56th Annual Meeting of the European Association for Animal Production, Uppsala, Sweden, June 5th – 8th, 2005, Commission on Cattle Production, Session CG2.5

Survival Analysis for Genetic Evaluation of Mastitis in Dairy Cattle: A Simulation study

E. Carlén[1*], U. Emanuelson[2], and E. Strandberg[1]

[1]Department of Animal Breeding and Genetics, and [2]Department of Clinical Sciences, Swedish University of Agricultural Sciences, SE-750 07 Uppsala, Sweden
Corresponding author, e-mail: Emma.Carlen@hgen.slu.se

Abbreviation key: EBV = estimated breeding value, **LM** = linear model, **SA** = survival analysis, **TBV** = true breeding value, **TFM** = time to first mastitis or censoring.

INTRODUCTION

Mastitis is a major problem in the dairy cattle industry since it causes great economic losses, risk for overuse of antibiotics, and reduced animal welfare. Because of the serious effects of this disease and its unfavorable genetic correlation with milk production (Emanuelson et al., 1988; Rauw et al., 1998; Heringstad et al., 2000; Carlén et al., 2004), many countries have implemented genetic evaluation for improved mastitis resistance during the last decades. Indirect measures correlated to mastitis, usually somatic cell counts in milk, are being used in most of these countries as clinical mastitis cases are not commonly recorded. However, veterinary treated cases of clinical mastitis are recorded in Sweden and other Nordic countries (Heringstad et al., 2000; Interbull, 2005), and is considered in other countries as well (e.g. Zwald et al., 2004). This information can be used as direct measures of the disease, but it is important to use the best available methodology and an appropriate trait definition in the genetic evaluation.

The approach currently used for genetic evaluation of clinical mastitis, is to apply a linear model (**LM**) to an all-or-none trait (Interbull, 2005). The approach is relatively straightforward, but it involves some obvious disadvantages. Mastitis is defined as a binary trait, distinguishing between cows with at least one case of mastitis (1) and cows without cases (0). This trait definition means that only the first case of mastitis is considered, and that there is no difference between cows with a case of mastitis early or late in lactation (equally "bad"). By excluding cases other than the first, and by ignoring the timing of the case or the period at risk, some of the available information is lost. In addition, with this methodology incomplete and ongoing records can not be treated properly, which might further limit the amount of information used or, even worse, introduce potential bias in the genetic evaluation. Loss of information occurs by treating cows culled before the end of the observation period as missing. However, if these cows instead are included in the analysis by treating them as healthy observations, which is currently being done e.g. in the Swedish genetic evaluation, bias might be introduced if the culling reason is correlated to mastitis. Moreover, such observations are not distinguished from cows that did not contract mastitis during the whole period. Another disadvantage with the traditional LM methodology for analysis of binary mastitis data is that the assumption of normally distributed observations is not fulfilled. A non-linear threshold model would be theoretically better. However, the problem with the inefficient use of available information and the risk for introducing bias due to handling of

incomplete and ongoing records is not expected to be solved by using threshold models instead of LM.

Some of the mentioned disadvantages connected to the traditional LM when mastitis is analyzed, are expected to be overcome when the method of survival analysis (**SA**) is used. Survival analysis is a statistical method for studying the occurrence and timing of specific events, where the analyzed response time equals the time elapsed from a starting point until the occurrence of the event of interest (Ducrocq, 1987). Observations where a competing event occurs before the event of interest can still be included in the analysis by treating them as censored. Another positive feature of SA is the possibility to include time-dependent covariables to model environmental effects, such as stage of lactation, more precisely. Within the field of genetic evaluation of dairy cattle, SA has been successfully used for traits with a longitudinal character such as longevity traits, for which many countries currently use this method in routine genetic evaluations (Interbull, 2005), and interval fertility traits, such as calving to last insemination (Schneider et al., 2005).

