Improving Prediction of National Evaluations by Use of Data from Other Countries

R. L. POWELL,* H. D. NORMAN,* and G. BANOS†
***Animal Improvement Programs Laboratory, Agricultural Research Service, USDA, Beltsville, MD 20705-2350
† Interbull Centre, Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, S-750 07 Uppsala, Sweden

---

Received April 22, 1999.
Accepted November 15, 1999.
Corresponding author: R. Powell; e-mail: rpowell@aipl.arsusda.gov.

2000 J. Dairy Sci. (Febr.)

ABSTRACT

National and international Holstein bull evaluations from Canada, France, Germany, Italy, The Netherlands, and the US were examined to determine whether inclusion of data from other countries increased the accuracy of prediction of national evaluations for milk, fat, and protein yields. The six national and six international evaluations from February 1995 were compared with national evaluations in January and February 1999. The later national evaluations were assumed to be improved estimates of true genetic merit because of added data. Correlations with later national evaluations generally were larger for earlier national evaluations than for international evaluations, probably because of the larger part-whole relationship between earlier and later national evaluations. However, standard deviations of difference of 1995 evaluations from later national evaluation were lower for international evaluations than for earlier national evaluations, which suggested improved prediction from inclusion of multinational data. For bulls with substantial increases in daughters, nationally and internationally, correlations were higher, and standard deviations of differences were lower for international evaluations compared with earlier national evaluations. Inclusion of multinational data improved the prediction of future national evaluations, especially for countries that import genetics of dairy cattle.
(**Key words:** genetic evaluation, multinational data, international evaluation)

Abbreviation key: Interbull = International Bull Evaluation Service, **I95** = international evaluations calculated from February 1995 data by the Interbull Centre, **N95** = national evaluations used as data for I95 evaluations, **N99** = national evaluations from January and

February 1999.

INTRODUCTION

The goal of most research on animal breeding is the improvement of prediction of the true genetic merit of animals. However, because true merit is never perfectly known, proposed improvements in evaluation procedures often are assessed by their ability to predict genetic estimates from added or independent data.

Predictions of genetic merit can be improved by new evaluation methodology and by adjustment of data. In some cases, those predictions also can be improved by inclusion of additional data. However, inclusion of additional data that are less accurate than previous data may not necessarily improve prediction of genetic merit unless proper editing and weighting are applied. Powell and Norman (3) presented a method that could be used to judge the benefit of multinational data in the prediction of national evaluations. However, in their study, the time between national evaluations was insufficient to produce a conclusive assessment. Nevertheless, the results of the study generally supported the inclusion of data from other countries.

National bull evaluations have been combined into international evaluations by the International Bull Evaluation Service [**Interbull** (1)] since August 1994. Data from North America were first included in February 1995. A quarterly schedule for routine Interbull evaluations was established in November 1998. By May 1999, the Interbull evaluation effort had expanded to include six breeds, 22 countries, and 60 breed-country combinations. All data are considered simultaneously in the system of evaluation across countries (4) that is used by the Interbull Centre [Uppsala, Sweden (1)]; therefore, our hypothesis is that the merits of individual bulls are more accurately represented than is possible by using data from only one country or by combining evaluations through conversion equations.

Interbull evaluations are expressed on the scale of each country. Since August 1995, the use of genetic correlations that were <1.0 has produced evaluations with different rankings on each country's scale. By international agreement, each country has the prerogative and the responsibility to determine what use to make of the Interbull evaluations on its scale. Specifically, each country determines which, if any, Interbull evaluations are official for that country. Official generally means that the Interbull evaluations are made public and are available as information on which to base breeding decisions. In countries other than the US, the international evaluation is public information only if it is official. Holstein evaluations from Interbull were first accepted by Italy in 1999 and are not yet accepted as official in the United Kingdom.

Both Interbull and national evaluations on the US scale are available on the internet for many bulls. However, only one is designated as official. For the US, the official status of Interbull evaluations differs by breed (6). For breeds other than Brown Swiss, an Interbull evaluation is official in the US if the Interbull evaluation has information from more daughters than did the national evaluation, if the reliability of the national evaluation is <85%, and if the reliability of the Interbull evaluation is equal to or greater than the national reliability; for Brown Swiss, an Interbull evaluation is official if the Interbull reliability is at least 5% greater than the national reliability.

