COMPETING RISKS ANALYSIS OF REASONS FOR DISPOSAL IN QUEBEC DAIRY HERDS

J.W. Dürr[1], M.P. Schneider[2], H.G. Monardes[3] and R.I. Cue[3]

[1] Universidade de Passo Fundo, Bairro São José, C.P. 611, 99001-970, Passo Fundo, RS, Brazil
[2] A.L.E.Co.L., Rivadavia 1350, 3080, Esperanza, Santa Fe, Argentina
[3] McGill University, 21,111 Lakeshore, Ste-Anne-de-Bellevue, QC, H9X 3V9, Canada

INTRODUCTION

Different aspects of survival analysis applied to dairy cows have been discussed in the literature, but in most cases the reason for failure is not taken into consideration. When failure reasons are of interest and must be accounted for in the analysis, the method of competing risks (Kalbfleisch and Prentice, 1980) offers an intuitive but powerful way of handling survival times. The general concept underlying competing risks analysis is that the occurrence of one type of event removes the individual from risk of all other types of events. In this case, the overall hazard function is the sum of all the type-specific hazards, which can be obtained in the same way as the regular hazard function, just regarding all failures of types other than the recorded reason for failure as censored at the individual's failure time (Allison, 1995). The competing risks approach is of particular interest in the study of reasons for disposal in dairy cattle. Cows being culled for different reasons were certainly affected by a distinct set of covariates or by the same covariates in different degrees (Dürr *et al.*, 1997), and modeling their hazards in a competing risks framework seems to be a very appropriate approach.

MATERIALS AND METHODS

The lifetime records used in this study were the same 331147 records used by Dürr *et al.* (1999). Disposal reasons were defined based on the Programme d'Analyse des Troupeaux Laitiers du Québec (PATLQ). Only those disposal reasons of higher incidences were studied, namely culling due to low milk or low fat production (LOWP), culling due to reproductive problems (REPRO), culling due to mastitis and/or high cell counts (MAST), culling due to udder breakdown and milking problems (UDBR) and culling due to feet and leg problems (F&L). A sixth class (INVOL) of culling reasons was defined including all disposal codes but LOWP. This general reason is a crude approximation of involuntary culling, if voluntary culling is assumed to be based only on production. Note that INVOL is a competing risk only for LOWP, since it includes all the other reasons. A parametric model was used to analyze the effect of different covariates on the failure time of each competing risk: $\lambda_d(t) = \lambda_{0d}(t) \exp\{y_i(t')+p_j(\tau)+z_k(t')+o_l+a_m+w_r(\zeta)+h_n(t')+s_q\}$, where $\lambda_d(t)$ is the cause-specific hazard function for disposal reason d, at time t ; $\lambda_{0d}(t)$ is a cause-specific Weibull baseline hazard function with scale parameter λ_d and shape parameter ρ_d; $y_i(t')$ is the effect of year i (i = 1981,...,1994), assumed to be piecewise constant with jumps arbitrarily chosen to occur at t'= March 1 of each year; $p_j(\tau)$ is the effect of lactation number and stage of lactation combined into j = 16 classes (lactation 1, 2, 3 and 4 × four stages), assumed to be piecewise constant with changes occurring at = 0, 120, 240, and 305 days of each lactation; $z_k(t')$ is the effect of annual change in herd size k (k = 1, for a decrease in herd size of >25%; k = 2, for a decrease in herd size of 15 to 25%; k = 3, for a decrease in herd size of 5 to 15%; k = 4, for herd with no appreciable change (-5 to +5%), k = 5, for an increase in herd size of 5 to 15%; k = 6, for an increase in herd size of 15 to 25%; and k = 7, for an increase in herd size of >25%), which is assumed to be a time-dependent covariate, piecewise constant, and jumps happening at t'= March 1 of each year; o_l is the effect of the l milk recording option (l = 1, for owner-sampler herds; l = 2, for official herds); a_m is the effect of the m age at first calving (m = 1, for 17 months;...; m = 19, for 40 months); $w_r(\zeta)$ is the effect of the r^{th} within herd-year-parity class of milk production at 305 days of lactation (r = 1, for cows producing more than 1.5 standard deviations below the herd-year-parity average; r = 2, for cows producing between 1.5 and 0.5 standard deviations below the herd-year-parity average; r = 3, for cows producing between 0.5 standard deviation below and 0.5 above the herd-year-parity average; r = 4, for cows producing between 0.5 and 1.5 standard deviations above the herd-year-parity average; and r = 5, for cows producing more than 1.5 standard deviations above the herd-year-parity average), considered to be piecewise constant and changing value at ζ = beginning of a new lactation; $h_n(t')$ is the random effect of the herd-year n (n = 1,...,28629), assumed to be piecewise constant with changes at t' = March 1 of each year; and s_q is the random effect of sire q (q = 1,...,1664). The effect $w_r(\zeta)$ was not included in the model for LOWP, because it would be confounded with the dependent variable. The Weibull model was analyzed with the "SURVIVAL KIT" (Ducrocq and Sölkner, 1994). A log-gamma prior density function was assumed for the herd-year random effect and a multivariate normal distribution with covariance between levels being introduced by genetic relationships was assumed for the random effect of sire. The pedigree file included only information on male parents (sires) and included a total of 1875 animals (1664 with data). The sire variance σ^2_s was estimated as the mode of its marginal posterior density, which was approximated by Laplacian integration. The gamma parameter γ_h was estimated jointly with the other effects after exact algebraic integration of the log-gamma random effect of herd-year.