One advantage of SA for mastitis data is that more of the available information is used by including the timing of the case or the length of the opportunity period. Another advantage is that cows without cases are treated as censored observations and we only include the information that these cows did not contract mastitis until the time of censoring; after this point we have no more information. Censoring is a more proper way of treating incomplete and ongoing records, and it reduces the potential bias occurring with the traditional LM when cows culled before they got a chance to express mastitis are treated as healthy observations. The use of SA, or similar methodology, to analyze time to first mastitis (**TFM**) have also been reported (Saebø and Frigessi, 2004; Carlén et al., 2005; Saebø et al., 2005). In the study by Carlén et al. (2005), SA was shown to be an alternative for genetic evaluation of clinical mastitis since TFM in the field data analyzed with SA gave higher precision of estimated breeding values (**EBV**) than did a binary mastitis trait analyzed with LM. A simulation study where true breeding values (**TBV**) can be simulated and thereafter correlated with EBV from the different methods would give a complementary indication of the usefulness of SA for genetic evaluation of mastitis.

The objective of this study was therefore to investigate by simulation whether the trait time to first mastitis analyzed with survival analysis results in a more accurate genetic evaluation for mastitis resistance than the more commonly used linear model methodology where a binary mastitis trait is analyzed.

MATERIALS AND METHODS

Population Structure

Each replicate of simulated data consisted of 60,000 first parity cows, daughters of 400 unrelated sires and distributed over 1200 herds. The average daughter group size was 150, varying between 115 and 187 (SD 12.3), and the herd size was fixed to 50. In total 50 replicates were done and results presented are averages over these replicates.

Simulation Process

The purpose of the simulation was to create the possible event of a mastitis case within lactation for each cow. In order not to favor any of the statistical methods used for analyses, sire and cow breeding values for mastitis liability on the underlying scale were simulated rather than the actual traits to be used in analyses. The simulation process was a further development of the simulation by Schneider et al. (2005) where the reproductive cycle of first parity cows was simulated. The reason for including the reproductive cycle also in our simulation was to achieve the length of the calving interval alternatively the day of culling

because of fertility reasons. These events were connected to milk production of the cows such that a cow with higher production had more chances to become pregnant (a higher number of inseminations allowed), and if she reached either the maximum waiting period allowed or the maximum number of inseminations allowed without becoming pregnant, she was kept longer in the herd before she got culled for bad fertility.

Four traits were simulated: 305-d milk production (kg), interval between calving and first ovulation (days), conception liability and mean mastitis liability. The phenotypic mean values for milk production and interval between calving and first ovulation were 8000 kg (SD 1000) and 28 days (SD 15), respectively. Conception and mastitis were simulated as binary traits with underlying normally distributed liabilities for the respective event with phenotypic means of zero for both traits and standard deviations of 1 for conception and 0.6 for mastitis liability (i.e. ~N(0,1) and ~N(0,0.6), respectively). The defined threshold values, above which cows conceived or got mastitis, were set to zero (50 % conception rate) and 3.0 (about 0.12 % average mastitis risk per day), respectively. Note that the average risk to contract mastitis a given day is low but to this a daily random liability variation component was added, which is described below.

Genetic parameters used to simulate breeding values for the parents and Mendelian sampling terms are shown in Table 1. For milk production and the 2 fertility traits, parameters were identical to those in Schneider et al. (2005). For mastitis liability we assumed a higher heritability than literature estimates from LM, but in agreement with reported estimates from threshold models where, in similarity to our study, mastitis was considered on the underlying scale (Heringstad et al., 1997; Kadarmideen et al., 2000). Herd variance as proportion of the phenotypic variance was 9 % for fertility traits, 20 % for mastitis liability and 30 % for milk production.

