Implementation and uptake of genomic selection in Irish Holstein-Friesian Dairy Cattle

J.F. Kearney[*], A.R. Cromie[*] and D.P. Berry[†]

Introduction

Genomic estimated breeding values (GEBVs) were introduced for Holstein-Friesian dairy cattle in Ireland in February 2009. At that time GEBVs were from SNP effects estimated from a training population of only AI bulls with daughters in Ireland. However, because of the relatively small size of the national dairy herd and the low usage of AI in Ireland, the potential size of the training population would be limited if restricted to only AI sires with daughters in Ireland. Since 2008 Ireland has been continuously swapping genotypes with interested parities with a view of including AI sires with MACE evaluations in the training population. The objective of this paper is to outline the impact of including bulls with MACE evaluations for milk and fertility in genomic evaluations in Ireland and also to retrospectively evaluate the uptake of genomic selection in Ireland since its introduction. We also compare the GEBVs of the first group of Holstein-Friesian bulls that were genomically selected in 2009 but now have milk production estimated breeding values (EBVs) based on actual daughter performance.

Material and methods

Estimation of GEBV. Genotype data on 54,001 single nucleotide polymorphisms (SNPs) from the Illumina Bovine50 Beadchip on 4,318 Holstein-Friesian AI sires with daughters in Ireland were available. A total of 2,419 SNPs on the X-chromosome or of no known position on the genome were discarded; a further 234 SNPs that did not follow Mendelian inheritance patterns between sires and sons were also discarded. None of the 4,318 individuals had >2% genotype discrepancies between sire-son pairs of genotypes. Of the remaining SNPs, 3784 were monomorphic and a further 3,415 had a minor allele frequency of <2%; these SNPs were discarded. SNPs were also discarded (in this order) if greater than 5% of calls were missing (n=1,518), SNPs were not segregating (n=15), SNP call clustering was poor (n=15), and SNP calls deviated ($P<0.1 \times 10^{-7}$) from hardy-Weinberg equilibrium (n=696). Following all edits, 41,920 SNPs remained for inclusion in the analysis.

The methodology used for genomic selection in 2009 described in detail by Berry et al. (2009) has remained largely unchanged for the estimation of GEBVs in 2010 in Ireland. Direct genomic values (DGVs) are estimated in Ireland using mixed models equations by replacing the traditional numerator relationship matrix with a genomic relationship matrix (VanRaden (2008)). The dependent variables included in the genomic evaluation are the deregressed traditional EBVs using the full pedigree of the animal (Berry et al., (2009)).

---

[*] Irish Cattle Breeding Federation, Highfield House, Bandon, Co. Cork, Ireland
[†] Moorepark Dairy Production Research Center, Fermoy, Co. Cork, Ireland

Blending of DGVs with traditional EBVs is undertaken using the method described by Berry et al., (2009) which is equivalent to that described by VanRaden (2008).

Testing of genomic selection. To test the accuracy of genomic selection using Irish data only genotyped sires with at least 40 milking daughters in Ireland with a reliability of at least 40% (excluding the parental contribution) were retained. The validation dataset were sires born after 1996 for milk production, 1995 for calving interval and 1994 for survival and had at least 65% reliability. When evaluating the impact of using MACE evaluations, sires in the training population had to have a reliability, excluding parental contribution of >40%. The same validation population was used.

Implementation. For the Spring 2009 breeding season it was decided to publish GEBVs on individual genomcially selected (GS) bulls, and the top GS bulls were included on the Active bull list (i.e., list of the top 75 bulls ranked on EBI with semen available). GS bulls included on the active bull list had to have sufficient progeny born to have a reliability for direct calving difficulty of ≥50% in the country of origin. Also the reliability of the GEBV for EBI, had to be ≥35%. In order to reduce the risk associated with farmers using just the top GS bulls, a limit was placed on the number of straws to be sold from any one GS bull.

Uptake and accuracy of GEBVs. The uptake of the use of GS bulls was assessed on 492,000 AI insemination records collected during 2009. Initial results on the predictive ability of genomic selection was assessed by correlating parental average (PA), DGVs, and GEBVs estimated in February 2009 on 35 young test sires with their daughter proven EBV estimated in August 2009.

Results and discussion

The intercept and linear regression coefficient of the number of alleles shared per individual on the additive relationship matrix was 11497 and 17245, respectively; the correlation between the genomic relatedness between all pairs of sires and their respective expected additive relatedness was 0.77. Table 1 shows the correlation and mean bias obtained between routinely estimated breeding values and the DGV and GEBV for the main traits in the Irish total merit index, the economic breeding index (EBI). Better predictive ability, as evidenced by stronger correlations and reduced absolute bias, was obtained when the larger training population was used. As a result the reliability of the young bulls is higher by about 10 percentage units when using the larger training population (Figure 1). The average reliability of young bulls with no genomic information is approximately 0.32. The lower number of records in the training population for calving interval and survival is attributable to their lower heritability and therefore the inability of some bulls to reach the reliability threshold to be included in the training population.

Uptake of GEBV bulls. Thirty four percent of inseminations in 2009 were to GS bulls, with 37% and 29%, to domestically and internationally proven bulls, respectively. The primary reason for this was that the GS bulls were of higher genetic merit than the proven bulls. The average EBI of the GS bulls used was approximately one standard deviation more than the average EBI of the proven bulls that were used. In line with recommendations on reducing risk associated with using GS bulls, farmers used, on average, at least 4 different GS bulls. This compared to 2.7 for proven bulls.

