Interbull Guidelines
for
National & International
Genetic Evaluation Systems
in Dairy Cattle
with Focus on Production Traits

November 2001

PREFACE

Along with the increasing international exchange of genetic material among various dairy cattle populations, there has also been an increasing need for international comparison of national genetic evaluation systems (GES) practiced in different countries. Naturally, each country's GES comprises many steps, each of which reflects special circumstances and needs prevailing in that country. In order to make the results obtained in one country's GES as transparent as possible to the dairy industry in another country, Interbull has made it one of its objectives to provide information on GES of different dairy countries and to give guidelines on how to interpret and utilize these results.

To achieve this aim Interbull has conducted several surveys and has published the results as six Interbull Bulletins (IBB) between 1986 and 2000, the last of which is IBB 24 (2000, available through www.interbull.org) comprising information on GES for production traits of dairy cattle from 31 countries. Results of IBB 24 clearly show that different countries' national GES demonstrate so much variation that it is almost impossible for any one responsible for one country's GES to have a clear and up-to-date picture of what the GES looks like in other countries. The main conclusion from this survey (IBB 24) is that understanding and interpretation of different countries' GES is extremely difficult because:

a) The "raw data" obtained from individual animals are subjected to treatments in many stages in which the biological and statistical justification of the treatment is not so obvious to (and, in some cases, even questionable by) others;
b) There are many differences in regard to the number, nature, model and method of such treatments that it is quite reasonable to claim that these differences are making major contributions to the apparent genotype-environment interaction through the so called genotype-model interaction.

Logical consequence of these conclusions (a and b, above) is that such differences among countries' GES, even though they may be considered justified *within* each country's GES, cause undue re-ranking of bulls *between* countries. The surveys conducted by Interbull had the goal of making national GES as transparent as possible so that the interpretation of the differences becomes possible. It was also hoped that the increased transparency would eventually lead to increased awareness of difficulties brought about by disparities of the various national GES, and this in turn would lead to a greater desire for international harmonization along the 'world's best practices' known at this stage.

The present document (IBB 28) is the result of extensive discussions between Interbull member organizations, coordinated by the Interbull Centre. For this purpose, staff from the Interbull Center (H. Jorjani) assisted by a sub-group of the Interbull Steering Committee (G. Averdunk, R. Powell and H. Wilmink) prepared a preliminary draft, which was presented to the Interbull Steering Committee and Interbull member organizations in September 2000 and thoroughly discussed in a workshop in Verden, Germany, October 2000. Based on the views expressed in Verden and the ensuing discussions a new version was prepared and was reviewed by Interbull membership and the Interbull Steering Committee meetings held in conjunction with Interbull Open Meeting in Budapest, Hungary, August 2001.

Interbull proposes that genetic evaluation centers and organizations trading genetic material internationally follow the recommendations outlined in the present document, which are believed to be based on the current world's best practices, as closely as possible. The word "should" is used in this Bulletin to indicate procedures to be followed in order to be in full compliance with recommendations. The stronger word "must" has been avoided in consideration of the cooperative spirit of this joint effort.

It is the view of Interbull that further improvements in the genetic evaluation methodology will come as a result of present and future research as well as empirical experience. However, our present recommendation would greatly facilitate correct interpretations and evaluations of the animals' breeding values beyond the domestic borders and enhance the improvement in all breeding system utilizing global genetic resources.

November 2001

Hossein Jorjani
Interbull Centre

Jan Philipsson
Interbull Secretary

Jean-Claude Mocquot
Interbull Chairman

Table of Contents

1 INTRODUCTION 1

2 GUIDING PRINCIPLES 2

3 GENERAL RECOMMENDATIONS 3

4 NATIONAL EVALUATIONS 3

4.1 PRE-EVALUATION STEPS 4
4.1.1 Assignment to a breed of evaluation 4
4.1.2 Animal ID 5
4.1.3 Pedigree information 5
4.1.4 Specific genetic defects 6
4.1.5 Number of generations of pedigree data 6
4.1.6 Sire categories 6
4.1.7 Traits of evaluation 7
4.1.8 Performance record 8
4.1.9 Time period for production data 9
4.1.10 Number of lactations included 9
4.1.11 Data quality 10
4.1.12 Inclusion and extension of records 10
4.1.13 Pre-adjustment of records 12
4.2 EVALUATION STEP 12
4.2.1 Number of statistical treatments and effects in the model 12
4.2.2 Effects in the genetic evaluation model 13
4.2.3 Genetic evaluation model 14
4.2.4 Model's unbiasedness 16
4.2.5 Genetic parameters 16
4.2.6 Use of phantom parent groups 16
4.3 POST-EVALUATION STEPS 17
4.3.1 Criteria for official publication of evaluations 17
4.3.2 System validation 18
4.3.3 Expression of genetic evaluations 19
4.3.4 Genetic base 19
4.3.5 Number of evaluations per year 20
4.3.6 Use of indexes 20
4.3.7 Anticipated change 21
4.3.8 Web site 21

5 INTERNATIONAL EVALUATIONS 21

5.1 PROBLEMS IN ACROSS COUNTRY COMPARISONS 22
5.1.1 Differences in climatic conditions 22
5.1.2 Differences in management practices 22
5.1.3 Differences in trait definitions 22
5.1.4 Differences in genetic evaluation systems 22
5.1.5 Other complicating factors 23
5.2 COMPARISON OF ANIMAL EVALUATIONS 23
5.2.1 Conversion equations 23
5.2.2 Multiple-trait across country evaluation 23
5.2.3 Choice of method for international comparisons 24
5.2.4 Minimum correlations and trait harmonization 24
5.2.5 Validity of MACE results 25
5.2.6 Publication of Interbull (MACE) evaluations 25

