Best Use of Conventional EBV of Bull Dams and Combination with Direct Genomic Values

S. Rensing¹, E Pasman¹ and F. Reinhardt¹

Vereinigte Informationssysteme Tierhaltung w.V. (vit), Heideweg 1, 27283 Verden, Germany

Abstract

Over-estimation of bull dam EBV– at least for production traits – has been obvious in many A.I. breeding programs. Main reasons for setting up nucleus herds or bull dam test stations were to achieve unbiased EBV for these highly selected cows. Because of overestimated dam EBV the parent average of test bulls is on average higher than the proof based on 1^{st} crop daughters.

This study is based on 1003 black-and-white Holstein bulls born 2001/2002 with dam proofs based on own performance in the domestic German evaluation August 2003. The parent averages for protein yield in August 2003 were about one standard deviation higher than the August 2009 proofs based on more than 100 daughters and with minimum 90% reliability. The realized August 2009 EBV were predicted more accurately by the sire parent average defined as 0.5*EBV of sire + 0.25*EBV of maternal-grand-sire + 0.125*EBV of maternal-grand-sire. In contrast to production traits the parent average for somatic cell score was an unbiased predictor for realized EBV and superior compared to sire parent average.

Therefore sire parent average instead of full parent average is used for combination of direct genomic values with conventional evaluation information for animals without daughter information to achieve unbiased genomic enhanced EBV. If bull selection for use on the entire cow population is based on genomic enhanced breeding values, the impact of the biased bull dam information on bull selection and genetic progress is larger. Therefore unbiased pedigree information in genomic <u>enhanced</u> EBV for solely genomic proven bulls is of high importance.

1. Introduction

Overestimation of highly selected bull dams is a problem in many traits since the introduction of animal models for genetic evaluation. The general assumption, that all animals within the same class of fixed effects (e.g. herd-yearseason) are treated equally, does not hold for (potential) bull dams. Because of the economic impact of cows meeting the high standards for bull dams, breeders pay more attention to those cows compared to the herd mates (preferential treatment). Since Holstein breeding programs have become international with similar goals including comparable traits the international focus within the leading Holstein populations is on the same cow families and extreme high figures.

Even though the resulting bias in EBV of highly selected cows is known, there are only few published studies on this effect. The reasons may be that there is no simple solution, all main breeding programs are affected and any result would not be positive for the image of the genetic evaluation and the breeding program. Nevertheless Holstein breeding programs take this matter seriously, as the establishment of bull dam test herds/stations or nucleus herds shows.

With the introduction of genomic selection the effect of overestimated bull dam proofs becomes a new dimension. In classical A.I. programs the main effect of overestimated bull dam proofs and with that of overestimated parent averages (**PA**) is a lower selection intensity for test bull candidates than expected. Proven bulls to be used on the entire cow population are selected based on daughter proofs. Even in animal models the influence of biased dam and grand dam proofs on published bull EBV based on daughter proofs becomes negligible. If bull selection for use on the entire cow population is based on genomic enhanced breeding values (**gEBV**), the impact of the parent information is substantial as long as daughter data is not available. Therefore knowledge on presence and extent of overestimated EBV of bull dams and detection of better alternatives has become more important.

2. Materials and Methods

2.1. Data materials

To follow up selection and breeding decisions in German Holstein A.I. programs since 2003 yearly data sets on all registered test bulls including EBV on two generation ancestors (August evaluation in that year) and performance data of dam and grand dam are stored. So the data set from 2003 gives a snapshot of test bulls born 2001 and 2002 at the moment breeding organizations started testing these bulls. Meanwhile these bulls have EBV based on 1st crop daughters. This study includes 1,003 black-and-white Holstein test bulls born 2001/2002 with dams having performance data in the German national genetic evaluation of August 2003. 483 of the 1,486 test bulls born in these years were excluded because they have foreign dams. The excluded 483 and the included 1,003 bulls had on average very similar sire and maternal grand sire EBV.

Table 1. Average number of daughters, EBV and reliabilities August 2009 for 1003 black-and-white Holstein A.I. bulls born 2001/2002.

