Test Day Model for the Genetic Analysis of Udder Health Traits: Somatic Cell Score and Clinical Mastitis

E. Negussie, K. Vuori, I. Strandén and E. A. Mäntysaari
MTT Agrifood Research, Biotechnology and Food Research, Biometrical Genetics, 31600 Jokioinen, Finland

Abstract

Genetic associations between test-day (TD) somatic cell score (SCS) and two clinical mastitis (CM) traits (CM1: -7 to 30 and CM2: 31 to 300 DIM) were estimated using a random regression model (RRM) that combines information from two different data models. Multi-trait random regression models (mt-RRM) were then developed for the genetic evaluation of mastitis. Data from 1.65 million first-lactation Finnish Ayrshire cows were used. Heritabilities of CM1 and CM2 were 0.026 and 0.016, respectively, while for TD SCS they ranged from 0.06 to 0.11. During first lactation, genetic correlations between TD SCS and CM1 and between TD SCS and CM2 varied from 0.40 to 0.77 and from 0.34 to 0.71, respectively. In genetic evaluation of mastitis, the comparison between mt-RRM and multi-trait lactation average models (mt-LAM) showed that mt-RRM is better than mt-LAM and has added advantages of making efficient use of available TD SCS information and offers proofs for bulls and cows.

1. Introduction

Clinical mastitis evaluation benefits from multi-trait analyses with a correlated trait such as SCS and conformation traits (Negussie *et al.,* 2006a). The efficiency of these traits in increasing the accuracy of CM evaluation depends on the genetic and environmental associations between the traits, and wide range of values have been cited in the literature.

So far most studies reported on genetic correlations between CM and SCS, and between CM and other traits have generally been from lactation average models. A genetic analysis based on lactation average model does not utilize all information in the data, as it does not allow simultaneous estimation of stage of lactation effects (Ødegård *et al.,* 2003; Negussie *et al.,* 2006b). Moreover, since SCS vary with stage of lactation and with a test-day, the genetic association between CM traits and SCS may also differ during lactation.

Currently, most mastitis evaluations are based on multi-trait sire models and lactation average records. With a lactation average sire model, available information on udder health traits, particularly information on test-day SCS may not be effectively utilised and only proofs for sires can be calculated. The objective of the present study was to estimate the genetic association between test-day SCS and CM traits during lactation and subsequently to develop genetic evaluation model that combine information from both traits using random regression model.

2. Material and methods

Data were from the Finnish animal health and production database. Records of CM and test-day SCS from 1.6 million first-lactation Finnish Ayrshire cows with first calving from 1988 onwards were used. SCS was expressed as $\log_e$-transformed somatic cell count ($\log_e$SCC) from bi-monthly test days measured in 1000cells/ml. All cases of veterinary treated clinical mastitis in first lactation from early (CM1: -7 to 30 DIM) and late stages of lactation (CM2: 31 to 300 DIM) were considered. Within these intervals, the absence or presence of mastitis was scored as “0” or “1”, respectively. Finally, information on CM cases extracted from the database was merged with TD SCS records for analyses. A subset of data, including cows with first calving from 1995 to 2000 and with an average of 5 cows in herd-3-year classes was sampled for estimation of covariance components and associated parameters.

2.1. Data Analysis

2.1.1 Covariance components

A multi-trait random regression model (mt-RRM) that combines information from two different data models was used for estimation of covariance components of test-day SCS, CM1 and CM2 traits. Initially, univariate RRM analyses of the TD SCS were made to determine appropriate order of polynomials that are needed to describe the variance structure in the data sufficiently. In addition, eigenvalues of the covariance matrices were analysed to assess the importance of adding further parameters. Consequently, in the mt-RRM, the additive genetic and permanent environmental effects for test-day SCS were modeled by second-order orthogonal Legendre polynomials. Whilst for CM, only the intercept term was fitted.