6 REFERENCES 25

1 INTRODUCTION

In response to the need for across country comparisons the European Association for Animal Production (EAAP) set up a working group in 1975 to investigate the feasibility of standardization of breeding value estimation procedures across countries. At about the same time a working group under the auspices of the International Dairy Federation (IDF) was working towards similar goals. Continuous need for such activities led to the establishment of International Bull Evaluation Service (Interbull) in 1983 by EAAP, IDF and International Committee for Animal Recording (ICAR, formerly International Committee for Recording the Productivity of Milk Animals, ICRPMA), with support from the United Nations' Food and Agriculture Organization (FAO). Later, in 1988, Interbull was recognized as a permanent sub-committee of the ICAR. The work in Interbull gradually led to the recognition of the need for an operational unit. Consequently, in 1991 the Interbull Centre was established in Uppsala, Sweden. In 1996, the European Union (EU) appointed the Interbull Centre as its reference body for genetic evaluation of dairy cattle.

To achieve the standardization / harmonization goal the first step is to fully document different countries' national genetic evaluation systems (GES[1]). Accordingly, Interbull has acted dutifully to make different countries' GES as transparent as possible by conducting surveys on national GES for economically important traits practiced in its member countries (published as Interbull Bulletins (IBB) 2, 3, 5, 6, 13 and 24).

The second step to achieve the standardization / harmonization goal is naturally the development of a set of guidelines to encourage adoption of similar state of the art methods, 'world's best practice', by different national genetic evaluation centers (IBB 1 and 4). It is important to notice that Interbull's aim in preparing and publishing the previous set of guidelines (especially IBB 4, 1990) has been to prepare the results of national GES, specifically sire proofs for production traits, to be used in conversion equations for international use. However, the present set of guidelines (IBB 28) is mainly concerned about how the raw data are treated within each country. In other words the new Interbull guidelines addresses the issue of preparing national data to enter national and international genetic evaluations. The interest in preparation of national data has been brought about by two important developments that have occurred since late 1980's.

The first development has been an increasing trend in the number of traits in national GES, number of countries with national GES, and more importantly number of potential exporting countries. The volume of exchange of genetic material has increased and the pattern of exchange is more complex than before. Nowadays we cannot label specific countries as pure exporters or importers and virtually every country is both an importer and an exporter at the same time. Therefore, people involved in national GES must be observant of a large number of countries and this is a difficult task for anyone. It is in the interest of all countries to have a national GES that is easily understood, and also that any change in it is fully documented, so that the changes can be easily followed, should anyone wish to do so.

The second development has been the start of a truly international routine evaluation of bulls at the Interbull Centre in 1994. In the beginning, there were only a handful of countries submitting data for evaluation at the Interbull Centre, while in 2001 data from 25 countries, six breeds, and a large number of traits were sent to the Interbull Centre for routine evaluations, and it is likely that these numbers will grow over time.

The combined effect of these two developments is that the major dairy countries are today more interdependent on the decisions made at the national level in other countries' GES than at any time before. Any aspect of any country's GES may have an impact on the estimated breeding values (EBV) of bulls from all other countries. Several independent studies (Emanuelson *et al*., 1999, Jorjani, 1999, 2000, and Rekaya *et al*., 1999) suggest that structural and operational parameters of populations and national GES have a significant impact on genetic parameters. On the other hand, estimated genetic correlations between some countries are in some cases low to the point that they do not make much sense from a

[1] In this document Genetic Evaluation System (GES) is meant to include all aspects from population structure and data collection to publication of results. Each and every statistical treatment of the data that has a genetic-breeding motivation or justification is an integrated part of GES.

biological point of view. The logical conclusion is that we need to search within national GES and their structural and operational parameters for reasons of low genetic correlations.

The purpose of this set of guidelines is to facilitate a higher degree of harmonization in the things that *can* be harmonized and to encourage documentation of the things that *cannot* be harmonized at this juncture of time. It is also hoped that these guidelines act as an Interbull service to increase the quality and accuracy of evaluations at the national and international level. In other words, the aim is to increase clarity in showing the biological and statistical reasons for what is done in national GES.

2 GUIDING PRINCIPLES

Interbull, a sub-committee of ICAR, as an international non-governmental, non-profit organization based on voluntary cooperation of many countries / organizations and considering its history and its international patrons is devoted to and advocates:

Active utilization of domestic animal genetic resources; as well as

Effective use of genetic resources globally to obtain the largest sustainable genetic progress.

To achieve this Interbull encourages:

Development of national GES according to the world's best practices for a broad range of economically important traits to fit variable objectives in member countries; and also

Bi- and multi-lateral cooperation between national genetic evaluation centers.

In an across-country evaluation of animals, the ideal situation is characterized by the existence of a balance between *variation* and *resemblance* among various countries' national GES. In a global perspective certain levels of *variation* in national GES is essential for the evolution of any country's GES. We need to have the experience of doing things in different ways in order to decide on the modifications and improvements in each and every stage of national and international evaluations. However, too much *variation* causes the task of international comparisons to become very difficult, if not impossible. Hence, there should be enough *resemblance* among different countries' GES so that the information obtained for an animal from one country becomes directly utilizable and valuable in a different country.

What we have observed and concluded from Interbull surveys (especially IBB 24) is that diversification of national GES, leading to higher levels of *variation* in national GES, seems to be a process that needs no effort and input from outside. The diverse needs and problems within each country are strong enough forces to drive a country's GES towards adaptation to specific circumstances and adoption of measures that increase *variation*.

Resemblance among different countries' GES, however, needs much effort, input and coordination from outside. Naturally, an international cooperative venture, such as Interbull, is the appropriate organization to coordinate efforts that, based on biological and statistical properties of traits and data, will hopefully bring about a desirable level of *resemblance* among national GES.

