National genetic evaluation for somatic cell score

B.L. HARRIS, A.M. WINKELMAN and W.A. MONTGOMERIE

Livestock Improvement Corporation, Private Bag 3016, Hamilton, New Zealand

ABSTRACT

The method for the national genetic evaluation of somatic cell score (SCS) in New Zealand is described. Evaluation was done using a multiple-trait random regression test day animal model. Trait one was the first-lactation SCS record and trait two was the records from the second and third lactations. Third-order Legendre polynomials of days in milk were fitted to the genetic and cow effects. The evaluation was done across breed. In the first national evaluation, the pedigree contained 15.9 million animals. There were 22.7 million first lactation test day records and 35.4 million second/third lactation test day records. Solutions were obtained using a preconditioned conjugate gradient solver and iteration on data. The resulting breeding values (BVs) indicate a small annual increase (0.01 to 0.03 genetic standard deviation units) in SCS. The range in sire BVs indicate the potential to select for decreased SCS.

Keywords: somatic cell score; genetic evaluation; test day model; dairy cows.

INTRODUCTION

In New Zealand (NZ) herd testing is done for milk, fat, protein and somatic cell count (SCC). Routine genetic evaluation has been performed for milk, fat and protein yields since 1987. A system has been developed to use the SCC herd test data to calculate somatic cell score (SCS - log transformation of SCC) breeding values (BVs) for the dairy industry. The current evaluation of milk, fat and protein involves combining herd test results to predict a 270-day yield deviation, corrected for test day effects. These yield deviations are then used as the records for genetic evaluation. SCC is measured and reported as a concentration (1000s of cells per ml of milk). Hence, expressing the SCC as an accumulated yield deviation would be inappropriate. Many countries have implemented, or are moving towards, test day (TD) models for genetic evaluation of herd test data (Schaeffer *et al*., 2000; de Roos *et al*., 2003; Liu *et al*., 2003). The use of a TD model for genetic evaluation of SCS would not require the calculation of an accumulated or average yield as the record for evaluation. Additionally TD models have benefits of improved accuracy of evaluation by simultaneously accounting for the fixed effects and genetic, permanent environmental (PE) and temporary environmental (TE) random effects that affect the trait. Random regression (RR) TD models can accommodate changes in the genetic and PE effects over time by fitting polynomials of time to these effects. Recently, the changes over time have been modelled by fitting Legendre polynomials of days in milk (DIM) to the random effects in the model. Legendre polynomials are popular because they do not make prior assumptions about the shape of the lactation curve and the orthogonality of increasing orders of fit means that an increase in the order of the fit incorporates new information. Because of the benefits of using RR TD models, we have chosen to develop such a model for the genetic evaluation of SCS in NZ. This document describes the statistical model used for the calculation of BVs and the results of the RR TD model from the first national evaluations for SCS in NZ.

MATERIALS AND METHODS

Data

Herd test SCC data were extracted from the NZ national database. The extract was done in October 2004 and included all herd test SCC records from year 1992 and onwards. Data from lactations 1, 2 and 3 were used for BV estimation. However, for the genetic evaluation, records from lactations 2 and 3 were considered to be the same trait genetically (see below for explanation). A total of 22,714,465 first lactation and 35,378,989 second/third lactation test day records from 15,963,454 animals (cows and ancestors) were analysed. On average, cows had approximately 3.5 herd tests per lactation Contemporary group was defined in the model as herd-year-season-TD within each of lactations 1, 2 and 3. Season of calving was defined as spring or autumn. In NZ, the majority of cows (≥ 97%) calve in the spring. There was some variation in the numbers of cows tested within a herd-year-season. The average number of cows within a herd-year-season was 43.5 (range 1 to 869), 35.3 (range 1 to 632) and 23.4 (range 1 to 632) for lactations 1, 2 and 3, respectively. The breed compositions were very stable across lactation with approximately 56% Friesian, 17% Jersey, 22% Friesian-Jersey crosses and 1.5% Ayrshires with the remainder being other breeds and other crosses. The effect of induction of calving is included as a predictor in the animal model for genetic

