Journal of
Animal Breeding and Genetics

J. Anim. Breed. Genet. ISSN 0931-2668

ORIGINAL ARTICLE

An integration of external information for foreign stallions into the Belgian genetic evaluation for jumping horses

J. Vandenplas[1,2], S. Janssens[3], N. Buys[3] & N. Gengler[1]

1 Animal Science Unit, Gembloux Agro Bio-Tech, University of Liège, Gembloux, Belgium
2 National Fund for Scientific Research, Brussels, Belgium
3 Livestock Genetics, Department Biosystems, Katholieke Universiteit Leuven, Heverlee, Belgium

Keywords
Bayesian approach; external information; integration; jumping horses.

Correspondence
J. Vandenplas, Gembloux Agro-Bio Tech, University of Liege, Passage des Déportés-2, Gembloux 5030, Belgium.
Tel: +32 81 62 23 58; Fax: +32 81 62 21 15;
E-mail: jvandenplas@ulg.ac.be

Received: 3 May 2012;
accepted: 3 October 2012

Summary

The aim of this study was to test the integration of external information, i.e. foreign estimated breeding values (EBV) and the associated reliabilities (REL), for stallions into the Belgian genetic evaluation for jumping horses. The Belgian model is a bivariate repeatability Best Linear Unbiased Prediction animal model only based on Belgian performances, while Belgian breeders import horses from neighbouring countries. Hence, use of external information is needed as prior to achieve more accurate EBV. Pedigree and performance data contained 101 382 horses and 712 212 performances, respectively. After conversion to the Belgian trait, external information of 98 French and 67 Dutch stallions was integrated into the Belgian evaluation. Resulting Belgian rankings of the foreign stallions were more similar to foreign rankings according to the increase of the rank correlations of at least 12%. REL of their EBV were improved of at least 2% on average. External information was partially to totally equivalent to 4 years of contemporary horses' performances or to all the stallions' own performances. All these results showed the interest to integrate external information into the Belgian evaluation.

Introduction

The Belgian sport horse population is situated at the crossroads of different sport horse populations, which leads to a mix of the European genes. Artificial insemination facilitates the use of foreign stallions since the 1980s. For these reasons, Belgium seems to be one of the centres for European sport horse breeding (Ruhlmann *et al.* 2009a). Three Belgian studbooks of warmblood horses are involved, that is the Belgian Warmblood Horse Studbook (BWP), the Royal Belgian Sports Horse Society (sBs) and the Studbook Zangersheide. For all three, the improvement in the performances in show jumping is an important breeding objective (Koenen *et al.* 2004).

Since 1998, a genetic evaluation for show jumping horses is implemented in Belgium, and as in most other European countries, the estimated breeding values (EBV) are based on national information only (Koenen & Aldridge 2002; Janssens *et al.* 2007), whereas the Belgian sport horse population is clearly linked with other foreign studbooks. This may lead to a limited reliability (REL) of EBV for horses with few Belgian records and to inappropriate breeders' choices of a stallion on the international scene. Similar issues exist in other countries as sport horse breeding is very international. Based on experiences in dairy cattle breeding, for which 'Interbull' (Uppsala, Sweden) provides sire EBV for dairy cattle from different countries, an international group of scientists and breeding organizations, called 'Interstallion', was created in 1998 to achieve reliable breeding values across countries for sport horse stallions. Within the framework of this group, Ruhlmann *et al.* (2009b) concluded that an international evaluation of jumping horses is feasible. However, such an international genetic evaluation combining all information sources is not yet available, and one option is to integrate external information into the local genetic evaluation. Different theoretical approaches exist to do this. In the case of multibreed genetic evaluations for beef cattle, Klei *et al.* (1996) proposed a Bayesian approach where external information is considered as prior information for the local evaluation. Two different Bayesian derivations were proposed by Quaas & Zhang (2006) and Legarra *et al.* (2007). Recently, Vandenplas & Gengler (2012) proposed some improvements to these methods, especially to take into account the double counting among related external animals. However, some issues arise before the implementation of a Bayesian procedure, like the independence of the external evaluations from the internal one or the similarities between the external and internal evaluated traits (Gengler & Vanderick 2008).

The first aim of this study was to apply a Bayesian approach to integrate external information, i.e. foreign EBV and their associated REL, for stallions into the Belgian genetic evaluation of show jumping horses, and the second aim was to test the model adequacy and the predictive ability of the applied method.

Materials and methods

Performance data on show jumping were provided by the horse riding organization for national level competitions, the Royal Belgian Federation for Equestrian Sports (KBRSF), and by the horse riding organization for recreational level competitions, the Rural Riding Association (LRV). The available performance data (data I) included 710 212 performances from 44 755 competitive horses during the period 1991–2009. Performances in show jumping consisted of ranking of horses participating in show jumping competitions converted into normalized score by a Blom's approximation (Janssens *et al.* 2007). These performances were also considered as two traits in terms of competition levels, that is the KBRSF level and the LRV level. The KBRSF level was considered as the Belgian breeding goal trait (hereafter called Belgian trait). The pedigree file, a combination of pedigree records provided by sBs and BWP, included 101 382 registered horses.

