically unrelated events. Kadarmideen et al. (2000) estimated a sire variance of 0.04. Chang et al. (2004) estimated a similar sire variance for the first 200 d of lactation. These authors used a fourth-order Legendre polynomial in their model, instead of the function of Ali and Schaeffer (Ali and Schaeffer, 1987). Furthermore,

Table 6. Mean, standard deviation (SD), minimum, maximum, and range of relative breeding values (includes only sires with at least 30 daughters).¹

Data set	Model	Data recording period	Mean	SD	Minimum	Maximum	Max – Min
1	Lactation model	50 d	100	3.51	90	111	21
2	Lactation model	100 d	100	3.35	88	108	20
3	Lactation model	300 d	100	3.12	89	106	17
4	Test-day model	50 d	102	4.97	95	123	28
5	Test-day model	100 d	102	5.11	89	124	35
6	Test-day model	300 d	100	5.19	86	125	39

¹In data sets 1 to 3, the mastitis frequency is equal to the percentage of cows with mastitis; in data sets 4 to 6, the mastitis frequency is equal to the percentage of mastitis days.

Chang et al. (2004) observed an increasing sire variance at the end of lactation.

It should be noted that in this study, an animal threshold repeatability model was used and that mastitis was regarded as the same trait at each day in all lactations. In this study, no problems were observed in using an animal model. In a preliminary analysis (results not shown), results were checked by a second analysis with a sire model (70,000 iterations) and the results from both models agreed well. Another point that should be discussed is the assumption that mastitis is the same trait at each day of lactation in all lactations, because Heringstad et al. (2003b) estimated different genetic correlations between different stages of lactation. These authors find genetic correlations between different stages of lactation, which varies between -0.19 and 0.98 . However, in the analyses of Rekaya et al. (1998) and Kadarmideen et al. (2000), mastitis was treated as one trait in different stages of lactation. This approach was also taken in the current study.

The permanent environmental variance estimated with the data sets 4 to 6 (test-day model) in this study were large compared with that of Kadarmideen et al. (2000). Reasons for the higher permanent environmental variance could be seen in different models (e.g., Kadarmideen et al. (2000) divided the lactation into 10 periods of 30 d each). Heritabilities estimated with data sets 4 to 6 ($h^2 = 0.06$ to $h^2 = 0.09$) are in agreement with heritabilities estimated by Chang et al. (2004), Heringstad et al. (2003b), and Rekaya et al. (1998). The heritability estimated by Kadarmideen et al. (2000) was higher, but repeatabilities (0.56 to 0.74) estimated in this study were higher compared with Kadarmideen et al. (2000).

Nielsen et al. (1999) estimated the genetic correlations between mastitis in the first 3 lactations in Danish Holsteins. These authors found a genetic correlation for mastitis in different lactations between 0.85 and 0.98 .

Rank correlations between the relative breeding values of the data sets analyzed with a lactation model (data sets 1 to 3), shown in Table 4, were high ($r_s = 0.86$ to $r_s = 0.84$) and similar to those estimated by Heringstad et al. (2001b). The rank correlations between the relative breeding values estimated with data sets 4 to 6 (test-day model) were slightly lower ($r_s = 0.68$ to $r_s = 0.87$).

The results of the relative breeding value calculation (Table 6) showed that data sets should be analyzed with a test-day model. Maximum relative breeding values are higher (123 to 125) compared with the data sets analyzed with a lactation model (106 to 111). This resulted in greater differences between maximum and minimum relative breeding values, because the mini-

um relative breeding values were similar for all data sets. Therefore, more possibilities for selection decision were offered if relative breeding values for a given data set were estimated with a test-day model. Standard deviations of relative breeding values were higher (5.0 to 5.2) if data sets were analyzed with a test-day model. Relative breeding values for mastitis, estimated by Philipsson et al. (1995), ranged from 76 to 114 for Swedish Red and White or 71 to 115 for Swedish Friesian, respectively. It should be noted that the results of Philipsson et al. (1995) were based on information from 1462 different sires with 460,938 effective daughters for Swedish Red and White, and 911 different sires with 241,600 effective daughters for Swedish Friesian, respectively. The standard deviations of the relative breeding values of the data sets analyzed with a test-day model were similar to those given by Philipsson et al. (1995) and minimum and maximum relative breeding values were higher.

