

## Heritability, Reliability of Genetic Evaluations and Response to Selection in Proportional Hazard Models

M. H. Yazdi,\* P. M. Visscher,\* V. Ducrocq,† and R. Thompson‡

\*University of Edinburgh, Institute of Cell, Animal and Population Biology, West Mains Road, Edinburgh EH9 3JT, UK,

†Station de Génétique Quantitative et Appliquée—Institut National de la Recherche Agronomique, 78352 Jouy-en-Josas, France,

‡Rothamsted Experimental Station, Institute of Arable Crops Research, Harpenden, Hertfordshire AL5 2JQ, UK, and Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS

### ABSTRACT

The purposes of this study were 1) to investigate the heritability, reliability, and selection response for survival traits following a Weibull frailty proportional hazard model; and 2) to examine the relationship between genetic parameters from a Weibull model, a discrete proportional hazard model, and a binary data analysis using a linear model. Both analytical methods and Monte Carlo simulations were used to achieve these aims. Data were simulated using the Weibull frailty model with two different shapes of the Weibull distribution. Breeding values of 100 unrelated sires with 50 to 100 progeny (with different levels of censoring) were generated from a normal distribution and two different sire variances. For analysis of longevity data on the discrete scale, simulated data were transformed to a discrete scale using arbitrary ends of discrete intervals of 400, 800, or 1200 d. For binary data analysis, an individual's longevity was either 0 (when longevity was less than the end of interval) or 1 (when longevity was equal or greater than the end of interval).

Three different statistical models were investigated in this study: a Weibull model, a discrete-time model (a proportional hazard model assuming that the survival data are measured on a discrete scale with few classes), and a linear model based upon binary data. An alternative derivation using basic expressions of reliabilities in sire models suggests a simple equation for the heritability on the original scale (effective heritability) that is not dependent on the Weibull parameters.

The predictions of reliabilities using the proposed formulae in this study are in very good agreement with reliabilities observed from simulations. In general, the estimates of reliability from either the discrete model

or the binary data analysis were close to estimates from the Weibull model for a given number of uncensored records in this simplified case of a balanced design. Although selection response from the binary data analysis depends on the end of interval point, there is a relatively good agreement between selection responses in the Weibull model and the binary data analysis. In general, when the underlying survival data is from a Weibull distribution, it appears that the method of analyzing data does not greatly affect the results in terms of sire ranking or response to selection, at least for the simplified context considered in this study.

**(Key words:** genetic parameter, proportional hazard model, response to selection, survival analysis)

**Abbreviation key:** PEV = prediction error variance, PH = proportional hazard, R = reliability.

### INTRODUCTION

Longevity is a trait of interest for animal breeders in general and dairy breeders in particular because of its effect on economic performance (e.g., Allaire and Gibson, 1992; VanRaden and Wiggans, 1995; Strandberg, 1996). In dairy cattle, actual measurements of longevity are obtainable only when a cow is culled or disposed of or after selection decisions have been made. However, for rapid genetic progress, genetic evaluation should be carried out as early as possible during an animal's life. Several strategies have been suggested to estimate the breeding value of an animal still in the herd. The main approaches that have been proposed and applied are 1) a simple modeling of a 0-1 variable indicating whether the cow is still alive at any specific time (e.g., first, second, or third lactation); 2) a linear modeling of herd life combined with projected records for cows still alive at the time of genetic evaluation; and 3) modeling of the probability that a cow is culled given that she was still in the herd until that time. In the first approach, survival variable was considered as a binary trait and analyzed either using

Received October 10, 2000.

Accepted December 26, 2001.

Corresponding author: M. H. Yazdi.

a linear or threshold model (Madgwick and Goddard, 1989; Jairath et al., 1998; Vollema and Groen, 1998; Boettcher et al., 1999). Here, the use of a linear model for a binary trait is considered to be a relatively robust approximation when the proportion of zeros and ones is not too extreme. For the second approach, Brotherton et al. (1998) and VanRaden and Klaaskate (1993) developed simple methods to project the life expectancy of cows that are still in the herd, that is censored longevity records, based on current available information either in the number of lactations or months of productive life and then these records together with completed records were used in a linear model analysis. The last approach, the analysis of survival data using a proportional hazard (PH) model (Cox, 1972; Kalbfleisch and Prentice, 1980) was suggested for animal genetic evaluation by Smith (1983) and Smith and Quaas (1984). Proportional hazard models can be semiparametric ("Cox models") or parametric (e.g., Weibull models) (e.g., Kalbfleisch and Prentice, 1980), depending on whether the baseline hazard distribution (acting like a mean) is assumed arbitrary or parametric. One advantage of such an approach is that censored records are naturally included in the analysis. Survival models that include random effects, which is required in genetic evaluation, are sometimes called frailty models. Ducrocq (1987) and Ducrocq and Sölkner (1998) presented examples of development and adaptation to large-scale applications of Smith's work in dairy cattle.

The correlation between the selection criterion and the true breeding value is often termed the accuracy of prediction. In dairy cattle evaluation, the accuracy of evaluations is usually expressed in terms of reliability, which is the squared correlation between the selection criterion and the true breeding value, or the repeatability of (conceptual) repeated progeny tests. Reliability in linear models can also be calculated from the prediction error variance (PEV) as given by the diagonal term of the inverse of the information matrix (Henderson, 1975). In large national genetic evaluations, obtaining this term is not tractable. Traditionally, an approximate reliability is used. For linear models, many approximate formulae are available, all of which use the heritability of the trait under consideration and an approximation of the diagonal elements of the inverse of the coefficient matrix pertaining to the mixed model equations. Selection of superior animals are based on their EBV, which is usually released together with a measure of accuracy of the evaluation (reliability) from (inter) national breeding evaluation. From the breeder's perspective, reliability is fundamental for prediction of selection response, risk assessment and comparison of different selection strategies.

Results are presented in terms of reliability (**R**), but other commonly used terms including accuracy, and PEV can be expressed as transformations of **R**.

For estimation of the reliability in nonlinear models, such as survival models, the asymptotic prediction error variance could be calculated from the diagonal term of the inverse of the information matrix. However, in survival analysis, Ducrocq (1999a) suggested an approximate formula for reliability based on the number of uncensored records and on the heritability on the original scale, i.e., the scale on which time until failure (culling or death) is measured. For the Weibull frailty model, heritability has been defined on the log scale (Ducrocq and Casella, 1996) and transformed to the original scale using a Taylor series expansion (Ducrocq, 1999a). This formula served its intended purpose, i.e., gave good approximations of the true asymptotic prediction error variance of genetic evaluation of French dairy bulls based on longevity of their progeny. However, this heritability depends on only one of the two parameters of the Weibull distribution, a rather disturbing property given the relatively strong negative correlation observed in the estimates of these two parameters: Different combinations of the Weibull parameters can lead to a similar fit of the data, but with different heritability values. Also, the derivation and interpretation of this heritability on the original scale has been thought to be dubious (e.g., Korsgaard et al., 1999).

The purposes of this study were 1) to investigate the heritability, reliability, and selection response for survival traits following a Weibull frailty proportional hazard model, 2) to examine the relationship between genetic parameters from a Weibull model, a discrete proportional hazard model, and a binary data analysis using a linear model. Both analytical methods and Monte Carlo simulations were used to achieve these aims.

## MATERIALS AND METHODS

### Theory

**Statistical models.** Three different statistical models were investigated in this study: a Weibull model, a discrete-time model (i.e., a PH model assuming that the survival data are measured on a discrete scale with few classes; Prentice and Gloeckler, 1978), and a linear model based upon binary data. The first and second statistical models are based on the concept of a hazard function, the cows' limiting risk of being culled at time  $t$ , conditional upon survival to time  $t$  (Ducrocq, 1987). The binary data analysis is based on the regular mixed model analysis of a 0/1 trait, indicating whether the animal is alive at a given time,

ignoring the continuous nature of the trait. We are interested in functions of genetic parameters for these models (heritabilities, reliabilities) and the relationship between genetic evaluations based upon the different evaluation models. Throughout, we model the genetic component as the random effect of a sire, i.e., we deal with sire models only. This restriction results from the fact that a good knowledge of the reliability of sires' EBV is considered as critical. Sire reliabilities are expected to be very close in sire and animal models, at least when a progeny test is used.

