

## Development of a Survival Model with Piecewise Weibull Baselines for the Analysis of Length of Productive Life of Holstein Cows in Japan

Y. Terawaki,<sup>\*1</sup> T. Katsumi,<sup>†</sup> and V. Ducrocq<sup>‡</sup>

<sup>\*</sup>Dairy Science Institute, and

<sup>†</sup>The Graduate School of Dairy Science, Rakuno Gakuen University, 069-8501 Ebetsu, Japan

<sup>‡</sup>Station de Génétique Quantitative et Appliquée, Département de Génétique Animale, Institut National de la Recherche Agronomique, 78352 Jouy en Josas, France

### ABSTRACT

A total of 100,000 records of Holstein dairy cows in Hokkaido, Japan, were used to study the current characteristics of their length of productive life, to build up a suitable model for genetic survival analysis, and to apply the piecewise Weibull baseline model. The data for this analysis only included records of cows belonging to herds in which more than 60% of cows had type score. Compared with other studies, the proportions of cows with only 1 or 2 calvings were low (24 and 46%, respectively), indicating a very low culling rate in first or second parity in Hokkaido. The median length of productive life was about 1,250 d (3 yr and 5 mo). Four different definitions of stages of lactation were studied. The best fit of the data was obtained when stages of lactation were chosen with cutpoints at 0, 60, 250, and 350 d after calving. Whatever the definition of stages of lactation, the piecewise Weibull baseline model was better than a unique baseline model. The estimated shape Weibull parameters  $\rho$  for the different parity  $\times$  stage of lactation combinations varied greatly (1.0 to 5.0). The estimated hazard functions based on the piecewise Weibull baseline model were smooth and reflected the important changes in observed hazard within lactation. The results of this study are expected to contribute to the development of a genetic model for sire evaluation and to genetic improvement of length of productive life of Holstein cattle in Hokkaido, Japan. **Key words:** survival model, piecewise Weibull baseline, longevity, survival analysis

### INTRODUCTION

Genetic evaluations of length of productive life of dairy cows are receiving increasing attention worldwide. For example, the November 2005 international evaluation for direct longevity performed by Interbull

for the Holstein breed included 17 countries (Interbull, 2005). Before a genetic analysis for length of productive life of dairy cows is performed, it is important to focus on the modeling of the nongenetic factors that influence longevity. There is currently no information about longevity for Holstein cows in Hokkaido, Japan. A preliminary analysis is needed to understand the present status of longevity of Holstein cows in Hokkaido and to build a suitable survival model describing the hazard for a cow over time (i.e., her limiting probability to be culled at time  $t$ , given she was alive at time  $t - 1$ ).

Length of productive life is influenced not only by the cow's production ability but also by the herd management policy. This management policy has a larger influence on longevity than on production traits. Therefore, it is expected that the suitable model for genetic analysis differ over countries and regions that have different dairy production systems. For example, in countries of the European Union, a quota system has been applied since 1984. This has led to a clear seasonal pattern of cullings (Ducrocq, 1994; Vukasinovic et al., 2001). But it is not necessary to consider such a seasonal fluctuation of culling rate in Japan, which does not have a quota system. On the other hand, based on preliminary analysis, it appears that the presence of a type score considerably influences the cows' hazards in Hokkaido. Furthermore, the definition of stages of lactation to account for changes in risk of being culled during the lactation varied between studies (Ducrocq, 1994; Beaudeau et al., 1995; Gröhn et al., 1997; Strandberg and Rönström, 2000; Buenger et al., 2001; Larroque and Ducrocq, 2001; Rönström and Strandberg, 2002; Rönström et al., 2003; Caraviello et al., 2004). Recently, Ducrocq (2005) reported a genetic survival analysis of French Holstein cows using a within-lactation piecewise Weibull hazard model. It is believed that piecewise Weibull baseline hazard functions can cope better with changes in the baseline hazard rate over time, leading to a better fit.

The objectives of this study were to assess the current characteristics of length of productive life of Holstein cows in Hokkaido, Japan and to determine a suitable

Received December 26, 2005.

Accepted May 5, 2006.

<sup>1</sup>Corresponding author: terawaki@rakuno.ac.jp

model for genetic survival analysis using an adequately specified piecewise Weibull hazard model.

## MATERIALS AND METHODS

The datasets used were provided by the Hokkaido Dairy Milk Recording and Testing Association. Records were edited for survival analysis for an observation period defined to be from January 1, 1984 to December 31, 1999. Therefore, cows calving for the first time before January 1, 1984, were excluded from the analysis. Records of cows sold to other dairy herds were considered as censored at date of sale. Cows still alive on December 31, 1999, were considered as censored at that date. Longevity was defined as the number of days from first calving to culling or censoring date. After editing the data, cows belonging to herds in which more than 60% of cows had type score were selected. The resulting file included 131,499 cows in 981 herds. For computational limitations (which were later found to be easily avoidable), a subset of 100,000 cows from 871 herds was selected at random from all herds.

