

Genetic Analysis of Herd Life in Canadian Dairy Cattle on a Lactation Basis Using the Survival Kit

Asheber Sewalem¹, Gerrit Kistemaker² and Brian Van Doormaal²

¹ Agriculture and Agri-Food Canada, Guelph, ON, Canada, N1G 4T2

² Canadian Dairy Network, Guelph, ON, Canada, N1G 4T2

Introduction

Longevity is a highly desirable trait that considerably affects overall profitability. With increased longevity the mean production of the herd increases because a greater proportion of the culling decisions are based on production and the proportion of mature cows, which produce more milk than young, cows is increased. An increased longevity enables higher selection intensity of females and thus greater selection response.

Longevity can be measured in several ways and genetic evaluation systems are not standardized across countries making the comparison of sire rankings difficult. In Canada, for instance, the survival of cows in each of the first three lactations is recorded as a binary trait and evaluated with a multiple trait linear animal model in which survival in each lactation is considered as a separate trait (Jairath *et al.*, 1998). However, linear models are not appropriate for analysis of binary traits. Moreover, with the current genetic evaluation of herd life, lactation records are not used until two years after calving. This delay results in a lag time of two years or more. Survival analysis using the proportional hazard model is an alternative method for evaluation of herd life. It combines information on uncensored and censored records, which enables a proper statistical treatment of censored records and accounts for the non-linear characteristic of longevity data. Survival analysis also offers several advantages over the linear model that is currently used in Canada. 1) Precision can be increased by accounting for differences in days of productive life between cows that survive for the same number of lactations. 2) It includes censored records eliminating the need to wait for two years before using a lactation record. 3) It also tends to give higher estimates of heritability than does the linear model suggesting increased reliability of sires EBV.

In view of its advantages, routine genetic evaluation of sires based on survival analysis was implemented in several European countries, including France (Ducrocq, 1999), Germany (Pasma and Reinhardt, 1999), The Netherlands (Vollema *et al.*, 2000), Italy (Schneider *et al.*, 2000) and Switzerland (Vokasinovic *et al.*, 2001). In the above countries genetic evaluation of sires for survival of their daughters is based on fitting the model with stage by lactation as a time dependent covariate. Stage of lactation is included in the model to account for changes in culling risk within lactation (Ducrocq, 1997). The sole base line hazard function cannot account for these changes over time. Roxstrom *et al.*, (2003) reparameterized the model and defined the base line hazard function on a lactation basis. The results showed a better overall fit of the model and also a reduced number of stages of lactation.

The objective of this study were; a) to assess the most important factors influencing the herd life of Canadian dairy breeds, b) to estimate genetic parameter for longevity and c) to compare sire EBVs between the results of survival analysis and the current genetic evaluations for herd life.

Materials and Methods

The data for this analysis were obtained from lactation records extracted for the May 2002 genetic evaluation of Holstein breed. The first 10 lactation records on cows that calved from July 1985 to May 2002 were included for parameter estimation. Length of productive life *t* was defined as time from one calving to the next calving or death or culling. A lifetime record was considered to be completed (uncensored) if the cow received a termination code, indicating that she was either culled or died for any reason. Censored records

represented cows being sold, exported or leased to another herd or cows still in the herd. Cows changing herds during their productive life were considered right censored. Records associated with missing sire identification, incorrect calving dates, age at first calving outside 18-40 months and records from herds with less than 30 cows in total were excluded from the analysis. Only records from sires with 30 daughters in at least 20 herds were kept for the analysis. A total of 536,478 cows from 3,903 herds sired by 2,090 sires were used for parameter estimation.

The following model was used,

$$\lambda(t) = \lambda_{0,s}(t) \exp\{x'_m(t)\beta + z'_m u\}$$

where, $\lambda(t)$ is the hazard of a cow, i.e., her probability of being culled at time t given she is alive just before t ; $\lambda_{0,s}(t) = \lambda_0(t)^{p-1}$ is the Weibull baseline hazard function with scale parameter λ and shape parameter p and t is the time in days from one calving to the next calving for each stratum; β contains the possibly time dependent covariates affecting the hazard with $x'_m(t)$ being the corresponding design vectors; and, u is a vector of random variables with associated incidence vector z'_m . The independent covariates included in the model were as follows, **hys** is a random time dependent effect of herd-year-season class; where year of calvings were from 1985-2002 and season of calvings were (1=January-March; 2=April-June, 3=July-September and 4=October-December.). The **hys** effect was assumed to follow a log gamma distribution; **Seasp** is a fixed time dependent effect of season of production; (1= January-March; 2 = April-June, 3=July-September and 4=October-December.); **stage** is a fixed time dependent effect of stage of lactation in days (1=0-80; 2=81-235; 3 >235), **hs** is a fixed time dependent effect of the annual change in herd size with the three classes (*decreasing*, *nearly unchanged* and *increasing*); **sp** is a fixed time dependent effect of herd supervision with three classes (*unsupervised*, *supervised* and *unknown*); **afc** is a fixed time independent effect of age at first calving in months; **my**, **py** and **fy** are fixed time dependent effects of milk, fat and protein yields calculated within herd-year-parity deviation with three classes of each, *low*=for cows producing more than 0.4 standard deviation below the herd year parity