Phenotypic values for the cows were created as the sum of the mean, the herd effect, animal breeding value and an environmental value for respective trait. Breeding values for the cows were composed of half the breeding value of the sire, half the breeding value of the dam and a Mendelian sampling term. In the case of mastitis liability the environmental value corresponded to a permanent cow effect, thus the phenotypic value corresponded to a mean liability over the lactation. The mastitis history for each cow was created by allowing the mean mastitis liability to vary from day to day by adding a daily random variation component sampled from a normal distribution with mean 0 and standard deviation 0.8 (~N(0,0.8)). Note that the 2 variances contributing to the mastitis liability sum to 1. In addition, an extra liability (0.3) of getting mastitis was added to the first 10 days in lactation. If the resulting value was above the defined threshold of 3.0 for a given day the cow was considered to have her first mastitis case at that day in lactation and the sampling for this cow was interrupted, otherwise the daily sampling continued from the day of calving (day 0) until the day of next calving or the day of culling.

A few percent of all cows (1.8 %) were culled within d 10 of lactation because of calving related reasons before they got a chance to contract mastitis. This early culling was simulated using a Weibull distribution. The rest of the cows where tested for heat after a voluntary waiting period of 8 weeks. Cows detected in heat (60 % heat detection rate) were inseminated. The first insemination day was allowed to vary between herds with a mean value of 56 days (SD 3). If conception was above the threshold the cow was considered pregnant and the length of the calving interval was created as: the number of days until the last insemination plus a gestation period with mean value of 280 days (SD 5). The gestation period was only allowed to vary between 265 and 295 days. Cows that did not become pregnant either because the maximum waiting period or the maximum number of inseminations allowed was reached were culled for fertility reasons. The number of inseminations allowed was connected both to herd and to the milk production of the cow in

relation to her herd mates. The day of culling was also connected to production of the cow and calculated as a mean value of 240 days (SD 15) plus 20 times the production deviation (in SD) of this cow from the average production.

Definition of Traits

Phenotypic observations of mastitis were defined both as a binary trait and as a longitudinal trait. The binary trait distinguished between cows with a mastitis case (1) and cows without cases (0) during the 150 first days of lactation. Cows without mastitis cases that were culled within the defined period were included in the analysis as healthy. This definition is similar to what is used in the Swedish national genetic evaluation where the same restricted period was introduced to reduce bias because of culling. However, in the routine genetic evaluation, veterinary treatments of clinical mastitis and culling because of mastitis are recorded, whereas in the simulation we consider the actual cases of mastitis. The longitudinal trait TFM was measured as the number of days from calving to either the day of first mastitis case (uncensored observation) or to the day of next calving or culling (censored observation). Cows could be culled early in lactation (within 10 d) because of calving-related reasons or later on because of infertility. Because cows culled are indicated by the censoring variable when TFM is analyzed, it is not necessary to use a restricted time period to reduce bias.

Statistical Analysis

The binary mastitis trait was analyzed with LM analysis, whereas TFM was analyzed with SA (Weibull proportional hazards model). The same effects (a mean and random herd and sire effects) were used in both models to allow for a better comparison between them. The mean in the Weibull model corresponds to an average hazard over time defined as $\lambda_0(t)$ the Weibull baseline hazard function $\left(\lambda\,\rho\left(\lambda t\right)^{\rho-1}\right)$ with scale parameter λ and shape parameter ρ. A value of $\rho < 1$ indicates that the hazard decreases with time, whereas $\rho > 1$ means that the hazard increases with time. The herd effect was assumed to be normally distributed in the LM, whereas in SA it was assumed to follow a log-gamma distribution and was integrated out from the joint posterior density. The sire effect was assumed to be normally distributed for both models. Sires were assumed to be unrelated.

To obtain REML estimates of the variance components and breeding value predictions the DMU package (Madsen and Jensen, 2000) was used for the LM and the heritability was calculated as:

$$h^2 = 4\sigma_s^2 \Big/ \left(\sigma_s^2 + \sigma_e^2\right) \qquad [1]$$

Survival Kit V3.12 (Ducrocq and Sölkner, 1998) was used to estimate the variance components for sire and herd and predict breeding values. The heritability was calculated as:

$$h_{equ}^2 = 4\sigma_s^2 \Big/ \left[\sigma_s^2 + \left(1/\left(1-c\right)\right)\right] \qquad [2]$$

where *c* is the proportion of censored records.