Two parametric proportional hazards models for survival analysis were used. For both models, the genetic component was neglected to concentrate on the modeling of the environmental effects. In the first one, a unique baseline hazard function was defined for the whole productive life of a cow:

$$h(t) = h_0(t) \exp[\text{HY}_i(t') + \text{Y}_j(t') + \text{MILK}_k(t'') + \text{SIZE}_l(t') + \text{MP}_m(t'') + \text{PS}_n(t') + \text{SC}_o + \text{TS}_p]. \quad [1]$$

In this model,  $h(t)$  is the hazard function of the cow,  $t$  days after her first calving;  $h_0(t)$  is a Weibull baseline hazard function, with parameters  $\lambda$  and  $\rho$  ( $\lambda > 0$ ,  $\rho > 0$ ); and  $\text{HY}_i(t')$  is the time-dependent random effect of the  $i$ th herd (871 herds)  $\times$  year interaction at calendar time  $t'$ , assumed to be independently distributed and following a log gamma distribution. Without loss of generality, the 2 parameters of the log-gamma distribution are assumed to be equal to  $\gamma$ . In such a case, the mean and the variance of the herd-year effect are  $\Psi(\gamma) - \ln(\gamma)$  and  $\Psi^{(1)}(\gamma)$  where  $\Psi(\cdot)$  and  $\Psi^{(1)}(\cdot)$  are the digamma and the trigamma functions, respectively;  $\text{Y}_j(t')$  is the time-dependent fixed effect of  $j$ th year (16 levels). As a consequence,  $\text{HY}_i(t') + \text{Y}_j(t')$  has a constant within-year variance but a mean which may vary from one year to the next;  $\text{MILK}_k(t'')$  is the time-dependent fixed effect of the  $k$ th within-herd class of milk yield;  $\text{SIZE}_l(t')$  is the time-dependent fixed effect of the  $l$ th variation in herd size;  $\text{MP}_m(t'')$  is the time-dependent fixed effect of the  $m$ th interaction between within-herd class of milk yield and parity;  $\text{PS}_n(t')$  is the time-dependent fixed effect of the  $n$ th interaction between parity and stage of lacta-

tion;  $\text{SC}_o$  is the time-independent fixed effect of the other group of origin of the cow's sire (Japan for Holstein sires born or raised in Japan; US or Canada for imported semen); and  $\text{TS}_p$  is the time-independent fixed effect of presence of type scores for this cow.

For the second model, different Weibull baseline hazards are used depending on current parity and stage of lactation. Consequently, the interaction effect of parity and stage of lactation in the exponential part of [1] becomes redundant:

$$h(t) = h_{0,n}(\tau) \exp[\text{HY}_i(t') + \text{Y}_j(t') + \text{MILK}_k(t'') + \text{SIZE}_l(t') + \text{MP}_m(t'') + \text{SC}_o + \text{TS}_p], \quad [2]$$

where  $h_{0,n}(\tau)$  is the Weibull baseline for the  $n$ th subclass of parity by stage of lactation interaction,  $\tau$  denotes the number of days between the most recent calving and current time  $t$ , and other effects are as previously defined for model [1].

Details for classes of the effects included in the models are in Table 1. Exact lactation records were not available in the data set but only codes (1 to 7) for 305-d mature-equivalent milk yield class (see Table 1). To better account for the fact that voluntary culling is mainly based on within-herd level of milk production, these codes were transformed into within-herd  $\times$  year milk classes created as follows. The code's average was calculated for each herd and year, separately for first and later parities. For every lactation of each cow, the deviation between the cows' codes and the corresponding averages were computed and rounded to the nearest integer. The resulting number was finally transformed back to a scale from 1 to 7 by adding 4 (average class) and merging extreme values below 1 and above 7 to 1 and 7, respectively. The 5 classes of the SIZE effect reflected the change in herd size between the previous year and the current one. The change was calculated from the sampled population, as an approximation of what happened in the original population. It has been repeatedly found that the risk of being culled for each cow is higher in shrinking herds compared with expanding herds. Four different cases of stages of lactation for  $\text{PS}_n(t')$  in model [1] and for the baseline in model [2] were considered. The cutpoints for case A were 0, 60, 150, and 270 d after each calving date, as, for example, in Beaudeau et al. (1995) or Gröhn et al. (1997). The cutpoints for the other cases were 0, 10, 250, and 350 d (case B), 0, 60, 250, and 350 d (case C), and 0, 250, and 350 d (case D). They were determined after a careful examination of the nonparametric estimate of the within-lactation hazard function for the data in this study. Cases B and C differ by the choice of the first cutpoint, 10 d as in Rönström et al. (2003) or 2 mo