Estimated heritability of SCS ($h^2 = 0.16$) in this study was in agreement with literature reports given by Reents et al. (1995), Schutz et al. (1994), and Lund et al. (1994). Genetic correlation between mastitis and SCS ($r_g = 0.84$) is lower compared with Philipsson et al. (1995) and Mrode and Swanson (1996), but similar to that of Lund et al. (1994). The genetic correlation estimated in this study showed that mastitis and SCS are not the same trait and that mastitis had an unfavorable genetic correlation to milk, fat, and protein yield as well as fat and protein content (Table 7). The genetic correlation between mastitis and milk yield was in agreement with those reported by Emanuelsson (1988), but was below the average genetic correlation of 0.43 estimated by other authors (Sander Nielsen et al., 1996; Luttinen and Juga, 1997; Lund et al., 1999). There was also an unfavorable genetic correlation between mastitis and fat yield and fat content ($r_g = 0.20$ to $r_g = 0.30$), so that an increased fat yield would lead to increased mastitis. The genetic correlation between mastitis and protein yield was higher compared with Heringstad et al. (1999), and lower than that estimated by Lund et al. (1999).

CONCLUSIONS

Mastitis is still one of the major problems and causes of considerable economic loss in the area of commercial milk production. Test-day models should be used for the variance component and breeding value estimation, because repeated cases of mastitis are distinct events in these models. The use of test-day models led to higher additive genetic and permanent environmental variance. This resulted in slightly higher heritabilities and significantly higher repeatabilities in test-day models,

compared with those generated in lactation models. Recording of data should be concentrated within the first 100 d of lactation, because most cases of mastitis occur during this period.

Genetic correlation between mastitis and SCS was moderately high and suggested that mastitis and SCS are not the same parameters of udder health. In addition, the unfavorable genetic correlation between mastitis and milk production traits would lead to a further decrease in udder health if mastitis were ignored in commercial breeding programs. Establishment of data recording schemes for mastitis in addition to SCS would improve the possibility for genetic gain of udder health (e.g., mastitis resistance).

ACKNOWLEDGMENTS

This project was supported by the Federal Ministry of Consumer Protection, Food, and Agriculture, Berlin, and NOG Nord Ost Genetic GmbH & Co. Kg, Verden, Germany. Special thanks to Jörn Bennewitz for helpful discussion. Both reviewers chosen by the journal are acknowledged for helpful comments on earlier versions of the manuscript.

