655. Integration of beef cattle international estimated breeding values in the Italian evaluation

R. Bonifazi[1*], M.P.L. Calus[1], J. ten Napel[1], R.F. Veerkamp[1], S. Biffani[2], M. Cassandro[3,4] S. Savoia[5] and J. Vandenplas[1]

[1]*Animal Breeding and Genomics, Wageningen University & Research, P.O. Box 338, 6700 AH Wageningen, the Netherlands; [2]Istituto Di Biologia E Biotecnologia Agraria, Consiglio Nazionale Delle Ricerche, Via Edoardo Bassini 15, 20133 Milano, Italy; [3]Department of Agronomy, Food, Natural resources, Animals and Environment, University of Padova, Viale dell'Università 16, 35020 Legnaro, Italy; [4]National Federation of National Breeders Associations (FedANA), XXIV Maggio 43, 00187 Roma, Italy; [5]Interbull Centre, Department of Animal Breeding and Genetics, SLU, P.O. Box 7023, 75007 Uppsala, Sweden; renzo.bonifazi@wur.nl*

Abstract

Beef cattle international evaluations, led by Interbeef, combine data from different countries to compute international estimated breeding values (EBV_{INT}) which are expressed on the same country scale as national ones, allowing participating countries to compare national and foreign sires. These two EBV might be in disagreement due to differences in the information used in national and international evaluations. This can be overcome by integrating the EBV_{INT} into the national evaluations to obtain a 'blended' EBV. Here we tested and validated a general procedure to integrate publishable sires EBV_{INT} at the national level. Using evaluations for Limousin weaning weight and the integration for Italy as a case study, we show that the integration procedure accurately blends international information at the national level while avoiding double-counting of national information which are used as input to compute EBV_{INT} in Interbeef evaluations. This procedure can also be used with other traits and breeds.

Introduction

International evaluations allow comparing estimated breeding values (EBV) of foreign elite animals with national animals by computing international EBV (EBV_{INT}) and associated reliabilities (REL_{INT}). EBV_{INT} usually would have higher reliabilities than national EBV (EBV_{NAT}) as international evaluations take into account data from relatives recorded in other countries (Bonifazi *et al.*, 2020b). In beef cattle international evaluations as performed by Interbeef, each participating country receives a list of EBV_{INT} and REL_{INT} for: (1) all animals in the national pedigree; and (2) so-called 'publishable sires', i.e. sires that meet Interbeef publication rules, as described in Bonifazi *et al.* (2021). Thus, an individual could have two EBV at the national level: EBV_{NAT} and EBV_{INT}, which can differ due to differences in information between national and international evaluations (e.g. national evaluations are usually multi-trait and in some cases multi-breed while international evaluations are trait-specific and performed within-breed). To overcome this issue, the EBV_{INT} and its associated measure of precision (e.g. REL_{INT}) can be integrated into the national evaluations (Vandenplas and Gengler, 2015), resulting in a single 'blended' EBV that includes and propagates international information to all animals in the national evaluation. However, double-counting of national information needs to be removed from the EBV_{INT} as Interbeef uses national phenotypes as input for international evaluations. Thus, the objectives of this study were to test and validate the accuracy of a general procedure to integrate publishable sires' EBV_{INT} at the national level using Limousin weaning weight and Italy (ITA) as a case study for national evaluations.

Materials & methods

Data and software. A total of 441,691 Limousin males and females individual age-adjusted weaning weights (AWW) were available. AWW were recorded in eight different countries for a total of six populations: Denmark, Finland and Sweden (modelled as one population), Czech Republic, Ireland,

Germany, Switzerland, and Italy. A total of 7,506 AWW (of which 2,304 from ITA) were recorded from animals born between January and May 2019. The international pedigree included 683,317 animals. All analyses were performed using MiXBLUP (ten Napel *et al.*, 2020): the convergence criterion was 'CD≤10^{-5}', and reliabilities were computed using the Tier and Meyer (2004) algorithm.

