Applying single step genomic evaluation and integrating local with foreign EBVs to increase calving performance

S. Naderi, T. Pabiou, R.D. Evans
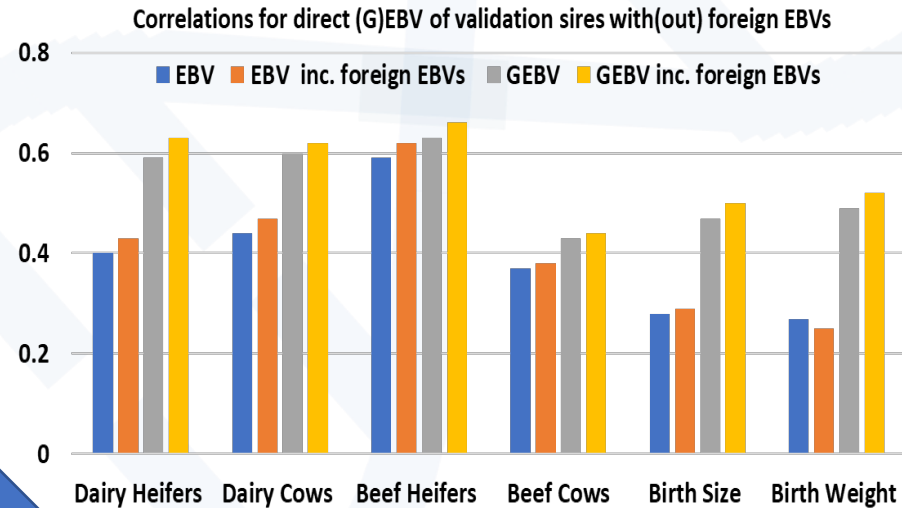
Objectives:

- Validate EBV and single step GEBV for calving traits
- Validate the integration of international EBVs

DATA & TRAITS

Traits	Records No.	Phenotypes with a genotype
Dairy Heifers	1,894,408	3%
Dairy Cows	6,347,887	3%
Beef Heifers	924,801	16%
Beef Cows	5,889,128	14%
Birth Size	1,932,005	31%
Birth Weight	199,759	19%
Animals in pedigree	25,957,178	
Animals with genotypes	1,775,185	

Methods:

- Multi-trait across breed linear mixed model in MiX99
- Interbull MACE calving ebvs for Dairy Cows & Interbeef calving and birthweight ebvs for Beef Cows were integrated into the model (Vandenplas et al. 2014 GSE)

Validation:

- Sires with at least 100 progeny for each trait
- Correlation between (G)EBVs and DYDs
- Slope of regressions of DYD on (G)EBV to measure bias

Results

Correlations for direct (G)EBV of validation sires with(out) foreign EBVs

Legend: EBV, EBV inc. foreign EBVs, GEBV, GEBV inc. foreign EBVs

Slopes of regressions of DYD on (G)EBV of the validation sires

Legend: EBV, EBV inc. foreign EBVs, GEBV, GEBV inc. foreign EBVs

Conclusions

Genomics and integration of external ebvs improved the prediction accuracy
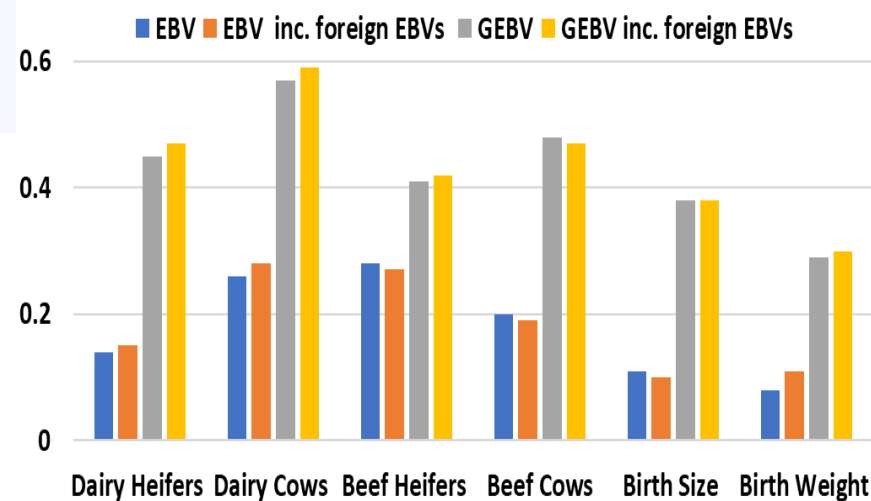
Improvement in slopes with the inclusion of genomic information