RESULTS AND DISCUSSION

The amount of censoring was really high for all competing risks, except for INVOL, which presents a reasonable proportion of uncensored records (35,2%). This low incidence for each individual culling code, however, did not prevent the Weibull model from detecting differences in the hazard rates and demonstrating how the failure time for each competing risk is affected by the covariates included in the model. Table 1 shows the likelihood ratio tests for the fixed effects included in competing risks analysis. Annual change in herd size had the smallest impact on the failure time of all competing risks, not reaching statistical significance for any of the disposal reasons. Age at first calving was also not significant at $P < 0.001$ for MAST and UDBR. The covariate with the largest impact was always lactation number × stage of lactation, followed by the effect of 305-day yield deviation. The change in the log-likelihood caused by the effect of 305-day yield deviation was 6-fold smaller in the model for INVOL than the change caused by the same covariate in the model for functional herd life in Dürr *et al.* (1999), demonstrating that censoring records of cows culled for LOWP drastically reduces the variation in the failure time explained by within herd-year-parity yield deviation. Interestingly, there is still a significant change in log-likelihood caused by 305-day yield deviation in the model for INVOL, indicating that culling due to low production is not the only disposal reason affected by production level.

Table 1. Results from the likelihood ratio tests comparing the full model with models excluding one effect at a time

Effect	-2 Change in Log-likelihood[a]						df
	LOWP	REPRO	MAST	F&L	UDBR	INVOL	
Year	2358.6*	139.8*	259.9*	501.1*	243.1*	612.2*	13
Lactation × Stage	44497.0*	23814.0*	1917.4*	1801.7*	5424.2*	15217.0*	15
Change in Herd Size	21.8	17.3	7.7	12.6	2.2	23.52	6
Age at First Calving	121.80*	341.5*	26.5	159.5*	29.2	503.4*	18
305d Yield Deviation[b]	-	289.5*	1406.2*	1117.9*	2174.8*	7709.9*	4
Recording Option	265.6*	6.9	31.7*	63.4*	61.7*	24.8*	1

[a]Logarithm of the marginal posterior odds ratio at the posterior mode.
[b]Effect not included in the model for LOWP.
*Significant at P<0.001.

