Nonlinear transformations models: Application to mortality of calves.

L. López-Segovia[1], A. Espinal[2], G. Gómez[3]

[1] Ciencias Básicas, UJAT
[2] Servei d'Estadística, UAB
[3] Statistics Dept., UPC

Abstract: We present an application of the proportional hazard-proportional hazard cure models to evaluate the environmental and genetic factors that affect both mortality and survival up to weaning of beef calves. We use the ntlm package (software R) to fit nonlinear semiparametric transformation models, which include the proportional hazard cure models as a special case. Results indicate that genetic factors such as calving month and calving difficulty are associated with the long-term survivorship of the calves.

Keywords: Proportional hazard cure (PHC) model; Extended hazard (EH) models; Proportional hazard-Proportional hazard cure (PH-PHC) model; Nonlinear transformations (NLT) models.

1 Introduction

Mortality of calves from birth to weaning (approximately at 180 days) reduces farm's income, and significantly increases cattle production costs (see Goyache et al. (2003) for a review). Thus, it is important to take into account the survival pattern of calves into the overall breeding. Tarrés et al. (2005) used a standard survival analysis to study how genetic and environmental factors influence mortality up to weaning. However, and due to the high proportion of censoring in the data, one could think of the presence of a mixture of two subpopulations of calves: those susceptible to die before weaning and those who don't. A binary mixture model, also known as cure model, (e.g. Farewell, 1982) which takes into account a fraction of *cure* individuals, could be appropriate in this situation. In this paper we show an application of the extended hazard (EHM) models proposed by Tsodikov (2002) which are developed to combine both *long-term* and *short-term* effects. EHM models include as a particular case the proportional hazard cure models. We fit a proportional hazard - proportional hazard cure (PH-PHC) model to fit both genetic and environmental factors and discriminate between mortality of calves effects (*short-term* effects) and survival or cure effects (*long-term* effects).

TABLE 1. Characteristics of Calves

Factors	herd1 (%)	herd3 (%)	herd7 (%)
`Lpl`			
$< 1300d$	360 (58.72)	263 (63.52)	498 (47.42)
$> 1300d$	253 (41.27)	151 (36.47)	552 (52.57)
`Month`			
sep-feb	204 (33.27)	319 (77.05)	650 (61.90)
mar-aug	409 (66.72)	95 (22.94)	400 (38.09)
`Gender`			
female	308 (50.24)	216 (52.17)	544 (51.80)
male	305 (49.75)	198 (47.82)	506 (48.19)
`Difficulty`			
without assistance	299 (48.77)	379 (91.54)	916 (87.23)
slightly assisted	38 (6.19)	25 (6.03)	50 (4.76)
strongly assisted	3 (0.48)	9 (2.17)	83 (7.90)
missing	273 (44.53)	1 (0.24)	1 (0.09)
total	613 (29.51)	414 (19.93)	1050 (50.55)

and difficulty at birth are the set of statistically significant factors for the nonsusceptible proportion (*long-term effects*) of calves for herd 1, calving difficulty is the only significant factor for herd 7, and there are no significant predictors among this set of covariates for herd 3. We point out that the interpretation of the regression parameters for the cure fraction $\pi(z)$ is such that a higher value for e^{β} would represent a lower probability of cure for the corresponding factor. Note that model (3), together with $\eta(z) = \exp(\beta_\eta z)$ and $\theta(z) = \exp(\beta_\theta z + \beta_c)$, implies that $\pi(z) = (\pi_0)^{e^{\beta}}$, where π_0 represents the probability of cure of the reference group. In particular, calves born in the period march-august have lower probability of cure than those born in september-february; and the probability of cure is a well lower for those that have had difficulties at calving for herd 1. For herd 7 the effect of `difficulty` is the same as for herd 1.

Regarding short-term (mortality) effects, we only find statistically significant predictors in herd 7 where the risk of death of calves born to older mothers, hence with a longer reproductive life, is twice the risk of death of calves born to younger mothers ($\beta_\eta = 0.89$, $e^{\beta_\eta} = 2.44$, p-value $= 0.056$). Due to a complete parametrization of the probability of cure (survival up to weaning) $\pi(z)$, we can estimate it for each of the categories of the significant covariates for the long-term effects given in Table 2. Table 3 displays the estimate, and confidence interval, for the probability of cure for the different groups for each herd, obtained through the relationship $\pi(z) = (\pi_0)^{e^{\beta}}$. We observe lower probabilities of cure for calves born between March and

TABLE 2. Statistical significant factors for mortality and cure for each herd using a PH-PHC model. Reference group for herd 1: calves born between September and February and without assistance, for herd 7: calves born without assistance.