The present document is the result of a continuing process initiated and maintained by a large number of breeding organizations in the Interbull member countries and coordinated by the Interbull Centre. The recommendations herewith are the result of a pragmatic compromise between dictates of several perspectives, among others, state of art in theoretical cattle breeding, and needs and capacity of individual farmers to incorporate the recommended changes into their operations. In few cases there is so much difference between practices in different countries that no compromise is in sight. In such cases, clear decisive leadership and simple to adopt and easy to follow recommendations are needed.

The present set of recommendations should ideally serve as a reflection of the best practices in today's world dairy cattle breeding and also should be able to give valuable insight about what is a sound practice, should any country start a national genetic evaluation system or change its current system to a different (and hopefully better) one.

Although we believe that the recommendations presented here are the best possible solutions for the current situation prevailing in the Interbull's member countries and for some foreseeable future (5-10 years), however, they are not to be considered as eternally sufficient. Therefore, there is a need for continuous monitoring and revision of these recommendations at least every five years.

3 GENERAL RECOMMENDATIONS

Interbull recommendations presented here are based on the latest Interbull survey, published as IBB 24, with information on GES in 36 organizations from 31 countries, and titled "National Genetic Evaluation Programmes for Dairy Production Traits Practiced in Interbull Member Countries 1999-2000" (also available through www.interbull.org). As such, these recommendations are concerned only with production traits. However, we have attempted to write it in such a way that it can be of use for other traits as well. At least, the same principles can in most cases be equally well applied to other traits.

Interbull and its parent organization ICAR are continually contributing to the development of guidelines for various stages of genetic evaluation systems. All Interbull member countries are recommended to follow Interbull and ICAR documents (also available through www.icar.org and www.interbull.org). It is recommended that countries that do not yet fulfill these standards move towards these as they change national GES. However, if ICAR and Interbull documents are silent about any matter, it is recommended that other international bodies be consulted before any decisions are made. In cases where no internationally recognized recommendation exists, adoption of procedures similar to those used in other countries is recommended. Of course, in any case detailed documentation of the adopted procedures cannot be overemphasized.

Interbull Recommendation
National genetic evaluation centers should update their GES in a cost-effective manner as the theoretical developments and computer capacity permit. They should also keep official, up-to-date and detailed documentation of all aspects of their GES on the Internet and also update their information on the Interbull web site as soon as any change has taken place.

Further, bilateral and multilateral cooperation between genetic evaluation centers is highly recommended. Cooperation may happen at a low level such as the sharing of computer codes or at a high level such as shared ownership of the genetic material and genetic evaluation systems.

Recommendations presented here should also be viewed holistically as a coherent system. Every specific recommendation pre-supposes acceptance and adherence to many other such specific recommendations. Therefore, and as an example, when "unique identification of all animals" is recommended in one section, then all further reference to "animals" is to be interpreted as "uniquely identified animals".

4 NATIONAL EVALUATIONS

In this document, different stages of national GES are divided into three parts: Pre-evaluation steps, genetic evaluation, and post-evaluation steps.

4.1 Pre-Evaluation Steps

All stages from the first collection of data, for example a milk sampling, until the time that the relevant numbers get ready to be used as inputs to the genetic evaluation procedures are considered to be pre-evaluation steps. This comprises collection of data on population structure parameters, data editing, preparatory statistical treatments of data, such as standardization or extension of records and pre-adjustment of data. For a general discussion on pre-evaluation steps see Averdunk & Dodenhoff (2000).

4.1.1 Assignment to a breed of evaluation

Interbull Centre conducts international genetic evaluations, among other traits, for production traits in six breeds: Ayrshire (AYS), Brown Swiss (BSW), Guernsey (GUE), Holstein (HOL), Jersey (JER) and Simmental (SIM). These evaluations are based on national GES for production traits in Interbull member countries. Some of these breeds, *e.g.* HOL, are found in many countries, and even some countries have two separate genetic evaluation systems for Black & White and Red Holstein. Other breeds, *e.g.* GUE, are less frequent and only a handful of countries may have any national genetic evaluation system for them.

Designation of breeds in individual countries may differ from the designations used by Interbull. For example, Holstein type breed may be called Holando Argentino, Danish Holstein, Holstein Friesian, Israeli-Holstein, Svensk Lågland Boskap and so on, to indicate both the origin of the population and also adaptation history of the population. This is not surprising because even a small population of an established breed imported to a new geographical location will start gradually to adapt to the local environment and production system and will be selected for different selection objectives, which will cause some degree of divergence from the parent population.

Assignment to a breed of evaluation is determined by several factors. A Red Holstein cow may be included in a predominately Black & White Holstein population or in a Red Holstein population. It is also possible for a population to be included in a specific breed of evaluation, say X, while its current level of genetic ties with other populations of breed X may be very low, because that population historically is considered to be of X origin. Consequently one can draw the conclusion that the assignment of individual animals or populations of cattle to any of the six breeds of evaluation is the result of an international consensus based on the origin, history of gene flow, and more importantly, current level of genetic ties in these populations.

In most of the countries genetic evaluations are conducted within breeds. However, multiple-breed (across breed) evaluations are also conducted in some countries. In addition to the Interbull's six breeds of evaluation there are a large number of local breeds and / or crosses of local with major breeds for which national genetic evaluations exist, but without ties to breeds in other countries.

Interbull Recommendation
All countries are recommended to establish national GES for all of their locally and internationally recognized breeds. Low number of animals per breed, poor production levels, type of production system and so on are important challenges, however, they should not be considered as hindrances. Assignment of an animal to a specific breed is justified if 75% of the animal's genes originate from that breed (or both sire and maternal grandsire are from the breed of evaluation).

