Improvements in genetic evaluation for longevity for the Italian Holstein.

***Fabiola Canavesi*[1]*, S. Biffani*[1] *and A.B. Samorè*[1]**
1 – Associazione Nazionale Allevatori Frisona Italiana, Cremona, Italy

---

Abstract

Transmitting abilities for longevity have been officially published for the Italian Holstein Friesian using survival analysis since August 2001. Productive life is calculated from national milk recording information as the number of days between first calving and the last milk recorded test day. The hazard function applied is explained by a baseline hazard function, only depending on time and equal for all animals, and fixed and random effects, possibly time dependent. Effects included in the model are stage of lactation by lactation number, age at first calving, annual change in herd size, herd-year-season (random), year-season (fixed), production level for milk, fat and protein, and random effect of the sire. This model, routinely used for the official sire proofs estimation for direct longevity, was recently improved: the sire model was modified into a sire maternal-grand-sire model and the milk production level was included in the model using a better measure of 305 days production and a higher number of production classes than before. The improvement in the adjustment for milk production led to an estimated heritability of 0.09. The introduction of the maternal grand sire increased the estimated heritability to 0.097 improving the overall accuracy of the system. Published longevity sire proofs are the result of the combination of direct longevity and type traits. The correlation with the official genetic sire evaluation was around 0.92.

Introduction

Longevity is nowadays the most important trait after production in the majority of selection indices all over the world. In spite of the complexity of the trait and of the difficulties in measuring it, the availability of tools like the Survival kit now allows to better estimate genetic values for this important economical trait in many countries (Vollema et al, 2000; Vukasinovic et al., 2001).

In Italy the current model using a Weibull Sire model is official since August 2001. The hazard function is described as the product of a baseline hazard function, depending only on time, and of an exponential function of risk factors, described by fixed and random effect, mostly generally time dependent. Fixed effect in the model are: stage of lactation within lactation number, age at first calving (time independent), year-season, 305d milk production class (5 classes), fat percent production class (5 classes), protein percent production class (5 classes), all computed within herd-year. Random effects considered are sire and herd-year-season effects (Schneider et al, 2000).

Voluntary culling for production is accounted for in the model using 305d production classified in 5 classes based on within herd-year deviations for milk, fat and protein percentages.

Recent studies by Samoré et al. (2003) have shown that the adjustment for production level in the model can be further improved. Preliminary tests by Canavesi et al. (2003) have shown some inaccuracies created by the 305 days ME production calculation system that does not guarantee a fair comparison for very short lactation in progress (< 30 DIM) and short closed lactations (< 305 DIM).

The objectives of this study were: a) to improve the official genetic evaluation model for direct longevity;
b) to verify the impact of the changes on bull rankings.

Material and methods

Data from the most recent August 2004 official genetic evaluation for functional longevity (Table 1) were used to compare the official method for genetic evaluation (OFF) with a new improved model (NEW) in terms of genetic correlations among proofs and re-ranking of bulls and cows.

The data set consisted of 2,119,993 Holstein cows calving for the first time between January 1980 and December 1999. The length of productive life was defined as the number of days between first calving and last test day record information, assumed to be the culling date because culling dates and culling codes are not systematically recorded and reliable. Cows were considered uncensored (complete information) if test day records were not available at least 6 months before the end of the study period. The rest of the cows were assumed to be alive and were treated as censored. Cows in herds decreasing in size more than 50% were treated as censored. For cows changing herds during their productive life, only the first part of records was included and treated as censored. The proportion of right censored records was 30%. The age at first calving was restricted between 18 and 42 months of age. Only herds with an average of 15 cows were included. The pedigree file consisted of 5,316 Holstein sires. Sires were required to have 20 daughters in 10 herds. Survival Analysis was performed using the Survival Kit V3.0 and updates by Ducrocq and Sölkner (1998a). The following Weibull model is used for the official genetic evaluation (OFF):

$$\lambda(t) = \lambda_o(t)\ \exp\{ s_i(\tau) + m_j(\zeta) + p_k(\zeta) + f_l(\zeta) + a_m + z_n(t') + l_o(t') + h_p(t') + s_q \} \qquad \textbf{(1)}$$