Saeid Naderi: snaderi@icbf.com

Applying single step genomic evaluation and integrating local with foreign EBVs to increase calving performance

S. Naderi[*], T. Pabiou, R.D. Evans

The Irish Cattle Breeding Federation, Highfield House, Newcestown Road, Bandon, Cork, Ireland; snaderi@icbf.com

Abstract
The objective of this study was to validate single step genomic breeding values (ssGEBV) for direct calving difficulty compared with non-genomic breeding values (EBV) and integrated (ssG)EBV. Calving difficulty genetic evaluation are undertaken for dairy heifers (DH), dairy cows (DC), beef heifers (BH) and beef cows (BC) calving scores with birth weight and birth size included as predictor traits. Single step GTBLUP approach was applied for a 6-trait calving difficulty evaluation on 1.77 million genotyped animals. To incorporate the international EBVs into domestic evaluation a joint evaluation was run where deregressed EBVs (DRP) were used as extra records weighted by their associated effective record contributions (ERC). The correlation between (G)EBV from well proven sires and their adjusted phenotypes was considered as the accuracy. Including genomic information and applying ssGTBLUP improved the accuracy for all six traits. In addition, blending the international EBVs into the domestic evaluation further improved the accuracy.

Introduction
The current routine GEBV estimation for Irish dairy and beef calving performance traits are based on a two-step approach, in which a reference population is first built to estimate the SNP marker effects, which are then used to calculate GEBVs for young animals. Incorporating the genomic information into traditional animal models used in national evaluations and applying a single-step genomic evaluations (ssGBLUP) could reduce potential biases in the GEBV estimation by simultaneously combining genomic, pedigree and all local phenotypic information. ssGBLUP has not been broadly implemented in large-scale national dairy or beef cattle evaluations due to large computational cost. Mäntysaari et al. (2017) developed an efficient computational approach for ssGBLUP called ssGTBLUP which requires less computation than ssGBLUP in iteratively solving by a preconditioned conjugate gradient (PCG) method. Moreover, integrating foreign information (i.e., MACE) into the national evaluation is advantageous but poses further challenges. Implementing a joint evaluation with national phenotypes in a single evaluation would enable the propagation of foreign information into domestic animals' evaluations even if they did not participate in the international evaluation itself. Therefore, the objective of this study was to explore and validate a ssGTBLUP approach comparing it with a traditional evaluation but also exploring the benefit of integrating foreign EBVs into the Irish genetic evaluation for calving performance traits.

Materials & Methods
Calving difficulty in Ireland is assessed on a scale from 1 (i.e, no assistance) to 4 (i.e., with veterinary assistance). Four different calving traits are defined: dairy heifers (DH), dairy cows (DC), beef heifers (BH) and beef cows (BC) were defined. Birth weight (BW; kg) and birth size (BSIZE; 1 to 5 scale) are included in the multi-trait genetic evaluation as predictor traits. The number of records as well as some statistics for each trait is in Table 1. The total number

of animals in the pedigree was 25,957,178 while the number of genotyped animals in the evaluation was 1,775,185.

Table 1. Number of records and the statistics for six different traits of dairy heifers (DH), dairy cows (DC), beef heifers (BH), beef cows (BC), birth weight (BW; kg) and birth size (BSIZE).

Traits	MEAN	SD[1]	MIN[2]	MAX[3]	No.[4]
DH	1.31	0.60	1.00	4.00	1,894,408
DC	1.21	0.51	1.00	4.00	6,347,887
BH	1.40	0.71	1.00	4.00	924,801
BC	1.26	0.57	1.00	4.00	5,889,128
BSIZE	3.16	0.68	1.00	5.00	1,932,005
BW	41.47	7.91	20.00	115.00	199,759

[1] Standard deviation (SD), [2] Minimum (MIN), [3]Maximum (MAX) and [4] Number of phenotypic records (No.) for each trait.