The general description of the mt-RRM including test-day SCS (S), CM1 (C1) and CM2 (C2) traits was:

$$\begin{bmatrix} y_{Shijklmno} \\ y_{C1hijkmno} \\ y_{C2hijkmno} \end{bmatrix} = \begin{bmatrix} f_{Si} \\ f_{C1i} \\ f_{C2i} \end{bmatrix} + \begin{bmatrix} ym_{Sjk} \\ ym_{C1jk} \\ ym_{C2jk} \end{bmatrix} + \begin{bmatrix} hy_{Shj} \\ hy_{C1hj} \\ hy_{C2hj} \end{bmatrix} + $$

$$+ \begin{bmatrix} \Sigma_{r=0}^{4} \phi_{\pi}(d)_r b_{Slr} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} htd_{Sho} \\ 0 \\ 0 \end{bmatrix} +$$

$$\begin{bmatrix} \Sigma_{r=0}^{2} \phi_{\alpha}(d)_r p_{Smr} \\ p_{C1m} \\ p_{C2m} \end{bmatrix} + \begin{bmatrix} \Sigma_{r=0}^{2} \phi_{\alpha}(d)_r a_{Snr} \\ a_{C1n} \\ a_{C2n} \end{bmatrix} + \begin{bmatrix} e_{Shijklmno} \\ e_{C1hijkmn} \\ e_{C2hijkmn} \end{bmatrix}$$

where $y_{Shijklmno}$, $y_{C1hijkmn}$ and $y_{C2hijkmn}$ are test-day SCS(S), CM1(C1) and CM2(C2) observations, respectively, recorded in herd *h*, on TD *o*, in month *k*, of the year *j*, on a cow *m* belonging to the calving age class *i*, calving-year×calving-month class *l* and measured on DIM *d*. Fixed effects were age at calving (f), calving-year×calving-month (ym), herd×3-year period (hy) and regression coefficients (b) describing the shape of the lactation curve within calving-year×calving-season classes. he modeling of fixed effects was the same for all traits with the exception of the lactation curve, which was modeled only for test-day SCS.

The covariables for coefficients $b_{..r}$ (r=0,…,4) were:

$$\phi_{\pi}(d) = [c_0\ c_1\ c_2\ c_3\ \exp(wd)]^T,$$

where c_0 c_1 c_2 c_3 represent coefficients of the third-order orthogonal Legendre polynomial at DIM d and w =-0.09 is coefficient of the exponential term of the Wilmink function (Wilmink, 1987). The calving seasons were October to February, March to June, and July to September. The herd effect was modeled by a fixed herd-3-year and a random herd-test-day (*htd*) effects. The number of *hy* and random *htd* classes are in Table 1.

Table 1. Description of the alternative analyses.

	mt-RRM (millions)	mt-LAM (millions)
No. animals	1.62	1.57
No. observations	7.61	1.62
Htd	1.92	---
Hy	0.33	0.33
No. animal equations	10.62	6.37
No. total equations	20.85	6.76

Random genetic effects were a_{Snr}, a_{C1n} and a_{C2n}. The a_{Snr} had random regression genetic effects for test-day SCS with coefficients from a second-order orthogonal Legendre polynomial at DIM *d* as in $\phi_{\pi}(d)$. Random effects p_{Smr} were non-genetic animal effects for a cow *m* with $\phi_{\pi}(d)$ as in above for test-day SCS; and p_{C1m} and p_{C2m} were for CM1 and CM2, respectively. Random $e_{Shijklmno}$, $e_{Chijkmn}$ and $e_{Chijkmn}$ were measurement errors.

In order to facilitate accurate estimation, residual covariances between CM traits and test-day SCS were assumed to be zero, because daily residuals between test-day SCS and CM traits can not be estimated. In addition, with the mt-RRM it was only possible to estimate permanent environmental variance for the longitudinal trait. To ameliorate this, the residual variance of CM traits was set to operationally low value (about 10%) so that part of this variance entered the permanent environmental component. This facilitated estimation of a permanent environmental correlation between CM and the longitudinal trait. The resulting covariance components of the random regression coefficients for additive genetic and permanent environmental effects were then used for estimation of the necessary parameters. The covariance components were estimated using DMU package (Madsen and Jensen, 2006).

2.1.2 Genetic evaluation

Multi-trait sire and animal RRMs were developed for the genetic evaluation of udder health traits. These models combine information from TD SCS with CM traits from early and late lactation stages (CM1 and CM2). Animal model parameters were derived from sire model estimates to calculate mt-RRM animal model EBVs for all animals. Estimates of breeding values from the mt-RRM were then compared to the corresponding mt-LAM. Model predictive abilities, correlation between and standard deviation of EBVs, and some computational aspects were assessed.

Parameters for mt-LAM BLUP analyses were derived from the mt-RRM estimates by summation over 305 days. For comparison, mt-LAM parameters were also directly estimated from the lactation average performance records. These estimates were in general found to be similar to those derived from the mt-RRM estimates. The system was solved by preconditioned conjugate gradient (PCG) method (Strandén and Lidauer, 1999), using a 2.6 GHz AMD Opteron CPU with 4 Gb of RAM. For each model, solving time and random access memory (RAM) requirements for solving the mixed model equations were monitored.