This derivation for the heritability on the original scale, which is not dependent on the Weibull parameters, was suggested by Yazdi et al. (2002) as the equivalent heritability. They showed very good agreement between accuracy and selection response calculated using h_{equ}^2 and observed accuracies and responses calculated from simulation. The term equivalent refers to the fact that the EBV of a sire with n daughters would get the same reliability as if it were

evaluated on a linear trait with this heritability. An increase in the proportion of uncensored records with time implies that the equivalent heritability increases with time until it reaches the theoretical heritability [$h^2 = 4\sigma_s^2 \big/ \left(\sigma_s^2 + 1\right)$] one would get in the total absence of censoring. Because the amount of censoring is accounted for in the definition of h_{equ}^2 , it is possible to directly compare this heritability estimate with the estimate from the LM analysis.

To test the adequacy of applying a Weibull proportional hazards model (fully parametric), a Cox proportional hazards model (semiparametric) was run for one of the replicates and a Kaplan-Meier curve was created to check whether data followed the Weibull distribution. The assumption was assessed graphically from a plot of logs of the baseline survivor function (S (t)), against logs of time (i.e., $\ln\left[-\ln S\left(t\right)\right]$ against $\ln t$). If the Weibull assumption holds, the resulting graph should be linear. The graph from the Cox analysis was approximately linear after 10 days into lactation. Prior to that time point the curve was nonlinear due to the higher frequency of mastitis around calving (results not shown). Based on this we decided to run another Weibull model, in addition to the one presented above, to account for the early high risk. This model was identical to the previous Weibull model except that stage of lactation was included as a time-dependent effect with changing value at day 10.

Comparison of Methods

The main approach for comparison of the 2 different methods was to calculate Pearson product-moment correlations (SAS, 1999) between the sire EBV from LM and SA, respectively, and the sire TBV for mastitis liability. In addition we compared the average true genetic merit for the best and worst 10 % of bulls ranked on EBV from the different methods and also the proportion of the best or worst 10 % of bulls based on TBV that were correctly identified to be in the best or worst 10 % based on EBV from the different methods. Further, the theoretical accuracy (r) in selection was calculated for the different methods according to:

$$r = \sqrt{\frac{n}{n+k}} \qquad [3]$$

where *n* corresponds to the average number of daughters per sire (n=150) and $k = \frac{4-h^2}{h^2}$

RESULTS AND DISCUSSION

Descriptive Statistics

The average frequency of the binary mastitis trait was 0.107 (SD 0.309). This figure is in good agreement with a recent estimate of the frequency of mastitis in field data of Swedish first parity cows where mastitis was defined in a similar way (Carlén et al., 2005). For SA, the average proportion of uncensored records, i.e. cows with mastitis, was 16.7 % (varying from 15.7 to 17.7). The average failure time, that is the number of days until the first case of mastitis, was 124 and the average censoring time was 347. The corresponding figures from the study by Carlén et al. (2005) where SA was used to analyze TFM in field data were 15 % uncensored records, and 123 and 364 days for average failure and censoring time, respectively.

Estimates of Parameters

The Weibull parameter ρ was estimated to 0.61 on average from the Weibull model without a fitted lactation stage effect. This indicates a decreasing risk of getting mastitis with

time within lactation and it reflects the higher risk simulated for the first 10 days of lactation. A high risk in the beginning of the lactation followed by a low and nearly constant risk in the remaining part of the lactation has been shown in the literature (Barkema et al., 1998; Heringstad et al., 1999; Carlén et al., 2004). In field data of TFM ρ was estimated to about 0.6 in first lactation and around 0.7 for second and third lactation (Carlén et al., 2005). From the Weibull model where a time-dependent effect of lactation stage was fitted, the value of ρ increased to 0.77 on average. The reason why ρ is getting closer to 1 is that the higher risk of getting mastitis in the beginning of lactation now is accounted for in the lactation stage effect instead of in the baseline hazard.