REFERENCES

- Ali, T. E., and L. R. Schaeffer. 1987. Accounting for covariances among test day milk yields in dairy cows. *Can. J. Anim. Sci.* 67:637–644.
- Chang, Y. M., D. Gianola, B. Heringstad, and G. Klemetsdal. 2004. Longitudinal analysis of clinical mastitis at different stages of lactation in Norwegian Cattle. *Livest. Prod. Sci.* 88:251–261.
- Dempster, E. R., and I. M. Lerner. 1950. Heritability of threshold characters. *Genetics* 35:212–235.
- Emanuelsson, U. 1988. Recording of production disease in cattle and possibilities for genetic improvements: A review. *Livest. Prod. Sci.* 20:89–106.
- Fournichon, C., H. Seggers, F. Beaudeau, L. Verfaillie, and N. Bareille. 2001. Health-control costs in dairy farming systems in western France. *Livest. Prod. Sci.* 68:141–156.
- Gianola, D., and J. L. Foulley. 1983. Sire evaluation for ordered categorical data with a threshold model. *Genet. Sel. Evol.* 15:201–224.
- Groeneveld, E. 1998. VCE User's guide. Version 4.2. Institute of Animal Husbandry and Animal Sciences, Mariensee, Germany.
- Heringstad, B., Y. M. Chang, D. Gianola, and G. Klemetsdal. 2003b. Genetic analysis of longitudinal trajectory of clinical mastitis in first-lactation Norwegian Cattle. *J. Dairy Sci.* 86:2676–2683.
- Heringstad, B., G. Klemetsdal, and J. Ruane. 1999. Clinical mastitis in Norwegian cattle: Frequency, variance components, and genetic correlation with protein yield. *J. Dairy Sci.* 82:1325–1330.
- Heringstad, B., G. Klemetsdal, and J. Ruane. 2000. Selection for mastitis resistance in dairy cattle—A review with focus on the situation in the Nordic countries. *Livest. Prod. Sci.* 64:95–106.
- Heringstad, B., G. Klemetsdal, and J. Ruane. 2001a. Variance components of clinical mastitis in dairy cattle—effects of trait definition and culling. *Livest. Prod. Sci.* 67:265–272.
- Heringstad, B., R. Rekaya, D. Gianola, G. Klemetsdal, and K. A. Weigel. 2001b. Bayesian analysis of liability of clinical mastitis in Norwegian Cattle with a threshold model: Effects of data sampling method and model specification. *J. Dairy Sci.* 84:2337–2346.
- Heringstad, B., R. Rekaya, D. Gianola, G. Klemetsdal, and K. A. Weigel. 2003a. Bivariate analysis of liability to clinical mastitis and to culling in first-lactation cows. *J. Dairy Sci.* 86:653–660.
- Kadarmideen, H. N., R. Rekaya, and D. Gianola. 2000. Genetic analysis of repeated mastitis episodes with Bayesian longitudinal threshold models. Page 78 in *Proc. 51th Annu. Mtg. of EAAP*, The Hague, Netherlands.
- Lund, M. S., J. Jensen, and P. H. Petersen. 1999. Estimation of genetic and phenotypic parameters for clinical mastitis, somatic cell production deviance, and protein yield in dairy cattle using Gibbs sampling. *J. Dairy Sci.* 82:1045–1051.
- Lund, T., F. Miglior, J. C. M. Dekkers, and E. B. Burnside. 1994. Genetic relationship between clinical mastitis, somatic cell count, and udder conformation in Danish Holsteins. *Livest. Prod. Sci.* 39:243–251.
- Luttinen, A., and J. Juga. 1997. Genetic relationship between milk yield, somatic cell count mastitis, milkability and leakage in Finnish dairy cattle; health. *Interbull Bull.* 15:78–83.
- Morant, S. V., and A. Gnanasakthy. 1989. A new approach to the mathematical formulation of lactation curves. *Anim. Prod.* 49:151–162.
- Mrode, R. A., and G. J. T. Swanson. 1996. Genetic and statistical properties of somatic cell count and its suitability as an indirect means of reducing the incidence of mastitis in dairy cattle. *Anim. Breed. Abstr.* 64:847–857.
- Nielsen, U. S., G. A. Pedersen, J. Pedersen, and J. Jensen. 1999. Genetic variation in disease traits and their relationships with survival in Danish dairy cattle. *Interbull Bull.* 21:170–177.
- Philipsson, J., G. Ral, and B. Berglund. 1995. Somatic cell count as a selection criterion for mastitis resistance in dairy cattle. *Livest. Prod. Sci.* 41:195–200.
- Reents, R., J. Jamrozik, L. R. Schaeffer, and J. C. M. Dekkers. 1995. Estimation of genetic parameters for test day records of somatic cell score. *J. Dairy Sci.* 78:2847–2857.
- Reinsch, N. 1996. Two Fortran programs for the Gibbs Sampler in univariate linear mixed models. *Arch. Tierzucht* 39:203–209.
- Rekaya, R., D. Gianola, K. Weigel, and G. Shook. 2003. Longitudinal random effects models for genetic analysis of binary data with application to mastitis in dairy cattle. *Genet. Sel. Evol.* 35:457–468.
- Rekaya, R., S. L. Rodriguez-Zas, D. Gianola, and G. E. Shook. 1998. Test-day models for longitudinal binary response: An application to mastitis in Holsteins. Page 44 in *Proc. 49th Annu. Mtg. of EAAP*, Warsaw, Poland.
- Sander Nielsen, U., G. A. Pedersen, J. Pedersen, and J. Jensen. 1996. Genetic parameters for mastitis, other diseases and somatic cell count in different parities in Danish dairy breeds. Page 10 in *47th Annu. Mtg. of EAAP*, Lillehammer, Norway.
- SAS Institute. 1999. SAS/STAT User's Guide Version 8. SAS Institute, Inc., Cary, NC.
- Schomaker, T., W. Junge, E. Stamer, N. Reinsch, and E. Kalm. 2002. Varianzkomponentenschätzung für die Merkmale Mastitis und Zellgehalt (SCS) und Vergleich verschiedener Auswertungsmodelle. *Züchtungskunde* 74:250–265.
- Schutz, M. M., P. M. VanRaden, and G. R. Wiggans. 1994. Genetic variation in lactation means of somatic cell scores for six breeds of dairy cattle. *J. Dairy Sci.* 77:284–293.
- Shook, G. E., and M. M. Schutz. 1994. Selection on somatic cell score to improve mastitis in the United States. *J. Dairy Sci.* 77:648–658.
- Simianer, H., H. Solbu, and L. R. Schaeffer. 1991. Estimated genetic correlations between disease and yield traits in dairy cattle. *J. Dairy Sci.* 74:4358–4365.
- Sorensen, D. A., S. Andersen, D. Gianola, and I. Korsgard. 1995. Bayesian inference in threshold models using Gibbs sampling. *Genet. Sel. Evol.* 27:229–249.
- Wood, P. D. P. 1967. Algebraic model of the lactation curve in cattle. *Nature* 216:164–165.
- Wright, S. 1934. An analysis of variability in number of digits in an inbred strain of Guinea pigs. *Genetics* 19:506–536.