Genetic Evaluation of Mastitis Resistance Using a First-Passage Time Model for Wiener Processes for Analysis of Time to First Treatment

S. Sæbø,¹ T. Almøy,¹ B. Heringstad,² G. Klemetsdal,² and A. H. Aastveit¹

¹Department of Chemistry, Biotechnology, and Food Science, and

²Department of Animal and Aquacultural Sciences, Agricultural University of Norway, N-1432 Ås, Norway

ABSTRACT

A Wiener process is a Brownian-motion process initiated in a certain state in a state space, and the first passage time is defined as the time of the process to reach a predefined absorbing state where the process stops. Time from 31 d prepartum to first treatment of clinical mastitis (CM) was modeled as first passage times of such Wiener processes. Two processes were used to allow for several risk factors, and for each process, initiation was at some arbitrary time point, in a certain health state with drift toward or away from absorption (disease). The drift parameter of each process was expressed as linear functions of covariates (year of calving and sire). First passage time was defined as the time from process initiation until the first health status process reached zero (absorption). The model was fitted to records for 36,178 first-lactation daughters of 245 Norwegian cattle sires using a Bayesian approach and Markov chain Monte Carlo methods. Genetic evaluation of sires was carried out by calculating the posterior probability of no CM (the value of the survival function) by d 331, i.e., 300 d after first calving. Alternatively, sire evaluation was based on the integrated area under the survival curve. These measures were highly correlated (0.999), which indicates a small degree of crossings of the sire-dependent survival curves. Hence, sire-specific hazards were close to proportional, resulting in a higher rank-correlation to sire evaluations from a survival model with proportional hazards than to the results from a multivariate threshold model.

(Key words: genetic evaluation, clinical mastitis, survival analysis, Wiener process)

Abbreviation key: AIS = adjusted and integrated survival function; CM = clinical mastitis; NRF = Norwegian cattle; S = survival function (probability of no treatment of CM by time T).

INTRODUCTION

Mastitis is one of the most frequent and costly diseases affecting dairy cattle. Recently, there has been an increased interest in including this trait in dairy cattle breeding programs. In Norway, all veterinary treatments have been recorded on an individual cow basis since 1975, and selection against clinical mastitis (CM) has been included in the breeding program of Norwegian cattle (NRF) since 1978. In initial studies, CM in a defined opportunity period of first lactation has been considered as a binary trait and analyzed by linear models (Heringstad et al., 1999). Further developments aimed at taking the binary nature of data into account by using threshold models for analyses (Heringstad et al., 2001). Recently, a multivariate threshold model (Chang et al., 2002) and a longitudinal threshold model (Heringstad et al., 2003) have been developed where the lactation period was divided into intervals, thereby making it possible to use information on multiple treatments in addition to taking the time of CM treatment into account.

The time aspect can probably be even better exploited by using survival time models. Survival models are commonly used for analyzing time to disease data, not only in human health surveys but also within the area of animal breeding. The analysis of disease resistance in pigs presented by Henryon et al. (2001) is a recent contribution in this category. So far, survival analysis with complex hierarchical structures has been carried out by means of proportional hazards models in the animal breeding context (e.g., Ducrocq and Casella, 1996; Korsgaard et al., 1998). Our alternative is to adopt a model based on first passage times of stochastic processes, or more precisely, the competing Wiener process model of Sæbø and Almøy (2004a) and Sæbø et al. (2005).

In this model, one assumes that every cow at a given time is in some health state with a certain distance from disease outbreak. The term “distance” should not be interpreted too literally, as it is a measure with no defined biological scale, but the disease resistance of cows can still be compared on this scale. The unobservable physiological fight against infection can be approxi-

Received January 19, 2004.

Accepted September 2, 2004.

Corresponding author: S. Sæbø; e-mail: solve.sabo@ikbm.nlh.no.

mated by a stochastic process operating in some kind of health state space. This state space describes the possible health states with regard to the disease in question and is assumed continuous with an absorbing state. An absorbing state is defined as a state from which the stochastic process cannot escape; hence, the process stops if it enters the absorbing state. In our context, we will let the event of process absorption represent CM contraction, and we will use Wiener processes with drift (Chhikara and Folks, 1989) to approximate the changes in a cow's health state over time. The time from when the process is initiated until it is absorbed (the first passage time) is a stochastic variable following some known probability distribution, as described later.