The following bivariate repeatability Best Linear Unbiased Prediction (BLUP) animal model was applied to perform the Belgian genetic evaluation (evaluation A; Janssens *et al.* 1997, 2007):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Zu} + \mathbf{Zp} + \mathbf{e} \qquad (1)$$

where $\mathbf{y}$ was the vector of performances, $\boldsymbol{\beta}$ was the vector of fixed effects, $\mathbf{u}$ was the vector of random additive genetic effects, $\mathbf{p}$ was the vector of random permanent environmental effects, and $\mathbf{e}$ was the vector of residuals. $\mathbf{X}$ and $\mathbf{Z}$ were incidence matrices relating performances to fixed effects and to random effects, respectively.

Fixed effects were the age of the participating horse, its sex and the show jumping event organized by the KBRSF or by the LRV in which it participated. Variance components for the random permanent environmental and genetic effects used for this study were those estimated by Janssens *et al.* (1997). Heritability was equal to 0.10 for performances at the KBRSF level (i.e. the Belgian trait) and to 0.11 for performances at the LRV level. Genetic correlation between these two traits was equal to 0.63.

Despite the fact that two traits are evaluated, only breeding values for performances in KBRSF level estimated using both performances in KBRSF level and LRV level (i.e. EBV estimated by the evaluation A for the Belgian trait; EBV_A) of stallions approved by BWP and/or SBS are published on a standardized scale following the recommendations of 'Interstallion' (Interstallion 2005; Janssens & Buys 2008).

Reliabilities of EBV_A based on data I (EBV_{AI}; REL_{AI}; Table 1) were computed using the equation:

$$REL = 1 - PEV/\sigma_g^2 \qquad (2)$$

where σ_g^2 is the genetic variance for the Belgian trait and PEV is the prediction error variance obtained from the diagonal element of the inverted left-hand side of the mixed model equation (1).

Available external information consisted of external EBV (EBV_E) and their associated external REL (REL_E) for stallions approved by BWP having a published Belgian index, born after 1978 and originally registered in a Dutch or a French studbook. External information on 98 French stallions and 67 Dutch stallions was provided by the Station de Génétique Quantitative

Table 1 Performed genetic evaluations, estimated breeding values (EBV) and reliabilities (REL)

Genetic evaluation[a]	Data sets[b]		
	I	II	III
A	Evaluation AI, EBV_{AI}, REL_{AI}	Evaluation AII, EBV_{AII}, REL_{AII}	Evaluation AIII, EBV_{AIII}, REL_{AIII}
B	Evaluation BI, EBV_{BI}, REL_{BI}	Evaluation BII, EBV_{BII}, REL_{BII}	Evaluation BIII, EBV_{BIII}, REL_{BIII}

[a]A = Belgian genetic evaluation; B = Bayesian evaluation.
[b]I = complete data; II = data for which all performances after 31 December 2005 were assumed to be missing; III = data for which all the French and Dutch stallions' own performances were assumed to be missing.

et Appliquée, Institut National de la Recherche Agronomique (France) and the Royal Dutch Sport Horse (the Netherlands), respectively. However, because $\mathrm{EBV_E}$ were not the same trait and not expressed on the same scale as the Belgian trait, pre-corrections were needed before its integration into the Belgian genetic evaluation.

First, $\mathrm{EBV_E}$ were converted to the Belgian trait and scale of the year 2009 for both countries. This conversion was performed separately for Dutch and French stallions following the method detailed by Goddard (1985) that regressed previously deregressed internal EBV on external EBV. The 2 samples to estimate conversion equations for the Dutch and French stallions included all Dutch stallions (i.e. 47) and French stallions (i.e. 93) having both an $\mathrm{EBV_{AI}}$ and an $\mathrm{EBV_E}$, respectively. REL of the converted EBV ($\mathrm{REL_c}$) were estimated from all the $\mathrm{REL_E}$ provided by France and by the Netherlands following the method detailed by Goddard (1985) that took into account the error in estimating the true regression equation and the variance of the converted EBV ($\mathrm{EBV_c}$) about the true regression equation. External information with a $\mathrm{REL_c}$ lower than 0.01 was set to missing. It is noted that genetic correlation coefficients for traits between Belgium and the exporting countries were needed for the conversion following Goddard (1985). Because no genetic correlation coefficient was available for the pair Belgium/the Netherlands (Ruhlmann *et al.* 2009b), the genetic correlation coefficients for traits were approximated by the Pearson correlation coefficient between Dutch $\mathrm{EBV_E}$ and $\mathrm{EBV_{AI}}$ estimated for the Dutch stallions for the pair Belgium/the Netherlands and by the Pearson correlation coefficient between French $\mathrm{EBV_E}$ and $\mathrm{EBV_{AI}}$ estimated for the French stallions for the pair Belgium/France. Second, $\mathrm{EBV_E}$ had to be free from internal information to avoid double counting between external and internal information (Gengler & Vanderick 2008). The literature review of Koenen (2002) and van Veldhuizen (1997) for the Dutch genetic evaluation and of Tavernier (1991) and Ricard (1997) for the French genetic evaluation showed that France and the Netherlands never use the same phenotypic information (i.e. same show jumping competitions) as Belgium for their respective genetic evaluations for show jumping. Following the literature, this second condition seemed to be respected. Third, as external information was associated with related stallions, double counting of information among related external stallions could exist. Therefore, the integration of external information was performed following the second version of modified Bayesian evaluation detailed by Vandenplas & Gengler (2012). This approach allows simplifications of the computational burden and takes into account double counting among related animals thanks to the estimation of the contributions due to relationships. These contributions were estimated by a two-step algorithm taken into account all relationships between the foreign stallions and their ancestors.