evaluation. Less than 0.5% of first lactation cows and around 5% of second and third lactation cows were induced. The distribution of the SCC measurements was positively skewed. For this reason, the trait analysed was the log transformation (base 2) of SCC, referred to as somatic cell score (SCS). Lactation averages of SCS showed increasing overall levels with increasing lactation number. The average SCS was 5.87 and 6.14 with a standard deviation of 1.06 and 1.10 for first lactation and second/third lactation records, respectively.

Genetic parameter estimation

Initial investigations of models for the estimation of genetic parameters of SCS were done using herd test data obtained from first-lactation cows in progeny test herds (Winkelman, 2002). Subsequent work investigated genetic parameters for herd test SCS measured in the second and third lactations. A detailed description of this investigation is contained in Winkelman and Harris (2004). From the results of these studies, a model for the genetic evaluation of SCS in the NZ dairy population was chosen. The decision was made to use third-order Legendre polynomials to model heterogeneity of genetic and cow effects over time in the national evaluation. Attempts were made to use a multiple trait (MT) RR model to estimate variance components for SCC in lactations 1, 2 and 3. Using Legendre polynomials with orders of fit of 2 and 3 resulted in singularities in the information matrix. Genetic correlations estimated using a MT repeatability model were 0.95 between lactations 1 and 2 and 0.99 between lactations 2 and 3. The work suggested that second and third lactation SCS is genetically almost the same trait. Therefore, the model chosen to estimate variance components for the national evaluation of SCS was a MT RR model fitting third-order Legendre polynomials to the sire and cow effects in lactations 1 and 2/3.

Statistical model for national genetic evaluation

The model for the national genetic evaluation of SCS was a MT RR TD model where lactation 1 and lactations 2/3 are modelled as different traits. A TD SCS record, measured on day *d* (3 to 270) of lactation, was modelled as:

$$y_{ijklmno} = htd_{ij} + age_{ikl} + ind_{im} + \sum_{s=1}^{6} ht_{is} w_{ins}^{h} + \sum_{s=1}^{6} rc_{is} w_{ins}^{r} + \sum_{t=2}^{5} \phi_t(d) s_{it} + \sum_{t=1}^{3} \phi_t(d) a_{in} + \sum_{t=1}^{3} \phi_t(d) p_{in} + e_{ijklmno}$$

where,

$y_{ijklmno}$ is the TD record for SCS,

i denotes trait 1 or 2,

htd_{ij} is the j^{th} herd-year-season-TD fixed effect for trait *i*, with season referring to spring or autumn calving period,

age_{ikl} is the k^{th} month class for the age at calving fixed effect for trait *i* nested within the l^{th} breed class, with breed classes Holstein-Friesian (HF), Jersey, Ayrshire, Holstein-Friesian x Jersey crossbred, and other,

ind_{im} is the m^{th} induced lactation fixed effect class for trait *i* with induced classes representing induced or not induced calving,

ht_{is} is the s^{th} breed heterosis (*h*) with covariate value w_{ins}^{h} for animal *n*, for trait *i*,

rc_{is} is the s^{th} recombination loss (*r*) with covariate value w_{ins}^{r} for animal *n*, for trait *i*,

$\phi_t(d)$ is the standardised Legendre polynomial of order *t* calculated at days in milk *d*,

s_{it} is the stage of lactation fixed effect for days in milk *d* and order *t* for trait *i*,

a_{in} is the random additive genetic effect for animal *n* for trait *i*,

p_{in} is the random non-additive genetic and permanent environment effect for animal *n* for trait *i*, and