The causative background of CM is complex and thus, advantageously, can be modeled by assuming several latent Wiener processes competing in reaching the absorbing state. It is well known from previous studies that the risk of CM is highest in early lactation (Pösö and Mäntysaari, 1996; Heringstad et al., 1999; Nash et al., 2000). The elevated risk in the first days after calving is probably connected to physiological changes known to be initiated a few days before calving (e.g., Kehrli et al., 1989). One theory is that increased levels of certain hormones result in a redistribution of blood cells in the body with increased levels in the bloodstream but reduced levels in other tissues such as the mammary glands (Kulberg et al., 2002). This will result in increased susceptibility to infections and mastitis. These physiological changes put the cow in a position where she is easily infected, i.e., in a health state close to disease.

There are several other risk factors in addition to the suppression of the immune system around calving. Many of these factors are related to the cow's environment, such as hygiene and milking technique. Unfavorable levels of environmental factors may be more uniformly distributed over the lactation, resulting in a fairly constant risk of infection and disease.

The 2-fold risk pattern described previously, with a temporary increased risk around calving and a more or less constant risk throughout the lactation, was modeled by assuming 2 latent Wiener processes.

The objectives of this study were to analyze time to first treatment of CM for first-lactation NRF cows using a first passage time model and to examine relevant selection criteria for ranking and selection purposes.

MATERIALS AND METHODS

Data

The data were from first-lactation daughters of 245 NRF sires, progeny tested in 1991 and 1992. Only re-

cords of cows with a first calving in 1990 through 1992 and from herds with at least 5 first-lactation daughters of these test bulls were kept. The resulting data set included 36,178 first-lactation cows from 5286 herds. The pedigree file contained 302 males, including the 245 sires with daughter records. With exception of one sire with only 22 daughters, the minimum number of daughters per sire was 104.

For each cow, the time variable was set to zero at d 31 before first calving. The observed event time was the day of first veterinary treatment of CM. A total of 24.5% of the cows had at least one case of CM during first lactation. Cows with no treatment of CM prior to 31 d before the second calving were right censored. Additionally, there were random right-censored observations caused by culling. Culling rates and CM frequency in the data are described by Heringstad et al. (2003).

Survival Model

In the competing process model (Sæbø and Almøy, 2004a; Sæbø et al., 2005) a Wiener process, $W(t)$, is assumed initiated at time zero, in a "health state" c (with $c > 0$). The first passage time of the process is the time of the process reaching zero, which is an absorbing state, as described previously. A Wiener process may have drift toward the absorbing state. For a Wiener process with drift, the process increment from time t to $t + \Delta t$ is normally distributed with a mean, depending on the drift. More specifically, it is assumed that $W(t + \Delta t) - W(t) \sim N(-\mu\Delta t, \sigma^2\Delta t)$, where $\mu > 0$, is the drift parameter of the process, and σ^2 is the process variance coefficient. When $\mu > 0$, we refer to the drift as positive (toward the absorbing state 0).

Under these conditions, the first passage time T (the time of process absorption) follows an inverse Gaussian distribution with parameters c and μ (Chhikara and Folks, 1989; Aalen and Gjessing, 2001). For a change in σ^2 , there exist other values of μ and c giving equal distribution for T ; hence, σ^2 is dependent on μ and c . Therefore, for simplicity, σ^2 was restricted to one in our analyses. The rate by which such processes are absorbed depends on how far from absorption the processes are initiated (value of c) and how "fast" the processes approach zero (value of μ). The risk of absorption increases if c is moved toward zero and/or μ is increased.