Predictors	β	e^{β}	se(β)	p	$L_{.95}$	$U_{.95}$
			herd1			
Long term predictor						
`Month`						
mar-aug	1.96	7.097	0.989	0.047	1.022	49.263
`Difficulty`						
slightly assisted	1.87	6.476	0.474	0.000	2.557	16.401
strongly assisted	2.17	8.716	1.051	0.039	1.110	68.386
			herd7			
Long term predictor						
`Difficulty`						
slightly assisted	0.01	1.007	1.027	0.990	0.134	07.549
strongly assisted	1.33	3.798	0.471	0.004	1.507	09.569
Short term predictor						
`Length productive`						
>1300 days	0.89	2.440	0.466	0.056	0.978	06.080

August for herd 1 and for calves born with assistance for herds 1 and 7. Furthermore, note that herd 7 is the only herd for which the length of productive live of the cow has an influence on the risk of death of the calves, and this short-term effect is influencing the probability of cure (survival up to weaning) in such a way that the confidence interval for those calves born with strong assistance $(.887, .953)$ is strictly below the confidence interval for calves born without assistance $(.968, .987)$. Thus,the probability of survival up to weaning of calves born without assistance is significatively higher than the probability of survival up to weaning of calves born with strong assistance.

Concluding, we point out that the PH-PHC model is an alternative to the standard Proportional Hazards model when there is a proportion of non-susceptible individuals in the population. This model allows us to jointly estimate the proportion of cure (survival up to weaning) and the effect of different set of covariates for short and long-term on individuals in a heterogeneous population. Moreover, we have been able to use the same approach for the three different herds, providing a unified method for situations, such as the one described in this paper, where the initial set of covariates has different short-long effects on each herd.

TABLE 3. Estimates of the Probability of Cure $\pi(z)$ and 95% Semiparametric Likelihood Ratio Confidence Intervals (in parentheses). Reference group for herd 1: calves born between September and February and without assistance, for herd 7: calves born without assistance.

Predictors	herd1 $(L_{.95}, U_{.95})$	herd3 $(L_{.95}, U_{.95})$	herd7 $(L_{.95}, U_{.95})$
Reference Group	.993 (.955, .999)	.975 (.955, .986)	.980 (.968, .987)
Month			
mar-aug	.953 (.723, .992)		
Difficulty			
slightly assisted	.957 (.743, .993)		.981 (.968, .987)
strongly assisted	.942 (.671, .991)		.930 (.887, .953)

Acknowledgments: This research was partially supported by Grants MTM2008-06747-C02-00 and MTM2009-10893 from the Ministerio de Ciencia e Innovación.

References

Farewell, V.T. (1982). The Use of Mixture Models for the Analysis of Survival Data with Long-Term Survivors. *Biometrics*, **38**, 1041-1046.

Goyache, F., Gutiérrez, J.P., Alvarez, I., Fernández, I., Royo, L.J., and Gómez, E. (2003). Genetic analysis of calf survival at different preweaning ages in beef cattle. *Livest. Prod. Sci*, **83**, 13-20.

Sy, J.P., and Taylor, J.M.G. (2000). Estimation in a Cox Proportional Hazards Cure Models. *Biometrics*, **56**, 227-236.

Tarrés, J., Casellas, J., and Piedrafita, J. (2005). Genetic and envieromental factors influencing mortality up to weaning of bruna dels pirineus beef calves in mountain areas. A survival analysis. *Animal Science*, **83**, 543-551.

Tsodikov, A.D. (2002). Semi-parametric models of long- and short-term survival: an application to the analysis of breast cancer survival in Utah by age and stage *Statist. Med.*, **21**, 895-920.

Tsodikov, A.D., Ibrahim, J.G., and Yakovlev, A.Y. (2003a). Estimating Cure Rates From Survival Data: An Alternative to Two-Component Mixture Models. *Journal of the American Statistical Association*, **98**, 464, 1063-1078.

Tsodikov, A.D. (2003b). Semiparametric models: a generalized self-consistency approach. *Journal Royal Statistical Society B*, **65**, Part 3, 759-774.