$e_{ijklmno}$ is the TE effect associated with record $y_{ijklmno}$

For the national SCS evaluation, genetic groups were assigned by breed, gender of missing parent, birth year and country of origin. Four breed classes were assigned genetic grouping, namely, HF, Jersey, Ayrshire-Red, and other breeds. Genetic groups were assigned in 5 year intervals from 1960 to 1980 then yearly, with the first birth year group being prior to 1960. Country of origin was defined as NZ, North American and Other. Gender of missing parent was defined as female or male. If a genetic group had less than 200 animals per group, birth years were clustered. No clustering occurred across breed, origin or gender genetic groups. To remove the effects of scaling, which is a phenomenon where contemporary groups with higher mean SCS have greater SCS phenotypic standard deviations, the method of Meuwissen *et al*. (1996) for joint estimation of BVs and heterogeneous variances was used. Essentially, the heterogeneous variances model is a multiplicative mixed model which scales the phenotypic record toward a common base phenotype variance. The mixed model equations were solved using a preconditioned conjugate gradient solver (Stranden & Lidauer, 1999) and iteration on data. Preconditioned conjugate gradient solver was used

because it is the preferred method for solving sparse systems of linear equations iteratively. BVs were calculated for days 3 to 270 within each lactation. Results for the BV averaged over lactations 1 and 2/3 are reported.

RESULTS AND DISCUSSION

The estimates for breed heterosis show that the first cross HFxJersey cow has a small advantage (reduced SCS) of 0.8% and 0.6% of the SCS mean compared with the average of the parental breeds for first lactation and second/third lactation records, respectively. The estimated average SCS lactation curves for lactations 1 and 2/3, obtained using the solutions for the Legendre polynomials, are shown in Figure 1. Because fitting the TD effect partially removed the effect of DIM, the curves differ slightly from those obtained from the phenotypic data.

FIGURE 1: Estimated average lactation curve for log_2 somatic cell count for lactation 1 and lactations 2/3.

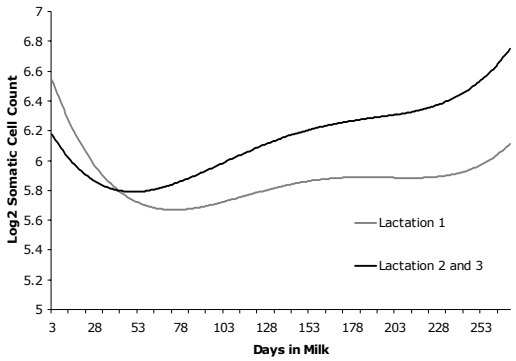
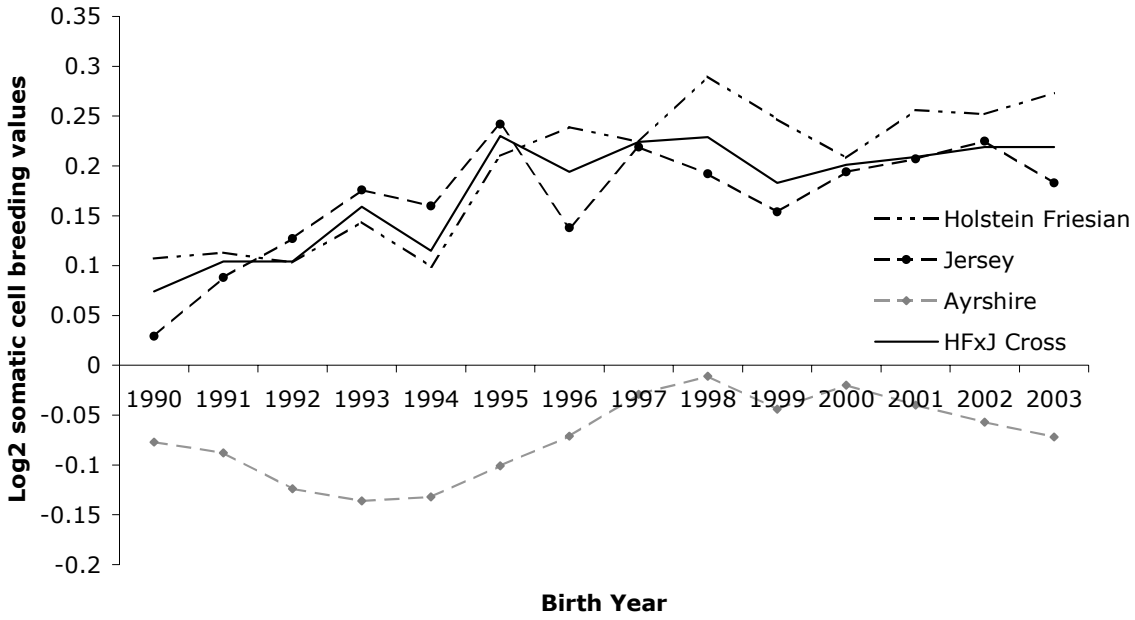


