Use of international clinical mastitis data as independent trait in the US evaluation system

R.R. Mota[1], E. Nicolazzi[1], J.H. Megonigal[1], K.L. Gaddis[1], J.W. Durr[1], P.M. VanRaden[2]
[1] CDCB - Council on Dairy Cattle Breeding, Bowie, MD, USA
[2] Animal Genomics and Improvement Laboratory, ARS, USDA, Beltsville, MD, USA

Abstract

Since its first publication in Holsteins (**HOL**) in April 2018, and later inclusion of Jerseys (**JER**) in April 2020, US bull evaluations for mastitis resistance (MAS) have been exchanged with Interbull participating countries. Foreign phenotypes for MAS have been used since then to enhance the domestic reference population. Prior to April 2021, countries had the choice of exchanging pure clinical mastitis, somatic cell score, or a combination of methodologies where clinical/subclinical mastitis or a multi-trait approach using multiple sources of information. Hence, only IDs of bulls coming from certain countries with similar trait definitions and if the country of most daughters does not send only SCS were being used in the US. Effective in the April 2021 routine run, Interbull introduced a new trait group named SNP training for clinical mastitis to better estimate SNP effects specifically for clinical mastitis (**CMA**). The new edits in the US were validated in January 2021 in a full test run. In the April run, genotyped bulls with an international evaluation from the other participating countries used CMA results, whereas the previous trait (called hereafter as the MAS), that combines mastitis from some countries and correlated SCS from others continued to be used for non-genotyped animals. In order to evaluate the impact on the US evaluations, this study aimed to compare predicted transmitting abilities (**PTA**) and reliabilities (**REL**) between December 2020 routine run (**2012r**) and January 2021 test run (**2101t**) and April routine run (**2104r**), for both HOL and JER breeds, but now taking into account the aforementioned criteria. Descriptive statistics, Pearson and Spearman correlations (r_g) as well as regression coefficients (b_1) by predicting MAS on CMA to measure potential biases, were calculated. The 2104r PTA means were slightly smaller for HOL and relatively smaller for JER in all scenarios. Pearson and Spearman correlations were always higher than 0.90 in all scenarios for both breeds, no matter the evaluation set of comparison. These results demonstrate a practically null impact on the US evaluations. By comparing 2012r and 2101t, the b1 values were, in general, close to 1 (range 0.98-1.06). On the other hand, a bit more bias can be seen by comparing 2012r with 2104r and 2101t with 2104r. These results may have been partially due to >30,000 corrected phenotypes received in the U.S. April evaluation. Our results suggest minor impact and genetic progress improvement enabled by the implementation of the new international trait SNP training for clinical mastitis. On the other hand, identifying which countries or individual bulls had direct MAS or only correlated SCS was previously difficult to automate correctly at the national level. Therefore, it is of most importance that Interbull continues to keep track of CMA countries, so several benefits such as gains in REL for bulls with many CMA daughters, would be achieved.

Key words: bias, correlations, evaluation exchange, participating countries, reliabilities, udder health

Introduction

The mastitis trait is one of the most common and costly health events that impact dairy herds. Defined as an infection disease that causes mammary gland inflammation, mastitis has been reported as a low heritable trait (~3%), which selection evidence firstly reported about four decades ago. However, with the introduction of genomics in 2009 (Van Raden et al., 2009) a faster progress can already be seen not only for mastitis but also for all low heritable traits, in general. This trait has unfavorable correlations with milk yield (range 0.24-0.55; Heringstad et al., 2000). This is mainly due to the fact that most countries report mastitis resistance (**MAS**) as a trait instead of clinical mastitis (**CMA**). In other words, MAS and CMA are usually the same trait, but Interbull uses MAS for the trait group that allows SCS as a substitute and CMA for the trait group that excludes SCS.

The evaluation of health events was first introduced in the United States (**US**) in April 2018 for Holsteins (**HOL**; Gaddis et al., 2020) and in April 2020 for Jerseys (**JER**; Jensen et al., 2019). For both breeds CMA predicted transmitting abilities (**PTA**) are presented as percentage points from resistance events above or below breed average. The number of records in routine runs has increased quickly from 1.8 million records in April 2018 to 5.1 million in April 2021, for both breeds combined. So far, the HOL database contains 4.2 million records whereas JER contains 582K.