Traditional breeding values (EBVs) as well as genomic breeding values (GEBVs) were estimated for each of the six traits using the following multiple linear mixed model:

$$y_i = \mathbf{X}b + Z_a u_a + Z_m u_m + Z_{mpe} u_{mpe} + \mathbf{e}$$

where y_i is a vector of phenotypes record for all animals with records and for trait i (i =1: 6), $\mathbf{X}$ is the incidence matrix for fixed effects, $\mathbf{b}$ is the vector of fixed effects consisting of breed proportions and specific heterosis effects, birth year, sex of calf and age of dam, Z_a, Z_m and Z_{mpe} are the incidence matrices relating the random effects u_a for direct additive genetic, u_m for maternal genetic and u_{mpe} for maternal environmental effects to observations in $\mathbf{y}$; and $\mathbf{e}$ is a vector of random residuals. For the traditional genetic evaluation, it was assumed that $u_{a(m)} \sim N\left(0,\ \mathbf{A}\sigma^2_{ua(m)}\right)$ where $\sigma^2_{ua(m)}$ are the direct(maternal) additive genetic variances and $\mathbf{A}$ is pedigree relationship matrix, and $u_{mpe} \sim N\left(0,\ \mathbf{I}\sigma^2_{umpe}\right)$ where σ^2_{umpe} are the maternal environmental variance For the estimation of GEBV, an efficient computational approach presented by Mäntysaari et al. (2017) called ssGTBLUP was used where the genomic relationship matrix was constructed based on the following equation:

$$\mathbf{G}_C = \mathbf{G}_0 + \mathbf{C}$$

where $\mathbf{C} = w\mathbf{A}_{gg}$, w is the residual polygenic proportion, $\mathbf{A}_{gg}$ is the pedigree relationship matrix among the genotyped animals and $\mathbf{G}_0$ is the genomic relationship matrix constructed based on VanRaden (2008) method 1. The matrices and vectors which are describing the linear equations system as well as efficient computation method with low memory cost for solving the equations in ssGTBLUP is outlined in detail by Mäntysaari et al. 2020.

To integrate direct Interbull EBVs and reliabilities into the Irish dairy evaluation as well as direct and maternal Interbeef EBVs and reliabilities, direct and maternal DRP were included into the model as extra traits weighted with their corresponding ERC. The ERC for both direct and maternal EBVs were computed using the reversed Harris et al. (1988) algorithm based on the reliability, pedigree and the traits' heritability. Direct and maternal EBVs were deregressed into direct and maternal DRP using the methods proposed by Jairath et al. (1998). The direct ERC and DRP for Interbull information and ERC and DRP for direct and maternal effects using Interbeef information were computed independently. Details on avoiding double-counting of Irish and foreign information and calculating adjusted ERCs and DRPs as well as the modelling are described by Vandenplas et al. (2014) and Pabiou et al. (2018). For all the computations, the MiX99 software (MiX99 development team, 2020) was used.

Validation. To assess the model prediction accuracy a group of "well" proven sires with at least 100 progeny in the domestic evaluation were identified for each trait. The progeny records of these sires were masked in all evaluations. The correlation between the (G)EBVs of those proven sires with their daughter yield deviation (DYD) reflected the prediction performance. DYDs were calculated based on the average yield deviation from progenies of each sire including all phenotypic records (full data set). The correlation between (G)EBV from validation animals (progeny of the proven sires) and their phenotypes which were corrected for all nongenetic effects was considered as prediction accuracy at animal level. The validation of the integrated (G)EBV was performed by comparing integrated (G)EBV with non-integrated (G)EBV for the corresponding evaluations. The slopes of the regressions of DYD on (G)EBV of the validation sires was computed to measure bias of the evaluations.

Results and Discussion

Validation results from four different models at sire level as well as number of validation sires for each trait are presented in Table 2. The results demonstrate that including genomic information and applying the single step genomic approach improved the prediction accuracy for all six different traits. For instance, the prediction accuracy for DH increased from 0.40 to 0.59 by incorporating the genomic information into the model.

Table 2. Correlations and slope for direct (G)EBV of validation sires with(out) foreign EBVs for six different traits of dairy heifers (DH), dairy cows (DC), beef heifers (BH), beef cows (BC), birth weight (BW; kg) and birth size (BSIZE).