Because time of process initiation ($T = 0$) is unlikely to be equal to the starting point of any health process, the Wiener processes are assumed initiated at some arbitrary time point ϕ . Then, the first passage time T follows a 3-parameter inverse Gaussian distribution given by (Chhikara and Folks, 1989)

Figure 1. Simulated health states of 2 latent processes for 2 cows. For both cows, process 1 was initiated at $\phi_1 = 30.4$ d in health state $c_1 = 1.75$ with drift $\mu_1 = -0.72$, whereas the values for process 2 were $\phi_2 = -16.2$ d, $c_2 = 14.6$, and $\mu_2 = -0.05$. For each cow, the “observed” time of disease is the time of first absorption in state 0 for either of the 2 latent processes (t_1 and t_2 for cow 1 and 2, respectively). The hazard function corresponding to the chosen parameter values is also given (long-dashed line).

$$f(t | c, \mu, \phi) = \frac{c}{\sqrt{2\pi}} (t - \phi)^{-3/2} \exp \left[-\frac{(c - \mu(t - \phi))^2}{2(t - \phi)} \right] \quad [1]$$

In the case of negative drift, i.e., $\mu < 0$, the process drifts away from the absorbing state, and the distribution in equation 1 is not a usual probability density, as it will integrate to a value < 1 . In a disease context, this would represent a situation where an individual is close to disease (small value of c) at the time of process initiation, but where the risk of absorption decreases as time passes.

The survival function (**S**) expressing the probability of no absorption (disease) by time t , is given by

$$\Pr(T > t) = S(t | c, \mu, \phi) = 1 - F(t | c, \mu, \phi) = 1 - \int_0^t f(u | c, \mu, \phi) du \quad [2]$$

An exact expression for equation 2 can be found in Aalen and Gjessing (2001). The hazard function, given by $h(t | c, \mu, \phi) = f(t | c, \mu, \phi) / S(t | c, \mu, \phi)$, expresses the instantaneous risk of absorption at time t given that the process has not been absorbed prior to t .

If several risk factors are involved, it is convenient to approximate each factor's impact on health by a Wiener

process. In general, it can be assumed that J independent latent processes compete in reaching the absorbing state (disease). The observed first passage time t_i for an individual i is equal to $\min(t_{i1}, \dots, t_{iJ})$, i.e., the time of the first process to be absorbed. It can be shown (Sæbø et al., 2005) that the overall hazard at any time point t will consist of the sum of the process-specific hazards of all processes initiated by time t , i.e.,

$$h(t | \theta) = \sum_{j=1}^J h(t | \theta_j) I(t > \phi_j) \quad [3]$$

where $\theta = \{\theta_1, \dots, \theta_J\}$, with $\theta_j = \{c_j, \mu_j, \phi_j\}$ and $I(\cdot)$ is the indicator function taking the value 1 if its argument is true and 0 otherwise. Correspondingly, the overall survival function will be the product of the process-specific survival functions of all active processes at time t :

$$S(t | \theta) = \prod_{j=1}^J S(t | \theta_j)^{I(t > \phi_j)} \quad [4]$$

The overall first passage time distribution is given by the relation

$$f(t | \theta) = h(t | \theta) S(t | \theta). \quad [5]$$

The number J of latent processes may be determined by the data as the number maximizing some model-fitting criterion (Sæbø and Almøy, 2004a). Alternatively, the decision can be based on prior knowledge of the phenomenon. In the introduction, we argued for using 2 latent processes for describing the infection pattern of CM, and in the following, we, therefore, use $J = 2$.

The physiological struggle against disease in the period with elevated risk around calving was approximated by one Wiener process (process 1), initiated at an unknown time, $t = \phi_1$. A second Wiener process (process 2), independent of process 1, initiated at another unknown time point $t = \phi_2$, was used to model the constant fight against disease during lactation. Absorption of any of the 2 processes represents a case of CM.

For illustration purposes, the former 2 processes were simulated for 2 cows. Assuming parameter values as estimated in the results section, one obtained the individual health state values of Figure 1. The first passage time, t_{ij} , specific for process j ($j = 1, 2$) of cow i ($i = 1, 2$) is the time from process initiation until the health status value reached zero (absorption). For cow 1, the overall first passage time $t_1 = \min(t_{11}, t_{12})$ was the time of latent process 2 to be absorbed ($t_1 \approx 27$ d). For this cow, process 1 was never absorbed, as it drifted quickly away from absorption after it was initiated at time $t = \phi_1$. For cow 2, process 2 drifted away from zero, whereas process 1 immediately after initiation was absorbed, giving a first passage time $t_2 \approx 34$ d. For the parameter values chosen for this simulation, the overall hazard function is also indicated in Figure 1, and one notes the dramatic, yet temporary, increase in risk of absorption just after the initiation of process 1, which is initiated close to the absorbing state but with a strong drift away from absorption.

