The Interbull Audit Project Part I
Results of an International Survey

***S. König*[1]*, H. Täubert*[1]*, H. Simianer*[1] *and H.H. Swalve*[2]**

[1] *Institute of Animal Breeding and Genetics, University of Göttingen, Göttingen, Germany*
[2] *Institute of Animal Breeding and Husbandry, University of Halle, Halle/Saale, Germany*

Abstract

An Interbull project was initiated in 2001 to develop software that could be used to provide a simulation environment for testing and comparing breeding value estimation programmes.

A questionnaire, called “Informal Survey Questionnaire” (ISQ) was sent out to eight selected countries to collect information on model specifications, genetic structure and structure of environmental effects in addition to Interbull Bulletin No. 24. The aim of this contribution is an attempt to summarize some of the results from the ISQ and highlight important points that should be considered when generating data according to a country’s specific structure. The results of the survey show distributions of levels across environmental and genetic effects, e.g. calving season, size of herd-calving-years, use of sires and of young bulls and distribution of their daughters across herds, which should be accounted for in the simulation program.

Key words: Informal Survey Questionnaire (ISQ), model specifications, environmental effects, genetic structure.

Introduction

At the 1998 Interbull meeting in Rotorua, the Audit expert Group presented a report to the Interbull Steering Committee on ‚Quality assurance systems for international breeding value estimation'. This report included the proposal to develop software that could be used to provide a simulation environment for testing and comparing breeding value estimation programmes.

The software, once in place, should provide test data for individual countries that closely matches the existing data structure. To develop the software, it is therefore essential to obtain a more detailed account of the data structure that is prevailing in countries subscribing to the international genetic evaluations by Interbull. In the ongoing project which was initiated in 2001, 8 selected countries were asked for their co-operation to provide details to be used as the basis for developing the software, i.e. to ensure that the software will be powerful and flexible enough to cope with at least current, if not also future, conditions for a variety of countries and their data structure. Countries were selected in such a way that a large variety of conditions and structural differences would be presented. A questionnaire, called “Informal Survey Questionnaire” (ISQ) was sent out to the selected countries. In the questionnaire, countries were given the options to provide original data or answer the questions pertaining to their data structure.

The questionnaire contained questions related to four categories: a) General set-up of the system of genetic evaluation including model specifications, b) Use of differential weights for different types of observations, c) Structure of environmental effects, d) Structure of genetic effects (namely use of bulls from different categories).

The aim of this contribution is an attempt to summarize some of the results from the ISQ and highlight important points that should be considered when generating data according to a country’s specific structure.

Material and Methods

Of the eight countries asked for their participation, two countries sent pedigree files and data files with lactation records, including the information about environmental effects. Five countries answered the questions of the ISQ and sent cross tables representing the distribution of class levels across groups of cell sizes for random and fixed effects. One country did not respond.

The data collected from the questionnaire can be devided into the following topics:

1. A general description of the evaluation procedures practised in Interbull member countries can be found in the Bulletin No. 24. As for a deeper insight, detailed information of the evaluation model like differential weights used for specific observations was collected.
2. Information about the genetic structure. This comprises the information of complete-ness of pedigrees, the distribution of no. of daughters per bulls for different categories of bulls, the distribution of bulls over herds, the distribution of daughters of bulls over regions within the same country stratified by classes of bulls and the no. of bull sires annually and distribution of no. of sons per bull sire.
3. The structure of environmental effects includes the distribution of records across levels of all environmental effects and the distribution of records for continuous variables.

The two sets of original data can be characterized as follows: The first data set comprised 305-day first lactation records from registered Holstein-cows calving between 1997 and 1999 from a total of six regions within this country. The data set from the second country comprised 305-day lactation records from parity one to three from cows of different breeds (Holstein, Simmental and Jersey).

Results and Discussions

The detailed specification of the evaluation model can be summarized as follows: The dairy evaluation system for four countries is based upon a single trait repeatibility model applied to lactation data. In these countries, different weights for specific observations are used, i.e. different weights according to the recording scheme. In one country, a weight of 1 for all traits is given for A4 – recording schemes. For alternate or more spaced designs, the weights are multiplied by a coefficient smaller than one and ranging from .91 to 1. For in progress lactations with less than 180 days in milk (DIM), the weight is multiplied by the coefficient 1.6 DIM / (DIM + 108). Prior to further processing, completed lactations shorter than 305 days are adjusted for days in milk. In another country , for partial lactation information, an expansion factor is used within the animal model. The expansion factor is defined as b = Corr(y_p, y_c)^2. The diagonal of the R^{-1} has elements equal $\lambda_e = \dfrac{(1-r)b^{-2}}{b^{-1}-r}$. In the third country, different weights based for number of tests within lactation (0.3, 0.65, 0.75, 0.80, 0.85, 0.90, 0.90, 1, 1) for 1 to 9 tests are used. In country G, weights of lactation records are determined by the testing characteristics and number of tests. The weight is reduced from the standard of 100 for not recording all milkings on test day (am – pm schemes), owner - sampler testing, fewer than ten tests per lactation and / or if tests are not uniformly distributed across lactation.

It proved to be rather difficult to compare genetic structures across countries since editing requirements vary largely between countries. The completeness of pedigree information for cows to be included in the genetic evaluation ranges from 0 % missing dams and sires to 27.8 % missing dams and 33.4 % missing sires.