International and national evaluations. The direct and maternal EBV_{INT} and REL_{INT} were estimated using the Interbeef model which is equivalent to a multi-trait animal model with maternal effects where each population national model was fitted as a different correlated trait (see Bonifazi *et al.* (2020a) for a complete description). Direct and maternal genetic effects were modelled using an across-populations genetic (co)variance matrix following the Interbeef procedure as described in Bonifazi *et al.* (2021). Random environmental, maternal permanent environmental, and residual effects were modelled using block-diagonal variance matrices. Both genetic and environmental variances were the same as those used in the national genetic evaluations of participating countries. The direct and maternal EBV_{NAT} and REL_{NAT} for ITA were obtained by running a single-trait evaluation using only ITA submitted phenotypes and the same national model as the one used for the international evaluations.

Integration. After EBV_{INT} and EBV_{NAT} and their associated REL (i.e. REL_{INT} and REL_{NAT}) on a 0 to 1 scale were obtained following the aforementioned models, publishable sires' international information was integrated into national evaluations in four steps as follows.

Step 1. For each publishable sire *i*, direct and maternal effective record contributions (ERC) associated with REL_{NAT} and REL_{INT} (ERC_{NAT} and ERC_{INT}, respectively) were computed as: $ERC_i = \lambda \cdot REL_i / (1 - REL_i)$, where $\lambda = \sigma^2_{residual} / \sigma^2_{genetic}$ with $\sigma^2_{genetic}$ being either the direct or maternal genetic variance.

Step 2. For each publishable sire *i*, direct and maternal de-regressed proofs for both national and international EBV (DRP_{NAT} and DRP_{INT}, respectively) were computed following Garrick *et al.* (2009): $DRP_i = PA_i + (EBV_i - PA_i) / REL_{i(o+p)}$, where PA_i is the parent average EBV of the individual computed as $(EBV_{sire} + EBV_{dam})/2$, and $REL_{i(o+p)}$ is the reliability due to the individual own performance (o) and its progeny (p) computed as $dERC_i / (dERC_i + \lambda)$. The de-regressed ERC for the individual *i*, $dERC_i$, is computed as $ERC_i - ERC_{PA}$, with ERC_{PA} being the ERC calculated from parent average reliability defined as $(REL_{sire} + REL_{dam})/4$. If the $dERC_i$ was ≤0, both the $dERC_i$ and its associated DRP_i were set to 0.

Step 3. For each publishable sire *i*, to avoid double-counting of national information, the direct and maternal adjusted DRP (DRP^*) and its associated weight ($dERC^*$) adjusted for national data were computed based on Vandenplas *et al.* (2014): $DRP^*_i = ((dERC_{INTi} \cdot DRP_{INTi}) - (dERC_{NATi} \cdot DRP_{NATi})) / dERC^*_i$, where $dERC^*_i = dERC_{INTi} - dERC_{NATi}$. If $dERC^*_i$ was ≤0 or if the gain in reliability (defined as the difference between REL_{INT} and REL_{NAT}) was smaller than 0.01, both the $dERC^*_i$ and its associated DRP^*_i were set to 0.

Step 4. The direct and maternal blended EBV were computed with a national evaluation using national phenotypes and publishable sires' direct and maternal DRP^* as pseudo-phenotypes. The $dERC^*$ were used as weights for the DRP^*. The maternal DRP^* were associated with the maternal effect of the animal itself and not of its dam. The direct and maternal DRP^* were modelled as additional phenotypes for the same trait and fitted using one general mean for each DRP^*.