The equations system of the Belgian model (1) integrating external information (evaluation B) can be written as:

$$\begin{bmatrix} \mathbf{X}^{\mathbf{prime}}\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}^{\mathbf{prime}}\mathbf{R}^{-1}\mathbf{Z} & \mathbf{X}^{\mathbf{prime}}\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}^{\mathbf{prime}}\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}^{\mathbf{prime}}\mathbf{R}^{-1}\mathbf{Z}+\mathbf{G}^{*-1} & \mathbf{Z}^{\mathbf{prime}}\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}^{\mathbf{prime}}\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}^{\mathbf{prime}}\mathbf{R}^{-1}\mathbf{Z} & \mathbf{Z}^{\mathbf{prime}}\mathbf{R}^{-1}\mathbf{Z}+\mathbf{P}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \\ \hat{\mathbf{p}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^{\mathbf{prime}}\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}^{\mathbf{prime}}\mathbf{R}^{-1}\mathbf{y}+\mathbf{G}^{*-1}\boldsymbol{\mu}_0 \\ \mathbf{Z}^{\mathbf{prime}}\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} \quad (3)$$

where $\mathbf{R}$ was the residual (co) variance matrix, $\mathbf{P}$ was the (co)variance matrix for the permanent environment, $\boldsymbol{\mu}_0$ was the vector of $\mathrm{EBV_c}$, and $\mathbf{G}^*$ was the matrix of prediction error (co)variances of $\mathrm{EBV_c}$.

Because the second version of modified Bayesian evaluation (Vandenplas & Gengler 2012) was applied, the inverse of $\mathbf{G}^*$ was equal to $\mathbf{G}^{*-1} = \mathbf{G}^{-1} + \boldsymbol{\Lambda}$ where the matrix $\mathbf{G}^{-1}$ is the inverse of the additive genetic (co)variances matrix and the matrix $\boldsymbol{\Lambda}$ was a block diagonal variance matrix with one block diagonals per horse. For the N stallions associated with external information, the different block diagonals were equal to $\boldsymbol{\Delta}_i\mathbf{G}_0^{-1}\boldsymbol{\Delta}_i$ for $i = 1, \ldots, N$. The matrix $\mathbf{G_0}$ was the matrix of genetic (co)variances among traits, and $\boldsymbol{\Delta}_i$ was a diagonal matrix with elements equal to $\sqrt{\mathrm{RE_{ij}} * \frac{\sigma^2_{u_j}}{\sigma^2_{e_j}}}$ for $j = 1$, 2 traits where $\sigma^2_{e_j}$ was the error variance of this jth trait, $\sigma^2_{u_j}$ was the genetic variance for the jth trait, and $\mathrm{RE_{ij}}$ was equal to the value of records equivalents (RE) only because of own records for the jth trait. RE were estimated thanks to the algorithm taking into account double counting among related animals (Vandenplas & Gengler 2012).

To approximate REL of breeding values estimated by the evaluation B based on data I for the Belgian trait ($\mathrm{EBV_{BI}}$; $\mathrm{REL_{BI}}$), the following procedure was applied. First, for each stallion that had external information, one virtual performance was added to the performance data and weighted by the value of RE independent from contributions due to relationships. The weight for the real performances in the performance data was equal to 1. An additional level for each fixed effect of the Belgian genetic evaluation was created and assigned to the virtual performances to ensure that they had no influence on the genetic evaluation. PEV were estimated by the inversion of the left-hand side of the mixed model equations, and REL_{BI} were calculated using the equation (2). All the genetic evaluations and computations of PEV were performed using the BLUPF90 program family (Misztal 2012) modified to integrate external information by taking into account double counting among related animals.

Descriptive statistics were computed to characterize integrated external information and its influence on the ranking of the horses. With regard to REL_E, REL_c and RE, mean and standard deviations (SD) as well as the number of foreign stallions associated with a non-zero REL_E, a non-zero REL_c and non-zero RE were described. Pearson correlation coefficients between EBV_{AI} and EBV_E and the coefficients of determination of the conversion equations for the Dutch and French stallions were also computed, as well as Spearman rank correlation coefficients among EBV_{AI}, EBV_{BI} and EBV_E, for all the horses, for the French stallions, for the Dutch stallions and for the 100 best stallions. This latter group included the 100 best-ranked stallions following the evaluation AI, born after 1979 and associated with a REL_{AI} equal or higher than 0.75.

The model adequacy was tested by the comparison of accuracy and precision of the evaluations AI and BI through comparisons of mean bias (MB), mean square error of prediction (MSEP) and Pearson correlation coefficients between observed and estimated performances associated with the Belgian trait ($r_{y:\hat{y}}$; Tedeschi 2006). Considering all the horses, MB and MSEP were expressed as a percentage of the average performance of all the performances. Considering the 100 best stallions, the Dutch and French stallions, MB and MSEP were expressed as a percentage of the average performance of their performances, respectively.