The effects of covariates on the risk of CM were taken into account by expressing the drift parameter of each process as linear functions of the covariates. By this, we assume that these variables in some way influence how fast an individual drifts toward or away from disease. For a daughter of sire k ($k = 1, 2, \dots, 245$) with first calving in year h ($h = 1990, 1991, 1992$), the drift parameter of process j ($j = 1, 2$) was assumed to be

$$\mu_{jkh} = \xi_j + s_{jk} + \alpha_{jh}, \quad [6]$$

where α_{jh} is the fixed effect of year of first calving and s_{jk} is the random effect of sire k . The first level of the year effect was restricted to zero; thus, the actual effect of the first year was included in the constant term ξ_j . The 245-element vector of sire effects was assumed to be multivariate normal (MVN). Hence, $\mathbf{s}_j \sim \text{MVN}(\mathbf{0}, \mathbf{A}\sigma_{sj}^2)$, where σ_{sj}^2 is the sire variance of process j and

$\mathbf{A}$ is the additive genetic relationship matrix. The matrix $\mathbf{A}$ was constructed as follows. First, the (302×302) matrix $\tilde{\mathbf{A}}$ of additive relationships between all 302 sires in the pedigree file was found. From this matrix, the (245×245) submatrix $\mathbf{A}$ was extracted, corresponding to the 245 sires with daughter records. As commented by Korsgaard et al. (1998), this can be an alternative to using the entire relationship matrix $\tilde{\mathbf{A}}$ and the 302-element vector $\tilde{\mathbf{s}}$ of sire effects with augmented effects of sires without daughter records. The only cost is the loss of predicted values of sires without records.

All unknown quantities were estimated using Markov chain Monte Carlo (MCMC) methods. This requires construction of a likelihood function based on the observed data. The variable y_i was defined to be either the time of first veterinary treatment of CM ($t_{i,m}$) for infected cow i , or the censoring time ($t_{i,c}$) if the cow did not contract mastitis. The observed data for cow i is defined as (y_i, δ_i) , where $\delta_i = I(y_i = t_{i,m})$ is the censoring indicator, taking the value 1 if its argument is true, and 0 otherwise. A cow with observed CM will contribute $f(t_{i,m}|\theta)$ to the likelihood (or equivalently $h(t_{i,m}|\theta)S(t_{i,m}|\theta)$ as given by equation 5), whereas a censored cow will have a contribution equal to $S(t_{i,c}|\theta)$. The conditional likelihood may be constructed by taking the product over all individuals, conditioning on the sire effects $\mathbf{s} = \{\mathbf{s}_1, \mathbf{s}_2\}$. Using equations 3 and 4, the conditional likelihood is given by

$$p(y, \delta | \theta, \mathbf{s}) = \quad [7]$$

$$\prod_{i=1}^n \left[\sum_{j=1}^2 \{h(y_i | \theta_j) I(y_i > \phi_j)\}^{\delta_i} \left\{ \prod_{j=1}^2 S(y_i | \theta_j)^{I(y_i > \phi_j)} \right\} \right].$$

A function proportional to the posterior density of θ , $\mathbf{s}$, and $\sigma_s^2 = \{\sigma_{s1}^2, \sigma_{s2}^2\}$ was constructed by

$$p(\theta, \mathbf{s}, \sigma_s^2 | y, \delta) \propto p(y, \delta | \theta, \mathbf{s}) p(\mathbf{s} | \sigma_s^2) p(\sigma_s^2) p(\theta) \quad [8]$$

where the priors $p(\theta)$ and $p(\sigma_s^2)$ express our prior beliefs regarding the components of θ and of the sire variance components σ_s^2 . Except for the elements of the sire vectors $\mathbf{s}_j$ ($j = 1, 2$), all unknown quantities were assumed independent. Hence, the priors $p(\theta)$ and $p(\sigma_s^2)$ in equation 8 can be replaced by the product of the prior densities of all of the independent components of θ and σ_s^2 . Proper, vague priors were assumed for all parameters. For the inverse of σ_{sj}^2 , gamma priors were used (Gamma(0.001, 0.001)), and for c_j , log-normal priors were assumed (mean 0 and variance 1000 for the natural logarithm of c_j). All other parameters in the model were assumed normally distributed with mean 0 and

Table 1. Means, standard deviations, and percentiles of the estimated posterior distributions for the parameters of the first passage time model with 2 latent Wiener processes.