**Table 1.** Details for classes of effects included in the models [1] and [2]

Effect	Details
$HY_i(t')^1, Y_i(t')$	Calendar year 1984, 1985, 1986, ..., 1999
$MILK_k(t'')^1$	Codes: 1 = $\leq 5,000$ kg, 2 = $\leq 6,000$ kg, 3 = $\leq 7,000$ kg, 4 = $\leq 8,000$ kg, 5 = $\leq 9,000$ kg, 6 = $\leq 10,000$ kg, 7 = $> 10,000$ kg in the original data. These codes were transformed into within herd $\times$ year classes <sup>2</sup>
$SIZE_l(t')$	1 = Variation in herd size between 2 consecutive years $< -15\%$ ; 2 = $-15\% \leq$ to $\leq 5\%$ , 3 = $5\% <$ to $< +5\%$ ; 4 = $+5\% \leq$ to $\leq +15\%$ ; 5 = $+15\% <$
$MP_m(t'')$	Milk class (1, 2, ..., 7) $\times$ parity (1, 2, ..., <6)
$PS_n(t'')$	Parity (1, 2, ..., <6) $\times$ stage of lactation <sup>3</sup>
$SC_o$	Origin of sire: 1 = born or raised in Japan; 2 = imported semen from US; 3 = imported semen from Canada
$TS_p$	1 = presence of type score; 2 = absence of type score

<sup>1</sup> $t'$  = Calendar time;  $t''$  = time since the most recent calving.

<sup>2</sup>See Materials and Methods for detailed explanation.

<sup>3</sup>Case A had cutpoints of 0, 60, 150, and 270 d after each calving. Other cases had cutpoints of 0, 10, 250, and 350 d (B), 0, 60, 250, and 350 d (C), and 0, 250, and 350 d (D).

after calving. For simplicity, case D ignores potential differences between the first 2 stages of cases B and C.

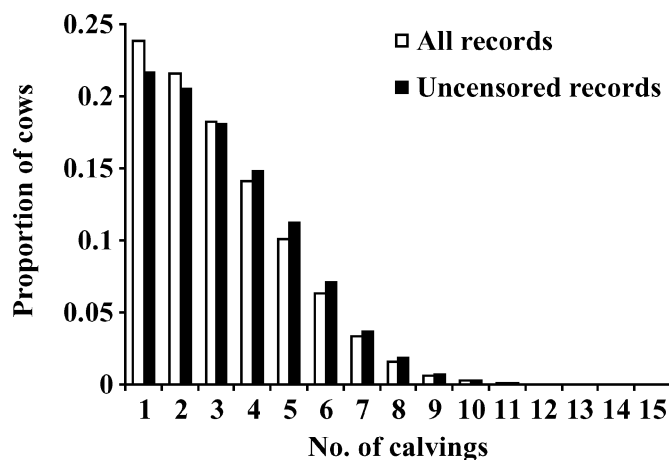
The goodness of fit of the different nonnested models was compared using the Akaike information criterion (AIC). All computations were carried out using the Survival Kit (Version 5.0) software (Ducrocq and Sölkner, 1998).

**RESULTS**

The distribution of number of calvings per cow is presented in Figure 1. The proportion of cows that experienced only one calving was about 24% overall and about 22% among uncensored cows. The proportion of cows having only 1 or 2 calvings was about 46 and 43% overall and for uncensored cows, respectively. Sample statistics for this data are shown in Table 2. The proportion of right-censored records was quite low (27.6%). This is due to the relatively long period of observation considered here (from 1984 to 1999) compared with

other reports (Dürr et al., 1999; Caraviello et al., 2004; Ducrocq, 2005). The Kaplan-Meier estimate of the overall survival function in the Japanese population is displayed in Figure 2 and indicates a high median length of productive life of about 1,250 d. Figure 3 shows the nonparametric estimate of the within-lactation hazard functions for our data set. Although different across parities, these hazard functions show a similar pattern: initially, the hazard strongly increased just after calving for all parities. In first parity, the hazards slightly decreased from about 60 d after calving until about 250 d after calving. On the other hand, a slight increase was observed in later parities. The slopes of the within-lactation hazards were steep after about 250 d, and became even steeper after about 350 d. The cutpoints (0, 60, 250, and 350 d) for case C were determined based on these observations.

All effects in models [1] and [2] were highly significant ( $P < 0.001$ ). The AIC values for models [1] and [2] for different cases are reported in Table 3 for the different choices of cutpoints. Case C, having 0, 60, 250, and 350 d of lactation as cutpoints, had the smallest AIC, indicative of the best goodness of fit for both models. Model [2] led to smaller AIC values than model [1] in all cases. This indicated that the piecewise Weibull baseline model was better than a model with a unique overall baseline. The differences between AIC values for models [1] and [2] were largest in case D. The difference



**Figure 1.** Number of calvings per cow.

**Table 2.** Descriptive statistics of censoring and failure time (d)

Right-censored records (%)	27.6
Minimum censoring time	109
Maximum censoring time	5,597
Average censoring time	1,035.8
Uncensored records (%)	72.4
Minimum failure time	9
Maximum failure time	5,474
Average failure time	1,217.6