To test the predictive ability of the applied method, subsets II and III were created and evaluations A and B were performed based on these two subsets (i.e. evaluations AII, AIII, BII and BIII). Resulting EBV for the Belgian trait are called EBV_{AI}, EBV_{AII}, EBV_{AIII}, EBV_{BI}, EBV_{BII} and EBV_{BIII} (Table 1). Subset II consisted of all performances before 31 December 2005 included. All other performances were assumed to be missing (i.e. 34.5% of all the performances). It simulated the predictive ability of the method routinely applied. Subset III consisted of the complete data except for all French and Dutch stallions' own performances assumed to be missing (i.e. 0.27% of all the performances). It simulated the predictive ability of the method applied for stallions with no own Belgian performances. REL of EBV_{AII}, EBV_{AIII}, EBV_{BII} and EBV_{BIII} (REL_{AII}, REL_{AIII}, REL_{BII} and REL_{BIII}, respectively) were approximated as described previously for evaluations A and B. Pearson correlation coefficients (r_{II}) between EBV_{AI} and EBV_{AII}, and EBV_{BI} and EBV_{BII}, as well as Pearson correlation coefficients (r_{III}) between EBV_{AI} and EBV_{AIII}, and EBV_{BI} and EBV_{BIII}, were estimated. Furthermore, variances of the differences (VAR_{II}) between EBV_{AI} and EBV_{AII}, and EBV_{BI} and EBV_{BII}, as well as variances of the differences (VAR_{III}) between EBV_{AI} and EBV_{AIII}, and EBV_{BI} and EBV_{BIII}, were also estimated. r_{II}, r_{III}, VAR_{II} and VAR_{III} were estimated for all the horses included in the genetic evaluations, for the 100 best stallions and for the French and Dutch stallions.

Means and SD of REL_{AI}, REL_{AII}, REL_{AIII}, REL_{BI}, REL_{BII} and REL_{BIII} were computed for all horses, for the French stallions and for the Dutch stallions.

Results and discussion

Descriptive statistics

Among the 98 French and 67 Dutch stallions associated with external information, only 97 French and 54 Dutch stallions had a non-zero REL_c. Furthermore, REL_c decreased at least by 57% compared with REL_E (Table 2). This decrease was expected because, first, the coefficients of determination of the conversion

Table 2 Coefficients of determination (R^2) of the conversion equations and Pearson correlation coefficients (r) between internal and external estimated breeding values. Means and standard deviations (SD) of non-zero external reliabilities (REL_E), non-zero reliabilities of a converted estimated breeding value (REL_c) and non-zero record equivalents free of contributions due to relationships (RE) as well as the number of foreign stallions (Nb) associated with non-zero REL_E, REL_c and RE

	French stallions	Dutch stallions
R^2	0.46	0.28
r	0.71	0.59
REL_E		
Nb	98	67
Mean	0.59	0.42
SD	0.19	0.18
REL_c		
Nb	97	54
Mean	0.27	0.07
SD	0.10	0.06
RE		
Nb	97	50
Mean	2.75	0.51
SD	1.44	0.52

equations and, second, the Pearson correlation coefficients between EBV_{AI} and EBV_E in the conversion equations were low to moderate (Table 2). Powell *et al.* (1994) surveyed different countries to determine expressions of REL_c associated with EBV_c in the context of dairy cattle. Several countries considered that REL_c were equal to REL_E. However, because REL_c must be integrated into a genetic evaluation, the variances of the sample regression equations must be taken into account as it was detailed by Goddard (1985). Therefore, lower variances of the sample regression equations and higher coefficients of determination, that is more accurate conversion equations, would be desirable to estimate reliable EBV_c (i.e. to obtain higher REL_c). Restrictions on data used for the comparisons of genetic evaluations between countries were formulated to improve the accuracy of conversions methods (e.g. Powell & Sieber 1992; Powell *et al.* 1994). However, owing to the low number of foreign stallions having both an EBV_{AI} and an EBV_E (i.e. 93 French stallions and 47 Dutch stallions) and the low average reliabilities associated with EBV_{AI} and EBV_E, most of the recommended restrictions could not be respected. Furthermore, the use of the Pearson correlation coefficients between EBV_{AI} and EBV_E for the French stallions and for the Dutch stallions led to an underestimation of REL_c because Calo *et al.* (1973) showed that genetic correlation coefficients for traits were higher than the corresponding Pearson correlation coefficients. This can be confirmed for the pair Belgium/France for which the genetic correlation coefficient was previously estimated between 0.76 and 0.88 (Ruhlmann *et al.* 2009b), while the Pearson correlation coefficient was equal to only 0.71 (Table 2). Therefore, it would be interesting to have estimates of the genetic correlation coefficients for the missing pair Belgium/the Netherlands because imprecise estimates of genetic correlation coefficients could lead to inexact REL_c (Calo *et al.* 1973; Powell *et al.* 1994).

As shown before, REL_E and the accuracy of the conversion equations have an effect on REL_c, but also on the improvement in the genetic evaluation. Indeed, a simplified system of mixed model equations ($\mathbf{Cs} = \mathbf{r}$) integrating external information (i.e. prior information) can be written as:

$$\left(\mathbf{C} + \mathbf{V}^{-1}\right)\mathbf{s} = \mathbf{r} + \mathbf{V}^{-1}\boldsymbol{\mu} \qquad (4)$$

where $\mathbf{C}$, $\mathbf{r}$ and $\mathbf{s}$ are the left-hand side, the right-hand side and the vector of solutions of the mixed model equations, respectively, $\boldsymbol{\mu}$ is the vector of mean prior information, and $\mathbf{V}$ is the prior variance matrix.