Heritabilities and variance components estimated from the different methods are presented in Table 2. The estimates of variance components from the 2 different Weibull models, with and without lactation stage effect, were nearly identical and only results from the Weibull model without an effect of lactation stage are therefore presented. The heritability estimate from SA (3.7 %) can be directly compared with the estimate from the LM (2.7 %), because the proportion of uncensored records is accounted for.

Comparison of Methods

Results of all approaches for comparing the methods were nearly identical for the 2 different Weibull models, with and without lactation stage effect. Results presented are from the Weibull model without an effect of lactation stage. Correlations between TBV for mastitis liability and EBV from the different methods are shown in Table 3. A 8 % higher correlation with TBV was found for breeding values predicted with SA (0.76) than for breeding values predicted with the LM (0.70)

The theoretical accuracy for LM (r = 0.71) and for SA (r = 0.76) was in very good agreement with the corresponding correlations obtained between true and estimated breeding values, which can be seen as the true accuracy. This was a proof that the simulation worked satisfactorily and for SA it verified the usefulness of the defined equivalent heritability.

The average TBV for mastitis liability of the best and worst 10 % of bulls ranked on EBV from the different methods are presented in Table 4. The best bulls ranked on SA had an average true genetic merit that was lower than the best bulls ranked on the LM, and the opposite was true for the worst bulls. Lower genetic merit implies less mastitis, thus a better mastitis resistance. The proportion of the best or worst 10 % of bulls based on TBV that were correctly identified to be in the best or worst 10 % based on EBV from the different methods can be seen in Table 5. Survival analysis ranked 51 % of the best bulls correctly compared to a correct ranking of 44 % for the LM. The difference was similar for the worst bulls. These results correspond to the results from both the correlations and the average true genetic merit, and indicate that SA was better to identify the truly best and worst bulls in regard to mastitis resistance.

Sires with a large proportion of daughters with mastitis get more accurate breeding values regardless of method used. For SA the proportion of censored daughters is accounted for in the calculation of the heritability, see formula [2], and therefore it will affect also the accuracy in selection. A larger proportion of daughters with mastitis (uncensored) will give a more accurate sire EBV. For a categorical trait, the heritability, and consequently the accuracy of EBV, will vary with the frequency. Sires with a high mastitis liability have a larger proportion of daughters with mastitis, i.e. closer to an even distribution between healthy and diseased daughters where the heritability and accuracy is the highest. The higher accuracy implies less regression towards a population average and this can be seen as higher absolute values for average TBV for the worst 10 % of bulls than for the best 10 % of bulls (Table 4). It is also shown as a higher proportion of correctly ranked sires among the worst sires than among the best sires (Table 5).

Results from all 3 approaches for comparison of methods are in agreement and indicate that, with the given trait definitions and data structure, SA is a better method than the traditional LM for accurately detecting bulls with the best genetic merit for mastitis liability. By selecting bulls based on EBV from SA a gain in genetic progress can therefore be expected. This confirmed the conclusion made from a previous study where field data were used to compare SA and LM for genetic evaluation of clinical mastitis (Carlén et al., 2005). In that study the accuracy in selection was about 3 and 25 % higher for first and later lactations, respectively, for TFM analyzed with SA than for a binary mastitis trait analyzed with LM. One advantage of SA is that more of the available information is used with the trait TFM, since cows with a case, and cows without cases, get different TFM values. The differences in results in this study could partly be explained by the differences in trait definition; TFM being a more continuously distributed trait and including cases after d 150 in lactation. This trait could however not be analyzed with the traditional LM methodology, since it is intimately linked to the possibility to handle censoring. Analyzing the TFM value alone without a censoring variable makes it impossible to distinguish cows that actually contracted mastitis from cows without mastitis cases but with a following calving or culling date. Another advantage of SA is that culled cows are treated properly, which reduces potential bias. With the traditional LM cows culled for other reasons than mastitis before reaching the end of the defined period can be included in the analysis as healthy observations. This might introduce a bias in the genetic evaluation if the culling reason is something correlated to mastitis. A re-run of the LM where cows without mastitis cases that were culled before lactation d 150 were excluded from the analyses did however not change the results from this simulation study. This was probably because only a small proportion of the cows were culled before lactation d 150: for calving-related reasons less than 2 % and for fertility reasons only a few cows.