For further analysis, the stratification of daughters by sire categories was done as follows: natural service bulls (NS), progeny test bulls (PT), proven foreign bulls (PF) and proven domestic bulls (PD).

The percentage of cows in different classes of daughters per bull category revealed differences with respect to country and sire category. This is illustrated in Table 1.

Table 1. Means $\bar{x}$, and ranges R for no. of daughters per bull for 6 countries (different lines for different countries).

NS		PT		PD		PF	
$\bar{x}$	R	$\bar{x}$	R	$\bar{x}$	R	$\bar{x}$	R
7	198	74	499	1590	20000	200	2000
5	*	91	*	4388	*	259	*
*	123	*	179	*	19000	*	4000
8	150	101	356	39	12000	69	3500
8	2000	57	750	413	15000	1049	15000
*	*	113	597	526	10966	413	11531

* = information not available.

For one country, the stratification was only possible according to the origin of sires. The parameters are $\bar{x}$, =33;R=1322 for foreign sires and $\bar{x}$, =32;R=1386 for domestic sires. While the no. of daughters per NS-bull is nearly the same in all countries, i.e. up to 80 % of the NS-bulls have less than 5 daughters, the comparison for the other sire categories is quite more heterogeneous. In one country, there is a trend to produce relatively large numbers of daughters from PT bulls, i.e. 48.8 % of these young bulls have more than 100 daughters. Other countries follow with 21.2 % and 13.0 %, respectively. A wide range from one to 10000 daughters is found for the categories PD and PF.

In many countries, AI – organisations within country operate regionally. A consequence is that most young bulls only have daughters in one region. A genetic connectedness in this case can only be ensured by widespread daughters of PD and PF bulls. Joint testing schemes in which PT bulls are used in more than one region should be the preferred approach to improve connectedness, since possible disturbing effects like preferential treatment or selective mating are reduced to a minimum. The distribution of daughters across five regions within one country is given in table 2 for PT-bulls and for PD-bulls. A region is identical to the area of an AI-organisation (= origin of the bulls).

Table 2. Distribution of daughters of PT-bulls (first row) and PD-bulls (second row) in percent across regions.

Origin of bulls	Origin of daughters of bulls: A	B	C	D	E
A	94.0	.6	.8	.5	4.0
	88.1	1.4	1.4	.2	7.2
B	1.9	95.4	1.5	.9	.9
	1.9	92.7	1.3	3.3	.7
C	6.1	4.5	76.3	.8	.9
	9.9	7.3	80.7	.4	1.7
D	5.7	11.6	6.8	72.7	3.1
	6.9	51.2	6.0	31.6	4.2
E	2.0	1.4	2.5	1.1	93.1
	3.6	1.8	4.2	.1	90.2

Similar results, namely a lack in the connectedness due to PT-bulls, were found for the second country for which five regions were identified.

Another point related to genetic connectedness is the no. of bull sires used annually and the distribution of the no. of sons per bull sires (Table 3). The stratification of the data in birth years of sons was possible for two countries and three birth years.

Table 3. Distribution of bull sires (in percent) for different classes of no. of sons per sire for countries A and B.

		no. of sires	no. of sons / sire		
			1-5	6-20	>20
1993	A	78	76.9	12.8	10.2
	B	166	65.8	20.4	13.8
1994	A	79	69.5	20.3	10.1
	B	168	68.4	16.7	14.9
1995	A	91	69.3	15.4	15.4
	B	155	66.5	18.0	14.5

A concentration on a small group of dominant bull sires results in increasing connectedness and an increase in genetic trend. However, also relatedness and possibly inbreeding may be increased. The rapid slope in inbreeding for the last three decades in most countries important for Holstein genetics and its effects have been shown by Miglior (2000).

In large dairy cow herds, it is common practice that animals are separated in several sub-groups. This has many advantages for the farm management. The results of the survey reveal clear differences for herd sizes across countries and across regions within countries and therefore a subdivision in management groups can be expected under large scale farm conditions.

The structure of environmental effects was more readily identifiable and comparable than the genetic structure. The number of lactations per month of calving is quite constant in two countries, but in two other countries, there is a trend for seasonal calving from September to January. Differences in the calving management (i.e. calving season and age of first calving) are existent within countries and depending on location and type of farm management (mixed farming incl. crop farminge, large scale farm conditions, pasture).

Conclusions

The results from the ISQ provides a background for the simulation environment for testing and comparing breeding value estimation programs. Data quality issues, however, are not adressed within the current project. This is a field of another problem, which has been discussed in detail by Averdunk and Dodenhoff (2000).

In spite of the fact that the answers to the questions were not as precise and standardized as was hoped for, this survey revealed information on model specifications, genetic structure and structure of environmental effects in addition to Interbull Bulletin No. 24. The results of the survey show distributions of levels across environmental and genetic effects, for example calving season, use of sires and of young bulls, which should be accounted for in the simulation program, presented by Täubert *et al.* (2002).

References

Averdunk, G. & Dodenhoff, J. 2000. Data quality issues: From insemination to genetic analysis. *Interbull Meeting*, Verden.

Miglior, F. 2000. Impact of inbreeding – Managing a declining Holstein gene pool. *10th World Holstein Friesian Conference*, Sydney, Australia.

Täubert, H., Simianer, H. & Swalve, H.H. 2002. The Interbull audit project part II: Development of a programm for auditing breeding value estimation programs. *Interbull Meeting*, Interlaaken.