4.1.2 Animal ID

Genetic evaluations have an absolute dependence on certain, positive, undisputed identification of animals. With no identification or wrong identification, the estimated values for components of variance as well as EBVs will be doubtful.

Interbull Recommendation
Each animal's ID should be unique to that animal, given to the animal at birth, never be used again for any other animal, and be used throughout the life of the animal in the country of birth and also by all other countries. The following information should be provided for each animal:

Breed code	Character 3
Country of birth code	Character 3
Sex code	Character 1
Animal code	Character 12

The recommendations presented here are for the purpose of international genetic evaluation of dairy cattle. Accordingly, the Animal ID given to them at birth should identify animals used in countries other than their country of birth. All parts of an Animal ID should be kept intact. If, for any reason, modification of the original Animal ID is necessary, it should be considered as a re-registration and fully documented by a cross-reference table relating the original (and intact) Animal ID and the Animal ID given to the animal in the importing country. The cross-reference table should be made available to other interested parties, particularly to the Interbull Centre. All communications at the international level should always use the Animal ID given to the animal in the country of birth.

To maintain the unique identification of all animals it is recommended that simple, software based tests be implemented so that a unique ID can be established as early as possible for all animals whose information is utilized across countries. The information available through Interbull (including the information on young bulls) can be of value in this respect.

4.1.3 Pedigree information

As mentioned before, genetic evaluations rely on identification of individuals. Genetic evaluations would find their true meaning and significance when the resemblance among individuals can be traced through the information on their parentage.

According to the survey results (IBB 24) the percentage of sire and dam identified animals shows large variation among countries. In most countries, assignment of animals to different genetic groups depends on the available pedigree information for them. The requirement for records to be used for formation of contemporary groups (calculation of herd-mate averages) is weaker than for the records for which a breeding value is going to be produced. Traditionally more emphasis is put on sire information than on dam information and records with missing dam, on occasions, may find their way into the genetic evaluations. It is also possible that an animal lacks the information on birth date. Lack of information on pedigree and / or birth date, obviously, cast a shadow of doubt over the quality of data.

Interbull Recommendation
Birth date and sire and dam IDs should be recorded for all animals. Genetic evaluation centers should, in cooperation with other interested parties, keep track and report percentage of animals with missing ID and pedigree information. The overall quantitative measure of data quality should include percentage of sire and dam identified animals or alternatively percentage of missing ID's. Measures should be adopted to reduce the percentage of non-parent identified animals and missing birth information to very low numbers and ideally to zero. Examples of such measures are supervision of natural matings and artificial inseminations, avoidance of mixed semen, monitoring parturitions, comparison of birth date with calving date of dam, taking bull's ID from AI straws, etc. If there is the slightest doubt about parentage of a calf, utilization of genetic markers, *e.g.* micro-satellites, to ascertain parentage at birth is recommended. Until this goal is achieved, it is the Interbull recommendation that doubtful pedigree and birth information to be set to unknown (set parent ID to zero).

4.1.4 Specific genetic defects

Number of identified single genes that are classed as genetic defects is increasing. There is also a widespread interest among breeders to know if the AI bulls are carriers of such specific genes, such as red gene among black cattle, and the defects BLAD, DUMPS, CVM, SMA, Weaver, etc.

Interbull Recommendation
The information on the various genetic defects should be available internationally as soon as it is possible after their existence is discovered.

4.1.5 Number of generations of pedigree data

The information on the amount of pedigree data included in the evaluations is rather incomplete. As indicated by the recent survey (IBB 24), a common response to the question of number of generations of pedigree data is 2-3 generations, but actual years of pedigree information varies from about 12 to 60 years.

Interbull Recommendation
To ensure sufficient pedigree information it is recommended that, even if production traits/records are not available, the pedigree information from the animals born within a period equivalent to a minimum of 3 generation intervals be included in the evaluations.

4.1.6 Sire categories

Bulls entering genetic evaluations in different countries can be looked upon in different ways. They may be born by their genetic dam or be a bull calf born after embryo transfer. They may be used in a natural service (NS) system or in an artificial insemination (AI) system. They may be young bulls going through progeny testing by utilization of information from their first batch of daughters or they may be proven bulls used to produce their second batch of daughters or used even after the second batch. They may be tested in one country, simultaneously tested in two or more countries or they may be imported proven bulls.

Each of the above categories has some advantages and is tarnished by some disadvantages with respect to their use in a national GES. With natural service the decline in effective population size is expected to be minimal, but the rate of genetic progress will be low as well. With artificial insemination there is higher rate of inbreeding, but also better control of matings is possible. Imported proven bulls contribute to genetic links among countries and also to faster rate of genetic progress, however, from a statistical viewpoint their use has some negative impacts on the ease of across country evaluations. It seems that one has to make some compromise between how different sire categories are used.

Interbull Recommendation
Countries should clearly and correctly describe different sire categories, that is to distinguish between domestically proven bulls vs. imported bulls, young bulls with first batch of daughters vs. proven bulls with second batch of daughters, and most important of all between NS bulls vs. AI bulls. Quantitative measures should be employed to define AI bulls. Responsible organizations are recommended to strive for establishing daughters in a large number of herds (preferably > 10) for young AI bulls.

Young bulls may be used in simultaneous progeny testing in two or more countries with large enough number of daughters in each country to warrant an independent official evaluation. These bulls should clearly be classified as "simultaneously progeny tested bulls", *e.g.* by the "P" code in the international genetic evaluation of Interbull.

