J. Dairy Sci. 96:3272–3284
http://dx.doi.org/10.3168/jds.2011-4987

Effects of a national genomic preselection on the international genetic evaluations

Clotilde Patry,*†[1] Hossein Jorjani,‡ and Vincent Ducrocq*
*UMR 1313 Génétique Animale et Biologie Intégrative, Institut National de la Recherche Agronomique, 78 352 Jouy-en-Josas, France
†Union nationale des Coopératives d'Elevage et d'Insémination Animale, 149 rue de Bercy, 75 595 Paris Cédex 12, France
‡Interbull Centre, Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Box 7023, S-75007 Uppsala, Sweden

ABSTRACT

Genomic preselection of young bulls is now widely implemented in dairy breeding schemes, especially in the Holstein breed. However, if this step is not accounted for in genetic evaluation models, the national breeding values of bulls retained by a genomic preselection and of their progeny are estimated with bias. It follows that countries participating in international genetic evaluations will provide a selected and possibly biased set of data to the Interbull Centre (Swedish University of Agricultural Sciences, Uppsala, Sweden). The objective of the study was to show evidence of bias at the international level due to a genomic preselection step in national breeding schemes. The consequence of a genomic preselection for the international evaluations (i.e., using selected and biased national estimated breeding values) was simulated using actual national estimated breeding values as a proxy for genomically enhanced breeding values. Data were provided for 3 countries with a large population of Holstein bulls. International breeding values from simulated scenarios were compared with international breeding values using all available data, assumed to be complete and unbiased. Bias was measured among young bulls retained by a genomic preselection and their contemporaries in other countries. The results were analyzed by traits measured within each country and by country of origin of the young bulls. It turned out that sending preselected data, though based on genomic information, created bias in international evaluations, penalizing young bulls from the country sending the incorrect data. It also had an effect on the young bulls from the other countries. Sending biased data further affected the quality of international evaluations. This study underlines the importance of accounting for genomic preselection at the national level first. Moreover, submitting all available data appeared essential to maintain the quality of the international genetic evaluations after implementation of a genomic preselection step.

Key words: Interbull, international genetic evaluation, genomic preselection, selection bias

INTRODUCTION

Breeding strategies in dairy cattle are being transformed by the emergence of genetic evaluation tools combining new molecular technologies and advanced statistical analyses. Genomic selection is developing fast and propagating worldwide. In 2009, only a few countries computed genomically enhanced breeding values (**GEBV**) in the Holstein breed only; in 2011, data from 5 dairy breeds and 13 countries were provided for validation of national genomic evaluations at the Interbull Centre (Swedish University of Agricultural Sciences, Uppsala, Sweden; Loberg et al., 2011).

Since 1994, the Interbull Centre has been in charge of the routine international genetic evaluations that facilitate comparisons between bull breeding values across countries and promote international genetic exchanges. As proposed by Schaeffer (1994), international EBV (**I-EBV**) are computed from national EBV (**N-EBV**) using BLUP applied to a mixed linear sire model for multiple traits. The method is commonly known as multiple-trait across country evaluation (**MACE**). For the same selection objective, each trait evaluated in different countries can have different levels of heritability so that traits are considered to be different but correlated. It follows that international evaluations benefit from an increased amount of performances for bulls having progeny in several countries. Each bull gets a revised breeding value expressed in the unit of trait used in each participating country. Thus, bulls from all countries can be ranked according to the same base and scale in each country.

The recent changes in national evaluation systems, and the new breeding strategies due to genomic advances, have to be considered in international genetic evaluations to keep providing international comparisons of dairy bulls across a large number of countries. The

---

Received September 27, 2011.
Accepted January 31, 2013.
[1] Corresponding author: clotilde.patry@jouy.inra.fr

first challenge for the Interbull Centre is to develop a new evaluation methodology using genomic information. The second is to maintain the international genetic evaluations as they are currently delivered, without genomic information and only based on pedigree and phenotypes. The international evaluations is still useful for all of the countries participating in international genetic evaluations, for those not computing genomic evaluations (i.e., 15 out of the 28 countries), and for those with genomic evaluations. In fact, international evaluations provide a tool for fair international comparison that is not yet available based on genomic information. Moreover, I-EBV are not only required to include foreign information in classical national evaluations, but also in the genomic evaluation system. To make genomic evaluations more reliable, the reference population might include phenotyped and genotyped animals from foreign countries.