We assume that the hazard function of a cow can be expressed as:

$$\lambda_{ij}(t) = \lambda_0(t)\exp\{s_i\}, \tag{1}$$

where  $\lambda_{ij}(t)$  is the hazard function of an animal  $j$ , progeny of sire  $i$ , depending on time  $t$  through  $\lambda_0(t)$ , called the baseline hazard function. The sire additive genetic effect,  $s_i$ , was assumed to have a normal distribution,  $s_i \sim N(0, \sigma_s^2)$ , where  $\sigma_s^2$  is the sire variance. Sires were assumed to be unrelated, and no fixed effect was included other than an overall mean. The data modeled is the time to culling (or failure),  $t$ , or if the record is censored the time at censoring. Here we assumed that the baseline hazard function is described by a Weibull hazard distribution,  $\lambda_0(t) = \lambda\rho(\lambda t)^{\rho-1}$ . The Weibull parameters  $\rho$  and  $\lambda$  define the shape and scale of the hazard function, respectively. For example, if  $\rho > 1$  (respectively,  $\rho < 1$ ) the hazard of a daughter of an average sire ( $s_i = 0$ ), increases (respectively, decreases) with time. For  $\rho = 1$ , the baseline hazard is constant, and the Weibull model reduces to an exponential regression model. Weibull models can also be viewed as special cases of accelerated ( $\rho > 1$ ) or decelerated ( $\rho < 1$ ) failure time models. The Weibull model is a parametric proportional hazards model. Estimation of fixed and random effects and of genetic parameters is much less demanding with a Weibull model than with a semi-parametric model, such as the popular Cox model (Cox, 1972), for which the baseline hazard function is left arbitrary. The Weibull model is also a very flexible parametric model, in particular, when time-dependent covariates are included in the exponential part of [1]. Also, computational feasibility for very large datasets explain why Weibull frailty models are frequently used for the analysis of dairy cattle longevity data (Ducrocq, 1987; Ducrocq, 1994; Ducrocq and Casella, 1996; Vollema and Groen, 1998; Ducrocq, 1999a; Vukasinovic et al., 1999).

Longevity traits are not always collected on a scale that can be considered as continuous (e.g., length of life in days). Sometimes, a coarser scale (years or lactation) is used, which can only take a few distinct values.

Then, a survival analysis accounting for the discrete nature of the measure may seem preferable. A particular discrete-time model, the grouped data model (Prentice and Gloeckler, 1978), was studied here: the underlying (unobserved) continuous culling times, were grouped into intervals  $A_k = [a_{k-1}, a_k]$ ,  $k = 1, \dots, r$  with  $a_0 = 0$ , such that culling times were in fact observed on a discrete scale with  $r$  values. Censored times can be recorded on the same scale, indicating the latest interval that the cow is known to survive. Then, the (discrete) hazard function of an animal was modeled according to Ducrocq (1999b) and Ducrocq et al. (2001), extending the work of Prentice and Gloeckler (1978) to mixed models as follows:

For  $a_{k-1} \leq t < a_k$ :

$$\lambda_{ij}(t) = [1 - \alpha_k^{\exp(s_i)}] \prod_{l=1}^{k-1} \alpha_l^{\exp(s_i)} \tag{2}$$

where

$$\alpha_l = \exp\left\{-\int_{a_{l-1}}^{a_l} \lambda_0(u) du\right\} \tag{3}$$

is the conditional survival probability in  $A_l$  for an animal  $j$ , with sire  $i$  such that  $s_i = 0$ .  $\alpha_l$  is a simple function of the hazard integrated over the whole interval  $[a_{l-1}, a_k]$ , where  $u$  is a variable of integration representing time. The  $\alpha_l$  varies between 0 and 1, and it is more interesting to reparameterise this model into  $\xi_l = \ln(-\ln \alpha_l)$ , which can take any value in  $]-\infty, +\infty[$ . Equivalently,  $\alpha_l = \exp[-\exp(\xi_l)]$  and

$$\lambda_{ij}(t) = \left\{ 1 - \exp[-\exp(\xi_k + s_i)] \right\} \prod_{l=1}^{k-1} \exp[-\exp(\xi_l + s_i)] \tag{4}$$

We note that  $\{1 - \exp[-\exp(\xi_k + s_i)]\}$  is the probability  $p_{s_i}$  that a progeny of sire  $i$  alive at time  $a_{k-1}$  is dead by time  $a_k$ .

The binary data analysis is based on the following linear model,

$$y_{ij}(T) = \mu + s_i + \varepsilon_{ij} \tag{5}$$

where  $y_{ij}(T)$  is 0 if the  $j$ th progeny of sire  $i$  is not alive at time  $T$ , and 1 otherwise;  $\mu$  is the population mean, and  $s_i$  is the breeding value of sire  $i$  on this binary scale. The reason for including this model is that it is simple but makes many incorrect assumptions when the true survival times are continuous and nonnormally distributed. Therefore, it should give an indication on how much information is lost when failure time information is grouped and data is censored at

different times compared with a proportional hazard model. If the value of T is chosen such that the proportion of zeroes or ones is not too extreme, we expect that the analyses will be at most marginally affected by the use of a linear model instead of a more proper threshold model.

**Genetic parameters.** Most quantitative traits in animal genetics are approximated using normal distributions, and classical procedures can be used for the estimation of (functions of) genetic parameters (e.g., heritability). The parametric distributions used for longevity traits often come from the family of generalizations of exponential distributions (e.g., the Weibull distribution). Ducrocq (1987) derived the heritability on the logarithmic scale ( $h_{\log}^2$ ) for Weibull sire model as follows:

$$h_{\log}^2 = 4 \sigma_s^2 / (\sigma_s^2 + \pi^2/6) \quad [6]$$

Expression [6] is also the pseudo-heritability defined by Smith (1983, p. 125) and derived by Korsgaard et al. (1999) for Cox frailty models. The value of  $\pi^2/6$  in [6] is the variance of an extreme value distribution, the distribution of the residual term of log (t) in PH models. However, Ducrocq (1999a) showed that this heritability is not convenient for the calculation of reliability, leading to strongly biased values when the reliability (R) of a particular sire is computed using the usual reliability formula (for example by following index selection arguments) for sire evaluation:

$$R = \frac{n h^2}{(n-1)h^2 + 4} \quad [7]$$

where n is the total number of progeny of the sire and  $h^2 = h_{\log}^2$ . Moreover, Ducrocq (1999a) found that it appeared much more appropriate to use  $n_{\text{uncen}}$ , the number of *uncensored* progeny in [7] as well as the heritability of the trait transformed onto the original scale ( $h_{\text{ori}}^2$ ), for which he proposed an approximate formula:

$$h_{\text{ori}}^2 = (\exp(\nu/\rho))^{-2} h_{\log}^2 \quad [8]$$

where  $\nu = -\text{Euler's constant} = -0.5772$  and  $\rho$  is the shape parameter of the baseline Weibull distribution. Formula [8] was obtained as the ratio of the variance on the observed scale of the sire effect on one hand and of the total variance on the other hand. These variances were derived using the delta method from a Taylor series of the variables around their mean on the log scale. Note that this prediction is a function of the shape parameter  $\rho$ : for example, the heritability on the logarithmic scale for sire variances of 0.022 is

0.05, while the approximate heritability on the original scale given by formula [8] is 0.17 with  $\rho = 1.0$  (hazard constant over time), and 0.09 with  $\rho = 2.0$ , respectively (hazard increasing with time). The agreement between reliabilities computed from the true asymptotic standard errors of prediction and the approximate ones computed using selection index theory and formula [8] was excellent for  $\rho = 2.0$  in a real case situation (Ducrocq, 1999a). However, the relationship between  $h_{\text{ori}}^2$  in [8] and  $\rho$  is not easy to explain, and the formula did not work at all in some other situations characterized by  $\rho$  values close to 1 or less than 2 (Ducrocq, 1999a). In practice, a value of  $\rho = 2.0$  has been often used for the analysis of length of productive life (Ducrocq, 1999a; Schneider and Miglior, 1999; Buenger et al., 2001; Larroque and Ducrocq, 2001). Values in the range from 1 to 2 have also been reported (Ducrocq, 1994, 1999a; de Jong et al., 1999; Durr et al., 1999; Vukasinovic, 1999; Neerhof et al., 2000).

By using Taylor series approximations for the ratio of genetic variance to phenotypic variance, Korsgaard et al. (1999) suggested that heritability on the log scale is also appropriate on the original scale. These observations motivated the investigation of a more appropriate formula to be used for  $h^2$  in reliability calculations.

An alternative derivation using basic expressions of reliabilities in sire models suggests a simple equation for the heritability on the original scale (called hereafter effective heritability,  $h_{\text{eff}}^2$ ) that is not dependent on the Weibull parameters (see Appendix A),

$$h_{\text{eff}}^2 = 4 \sigma_s^2 / (\sigma_s^2 + 1). \quad [9]$$

For example, using a sire variance of 0.022 in this formula predicts a heritability of 0.09. The differences between these heritabilities and those from using the previous equation are small when  $\rho = 2.0$  (as in Ducrocq, 1999a) but can be huge when  $\rho$  is far from this special value such that the correcting term in [9] is close to  $\pi^2/6$ . The predicted reliability from the Weibull models ( $R_{\text{Wei}}$ ) is calculated using the number of uncensored progeny ( $n_{\text{uncen}}$ ) and the sire variance (Appendix A),

$$R_{\text{Wei}} = n_{\text{uncen}} / (n_{\text{uncen}} + 1/\sigma_s^2). \quad [10]$$

We further predict reliabilities when records are censored at time T as a function of n, the total number of progeny of a sire. If the proportion of its progeny that have not survived until a given time t is  $\bar{p}$ , an equivalent heritability ( $h_{\text{equ}}^2$ ) can be defined as the value of  $h^2$  such that the reliability can be computed



using the index selection formula [7] (with  $n$  instead of  $n_{uncen}$  as in [10]). Equating formula [7] with  $h^2 = h_{equ}^2$  to:

$$R_{Wei} = n\bar{p}/(n\bar{p} + 1/\sigma_s^2) \quad [11]$$

leads to (see also Appendix A for a full derivation):

$$h_{equ}^2 = 4 \sigma_s^2/(\sigma_s^2 + 1/\bar{p}). \quad [12]$$

**Binary analysis.** The binary data situation is a special case of a discrete measure of survival with only two possible values: 0 or 1, depending on whether the animal is alive at a given time (or, “threshold”). The genetic parameters from the linear model on binary data can be derived using a first-order Taylor series expansion applied to the hazard function given in equation [4]. When survival time follows a Weibull distribution, the relationship between survival to time  $T$  and breeding value  $s_i$  of sire  $i$  is described by a complementary log-log function (e.g., Madgwick and Goddard, 1989).

