Accuracy Of Genomic Breeding Values For Non-Production Traits And Its Relationship With The Accuracy Of The 'Phenotype' Used

M. Haile-Mariam[*], P.J. Bowman[*], A.C. Chamberlain[*], M.E. Goddard[‡,*],
H.W. Raadsma[†] and B.J. Hayes[*]

Introduction

Genomic selection can accelerate genetic change for economic traits ((Meuwissen, et al. (2001)). The increase in genetic gain which can be achieved largely depends on the accuracy of direct genomic values (DGV) of selection candidates. For Australian dairy cattle, accuracy of DGVs for production traits and the economic index range from 0.3 to 0.8 (e.g. Hayes et al. (2009a,b); Moser et al. (2009)). However, accuracy of DGV for other traits have not been previously reported. Our aim was to evaluate the accuracy of DGVs for non-production traits for Holstein bulls, as well as investigating ways of enhancing the accuracy of DGVs of traits such as type and fertility which are usually recorded on limited number of progeny per bull. For example, a decision can be made to use either the most reliably evaluated bulls for a trait or all the available bulls, and this will affect the number of animals to be included in the reference set. This is important because the accuracy of DGVs is influenced both by the accuracy of the phenotype record used and by the number of animals in the reference set (Goddard (2008).

Materials and Methods

About 2,500 Holstein bulls were genotyped for the Illumina Bovine50K array, which includes 54,001 SNP markers. Details on quality control of the SNP data are given by Hayes et al. (2009a,b). Briefly, criteria for selecting SNPs were; less than 5% pedigree discordants, 90% call rate, MAF>1%, Hardy Weinberg P<0.00001. After these checks a small number of SNP that could not be mapped were omitted from the final data set. Parentage checking was then performed again, and any genotypes incompatible with pedigree were set to missing. The genotypes were ordered by chromosome position and were then submitted to fastPHASE (Scheet and Stephens, (2006)). The genotypes were taken as those filled in by fastPHASE. Finally 43,115 SNPs that met all of these criteria were used for analyses.

The phenotypes used were daughter 'yield' deviations (DYD) for type and workability traits, calving ease (CE), survival and daughter fertility. For each bull, DYDs were calculated as average performance of daughters' adjusted for fixed and non-genetic random effects of the daughters and genetic effects of his mates.

---

[*] Biosciences Res. Division, DPI, 1 Park Drive, Bundoora, Vic 3083 Australia
[†] ReproGen, Fac. Vet. Sci. Univ. Sydney, 425 Werombi Rd. Camden, NSW 2570, Australia
[‡] Dept. of Agriculture and Food Systems, The University of Melbourne, Vic 3010 Australia

To test if the number of daughters (i.e. reliability of the 'phenotype') per bull affects the accuracy of DGV, bulls with ≥1, ≥10, ≥20 or ≥30 progeny were used as reference set. SNP-BLUP software used for estimating SNP-effects is similar to GBLUP of Meuwissen et al. (2001) and is described by Konstantinov and Hayes (2010) and Moser et al. (2009). In SNPBLUP solutions for bulls with genotype are obtained from a generalised ridge regression equation using the following model $\mathbf{y} = \mathbf{1}_n \mu + X\mathbf{g} + \mathbf{e}$, where y is a vector of DYDs, μ is the mean, 1_n is a vector of 1s, X is an N_{bull} x N_{SNP} matrix of genotypes coded as 0, 1 or 2, g is a vector of breeding values and e is a vector of random normal deviates ~ $N(0, \sigma_e^2)$. The (co)variance matrix of the SNP effects is assumed to be $I\lambda$. The system of equations was solved iteratively by Preconditioned Conjugate Gradient Method (PCG). A 10-fold cross-validation procedure with golden segment search using the training data set was performed to locate the best λ within a given range. In addition similar analyses as above were performed using ASReml (Gilmour et al. (2002)) to obtain bull solutions adjusted for number of progeny and their standard error (s.e.). SNP-effects from the SNP-BLUP were used to derive DGVs for bulls in the validation set. The DGVs and bull solutions from ASReml were correlated with EBVs to obtain realised accuracy and s.e. were used to get the expected accuracy as described by Hayes et al (2009b). Expected accuracy was calculated for selected traits and were compared with the realised accuracy of DGV. Bulls born after 2002 with ≥30 progeny were used as a validation set.

Table 1: No. and mean reliability of EBVs of bulls in the reference and validation set.