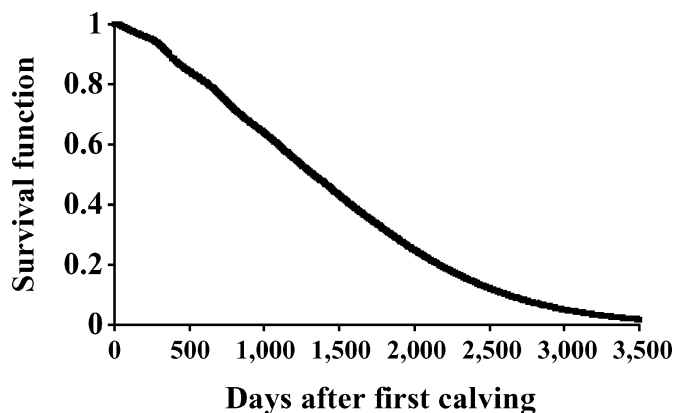


Figure 2. Kaplan-Meier estimate of survival function.

between AIC values for cases C and D decreased from 5,283.5 in model [1] to 2,154.1 in model [2].

Figure 4 has the values of the shape Weibull parameters,  $\rho$ , estimated using models [1] and [2]. A  $\rho$  value less than, equal to, or larger than 1 indicates a decreasing, constant, or increasing hazard over time, respectively. In Figure 4, the unique  $\rho$  value for model [1] is indicated with a straight dashed line. With model [2], for each parity  $\times$  stage of lactation interaction, the  $\rho$  estimates are connected by a different line for each parity. The  $\rho$  estimates obtained with model [2] were between 1.0 and 5.0, except for case B. The range of  $\rho$  was largest in first parity and became smaller with later parities. This indicates that the change of the slopes of the hazard was stronger in first parity and more moderate subsequently. Cases B, C, and D had the same cutpoints for the last 2 stages of lactation. Consequently, the  $\rho$

values for these stages were very similar. The  $\rho$  estimate (0.948) for the second stage in first parity in case A illustrates the ability of model [2] to cope with the decrease in hazard. The  $\rho$  values during the first stage (0 to 10 d after calving) in case B were very large compared with the other cases. The  $\rho$  estimate in fourth parity (30.3) was particularly surprising. These extremely variable and large values can be attributed to the very low number of cows apparently culled during this period in each parity (0 to 5 cows). Defining a very short stage of lactation just after calving might be unsuitable, at least for the data set in our study.

Table 4 shows the estimates of the  $\gamma$  parameter, their standard error and the corresponding herd-year variances using models [1] and [2]. A slightly smaller  $\gamma$  parameter; that is, a larger variance of herd  $\times$  year effect, was obtained with model [2] than with model [1].

Table 5 shows relative risk ratios of some effects estimated using models [1] and [2] with the definition of stages of lactation of case C. Risk ratios for both models were similar for all effects. Very high risk ratios were found for the lowest milk classes. Higher risk ratios for cows without type score were also observed.

The predicted hazard functions estimated using models [1] and [2] are shown in Figure 5, again with the definition of stages of lactation of case C. These hazard functions were computed considering a reference cow with a Japanese sire, with an average production (fourth class of milk yield) during all her herd life, with a type score and with a constant calving interval of 400 d. The hazard functions of model [1] were stair-shaped. On the other hand, the hazard functions predicted using model [2] were much smoother. Model [2] well reflected