average, *average*=for cows producing between 0.4 standard deviation below and 0.6 standard deviation above the herd year parity average and *high*=for cows producing 0.6 standard deviation above the herd year parity average; **sire** is a random time independent genetic effect of the cow's sire which is assumed to follow a multivariate normal distribution with mean zero and variance $A\sigma^2_s$, where σ^2_s is the variance among sires and A is the relationship matrix. Effects included in this study were very similar to those included in the current national genetic evaluation of herd life in Canada.

Preliminary analysis of the data was done with Cox model to check the validity of the Weibull assumption and to decide if stratification by lactation should be done. The model used in the Cox model was the same as described above for the Weibull model except that the baseline was left arbitrary. One baseline hazard function $\lambda_{0,s}(t)$ was defined for each lactation (subscript 0, designates a baseline hazard and subscript relates stratum s). Detailed description of the model and survival analysis of longevity data in dairy cattle on a lactation basis was described by (Roxstrom *et al.*, 2003).

Genetic parameters obtained from the Weibull model were used to estimate the breeding values of sires. The analyses were performed using the survival kit (Ducrocq and Solkner, 1998). Correlation among sire EBVs from the current genetic evaluation of herd life and the survival analysis were calculated.

Results

From among the 536,478 records, 73.4% were right censored. The mean failure time after first calving was 368 and 266 days for censored and uncensored cows, respectively. All effects included in the model were significant. The most important change in log likelihood was observed for the effect of milk yield. The relative culling rate for cows producing 0.4 standard deviation below the herd year parity mean have 1.6 and 1.5 fold higher risk of being culled than do average producers for milk and protein yields, respectively.

A linear increase of relative culling risk was observed as age at first calving increased. The

risk of being culled was higher for older heifers than heifers calving at an age between 18 and 24 months. The change in herd size variation had the least impact on the change in the log likelihood compared to the other main effects. Cows in herds of decreasing size have a relatively higher risk of culling than herds nearly unchanged in size. However, the differences among the three classes were small.

The relative culling risk associated with the milk recording system shows that cows in unsupervised herds had 1.12 times higher risk of being culled than cows in supervised herds. A systematic difference in relative risks between seasons of production was observed. Cows were 25% more likely to be culled just before the end of quota year (April to June) than after the beginning of the new quota year (July to September).

Genetic parameters

The sire variance was 0.046 and the Weibull parameters (ρ) were 1.67861, 1.7790, 1.67834, 1.60656, 1.53866, and 1.43232 for lactation 1, 2, 3, 4, 5 and 6+, respectively. The sire variance obtained in this analysis was almost similar to the results obtained using the entire length of herd life (0.049) Sewalem *et al.* (2001).

Breeding values for each sire were estimated using the above parameters. In the survival analysis EBVs of sires were estimated using the first three lactations and all lactations of each cow. This was an attempt to have an EBV also based on the same number of lactations as the current genetic evaluation of direct herd life. The correlation between EBVs obtained from the survival analysis using all lactations of each cow and the current genetic evaluation of direct herd life was 0.72. A similar correlation result (0.72) was reported by Boettcher *et al.* (1999) and Sewalem *et al.* (2001). When using the first three lactation in the survival kit the correlation between bulls EBVs increased to 0.83 for bulls with at least 250 culled daughters.

Even so, these correlations show that there is a difference between the two methods but it does not show which method is better. Therefore, consistency and predictability of the two methods were checked and compared.

Consistency:- In each of the two methods, the whole data set was randomly divided in two subsets by herd. Analysis was carried out for each subset of the data. Correlations between bull proofs from the two data sets using the same method and requiring different minimum number of daughters per sire were calculated. The correlation ranged from 0.56 to 0.96 in both methods depending on the number of daughters per sire culled for survival analysis and survived for the current genetic evaluation. The survival model seems slightly more consistent than the current methods but differences are small and cannot explain the low correlation between the two models.