Parameter ¹	Posterior			
	Mean	SD	2.5%	97.5%
ξ_1	-0.72	0.022	-0.77	-0.68
ξ_2	-0.049	0.0023	-0.054	-0.044
c_1	1.75	0.023	1.70	1.79
c_2	14.60	0.18	14.20	14.90
ϕ_1	30.4	0.018	30.4	30.5
ϕ_2	-16.2	1.00	-18.3	-14.3
σ_{s1}^2	0.0048	0.00096	0.0033	0.0069
σ_{s2}^2	0.000091	0.000014	0.000066	0.00012
$\alpha_{1,1991}$	0.022	0.014	-0.00041	0.049
$\alpha_{1,1992}$	0.039	0.019	0.0028	0.075
$\alpha_{2,1991}$	0.00093	0.0014	-0.0018	0.0037
$\alpha_{2,1992}$	0.0023	0.0020	-0.0017	0.0063

¹The parameter ξ_j represents the drift of latent process j for cows having their first calf in 1990 and a sire with sire effect equal to zero. The process was initiated in state c_j at time ϕ_j , where time zero was defined as 31 d before first calving. For each process, a sire effect with variance σ_{sj}^2 was added to the drift. The parameters $\alpha_{j,1991}$ and $\alpha_{j,1992}$ give the extra contribution to the drift for cows calving in 1991 and 1992, respectively.

variance 1000. Various choices for the dispersion of the priors were considered, but it appeared that the likelihood dominates the posterior distribution because of the amount of data for most parameters. Sæbø et al. (2005) found only minor prior sensitivity for the sire variance components.

Because of the complexity of the joint posterior distribution, conditional posterior distributions for single parameters cannot be obtained in closed form. Draws from the conditional posterior distributions were therefore obtained by combining single component and block-wise updating using the Metropolis-Hastings algorithm (Metropolis et al., 1953) at each step.

A total of 110,000 iterations of the Metropolis-Hastings algorithm were run. In each iteration, values of all unknown parameters were in turn drawn from their respective conditional distributions. Based on visual inspection of the trace plots, 10,000 iterations were discarded as burn-in. After burn-in, 100,000 iterations were run, where only every 10th value of each parameter was saved to save computer space. Statistical inference was based on the resulting 10,000 samples.

Sire Evaluation Criteria

The predicted sire effects can be used to identify sires with daughters showing the weakest drift toward disease for each latent process. For sire-ranking purposes, it would, however, be useful to summarize the information from the 2 processes into a single measure. Two such measures based on the predicted sire effects were

computed: the probability of no CM by a defined time τ and an integrated survival function. Sire ranking based on these 2 measures were compared with corresponding rankings from the multivariate threshold model of Chang et al. (2002) and from the proportional hazards model of Sæbø and Frigessi (2004).

Probability of no CM by time τ . The probability of no CM during first lactation is given by the value of the survival function at the end of the lactation. Here, we define $t = \tau$ as the lactation end point (for some reasonable $\tau > 0$). Posterior distributions of these probabilities may be computed by exploiting the output from the MCMC algorithm. At iteration g of the algorithm ($g = 1 \dots 10,000$), a set of sire effects $s_{1,k}^g$ and $s_{2,k}^g$ for sire k , are drawn from the respective posterior distributions. In addition, values of all other model parameters are also obtained. Let θ_k^g comprise the sampled values of all parameters relevant for sire k at iteration g , including the sire effects. Sire-specific survival probabilities are computed as

$$S_k^g(\tau) = S(\tau | \theta_k^g) \quad [9]$$

for all g . From these values, the posterior density of $S_k(\tau)$ can be estimated for all sires, and the estimated posterior means may be used as a selection criterion.