The benefit of using data from multiple countries is still being questioned. No country accepts all Interbull evaluations as official (2). However, for countries that accept Interbull evaluations as official for all bulls from other countries, the result is essentially the same as accepting all Interbull evaluations as official because national and Interbull evaluations are nearly the same for bulls with daughters from only that country. Most countries accept Interbull evaluations as official only for a bull that does not meet a minimum reliability requirement for its national evaluation. Thus, a domestic evaluation rather than the international evaluation commonly is designated as official, even though the domestic evaluation may be based on fewer data than were available for the international evaluation. Although that practice is due partly to the timeliness of national results, the value of additional foreign data for bulls with national evaluations has not been demonstrated empirically.

The acceptance of Interbull evaluations as official unless a domestic evaluation of a specified reliability exists means that foreign bulls initially have Interbull evaluations designated as official, and then their domestic evaluations become official when information from enough local daughters becomes available. The instability that results from this change in the source of official evaluations can damage the credibility of evaluation procedures in general.

The theory of combining data across countries has been presented by other researchers (1, 4). However, addition of data from other countries does not necessarily increase the accuracy of a national evaluation. The objective of this study was to determine whether inclusion of multinational data through the Interbull evaluation process improved the prediction of future national evaluations. Evidence of improved predictions would provide support for the use of Interbull evaluations in preference to national evaluations; failure to improve predictions would indict that policy.

MATERIALS AND METHODS

Holstein bull evaluations for Canada, France, Germany, Italy, The Netherlands, and the United States were compared with Interbull evaluations for those countries. Since the first Interbull evaluations that included US Holstein data were released in February 1995, methodology has improved to include genetic correlations (1), edits for minimum birth year to improve estimates of genetic variance (5), and improved procedures for deregression of national bull evaluations (1). Therefore, Holstein evaluations from February 1995 data were recalculated by the Interbull Centre to provide international (**I95**) evaluations for milk, fat, and protein yields on each national scale. The current methodology for Interbull evaluations (1, 4) was used; however, the genetic correlations and variances were those used for August 1995 Interbull evaluations. Canadian evaluations for 1995, which had been reported in units of breed class averages and transmitting abilities, were changed to kilograms of breeding value, which is how Canadian evaluations currently are expressed. Interbull evaluations were in transmitting ability for the US and in breeding value for other countries.

For Holstein bulls with I95 evaluations that included daughter information from that country and at least one other country, a data set was created for each of the six countries. Those data sets consisted of national (**N95**) evaluations provided as the input for I95 evaluations, the I95 evaluations on the scale of that country, and recent national (**N99**) evaluations from January and February 1999. The numbers of bulls included in the data sets are in Table 1. Because French evaluations are released only for bulls evaluated there as young bulls, most bulls from other countries that were used in France did not have French evaluations. Medians for the numbers of daughters in N95 and percentage increases to N99 and I95 are also presented to describe the data.

Comparisons of the merits of N95 and I95 evaluations as predictors of N99 evaluations would not be informative if the N95 and I95 evaluations were based on essentially the same data (i.e., relatively few data from other countries). If N99 data represented only slight increases in daughters from N95 data, N95 evaluations would be expected to predict N99 evaluations better than I95 evaluations even if N95 evaluations were not as closely related to true genetic merit as I95 evaluations were. Because of the part-whole relationship between earlier and later national evaluations and the need for added foreign data to differentiate between N95 and I95 evaluations, a subset of the data set for each country was created for bulls that had substantial increases in daughter data for both N99 and I95 evaluations compared with N95 evaluations. Those subsets included only bulls with evaluations that had information from twice as many daughters for I95 as for N95 evaluations among the 30% of bulls with the largest increases in daughter numbers between N95 and N99 evaluations. Because of the limited data available for France, the required increase in additional daughters for I95 evaluations was reduced to 50%.

Correlations of N99 evaluations with N95 and I95 evaluations and standard deviations of differences of N99 evaluations from N95 and I95 evaluations were used to assess the usefulness of international evaluations as predictors of later national evaluations. Although the genetic bases in Canada, France, and Italy had changed between N95 and N99 evaluations, correlations and standard deviations of differences should have been unaffected. Changes in national evaluation systems for N95 and N99 evaluations would make the conclusions less applicable but still useful. From 1995 to 1999, all national systems changed to varying degrees, most notably for Canada and Germany, which introduced methodology that was based on a test-day model. Those implementations reduced the numbers of bulls in the study because test-day data were not available for earlier daughters. In addition to the fewer bulls that qualified for the data subsets, the percentage increases were lower for the numbers of added daughters for those qualifying bulls.