On the one hand, in the case of highly accurate prior information (i.e., in this case, high REL_c needed high REL_E and accurate conversion equations), the value of $\mathbf{V}$ will be close to zero, $\mathbf{V}^{-1}$ will be large, and therefore, the equations system (4) will tend to $\mathbf{V}^{-1}\mathbf{s} = \mathbf{V}^{-1}\boldsymbol{\mu}$ and the estimate of $\mathbf{s}$ to $\hat{\mathbf{s}} \approx \boldsymbol{\mu}$. On the other hand, in the case of low accurate prior information (i.e. in this case, low REL_c estimated from low REL_E and/or poorly accurate conversion equations), the prior information is non-informative, the value of $\mathbf{V}$ will be large, $\mathbf{V}^{-1}$ will tend to zero, and therefore, the equations system (4) will tend to $\mathbf{Cs} = \mathbf{r}$ and the estimate of $\mathbf{s}$ to $\hat{\mathbf{s}} \approx \mathbf{C}^{-1}\mathbf{r}$, similarly to a system that does not integrate prior information. Between these two extreme cases, $\hat{\mathbf{s}}$ can be considered as a weighted average of the combination of data and prior information (Klei *et al.* 1996).

With regard to RE (Table 2), the number for each stallion was a function of REL_c and of the relationships with other stallions. The low number of RE can be explained by the low-to-moderate REL_c but also by the low-to-moderate coefficients of determination of the conversion equations and Pearson correlation coefficients between EBV_{AI} and EBV_E. Furthermore, 4 Dutch stallions that were highly related to other ones were associated with RE equal to 0 after estimation of contributions due to relationships, while this was not the case for the French stallions. Hence, considering contributions due to relationships seems necessary to avoid double counting of external information.

Concerning the rankings of the horses, the rank correlation between EBV_{AI} and EBV_{BI} for all horses (Table 3) showed that the integration of external information for foreign stallions into the Belgian genetic evaluation influenced very slightly the ranking of the whole population. These modifications can

Table 3 Spearman rank correlation coefficients between breeding values (EBV) estimated by the Belgian genetic evaluation based on the complete data I (EBV_{AI}) and EBV estimated by a Bayesian[a] evaluation based on data I (EBV_{BI}), Spearman rank correlation coefficients between EBV_{AI} and external EBV (EBV_E), and Spearman rank correlation coefficients between EBV_{BI} and EBV_E for all horses, for the 100 best stallions, for the French stallions and for the Dutch stallions

		Spearman rank correlations		
Group of horses	Nb	EBV_{AI}/EBV_{BI}	EBV_{AI}/EBV_E	EBV_{BI}/EBV_E
All horses	101 382	>0.99	–	–
Best stallions	100	0.98	–	–
French stallions	98	0.87	0.69	0.90
Dutch stallions	67	0.95	0.61	0.73

[a]Bayesian: Belgian genetic evaluation integrating external information by a Bayesian approach.

be explained by the animal model: all the relationships between the foreign stallions and other horses were taken into account. These relationships caused an effect of foreign information on related horses through the foreign stallions. However, these modifications were small because external information was integrated only for about 0.2% of the horses. Concerning the foreign stallions, the integration of external information led to a change of their rankings according to the rank correlations between EBV_E and EBV_{AI} or EBV_{BI} (Table 3). The Belgian ranking of the foreign stallions was more similar to the ranking in their country of birth when the external information was integrated, as expected and shown by Quaas & Zhang (2006). However, because average quantity of French information was higher in terms of RE than the Dutch information (Table 2), the increase in the rank correlation coefficients between EBV_E and EBV_{AI} and between EBV_E and EBV_{BI} for the French stallions was higher than the corresponding increase for the Dutch ones, as according to the theory (Zhang *et al.* 2002; Legarra *et al.* 2007). Finally, because the Belgian ranking of foreign horses was more similar to the ranking in their country of birth and because 24 foreign stallions were considered in the group of the 100 best stallions, the rank correlation between EBV_{AI} and EBV_{BI} for the best stallions (Table 3) showed a change in their ranking of 2%. The reranking was mainly due to the reranking of the foreign stallions, but also due to the reranking of foreign stallions' relatives. The best stallions associated with external information gained four ranks on average in the ranking of the 100 best stallions, and a gain of 23 ranks was the largest reranking for a foreign best stallion. Integration of external information also led to a gain of 18 ranks for a stallion not associated with external information but related to several foreign stallions.

Model adequacy

Considering all horses, as well as the best stallions, the comparison of MB, MSEP and $r_{y:\hat{y}}$ did not show a reduction in the level of precision and accuracy of the Belgian model when external information was included because MB were close to 0% and MSEP were equal for the two genetic evaluations. Similar values were also estimated for $r_{y:\hat{y}}$ associated with the evaluations AI and BI (Table 4).

With regard to the French stallions, the bias of 1.32% for the evaluation AI was reduced when external information was integrated (Table 4). However, it was not confirmed by the associated MSEP, and thereby, the adequacy of the model for the French stallions was not improved, but also not diminished, by the integration of external information.