Adjusted integrated survival function. An alternative to the $S_k(\tau)$ criterion would be to evaluate sires according to a criterion based on the area under the survival curve. It can be shown (Sæbø and Almøy, 2004b) that an adjusted integrated survival function $AIS_k(\tau)$ has the following pleasant property:

$$AIS_k(\tau) = \frac{1}{(1 - S_k(\tau))} \int_0^\tau S(t | \theta_k) dt = E \left[\frac{1}{h_k(t | \theta_k)} \mid 0 < t < \tau \right]. \quad [10]$$

Hence, $AIS_k(\tau)$ is equal to the conditional expectation of the inverted hazard function. A large value of $AIS_k(\tau)$ means that sire k has a favorable hazard pattern over the time interval $[0, \tau]$. The value of $AIS_k(\tau)$ may be computed by a simple numerical integration routine (e.g., the rectangle rule) at each step of the MCMC algorithm and estimated posterior means may be used for sire evaluations.

RESULTS AND DISCUSSION

Posterior Estimates

Monte Carlo estimates of posterior means, standard deviations, and percentiles of model parameters are

given in Table 1. The posterior mean estimate of ϕ_1 is 30.4, indicating that process 1 is initiated just before calving. The initial health state of this process is $\hat{c} = 1.75$, which is close to absorption relative to process 2, for which the estimated initial state is $\hat{c}_2 = 14.6$. Because of the small distance to the absorbing state at the time of calving for process 1, the risk of process absorption (disease) is high during this period (Figure 1). Insertion of estimated posterior means for relevant parameters into equation 6 gives negative drift parameters of process 1 for all individuals, implying that the processes are expected to drift away from disease. This means that the risk of absorption (CM) is highest immediately after process initiation and then decreases toward zero as the process is brought away from the absorbing state. This harmonizes with the immune system being temporarily suppressed around calving because of physiological changes (Kulberg et al., 2002).

The estimate of ϕ_2 indicates that process 2 is initiated approximately 47 d before calving. Also, for this process, the drift parameter (from equation 6) turns out to be negative for all individuals, but the expected drift away from the absorbing state for this process is much weaker than for process 1. The corresponding process-specific hazard function shows a small increase initially, after which it slowly decreases toward zero (Figure 1). One explanation for the decrease may be reduced milk production toward the end of lactation. The hazard at the population level may also decrease, even if there are constant or increasing hazards at the individual level, if there is a heterogeneity across the population and the most 'frail' animals become ill first and only the most resistant animals contribute to the hazard estimate at later time points (Vaupel et al., 1979).

For the year effect, the estimates indicate an increase in the risk of CM over the yr 1990 to 1992. This is in accordance with the observed increase in CM frequency during this period (Heringstad et al., 1999). For further discussion with regard to the fit of the competing process model to the mastitis data, we refer to Sæbø and Almøy (2004b) and Sæbø et al. (accepted).

Sire Evaluation

A low value of the sire effect would be preferred for each of the latent processes, as it reduces the drift toward disease, or more correctly, increases the drift away from disease. However, the Pearson's correlation between s_1 and s_2 was only 0.36, indicating that a top-ranked sire based on process 1 may not necessarily be among the top ranked sires for process 2.

To summarize the information from the 2 processes into a single measure, $S_k(331)$ and $AIS_k(331)$ were calculated. The top 10 sires for both measures are listed

Table 2. The identification numbers of the 10 highest ranking sires from $S_k(331)$, $AIS_k(331)$, $MTM(331)$, and PHM .¹ Spearman's rank correlations between sire evaluations are also given.