RESULTS

Correlations of N99 evaluations with N95 and I95 evaluations (Table 2) ranged from 0.875 to 0.984 for all bulls with an evaluation in at least one country besides the country of the national evaluation. The much stronger part-whole relationship between N95 and N99 evaluations than between I95 and N99 evaluations presented a situation in which N95 evaluations would be expected to be the better predictor of N99 evaluations unless the international data in I95 evaluations were at least moderately useful in improving evaluation accuracy. For France, The Netherlands, and the US, the correlations between N95 and N99 evaluations for all yield traits were higher than correlations between I95 and N99 evaluations. However, for Canada, Germany, and Italy, correlations between N95 and N99 evaluations were lower than between I95 and N99 evaluations. The median increase in daughter numbers from N95 to N99 was 22% for Germany and 29% for Italy compared with <1 to 5% for the other countries (Table 1). Thus, more new data were available for German and Italian national evaluations relative to the other countries. The correlations in Table 2 did not indicate clearly that the use of multinational data (I95 evaluations) was a benefit in the prediction of later national (N99) evaluations.

Based on standard deviations of differences, however, international evaluations were more predictive of later national evaluations for all yield traits. The standard deviations of differences of N95 and I95 evaluations from N99 evaluations (Table 3) were smaller for I95 evaluations for all countries but especially for Germany and Italy; standard deviations of differences of I95 from N99 evaluations were less than corresponding differences for N95 evaluations by 69 to 73% for Germany, 45 to 65% for Italy, 28 to 34% for Canada, 28 to 30% for The Netherlands, 7 to 30% for France, and 7 to 19% for the US.

For bulls with substantial increases in daughter numbers nationally and internationally, correlations of N99 evaluations with N95 and I95 evaluations are in Table 4, and standard deviations of differences of N95 and I95 evaluations from N99 evaluations are in Table 5. Those statistics should be compared only within country and not across countries. For France, evaluation information was available from only 24 bulls. As expected, correlations generally were less than those in Table 2, especially those between N95 and N99 evaluations because bulls in the data subset had relatively more new data than for the full data set. Both the correlations and standard deviations of differences indicated better prediction of N99 evaluations from I95 evaluations than from N95 evaluations for all traits and all countries. The reduction of standard deviation differences from N99 evaluations for I95 evaluations relative to N95 evaluations was considerable: 28 to 36% for Canada, 26 to 51% for France, 70 to 74% for Germany, 56 to 64% for Italy, 55 to 58% for The Netherlands, and 7 to 24% for the US. The standard deviations of differences between evaluations were viewed as more informative than the correlations between evaluations because those standard deviations more directly measured closeness to later national evaluations and not just the ranking as with correlations.

Inclusion of data from other countries did not appear as useful for Canada and especially for the US as for Germany, Italy, and The Netherlands as judged by the standard deviations in Tables 3 and 5. One possible cause for this difference among countries is that Canada and the US have a longer history of exportation of genetics for dairy cattle. Therefore, data from other countries would have provided a smaller proportional increase in data for Canadian and US national evaluations. This conclusion was supported by the standard deviations for France, which had a lower proportional reduction in standard deviation differences from N99 evaluations for I95 evaluations relative to N95 evaluations than did the other European countries. Because information from bulls outside France was largely excluded from French national evaluations, France appeared to be an exporting country. Conversely, for the other European countries, inclusion of data from other countries generally meant adding information from large numbers of North American first- and second-crop daughters.

For the data sets with all bulls, most bull sires (99%), bull dams (87%), and bull maternal grandsires (99%) were North American, and most of those ancestors were from the US. For the data subsets of bulls with the largest increases in daughter numbers nationally and internationally, all sires and maternal grandsires and 94% of dams were North American. Use of multinational data that included ancestors would benefit European countries more than Canada and especially the US. The proportion of bulls in the full data set for each country that were from that country was 83% for the US, 39% for Canada, 26% for France, 22% for Germany, 18% for The Netherlands, and 6% for Italy; corresponding values for the data subsets were 31, 6, 8, 0, 4, and 0%. The origin of bulls and their ancestors can indicate the relative benefit to an individual country of using multinational data: the greater the proportion of bulls and ancestors from a country, the less benefit from using international evaluations.