International breeding evaluations are dependent on genetic links among countries. To ensure sufficient level of connectedness, Interbull encourages all its members to prepare an action plan for the exchange of young AI bulls between countries and within countries, wherever regional GES with weak genetic links are in effect. An example of a measure to build up genetic links is the simultaneous progeny testing of young bulls mentioned above (for more information refer to "General Information" on the Interbull's web site at www.interbull.org).

4.1.7 Traits of evaluation

At the moment international evaluation for the following categories of traits exists at the Interbull Centre:

a) Production traits (milk, protein and fat yield);
b) Conformation traits (18 traits, visit www.interbull.org for a complete list);
c) Health traits (somatic cell, mastitis).

Interbull member countries, in addition to these traits, may have evaluations for many other traits, including the two composition traits fat and protein percent and different functional traits. Different categories of traits (for an example of how to categorize traits see Groen *et al*. 1997) may be evaluated separately or together. Different countries or breeds have different priorities for their traits of interest and it is conceivable that for a trait of interest in Country A, there might not be an official, national evaluation in Country B.

The yield traits are most often expressed as 305-day production in kilogram (kg) or liter (l), but pound (lb) is also used. Evaluation for fat and protein percentages are usually calculated indirectly, using yield evaluations and phenotypic values.

Interbull's parent organization, ICAR, has extensive guidelines for different types of milk recordings (see for example www.icar.org/recordin.htm). The number of samplings per day and the interval between two milk samplings are among the factors distinguishing different ICAR schemes. Responsible organizations in each country, based on the special circumstances prevailing in that country, adopt one or more of the ICAR approved recording schemes. For more information on these schemes visit www.icar.org (see also Wilmink *et al*. 1998 for a discussion on the impact of milk recording scheme upon accuracy).

Irrespective of the recording schemes used, the starting point is the recording and collection of supervised, or otherwise approved, sampled milk from individual cows. This will first lead to a direct, or estimated, 24-hour milk production, as well as fat and protein yield and/or fat and protein percent, and in models based on lactation records eventually to estimated values for the entire production period (305 days). Parameters of the lactation curve or breeding values are most often expressed for a 305-day production period regardless of whether a test-day or a lactation model is used.

Interbull Recommendation
Direct measurement of traits and utilization of the metric system is encouraged. Recording organizations are recommended to adopt recording schemes that ensure accurate collection and reporting of all data. It is also recommended that national genetic evaluation centers provide detailed definitions of traits on their web sites. The definitions should include all data checks and edits, such as range of acceptable phenotypic values, age, parity, etc.

4.1.8 Performance record

After establishing the quality of record and identity of the animal that the record is coming from, one must be able to merge these two sources of information with the phenotypic performance of each animal in order to partition the variance to its causal and observational components.

Interbull Recommendation
As regards the data requirement for various traits of interest, Interbull recommendations are as follows:

a) Records of all animals with known Animal ID should be included in the genetic evaluations;
b) All records should be accompanied by relevant dates (birth, calving, etc.);
c) All records should be accompanied by sufficient information for formation of contemporary groups, such as herd and geographical location of the herd (*e.g.* region); Information on internationally standardized method of recording should be included. An example for the production traits is ICAR A4, A6, B4, etc;
d) All other relevant information, depending on the trait of interest, should accompany the record. For production traits examples of relevant information are: recording scheme, number of milkings per day, production system (*e.g.* Alpine pasture, total mixed ration (TMR) or grazing), methods for estimation of 24-hour and 305-day yields, extension methods, adjustment methods, etc.

4.1.9 Time period for production data

Number of years of production data in the genetic evaluations in different countries varies approximately between 10 and 40 years. The minimum of these values, 10 years, translated into number of generations is at most equivalent to about two generations of data (assuming a generation interval of 5-7 years depending on the selection path).

Although it is highly desirable to include several generations of data in the evaluations, however, there may be also some problems associated with it. One problem is possible changes of trait definitions during long periods of time. Another problem is the lack of data for older animals for some of the traits of interest; one example is missing information on protein yield for older cows. (For the effect of time period of data on international evaluation see also the study by Weigel and Banos, 1997).

Interbull Recommendation
Number of years of production data to be included in the evaluations should desirably be equal to at least 3 generation intervals (≥ 15 years) of consistently recorded data.

4.1.10 Number of lactations included

Number of lactations included in the genetic evaluations in different countries varies between one lactation and all lactations. There are about 15 member countries that use 5 or more lactations in their evaluations. Considering the sharp drop of number of cows from the first lactation to the second and third lactations there are not many cows left in the population with a high number of lactations, say 4-5 lactations, depending on the population (age) structure. However, improvement of health conditions and consumer concerns, as well as the increase in computer capacity have facilitated utilization of data from a longer production life span and more countries are moving away from a “first lactation only” model to “multiple lactation” models.

Dividing the first lactation into part-lactations or gathering later lactations in one group is also encountered. Production of breeding values follows the same pattern as mentioned above. Often, only one breeding value is published for an animal, but publication of several breeding values for part-lactations or separate lactations is also observed.

Interbull Recommendation
Number of lactations to be included in the evaluations is recommended to be at least three lactations. Breeding values should be produced for the whole lactation period, separately for different lactations. Separate breeding values should then be combined into one single composite breeding value for each trait for the whole life, in which different lactations are given separate weights based on each lactation’s economic value.

Number of lactations desirable to be included in the genetic evaluations also depends, among other things, on the population structure, i.e. herd size. In a population with many small herds there may be a need to resort to measures that ensure large enough contemporary groups.