$$1 - E(y_{ij}(T)) = p_{s_i} = 1 - \exp[-\exp(\xi_{(T)} + s_i)] \quad [13]$$

where  $\{1-p_{s_i}\}$  is the probability that a progeny of sire  $i$  is still alive at time  $T$ ,  $\xi_{(T)} + s_i = \ln[-\ln(1 - p_{s_i})]$  and  $\xi_{(T)}$  is the complementary log-log function of the average conditional probability to survive at time  $T$ . Analysis on the 0/1-scale gives a phenotypic variance with expectation,  $\sigma_p^2 = \bar{p}(1 - \bar{p})$ , with  $\bar{p}$  the population mean, i.e., the proportion of animals culled by time  $T$ , and sire variance  $\sigma_{s(b)}^2$  obtained using the delta method:

$$\sigma_{s(b)}^2 = d^2 \sigma_s^2 \quad [14]$$

with  $d = (dp_{s_i}/ds_{s_i=0}) = -(1 - \bar{p})\ln(1 - \bar{p})$ . Then, the heritability on the binary scale ( $h_{bin}^2$ ) can be approximated as:

$$\begin{aligned} h_{bin}^2 &\approx \frac{4\sigma_{s(b)}^2}{\sigma_{s(b)}^2 + \bar{p}(1 - \bar{p})} \\ &= \frac{4\sigma_s^2}{\sigma_s^2 + \frac{\bar{p}}{(1 - \bar{p})[\ln(1 - \bar{p})]^2}}. \end{aligned} \quad [15]$$

From this, the reliability on the binary scale is predicted (using, for example, selection index theory) as:

$$R_{bin} = n/[n + (4 - h_{bin}^2)/h_{bin}^2]. \quad [16]$$

This reliability is in terms of the total number of progeny ( $n$ ) and the heritability on the binary scale, which is a function of the time  $t$  when the binary trait is measured (through  $\bar{p}$ ). We can also express the reliability as a function of the sire variance ( $\sigma_s^2$ ) and an effective number of progeny ( $n_{e(b)}$ ),

$$R_{bin} = n_{e(b)}/(n_{e(b)} + 1/\sigma_s^2), \quad [17]$$

where

$$n_{e(b)} = n\{d^2/[\bar{p}(1 - \bar{p})]\} \quad [18]$$

Obviously, the calculation of heritability on the binary scale in [15] or the effective number of progeny in [18] are dependent on the proportion of animals culled before time  $T$ . The heritability increases with increasing proportion of animals culled until the proportion is about 0.8 and then it declines (setting the first derivative of [15] with respect to  $\bar{p}$  to zero results in solving the equation  $\ln(1 - \bar{p}) + 2\bar{p} = 0$ ).

**Discrete analysis.** Ducrocq (1999b) showed that the likelihood analysis for a Weibull model can be easily tailored to analyze discrete data using Prentice and Gloeckler’s (1978) grouped data model. The derivation of the reliability for the grouped data model is similar to the binary case. When there are  $k$  discrete intervals  $A_j = [a_{j-1}, a_j]$ , this leads to  $k$  distinct possible values of survival time. In this paper, we consider the cases in which survival times greater than  $a_k$  are considered censored at  $a_k$  leading to effectively  $k + 1$  groups.

If  $p_k$  is the proportion of the population that does not survive until  $a_k$ , the end of the  $k$ th interval (with  $p_0 = 0$ ), one gets:

$$R_{dis_k} = n_{e(dis_k)}/(n_{e(dis_k)} + 1/\sigma_s^2) \quad [19]$$

where:

$$n_{e(dis_k)} = n \sum_{i=1}^k \{(1 - p_{i-1})d_i^2/[c_i(1 - c_i)]\} \quad [20]$$

with

$$c_i = (p_i - p_{i-1})/(1 - p_{i-1}) \quad [21]$$

and

$$d_i = -(1 - c_i)\ln(1 - c_i). \quad [22]$$

This effective number  $n_{e(dis_k)}$  has  $k$  components, one for each time grouping, each of the same form as  $n_{e(b)}$  but with a multiplier  $(1 - p_{i-1})$  which indicates the proportion of animals surviving until  $a_i$ ;  $c_i$  is the conditional probability of being culled between  $a_{i-1}$  and  $a_i$

given survival until  $a_{i-1}$  (i.e., a discrete version of the hazard rate at 'time'  $i$ ).

Expanding  $d_i$  as a Taylor series suggests that

$$d_i^2/[c_i(1 - c_i)] \approx c_i[1 - c_i^2(1 + c_i)/12] \quad [23]$$

is an excellent approximation (proportional error <1%) for small  $c_i$  ( $c_i < 0.4$ ). For example, for one group, the reliability is predicted as,

$$n_{e(\text{dis}_1)} = np[1 - \bar{p}^2(1 + \bar{p})/12]. \quad [24]$$

In the Weibull analysis until time  $T$ , the number of uncensored progeny is  $n\bar{p}$ , and, therefore, to a first-order approximation, the loss of 'progeny' by performing a discrete analysis is approximately  $n[\bar{p}^2(1 + \bar{p})/12]$ , which is small for small  $\bar{p}$ . Note that  $\bar{p}$  is defined as the proportion of animals that did not survive in a particular time period, and that a small value of  $\bar{p}$  is relevant for dairy cattle data when we consider a short period, for example, one lactation or less.

For a single group only, the resulting reliability is equivalent to the reliability for the binary analysis (see results above), and so the heritability becomes, approximately,

$$h_{\text{dis}_1}^2 = 4\sigma_s^2/[\sigma_s^2 + \{\bar{p}(1 - \bar{p}^2(1 + \bar{p})/12)\}^{-1}]$$

**Response to selection.** For the continuous (Weibull) case, we derive in Appendix B the prediction of the average lifespan in a base population of progeny from unselected sires. To predict the average lifespan of progeny of selected sires, a simple approximation suggests, for mean breeding value and assuming a small variance of the selected sires ( $\sigma_s^{2*}$  in Appendix B),

$$E(t)_{\text{offspring}} \approx \int_0^{\infty} tf(t)\exp(\overline{BV}) dt \approx \lambda^{-1} \Gamma[1 + (1/\rho)] \exp(-\overline{BV}/\rho) \quad [25]$$

where  $t$  is the lifespan of the animal, and  $f(\cdot)$  is the density function of a Weibull distribution with parameters  $\lambda$  and  $\rho$ . The average breeding value can be predicted using  $\overline{BV} = r \times i \times \sigma_s$ , with  $r$  being the square root of reliability ( $R$ ) based on the sire variance in the Weibull frailty model and the number of progeny per sire;  $i$  is the selection intensity, and  $\sigma_s$  is the square root of the sire variance. For unselected sires, the prediction in Appendix B reduces to:

$$E(t)_{\text{population}} \approx \lambda^{-1} \Gamma[1 + (1/\rho)] \exp(\sigma_s^2/2\rho^2). \quad [26]$$

Hence, using the above approximations, a simple prediction of reliability and response to selection in the Weibull frailty models can be made that only uses the sire variance, the number of uncensored progeny per sire, and the population parameters of the Weibull distribution. For either the binary or the discrete analysis models, assuming that the data are discrete measures from an underlying continuous time scale with a Weibull hazard distribution, we will use the same predictions as above. The only parameter that changes from the Weibull model prediction is the reliability.

## Simulation

Data roughly similar to length of productive life of dairy cows were simulated using the Weibull frailty model described in [1]. To consider a plausible survival rate among cattle populations, two different  $\rho$  values of 1.0 (constant hazard), and 2.0 (increasing hazard) were used. This range includes all rates (except one) found in the literature (Ducrocq, 1994, 1999a; de Jong et al., 1999; Durr et al., 1999; Schneider and Miglior, 1999; Vukasinovic, 1999; Neerhof et al., 2000; Buenger et al., 2001; Larroque and Ducrocq, 2001). These values and an arbitrarily chosen median time of 730 were used as the Weibull parameters for simulation. Using formulae B-5 in Appendix B, the corresponding expected values of the simulated longevity measure for the two sets of values are 1053 and 777 d, respectively, i.e., between two and three lactations. Breeding values of 100 unrelated sires were generated from a normal distribution with  $\mu = 0$  and two different sire variances (either 0.022 or 0.09), indicating a low and high genetic variation  $h_{\text{eff}}^2 = 0.086$  to 0.33 with the proposed formula, covering a slightly wider range than the one reported in the literature (see Ducrocq, 1994, 1999a; Schneider and Miglior, 1999; Vukasinovic et al., 1999; Neerhof et al., 2000; Buenger et al., 2001).