Table 4 Number of performances (Nb) associated with the Belgian trait for all horses, the 100 best stallions, the French stallions and the Dutch stallions. Mean bias (MB), mean square errors of prediction (MSEP) and Pearson correlation coefficients between observed and estimated performances ($r_{y:\hat{y}}$) for the evaluation AI[a] and the evaluation BI[b] applying all the performances, performances associated with the 100 best stallions, French performances and Dutch performances for the Belgian trait. MB and MSEP are expressed as a percentage of the average performance of all the performances and performances associated with the 100 best stallions, the French and Dutch stallions, respectively

	All horses		Best stallions		French stallions		Dutch stallions	
	AI	BI	AI	BI	AI	BI	AI	BI
Nb	350 907		2749		1322		414	
MB	2.00e-3	−4.00e-8	0.06	0.06	1.32	1.06	1.49	1.36
MSEP	2.74	2.74	0.67	0.67	2.26	2.26	2.38	2.37
$r_{y:\hat{y}}$	0.50	0.50	0.26	0.26	0.36	0.36	0.49	0.50

[a]Belgian genetic evaluation based on the complete data I.
[b]Bayesian evaluation based on the complete data I.

With regard to the Dutch stallions, it seems that the integration of Dutch information improved the adequacy of the model for their genetic evaluation because MB and MSEP of the evaluation BI were slightly lower than MB and MSEP of the evaluation AI (Table 4). Furthermore, $r_{y:\hat{y}}$ of the evaluation BI was slightly higher than the one of the evaluation AI.

Finally, the integration of external information did not diminish the adequacy of the Belgian model for all horses, for the best stallions and for foreign stallions. However, owing to the low amount of external information (only for about 0.2% of the horses) and to the low average REL_E, the model adequacy was not improved or only weakly for the Dutch stallions.

Predictive ability

Considering all horses, the evaluation BII had a similar or slightly worse predictive ability than the evaluation AII according to r_{II} and VAR_{II} (Table 5). This low difference between predictive abilities was expected because external information was integrated into the evaluation for only 165 foreign stallions, whereas 44 755 horses have performances among the 101 382 horses in the pedigree. However, if the evaluations were based on subset III, i.e. if only performances of the foreign stallions were assumed to be missing, there was a slight advantage for the evaluation B, especially in terms of VAR_{II}. Furthermore, r_{III} close to 1 can be explained by the fact that only 0.27% of all the performances were assigned to missing values in

Table 5 Pearson correlation coefficients (r_{II}) and variances of differences (VAR_{II}) between EBV_{AI}[a] and EBV_{AII}, and EBV_{BI} and EBV_{BII}, and Pearson correlation coefficients (r_{III}) and variances of differences (VAR_{III}) between EBV_{AI} and EBV_{AIII}, and EBV_{BI} and EBV_{BIII} for all horses, for the 100 best stallions, for the French stallions and the Dutch stallions

		Data sets			
		I-II		I-III	
Group of horses	Genetic evaluation	r_{II}	VAR_{II} ($\times 10^{-3}$)	r_{III}	VAR_{III} ($\times 10^{-3}$)
All horses	A	0.89	4.49	>0.99	0.10
	B	0.89	4.58	>0.99	0.06
Best stallions	A	0.80	9.34	0.98	1.18
	B	0.82	10.24	0.99	1.02
French stallions	A	0.96	3.37	0.89	8.05
	B	0.99	2.23	0.98	3.48
Dutch stallions	A	0.93	7.12	0.95	5.73
	B	0.95	6.58	0.97	4.31

[a]EBV_{ij}: Estimated breeding values where *i* refers to the type of the genetic evaluation (i.e. A = Belgian genetic evaluation and B = Bayesian evaluation) and *j* refers to the used data (i.e. I = complete data, II = data for which all performances after 31 December 2005 were assumed to be missing, and III = data for which all French and Dutch stallions' own performances were assumed to be missing).

subset III (Table 5), hence their limited overall influence.

Considering the French stallions, r_{II}, VAR_{II}, r_{III} and VAR_{III} showed that the predictive ability of breeding values was improved when French information was integrated (Table 5). The high r_{II} (0.96) between EBV_{AI} and EBV_{AII} also showed that performances after 2005 influenced less the genetic evaluation of the French stallions compared with the genetic evaluation for all horses. It can be explained by the facts that the French stallions had few performances after 2005 and few relationships with Belgian horses competing after 2005. Regarding the evaluations based on the subset III, there were an increase of r_{III} around 10% and a reduction of VAR_{III} of 57% for the evaluations B in comparison with r_{III} and VAR_{III} of the evaluations A, respectively. Following these results, integrated external information was almost equivalent to the French stallions' own performances. These results suggest that the integration of external information could be very interesting in the case of imported stallions having no or few own Belgian performances (e.g. young imported stallions or confirmed foreign stallions imported through their semen).

Considering the Dutch stallions, the lower r_{II} and higher VAR_{II} compared with the r_{III} and VAR_{III}, respectively, for the evaluations A led to higher influence of performances recorded after 2005 for their genetic evaluation compared with their own records (Table 5). The low number of Dutch stallions with performances, that is <60% of the Dutch stallions, can explain this observation. This is also explained by the high r_{III} between EBV_{AI} and EBV_{AIII} (0.95). However, following r_{III} and VAR_{III} (Table 5), the predictive ability was improved when external information was integrated into the evaluation based on the subset III. Regarding the evaluations based on the subset II, the increase in r_{II} and the slight improvement in VAR_{II} also confirmed the improvement in the predictive ability when Dutch information was integrated into the Belgian genetic evaluation.