4.1.11 Data quality

In each country a number of criteria are used to edit the data in order to exclude logical inconsistencies and those records that make little biological sense. Examples of logical inconsistencies are ID checks to ensure that animal's ID and her parents' ID are different or the animal has been born after her parents. Examples of biologically peculiar records are very low calving age or short calving intervals. However, among the checks for data edits there may exist some production level checks as well. This may lead to inadvertent exclusion of perfectly legitimate records, one important example of which is the exclusion of records from culled cows.

One major source of concern about data edits is the exclusion of short lactations, and to a lesser degree records in progress. Different countries impose different restrictions for required minimum number of days in lactation for inclusion of a record and this value can be as low as 5 days. So, what is rejected in one country may be considered as a completely legitimate record in another country. In other words, the minimum number of days required in one country may be considered as pre-selection of data and introduction of bias in another country.

Obviously the data entering a country's GES should have high quality, irrespective of how "quality" is defined. The quality of records should be acknowledged in genetic evaluations through some quantitative method of data quality assessment. Examples of measures that can be used are: percentage of animals in the national recording system, percentage of sire and dam identified animals, frequency of supervised or otherwise verified milk recordings, percentage of culled cows (specially in the first lactation), average DIM, and so on. Genetic evaluation centers should continuously strive for improving the overall data quality. The documentation and monitoring of how frequent different kinds of records (approved/verified and others) are and how they are treated are also very important.

It is also important to make sure that data edits do not introduce selection or bias of any kind. Pre-selection of data needs to be evaluated thoroughly. For production traits it is important to include all records. However, records of poor quality should be given lower weights.

Interbull Recommendation
It is desirable that all data related to all animals (herd book, insemination, milk recording, veterinary practices, etc.), irrespective of their sources, be available to the genetic evaluation centers in form of an integrated data-base. A complete documentation of data checks, including data edits conducted by milk recording organizations, is essential. Interbull recommends that quantitative measures of assessing data quality to be adopted by member organizations / countries. National genetic evaluation centers are also recommended to devise simple methods of checking for detection of outliers and exclusion of logical inconsistencies in the input data. Biological improbabilities should also be checked. However, extra precautions should be employed so that no inadvertent selection of data or introduction of bias becomes possible. Poor quality data should be excluded from genetic evaluations. Complete documentation of all procedures to check and edit the data is very important. National genetic evaluation centers are encouraged to have quality assurance systems implemented.

4.1.12 Inclusion and extension of records

Inclusion and extension of lactation records, or alternatively exclusion and truncation of lactations, is probably the area that the practices of different countries differ most and is judged to have the largest negative impact on the transparency of comparisons, not to mention the claim that this is the stage in which most of the bias and inadvertent selection is introduced into the GES.

The information contained in IBB 24 shows that in different countries several criteria are used to categorize lactations and what to do with lactations in each category. One criterion to categorize lactations is the lactation number. Thus, first lactation records are usually treated differently from later lactation records. Later lactation records (≥2) are usually treated alike. It may also happen that the first lactation is divided into part lactations. Another criterion used is type of lactation and if the lactation is a record in progress (RIP), from a culled cow, a dried off cow or a naturally terminated lactation of length shorter or longer than 305 days. Number of days in milk (DIM) and if a minimum or a maximum number of DIM is imposed for inclusion and extension or absence thereof is another criterion. Alternatively, instead of DIM the number of test days, or a combination of the two, may be used for categorization of lactations. When the decision is made to include a lactation record, then decisions are made for extension of the lactation and the choice of an extension method. The parameters (factors) of extension may be different for different kinds of lactations. Of course, extension of lactations is not an issue in the test-days models.

Even a cursory review reveals the fact that it is hardly possible to find two countries that treat these different categories of lactations in the same way. Consequently, it is easy to see that forming a transparent view of inclusion and extension practices in different countries is impossible. One example can be mentioned to demonstrate the graveness of the problems.

The minimum number of days required in different countries for inclusion and / or extension shows large variation. So, for example, a naturally terminated (first) lactation of length 210 DIM can be excluded from evaluations in some countries. On the other hand, the same kind of lactations, with the same length, may be included without extension in some countries or included and extended in some other countries. The same is true for records in progress and records from culled cows. The matter is more complicated by the fact that lactations are not always extended to 305 days. Assuming that the extension of records has a significant effect on different countries' genetic parameters, it can be concluded that national and international organizations need to agree on a few rules for inclusion and extension, which addresses the issues raised in here. (However, the problems associated with extension may be a diminishing problem, because more and more countries are moving to test day models.)

Interbull Recommendation

Different kinds of lactations, *i.e.* records in progress, records from culled cows, records of dried-off cows (*i.e.* lactations of cows remaining in the herd but terminated artificially because of a new pregnancy or any other management reasons), naturally terminated lactations shorter than 305 days and finally, lactations longer than 305 days should be identified in the system and treated differently.

All records with ≥45 DIM or two test days should be included in the evaluations. Extension or lack thereof should be decided upon after enough scientific / empirical justifications have been established for each kind of lactation. Records in progress and short lactations from culled cows should normally be extended. Lactations of cows dried-off before 305 days and naturally terminated lactations shorter than 305 days may be extended provided adjustment for days open and / or current calving interval have not been satisfactory. Data from lactations longer than 305 days should be cut at 305 days.

Extension methods and factors should be re-evaluated continually to ensure that they are up to date and that no unplanned selection of data occurs. Extension factors should be re-estimated at least every 5 years. Different kinds of lactations should be extended using the same extension method and different extension factors. Extension rules and methods should be the same across lactations. When ever the data span over many years the extension rules and factors should be appropriate and specific to the various time periods.