The total number of progeny per sire ranged from 50 (without censoring) to 100 (with censoring). In the case of censoring, three levels of random censoring (20, 50, and 80%) were simulated for a total number of 100 progeny, resulting in, on average, 80, 50, and 20 uncensored progeny per sire. To estimate the empirical reliability of predicted breeding values, we simulated 20 progeny groups for each sire.

For analysis of longevity data (similar to length of productive life here) on the discrete scale, data were simulated for 100 sires with 100 progeny each in different combinations of sire variances and the Weibull parameter  $\rho$ , i.e., assuming that the underlying true distribution of lifespan is Weibull. The longevity measure was transformed to a discrete scale using arbitrary end of discrete intervals at (depending on the

analysis) 400, 800, or 1200 d. These values were chosen to roughly approximate survival for 1, 2, or 3 lactations and correspond to a high, moderate, and low mean survival probabilities. In the case of two discrete classes, the discrete value of an individual's longevity was either 0 (when longevity was less than the end of interval) or 1 (when longevity was equal or greater than the end of interval). These data were analyzed both with a linear model and with the grouped data model. An additional analysis using discrete data with values 0, 1, 2, and 3, corresponding to culling in three periods using end of intervals of 400, 800, and 1200 d and surviving for over 1200 d, respectively, was also performed to check the derived reliabilities for discrete analyses and evaluate the advantage in using four classes instead of two.

To calculate parameters for response to selection, additional progeny of selected sires were simulated and their mean failure (culling) time calculated. The number of replicates was 100 for all analyses. All simulations and estimations of EBV for the proportional hazard models and for the grouped data model based on two or four classes (Prentice and Gloeckler, 1978; Ducrocq, 1999b) were performed using the Survival Kit software (Ducrocq and Sölkner, 1998), which uses a Bayesian approach and a Laplacian approximation. The analysis of binary data (discrete data with two classes) using a linear model was performed by constructing a one-way analysis of variance using SAS software (1990) and also authors' Fortran program.

## RESULTS

### Reliability

Estimates of observed reliabilities, using survival time as the dependent variable in the Weibull model, in the different scenarios (combinations of two different sire variances, two different Weibull parameters  $\rho$ , and four different censoring rates) and calculated reliabilities from the proposed formula for estimation of heritability are presented in Table 1. The observed reliability increased from 0.31 to 0.64 and from 0.65 to 0.88 (on average) when the number of uncensored progeny increased for the low and high sire variance, respectively. With the same total number of progeny (100) but with increasing number of uncensored records from about 20 to 80 on average, the reliability almost doubled in the case of low sire variance and it was increased by about 60% with high sire variance. The different  $\rho$  values did not affect the observed reliability either in low or high sire variances, confirming that this parameter should not appear in the formulae. The predictions of reliabilities using the proposed formulae [9 and 10] in this study are in very good

agreement with observed reliabilities (Table 1), and clearly superior to [7].

Results of reliability estimates from the linear model and grouped data model analysis on discrete data are presented in Table 2. Given the number of uncensored records (ranging from 19 to 84), the range of observed reliabilities was between 0.28 and 0.64 and between 0.63 and 0.88 for low and sire variances, respectively. There was the same pattern as the Weibull model estimates between estimates of reliabilities from low or high sire variance. Reliability in the discrete model was higher when survival time was transformed to several time intervals (in the case of four discrete values corresponding to end of intervals of 400, 800, and 1200 d, respectively). In general, the estimates of reliability were close to estimates from the Weibull model for a given number of uncensored records. This may indicate that the underlying (unobserved) hazard distribution was approximately Weibull and that little information was lost by going from a continuous time scale to a scale of the number of lactations.

The estimates of reliability from the linear model analysis on discrete data with two classes are very close to the results of the grouped data model for those data, and ranged from 0.23 to 0.57 and from 0.63 to 0.84 for low and sire variances, respectively. It should be noted that the number of progeny records per sire in the binary case is the same as the number of uncensored progeny records in the discrete model analysis (no censoring in the latter case). As the results show, the reliability increases when the number of uncensored progeny records per sire (i.e., the number of animals that are no longer alive at the end of interval point) increases. In this approach, the genetic variation among sires increases and reaches its highest value at the average proportion of culled progeny of approximately 0.8, as predicted from differentiating equation [15] with respect to the average proportion culled (derivation not shown). In general, when the underlying survival data is from a Weibull distribution, it appears that the method of analyzing data does not greatly affect the results in terms of sire ranking or response to selection, at least for the simplified situation considered in this study.

### Selection Response

Observed and expected responses from different combinations of sire variances and  $\rho$  values are presented in Table 3 when a Weibull analysis is carried out, in a case of no censoring. The observed population means ranged from 779 to 1065 d and from 788 to 1098 d when  $\rho$  varies from 1.0 to 2.0 in low and high sire variance situations, respectively. The predicted value

**Table 1.** Estimates of reliability<sup>1</sup> (R, mean  $\pm$ SE<sup>2</sup>) for different sire variances ( $\sigma_s^2$ ), Weibull shape parameter ( $\rho$ ) and number of uncensored progeny records per sire in the Weibull model.

	Expected level of censoring (%)	n <sup>3</sup>	1.0	2.0	R <sub>exp</sub> <sup>4</sup>
$\sigma_s^2 = 0.022$	00.0	50.0	0.524 <i>0.004</i>	0.525 <i>0.004</i>	0.524
	20.0	79.8	0.636 <i>0.003</i>	0.630 <i>0.004</i>	0.637
	50.0	50.2	0.531 <i>0.004</i>	0.523 <i>0.004</i>	0.525
	80.0	20.2	0.313 <i>0.003</i>	0.309 <i>0.004</i>	0.308
$\sigma_s^2 = 0.09$	00.0	50.0	0.815 <i>0.002</i>	0.817 <i>0.002</i>	0.818
	20.0	79.1	0.876 <i>0.002</i>	0.874 <i>0.002</i>	0.877
	50.0	50.5	0.815 <i>0.002</i>	0.820 <i>0.002</i>	0.820
	80.0	20.7	0.656 <i>0.003</i>	0.647 <i>0.004</i>	0.651

<sup>1</sup>Observed reliability (average over 100 replicates), calculated as the intra-class correlation between estimated sires breeding values from 20 progeny groups.

<sup>2</sup>The SE are presented in *italic* form.

<sup>3</sup>n = Average number of uncensored progeny records/sire.

<sup>4</sup>Expected reliability ( $R_{Wei} = n_{uncen}/(n_{uncen} + 1/\sigma_s^2)$ ).

of the overall mean depended on sire variance (equation [26]) and ranged from 781 to 1102 d for different  $\rho$  values and sire variances. Breeding values of selected sires are presented in sire standard deviations and are negative for favorable bulls, i.e., bulls with daughters that have their hazard function in [1] inferior to the baseline hazard. The observed average EBV for the top 5 and 10% selected bulls were in close agreement with the expected values of  $-1.67$  and  $-1.43$  for low and  $-1.91$  and  $-1.64$  for high sire variances, respectively. There is a large increase in the mean longevity of progeny of selected bulls, whatever the sire variances and  $\rho$  values. The (observed) selection responses varied from about 86 to 282 d, depending on selection intensity and the distribution of survival time for low sire variance. The corresponding responses are higher for high sire variance than for low sire variance, as expected. The observed and expected selection responses are not statistically different in both cases (low and high sire variances). For the low variance value (lower than the estimates found in the literature for dairy cattle),  $\rho$  equal to 2 and a selection intensity of 10%, a response of about 3 mo is rather promising for improvement of longevity in dairy cattle.

Selection responses (observed and expected) from a linear model analysis on two discrete data classes (i.e., binary data) from different combinations of sire variances and  $\rho$  values for the Weibull distribution used

to simulate the discrete data are presented in Table 4. Breeding value of sires and selection responses were estimated at two different ends of interval points (400 and 1200 d). Observed EBV of sires are less (underestimated) than expected EBV at 400-d evaluation. The corresponding EBV are higher (overestimated) than expected at 1200 d evaluation. Similar to response to selection in the Weibull model (Table 3), there is a wide range of selection responses depending on sire variance, underlying distribution of survival time and truncation time points. There are no significant differences between the observed and expected responses at the end of the 400-d interval point for low and high sire variances. However, there are relatively large differences between the observed and expected responses at the end of the 1200-d interval point, where differences were significant when parameter  $\rho$  was equal to two. Although the selection response from the binary data analysis depends on the end of interval point, there is a relatively good agreement between selection responses in the Weibull model and binary data analysis.

Results for the Weibull frailty model with systematic censoring at 400 and 1200 d similar to the binary analysis truncation time points in the case of low sire variance (0.022) are presented in Table 5. Observed and expected EBV are close at both censoring times. These results are comparable to results in the binary data



**Table 2.** Observed<sup>1</sup> (mean<sup>2</sup> ±SE) and expected (exp) reliabilities<sup>3</sup> for discrete-time proportional hazard model (obs\_dis) and binary data analysis (obs\_bin).