HOL bulls	average	S
n daughters	131.0	82.4
rel.% production	93.5	2.0
EBV protein kg	24.3	17.1
rel.% SCS	87.4	3.6
EBV RZS*	100.9	11.7

*) RZS= relative breeding value SCS (\emptyset 100; sg 12, low cells high value)

Table 1 shows that official proofs of August2009 from the Random-Regression-Test-Day-

Animal-Model (www.vit.de, 2009) include over 100 daughters with performance data for production traits and SCS in first and second lactation resulting in high reliabilities of 93.5% respectively 87.4%. The high standard deviation of number of daughters is caused by a small number of bulls born 2001 having already many 2nd crop daughters. The standard deviation of EBV for protein kg with 17.1kg represents about one genetic standard deviation. The same is true for SCS that is displayed as relative breeding value RZS with average 100 and a genetic standard deviation of 12 (high figures = low SCS).

2.2. Statistical model

Data were analysed with one-way analysis of variance using SAS. The contribution of full parent average (PA) and the sire parent average (sPA) to realized EBV of August 2009 were evaluated with the following models:

PA-model: $EBV_{.09} = b_1 * EBVsire_{.03} + b_2 * EBVdam_{.03} + e$ sPA-model: $EBV_{.09} = b_1 * EBVsire_{.03} + b_2 * EBVmgs_{.03} + b_3 * EBVmggs_{.03} + e$

3. Results and Discussion

In table 2 comparisons of average realized EBV August 2009 with August 2003 parent averages are shown. The August 2003 figures were corrected for the base change in 2005 (all figures expressed on "cows born 2000"). The Holstein test bulls born 2001/2002 were on average hardly selected for SCS. For production traits represented by kg protein the test bulls have a PA about 2.5 standard deviations above the current base and about three standard deviations above the August 2003 base (cows born 1995) respectively. For sires and dams almost the same selection intensity was applied (sire – dam EBV = 4 kg protein).

1,003 HOL bulls	EBV RZS*	EBV kg			
		protein			
08-2009	100.9	24.3			
PA 08-2003	100.9	40.3			
Diff. PA - EBV	0.0	-16.0			
Sire PA 08-2003	100.9	26.0			
Diff. sPA - EBV	0.0	-1.7			

Table 2. Average daughter based EBV 08-2009 compared with full PA and sPA 08-2003 (all figures on current base "cows born 2000").

*) RZS= relative breeding value SCS (\emptyset 100; sg 12, low cells high value)

Realized daughter based proofs show that PA for somatic cell score is an unbiased predictor. For kg protein this was not observed. Average realized EBV are with 24.3 kg almost one standard deviation lower than predicted by PA. So a substantial overestimation of PA is present.

Furthermore table 2 shows that the sire PA is a more realistic predictor of average realized EBV compared to full PA.

Comparison of predictability of average PA and sPA indicates already that dams and grand dams are sources of overestimation of PA. Table 3 give the result of the variance analyses for SCS and protein kg with the PA model.

Table 3. Regression of 08-2003 sire and damEBV on realized EBV of 08-2009 (Model PA).

	RZS*		kg P				
\mathbb{R}^2	0.275		0.218				
Effects	sig. (p)	regr.	sig. (p)	regr.			
Sire	< 0.000	0.493	< 0.000	0.473			
2003	1		1				
Dam	< 0.000	0.519	< 0.000	0.177			
2003	1		1				

*) RZS= relative breeding value SCS (\emptyset 100; sg 12, low cells high value)

The correlation of 2003 PA with 2009 EBV is 0.52 (R^2 0.275) for SCS and 0.47 (R^2 0.218) for protein kg. With that the correlation for SCS is only slightly lower than the expected correlation from computed PA reliability of about 30%. For protein the correlation is remarkably lower than the correlation expected from computed PA reliability of about 0.33. The low correlation for protein yield is due to the dam EBV. According to expectation

differences in both sire and dam EBV contribute by 0.5 to the realized EBV. Table 3 shows that this is nearly fulfilled for SCS and protein sire EBV. Realized EBV reflect the differences in dam EBV only for SCS but not for protein. The regression factor of dam proof on realized EBV (under 0.2) is less than half of the expected 0.5.