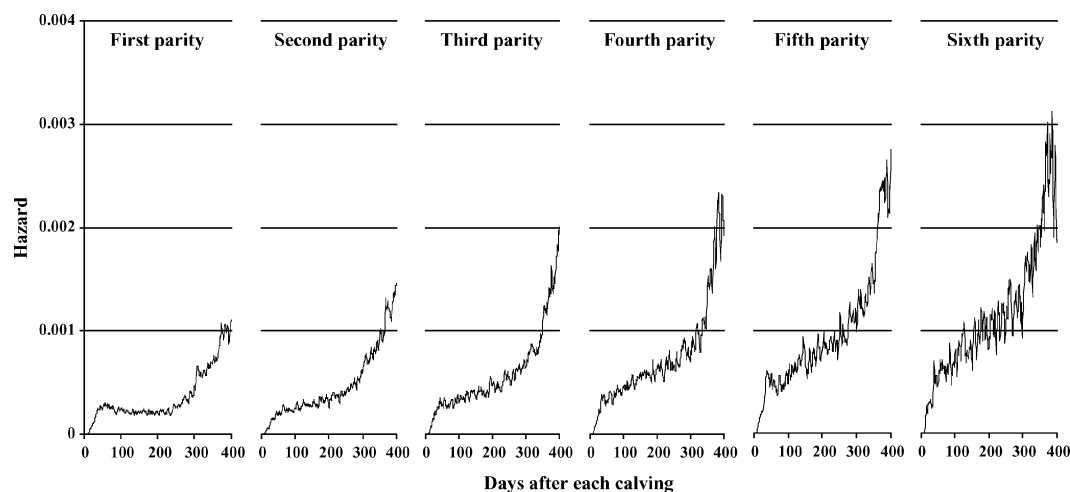
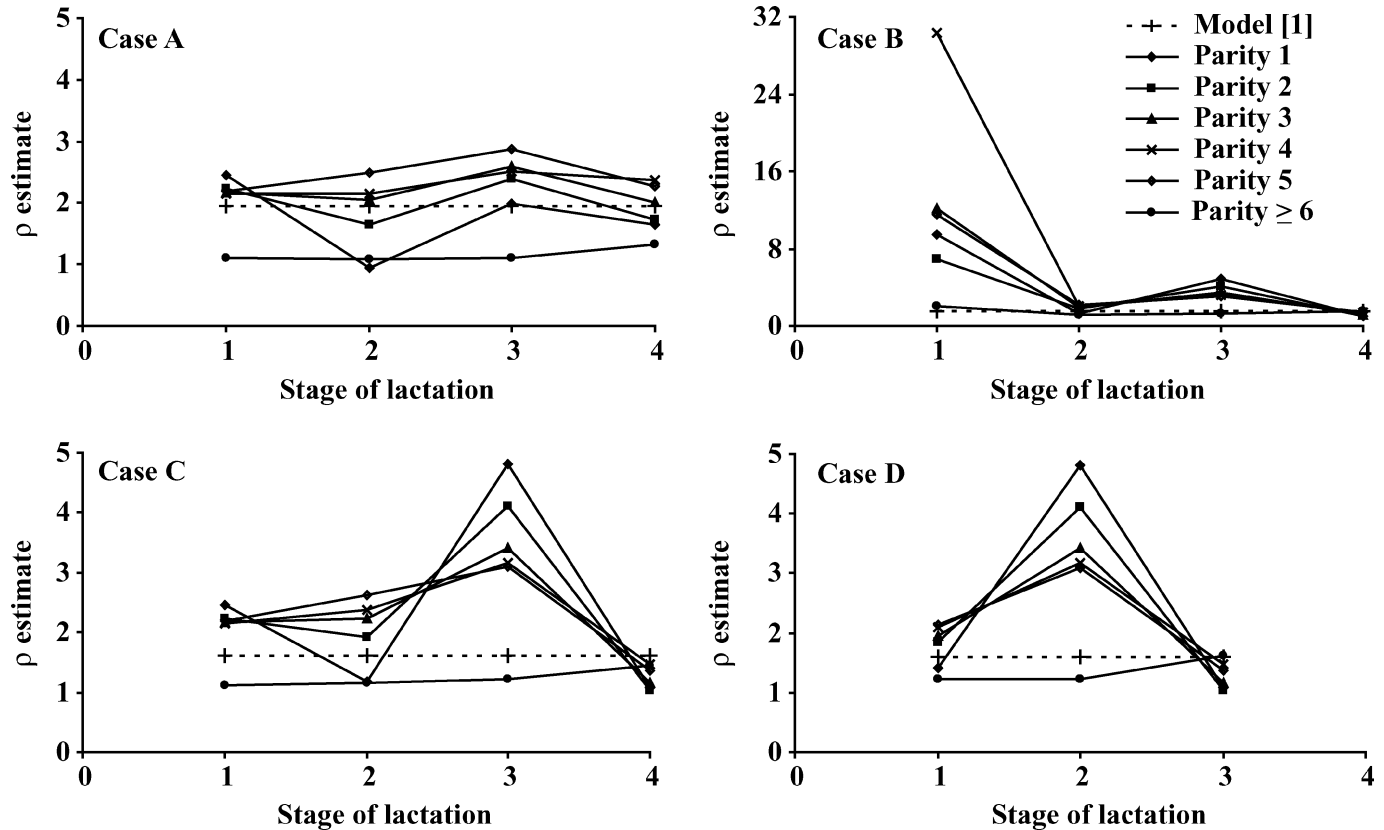


Figure 3. Within-lactation nonparametric estimate of the hazard function (derived from the Kaplan-Meier estimate of the within lactation survival function).

**Table 3.** Comparison of the Akaike information criterion (AIC) for models [1] and [2] with 4 different definitions (cases A to D) of stages of lactation

Case (cutpoints)	Model [1]	Model [2]	Difference
A (0, 60, 150, and 270 d)	7,017.8	3,911.6	3,106.2
B (0, 10, 250, and 350 d)	7,402.1	1,038.9	6,362.9
C (0, 60, 250, and 350 d)	4,965.8	0 <sup>1</sup>	4,965.8
D (0, 250, and 350 d)	10,249.3	2,154.1	8,095.2

<sup>1</sup>Used as reference (AIC = 1,076,152.1).



**Figure 4.** Weibull parameter ( $\rho$ ) estimates in model [1] and for parity  $\times$  stage of lactation subclasses in model [2]. Cutpoints for stages of lactation were: 0, 60, 150, and 270 d after each calving (case A); 0, 10, 250, and 350 d (B); 0, 60, 250, and 350 d (C); and 0, 250, and 350 d (D). Model [1] and Parity 1 to  $\geq 6$  indicate  $\rho$  estimates in models [1] and [2], respectively.

