J. Dairy Sci. 96:8002–8013
http://dx.doi.org/10.3168/jds.2013-6830

Comparison between sire-maternal grandsire and animal models for genetic evaluation of longevity in a dairy cattle population with small herds

J. Jenko,*[1] G. Gorjanc,† M. Kovač,† and V. Ducrocq‡
*Agricultural Institute of Slovenia, Hacquetova ulica 17, 1000 Ljubljana, Slovenia
†Department of Animal Science, Biotechnical Faculty, University of Ljubljana, Groblje 3, 1230 Domžale, Slovenia
‡INRA, UMR 1313 Génétique Animale et Biologie Intégrative, 78350 Jouy-en-Josas, France

ABSTRACT

Survival analysis techniques for sire-maternal grandsire (MGS) and animal models were used to test the genetic evaluation of longevity in a Slovenian Brown cattle population characterized by small herds. Three genetic models were compared: a sire-MGS model for bulls and an approximate animal model based on estimated breeding values (EBV) from the sire-MGS model for cows, an animal model, and an animal model based on the estimated variance components from the sire-MGS model. In addition, modeling the contemporary group effect was defined as either a herd or a herd-year (HY) effect. With various restrictions on the minimum HY group size (from 1 to 10 cows per HY), changes in estimates of variance components, and consequently also in EBV, were observed for the sire-MGS and animal models. Variance of contemporary group effects decreased when an HY effect was fitted instead of a herd effect. In the case of a sire-MGS model, estimates of additive genetic variance were mostly robust to changes in minimum HY group size or fitting herd or HY effect, whereas they increased in the animal model when HY instead of herd effects was fitted, possibly revealing some confounding between cow EBV and contemporary group effect. Estimated heritabilities from sire-MGS models were between 0.091 and 0.119 and were mainly influenced by the restriction on the HY group size. Estimated heritabilities from animal models were higher: between 0.125 and 0.160 when herd effect was fitted and between 0.171 and 0.210 when HY effect was fitted. Rank correlations between the animal model and the approximate animal model based on EBV from the sire-MGS model were high: 0.94 for cows and 0.93 for sires when a herd effect was fitted and 0.90 for cows and 0.93 for sires when an HY effect was fitted. Validation performed on the independent validation data set revealed that the correlation between sire EBV and daughter survival were slightly higher with the approximate animal model based on EBV from the sire-MGS model compared with the animal model. The correlations between the sire EBV and daughter survival were higher when the model included an HY effect instead of a herd effect. To avoid confounding and reduce computational requirements, it is suggested that the approximate animal model based on EBV from the sire-MGS model and HY as a contemporary group effect is an interesting compromise for practical applications of genetic evaluation of longevity in cattle populations.

Key words: longevity, survival analysis, genetic evaluation, dairy cattle

INTRODUCTION

Longevity in dairy cattle is most commonly defined as the length of productive life (Ducrocq, 2005; Terawaki et al., 2006; Chirinos et al., 2007). Common approaches for the analysis of such data are survival analysis (Forabosco et al., 2009), due to a finer time scale (in days rather than in months, years, or lactations), the ability to treat living cows as censored, and the effective handling of time-dependent explanatory variables (Smith and Quaas, 1984; Ducrocq, 1987). Genetic evaluation of longevity based on survival analysis is commonly performed with sire or sire-maternal grandsire (**MGS**) models (Forabosco et al., 2009) as implemented in the Survival Kit software (Mészáros et al., 2013). This software implements a Bayesian approach to estimate parameters of frailty (i.e., mixed survival) models, assuming a conjugate log-gamma or nonconjugate normal prior distribution for the frailty (random) terms. Inferences on random hyperparameters are drawn from their marginal posterior density using Laplacian approximation (Ducrocq and Casella, 1996; Mészáros et al., 2013).

The main reason for using sire and sire-MGS versus animal models is computational. The joint maximization of a complex nonlinear function with many parameters such as the one used in survival mixed models is very demanding. Efficient optimization of such functions requires repeated computation and storage

---

Received March 19, 2013.
Accepted August 20, 2013.
[1] Corresponding author: janez.jenko@kis.si

of a Hessian matrix, which is usually denser than in standard linear mixed models. The use of AI in dairy cattle leads to a much smaller number of sires than the number of recorded animals, which makes sire and sire-MGS models the only computationally tractable ones in very large populations, in contrast to animal models.