One drawback with TFM analyzed with SA is that, in similarity to the binary mastitis trait, only the first case of mastitis within lactation is considered. Another drawback with SA is that multi-trait analysis is not easy, and for an efficient genetic evaluation of clinical mastitis data a multi-trait analysis together with somatic cell count would be desirable. There are however ways to get around this problem. In the French national genetic evaluation breeding values for direct longevity are estimated with SA and thereafter combined longevity, based on direct longevity and several other traits, is computed using multi-trait analysis (Interbull, 2005). Recent work has also shown that it is possible to analyze a survival trait together with a normally distributed continuous trait or a threshold trait using a Bayesian approach and applying Gibbs sampling (Damgaard, 2005).

CONCLUSIONS

The correlation between sire TBV for mastitis liability and sire EBV from SA was considerably higher than the corresponding correlation from LM. The increased precision of sire breeding values predicted by SA could be translated into a higher genetic progress and indicates that SA is a better method than LM for genetic evaluation of mastitis data.

ACKNOWLEDGEMENTS

The authors are very grateful to Maria del Pilar Schneider for the contribution to the simulation program and for assistance with the survival analysis.

REFERENCES

Barkema, H. W., Y. H. Schukken, T. J. G. M. Lam, M. L. Beiboer, H. Wilmink, G. Benedictus, and A. Brand. 1998. Incidence of clinical mastitis in dairy herds grouped in three categories by bulk milk somatic cell counts. J. Dairy Sci. 81:411-419.

Carlén, E., M. del P. Schneider, and E. Strandberg. 2005. Comparison between linear models and survival analysis for genetic evaluation of clinical mastitis in dairy cattle. J. Dairy Sci. 88:797-803.

Carlén, E., E. Strandberg, and A. Roth. 2004. Genetic parameters for clinical mastitis, somatic cell score, and production in the first three lactations of Swedish Holstein cows. J. Dairy Sci. 87:3062-3070.

Damgaard, L. H. 2005. Quantitative genetic analysis of survival, linear Gaussian and ordered categorical traits. Ph.D. Thesis, Dept. Large Anim. Sci., The Royal Vet. And Agric. Univ., Frederiksberg, Denmark.

Ducrocq, V. 1987. An analysis of length of productive life in dairy cattle. Ph.D. Diss., Cornell Univ., Ithaca, NY.

Ducrocq, V., and J. Sölkner. 1998. The Survival Kit – A Fortran package for the analysis of survival data. Proc. 6th World Congr. Genet. Appl. Livest. Prod., Armidale, Australia. 27:447-448.

Emanuelson, U., B. Danell, and J. Philipsson. 1988. Genetic parameters for clinical mastitis, somatic cell counts, and milk production estimated by multiple-trait restricted maximum likelihood. J. Dairy Sci. 71:467-476.

Heringstad, B., A. Karlsen, G. Klemetsdal, and J. Ruane. 1997. Preliminary results from a genetic analysis of clinical mastitis data. Proceedings International workshop on genetic improvement of functional traits in cattle; health. Uppsala, Sweden, June, 1997. Interbull Bull. 15:45-49.

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.

Interbull. 2005. Description of National Genetic Evaluation Systems for dairy cattle traits as applied in different Interbull member countries. [Online]. Available: http://www-interbull.slu.se/national_ges_info2/framesida-ges.htm. Accessed January 2005.

Kadarmideen H. N., R. Thompson, and G. Simm. 2000. Linear and threshold model genetic parameters for disease, fertility and milk production in dairy cattle. Anim Sci. 71:411-419.

Madsen, P., and J. Jensen. 2000. A users's guide to DMU. A package for analyzing multivariate mixed models. Version 6, release 4. Danish Institute of Agricultural Sciences, Denmark. 19 pp.

Rauw, W. M., E. Kanis, E. N. Noordhuizen-Stassen, and F. J. Grommers. 1998. Undesirable side effects of selection for high production efficiency in farm animals: a review. Livest. Prod. Sci. 56:15-33.