**Table 4.** Estimates of the gamma parameters, standard errors, and corresponding herd-year variances for models [1] and [2] with 4 different definitions (cases A to D) of stages of lactation<sup>1</sup>

Casae	Model [1]			Model [2]		
	Gamma parameter	Standard error	Herd-year variance	Gamma parameter	Standard error	Herd-year variance
A	4.515	0.024	0.2478	4.205	0.023	0.2683
B	4.999	0.025	0.2214	4.503	0.024	0.2485
C	4.731	0.024	0.2353	4.427	0.024	0.2533
D	5.060	0.025	0.2184	4.518	0.024	0.2476

<sup>1</sup>Cutpoints to define stages of lactation were 0, 60, 150, and 270 d (case A); 0, 10, 250, and 350 d (case B); 0, 60, 250, and 350 d (case C); or 0, 250, and 350 d (case D).

**Table 5.** Relative risk ratios for some effects in models [1] and [2] (case C)

Effect	Class	Model [1]	Model [2]
Class of milk yield	1	17.948	17.011
	2	6.137	6.163
	3	2.079	2.076
	4	1.000 <sup>1</sup>	1.000 <sup>1</sup>
	5	0.768	0.744
	6	0.634	0.608
	7	0.579	0.540
Class of variation in herd size	1	1.728	1.810
	2	1.156	1.157
	3	1.000 <sup>1</sup>	1.000 <sup>1</sup>
	4	0.877	0.873
	5	0.695	0.692
Origin of the cow's sire	Japan	1.000 <sup>1</sup>	1.000 <sup>1</sup>
	United States	0.878	0.878
	Canada	0.921	0.919
Presence of type score	Yes	1.000 <sup>1</sup>	1.000 <sup>1</sup>
	No	2.361	2.396

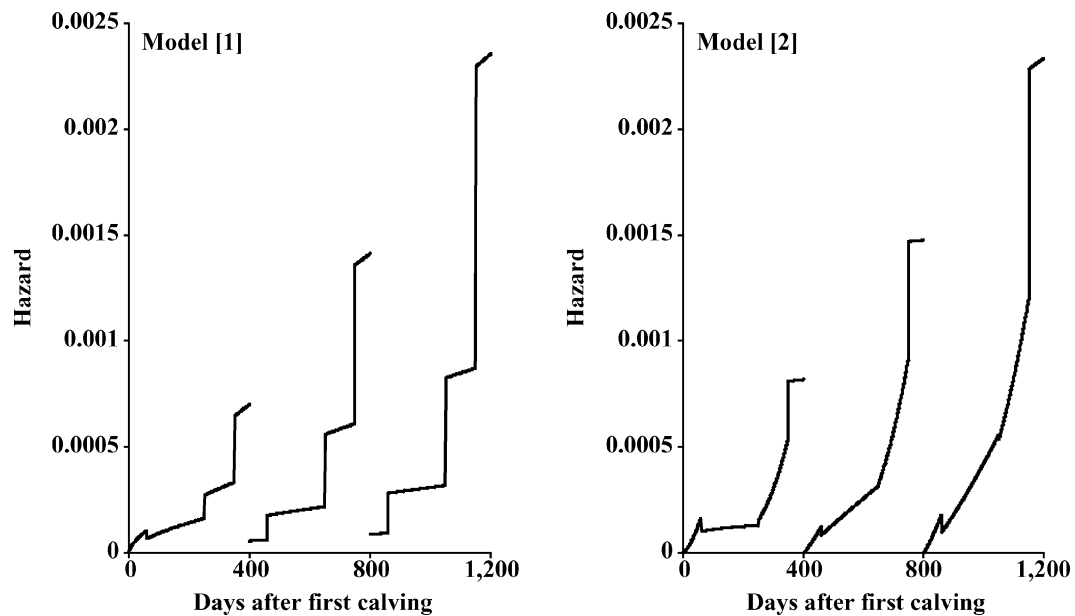
<sup>1</sup>Used as reference.

the changes in the nonparametrical hazards (Figure 3) during the lactation.

## DISCUSSION

A number of observations show that Japanese Holstein cows live longer, on average, than their counterparts in North American and European Holstein populations. For example, the proportion of cows culled during their first or second lactation (46%) is relatively low; the proportion of cows experiencing 4 or more calvings

is large (36%). Also, the average failure time and the average censoring time were large compared with other reports (e.g., Dürr et al., 1999) but these values are at least partly related to the relatively long period of observation (from 1984 to 1999). The Kaplan-Meier estimate of the overall survival function (Figure 2) can be used to find a more unbiased comparison criterion: the median length of productive life; that is, the number of days after calvings when half of the uncensored cows have died. This median value is 1,250 d for the Japanese Holstein, which is larger than that reported by Dürr et



**Figure 5.** Predicted hazard functions for models [1] and [2] (case C) for a cow with average milk yield, with a type score, with a Japanese sire, and constant calving intervals of 400 d.

al. (1999) for Canadian Holstein (1,078 d) or by Ducrocq (unpublished results) for French Holstein (1,030 d). The diverse management systems and economic contexts are, of course, largely responsible for these differences.