Trait	No. of sires	Statistic	EBV	EBV inc. foreign EBVs	GEBV	GEBV inc. foreign EBVs
DH	1,143	Cor.	0.4	0.43	0.59	0.63
		Slope	0.14	0.15	0.45	0.47
DC	5,194	Cor.	0.44	0.47	0.60	0.62
		Slope	0.26	0.28	0.57	0.59
BH	343	Cor.	0.59	0.62	0.63	0.66
		Slope	0.28	0.27	0.41	0.42
BC	9,289	Cor.	0.37	0.38	0.43	0.44
		Slope	0.20	0.19	0.48	0.47
BSIZE	2,235	Cor.	0.28	0.29	0.47	0.50
		Slope	0.11	0.10	0.38	0.38
BW	363	Cor.	0.27	0.25	0.49	0.52
		Slope	0.08	0.11	0.29	0.30

The greatest improvement in accuracy was for BW where the traditional approach with only pedigree information had a correlation of only 0.27 which improved to 0.49 when the ssGBLUP approach was applied. Across all traits, the slope of DYD on (G)EBV of validation sires where the expectation of the slope was 0.5, improved with the inclusion of genomic information into the model indicating ssGBLUP could reduce potential biases.

The integration of Interbull and Interbeef information into the domestic evaluation also improved prediction accuracy at both the EBV or GEBV level and across all six different traits; however, the gain in prediction performance was not huge and in context of bias reduction only small improvement was observed. Small number of sires in the validation set which had a foreign sire with a MACE or an Interbeef proof could explain this modest improvement (i.e., only 256 sires for DC). Prediction performance at the animal level followed the same pattern as observed at sire level with the inclusion of genomic information or foreign information

improving the accuracy (Table 3). The gain in prediction accuracies was more pronounced for incorporating the genomic information rather than integrating foreign EBVs into the model.

Table 3. Correlations for direct (G)EBV of validation animals with(out) foreign EBVs for six different traits of dairy heifers (DH), dairy cows (DC), beef heifers (BH), beef cows (BC), birth weight (BW; kg) and birth size (BSIZE).

Trait	EBV	EBV inc. foreign EBVs	GEBV	GEBV inc. foreign EBVs
DH	0.17	0.18	0.2	0.22
DC	0.14	0.16	0.17	0.18
BH	0.15	0.16	0.17	0.19
BC	0.13	0.14	0.2	0.21
BSIZE	0.12	0.12	0.2	0.2
BW	0.18	0.18	0.26	0.27

Conclusion. The result from present study is encouraging in that the single step genomic approach and using foreign breeding values in the Irish domestic genetic evaluation improved prediction performance and reduced the bias in the evaluations. However, before implementing, the validation results should be compared with the current routine two-step approach.

References
Harris B. and Johnson D. (1998). J. Dairy Sci. 81(10): 2723–2728. https://doi.org/10.3168/jds.S0022-0302(98)75829-1.
Jairath L., Dekkers J. C. M., Schaeffer L. R., Liu Z., Burnside E.B., and et al. (1998). J. Dairy Sci. 81(2): 550– 562. https://doi.org/ 10.3168/jds.S0022-0302(98)75607-3.
MiX99 Development Team, (2020) MiX99: Natural Resources Institute Finland (Luke). Jokioinen, Finland. URL: http://www.luke.fi/mix99.
Mäntysaari E.A., Evans R. D., and Strandén I. (2017). J. Anim. Sci. 95(11): 4728:4737. https://doi.org/ 10.2527/jas2017.1912
Mäntysaari E.A., Koivula M., and Strandén I. (2020). J. Dairy Sci. 103(6):5314-5326. https://doi.org/ 10.3168/jds.2019-17754.
Pabiou T., Pitkanen T., Evans R. D., Herpje E., and Vandenplas J. (2018). Proc. Of WCGALP, New Zealand, Auckland.
Vandenplas J., Colinet F.G., and Gengler N. (2014). Genet. Sel. Evol. 46 (1): 46-59. https://doi.org/ 10.1186/s12711-014-0059-3.
VanRaden P.M. (2008). J. Dairy Sci. 91 (11): 4414-4423. https://doi.org/ 10.3168/jds.2007-0980.