Traits	Validation		Reference							
	No.	Rel.	≥1 progeny		≥10 progeny		≥20 progeny		≥30 progeny	
			No.	Rel.	No.	Rel.	No.	Rel.	No.	Rel.
Type traits	63	75	1680	66(12)	1354	73(39)	1098	77(57)	800	81(64)
Workability	296	78	1877	80(31)	1861	81(40)	1835	81(58)	1768	82(63)
Calving E.	70	45	1718	37(10)	1146	38(10)	647	44(12)	350	53(20)
Fertility	242	50	1820	61(10)	1745	63(10)	1731	63(11)	1700	64(17)
Survival	200	51	1845	70(20)	1842	70(20)	1838	70(20)	1829	70(41)

[†] Values in parenthesise are the minimum reliability.

Results and discussion

The number of bulls in the validation set with their average reliability of the EBVS is shown in Table 1. Table 1 also shows the number of reference bulls when the number of progeny per bull increased from ≥1 to ≥30. Table 2 shows h^2 and realised accuracy of DGVs when SNP-effects calculated based on different set of reference bulls are used. Increasing the number of progeny per bull from ≥1 to ≥10 increased the accuracy of the DGV from an average of 0.4 to 0.51 for type traits (Table 2) despite the decrease in the number of bull in the reference population by about 330 to 380 (Table 1). Similarly the accuracy of DGV of bulls for CE increased by 0.03 (Table 2) when the number of progeny required per bull increased from ≥1 to ≥10. Although this increase is marginal, it is observed despite the decrease in the number of bulls by 572 (Table 1). For other traits size of progeny group did not affect accuracy of DGVs because most bulls were evaluated on reasonably large number of daughters. When the number of progeny per bull increased from ≥10 to ≥20 or ≥ 30 and the size of the reference set decreased, the accuracy of DGVs of the bulls in the validation set fluctuated with no clear trend (Table 2). Table 2 also shows that correlation between DGVs and DYD are almost always lower than the correlation of DGVs with EBVs.

Table 2: Heritability and realised accuracy of DGV for bulls in the validation set when different set of bulls with different number of daughters are used as reference set.

Traits	h^2	≥1 progeny	≥10 progeny	≥20 progeny	≥30 progeny
Body Depth	0.32	0.36(0.36)	0.58(0.55)	0.55(0.51)	0.60(0.54)
Foot Angle	0.13	0.31(0.16)	0.50(0.37)	0.54(0.40)	0.55(0.41)
Udder Depth	0.33	0.38(0.34)	0.41(0.39)	0.42(0.39)	0.42(0.39)
Teat Length	0.35	0.42(0.40)	0.53(0.47)	0.56(0.49)	0.49(0.42)
Mam. System	0.20	0.68(0.60)	0.70(0.62)	0.68(0.63)	0.61(0.58)
Stature	0.35	0.38(0.32)	0.50(0.45)	0.48(0.42)	0.55(0.49)
Udder Texture	0.16	0.56(0.52)	0.56(0.51)	0.61(0.57)	0.67(0.62)
Bone quality	0.23	0.64(0.58)	0.62(0.56)	0.67(0.61)	0.68(0.63)
Angularity	0.19	0.17(0.18)	0.42(0.39)	0.46(0.42)	0.56(0.51)
Chest Width	0.20	0.38(0.35)	0.44(0.39)	0.42(0.37)	0.36(0.31)
Pin Width	0.29	0.35(0.34)	0.62(0.57)	0.63(0.57)	0.66(0.61)
Pin Set	0.32	0.33(0.30)	0.49(0.44)	0.45(0.39)	0.50(0.43)
Rear set	0.15	0.21(0.18)	0.33(0.23)	0.43(0.30)	0.43(0.28)
Fore Attachment	0.17	0.38(0.31)	0.38(0.32)	0.32(0.26)	0.27(0.22)
Rear A. Height	0.21	0.39(0.39)	0.43(0.42)	0.42(0.39)	0.34(0.32)
Rear A. Width	0.18	0.27(0.26)	0.42(0.37)	0.50(0.45)	0.43(0.39)
Central Ligament	0.17	0.50(0.46)	0.52(0.48)	0.45(0.42)	0.45(0.41)
Teat Placement	0.29	0.50(0.33)	0.54(0.47)	0.54(0.47)	0.52(0.45)
Average	0.23	0.40	0.51	0.51	0.51
Milking speed	0.19	0.41(0.31)	0.40(0.31)	0.38(0.31)	0.37(0.30)
Temperament	0.16	0.43(0.32)	0.42(0.31)	0.43(0.31)	0.43(0.31)
Likeability	0.19	0.33(0.26)	0.32(0.26)	0.30(0.26)	0.31(0.26)
Average	0.18	0.39	0.38	0.37	0.37
Calving ease	0.05	0.36(0.20)	0.39(0.18)	0.37(0.23)	0.40(0.27)
Fertility	0.04	0.35(0.21)	0.31(0.21)	0.32(0.22)	0.33(0.22)
Survival direct	0.03	0.38(0.13)	0.42(0.18)	0.42(0.18)	0.43(0.19)
Average	0.04	0.36	0.38	0.37	0.39

Correlations in parenthesis are between DGV and DYD.