Estimates for the Fixed Effects. The hazard functions showed that the risk of failure due to low production is very high in the first 240 days of first lactation, and then decreases to a very low level for the rest of the lactation. In second and later lactations, the hazard rate for LOWP starts at a low level and then becomes high from 121 to 240 days in milk, when it drops and stays low until the cow reaches the same stage in the next parity. Since official production certificates in Quebec are only issued once cows have reached 240 days in milk, and also because the official herd production average includes only cows with more than 240 days in milk, herd managers do cull their poor producers before 240 days, using a legitimate marketing strategy to make their herds look better. The estimated hazard rate for INVOL is similar to the estimates for functional herd life (Dürr *et al.*, 1999), except that the hazard from 121 to 240 days after calving is much lower for INVOL. The explanation is simple: the higher hazard between 121 and 240 days for functional herd life was due to the higher risk of being culled for low production, which is not present in the hazard estimates for INVOL. This is a very important finding, because it indicates that the adjustment for herd-year-parity class of milk production at 305 days does not account for all voluntary culling based on low production. In other words, using the competing risks approach might be a better representation of the hazard experienced by a dairy cow, regardless her production, than Functional Herd Life. While the hazard rate is highest at the end of the lactation and dry period for INVOL, cows are at a higher risk of being culled for MAST, UDBR and F&L between 121 and 240 days after calving. This difference happens because INVOL includes REPRO, which is highly concentrated at the end of the lactation and has a higher incidence than the other reasons for disposal. Estimates of the year effect show that the only competing risk with a clearly descending trend for culling rates in the period studied is LOWP as it had been already described by Dürr *et al.* (1997). The relative culling rate for INVOL had a conclusively ascending trend from 1982 to 1994. The effect of age at first calving for LOWP is higher than the average for heifers calving at a very young age (17 to 22 months), then decreases and stays constant from 23 months to 34 months of age. After 34 months, the relative culling rate increases rapidly up to 30% above the average for cows calving at 39 months of age. The effect of age at first calving did not significantly affected MAST and UDBR, but it had a clearly ascending trend for INVOL, REPRO and F&L. Estimates of the effect of 305-day herd-year-parity yield deviation for the different competing risks show that yield deviation from the herd-year-parity average significantly affects culling for reasons other than low production. The lower the relative production level of the cow, the higher the risk of being culled for whatever reason. The risk of being discarded for REPRO, LOWP, UDBR and F&L is higher in official than in owner sampler herds and MAST is the only competing risk in which the risk of being culled is higher in owner sampler herds.

Estimates for the Random Effects. The estimates of the sire variance ranged from 0.03 to 0.15, and the gamma parameter for the random effect of herd-year from 1.38 to 7.17. With the exception of the estimates for INVOL, these values must be interpreted with extreme caution. LOWP has only 20% of records which are uncensored, and this proportion drops to approximately 10% for REPRO and 5% for MAST, UDBR and F&L. Therefore, the amount of information available to estimate sire variances is really limited and little confidence could be granted to genetic evaluations based on these information.

CONCLUSION

Competing risks analysis is well suited for studies of culling trends in dairy cattle populations, providing an intuitive way of describing the impact of different covariates on the failure time and, at the same time, a solid theoretical framework for hypothesis testing.

REFERENCES

Allison, P. D. (1995) “Survival Analysis Using the SAS® System: A Practical Guide”. Cary, NC: SAS Institute Inc.

Ducrocq, V. P. and Sölkner, J. (1994) *Proc. 5th WCGALP* **XXII** : 51-52.

Dürr, J.W., Monardes, H.G., Cue, R.I. and Philpot, J.C. (1997) *Can. J. Anim. Sci.* **77** : 601-608.

Dürr, J.W., Monardes, H.G. and Cue, R.I. (1999) *J. Dairy Sci.* **82** : 2503-2513.

Kalbfleisch, J. D. and Prenntice, R. L. (1980). “The Statistical Analysis of Failure Time Data”. John Wiley & Sons, Inc., New York, NY.