Table 1: Correlations and mean bias between estimated breeding values using traditional methods, and both DGVs and GEBVs when the training population is based on AI sires with only daughters in Ireland (Domestic) or MACE evaluations (MACE).

	Domestic			MACE		
	N	DGV	GEBV	N	DGV	GEBV
Correlation						
Milk (kg)	883	0.66	0.69	3204	0.77	0.77
Fat (kg)	883	0.65	0.65	3204	0.73	0.72
Protein (kg)	883	0.67	0.69	3204	0.77	0.77
SCC (SCC units)	883	0.41	0.47	3204	0.60	0.63
Calving Interval (days)	542	0.66	0.69	1447	0.73	0.73
Survival (%)	542	0.55	0.58	1108	0.58	0.59
Bias						
Milk (kg)	883	36	46.3	3204	35.8	37.6
Fat (kg)	883	-0.88	-0.58	3204	-0.36	-0.25
Protein (kg)	883	-0.14	0.3	3204	0.4	0.62
SCC (SCC units)	883	0.02	0.02	3204	0.02	0.02
Calving Interval (days)	542	-1.13	-1.29	1447	-1.11	-1.17
Survival (%)	542	0.54	0.76	1108	0.6	0.71

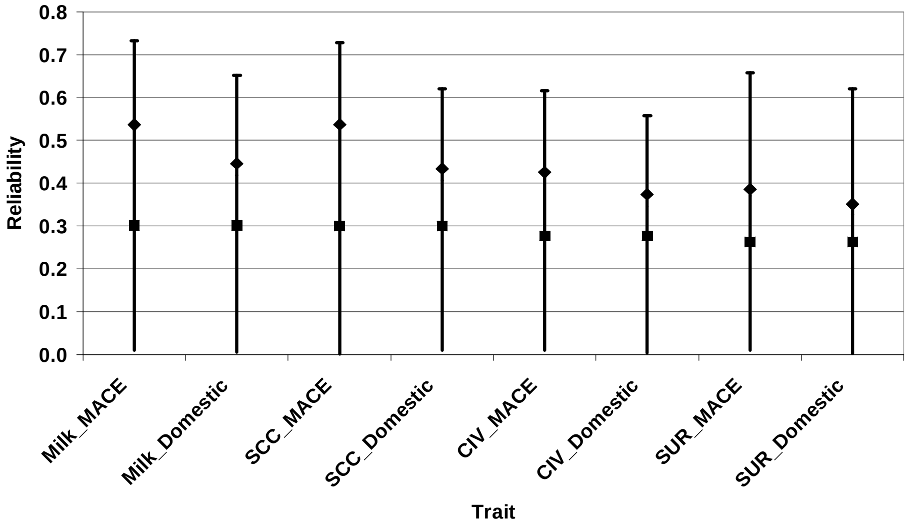

Figure 1: Predicted reliability for parent average (n) and GEBV (u) for young bulls when using MACE (*trait*_MACE) or domestic proven (*trait*_Domestic) bulls only. Error bars represent the minimum and maximum genomic relaitbility for individual bulls.

Accuracy of GEBV prediction in sires proven in 2010. The correlation and mean difference between parent average, DGV, GEBV and daughter EBV are given in Table 3. The correlation between parent average and daughter based EBVs were consistently lower than the respective correlations with the GEBV and DGV. At this stage the DGVs are the best predictors of progeny performance however one must recognize the limitations of this analysis due because only 35 bulls being included in this comparison, the average reliability of the sires is 80%, and the daughter records are not completed lactation records. No data on fertility is currently available.

Table 3: Correlations and mean difference between daughter proofs and GEBV, DGV, and parent average for 35 GS bulls selected when in lay-off in Spring 2009 but now with greater than 70% reliability for production traits based on milking daughters.

	Correlation			Mean Difference		
	GEBV	DGV	PA	GEBV	DGV	PA
Milk (kg)	0.64	0.65	0.63	65	50	77
Fat (kg)	0.51	0.57	0.4	2	2	3
Prot (kg)	0.59	0.65	0.53	2	1.5	2.2
SCC	0.69	0.69	0.61	0.02	0.01	0.02
Fat %	0.79	0.73	0.76	0.01	-0.01	-0.01
Protein %	0.70	0.70	0.69	0.003	0.003	0.003

Conclusions

Overall the implementation of genomic evaluations in Ireland has been very successful. The exchange of genotypes with other countries has resulted in an increase in reliability for young bulls. The uptake of the GS bulls has been very encouraging with farmers using several bulls as recommended to reduce the risks of using less reliable bulls. Simulations using Irish population parameters shows genetic gain could be increased by 50% with the same number of bulls progeny tested (Sonesson et al., (2008)).

Acknowledgements

Research funding was provided by the Irish Dairy Research Trust and the Research Stimulus Fund (RSF-06-0353; RSF-06-0409) from the Department of Agriculture, Food and Fisheries, Ireland.

References

Berry, D.P., Kearney, J.F., and Harris, B. (2009). *Interbull Bulletin*. 39:34-37.

Sonesson, A., Meuwissen, T.H.E., and Cromie, A.R. (2008). Genomic Selection in the Irish dairy cattle breeding scheme. *Technical Report*. (www.icbf.com)

VanRaden, P.M. (2008) *J. Dairy Sci*., 91: 4414-4423.