Reliabilities

Means and SD of REL were calculated for all the genetic evaluations (Table 6). All the genetic evaluations had a minimum and a maximum REL equal to 0.00 and 0.99, respectively. As expected since external information was only integrated for 165 stallions, the integration of external information did not influence on average the genetic evaluation for all the horses.

It is noted that the accuracy of the procedure applied to estimate REL_{BI}, REL_{BII} and REL_{BIII} depended on the accuracy of REL_c and therefore on accuracies of the conversion equations and of the genetic correlation coefficients for traits between Belgium and the exporting countries. Because the genetic correlation coefficient was unknown for the pair Belgium/the Netherlands, they were estimated from EBV_{AI} and EBV_E of the foreign stallions and errors linked with the estimation of REL_c were also

Table 6 Means and standard deviations (SD) of reliabilities $(REL)_{AI}$[a], REL_{AII}, REL_{AIII}, REL_{BI}, REL_{BII} and REL_{BIII} for all horses, for the French stallions and for the Dutch stallions

		Data sets					
		I		II		III	
Group of horses	Genetic evaluation	Mean	SD	Mean	SD	Mean	SD
All horses	A	0.21	0.17	0.17	0.16	0.21	0.17
	B	0.21	0.17	0.17	0.16	0.21	0.17
French stallions	A	0.58	0.23	0.54	0.24	0.52	0.26
	B	0.61	0.20	0.58	0.20	0.57	0.21
Dutch stallions	A	0.51	0.26	0.46	0.26	0.47	0.27
	B	0.52	0.25	0.47	0.25	0.49	0.26

[a]REL_{ij}: Reliabilities of estimated breeding values where *i* refers to the type of genetic evaluation (i.e. A = Belgian genetic evaluation and B = Bayesian evaluation) and *j* refers to the used data (i.e. I = complete data, II = data for which all performances after 31 December 2005 were assumed to be missing, and III = data for which all French and Dutch stallions' own performances were assumed to be missing).

introduced into the estimation of REL associated with EBV_B. Again, these errors show the need to estimate genetic correlation coefficients for traits between countries to perform an unbiased genetic evaluation.

Regarding to the Dutch and French stallions, the average REL_{BI} were improved compared with the average REL_{AI}. Additionally to the influence of the imprecision because of accuracy of regression equations and of the unknown genetic correlations as explained previously, the improvement in average REL also depended on the range of REL_c and thereby on the range of REL_E. Zhang *et al.* (2002) also concluded that the amount of improvement depends on REL_E for a simulation for beef cattle.

For the particular case of the French stallions, the integration of external information led to an increase (5%) of the average REL for the genetic evaluations based on data I (Table 6). Then, the average REL_A decreased when performances were assumed to be missing, as it was expected. This reduction was higher for REL_{AIII}. The own performances were more informative for the French stallions than the contemporary horses' performances recorded after 2005, as already observed. Nevertheless, the integration of French information led to an average REL_{BII} and REL_{BIII} equal or close to the average REL_{AI}. So, external information was on average at least equivalent to the Belgian performances related to these stallions.

Regarding to the Dutch stallions, average REL_{AII} and REL_{AIII} confirmed the higher influence of performances after 2005 for their genetic evaluation compared with their own performances (Table 6). Furthermore, average REL_{BII} and REL_{BIII} compared with average REL_{AI} showed that the integration of Dutch information was not totally equivalent to the missing information, despite it was not insignificant in each case. Indeed, there was an increase of the average REL of at least 2% when Dutch information was integrated into the Belgian evaluation.

Conclusion

According to these results, external information, that is foreign EBV and their associated REL, for French and Dutch stallions was partially to totally equivalent to 4 years of contemporary horses' performances or to their own performances in show jumping. Its integration did not diminish the adequacy of the Belgian model for all horses, as well as for foreign stallions. It also improved the predictive ability and the accuracy of EBV for the foreign stallions. The resulting Belgian ranking of the foreign stallions was more similar to their foreign ranking according to their country of birth, according to the Spearman rank correlations. All these results showed the interest to integrate external information into the Belgian genetic evaluation for show jumping, especially for imported stallions having no or few Belgian performances (e.g. young imported stallions or confirmed foreign stallions imported through their semen). However, estimates of genetic correlations for traits among countries, as well as accurate conversion equations, are needed for a more accurate Belgian Bayesian evaluation.

Acknowledgements

Jérémie Vandenplas, as a Research Fellow, acknowledges the support of the National Fund for Scientific Research (Brussels, Belgium) for this position and for the additional grants 2.4507.02F(2) and F4552.05 (2.4.623.08.F). The authors acknowledge Anne Ricard (Station de Génétique Quantitative et Appliquée, Institut National de la Recherche Agronomique, France) and Danielle Arts (Royal Dutch Sport Horse, the Netherlands) to provide the external information to realize this research. The authors are also grateful to University of Liege (SEGI facility) for the use of NIC3 supercomputer facilities. The authors thank the anonymous reviewers for the useful comments.