To obtain EBV also for cows in the case of large populations, a 2-step approximate procedure was developed using results obtained from sire-MGS models (Ducrocq, 2001). This approximate animal model procedure is incorporated in the Survival Kit software (Mészáros et al., 2013). In the first step, a sire-MGS survival model is applied to obtain sire EBV. Then, in the second step, cows are evaluated assuming that other effects are known (Ducrocq, 2001).

Even if computationally possible, the application of an animal model for genetic evaluation of longevity in populations with small contemporary groups is questionable because of the lack of information for decisive separation of genetic and environmental effects. Namely, the amount of information per individual cow is limited in the presence of a single (potentially censored) record per cow for a lowly heritable trait. Therefore, modeling the contemporary group effect and its interplay with genetic effects is of considerable importance. Some authors distinguished variation in culling policy only between herds (Potočnik et al., 2011). However, within-herd culling policy also varies between years and between seasons within year, leading to the standard use of time-dependent herd-year (**HY**; e.g., Dürr et al., 1999; Terawaki et al., 2006) or herd-year-season (e.g., Ducrocq, 2005; Sewalem et al., 2005; Chirinos et al., 2007) interaction effects in survival analysis models. Modeling contemporary groups in a time-dependent manner within herd can better account for the conditions around the time of culling, but no clear guidelines exist to determine which approach is favorable in the case of limited amount of information per contemporary group.

In Slovenia, cattle populations are small with rather small herds, representing an example where sire-MGS and animal models can be tested and the interaction with contemporary group definition can be explored. The objective of this research was therefore to compare the application of different genetic models and different definitions of contemporary groups for genetic evaluation of longevity of Slovenian Brown cattle.

MATERIALS AND METHODS

Data

Data on longevity defined as time from the first calving until culling or the sixth calving of Brown cows between January 1998 and December 2008 in Slovenia were used. In the data preparation procedure, records from cows <20 mo and >48 mo at the time of first calving or with the missing date of the first calving were discarded. Lengths of productive life beyond lactation 5 were censored to concentrate the study on the risk of being culled early in life, because culling of older cows may involve reasons different from those in younger cows and early cullings are the most expensive. Censoring was also applied to cows changing herds during their lifespan, with censoring occurring at the date of last milk recording in their first herd. The final data set included 42,120 cows with phenotype data. A relatively small herd size (from 6.7 on average in 1999 to 8.7 Brown cows per herd in 2008) is an intrinsic feature of herds of the Brown cattle breed in Slovenia. Therefore, a series of 5 data sets were prepared for the comparison of estimated variance components under different restrictions on minimum number of cows per HY group (1, 2, 3, 5, or 10). A pedigree file with 5 generations was created for each data set separately. To test the predictive ability of the models, the data were divided into a training set for the estimation of variance components and the prediction of breeding values and a validation data set for testing the predictive ability of survival in cows. The training data set comprised 90% of all the cows with records (37,908 cows), whereas the remaining 10% (4,212 cows) were used for validation. Cows in the validation data set came from larger herds (≥10 cows per HY class) and were second-crop progeny of 63 sires.

Base Model

The analysis was based on a proportional hazards model with Weibull distributions:

$$t_i \sim \text{Weibull}\left(\lambda_\text{i},\, \rho_i\right), \tag{1}$$

where t_i is the longevity record of the ith cow, and λ_1 and ρ_i are scale and shape parameters, respectively, of the Weibull distributions that were modeled. Proportional hazards provide a flexible framework to describe hazard functions $h(t_i)$ separately for the general aging process—the so-called baseline hazard function, $h\left(t_i\right)_0$—and the multiplicative effect of a function of genetic and nongenetic factors $f(t_i)$:

$$h\left(t_i\right) = h\left(t_i\right)_0 \times f\left(t_i\right). \tag{2}$$

The baseline hazard function was modeled in a piecewise fashion, assuming a different Weibull hazard function for each combination of lactation number and stage of lactation, as in Terawaki et al. (2006). The