Table 1 provides the summary statistics for sire BVs within and across breed. The magnitude of the heritabilities of SCS (0.06 to 0.21) and the range in sire BVs indicate that selection for reduced SCS would be effective. The national genetic trend in the cow population for SCS by breed is given in Figure 2. The rates of genetic trend for SCS are 0.03, 0.02, 0.01 and 0.02 genetic standard deviation units per year for HF, Jersey, Ayrshire and HFxJersey crossbred cows, respectively. Compared with rate of genetic gain for milk production, which is close to 0.2 genetic standard deviation units per year, the values for SCS are low. The positive trend genetic trend in SCS shown in Figure 2 is undesirable and provides further justification for including SCC in the national economic selection index. The reported BVs were averages over each day of the lactation for lactations 1 and 2/3. However, if the need arises to increase selection pressure on SCS at specific stages of the lactation, the RR model can accommodate this by providing BVs for the part of the lactation of interest.

TABLE 1: Summary statistics for log_2 somatic cell count sire BVs.

	Number	Mean	SD	Minimum	Maximum
All breeds	15556	0.110	0.301	-0.900	1.732
Holstein-Friesian	7986	0.153	0.290	-0.852	1.732
Jersey	5919	0.077	0.297	-0.900	1.573
Ayrshire	1185	-0.051	0.287	-0.889	1.630
Other breeds	466	0.204	0.380	-0.751	1.705

FIGURE 2: Genetic trend for somatic cell count in the national cow population.



ACKNOWLEDGEMENTS

This work was funded by Dairy InSight. The authors thank Nicolas Gengler (Belgium) and George Wiggans (USA) for helpful comments.

REFERENCES

Liu, Z., Reinhardt, F., Bunger, J., Jaiter, J., Reents, R. (2003) Genetic evaluation of somatic cell scores using a random regression test day model for a very large dairy cattle population. *Interbull bulletin 31:* 88-91

Meuwissen T.H.E., de Jong, G., Engel, B. (1996) Joint estimation of breeding values and heterogeneous variances of large data files. *Journal of dairy science 79:* 310-316

de Roos, A.P.W, Harbers, A.G.F. and de Jong, G. (2003) Genetic parameters of testday somatic cell score estimated with a random regression model. *Interbull bulletin 31:* 97-101

Schaeffer, L.R., Jamrozik, J., Kistemaker, G.J. and Van Doormaal, B.J. (2000) Experience with a testday model. *Journal of dairy science 83:* 1135-1144

Stranden, I. and Lidauer, M. (1999) Solving large mixed linear models using preconditioned conjugate gradient iteration. *Journal of dairy science 82:* 2779-2787

Winkelman, A.M. (2002) Random regression models for genetic evaluation of somatic cell score in New Zealand dairy sires. *Proceedings of the 7th World Conference on Genetics Applied to Livestock Production*: CDROM Communication 09-27

Winkelman, A.M. and Harris B.L. (2004) Somatic cell score testday model for national genetic evaluation. *Livestock improvement mimeograph* (www.aeu.org.nz)