Finally, from mt-RRM, an animal gets 5 RRM breeding value coefficients from which 305-breeding values can be computed. Corresponding EBVs for TD SCS and CM1 and CM2 traits of animal *i* were calculated as: $EBV_{Si} = \sum_{d=8}^{312} \phi_{\alpha}^{T}(d)\hat{\mathbf{a}}_i$ and $EBV_{C1i}=\hat{\mathbf{a}}_i$ and $EBV_{C2i}=\hat{\mathbf{a}}_i$, respectively. Estimated breeding values of CM traits: CM1 and CM2 were combined into an index (CM_{com}) by giving equal weight to the traits.

3. Results and discussion

3.1 Genetic parameters

Heritabilities of CM1 and CM2 from the mt-RRM were 0.026 and 0.016, respectively (Table2).

Table 2. Estimated heritabilities (diagonal), genetic (below diagonal) and phenotypic correlations (above diagonal) for selected DIM of test-day SCS, CM1 and CM2 traits by mt-RRM in first lactation.

Traits		SCS							**CM1**	**CM2**
	DIM	30	60	110	160	210	260	310		
	30	**0.07**	0.62	0.58	0.53	0.48	0.43	0.38	0.02	0.19
	60	0.99	**0.08**	0.64	0.61	0.57	0.52	0.45	-0.01	0.20
	110	0.96	0.99	**0.09**	0.67	0.65	0.61	0.52	-0.03	0.20
SCS	160	0.93	0.97	0.98	**0.09**	0.68	0.65	0.56	-0.05	0.19
	210	0.90	0.94	0.97	0.99	**0.10**	0.68	0.60	-0.05	0.17
	260	0.85	0.90	0.94	0.97	0.98	**0.11**	0.64	-0.04	0.15
	310	0.78	0.83	0.87	0.91	0.95	0.98	**0.10**	-0.02	0.11
CM1		0.77	0.66	0.60	0.55	0.50	0.45	0.41	**0.026**	0.09
CM2		0.34	0.49	0.54	0.58	0.63	0.68	0.71	0.51	**0.016**

The estimates fall within the range of most reported values (0.02-0.03) from analyses with traditional linear models based on data from the Nordic health-recording systems (Heringstad *et al.,* 2000; Negussie *et al.,* 2006a).

Heritability of test-day SCS during first-lactation ranged from 0.06 to 0.11 (Table 2). The estimates were slightly lower in early lactation and increased gradually towards the late part of mid lactation. A possible explanation could be a large environmental variation during the early stages of lactation, or a low genetic variance. The estimates are in line with earlier studies (Koivula *et al.,* 2005; Negussie *et al.,* 2006b).

During lactation, genetic correlations between test-day SCS and CM1 and between test-day SCS and CM2 ranged from 0.40 to 0.77 and from 0.34 to 0.71, respectively (Table 2). The difference in the genetic association between TD SCS and CM traits during the different stages of lactation (early vs. late) suggest that the two CM traits measure different aspects of mastitis. Hence, combining information from both sources is essential in the genetic evaluation of animals for mastitis resistance. The practical implication of this study is therefore the development of a test-day evaluation model that combines test-day SCS information with CM traits from different stages of lactation. This will lead to a better use of udder health information. In addition, the test-day model allows the calculation of different selection criteria, which enables the testing of young bulls at an early age offering an early prediction of animals genetic merit.

3.2 Genetic evaluations

Standard deviations (SD) of EBVs for CM_{com} from mt-RRM and mt-LAMs are in Table 3. For the different groups of animals, the SD of EBVs from the mt-RRM was higher than that from mt-LAM. This could be explained by better utilization of information by the test-day model which in turn revealed more genetic variation. The increase in the SD of EBVs was relatively higher for old bulls and cows than for young bulls and cows.

Table 3. Standard deviations (SD) of EBVs for CM traits (CM_{com}) † from mt-RRM and mt-LAM for different groups of bulls: old bulls (born from 1992-94, with at least 100 daughters), young bulls (born from 1996-98, with at least 50 daughters), and old and young Ayrshire cows (born in 1995 and 1999, respectively).