4.1.13 Pre-adjustment of records

While some countries do not perform any pre-adjustment of the records, many other countries have pre-adjustment of data for a few environmental effects. Both additive and multiplicative adjustments are in common use. Adjustment for environmental effects with a temporal component (*e.g.* calving age, calving interval, days open, month of calving and so on) is very common. An effect whose pre-adjustment may be the subject of a debate is heterogeneity of variances. While some have pre-adjustment for it, others prefer to handle it in the model and in association with the estimation of random effects (*i.e.* breeding values and residuals).

Interbull Recommendations
Generally, all effects should preferably be accounted for in the evaluation model. However, if records are to be pre-adjusted, it is more justifiable to do so for those environmental effects that are in need of multiplicative adjustments. Effects in need of additive adjustments should be considered in the model. In any case, adjustment should be made to the population mean and not to an extreme class. Further, pre-adjustment factors should be updated as often as possible (at least once per generation), and be specific to different time periods.

Evaluation centers should continually review their reasons for pre-adjustment of records to ensure that there is enough theoretical justification for the continuation of this practice and consider the possibility of whether the effect can be better dealt with in the evaluation model. Choice of the environmental effects should also be assessed carefully so that the assumptions of the models are not violated.

4.2 Evaluation Step

In an ordinary sized genetic evaluation system millions and millions of numbers obtained from a large number of animals and environments go through a number of seemingly genetically motivated statistical treatments until they are summarized in a few estimated genetic parameters for a population and a few estimated breeding values for each animal. The summarization takes place in a collection of integrated genetical and statistical algorithms commonly known as method and / or model of evaluation. “Evaluation step” is the designation used in the present document for setting up and running of these algorithms.

The genetic evaluation step (method / model) can be branded as the most sensitive part of the genetic evaluation system, “sensitive” in two respects. First, it is a sensitive step because choice of the method and the model, and the effects and parameters therein, has potentially large effects on estimated population parameters and estimated breeding values. Second, it is sensitive because huge amounts of thoughts and theoretical considerations are invested in them. For a general discussion on evaluation step see Wiggans (2000).

4.2.1 Number of statistical treatments and effects in the model

Number of statistical treatments that each piece of information is subjected to is both large and shows much variation between countries. Further, some effects may be accounted for more than once. Double or multiple counting of an effect may occur when two or more biologically close effects are used together (*e.g.* days open and calving interval or age and lactation number).

To decide on the number of statistical treatments and effects in the model several questions should be answered:

a) How large are (contemporary) group sizes?
b) Are the estimates of parameters constant over time?
c) Are multiplicative adjustment factors necessary?
d) What are the consequences of the environmental effects being adjusted for or included in the model for components of variance?
e) Is the effect to be estimated from the data or from the main random effects included in the model (breeding values, residuals)?
f) What are the effects of different combinations of parameters on the degree of freedom and of the fit of the model?

To make national GES more transparent the number of statistical treatments should be reduced as much as possible. Ideally it is desirable to take care of all statistical treatments in one single stage in the evaluation model. However, this may not be practically feasible in all situations. There are several stages at which the reduction in number of statistical treatments can occur. The easiest and most obvious is at the stage of extension of lactation records. Extension methods need to be harmonized between countries and some kinds of records could be considered for exclusion from the extension process (*e.g.* naturally terminated records).

Another measure to simplify evaluations is at the pre-adjustment stage, in which number of pre-adjustments can be reduced or eliminated altogether. One argument against the use of adjustment factors is that these are in danger of not being updated regularly. To reduce the number of statistical treatments one suggestion is to combine the extension and pre-adjustment steps with each other and then integrate them into the evaluation model by utilization of all records accompanied by the following information as co-variates (wherever applicable): birth date, calving date, culling date / drying date, number of days in milk (DIM), age, lactation number, number of milkings per day, etc.

Another advantage of bringing all of the genetically motivated statistical treatments of data into one single step, *i.e.* the evaluation model, is that there will be less confusion with regard to analysis and interpretation of some phenomena that have dubious biological background. One example is higher milk production in daughters of imported bulls that may be attributed to heterosis, genotype-environment correlation or genotype-environment interaction.

Dealing with all statistical treatments in one single step was not possible with older generation of computers. Fortunately, today’s computers pose fewer problems in this respect.

Interbull Recommendation
Organizations responsible for national GES should strive for simplicity of the analysis model and avoid amendments that reduce simplicity and clarity of the analysis model. This is not to claim that the simplest model is always the best. The best model should be decided upon considering the fit and predictive ability of the model.

4.2.2 Effects in the genetic evaluation model

After going through several stages of statistical treatments in the pre-evaluation steps (from simple data checks and edits to extension of lactation records) data are used as input variables in the national genetic evaluation models. A potentially large number of effects are used as fixed or random effects in the national genetic evaluation models. Some of these effects, usually involving herd, year, season, lactation number and calving age or calving date/month, are used in many countries. Some other effects, such as days in milk or calving interval, are used only in a few countries.

However, there is little consensus among countries as to which effects should be treated as fixed and which effects as random. This is to a certain degree understandable, because of theoretical considerations and also because of differences in population structures in different countries. The choice of the effects to be included in the model and to treat them as fixed or random should be based on sound biological and statistical grounds aimed at increasing the accuracy and avoiding the bias.

It should be remembered that whether to treat an effect as fixed or random can be considered as a two-step process. First, one must decide if the effect under consideration can be considered as a random variable. Implied in this first step is the question if one wishes to make inference about the levels of the effect not included in the data/model. Then, in the second step, the association of the effect with the main random effect considered in the model (usually animal or sire) is of decisive role (for more detailed discussions see: Schaeffer, 1999). If the association is non-random, then the effect should be treated as a fixed effect (see also Wiggans, 2000).

Interbull Recommendation
In considering an effect as fixed or random the following should be taken into consideration:

- If there is enough evidence to suggest that the effect is non-randomly associated with the main random effect;
- If number of levels is small;
- If size of groups is large ;
- If the effect has a repeating nature;
- If the effect is used to elucidate the time trend.