Studies that assess the accuracy of DGVs usually use bulls that have reliable EBVs as reference set (Berry et al. (2009); Gredler, et al. (2009); Hayes et al. (2009a,b); VanRaden et al. (2009)). For example, reference bulls used by VanRaden et al. (2009) were required to have a reliability of at least 75%, and bulls in the validation set were required to have information from ≥10 daughters in their evaluations. In this study including bulls with less than 10 progeny had decreased the accuracy of DGVs for some traits. However, both expected and realised accuracy in Table 3 shows data of all bulls can be used without affecting accuracy of DGVs if the number of progeny is accounted for in the model. In agreement to other studies (Hayes et al. (2009b)) Table 3 shows that expected accuracy are marginal higher than realised accuracy for all traits except CE.

In most cases for a given trait the h^2 and the accuracy of DGV are related as expected (Hayes et al. (2009a)). However, the accuracy of DGVs for stature which is the most heritable trait is similar to that of udder texture (Table 2). The reasons for such discrepancy could be related with the fact that traits like stature, body depth and chest width as predictors of liveweight could be under some form of selection (VanRaden et al. (2009)). Recently a number of studies have reported accuracy of DGVs for several dairy traits (Berry et al. (2009); Gredler, et al. (2009); Hayes et al. (2009,a,b); VanRaden et al. (2009)). Although it is difficult to compare reliabilities from different studies due to difference in methods used to calculate reliability and sizes of reference data, the accuracy estimated in the current study are within the range of those reported in the literature.

Putting restriction on the number of progeny per bull (e.g. ≥10) for some traits increased accuracy of DGV compared to using all the data. Further analyses showed that data of bulls even with one progeny can be included in the reference set without affecting accuracy of DGVs if the number of progeny per bull is considered in the model. Proper weighing of data could also enable the use of cow phenotype and genotype data in the reference set for estimating SNP effects. Possible advantages or disadvantages of using bulls with more (e.g ≥ 10) or less (e.g. ≥1) reliable 'phenotype' on bias and other related statistics should be investigate before making decisions.

Table 3: Expected accuracy of bull's DGVs in the validation set when bulls with different number progeny are included in the reference set.

Traits	≥1 progeny	≥10 progeny	≥20 progeny	≥30 progeny
Udder Depth	0.43(0.55)[0.57]	0.42(0.45)[0.53]	0.42(0.44)[0.50]	0.41(0.42)[0.48]
Mam System	0.44(0.55)[0.70]	0.50(0.54)[0.71]	0.53(0.54)[0.69]	0.54(0.53)[0.63]
Stature	0.47(0.57)[0.51)	0.54(0.56)[0.52]	0.54(0.56)[0.51]	0.53(0.54)[0.56]
Angularity	0.39(0.54)[0.51]	0.43(0.52)[0.55]	0.44(0.51)[0.54]	0.46(0.50)[0.57]
Calving ease	0.23(0.36)[0.42]	0.32(0.37)[0.41]	0.30(0.35)[0.40]	0.31(0.33)[0.43]
Fertility	0.51(0.53)[0.55]	0.43(0.53)[0.55]	0.45(0.53)[0.54]	0.43(0.53)[0.55]
Likeability	0.52(0.59)[0.31]	0.54(0.59)[0.31]	0.52(0.58)[0.31]	0.51(0.58)[0.31]

Expected accuracy in parenthesis of DGVs adjusted for number of progeny and values in square bracket are realised accuracy when DYD are adjusted for number of progeny.

References

Berry, D.P., Kearney, F., and Harris, B.F. (2009). Interbull Bull. 39:29-34. Uppsala, Sweden.
Goddard, M.E. (2008). *Genetica* 136:245.
Gredler, B., Nirea, K.G, Solberg, T.R. *et al.* (2009). In *Proc. 18th AAABG 568-571.*
Hayes, B.J., Bowman, P.J., Chamberlain, A.C. *et al.* (2009a). *J. Dairy Sci.* 92:1313.
Hayes B.J., Bowman, P.J., Chamberlain, A.C. *et al.* (2009b). *Genet. Sel. Evol.,* 41:51.
Konstantinov, K.V., and Hayes, B.J. (2010). In this proceeding.
Meuwissen, T. H. E., Hayes, B. J. and Goddard, M.E. (2001). *Genetics,* 157:1819–1829.
Scheet, P., and M. A. Stephens. (2006). *Am. J. Hum. Genet*. 78:629–644.
Van Raden, P. M., Van Tassell, C. P., Wiggans, G. R., *et al.* (2009). *J. Dairy Sci.,* 92:16–24.
Moser, G., Tier B., Crump, R. E., *et al.* (2009). *Genet. Sel. Evol.*, 41:56.