Where, $\lambda(t)$ is the hazard of a cow, t days after calving,
- $\lambda_o(t) = (\lambda\rho\ (\lambda t)^{\rho-1})$ is the Weibull baseline hazard function, with scale parameter λ, and shape parameter $\rho = 2$,

- $s_i(\tau)$ is the time-dependent effect of lactation number by stage of lactation i (lactation 1, 2, 3, 4, 5 and 6 + and 5 stages of lactation), assumed to be piecewise constant with changes occurring at 0, 90, 180, 270 and 365 days of each lactation,

- $m_j(\zeta)$, $p_k(\zeta)$ and $f_l(\zeta)$ are the time-dependent classes of within herd-year deviations of milk yield j (305-day), protein content k (305-day) and fat content l (305-day), which are assumed to be a time-dependent covariate piecewise constant and changing value at the beginning of a new lactation. Five classes were considered, < -1.5 standard deviation, between –1.5 and –0.5 standard deviation, between –0.5 and +0.5 standard deviation, between +0.5 and +1.5 standard deviation and > +1.5 standard deviation.

- a_m is the time-independent effect of age at first calving (25 classes, from 18 to 42 month of age),

- $z_n(t')$ is the time-dependent effect of annual change in herd size n. Decrease in herd size > 50%, decrease of 15 to 50%, stable size (-15 to +15%), increase of 15 to 50%, and increase > 50%; this variable is assumed to be a time-dependent covariate, piecewise constant, with changes on April 1st of each year,

- $h_p(t')$ is the random time-dependent effect of herd-year-season p, assumed to be piecewise constant with changes at January 1st, April 1st, July 1st and October 1st. Herd-year-season effects were assumed to follow a log-gamma distribution, which was algebraically integrated out during the analysis.

- $l_o(t')$ time-dependent effect of year-season l, assumed to be piecewise constant with changes at January 1st, April 1st, July 1st and October 1st. Year-season is included in the model simultaneously with $h_p(t')$ in a way to specify that the mean of the random herd-year-season can vary with time (Ducrocq, 1999).

-s_q is the random time-independent effect of sire of the cow. Sire effects were assumed to follow a multivariate normal distribution with variance-covariance $\mathbf{A}\,\sigma^2_s$.

Sire variance for the official model was estimated equal to 0.021 that corresponded to an heritability on the actual scale of 0.082.

The new model was an improvement of the above with few changes, namely:

a) 305 ME milk production for all lactation records used in the evaluations;
b) 9 classes of production level for milk kg;
c) maternal grand-sire effect added to the model.

In the new model all records, including lactation in progress with less than 30 DIM and closed lactation with less than 305 DIM, are projected, while in the official model this two categories of lactation are not projected.

In the new model nine classes for production are identified based on deviations from the within herd-year mean such as class 1 corresponds to animals that have a production level below -1.5 SD from the mean, class two was between –1.5 and –1.0 SD, class 3 between –1.0 and -0.6 SD, class 4 between -0.6 and –0.3 SD, class 5 between –0.3 and 0.3, class 6 between –0.3 and 0.6 SD, class 7 between 0.6 and 1.0, class 8 between 1.0 and 1.5 and class 9 above 1.5 SD.

The new model was as:

$$\lambda(t) = \lambda_o(t)\ \exp\{ s_i(\tau) + m_j(\zeta) + p_k(\zeta) + f_l(\zeta) + a_m + z_n(t') + l_o(t') + h_p(t') + s_q + \tfrac{1}{2}\, mgs_r\}$$

Where all the effects are the same as in equation (1) and where s and mgs are random effects of the sire and of the maternal grandsire of the cow. When the maternal grandsire is not known, the term ½ mgs is ignored. The sire and mgs effects are grouped in a vector **s** and are assumed to follow a multivariate normal distribution with a variance-covariance matrix $\mathbf{A}\,\sigma^2_s$.