References

Calo L.L., McDowell R.E., VanVleck L.D., Miller P.D. (1973) Genetic aspects of beef production among Holstien-Friesians pedigree selected for milk production. *J. Anim. Sci.*, **37**, 676–682.

Gengler N., Vanderick S. (2008) Bayesian inclusion of external evaluations into national evaluation system: application to milk production traits. *Interbull Bull.*, **38**, 70–74.

Goddard M. (1985) A method of comparing sires evaluated in different countries. *Livest. Prod. Sci.*, **13**, 321–331.

Interstallion (2005) Recommendations on Choice of Scale and Reference Population for Publication of Breeding Values in Sport Horse Breeding. Interstallion, Leuven, Belgium (available at: http://www.biw.kuleuven.be/genlog/livgen/docs/publicationscale.pdf; last accessed 19 March 2012).

Janssens S., Buys N. (2008) Genetische Index Voor Springen – GSIn2008. Katholieke Universiteit Leuven, Leuven, Belgium (available at: http://www.biw.kuleuven.be/genlog/livgen/chgs_ebv/GenetischeSpringindex_2008.pdf; last accessed 19 March 2012).

Janssens S., Geysen D., Vandepitte W. (1997) Genetic parameters for show jumping in Belgian sport horses. In: Proceedings of the 48th Annual Meeting of the European Association for Animal Production. Vienna (Austria), 25-28 August 1997. 5 pp.

Janssens S., Buys N., Vandepitte W. (2007) Sport status and the genetic evaluation for show jumping in Belgian sport horses. In: Proceedings of the 58th Annual Meeting of the European Association for Animal Production. Dublin (Ireland), 26-29 August 2007, 26–29

Klei L., Quaas R.L., Pollak E.J., Cunningham B.E. (1996) Multiple breed evaluation. In: Proceedings of Beef Improvement Federation 28th Annual Research Symposium and Annual Meeting. Birmingham, USA, pp. 93–105.

Koenen E.P.C. (2002) Genetic evaluations for competition traits of warmblood sport horses. Note presented at World Breeding Federation for Sport Horses seminar. Budapest (Hungary). 5 November 2002.

Koenen E.P.C., Aldridge L.I. (2002) Testing and genetic evaluation of sport horses in an international perspective. In: Proceedings of 7th World Congress Applied to Livestock Production. Montpellier (France), 19-23 August 2002, pp. 19–23.

Koenen E.P.C., Aldridge L.I., Philipsson J. (2004) An overview of breeding objectives for warmblood sport horses. *Livest. Prod. Sci.*, **88**, 77–84.

Legarra A., Bertrand J.K., Stradel T., Sanchez J.P., Misztal I. (2007) Multi-breed genetic evaluation in a Gelbvieh population. *J. Anim. Breed. Genet.*, **124**, 286–295.

Misztal I. (2012) BLUPF90 Family of Programs. University of Georgia, Athens, GA, USA (available at http://nce.ads.uga.edu/~ignacy/newprograms.html; last accessed at 21 March 2012).

Powell R.L., Sieber M. (1992) Direct and indirect conversion of bull evaluations for yield traits between countries. *J. Dairy Sci.*, **75**, 1138–1146.

Powell R.L., Wiggans G.R., VanRaden P.M. (1994) Factors affecting calculation and use of conversion equations for genetic merit of dairy bulls. *J. Dairy Sci.*, **77**, 2679–2686.

Quaas R.L., Zhang Z. (2006) Multiple-breed genetic evaluation in the US beef cattle context: methodology. In: Proceedings of 8th World Congress on Genetics Applied to Livestock Production. Belo Horizonte (Brazil). 13-18 August 2006, pp. 13–28.

Ricard A. (1997) Breeding evaluations and breeding programs in France. In: Proceedings of the 48th annual meeting of the European Association for Animal Production. Vienna (Austria), 25-28 August 1997, 9 pp.

Ruhlmann C., Bruns E., Fraehr E., Philipsson J., Janssens S., Quinn K., Thorén Hellsten E., Ricard A. (2009a) Genetic connectedness between seven European countries for performance in jumping competitions of warmblood riding horses. *Livest. Sci.*, **120**, 75–86.

Ruhlmann C., Janssens S., Philipsson J., Thorén Hellsten E., Crolly H., Quinn K., Manfredi E., Ricard A. (2009b) Genetic correlations between horse show jumping competition traits in five European countries. *Livest. Sci.*, **122**, 234–240.

Tavernier A. (1991) Genetic evaluation of horses based on ranks in competitions. *Genet. Sel. Evol.*, **23**, 159–173.

Tedeschi L.O. (2006) Assessment of the adequacy of mathematical models. *Agric. Syst.*, **86**, 225–247.

Vandenplas J., Gengler N. (2012) Comparison and improvements of different Bayesian procedures to integrate external information into genetic evaluations. *J. Dairy Sci.*, **95**, 1513–1526.

van Veldhuizen A.E. (1997) Breeding value estimation for riding horses in Netherlands. In: Proceedings of the 48th annual meeting of the European Association for Animal Production. Vienna (Austria), 25-28 August 1997, 8 pp.

Zhang Z.W., Quaas R.L., Pollak E.J. (2002) Simulation study on the effects of incorporating external genetic evaluations results. In: Proceedings of 7th World Congress on Genetics Applied to Livestock Production. Montpellier (France), 19-23 August 2002. Communication N°20-14 in CD.