However, the implementation of a genomic preselection (**GPS**) step for young bulls (**YB**) at the national level may threaten the correctness of international genetic evaluations. Two reasons for this exist. The first reason is that, in the near future, each country implementing GPS might send to the Interbull Centre biased data for bulls retained by GPS. Patry and Ducrocq (2011 b) showed by simulations that once the YB selected from genomic information have daughters, their BLUP solutions are, on average, underestimated when compared with the simulated true breeding values. The risk is that such bias could be propagated at the international level to related animals and correlated traits. Biased N-EBV may be avoided if all genotyped and nongenotyped animals are included in the BLUP evaluations. Two approaches might be implemented in national evaluations. The first one is based on single-step evaluations where the relationship matrix is modified to include all animals, genotyped or phenotyped. (Legarra et al., 2009, Misztal et al., 2009, Aguilar et al., 2010, Christensen and Lund, 2010). The second type of approach is based on multi-step evaluations where GEBV are computed first and then included as weighted-deregressed performances (Ducrocq and Liu, 2009, Patry and Ducrocq, 2011a) besides the actual ones in BLUP evaluations. The GEBV can also be considered as a correlated trait (Mäntysaari and Strandén, 2010, Stoop et al., 2011) for all genotyped candidates. All of these approaches better describe the process of GPS and avoid EBV from being biased; however, all BLUP solutions will include genomic information from that moment.

The second reason is that countries might send to the Interbull Centre performances for selected bulls only and BLUP solutions for culled candidates would be missing. Selection based on genomic information was shown to affect the distribution of the Mendelian sampling (**MS**) term (Patry and Ducrocq, 2011b) and violate the BLUP hypotheses assuming that the MS term averages to zero. It is thus feared that MACE results computed from a selected subpopulation (e.g., selected YB) could also be biased (Henderson, 1975) too.

Because of those 2 reasons, it was justified to wonder the effect of selected N-EBV on international evaluations. The objective of this study was to describe and assess the effect of GPS on international evaluations due to (1) the propagation of bias from the national to the international level and (2) the creation of a bias in MACE solutions while using selected N-EBV. Three types of GPS strategies could be implemented by the participating countries and were considered for this study. Countries implementing GPS might deliver biased and selected N-EBV to the Interbull Centre (strategy 1). Countries might account for GPS in their national evaluation model and prevent N-EBV from being biased. Therefore, countries might send unbiased N-EBV for either only selected YB (strategy 2) or for all YB (strategy 3) to the Interbull Centre. The bias (i.e., the systematic under- or overestimation of I-EBV) was therefore measured under the alternative strategies, which were simulated based on real data. Consequences on rankings for international comparisons were also analyzed. Solutions to implement at the national and international levels and preventing biased I-EBV were put forward.

MATERIALS AND METHODS

Material

To participate in the international evaluations, each country provided the Interbull Centre N-EBV information on performance and their effective daughter contribution as weight. The national data sets required to run the August 2010 routine evaluation were used to simulate GPS and to assess its effect on the international level. In the present study, we focused on one production trait (i.e., protein yield in the Holstein breed), restricted to animals measured and evaluated in only 3 large countries, denoted hereafter as "country A," "country B," and "country C." A total of 57,688 bulls were considered out of the about 116,000 bulls with protein yield evaluations from 27 countries. The 3 countries in this study delivered national genetic evaluations from BLUP applied to a single-trait animal model. Note that country A used a random regression test day model, whereas countries B and C used a repeatability model.

Genetic parameters are displayed in Table 1. Heritability values were sent by the national evaluation