

EMPIRICAL BIAS IN THE PEDIGREE INDICES OF HEIFERS EVALUATED USING TEST DAY MODELS

J. Pösö, E.A. Mäntysaari, M. Lidauer, I. Strandén and A. Kettunen

Agricultural Research Centre, FIN-31600 Jokioinen, Finland

SUMMARY

Bias was estimated in test day and lactation model evaluations for first lactation milk yield from a data set of 675,283 cows. Empirical bias was defined as an average difference between estimated breeding values and pedigree indices. Statistically significant, but, in practice, negligible bias was found for daughters of both proven and young sires, and for daughters of bull dam candidates. The amount of bias in test day evaluations was unaffected by the comparison group definition used. **Keywords:** Ayrshire, breeding value estimation, milk production

INTRODUCTION

Test day model approach, where each test day measurement is considered as an observation, gives a possibility to better describe the seasonal variation both over and within herds than the traditional lactation model. Test day model can account for short-term seasonal effects associated with actual time of production, whereas in the lactation model these have to be fitted through factors associated with the time of calving. However, small herd sizes will lead to even smaller comparison groups in test day models, which may induce bias in genetic evaluations. The objective of the study was to estimate the effect of two different contemporary comparison group definitions to the magnitude of bias in test day model evaluations and to compare the bias to that found when the lactation model is used.

MATERIALS AND METHODS

Data consisted of 6,741,565 first lactation test day records on milk yield from 675,283 cows that calved between January 1988 and October 1996. The cows represented the three national dairy breeds (Ayrshire, Holstein-Friesian and Finncattle) in Finland and were from 25,114 milk recorded herds. With the pedigree file the data contained information from 7797 bulls, 1,091,816 cows and 193 phantom parent groups. Pedigree indices (PI) were calculated from a reduced data set formed by excluding records of cows that calved between April 1995 and March 1996. For these 82,253 cows PIs were formed as an average of their parents' estimated breeding values (EBV). A simple single trait animal model was assumed:

$$y_{ijklmnopq} = age_i + dcc_j + year*month_k + herd_l + \sum b_{r(m)} X_r + CCG_n + pe_{o(p)} + a_p + e_{ijklmnopq}$$

where $y_{ijklmnopq}$ is a test day observation on milk yield, age_i is the fixed effect of calving age class, dcc_j is the fixed effect of days carried calf class, $year*month_k$ is the fixed effect of test year-month, $herd_l$ is the fixed effect of herd of production, $b_{r(m)}$ are the regression coefficients on the

functions of DIM that describe the shape of lactation curves within calving month classes, ($X_1=1$, $X_2=DIM/c$, $X_3=(DIM/c)^2$, $X_4=\ln(c/DIM)$, $X_5=(\ln(c/DIM))^2$, DIM is the days in milk for a particular test day, $c=305$), CCG_n is the effect of contemporary comparison group (either herd-year (HY) or herd-test-month (HTM)), $pe_{o(p)}$ is the permanent environmental effect of a cow p , a_p is the additive genetic effect, and $e_{ijklmnopq}$ is the residual. The EBVs and PIs from test day models were transformed to 305 d scale.

First lactation 305 d milk yield records were found for 666,196 cows. The following linear animal model was assumed for estimation of EBVs:

$$y_{ijklmn} = \text{month*year}_i + \text{age*DO}_j + \text{herd}_k + CCG_l + a_m + e_{ijklmn}$$

where y_{ijklmn} is 305 d milk yield, month*year_i is the fixed effect of calving month*calving year, age*DO_j is the fixed effect of calving age*days open interaction, herd_k is the fixed effect of calving herd, CCG_l is the effect of contemporary comparison group (herd*calving year), a_m is the additive genetic effect, and e_{ijklmn} is the residual. The variance components used in the breeding value estimations were obtained from a subset of 6,310 first lactation Ayrshire cows that were from 78 relatively large herds using the models described above (Table 1).

The EBVs and PIs of Ayrshire cows were selected for closer inspection. The cows were grouped either as 1) second crop daughters of proven sires (more than 100 daughters/sire in the reduced data set), 2) daughters of young unproven sires (sires with less than 10 daughters in the reduced data set but more than 100 daughters in the full data set), or 3) daughters of bull dam candidates (dams' indices for test day milk yield two standard deviations above the Ayrshire population mean). The empirical bias was defined as an average difference between the EBVs and PIs. The significance of possible bias was tested with t-test and by fitting linear regression of EBV on PI (Reverter *et al.* 1994).

Table 1. Variance components used in estimation of breeding values and pedigree indices and their functions (heritability within comparison group (h^2) and repeatability (r))

	comparison group	permanent environment	additive genetic	residual	h^2	r
HTM ^A	1.39	4.55	3.15	4.87	0.25	0.61
HY ^B	0.76	3.91	3.25	6.12	0.24	0.54
305d ^C	79,834	-	265,841	431,102	0.38	-

^A Test day model with herd-test-month as contemporary comparison group

^B Test day model with herd-year as contemporary comparison group

^C Lactation yield records

RESULTS AND DISCUSSION

With both test day models statistically significant empirical bias was found for all three groups of cows (Table 2). Daughters of young bulls had the largest bias, -53 kg (HTM-model), which in practice, however, corresponds to only 0.23 units of standard deviation of PIs. The estimated bias was negative for all categories, i.e., pedigree indices were larger than the estimated breeding values after first lactation. Having HTM or HY in the model as CCG gave similar results, although the HTM effect did have a tendency of giving slightly larger bias than the HY effect.