Predictability:- Data, 3 years apart, extracted for February 2000 and May 2003 genetic evaluation, were used for each of the two methods to estimate breeding value of sires. In the current method the correlation ranged from 0.87 with all sires to 0.98 when the number of daughters reached to 500. The corresponding figures for the survival analysis were 0.89 and 0.96.

Genetic trend

Genetic trends in average breeding values for the two methods are shown in Figure 1. Both methods show that there is a favorable trend for herd life over time. Moreover, in the survival analysis there seems to an over-estimation of breeding values for young bulls. This might be due to censored records. In the last ten years (Figure 1) there is a gain of 0.158 genetic standard deviation units per year using the survival kit compared to 0.038 in the current evaluation of herd life.

Conclusions

1. All effects fitted in the model had a significant effect on herd life.

2. Analysis of herd life data on a lactation basis seems more correlated with the current national genetic evaluation of herd life than analysis of herd life based on the entire length of productive life.
3. The two methods have a similar consistency and predictability.
4. The genetic trend obtained from the survival kit seems to be over estimated.

References

- Boettcher, P.J., Jairath, L.K. & Dekkers, J.C.M. 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.
- Ducrocq, V. & Sölkner, J. 1998. “The Survival Kit–V3.0”, a package for large analyses of survival data, in: *Proceeding of the 6th World Congress on Genetics Applied to Livestock Production*, 11-16 January, Vol. 27, University of New England, Armidale, pp. 447-448.
- Ducrocq, V. 1997. Survival analysis, a statistical tool for longevity data, Paper presented at 48th Annual Meeting of the European Association for Animal Production (EAAP), 25-28 August. Vienna, Austria.
- Ducrocq, V. 1999. Topics that may deserve further attention in survival analysis applied to dairy cattle breeding: some suggestions, in: Proceedings of Workshop on genetic improvement of functional traits in cattle (GIFT), Jouy-en-Josas 1999, *Interbull Bulletin* 21, 181-189.
- Jairath, L., Dekkers, J.C.M., Schaeffer, L.R., Liu, Z., Burnside, E.B. & Kolstad, B. 1998. Genetic Evaluation for herd life in Canada. *J. Dairy Sci.* 81, 550-562.
- Pastman, E. & Reinhardt, F. 1999. Genetic evaluation for length of productive life of Holstein Cattle in Germany. *Interbull Bulletin* 21, 55-59.
- Roxstrom, A., Ducrocq, V. & Strandberg, E. 2003. Survival analysis of longevity in dairy cattle on a lactation basis. *Genet. Sel. Evol.* 35, 305-318.
- Schneider, M.P., Canavesi, F. & Samore, A.B. 2000. Genetic evaluation for functional herd life in Italian Holstein. 51th Annual Meeting of EAAP, Den Haag (Book abstract) 6:34.
- Sewalem, A. & Kistemaker, G. 2003. Genetic evaluation of herd life in Canadian dairy cattle. Technical Report to the Genetic Evaluation Board.
- Vukasinovic, N., Moll, J. & Casanova, L. 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.
- Vollema, A.R., van der Beek, S., Harbers A.G.F. & De Jong, G. 2000. Genetic evaluation for longevity of Dutch dairy bulls, *J. Dairy Sci.* 83, 2629-2639.

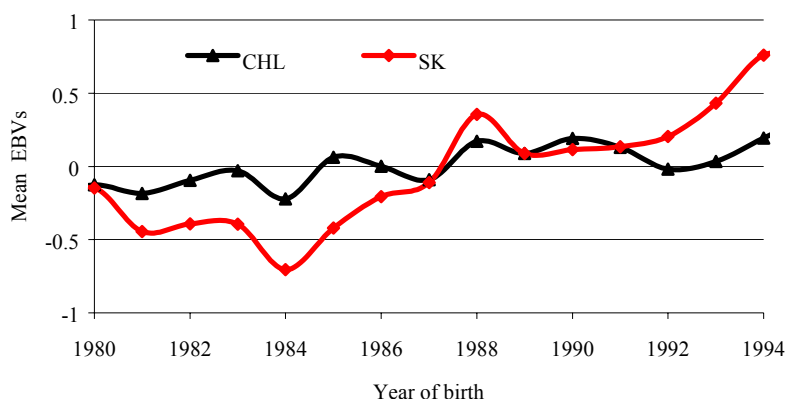


Figure 1. Genetic trends of longevity (in genetic standard deviations) using the survival kit (SK) and the current genetic evaluation of herd life in Canada (CHL).