Saebø, S., T. Almøy, B. Heringstad, G. Klemetsdal, and A. H. Aastveit. 2005. Genetic evaluation of mastitis resistance using a first-passage time model for Wiener processes for analysis of time to first treatment. J. Dairy Sci. 88:834-841.

Saebø, S., and A. Frigessi. 2004. A genetic and spatial Bayesian analysis of mastitis resistance. Genet. Sel. Evol. 36:527-542.

SAS 1999. SAS Release 8.2. 1999-2001. SAS Inst. Inc., Cary, NC, USA.

Schneider, M. del P., E. Strandberg, V. Ducrocq, and A. Roth. 2005. Survival analysis applied to genetic evaluation for female fertility in dairy cattle. J. Dairy Sci. 88:2253-2259.

Yazdi, M. H., P. M.Visscher, V. Ducrocq, and R. Thompson. 2002. Heritability, reliability of genetic evaluations and response to selection in proportional hazard models. J. Dairy Sci. 85:1563-1577.

Zwald, N. R., K. A. Weigel, Y. M. Chang, R. D. Welper, and J. S. Clay. 2004. Genetic selection for health traits using producer-recorded data. I. Incidence rates, heritability estimates, and sire breeding values. J. Dairy Sci. 87:4287-4294.

Table 1. Assumed heritabilities [1] (on diagonal), genetic correlations (above diagonal), and environmental correlations (below diagonal) for the four simulated traits: 305-d milk production, interval between calving and first ovulation (CFO), conception liability and mastitis liability.

	305-d milk production	CFO	Conception liability	Mastitis liability
305-d milk production	0.30 (0.43)	0.10	-0.10	0.20
CFO	0.00	0.20 (0.22)	0.00	0.00
Conception liability	0.00	0.00	0.05 (0.055)	0.00
Mastitis liability	0.00	0.00	0.00	0.10 (0.13)

[1] Heritabilities with herd variance included in and excluded from (within brackets) the phenotypic variance.

Table 2. Estimates of heritability and variance components for simulated mastitis data analyzed as a binary trait with linear model (LM) and as time to first mastitis with survival analysis (SA) (mean and standard error based on 50 replicates).

	Method	
Estimates	LM	SA
Heritability	$0.027_{0.000}$	$0.037_{0.001}$
Sire variance	$0.00060_{0.000}$	$0.056_{0.001}$
Herd variance [1]	$0.0053_{0.000}$	$0.592_{0.005}$
Residual variance	$0.089_{0.000}$	

[1] Herd variance is not included in the phenotypic variance for presented heritabilities.

Table 3. Correlations between sire true breeding values (TBV) for mastitis liability, and estimated sire breeding values from linear models (LM) and survival analysis (SA) (mean and standard error based on 50 replicates).

	Estimated sire breeding values	
	LM	SA
TBV	$0.700_{0.004}$	$0.757_{0.003}$
LM		$0.902_{0.001}$

Table 4. Average true breeding value (TBV) for mastitis liability [1] of the best and worst 10 % of sires [2] ranked on estimated breeding values from linear models (LM) and survival analysis (SA), respectively (mean and standard error based on 50 replicates).

	Average TBV	
Method	Best 10 % of sires	Worst 10 % of sires
LM	-0.2236 $_{0.0037}$	0.2412 $_{0.0028}$
SA	-0.2457 $_{0.0032}$	0.2611 $_{0.0031}$

[1] A lower breeding value indicates a better mastitis resistance.
[2] Total number of sires = 400.

Table 5. The proportion (percentage) of the best or worst 10 % of sires [1] based on true breeding values for mastitis liability that were correctly identified to be in the best or worst 10 % of sires based on estimated breeding values from linear models (LM) and survival analysis (SA), respectively (mean (%) based on 50 replicates).

Method	Best sires	Worst sires
LM	44.2	49.4
SA	50.5	54.7

[1] Total number of sires = 400.