For lactation yield records, only PIs of daughters of proven sires were found to be significantly biased according to t-test. Again the bias was practically negligible (0.03 standard deviations). The regression coefficients were statistically different from the expected value of 1 for all three categories implying existing bias also for other groups of cows than daughters of proven sires. In contrast to this study, Lidauer and Mäntysaari (1996) found distinctly biased PIs for daughters of young sires for 305 d milk yield. However, the statistical model used was not exactly the same as the one used here and, perhaps more importantly, the PIs in their study were estimated using records from the first three lactations. Earlier, Mäntysaari and Sillanpää (1993) demonstrated a smaller bias for young bulls when the second and third lactation records were excluded from the evaluations.

Table 2. Means and standard deviations (SD^A) of differences between estimated breeding values (EBV) and pedigree indices (PI), standard deviations of PIs, and regression coefficients $b_{(EBV,PI)}$ for three groups of primiparous Ayrshire cows

	Daughters of proven sires			Daughters of young sires			Daughters of bull dam candidates		
	HTM ^B	HY ^C	305d ^D	HTM	HY	305d	HTM	HY	305d
Cows (sires)	37,160 (426)			15,712 (74)			1,933 (294)		
Mean _(EBV-PI)	-19.3 ^{***}	-18.7 ^{***}	6.9 ^{***}	-53.0 ^{***}	-44.5 ^{***}	4.1 ^{ns}	-28.1 ^{***}	-26.0 ^{***}	10.9 ^{ns}
SD _{PI}	267.7	268.8	251.6	227.1	229.1	211.0	194.8	196.3	189.3
SD _(EBV-PI)	228.1	236.2	209.4	223.2	230.5	203.7	255.8	266.0	233.4
$b_{(EBV,PI)}$	0.975 ^{***}	0.978 ^{***}	0.989 [*]	0.789 ^{***}	0.794 ^{***}	0.806 ^{***}	0.924 [*]	0.927 [*]	0.921 [*]

Significance levels: *** $P \leq 0.001$, * $P \leq 0.05$, ^{ns} not significant

^A Calculated within sires

^B Test day model with herd-test-month as contemporary comparison group

^C Test day model with herd-year as contemporary comparison group

^D Lactation yield records

Correlations between evaluations with different models were very high in all groups of cows. The correlation between HTM and HY was 0.99 and even between test day models and lactation model 0.96 and 0.97, respectively.

The results suggest that the lactation model leads to a smaller bias than test day models. The results should, however, be viewed with caution. The lactation records used were compiled some 6 months later than the test day data were received. As a consequence some of the animals had a lactation record based on more test day observations than they had in the actual test day data, and the magnitude of bias was found to be dependent on the number of test day observations (Table 3). The average difference EBV-PI diminished when the number of test day observations increased. The smallest bias was found when cows had 10 test day observations whereas with more than 10 observations the bias turned positive and started to increase again. This pattern is likely to reflect the rigid and unrealistic assumption of uniform correlation structure among test day observations in this kind of simple repeatability model which ignores, e.g., individual lactation curves.

REFERENCES

- Lidauer, M. and Mäntysaari, E.A. (1996) *Agric. Food Sci. Finl.* 5:387-397.
 Mäntysaari, E.A. and Sillanpää, M.J. (1993) 44th Annual Meeting of EAAP, Aarhus. Mimeo.
 Reverter, A., Golden, B.L., Bourdon, R.M. and Brinks, J.S. (1994) *J. Anim. Sci.* 72:34-37.

Table 3. Mean differences between estimated breeding values and pedigree indices for milk yield of three groups of primiparous Ayrshire cows with different number of test day observations

		Number of test day observations				
		4	6	8	10	12
Daughters of proven bulls	HTM ^A	-223.7***	-103.6***	-63.7***	-12.7***	45.2***
	HY ^B	-229.5***	-105.9***	-63.4***	-11.9***	47.7***
Daughters of young bulls	HTM ^A	-235.0***	-143.9***	-99.5***	-33.0***	39.1***
	HY ^B	-230.3***	-138.4***	-90.9***	-24.2***	50.9***
Daughters of bull dam candidates	HTM ^A	-298.2***	-92.2*	-92.1***	1.9 ^{ns}	25.8 ^{ns}
	HY ^B	-304.9***	-92.5*	-89.9***	3.8 ^{ns}	29.0 ^{ns}

Significance levels: *** $P \leq 0.001$, ** $P \leq 0.01$, * $P \leq 0.05$, ^{ns} not significant

^A Test day model with herd-test-month as contemporary comparison group

^B Test day model with herd-year as contemporary comparison group