Scenarios and validation. We applied the above procedure for ITA where national and international data differ due to differences in publication calendars. We integrated EBV_{INT} from an international evaluation in January 2019 into a national evaluation in May 2019. All analyses used the complete international pedigree. Hereafter, we will refer to ITA phenotypes as 'national', and to phenotypes from other countries

as 'foreign'. First, the two following evaluations were performed and their outputs were used as inputs for the integration procedure:

- INT_{JAN}. An international evaluation using both national and foreign phenotypes of animals born up to January 2019. Publishable sires to be integrated are selected from this evaluation.
- NAT_{JAN}. A national evaluation using only national phenotypes of animals born up to January 2019. EBV and REL from this evaluation are used during the integration procedure to avoid double-counting of national information.

Then, three scenarios were implemented and validated:

- NAT_{MAY}. A national evaluation using only national phenotypes of animals born up to May 2019. This scenario is used for comparison with $BLEND_{MAY}$ and GOLD to illustrate the benefits of the integration procedure over a national evaluation without any integration.
- $BLEND_{MAY}$. A blended national evaluation using national phenotypes of animals born up to May 2019 with the integration of publishable sires' information from INT_{JAN}.
- GOLD. As $BLEND_{MAY}$ but integrating information of publishable sires from an international evaluation using only foreign phenotypes of animals born up to January 2019. This international evaluation avoids double-counting of national information when integrating publishable sires' information in GOLD, providing a 'gold standard' for comparison with $BLEND_{MAY}$.

We validated the integration procedure by comparing publishable sires' EBV obtained with either NAT_{MAY}, $BLEND_{MAY}$, or GOLD with their EBV from a reference scenario (REF) that is an international evaluation using national phenotypes of animals born up to May 2019 and foreign phenotypes of animals born up to January 2019. Scenario REF uses the same sources of information as in $BLEND_{MAY}$ and GOLD, but without approximating foreign information by DRP* and dERC*. Therefore, the more accurate the integration procedure, the closer the EBV are to those of REF. As validation metrics, we computed the Pearson correlations (ρ) between EBV, and the slopes (b_1) of the regression of REF EBV on the EBV of the validated scenarios. Publishable sires were divided into three groups: (A) domestic publishable sires with ≥15 recorded offspring in ITA; (B) domestic publishable sires with <15 recorded offspring in ITA; and (C) foreign publishable sires with no recorded offspring in ITA.

Results

The number of publishable sires in groups A, B and C was 1,382, 94 and 3,470 for direct EBV, and 491, 51, and 1,165 for maternal EBV, respectively. The mean dERC* of all publishable sires and the number of dERC*>0 were similar between $BLEND_{MAY}$ and GOLD (Table 1). As expected, for direct EBV, group A of publishable sires had a smaller amount of international information to be integrated (mean dERC* equal to 0.5 for $BLEND_{MAY}$) compared to sires in groups B and C (mean dERC* equal to 5.2 and 2.5, respectively, for $BLEND_{MAY}$) (Table 1). For all groups, direct EBV ρ and b_1 improved in both $BLEND_{MAY}$ (0.98 and 1.02, respectively, on average across groups) and GOLD (0.98 and 1.06 on average) compared to NAT_{MAY} (0.60 and 0.77 on average) (Table 1). For maternal EBV, $BLEND_{MAY}$ and GOLD resulted in similar ρ and b_1 as NAT_{MAY} for all groups except for foreign sires (Table 1). For foreign sires, $BLEND_{MAY}$ and GOLD had ρ closer to 1 but worse b_1 compared to NAT_{MAY}. Overall, $BLEND_{MAY}$ gave similar ρ and similar or slightly worse b_1 than GOLD: on average across groups and effects, ρ and b_1 were 0.93 and 0.88 for $BLEND_{MAY}$, and 0.93 and 0.94 for GOLD.