Risk rate for milk production levels resulting from the two models were compared. Impact on bull ranking was estimated calculating simple and rank correlations between production and longevity EBVs. Mace procedures were also used to estimate genetic correlations between production and direct longevity proofs.

Table 1 – Descriptive statistic of the data set used for the two evaluations.

	N
Total number of elementary records	68,157,394
Number of cows	3,080,175
Number of sires	8349
Right censored records	799424
% censored records	25.9

Results and discussion

Sire variance for the new sire-maternal grandsire model was estimated from subset of the data and resulted equal to 0.02489, much higher than the estimates obtained using the official model (Schneider et al., 1999). This resulted in an heritability of 0.097 on the normal scale. An heritability estimated of around 0.09 was obtained from a sire model using the same structure with respect to 305 DIM projection of records and number of production used in the model for milk kg as in the new model.

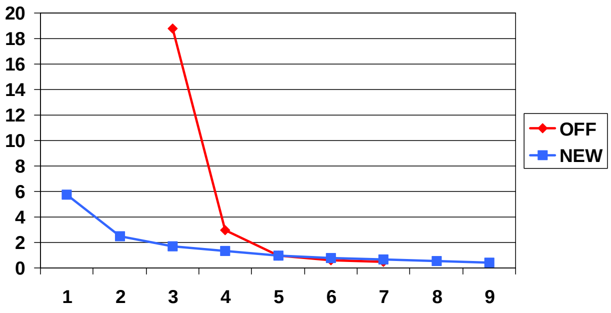


Figure 1 – Risk ratio for the five classes of production.

Figure 1 reports the relative culling rate for the five or nine classes of production. The two sets of solutions are aligned with respect to the middle class. It is possible to observe that the official model gives the lowest production class a 19-fold higher risk of being culled compared to the average class. With the new model the low production class has a 5.77 higher risk of being culled compared to the average class. This value appears much more reasonable and more comparable with values found for other populations (Vollema and Groen, 1998; Vukasinovic et al., 1997) . It is difficult to demonstrate whether this was due more to the fact that many lactation records were not projected to 305 DIM and thus not fairly compared with the others or to the class structure.

Table 2 illustrates correlations between the ETA for the direct functional longevity estimated with the official and the new model and production traits EBVs. Data used were restricted to national proven bulls. Functional longevity is expressed as relative culling rate (RCR) with a reversed sign indicating bull with higher values as the bulls with lower RCR values.

Proofs correlations are moderate but negative with production traits indicating an unfavourable correlation between high production and functional longevity and they do increase slightly with the improved model.

Table 2 – Correlation of direct longevity proofs with production traits EBVs.

	Milk	**fat**	**protein**
OFF	-0.21	-0.17	-0.25
NEW	-0.25	-0.21	-0.31

Genetic correlations among functional longevity and production traits were estimated using bulls proofs and programs provided by Interbull. Table 3 shows the estimated genetic correlations between milk and direct functional longevity proofs. When estimated at genetic level the correlation tend to decrease with the new model although the difference is very small and not significant.

Table 3 – Estimated genetic correlation of direct functional longevity with milk production EBVs.

	OFF	**NEW**
MILK	-0.238	-0.218

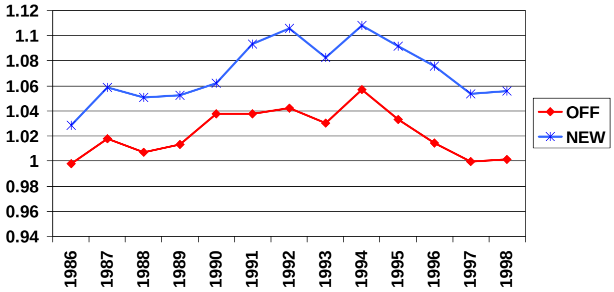
Correlation between old and new model ETA for direct functional longevity proofs is around 0.90. Rank correlation is 0.89.