		$\rho$					
		1.0			2.0		
Threshold time point <sup>4</sup>		obs_dis	obs_bin	Exp	obs_dis	obs_bin	Exp
$\sigma_s^2 = 0.022$	400	0.420	0.415	0.407	0.298	0.275	0.292
		<i>0.004</i>	<i>0.008</i>		<i>0.004</i>	<i>0.010</i>	
		32			19		
	800	0.534	0.526	0.527	0.543	0.536	0.540
		<i>0.004</i>	<i>0.006</i>		<i>0.004</i>	<i>0.007</i>	
		53			57		
	1200	0.578	0.567	0.573	0.583	0.564	0.584
		<i>0.004</i>	<i>0.006</i>		<i>0.004</i>	<i>0.007</i>	
		68			84		
	400,800,1200	0.603		0.597	0.641		0.640
		<i>0.003</i>			<i>0.003</i>		
		68			84		
$\sigma_s^2 = 0.09$	400	0.733	0.737	0.738	0.634	0.632	0.628
		<i>0.003</i>	<i>0.004</i>		<i>0.004</i>	<i>0.006</i>	
		32			20		
	800	0.813	0.812	0.820	0.824	0.819	0.828
		<i>0.002</i>	<i>0.003</i>		<i>0.002</i>	<i>0.003</i>	
		54			57		
	1200	0.838	0.834	0.846	0.842	0.838	0.852
		<i>0.002</i>	<i>0.002</i>		<i>0.002</i>	<i>0.002</i>	
		68			84		
	400,800,1200	0.852		0.858	0.877		0.879
		<i>0.002</i>			<i>0.002</i>		
		68			84		

<sup>1</sup>The vertical triplets comprise observed reliability, standard error (in *italic*), and number of uncensored records out of 100 progeny records, respectively, estimated from simulation.

<sup>2</sup>Mean over 100 replicates.

<sup>3</sup>Expected reliability  $R_{dis_k} = n_{e(dis_k)} / (n_{e(dis_k)} + 1/\sigma_s^2)$ .

<sup>4</sup>The longevity at end of interval times 400, 800 and 1200 d were assigned to 1 and in the case 400,800,1200 the longevity values varied from 1 to 3 depending on the survival time. The record of an animal was uncensored if the survival time was less than the end of interval time.

analysis for low sire variance. This is a proper comparison between the Weibull model and the binary data analysis when observations are restricted to the end of the first lactation.

The relationship between time and heritability on the binary scale for two different sire variances (0.022 and 0.09) by using [15] is illustrated in Figure 1. When time increases, the heritability also increases and reaches its maximum values 0.056 and 0.220 at about 1100 d corresponding to a proportion of uncensored records (or of 0 on the binary scale when the animal is not alive) of ~0.8 for low and high sire variances, respectively. The equivalent heritability from the binary data analysis depends on the reliability on the binary scale. As the proportion of animals culled increases, the equivalent heritability increases. The rate of increase is reduced as the proportion of animals culled exceed about 50% (~730 d) and reaches asymptotically the effective heritability (formula [9]) on the observed scale for the Weibull distribution.

## DISCUSSION

In this study, the relationship between heritability and reliability was examined for proportional hazard models. The estimation of *effective* heritability used here was different from what was proposed by Ducrocq (1999a) as the heritability on the original scale. In the scenarios considered here, the results of observed reliabilities from survival analysis using simulation data are in very good agreement with the calculation of reliability using the proposed effective heritability. This estimation depends only on the sire variance and the number of uncensored records. This new proposed heritability is clearly superior to the one proposed of by Ducrocq (1999a) and its use should be recommended for national evaluations of longevity using proportional hazard methodology.

The simulations look at a limited number of scenarios, with equal numbers of daughters, no fixed effects, and a limited censoring regime. The simulations were

**Table 3.** Observed (obs, mean<sup>1</sup> ± SE<sup>2</sup>) and expected (exp) responses to selection<sup>2</sup> with different sire variances ( $\sigma_s^2$ ), the shape parameter of underlying Weibull distribution ( $\rho$ ), and selection intensities.

$\sigma_s^2$ $\rho$	0.022				0.09			
	1.0		2.0		1.0		2.0	
	Obs	exp	Obs	exp	Obs	exp	Obs	exp
Population mean <sup>4</sup>	1065 ± 2	1065	779 ± 1	781	1098 ± 3	1102	788 ± 1	795
EBV (5% top)	-1.664 ± 0.030	-1.673	-1.646 ± 0.024	-1.673	-1.925 ± 0.024	-1.914	-1.847 ± 0.023	1.914
EBV (10% top)	-1.421 ± 0.023	-1.435	-1.404 ± 0.020	-1.435	-1.669 ± 0.019	-1.641	-1.594 ± 0.017	-1.641
Future progeny mean of selected bulls (5% top) <sup>5</sup>	1346 ± 9	1350	882 ± 3	880	1884 ± 19	1870	1031 ± 5	1036
Future progeny mean of selected bulls (10% top) <sup>5</sup>	1307 ± 6	1303	865 ± 2	864	1744 ± 13	1723	993 ± 3	994
Response in days (selected 5% top)	282 ± 9	285	103 ± 3	98	786 ± 18	769	243 ± 5	241
Response in days (selected 10% top)	243 ± 6	238	86 ± 2	83	646 ± 11	622	205 ± 3	199

<sup>1</sup>Mean over 100 replicates.

<sup>2</sup>The SE figures are shown in *italic*.

<sup>3</sup>Population mean, future progeny means, and their standard errors are presented in days.

<sup>4</sup>Calculated from  $E(t)_{\text{population}} \approx \lambda^{-1} \Gamma [1 + (1/\rho)] \exp((\sigma_s^2/2\rho^2)$ .

<sup>5</sup>Calculated from  $E(t)_{\text{offspring}} \approx \lambda^{-1} \Gamma [1 + (1/\rho)] \exp(-BV/\rho)$ .

carried out because it seemed important to clarify the important principles raised, given the complexity and disagreement in the literature. Again because of the complexity, simpler estimation procedures including discrete and binary analyses were also used, in an attempt to better understand the results. These were thought to be successful especially in quantifying information from censored records and showing the potentially small loss from using simplified analyses.

Expansions to deal with fixed effects, genetic relationships, and other correlated traits could perhaps build on this work and approximations suggested for prediction error variances in similar circumstances for linear models. The Weibull distribution was used because it is commonly used in modeling survival data and heritabilities for this type of data have been suggested. United Kingdom dairy cattle survival data have unfortunately only been available in discrete

**Table 4.** Observed (obs<sup>1</sup>, mean ± SE<sup>2</sup>) and expected (exp) responses to selection with different sire variance ( $\sigma_s^2$ ), the shape parameter of underlying Weibull distribution ( $\rho$ ), and selection intensities for binary data analysis at end of interval times 400 and 1200 d.

$\sigma_s^2$ $\rho$	0.022				0.09			
	1.0		2.0		1.0		2.0	
	obs	exp	Obs	exp	obs	exp	Obs	Exp
400 d								
EBV (selected 5% top)	-1.233 ± 0.022	-1.288	-0.970 ± 0.024	-1.090	-1.687 ± 0.025	-1.733	-1.518 ± 0.019	-1.599
EBV (selected 10% top)	-1.058 ± 0.020	-1.104	-0.826 ± 0.022	-0.934	-1.441 ± 0.019	-1.486	-1.300 ± 0.017	-1.370
Response in days <sup>3</sup> (selected 5% top)	223 ± 9	210	60 ± 3	61	642 ± 19	670	182 ± 5	193
Response in days (selected 10% top)	185 ± 6	176	53 ± 2	51	541 ± 12	543	160 ± 4	160
1200 d								
EBV (selected 5% top)	-1.571 ± 0.019	-1.528	-1.645 ± 0.023	-1.542	-1.870 ± 0.024	-1.856	-1.896 ± 0.021	-1.862
EBV (selected 10% top)	-1.335 ± 0.016	-1.310	-1.412 ± 0.019	-1.322	-1.595 ± 0.020	-1.591	-1.637 ± 0.016	-1.596
Response in days (selected 5% top)	269 ± 8	256	100 ± 3	90	741 ± 18	736	243 ± 5	233
Response in days (selected 10% top)	218 ± 6	214	83 ± 2	76	606 ± 11	596	206 ± 3	193

<sup>1</sup>Mean over 100 replicates which are estimated from simulation.

<sup>2</sup>The SE figures are shown in *italic* format.

<sup>3</sup>See Table 3 for calculation.

**Table 5.** Observed (obs<sup>1</sup>, mean ± SE<sup>2</sup>) and expected (exp) responses to selection with sire variance ( $\sigma_s^2 = 0.022$ ) and two different shape parameter of underlying Weibull distribution ( $\rho$ ), and selection intensities in the Weibull model with censoring at end of interval times 400 and 1200 d.

$\rho$	1.0		2.0	
	obs	Exp	obs	Exp
400 d				
EBV (selected 5% top)	-1.236 ± 0.110	-1.292	-1.008 ± 0.092	-1.091
EBV (selected 10% top)	-1.063 ± 0.078	-1.108	-0.877 ± 0.074	-0.936
Response in days (selected 5% top) <sup>3</sup>	220 ± 9	211	56 ± 3	63
Response in days (selected 10% top) <sup>3</sup>	186 ± 7	176	48 ± 2	54
1200 d				
EBV (selected 5% top)	-1.543 ± 0.154	-1.562	-1.616 ± 0.158	-1.628
EBV (selected 10% top)	-1.327 ± 0.116	-1.339	-1.384 ± 0.113	-1.395
Response in days (selected 5% top) <sup>c</sup>	273 ± 9	263	97 ± 3	98
Response in days (selected 10% top) <sup>3</sup>	236 ± 7	220	82 ± 2	83

<sup>1</sup>Mean over 100 replicates which are estimated from simulation.