The nonparametric estimates of the within-lactation hazard functions illustrate the within-lactation cyclic pattern of culling risks already reported by Ducrocq (1994, 2002, 2005), Strandberg and Röxström (2000), Vukasinovic et al. (2001), Röxström and Strandberg (2002), or Röxström et al. (2003). This cyclic pattern is at the origin of the introduction of the parity by stage of lactation effect in model [1]. It is also the main motivation for proposing model [2] for the analysis of length of productive life, by directly accounting for it in the baseline hazard function. Here, as in other studies, culling risk is relatively low during most of the lactation, remaining quite stable in first parity but moderately increasing for later parities. Clearly, culling during the early part of the lactation corresponds to emergency cases (very low milk production, severe functional problems). At the end of the lactation, culling risk steadily increases: most of the animals are culled when milking them is no longer profitable. Note that the data at hand do not reveal when the culling decision was actually made. Poor production or functional weaknesses may have been detected in early lactation, leading to the decision of, for example, not breeding the cow again, with culling occurring much later. The hazard function during the first few days after each calving seems to indicate an initial absence of any risk of being culled. This observation is similar to situations in which length of productive life is computed exclusively from milk recording data. In such a system, cows that have been culled just after calving but before being milk recorded are, in practice, declared as culled at the end of the previous lactation. However, culling dates are recorded based on the reports from herd managers in Hokkaido, Japan. When exact culling dates are available, a decreasing trend in the hazard is often observed just after calving (see for example, Strandberg and Röxström, 2000). The culling pattern just after calving in Hokkaido might be quite different from other countries.

In studies using models describing the baseline hazard over the whole herd life of the cows, the limits of stages of lactation have been arbitrarily set, without real justification. The use of the nonparametric estimates of the within-lactation hazard function eliminates this arbitrariness. Cutpoints can be precisely chosen to optimize goodness of fit. However, it is important to remember that, whatever the model, increasing the number of stages of lactation leads to a corresponding increase in the size of the already large recoded data file (Ducrocq, 2005). A compromise has to be found between number of cutpoints and goodness of fit. Attempts to

improve goodness of fit while maintaining or decreasing the number of stages have first concentrated on the use of a different Weibull hazard function for each lactation (Röxström et al., 2003; Sewalem et al., 2005), while keeping a stage of lactation effect in the exponential part of the model. The results obtained here confirm those of Ducrocq (2002, 2005). The definition of a different Weibull baseline hazard function for each combination of parity and stage of lactation leads to a substantial improvement of goodness of fit. Interestingly, the need for an "early" stage of lactation effect (0 to 10 d after each calving as in Röxström et al. (2003), or 0 to 60 d as in case C) is not essential with the piecewise Weibull model. If computing resources are limiting, case D is a reasonable compromise between number of stages and goodness of fit. Finally, an appealing feature of the piecewise Weibull model is its ability to describe the within-lactation hazard function in a smoother fashion than previous models.

The Weibull  $\rho$  parameters estimated by model [1] were similar to other comparable studies (Ducrocq, 1994; Dürr et al., 1999; Vollema et al., 2000; Vukasinovic et al., 2001; Röxström and Strandberg, 2002): they were between 1.5 and 2.0. The  $\rho$  values for model [2] logically reflected the various characteristics of the hazards in different parities and stages of lactations seen in Figure 3. Ducrocq (2005) examined in detail the piecewise Weibull  $\rho$  estimates in the French Holstein population. Because case D considered here was similar to Ducrocq's study in terms of cutpoint definitions for stage of lactation, both sets of  $\rho$  parameters can be compared. A common finding is the fact that  $\rho$  values were small for the first stage and were substantially larger in the next stage. But the  $\rho$  values in the last stage in this study were slightly above 1 whereas they were much higher ( $>3$ ) in Ducrocq's study. This fact would indicate that culling is more concentrated over time in the Japanese context: after 350 d, the risk of being culled is rather stable in Hokkaido in contrast with a continuously sharp increase observed in France. In the piecewise Weibull model of Ducrocq (2005), the Weibull baseline hazard functions were also different for each year-season combination. The inclusion of such an interaction was justified because changes in culling practices over time led to biased genetic trends (Ducrocq, 2004). The need for this interaction will have to be checked when the models considered here for nongenetic effects are extended to mixed models.

The estimates of the fixed effects included in the models studied here are consistent with other reports. The most important effect is the within herd class of milk production effect. Its purpose in a genetic evaluation is to adjust the sire's genetic value for longevity for the main reason of voluntary culling (low milk production)

to approximate functional longevity. Obviously, the extension of the model to a mixed model to be applied for a national genetic evaluation requires a more direct and precise calculation of the within-herd deviation than the one used here. This was not done here because the information on milk production was incomplete. Again, the possible existence of an interaction between the effect of within-herd milk production and time will have to be checked. Ducrocq (2005) found an important decrease over time in culling for low milk production, which could be confounded with genetic trend if not properly accounted for. The effect describing the impact of having a type score on culling risk was also found to be substantial in this data set. This is at least partly related to the fact that cows culled very early in first lactation do not have the opportunity to be scored. But it may also be due to a preference from the dairymen toward cows with a particular morphology, leading to these type-scored cows to have some level of "preferential treatment" with respect to the culling decision. Whether this effect should be kept in a genetic evaluation model is unclear.