Table 1. Validation results for direct and maternal EBV of publishable sires.[1] All scenarios are validated against EBV from scenario REF.[2]

Sire group	Scenario[2]	Direct EBV				Maternal EBV			
		ρ	b_1	Mean dERC*	n. dERC* >0	ρ	b_1	Mean dERC*	n. dERC* >0
A (Domestic ≥15 off)	NAT$_{MAY}$	0.95	0.96	-	-	0.99	1.01	-	-
	BLEND$_{MAY}$	0.99	0.96	0.5	97	0.98	0.96	0.1	52
	GOLD	0.99	1.01	0.3	121	0.98	1.02	0.1	70
B (Domestic <15 off)	NAT$_{MAY}$	0.63	0.66	-	-	0.86	0.80	-	-
	BLEND$_{MAY}$	0.97	0.96	5.2	94	0.84	0.72	0.7	51
	GOLD	0.97	1.03	4.0	94	0.86	0.87	0.7	51
C (Foreign)	NAT$_{MAY}$	0.24	0.69	-	-	0.51	0.98	-	-
	BLEND$_{MAY}$	0.97	1.13	2.5	3,469	0.83	0.57	0.5	1,127
	GOLD	0.97	1.14	2.4	3,470	0.81	0.58	0.4	1,165

[1] ρ = Pearson correlation; b_1 = slope; dERC* = adjusted de-regressed ERC; n. = number; off = recorded offspring in ITA.

[2] REF = reference international evaluation; NAT$_{MAY}$ = national evaluation without integration; BLEND$_{MAY}$ = blended evaluation with integration of publishable sire's information; GOLD = as BLEND$_{MAY}$, but integrating publishable sire's information computed without national data.

Discussion

Our results show that a blended evaluation is especially beneficial for sires in groups B and C which have only a few or no recorded offspring at the national level. Also, double-counting of national information during blending was properly removed as shown by the good agreement of BLEND$_{MAY}$ and GOLD. Scenario GOLD completely avoids double-counting of national information but would require Interbeef to run and distribute for each population EBV and REL from an international evaluation from which national phenotypes are removed. Instead, BLEND$_{MAY}$ can be directly applied by countries, assuming that the publishable sires' parental EBV$_{INT}$ and REL$_{INT}$ are known and that national information submitted for the international evaluation (here NAT$_{JAN}$) are available. Overall, maternal EBV showed lower ρ and b_1 compared to direct EBV, which could be due to the lower gains in REL associated with maternal EBV compared to direct EBV. The outlined integration procedure enables the integration of publishable sires' EBV$_{INT}$ at the national level in a single blended EBV. Results show that the integration procedure yields accurate results for both direct and maternal EBV.

References

Bonifazi R., Vandenplas J., ten Napel J., Matilainen K., Veerkamp R.F., *et al.* (2020) Genet. Sel. Evol. 52(32):1-16. https://doi.org/10.1186/s12711-020-00551-9

Bonifazi R., Vandenplas J., ten Napel J., Cromie A., Veerkamp R.F., *et al.* (2020b) Acta Fytotech. Zootech. 23:144-155. https://doi.org/10.15414/afz.2020.23.mi-fpap.144-155

Bonifazi R., Vandenplas J., ten Napel J., Veerkamp R.F., and Calus M.P.L. (2021) J. Anim. Sci. 99(9):1-14. https://doi.org/10.1093/jas/skab222

Garrick D.J., Taylor J.F., and Fernando R.L. (2009) Genet. Sel. Evol. 41(55):1-8. https://doi.org/10.1186/1297-9686-41-55

Ten Napel J., Vandenplas J., Lidauer M., Stranden I., Taskinen M., *et al.* (2020) MiXBLUP, user-friendly software for large genetic evaluation systems.

Tier B., and Meyer K. (2004) J. Anim. Breed. Genet. 121(2):77-89. https://doi.org/10.1111/j.1439-0388.2003.00444.x

Vandenplas, J., Colinet, F.G., and Gengler, N. (2014) Genet. Sel. Evol. 46(59):1-15. https://doi.org/10.1186/S12711-014-0059-3

Vandenplas J., and Gengler N. (2015) Livest. Sci. 181:121-130. https://doi.org/10.1016/j.livsci.2015.09.012