<sup>2</sup>The SE figures are shown in *italic* form.

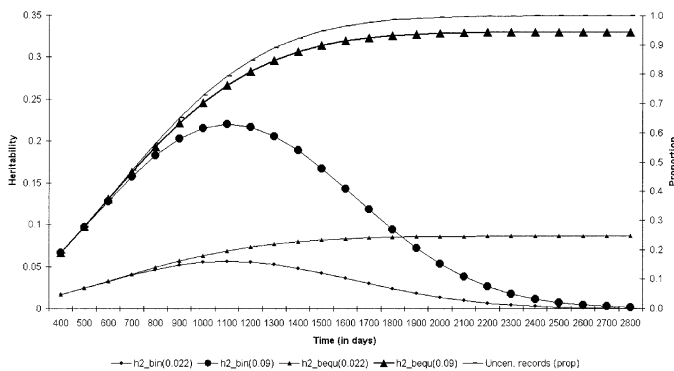
<sup>3</sup>See Table 3 for calculation.

form, so there was little impetus to investigate Cox models. In one sense, these models are the limiting case of discrete models if the  $\alpha_i$  parameters in the discrete model are considered as independent parameters not constrained by a Weibull distribution.

The introduction of animal models, while theoretically de rigueur, still seems problematic in that approximations are not altogether satisfactory and Markov Chain Monte Carlo techniques that have been developed are not yet computationally feasible to be in routine use for animal evaluation.

Previous expressions of the heritability for survival traits have been on both a logarithmic and original scale (e.g., Ducrocq and Casella, 1996). One intuitive explanation of the difference between the suggested heritability of  $h_{\text{eff}}^2 = 4 \times \sigma_s^2 / (\sigma_s^2 + 1)$  and the previously derived  $h_{\text{log}}^2 = 4 \times \sigma_s^2 / (\sigma_s^2 + \pi^2/6)$  (Ducrocq, 1987) is that these heritabilities are relevant to breeding values estimated in different ways on different scales. The latter heritability is relevant to a linear analysis on a log scale, i.e., transforming the data to a log scale and assuming normality of the random effects. This, however, is not the most efficient estimation procedure (in the sense of how well the data are utilized in estimating parameters) and taking fully account of the Weibull distribution leads to a more efficient analysis. In fixed effects analysis, Cox and Hinkley (1968) show that the linear analysis has efficiency of  $6/\pi^2$ . So, replacement of  $\pi^2/6$  by 1 could be therefore thought of a consequence of carrying out a more efficient analysis.

Several statistical approaches have been examined for the analysis of survival time. As far as heritability or reliability is concerned, the differences that were found here between the analysis of a continuous measure of longevity and a discrete (possibly binary) one do not appear large. Yazdi et al. (1999) analyzed the number of lactation records on UK dairy cattle data using the grouped data model and the Weibull model, and found high correlations between predicted breeding values of bulls. However, these results do not imply that proportional hazards models are an unnecessarily complex methodological refinement: the simulated sit-



**Figure 1.** Relationship between time and heritability on binary scale<sup>1</sup> and equivalent<sup>2</sup> heritability for two different sire variances (0.022 and 0.09).

$${}^1h_{\text{bin}}^2\{\sigma_s^2\} \approx \frac{4\sigma_s^2(b)}{\sigma_s^2(b) + \bar{p}(1 - \bar{p})} = \frac{4\sigma_s^2}{\sigma_s^2 + \frac{\bar{p}}{(1 - \bar{p})\ln(1 - \bar{p})^2}}$$

$${}^2h_{\text{bequ}}^2\{\sigma_s^2\} = 4\sigma_s^2/(\sigma_s^2 + 1\bar{p}).$$

uations considered here were idealized (no fixed effect other than the mean, no time-dependent covariate, nearly balanced design). It has been documented that survival analysis has advantages over linear model methodology (e.g., Ducrocq, 1987; Vukasinovic et al., 1999). It uses both information of those animals that were culled (uncensored) as well as records of those animals that are still alive and productive (censored). It also allows for changes in the culling policy and environmental factors to be accounted for in the survival analysis by treating them as time-dependent covariates in the proportional hazard model. However, for our highly simplified case of a balanced design, no fixed effects, no time-dependent covariates, equal expectation of the number of censored records per sire, and assuming a Weibull underlying distribution of the survival trait, it appears the method of analyzing survival data (either using the Weibull or grouped data models) does not greatly influence the results in terms of sire ranking.

Classical selection response depends on the genetic variance and selection intensity. In survival analysis, the shape of the survival function affects the genetic variance and hence the response to selection. A value of  $\rho$  above 1 in the Weibull model indicates that the baseline hazard rate increases with time. In dairy cattle, it corresponds to the risk of culling and it increases as a cow gets older because she becomes more subject to the degenerative effects of aging. Values of  $\rho$  in the range indicated here (1 to 2) are used in several countries (Ducrocq and Sölkner, 1998; de Jong et al., 1999; Ducrocq, 1999a; Durr et al., 1999; Pasman and Reinhardt, 1999; Pedersen, 2001, personal communication, for Denmark; Schneider and Miglior, 1999; Vukasinovic et al., 1999) using survival analysis approach for estimation of sire breeding values for longevity (e.g., Ducrocq and Sölkner, 1998; Vukasinovic et al. 1999). However, in these analyses, time-dependent covariates such as stage of lactation  $\times$  lactation number are fitted, whereas in our simulations and analyses time-dependent covariates were not included. Our predictions are appropriate for the situation when the overall hazard can be approximated by a Weibull curve with the value of  $\rho$  estimated from the overall survival curve ignoring time-dependent covariates. By considering  $\rho$  values between 1.0 and 2.0 and selection of the top 5%, the expected responses would be between the range of 3 to 8 mo depending on the sire variance. We used extreme values of sire variances and intensity of selection in the simulation and prediction of selection response to test the accuracy of the nonlinear predictions, rather than to imply a practical scenario. Although, in practice, selection of dairy bulls and cows will be on multiple traits, with a large emphasis on

milk production, the values of 3 to 8 mo point to what selection response could be achieved if all selection pressure was on survival. It shows that despite the low heritability of survival traits, a relatively large response to selection could be achieved.

## REFERENCES

- Allaire, F. R., and J. P. Gibson. 1992. Genetic value of herd life adjusted for milk production. *J. Dairy Sci.* 75:1349–1356.
- Boettcher, P. J., L. K. Jairath, and J. C. M. Dekkers. 1999. Comparison of methods for genetic evaluation of sires for survival of their daughters in the first three lactations. *J. Dairy Sci.* 82:1034–1044.
- Brotherstone, S., R. F. Veerkamp, and W. G. Hill. 1998. Predicting breeding values for herd life of Holstein-Friesian dairy cattle from lifespan and type. *Anim. Sci.* 67:405–411.
- Buenger, A., V. Ducrocq, and H. H. Swalve. 2001. An analysis of survival in dairy cows using supplementary data on type scores and housing systems from a region of North West Germany. *J. Dairy Sci.* 84:1531–1541.
- Cox, D. R. 1972. Regression models and life tables (with discussion). *J. R. Stat. Soc. Ser. B* 34:187–220.
- Cox, D. R., and D. V. Hinkley. 1968. A note on the efficiency of least squares. *J. R. Stat. Soc. B* 30:284–289.
- de Jong, G., A. R. Vollema, S. van der Beek, and A. Harbers. 1999. Breeding value for functional longevity in the Netherlands. Page 68 in *Proc. Int. Workshop on Genetic Improvement of Functional Traits in cattle, Longevity, Jouy-en-Josas, France. INTERBULL Bull. No. 21. Int. Bull Eval. Serv., Uppsala, Sweden.*
- Ducrocq, V. 1987. An analysis of productive life in dairy cattle. Ph.D. Diss., Cornell Univ., Ithaca, NY.
- Ducrocq, V. 1994. Statistical analysis of length of productive life for dairy cows in the Normande breed. *J. Dairy Sci.* 77:855–865.
- Ducrocq, V. 1999a. Two years of experience with the French genetic evaluation of dairy bulls on production-adjusted longevity of their daughters. Page 60 in *Proc. Int. Workshop on Genetic Improvement of Functional Traits in cattle, Longevity, Jouy-en-Josas, France. INTERBULL Bull. No. 21. Int. Bull Eval. Serv., Uppsala, Sweden.*
- Ducrocq, V. 1999b. Extension of survival analysis models to discrete measures of longevity. Page 41 in *Proc. Int. Workshop on Genetic Improvement of Functional Traits in Cattle, Longevity, Jouy-en-Josas, France. INTERBULL Bull. No. 21. Int. Bull Eval. Serv., Uppsala, Sweden.*
- Ducrocq, V., and G. Casella. 1996. A Bayesian analysis of mixed survival models. *Genet. Sel. Evol.* 28:505–529.
- Ducrocq, V., and J. Sölkner. 1998. "The Survival Kit V3.0", a package for large analyses of survival data. Pages 447–450 in *Proc. 6th World Congr. on Genetics Applied to Livestock Production, Armidale, Vol. 27.*
- Durr, J. W., H. Monardes, and R. I. Cue. 1999. Genetic analysis of herd life in Quebec Holsteins using Weibull models. *J. Dairy Sci.* 82:2503–2513.
- Henderson, C. R. 1975. Rapid method for computing the inverse of a relationship matrix. *J. Dairy Sci.* 58:1727.
- Jairath, L., J. C. M. Dekkers, L. R. Schaeffer, Z. Liu, E. B. Burnside, and B. Kolstad. 1998. Genetic evaluation of herd life in Canada. *J. Dairy Sci.* 81:550–562.
- Kalbfleisch, J. D., and R. L. Prentice. 1980. *The Statistical Analysis of Failure Time Data.* John Wiley & Sons, New York.
- Korsgaard, I. R., A. H. Andersen, and J. Jensen. 1999. Discussion of heritability of survival traits. Page 31 in *Proc. Int. Workshop on Genetic Improvement of Functional Traits in cattle, Longevity, Jouy-en-Josas, France. INTERBULL Bull. No. 21. Int. Bull Eval. Serv., Uppsala, Sweden.*
- Larroque, H., and V. Ducrocq. 2001. Relationship between type and longevity. *Genet. Sel. Evol.* 33:39–59.