Rank	$S_k(331)$	$AIS_k(331)$	$MTM(331)$	PHM
1	77	77	149	77
2	104	79	79	149
3	79	104	77	182
4	128	241	182	79
5	241	147	104	104
6	244	128	150	236
7	147	244	20	128
8	15	15	181	167
9	182	182	243	178
10	238	238	58	212
Rank correlations				
$S_k(331)$		0.999	0.806	0.914
$AIS_k(331)$			0.808	0.915
$MTM(331)$				0.905

¹Rankings in the table are based on $S_k(331)$ = competing process model [probability of no mastitis by time 331 (300 d after calving)]; $AIS_k(331)$ = competing process model (adjusted integrated survival function up to time 331); $MTM(331)$ = multivariate threshold model (probability of no mastitis by time 331); and PHM = proportional hazards model (estimates of sire effects).

in Table 2. The Spearman's rank correlations between $S_k(331)$ and $AIS_k(331)$ was 0.999, and the 2 criteria ranked the same set of sires among top 10 (Table 2). This means that there is a small degree of crossing between the sire-dependent survival curves. Posterior mean survival curves of the two sires ranking highest and lowest for $S_k(331)$ are given in Figure 2. Apart from a large difference in the drop of survival probability

Figure 2. Posterior mean survival curves of the two sires with highest (dotted line) and lowest (solid line) probability of no mastitis by time 331 (300 d after calving).

around calving, the curves show similar shapes. From this it seems that varying risk of CM around calving is the main cause of variation in survival probability at the end of the lactation; hence, the choice of lactation end point τ was not crucial for the sire ranking.

Among the top 10 sires for $S_k(331)$ and $AI S_k(331)$, only 4 and 5 were in common when compared with the corresponding ranking from the multivariate threshold model (Chang et al., 2002) and the proportional hazards model (Sæbø and Frigessi, 2004), respectively. A total of 4 sires were ranked top 10 for all criteria (77, 79, 104, and 182). Rank correlations between the 2 measures of the competing process model and the proportional hazards model were highest (0.91), indicating that sire-specific hazards were close to proportional, as was also observed in Figure 2. The multivariate threshold model and the competing process model had a less correlated ranking (0.81), but there is still considerable agreement on which sires to select. The fact that different statistical approaches point out similar sets of top-ranked sires strengthens our belief in these animals as genetically superior for CM resistance. Still, it would be of interest to compare the different approaches in terms of efficiency of selection. This may be accomplished by regressing CM incidence in second-crop daughters for elite sires on sire evaluations in first-crop daughters (Ødegård et al., 2003).

Costs attributable to CM obviously vary over lactation, and the $AI S_k(331)$ criterion may be improved by adjusting for external factors such as potential losses of milk yield from CM, leading to larger costs early in lactation than later. By including a weight function $v(t)$ expressing how the external factor varies with time, the adjusted, integrated, and weighted survival function ($AIWS_k(\tau)$) can be written:

$$AIWS_k(\tau) = \frac{1}{(1 - S_k(\tau))} \int_0^{\tau} v(t)S(t | \theta_k)dt = \quad [11]$$

$$E\left[\frac{v(t)}{h_k(t | \theta_k)} \mid 0 < t < \tau\right],$$

which equals the conditional expectation of the weight hazard ratio over time interval $[0, \tau]$.

Our competing process model is based on an assumption of independent latent processes, which may not hold. Further work is needed to investigate the effect of this assumption, and perhaps develop the model further to allow for dependent processes. In this first attempt toward genetic evaluation of sires with a competing process model, the number of latent processes was decided a priori. It would be desirable to conduct a further study to examine how the number of processes affects sire evaluation. Finally, there are other fixed

and random covariates that need to be accounted for, for example the herd effect. Because herd sizes were small (average of 6.8 cows per herd), the information on the herd effect in the data was limited. Hence, adjustments for herd effects were not carried out in this first analysis with the competing process model.

CONCLUSION

The competing process model adopted in this study is an alternative to the proportional hazards models. We find the process point of view appealing, as the properties of the estimated latent processes harmonize with prior biological knowledge. Further, no assumption of hazard proportionality needs to be done. However, the implementation of the competing process model is computationally demanding and increasingly so with the number of latent processes. By adopting a sire model, we have limited the number of parameters to be estimated. An animal model implementation would thus require even larger computer resources. Hence, more effective sampling algorithms would be desirable for larger implementations. However, the computer power is increasing, making such large-scale analyses feasible in future.

ACKNOWLEDGMENTS

This work is part of the “Healthy Cow” project financed by the Research Council of Norway. Access to the data was given by the Norwegian Dairy Herd Recording System and by the Norwegian Cattle Health Service in agreement number 011.2000 by October 12, 2000. GENO Breeding and A.I. Association are acknowledged for providing pedigree information on sires.