CONCLUSIONS

For all bulls with I95 evaluations that included national evaluation data from at least two countries, the generally smaller standard deviations of differences from N99 evaluations for I95 evaluations compared with N95 evaluations provided evidence of the value of including data from other countries. Improvement for the US was small compared with improvements for other countries. The part-whole relationship between N95 and N99 evaluations complicated the determination of whether I95 or N95 evaluations were better predictors of true genetic merit. Correlations with N99 evaluations were greater for N95 evaluations than for I95 evaluations for half of the countries.

By requiring that I95 and N99 evaluations contain specified increases in data from N95 evaluations, the impact of part-whole relationships was reduced. Those subsets showed clearly the benefit of including multinational data from international evaluations when predicting national evaluations. Improvements were obvious from all correlations and standard deviations of evaluation differences. As with the larger group of bulls, the benefit from including multinational data was greater for importing rather than exporting countries. However, the recent importation of European semen into North America will increase the benefit of adding data from other countries to future Canadian and US evaluations as well. Benefits of international data also would be expected to be greater for countries with smaller populations.

Positive indications in this study of the usefulness of international evaluations should not be taken as a suggestion that the Interbull system for calculation of international genetic evaluations is optimal. Improvements are needed in weighting of data, calculation of reliability, and, perhaps, in estimation of parameters. Opportunities also exist for improvements in the national data that are the input to international evaluations.

Based on this study, the use of national evaluations when international evaluations are available has little justification. The benefit to evaluations on the US scale from the use of multinational data will increase as more bulls from other countries are used in the US. The minimum reliability for a US evaluation to be considered official instead of the Interbull evaluation has increased from 80 to 85%. For Brown Swiss, the Interbull evaluation is official if its reliability is at least 5% greater than the US reliability (6). The effect of those changes is that more Interbull evaluations are accepted as official in the US. This study supports that policy direction and suggests that usage restrictions on international evaluations can be relaxed further or eliminated.

ACKNOWLEDGMENTS

The cooperation of the national agencies that calculate genetic evaluations in Canada, France, Germany, Italy, and The Netherlands in supporting the use of evaluations on the national scales of those countries was essential for this project and is gratefully acknowledged.

REFERENCES

1 Banos, G., and A. Sigurdsson. 1996. Application of contemporary methods for the use of

international data in national genetic evaluations. J. Dairy Sci. 79:1117-1125.

2 International Bull Evaluation Service. 1999. INTERBULL routine genetic evaluation for dairy production traits February 1999. http://www-interbull.slu.se/eval/feb99.html. Accessed May 3, 1999.

3 Powell, R. L., and H. D. Norman. 1998. Use of multinational data to improve national evaluations of Holstein bulls. J. Dairy Sci. 81:2257-2263.

4 Schaeffer, L. R. 1994. Multiple-country comparison of dairy sires. J. Dairy Sci. 77:2671-2678.

5 Weigel, K. A., and G. Banos. 1997. Effect of time period of data used in international dairy sire evaluations. J. Dairy Sci. 80:3425-3430.

6 Wiggans, G. R., R. L. Powell, and H. D. Norman. 1998. Current and planned changes in USDA-DHIA genetic evaluations (May 1998). AIPL Res. Rep. CH11(6-98). http://aipl.arsusda.gov/memos/pdf/chng985.pdf. Accessed May 3, 1999.

TABLE 1. Number of bulls, median number of daughters, and median percentage increases in daughters between evaluations for bulls with daughter information from both that country and at least one other country for February 1995 international (I95) evaluations and for a subset of those bulls with daughter data that increased substantially for 1999 national (N99)[1] and I95 evaluations compared with 1995 national (N95) evaluations.[2]

	All bulls				Bulls with increased daughter data[3]			
			Daughter increase from N95				Daughter increase from N95	
Country	Bulls	Daughters	N99	I95	Bulls	Daughters	N99	I95
	(no.)		(%)		(no.)		(%)	
Canada	315	163	3	418	66	104	127	1977
France	129	84	0	90	24	55	110	156
Germany	227	294	22	211	61	201	166	1118
Italy	206	428	29	765	62	466	143	935
The Netherlands	271	174	3	355	72	299	66	1604
US	605	370	5	66	71	96	78	235

[1]National evaluations released in January and February 1999.

[2]National evaluations used as data for I95 evaluations.

[3]Bulls with evaluations that had information from twice as many daughters (50% more for France) for I95 as for N95 evaluations among the 30% of bulls with the largest increases in daughters between N95 and N99 evaluations.