Taking into account maternal grand sire (MGS) and great grand sire (MGGS) instead of the dam (model sPA) should lead to slightly reduced overall R^2 and a regression of 0.25 for MGS and 0.125 for MGGS.

Table 4. Regression of 08-2003 sire, MGS and MGGS EBV on realized 08-2009 EBV (Model sPA).

5111).				
	RZS*		kg P	
\mathbb{R}^2	0.240		0.219	
Effects	sig. (p)	regr.	sig. (p)	regr.
Sire	< 0.000	0.504	< 0.000	0.468
2003	1		1	
MGS	< 0.000	0.265	0.0004	0.119
2003	1		0.0004	
MGGS	0.0004	0.116	0.0001	0.121
2003	0.0004		0.0001	

*) RZS= relative breeding value SCS (\emptyset 100; sg 12, low cells high value)

Table 4 indicates that overall R^2 of model sPA is as expected slightly lower for SCS (0.24 compared to 0.275 with model PA). Regressions factors of sire, MGS and MGGS were in the expected range. For protein-kg the overall R^2 of the sPA-model is not lower than that of PA-model, which is surprising. Still the explained amount of variation in the resulting proof of protein-kg is still considerably lower than expected. Regressions factors of sire and MGGS proofs on the later bull proofs for protein-kg were about as expected. The low regression factor for MGS (0.12) means, that the predictive value of the MGS-proof is much lower than expected. Most of the German Holstein test bulls born 2001/2002 have foreign sires as MGS that were mainly used in Germany to produce offspring for breeding. The proofs of daughters of these MGS may be biased in a similar way as the proofs of bull dams. The proofs of the MGS of the bulls under study may also be biased because they only have such daughters.

4. Conclusions

Full PA for production traits of 1,003 blackand-white German Holstein test bulls born 2001/2002 were overestimated by about one genetic standard deviation. Variance analyses indicate that almost solely the dam is the source of this overestimation. The correlation of sPA to realized EBV is about the same as full PA, but predicts the absolute level of realized EBV much more realistic.

Therefore in the German genomic evaluation system sPA is used as conventional genetic information for combination with direct genomic breeding values to genomic enhanced breeding values (gEBV) if conventional EBV is based on less than 10 EDC; i.e. young candidates and cows (Reinhardt et al., 2009). Because of the clear bias of at least production EBV of highly selected cows it is doubtful if including such genotyped cows in training sets for genomic formulas delivers the expected gain in accuracy.

For SCS full PA is the better predictor. Nevertheless overestimation for full PA can be expected for at least type traits, too. In Germany only very few dams of test bulls have EBV including own performance for type traits. Conventional genetic evaluation for type traits is based only on randomly sampled test daughters with herd mates from all herds with official milk recording. For all traits with low heritability (e.g. daughter fertility) full PA and sPA are expected to be very similar because of the limited impact of own performance on cow EBV. Using sPA for gEBV of candidates has the advantage of same conditions for all traits and animals regardless the source of the dam (domestic/foreign) and the used conventional evaluation model (animal model/sire model).

Biased bull dam EBV have highly negative impact on the true accuracy if used in gEBV. Unbiased gEBV are crucial for selection in the entire cow population if young bulls with gEBV are widely used. The impact of biased PA on gEBV of a young bull will first decrease when daughter information is considered. It remains almost constant for up to four years with the given limited reliability of the gEBV. Therefore analyses of extent of overestimated bull dam EBV for other traits (e.g. conformation) are needed. As long as daughter information is not available using sPA instead of full PA currently seems to be the best method to estimate unbiased gEBV.

5. References

- Reinhardt, F., Liu, Z., Seefried, F. & Thaller, G. 2009. Description of German genetic evaluation system for dairy breeds. <u>www.vit.de</u> 2009.
- Reinhardt, F., Liu, Z., Seefried, F. & Thaller, G. 2009. Implementation of genomic evaluation in German Holsteins. *Paper* presented at Interbull meeting August 21st-23rd, Barcelona/Spain.