Since its first publication in HOL and JER, US bull evaluations for MAS have been exchanged with Interbull participating countries. Foreign phenotypes for MAS have been used since then to enhance the domestic reference population. Prior to routine run in April 2021, countries had the choice of exchanging pure clinical mastitis (as the case in the US), somatic cell score (**SCS**) for countries without a health evaluation, or a combination of methodologies where clinical/subclinical mastitis or a multi-trait approach using multiple sources of information are used. Hence, only IDs of bulls coming from certain countries with similar trait definitions, e.g., BEL, CAN, CHE, DEU, DFS, FRA, GBR and NLD, and only if the country of most daughters did not send only SCS, e.g., USA bulls with daughters only in JPN were not considered and were not used in the USA MAS evaluations. The issue here is that if single-trait genomic methods are used the many historical records for a correlated trait (**SCS**) may outweigh the direct data for MAS. Then we should separate the correlated from direct data or use multi-trait methods to handle both traits properly. The concern about using SCS as a correlated, substitute trait for MAS reflects a much larger problem with evaluating new traits and trying to also use correlated traits.

Effective in the April 2021 routine run, Interbull introduced a new trait group named SNP training for clinical mastitis to better estimate SNP effects specifically for CMA. The new edits in the USA were validated in January 2021 in a full test run. Therefore, CMA values from April 2021 were used for genotyped bulls with an international evaluation from the other participating countries, whereas the previous mastitis trait was used for non-genotyped bulls. Most of the participating countries to the new trait have been already accounted for, thus, only a minor impact to the USA evaluations was expected. To confirm this hypothesis, this study aimed to compare predicted PTA and reliabilities (**REL**) values, for both HOL and JER breeds, and among the evaluations: December 2020 routine run (**2012r**), January 2021 test run (**2101t**) and April routine run (**2104r**), taking into account the aforementioned criteria.

Material & Methods

The data used in this study were Holstein and Jersey MACE values provided by the Interbull Centre, Uppsala, Sweden (https://interbull.org/ib/interbullactivities).

To infer about the impact of the inclusion of the new group trait, we investigated all bulls having MACE values for at least two of the runs. The adopted criteria was that genotyped bulls with

an international evaluation from the other participating countries, will be using the CMA results, whereas non genotyped animals will continue to use the previous MAS, that combines mastitis from some countries and correlated SCS from others.

Pearson and Spearman correlations were calculated as follows:

$$rg = \frac{\sigma ab}{\sqrt{\sigma^2 a * \sigma^2 b}},$$

where r_g is the genetic correlation, and a and b can be either the investigated runs (2012r, 2101t or 2104r). Regression coefficients by predicting the old trait MAS on the new CMA trait to measure potential biases by adding the new clinical mastitis trait were also calculated. The statistical analyses were done by using the R software (R Development Core Team, 2021, https://www.R-project.org/).

Finally, the investigation was conducted among the three evaluations in four different scenarios and by breed: s1- all bulls in common; s2- genotyped bulls in common; s3- high reliable genotyped bulls in common (REL > 95%); s4- non genotyped bulls in common. The total of bulls in common for each scenario by breed is presented in Table 1.

Results & Discussion

The 2104r PTA means were slightly smaller for HOL and relatively smaller for JER in all scenarios (Table 2). The April run also had higher standard deviations compared to 2012r and 2101t. These values can be explained by the fact that 2104r has four more months of data. CDCB received a great amount of data in April coming with more than 30k corrections which could have led to more variation.

Pearson (Figures 1, 2 and 3) and Spearman (results not shown) correlations were positive and high in all scenarios, no matter the evaluation set of comparison. These values were always higher than 0.90 for HOL and JER. These results demonstrate a practically null impact on the US evaluations since neither re-ranking nor high variations were observed, especially for high reliable bulls.

It can be noted in Figures 1 and 2 for the HOL breed, a set of animals that deviate from the others. For these animals, their 2012r PTA is constant whereas 2101t and 2104r values are not. The explanation for this is that these genotyped animals are from the Simmental breed that sometimes received two evaluations, as HOL and as Simmental themselves. In CDCB, the value to be used is the value with higher reliability.