TABLE 2. Correlations of 1999 national (N99) evaluations[1] with 1995 national (N95) evaluations[2] and February 1995 international (I95) evaluations for bulls with an evaluation in at least one country besides the country of the national evaluation.

Country[3]	N95 and N99 Evaluations			I95 and N99 Evaluations		
	Milk yield	Fat yield	Protein yield	Milk yield	Fat yield	Protein yield
Canada	0.928	0.917	0.927	0.939	0.925	0.938
France	0.958	0.962	0.958	0.955	0.948	0.951
Germany	0.918	0.908	0.912	0.923	0.919	0.926
Italy	0.875	0.937	0.880	0.908	0.951	0.914
The Netherlands	0.983	0.984	0.984	0.980	0.980	0.982
US	0.981	0.975	0.981	0.971	0.970	0.972

[1]National evaluations released in January and February 1999.
[2]National evaluations used as data for I95 evaluations.
[3]The number of bulls for each country is in the column for all bulls in Table 1.

TABLE 3. Standard deviations[1] of differences of 1999 national (N99) evaluations[2] with 1995 national (N95) evaluations[3] and February 1995 international (I95) evaluations for bulls with an evaluation in at least one country besides the country of the national evaluation.

Country[4]	N99 − N95 Evaluations			N99 − I95 Evaluations		
	Milk yield	Fat yield	Protein yield	Milk yield	Fat yield	Protein yield
	(kg)					
Canada	296	10.9	8.5	195	7.9	5.8
France	179	6.0	5.0	137	5.6	3.5
Germany	206	8.2	5.9	64	2.2	1.7
Italy	234	5.8	7.1	89	3.2	2.5
The Netherlands	118	4.0	3.3	85	2.9	2.3
US	70	2.7	2.1	65	2.2	1.8

[1]Transmitting ability for the US; breeding value for other countries.
[2]National evaluations released in January and February 1999.
[3]National evaluations used as data for I95 evaluations.
[4]The number of bulls for each country is in the column for all bulls in Table 1.

TABLE 4. Correlations of 1999 national (N99) evaluations[1] with 1995 national (N95) evaluations[2] and February 1995 international (I95) evaluations for bulls with daughter data that increased substantially[3] for both N99 and I95 evaluations compared with N95 evaluations.

	N95 and N99 Evaluations			I95 and N99 Evaluations		
Country[4]	Milk yield	Fat yield	Protein yield	Milk yield	Fat yield	Protein yield
Canada	0.833	0.876	0.849	0.915	0.930	0.922
France	0.795	0.818	0.684	0.866	0.827	0.683
Germany	0.877	0.827	0.852	0.886	0.856	0.874
Italy	0.823	0.890	0.799	0.862	0.921	0.842
The Netherlands	0.927	0.950	0.931	0.947	0.961	0.953
US	0.933	0.903	0.942	0.937	0.920	0.950

[1]National evaluations released in January and February 1999.

[2]National evaluations used as data for I95 evaluations.

[3]Bulls with evaluations that had information from twice as many daughters (50% more for France) for I95 as for N95 evaluations among the 30% of bulls with the largest increases in daughters between N95 and N99 evaluations.

[4]The number of bulls for each country is in the column for bulls with increased daughter data in Table 1.

TABLE 5. Standard deviations[1] of differences of 1999 national (N99) evaluations[2] with 1995 national (N95) evaluations[3] and February 1995 international (I95) evaluations for bulls with daughter data that increased substantially[4] for both N99 and I95 evaluations compared with N95 evaluations.

Country[5]	N99 – N95 Evaluations Milk yield	Fat yield	Protein yield	N99 – I95 Evaluations Milk yield	Fat yield	Protein yield
	(kg)					
Canada	367	13.0	10.4	236	9.3	7.3
France	337	9.9	9.1	186	7.3	4.5
Germany	254	10.7	7.7	77	2.9	2.0
Italy	253	7.0	8.1	98	3.1	2.9
The Netherlands	180	6.0	5.3	81	2.5	2.2
US	113	5.0	3.4	105	3.8	2.7

[1]Transmitting ability for the US; breeding value for other countries.

[2]National evaluations released in January and February 1999.

[3]National evaluations used as data for I95 evaluations.

[4]Bulls with evaluations that had information from twice as many daughters (50% more for France) for I95 as for N95 evaluations among the 30% of bulls with the largest increases in daughters between N95 and N99 evaluations.

[5]The number of bulls for each country is in the column for bulls with increased daughter data in Table 1.