Madgwick, P. A., and M. E. Goddard. 1989. Genetic and phenotypic parameters of longevity in Australian dairy cattle. *J. Dairy Sci.* 72:2624–2632.

Neerhof, H. J., P. Madsen, V. Ducrocq, A. R. Vollema, J. Jensen, and J. I. R. Korsgaard. 2000. Relationships between mastitis and functional longevity in Danish black and white dairy cattle using survival analysis. *J. Dairy Sci.* 83:1064–1071.

Pasman, E., and F. Reinhardt. 1999. Genetic evaluation for length-of-productive life of Holstein Cattle in Germany. Page 55 in Proc. Int. Workshop on Genetic Improvement of Functional Traits in cattle, Longevity, Jouy-en-Josas, France. INTERBULL Bull. No. 21. Int. Bull Eval. Serv., Uppsala, Sweden.

Prentice, R. L., and L. A. Gloeckler. 1978. Regression analysis of grouped survival data with application to breast cancer data. *Biometrics* 34:57–67.

SAS/STAT User's Guide, Version 6, Fourth edition, Volume 1: Statistics, Version 6. 1996. SAS Inst., Inc., Cary, NC.

Schneider, M. P., and F. Miglior. 1999. A proposal for genetic evaluation for functional herd life in Italian Holstein. Page 11 in 50th Annu. Mtg. of the European Association for Animal Production, Zurich, Switzerland. (Abstr.)

Smith, S. P. 1983. The extension of failure time analysis to problems of animal breeding. Ph.D. Diss., Cornell Univ., Ithaca, NY.

Smith, S. P., and R. L. Quaas. 1984. Productive life span of bull progeny groups: Failure time analysis. *J. Dairy Sci.* 67:2999–3007.

Strandberg, E. 1996. Breeding for longevity in dairy cows. Page 125 in Progress in Dairy Science. CAB Int., Wallingford, Oxon, United Kingdom.

VanRaden, P. M., and E. J. H. Klaaskate. 1993. Genetic evaluation of length of productive life including predicted longevity of live cows. *J. Dairy Sci.* 76:2758–2764.

VanRaden, P. M., and G. R. Wiggans. 1995. Productive life evaluation: Calculation, accuracy, and economic value. *J. Dairy Sci.* 78:631–638.

Vollema, A. R., and A. F. Groen. 1998. A comparison of breeding value predictors for longevity using a linear model and survival analysis. *J. Dairy Sci.* 81:3315–3320.

Vukasinovic, N. 1999. Application of survival analysis in breeding for longevity. Page 3 in Proc. Int. Workshop on Genetic Improvement of Functional Traits in cattle, Longevity, Jouy-en-Josas, France. INTERBULL Bull. No. 21. Int. Bull Eval. Serv., Uppsala, Sweden.

Vukasinovic, N., J. Moll, and N. Künzi. 1999. Genetic Evaluation for length of productive life with censored records. *J. Dairy Sci.* 82:2178–2185.

Yazdi, M. H., R. Thompson, V. Ducrocq, and P. M. Visscher. 1999. A comparison of two survival analysis methods with the number of lactations as a discrete time variate. Page 48 in Proc. Int. Workshop on Genetic Improvement of Functional Traits in cattle, Longevity, Jouy-en-Josas, France. INTERBULL Bull. No. 21. Int. Bull Eval. Serv., Uppsala, Sweden.

APPENDIX A

Derivation of an Approximate Heritability for the Computation of Reliabilities

Consider a simple Weibull frailty model, with the hazard of a daughter j of sire i defined as:

$$\lambda_{ij}(t) = \lambda_0(t)\exp\{s_i\} = \rho t^{\rho-1}\exp\{\mu + s_i\} \quad [A-1]$$

where  $s_i$  is the effect of sire i and  $\lambda_0$  is a Weibull baseline hazard function with parameters  $\rho$  and  $\lambda$ . No fixed effect other than the grand mean  $\mu$  ( $= \rho \log \lambda$ ) is considered here. Reliability can be calculated as (Henderson (1975):

$$R_{true} = 1 - [\text{Var}(\hat{s}_i - s_i)/\sigma_s^2] \quad [A-2]$$

where the 'exact' asymptotic prediction error variance is obtained from the diagonal term of the inverse of the information matrix. Unfortunately, this matrix is usually too large to be calculated and an approximate formula for the prediction error variance  $\text{Var}(\hat{s}_i - s_i)$  is needed. This approximation can be obtained from the diagonal element of the Hessian matrix of the log posterior density  $\log p(\mu, \rho, \sigma_s^2|y)$ . If sires are unrelated and  $s_i \sim N(0, \sigma_s^2)$ , this log posterior density has the form:

$$\log p(\mu, \rho, \sigma_s^2|y) = \text{constant} + n \log \rho + (\rho - 1) \sum_{j \in \{\text{uncen}\}} \log y_j + \sum_{j \in \{\text{uncen}\}} \{\mu + s_i\} - \sum_{j \in \{\text{uncen}, \text{cens}\}} y_j^\rho e^{\mu + s_i} - \sum_i (s_i^2/2\sigma_s^2) \quad [A-3]$$

where  $\{\text{uncen}\}$  and  $\{\text{cens}\}$  are the sets of uncensored and censored records and  $y_j$  is either the failure time or the censoring time of animal j.

The derivative of [A-3] with respect to the sire effect is:

$$\partial \log p(\mu, \rho, \sigma_s^2|y)/\partial s_i = n_{\text{uncen}} - \sum_{j \in \{\text{uncen}, \text{cens}, i\}} y_j^\rho e^{\mu + s_i} - s_i/\sigma_s^2 \quad [A-4]$$

where the summation is now over all daughters of sire i only and  $n_{\text{uncen}}$  is the number of uncensored progeny of sire i. At the mode, [A-4] is 0 and therefore:

$$\sum_{j \in \{\text{uncen}, \text{cens}, i\}} y_j^\rho e^{\mu + s_i} = n_{\text{uncen}} - \hat{s}_i/\sigma_s^2 \quad [A-5]$$

Ignoring uncertainty due to estimation of  $\rho$  and  $\mu$ , we have, asymptotically:

$$\text{Var}(\hat{s}_i - s_i) = [-\partial^2 \log p(\mu, \rho, \sigma_s^2|y)/\partial s_i^2]_{s_i = \hat{s}_i}^{-1} \quad [A-6]$$

The term to be inverted is:

$$-\partial^2 \log p(\mu, \rho, \sigma_s^2|y)/\partial s_i^2 = \sum_{j \in \{\text{uncen}, \text{cens}, i\}} y_j^\rho e^{\mu + s_i} + 1/\sigma_s^2 \quad [A-7]$$

Combining [A-5] and [A-7], we obtain:

$$-\partial \log p(\mu, \rho, \sigma_s^2|y)/\partial s_i^2 = n_{\text{uncen}} - \hat{s}_i/\sigma_s^2 + 1/\sigma_s^2 \quad [A-8]$$

which implies, for a sire  $i$  with  $n_{\text{uncen}}$  uncensored daughters:

$$\text{Var}(\hat{s}_i - s_i) = [n_{\text{uncen}} - \hat{s}_i/\sigma_s^2 + 1/\sigma_s^2]^{-1} = f(\hat{s}_i) \quad \text{[A-9]}$$

where  $f(\cdot)$  is a function of  $\hat{s}_i$ . But we also have (Henderson, 1975):

$$\text{Var}(\hat{s}_i - s_i) = \sigma_s^2 - \text{Var}(\hat{s}_i) \quad \text{[A-10]}$$

One can approximate the right hand side of [A-9] using a Taylor series expansion of  $f(\cdot)$  and equate it to the right hand side of [A-10] in order to find  $\text{Var}(\hat{s}_i)$  in [A-10]:

If  $f(\hat{s}_i) = \text{Var}(\hat{s}_i - s_i) = \frac{\sigma_s^2}{n_{\text{uncen}}\sigma_s^2 + 1 - \hat{s}_i}$ , the first and second derivatives of  $f(\cdot)$  with respect to  $\hat{s}_i$  are:  $f'(\hat{s}_i) = \frac{\sigma_s^2}{(n_{\text{uncen}}\sigma_s^2 + 1 - \hat{s}_i)^2}$  and  $f''(\hat{s}_i) = \frac{2\sigma_s^2}{(n_{\text{uncen}}\sigma_s^2 + 1 - \hat{s}_i)^3}$  and the Taylor series expansion of  $f(\cdot)$  around  $E[\hat{s}_i] = 0$  gives:

$$f(\hat{s}_i) \approx f(E[\hat{s}_i]) + f'(E[\hat{s}_i])(\hat{s}_i - E[\hat{s}_i]) + 0.5f''(E[\hat{s}_i])(\hat{s}_i - E[\hat{s}_i])^2 \quad \text{[A-11]}$$

Taking the expectation of [A-11] over all  $\hat{s}_i$  corresponding to sires with  $n_{\text{uncen}}$  uncensored daughters, one gets:

$$E[f(\hat{s}_i)] \approx \frac{\sigma_s^2}{n_{\text{uncen}}\sigma_s^2 + 1} + 0.5 * \frac{2\sigma_s^2}{(n_{\text{uncen}}\sigma_s^2 + 1)^3} \text{Var}(\hat{s}_i) \quad \text{[A-12]}$$

$$\sigma_s^2 - \text{Var}(\hat{s}_i) = E[f(\hat{s}_i)] \approx \frac{\sigma_s^2}{n_{\text{uncen}}\sigma_s^2 + 1} \left[ 1 + \frac{1}{(n_{\text{uncen}}\sigma_s^2 + 1)^2} \text{Var}(\hat{s}_i) \right] \quad \text{[A-13]}$$

It follows that:

$$\text{Var}(\hat{s}_i) = [n_{\text{uncen}}\sigma_s^4/(n_{\text{uncen}}\sigma_s^2 + 1)]/[1 + \{\sigma_s^2/(n_{\text{uncen}}\sigma_s^2 + 1)^3\}] \quad \text{[A-14]}$$

Then:

$$\text{Var}(\hat{s}_i - s_i) \approx \sigma_s^2(1 + r)/(n_{\text{uncen}}\sigma_s^2 + 1 + r) \approx \sigma_s^2/(n_{\text{uncen}}\sigma_s^2 + 1) \quad \text{[A-15]}$$

with  $r = [\sigma_s^2/(n_{\text{uncen}}\sigma_s^2 + 1)^2]$ .

$r$  is a function of  $\sigma_s^2$  that is always positive with a maximum at  $\sigma_s^2 = 1/n_{\text{uncen}}$ , the maximum being  $r = 1/4 n_{\text{uncen}}$  which is negligible compared to 1 for moderate  $n_{\text{uncen}}$ .

From [A-15], it is possible to derive an expression of an “effective” heritability  $h_{\text{eff}}^2$  that can be used in a context similar to classical mixed linear models. First, consider the usual approximate formula for reliability  $R$  of the proof of sire  $i$ , based on selection index theory using progeny information only and ignoring uncertainty due to estimation of fixed effects:

$$R = n_{\text{uncen}}h_{\text{eff}}^2/[(n_{\text{uncen}} - 1)h_{\text{eff}}^2 + 4] \quad \text{[A-16]}$$

Equating  $R$  to  $R_{\text{true}}$ , one gets:

$$h_{\text{eff}}^2 = 4[\sigma_s^2 - \text{Var}(\hat{s}_i - s_i)]/[\sigma_s^2 + (n_{\text{uncen}} - 1)\text{Var}(\hat{s}_i - s_i)] \quad \text{[A-17]}$$

Plugging expressions [A-15] of  $\text{Var}(\hat{s}_i - s_i)$  into the definition [A-17] of the effective heritability leads to:

$$h_{\text{eff}}^2 \approx [4\sigma_s^2/(\sigma_s^2 + 1 + r)] \approx [4\sigma_s^2/(\sigma_s^2 + 1)] \quad \text{[A-18]}$$

and the approximate reliability of its proof is simply:

$$R = [n_{\text{uncen}}/(n_{\text{uncen}} + 1/\sigma_s^2)]. \quad \text{[A-19]}$$

When  $n_{\text{uncen}}$  is not known, for example, when one wants to predict how reliability will change with time as more and more daughters are uncensored, it is possible to express  $R$  as a function of  $n$ , the total number of daughters of sire  $i$  and  $p$ , the proportion of its progeny that are no longer alive at a given time  $t$ . Then:

$$R = [pn/(pn + 1/\sigma_s^2)] = n/[n + 1/(p\sigma_s^2)] \quad \text{[A-20]}$$

[A-20] is a particular case of the expression of the reliability based on the usual index selection formula for sire evaluation  $R = n/(n + k)$ , where  $k$  is a function of the heritability of the trait ( $k = (4 - h^2)/h^2$ ). This similarity suggests the definition of an “equivalent” heritability  $h_{\text{equ}}^2$  as:

$$h_{\text{equ}}^2 = 4\sigma_s^2/[\sigma_s^2 + (1/p)]. \quad \text{[A-21]}$$

The term “equivalent” refers here to the fact that the proof of sire  $i$  with  $n$  daughters would get the same reliability as if it were evaluated on a linear trait with heritability  $h_{\text{equ}}^2$ . Since the proportion  $p$  changes over time, the equivalent heritability  $h_{\text{equ}}^2$  also increases with time until it reaches  $h_{\text{eff}}^2$ , the theoretical heritability that one would get in total absence of censoring.

APPENDIX B

Derivation of the Average Survival Time in Selected and Unselected Populations

Consider a Weibull regression model with an overall mean as the only fixed effect and a random sire effect. For a given sire  $i$  with effect  $s_i$ , the probability density function of failure (or culling here) at time  $t$  of his progeny is the product of the hazard function and the survival function,

$$f(t) = h(t)S(t) = \lambda \rho (-\lambda t)^{\rho-1} \exp(s_i) [\exp(-\lambda t) \rho]^{\exp(s_i)} \quad [B-1]$$

The expected survival time for progeny of sire  $i$  is

$$E_{s_i}(t) = \int_{t=0}^{\infty} t f(t) dt = \frac{\Gamma(1 + \frac{1}{\rho})}{\lambda} \exp(-\frac{1}{\rho} s_i). \quad [B-2]$$

The average survival time in the population is calculated by integrating out the random sire effect,

$$E(t) = \frac{\Gamma(1 + \frac{1}{\rho})}{\lambda} \int_{-\infty}^{+\infty} \exp(-\frac{1}{\rho} s_i) f(s_i) ds_i \quad [B-3]$$

If we assume a normal distribution with mean  $\mu$  and variance  $\sigma_s^{2*}$  for the sire effects, even after selection, i.e, if

$$f(s_i) = \frac{1}{\sqrt{2\pi\sigma_s}} \exp(-\frac{(s_i - \mu)^2}{2\sigma_s^{2*}})$$

then

$$E(t) = \frac{\Gamma(1 + \frac{1}{\rho})}{\lambda} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma_s}} \exp\left(-\frac{s_i}{\rho} - \frac{(s_i - \mu)^2}{2\sigma_s^{2*}}\right) ds_i$$

$$E(t) = \frac{\Gamma(1 + \frac{1}{\rho})}{\lambda} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma_s}}$$

$$\exp\left(-\frac{\left[(s_i - \mu + \frac{\sigma_s^2}{\rho})^2 + (2\mu\sigma_s^{2*}/\rho) - \left(\frac{\sigma_s^{2*}}{\rho}\right)^2\right]}{2\sigma_s^2}\right) ds_i$$

$$E(t) = \frac{\Gamma(1 + \frac{1}{\rho})}{\lambda} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma_s}} \quad [B-4]$$

$$\exp\left(-\frac{\left(s_i + \frac{\sigma_s^2}{\rho}\right)^2}{2\sigma_s^2}\right) \exp\left(\frac{\sigma_s^{2*}}{2\rho^2} - \frac{\mu}{\rho}\right) ds_i$$

Equation [B-4] can be simplified when noting that,

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma_s}} \exp\left(-\frac{\left(s_i + \frac{\sigma_s^2}{\rho}\right)^2}{2\sigma_s^2}\right) ds_i = 1.0$$

Then, the expected average survival time in the population reduces to:

$$E(t) = \frac{\Gamma(1 + \frac{1}{\rho})}{\lambda} \exp\left(\frac{\sigma_s^{2*}}{2\rho^2} - \frac{\mu}{\rho}\right)$$

This approximation can be used both for selected and unselected populations. In the unselected (base) population,  $\sigma_s^{2*} = \sigma_s^2$  and  $\mu = 0$ , so that the average lifespan in the population is,

$$E(t) = \frac{\Gamma(1 + \frac{1}{\rho})}{\lambda} \exp\left(\frac{\sigma_s^2}{2\rho^2}\right)$$

Usually, the sire variance is small ( $\ll 1$ ) relative to  $2\rho^2$ , so  $\exp(\sigma_s^2/2\rho^2) \sim 1$ , and the approximation of the expected survival time is the same as the exact value when sires are ignored,

$$E(t) \sim \frac{\Gamma(1 + \frac{1}{\rho})}{\lambda}. \quad [B-5]$$