Table 1. – Number of bulls in common between evaluations by breed in four scenarios: all bulls (s1), genotyped bulls (s2), high reliable bulls (REL > 95%; s3) and non-genotyped bulls (s4)

	HOL		
Scenario / run	**2012r vs. 2101t**	**2012r vs. 2104r**	**2101t vs. 2104r**
s1	69,285	67,742	96,762
s2	12,307	10,401	11,422
s3	425	439	419
s4	56,978	57,341	85,340
	JER		
s1	3,932	2,648	7,173
s2	1,955	672	712
s3	67	68	68
s4	1,977	1,976	6,461

2012r: December 2020 routine run; 2101t: January 2021 test run; 2104r: April routine run

Table 2. – PTA means (standard deviations) between evaluations by breed in four scenarios: all bulls (s1), genotyped bulls (s2), high reliable bulls (REL > 95%; s3) and non-genotyped bulls (s4)

	HOL		
	2012r vs. 2101t	**2012r vs. 2104r**	**2101t vs. 2104r**
s1	2.81 (23.07) vs. 2.81 (22.72)	3.08 (23.10) vs. 2.78 (23.62)	3.34 (22.29) vs. 3.15 (23.13)
s2	-2.82 (17.22) vs. -2.92 (16.73)	-1.94 (16.99) vs. -2.56 (17.87)	-1.85 (16.45) vs. -2.28 (17.80)
s3	-0.98 (18.83) vs. -1.01 (19.01)	-0.87 (18.71) vs. -1.50 (20.48)	-0.80 (18.94) vs. -1.35 (20.45)
s4	4.03 (23.97) vs. 4.05 (23.64)	3.99 (23.93) vs. 3.74 (24.39)	4.04 (22.87) vs. 3.88 (23.65)
	JER		
s1	8.65 (16.96) vs. 7.28 (16.40)	7.85 (18.14) vs. 5.46 (17.94)	5.65 (15.59) vs. 3.39 (16.38)
s2	6.44 (16.93) vs. 4.45 (16.74)	-0.92 (19.32) vs. -1.69 (19.75)	-1.35 (19.38) vs. -2.26 (19.99)
s3	0.30 (18.78) vs. 0.30 (19.18)	0.34 (18.65) vs. -0.29 (19.74)	0.12 (19.09) vs. -0.60 (19.85)
s4	10.84 (16.71) vs. 9.89 (15.63)	10.83 (16.71) vs. 7.89 (16.60)	6.42 (14.91) vs. 4.01 (15.81)

2012r: December 2020 routine run; 2101t: January 2021 test run; 2104r: April routine run

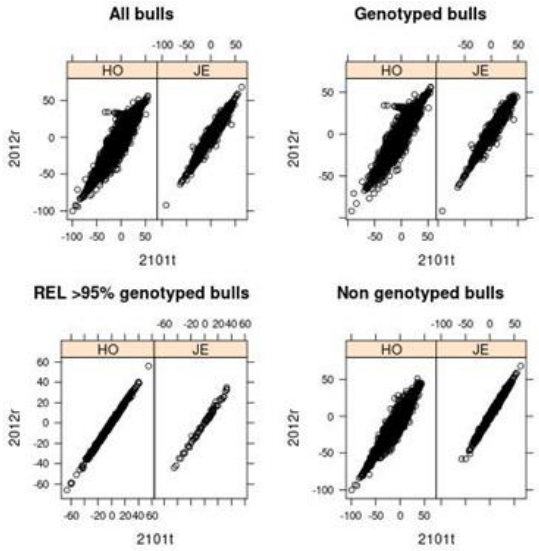

Figure 1. Pearson correlations between the December 2020 routine run (2012r) and January 2021 test run (2101t) evaluations by breed in four different scenarios: all bulls (s1), genotyped bulls (s2), high reliable bulls (REL > 95%; s3) and non-genotyped bulls (s4).

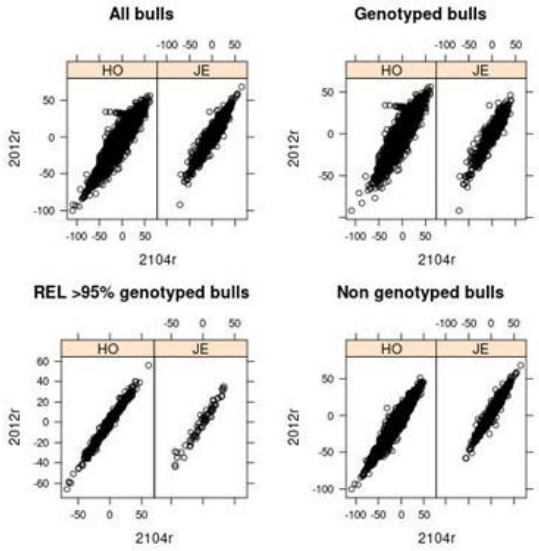

Figure 2. Pearson correlations between the December 2020 routine run (2012r) and April 2021 routine run (2104r) evaluations by breed in four different scenarios: all bulls (s1), genotyped bulls (s2), high reliable bulls (REL > 95%; s3) and non-genotyped bulls (s4).