REFERENCES

- Aalen, O. O., and H. Gjessing. 2001. Understanding the shape of the hazard rate: A process point of view. *Stat. Sci.* 16:1–22.
- Chang, Y. M., D. Gianola, B. Heringstad, and G. Klemetsdal. 2002. Correlations between clinical mastitis in different periods of first-lactation Norwegian Cattle using a multivariate threshold model. Pages 177–192 in *Case Studies in Bayesian Statistics*. Vol. 6. C. Gatsonis, A. Carriquiry, A. Gelman, D. Higdon, R. Kass, D. Pauler, and I. Verdinelli, ed. Springer-Verlag, New York, NY.
- Chhikara, R. S., and J. L. Folks. 1989. *The Inverse Gaussian Distribution: Theory, Methodology, and Applications*. Marcel Dekker, New York, NY.
- Ducrocq, V., and G. Casella. 1996. A Bayesian analysis of mixed survival models. *Genet. Sel. Evol.* 28:505–529.
- Henryon, M., P. Berg, J. Jensen, and S. Andersen. 2001. Genetic variation for resistance to clinical and sub-clinical diseases exists in growing pigs. *Anim. Sci.* 73:375–387.
- Heringstad, B., Y. M. Chang, D. Gianola, and G. Klemetsdal. 2003. 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 correlations with protein yield. *J. Dairy Sci.* 82:1325–1330.

- Heringstad, B., R. Rekaya, D. Gianola, G. Klemetsdal, and K. A. Weigel. 2001. Bayesian analysis of liability to clinical mastitis in Norwegian cattle with a threshold model: Effects of data sampling method and model specification. *J. Dairy Sci.* 84:2337–2346.
- Kehrli, M. E., B. J. Nonnecke, and J. A. Roth. 1989. Immunosuppression in dairy cows at calving. *Am. J. Vet. Res.* 50:207–214.
- Korsgaard, I. R., P. Madsen, and J. Jensen. 1998. Bayesian inference in the semiparametric log-normal frailty model using Gibbs sampling. *Genet. Sel. Evol.* 30:241–256.
- Kulberg, S., A. K. Storset, B. Heringstad, and H. J. S. Larsen. 2002. Reduced levels of total leucocytes and neutrophils in Norwegian cattle selected for decreased mastitis incidence. *J. Dairy Sci.* 85:3470–3475.
- Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. Equations of state calculations by fast computing machine. *J. Chem. Phys.* 21:1087–1091.
- Nash, D. L., G. W. Rogers, J. B. Cooper, G. L. Hargrove, J. F. Keown, and L. B. Hansen. 2000. Heritability of clinical mastitis incidence and relationships with sire transmitting abilities for somatic cell score, udder type traits, productive life, and protein yield. *J. Dairy Sci.* 83:2350–2360.
- Ødegård, J., G. Klemetsdal, and B. Heringstad. 2003. Genetic improvement of mastitis resistance: Validation of somatic cell score and clinical mastitis as selection criteria. *J. Dairy Sci.* 86:4129–4136.
- Pösö, J., and E. A. Mäntysaari. 1996. Relationships between clinical mastitis, somatic cell score, and production for the first three lactations of Finnish Ayrshire. *J. Dairy Sci.* 79:1284–1291.
- Sæbø, S., and T. Almøy. 2004a. Composite hazard functions reflecting competing latent processes. *Biometrical J.* 46:103–113.
- Sæbø, S., and T. Almøy. 2004b. A note on measures for treatment comparison based on parametric survival models with non-proportional hazards. Ph.D. Diss., Agricultural University of Norway, Ås.
- Sæbø, S., T. Almøy, and A. H. Aastveit. 2005. Disease resistance modeled as first passage times of genetically dependent stochastic processes. *Appl. Stat.* 54:273–285.
- Sæbø, S., and A. Frigessi. 2004. A genetic and spatial Bayesian analysis of mastitis resistance. *Genet. Sel. Evol.* 36:527–542.
- Vaupel, J. W., K. G. Manton, and E. Stallard. 1979. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* 16:439–454.