Groups		Bulls/cows No.	MODEL	
			mt-RRM	mt-LAM
Bulls				
	Old	437	0.042	0.030
	Young	441	0.041	0.029
Cows				
	Old	88474	0.031	0.019
	Young	80857	0.029	0.018

† Values for CM_{com} are based on combining EBVs for CM1 and CM2

Correlations between EBVs from mt-RRM and mt-LAM were assessed for different groups of bulls. For TD SCS, correlations between EBVs were ~0.98 for older group of sires and ranged from 0.95 to 0.97 for young cows and bulls, respectively. Correlations between EBVs for CM_{com} were also higher for older groups of bulls (~0.91) than for young cows and bulls (0.82 – 0.89). One of the reasons for this could be the use of test-day model (test-day SCS), which allows better modelling of the herd environment and thereby improves the accuracy of young cow and bulls EBVs. Older bulls with large numbers of daughters, however, receive relatively accurate EBVs from both models. Thus, the apparent advantage of mt-RRM comes from better evaluation of cows and young bulls with less numbers of daughters.

As a consequence, some changes would be expected in the ranking of young bulls and cows.

The predictive ability of models for CM_{com} was assessed using data splitting method. In this analysis, a slightly higher correlation between EBVs from split data sets was found for mt-RRM (0.73) than for mt-LAM (0.71) indicating better model prediction performance.

Table 4. Number of total equations (N_{eq}), iterations until convergence (N_{conv}), solving time[†] and RAM requirements (Mb) for the multi-trait random regression (mt-RRM) and lactation average models (mt-LAM)

Model	N_{eq} (mill.	N_{conv}	Solving time (Min.	RAM (Mb)
Mt-LAM (sire)	0.436	174	16	22
mt-LAM(animal)	6.76	1189	73	215
mt-RRM (sire)	10.31	502	48	323
mt-RRM(animal)	20.85	1300	205	644

[†] AMD Opteron CPU 2.6 Ghz running Linux

3.3 Computational aspects

Details of computation, i.e. solving time and random access memory (RAM) requirements for solving mixed model equations for the mt-RRM and mt-LAMs are in Table 4. The shortest and longest computing time was required by mt-LAM sire and by mt-RRM animal models, respectively. The mt-RRM animal model required 3.5 hrs of solving time, which is twice the computing time required for solving a corresponding mt-LAM (Table 4). The relatively slow convergence when solving RRM could probably be due to the complexity of the covariance matrices. Nevertheless, in view of fast computers and efficient algorithms this would not be prohibitive to the large-scale routine application of mt-RRM.

4. Conclusions

This study showed that genetic correlations between TD SCS and CM traits varied during lactation. In CM evaluation, the comparison between models showed that the mt-RRM is better than the corresponding mt-LAM. Moreover, mt-RRM has added advantages of a) making efficient use of available information on TD SCS and offers accurate evaluation; b) derivation of different selection criteria which would allow early testing of bulls and c) as a by-product of mt-RRM evaluation, herd-test day solutions can be utilized for herd management and decision-making purposes.

5. References

Heringstad, B., Klemetsdal, G. & Ruane, J. 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.

Koivula, M., Negussie, E. & Mäntysaari, E.A. 2004. Genetic parameters for test-day somatic cell score at different lactation stages of Finnish dairy cattle. *Livest. Prod. Sci. 90,* 145-157.

Madsen, P. & Jensen, J. 2006. *A Users's Guide to DMU. A package for analysing multivariate mixed models.* Version 6, release 4.6. Danish Institute of Agricultural Sciences, Denmark.

Negussie, E., Koivula, M. & Mäntysaari, E.A. 2006a. Genetic parameters and single- versus multi-trait evaluation of udder health traits in Finnish Ayrshire. *Acta. Agric. Scand. A. 56,* 73-82.

Negussie, E., Koivula, M., Mäntysaari, E. A. & Lidauer, M. 2006b. Genetic evaluation of somatic cell score in dairy cattle considering first and later lactations as two different but correlated traits. *J. Anim. Breed. Genet. 123,* 224-238.

Ødegård, J., Jensen, J., Klemetsdal, G., Madsen, P. & Heringstad, B. 2003. Genetic analysis of somatic cell score in Norwegian cattle using random regression test-day models. *J. Dairy Sci. 86,* 4103-4114.

Stranden, I. & Lidauer, M. 1999. Solving large mixed linear models using preconditioned conjugate gradient iteration. *J. Dairy Sci. 75,* 2017-2022.

Wilmink, J.B.M. 1987. Adjustment of test-day milk, fat, and protein yields for age, season, and stage of lactation. *Livest. Prod. Sci. 16,* 335-348.