By comparing 2012r and 2101t, the b1 values were, in general, close to 1 (range 0.98-1.06), with the highest bias (1.06) observed for non-genotyped JER animals (s4), where animals kept their combined MAS values (Table 3). On

the other hand, a bit more bias can be seen by comparing 2012r with 2104r and 2101t with 2104r. These results were somehow expected, due to the fact that in the April evaluation more than 30k corrections were applied.

However, less bias were observed by comparing evaluations when the CMA is already accounted for, i.e., 2101t and 2104r. For this set of comparison, the b1 value increased for high REL animals from 0.90 to 0.92 compared to 2012r vs 2104r. Also for the genotyped animals the b1 increased slightly from 0.89 to 0.90.

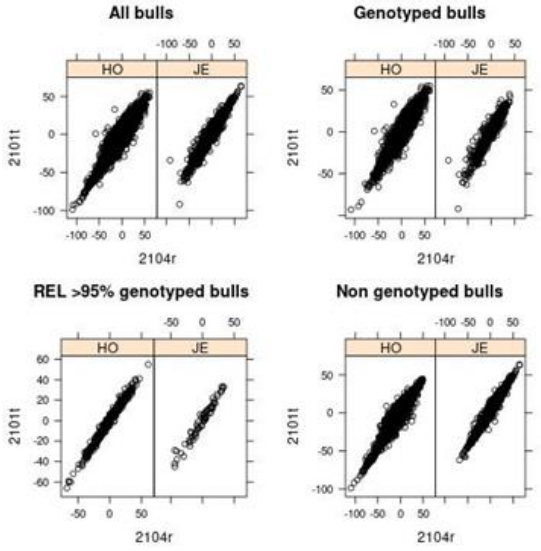

Figure 3. Pearson correlations between the January 2021 test run (2101t) and April 2021 routine run (2104r) evaluations by breed in four different scenarios: all bulls (s1), genotyped bulls (s2), high reliable bulls (REL > 95%; s3) and non-genotyped bulls (s4).

Table 3. – Predicted bias between evaluations by breed in four scenarios: all bulls (s1), genotyped bulls (s2), high reliable bulls (REL > 95%; s3) and non-genotyped bulls (s4)

	HOL		
Scenario / run	**2012r vs. 2101t**	**2012r vs. 2104r**	**2101t vs. 2104r**
s1	1.01	0.97	0.96
s2	0.99	0.89	0.90
s3	0.99	0.90	0.92
s4	1.01	0.98	0.96
	JER		
s1	1.02	0.98	0.94
s2	0.98	0.92	0.92
s3	0.98	0.92	0.94
s4	1.06	0.99	0.93

2012r: December 2020 routine run; 2101t: January 2021 test run; 2104r: April routine run

Conclusions

Our results suggest minor impact and genetic progress improvement enabled by the implementation of the new international trait SNP training for clinical mastitis. Identifying which countries or individual bulls had direct MAS or only correlated SCS was previously difficult to automate correctly at the national level. Therefore, it is of most importance that Interbull continues to keep track of CMA countries to increase the reference population that will allow several benefits such as gains in REL for bulls with many CMA daughters.

Acknowledgments

The authors acknowledge the financial and/or technical support of CDCB - Council on Dairy Cattle Breeding, AGIL - Animal Genomics and Improvement Laboratory, USDA and the Dairy producers. The authors also acknowledge the Interbull Centre for providing the data.

References

B. Heringstad, G. Klemetsdal, and J. Ruane. 2000. Selection for mastitis resistance in dairy cattle: a review with focus on the situation in the Nordic countries. Livestock Production Science. 64 (2-3), 95-106.

L. Jensen, K.L. Gaddis, D. Norman. 2019. Extending genomic evaluations to direct health traits in Jerseys. ADSA 2019 Annual Meeting. Cincinnati, Ohio, 2019.

K.P. Gaddis, P.M. VanRaden, J.B. Cole, H.D. Norman, E. Nicolazzi, and J.W. Dürr. 2020. Symposium review: Development, implementation, and perspectives of health evaluations in the United States. JDS 103: 5354-5365.

P.M. VanRaden, C.P. Van Tassell, G.R. Wiggans, T.S. Sonstegard, R.D. Schnabel, J.F. Taylor and F.S. Schenkel. 2009. Invited review: Reliability of genomic predictions for North American Holstein bulls. Journal of Dairy Science. 92 (1), 16-24.