### CONCLUSIONS

Whatever the stages of lactation definition, the piecewise Weibull baseline model led to a better fit than a unique baseline model. The piecewise Weibull baseline model was able to cope better with the cyclic change of hazard within lactations and showed smoother hazard functions. Among the scenarios considered, case C (0, 60, 250, and 350 d of cutpoints) was the best setting for stage of lactation in this study, but case D could be a simpler alternative. The next step is the extension of this model to a genetic model for sire evaluation. Potential interactions between the baselines or some of the fixed effects considered here and genetic trend should be carefully considered.

### ACKNOWLEDGMENTS

The authors thank the Hokkaido Dairy Milk Recording and Testing Association for providing data for this study. This work was carried out during the stay of the first author at the Station de Génétique Quantita-

tive et Appliquée, Institut National de la Recherche Agronomique in Jouy-en-Josas, France.

### REFERENCES

- Beauudeau, F., V. Ducrocq, C. Fourichon, and H. Seegers. 1995. Effect of disease on length of productive life on French Holstein dairy cows assessed by survival analysis. *J. Dairy Sci.* 78:103–117.
- Buenger, A., V. Ducrocq, and H. H. Swalve. 2001. Analysis of survival in dairy cows with supplementary data on type scores and housing systems from a region on northwest Germany. *J. Dairy Sci.* 84:1531–1541.
- Caraviello, D. Z., K. A. Weigel, and D. Gianola. 2004. Prediction of longevity breeding values for US Holstein sires using survival analysis methodology. *J. Dairy Sci.* 87:3518–3525.
- Ducrocq, V. 1994. Statistical analysis of length of productive life for dairy cows of the Normande breed. *J. Dairy Sci.* 77:855–866.
- Ducrocq, V. 2002. A piecewise Weibull mixed model for the analysis of length of productive life of dairy cows. *Proc. 7th World Congr. Genet. Appl. Livest. Prod., Montpellier, France* 32:505–508.
- Ducrocq, V. 2004. Illustration of a trend validation test for longevity evaluations. *Proc. Interbull Technical Workshop, Sousse, Tunisia. Interbull Bull.* 32:151–156.
- Ducrocq, V. 2005. An improved model for the French genetic evaluation of dairy bulls on length of productive life of their daughters. *Anim. Sci.* 80:249–256.
- Ducrocq, V., and J. Sölkner. 1998. The Survival Kit-V3.0: A package for large analyses of survival data. *Proc. 6th World Congr. Genet. Appl. Livest. Prod., Armidale, Australia* 27:447–448.
- Dürr, J. W., H. G. Monardes, and R. I. Cue. 1999. Genetic analysis of herd life in Quebec Holsteins using Weibull models. *J. Dairy Sci.* 82:2503–2513.
- Gröhn, Y. T., V. Ducrocq, and J. A. Hertl. 1997. Modeling the effect of a disease on culling: An illustration of the use of time-dependent covariates for survival analysis. *J. Dairy Sci.* 80:1755–1766.
- Interbull. 2005. Genetic evaluation. Direct Longevity. <http://www-interbull.slu.se/longevity/framesida-long.htm> Accessed Dec. 20, 2005.
- Larroque, H., and V. Ducrocq. 2001. Relationships between type and longevity in the Holstein breed. *Genet. Sel. Evol.* 33:39–59.
- Röxström, A., V. Ducrocq, and E. Strandberg. 2003. Survival analysis of longevity in dairy cattle on a lactation basis. *Genet. Sel. Evol.* 35:305–318.
- Röxström, A., and E. Strandberg. 2002. Genetic analysis of functional, fertility-, mastitis-, and production-determined length of productive life in Swedish dairy cattle. *Livest. Prod. Sci.* 74:125–135.
- Sewalem, A., G. J. Kistemaker, V. Ducrocq, and B. J. van Doormaal. 2005. Genetic analysis of herd life in Canadian dairy cattle on a lactation basis using a Weibull proportional hazards model. *J. Dairy Sci.* 88:368–375.
- Strandberg, E., and A. Röxström. 2000. Genetic parameters of functional and fertility determined length of productive life in Swedish dairy cattle. *Anim. Sci.* 70:383–389.
- Vollema, A. R., S. van der Beek, A. G. F. Harbers, and G. de Jong. 2000. Genetic evaluation for longevity of Dutch dairy bulls. *J. Dairy Sci.* 83:2629–2639.
- Vukasinovic, N., J. Moll, and L. Casanova. 2001. Implementation of a routine genetic evaluation for longevity based on survival analysis techniques in dairy cattle populations in Switzerland. *J. Dairy Sci.* 84:2073–2080.