Direct longevity is not officially published in Italy. Official proofs are the combined longevity using MACE procedures, considering different traits instead of same trait in different countries as it is for the international evaluation. This ignores non-genetic correlations among de-regressed proofs but helps in improving early prediction of functional longevity (Druet et al, 1999). Direct longevity ETAs are combined with udder traits and feet&legs functionality to give the final ETA for functional combined longevity. Final proofs are then standardized and have an average of 4.

The correlation among combined longevity proofs is around 0.92 due to the correlated type proofs that are the same in both indexes. Rank correlation is 0.915 and leads to change of 42 bulls out of the best 100 bulls for longevity.

Proofs resulting from the new model have been validated for genetic trend and have been submitted to Interbull Centre for test run on International genetic evaluation for longevity traits. Figure 2 shows genetic trend expressed as relative culling rate for the two models. The general trend with the new model seems to be slightly more negative with relative culling rate increasing more over time compared to the current one.

Figure 2 – Estimated genetic trend expressed as



relative culling rate (RCR).

The new model will be used for official genetic evaluation for functional longevity in Italy starting from November 2004.

Conclusions

The new model is an improvement of the survival analysis model used for the official genetic evaluation for direct longevity in the Italian Holstein. Improvement were made on the projection of lactation records to 305 days and in the structure of within herd-year production classes for milk kg. The maternal grandsire effect was also added to the random part of the model. This resulted in a higher sire variance estimate which should improve the accuracy of the evaluation.

Recent research and applied results have shown that the use of a baseline hazard function defined within lactation can further improve the model and lead to a better estimation of genetic trend with respect to changing survival condition over time. A new version of survival kit (5.1) is now available and it is already applied in France and under test in The Netherlands and Germany (Ducrocq, 2004; Van der Linde C., et al., 2004). Further improvement of the model will explore the possibility of using a more flexible baseline hazard function in the Italian population.

References

Druet T., Sölkner J., Groen Ab F. and Gengler N., 1999. Improved genetic evaluation of survival using MACE to combine direct and correlated information from yield and functional traits. Interbull bulletin n. 21:122-127.

Ducrocq, V, 2004. Illustration of a trend validation test for longevity evaluation. Interbull bulletin n. 32: 151-156.

Ducrocq, V., Sölkner, J., 1998. The Survival Kit – a Fortran package for the analysis of survival data. Proc. 6th World Congr. Genet. Appl. Livest. Prod., Armidale, Australia. 27, 447-448.

Samorè A.B., Schneider M. del P., Canavesi F., Bagnato A., Groen A.F., 2003. Relationship between somatic cell count and functional longevity assessed using survival analysis in Italian Holstein-Friesian. Livest. Prod. Sci. 80: 211-230.

Schneider, M.del P., Miglior F., 1999. A proposal for genetic evaluation for functional herd life in Italian Holstein. Proc. 50th Annual meeting of the European Association for Animal Production (EAAP), Zurich, Switzerland p.11.

Schneider, M.del P., Canavesi, F., Samoré, A.B., 2000. Genetic evaluation for functional longevity in Italian Holsteins. Proc. 51th Annual meeting of the European Association for Animal Production (EAAP), Den Haag, The Netherlands. p.34.

Van der Linde C., De Jong G. and A. Harbers, 2004. Using a piecewise Weibull mixed model in teh genetic evaluation for longevity. Interbull bulletin n. 32: 157-162.

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

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.

Vukasinovic, N., Moll, J., Künzi, N., 1997. Analysis of productive life in Swiss Brown Cattle. J. Dairy Sci. 80, 2572-2579.

Vukasinovic, N., Moll, J., Casanova, L., 2001. Implementation of a routine genetic evaluation for longevity based on survival analysis techniques in dairy cattle population in Switzerland. J. Dairy Sci. 84, 2073-2080.