The above recommendation has been prepared in order to facilitate further discussions on this seemingly difficult issue and to bring more transparency to national genetic evaluation systems. It is worth noting that the assignment of effects as fixed or random according to the above recommendation may be differently practiced in different countries depending on population structure.

4.2.3 Genetic evaluation model

Description of the model used in different countries requires distinguishing between several elements in the model. These are as follows:

- Single trait vs. multiple traits: With a few exceptions, most countries, at present, have separate evaluations for different traits, *e.g.* separate evaluations for milk yield, fat yield, protein yield, and other traits if there are other traits in the national GES;
- Single lactation vs. multiple lactations: At present approximately 1/5 of the participating countries in the survey use the data on the first lactation only and therefore have a single lactation model;
- Repeatability model vs. multiple trait model: For those countries that use multiple lactation model approximately 2/3 of them have a repeatability model and 1/3 multiple trait model. Thus, in about 10 countries the genetic correlation between a trait, say milk yield, in different lactations is utilized;
- Lactation model vs. test day model: Application of a test day model in routine genetic evaluations of domestic animals is a recent development and up to now only a handful of countries have practiced it. However, several more countries have plans to embark on using this model for milk production traits;
- Sire model vs. animal model: This is probably the biggest change since the previous Interbull survey (IBB 5, 1992). Nowadays, almost all countries have moved from sire model to animal model; and finally
- If the method can be considered as unbiased: All countries participating in the survey (IBB 24) declared that they are using a BLUP method.

The model applied must reasonably well consider many factors, among others: non-genetic factors influencing the records, the structure and distribution of records on management units, the methods of sampling of young bulls (in order to avoid any bias in the EBVs), the nature of traits under evaluation and so on.

Interbull Recommendation
For the purpose of international genetic evaluations unbiasedness should be considered as the most important single criteria. The use of an unbiased method is therefore recommended. For the choice of evaluation model national genetic evaluation centers are recommended to use the following set of priorities:

a) An animal model in contrast to a sire model;
b) A within lactation multiple trait model in contrast to a within lactation single trait model;
c) A multiple lactation model in contrast to a single lactation model;
d) A multiple trait multiple lactation model in contrast to a single trait repeatability model;
e) A test day model in contrast to a lactation model.

Examples from production traits for b-d above would be:

- One single evaluation for the three yield traits (milk, fat and protein) considering their relationships is preferred to three separate evaluations for milk, fat and protein;
- One single combined evaluation for milk yield in the three lactations (1^{st}, 2^{nd} and 3^{rd} lactations) is preferred to three separate evaluations for 1^{st}, 2^{nd} and 3^{rd} lactations;
- An evaluation for milk yield as three different traits in 1^{st}, 2^{nd} and 3^{rd} lactations considering genetic correlations between them to be less than 1 is preferred to an evaluation for milk yield as one trait and a repeatability of, say, 0.5 in the three lactations.

One complicating factor is that while it is easy to see the order of priority within each of the above five groups (a to e), it is not that obvious which order of priority should be adopted between groups. For example, should a “within lactation multiple trait” (b above) be implemented before a “multiple trait multiple lactation” (d above) model? The answer to such questions is not easy. In such cases ease of implementation (including computational demand), effects on the overall bias and prediction error variance, and consequences for use of results in selection should be the deciding factor. Sometimes utilization of a simple statistical technique may help to alleviate some problems. One example is the use of canonical transformation together with multiple trait models. Another complicating factor is related to the consequences of the choices we make within each group. For example for the (d) above, what are the effects on selection of bulls and ranking of such bulls and if it is really the case that bulls are selected on more than the first lactation results.

Another point in relation to the above recommendation is that it is almost exclusively concerned with milk production traits and does not take into consideration many aspects of genetic analysis models for other traits. However, points (a) to (e) collectively point to the general direction that in setting up a list of preferences and priorities and choice of the model, the guiding principle is to choose a model that is more capable of utilizing (or exposing) the genetic variation. In practice this guiding principle translates into the choice of models that have either theoretical superiority or enable us to obtain an estimate of an animal’s breeding value that encompass a larger proportion of animal’s genome and/or life time. Even though the theoretical expectations may not be realized under certain combinations of circumstances, however, Interbull recommends adherence to superior theoretical models and encourages identification of the practical circumstances under which the theoretical expectations are not realized, should this be the case. For example, one should check the data quality to see if requirements of the model are met.

4.2.4 Model's unbiasedness

At the time of publication of the previous "Interbull Recommendations" in 1990 a transition from pre-BLUP methods or sire BLUP models to more advanced BLUP methodology was under way. Therefore, it was natural to recommend the use of such models and methods that warrant unbiasedness. This recommendation seems redundant in year 2001. However, as regard to the international evaluations the current practice (Multiple trait Across Country Evaluation, MACE) requires national genetic evaluations to be unbiased.

> **Interbull Recommendation**
> For the purpose of international genetic evaluations unbiasedness should be considered as the most important single criteria.

In this regard two points are worthy of consideration. The first point is that enough attention should be paid to the prerequisites of unbiasedness and an unbiased method, such as BLUP, for simultaneous estimation of fixed and random effects. The second point is how the biasedness is being assessed. One suggestion for assessment of unbiasedness is the monitoring and utilization of the Mendelian sampling and residuals. Also, it is Interbull recommendation that validation tests (Boichard *et al.*, 1995; visit also www.interbull.org for an operational description) should be conducted regularly by genetic evaluation centers.

Further, there are circumstances under which some degree of compromise on biasedness can be envisaged, for example to avoid high prediction error variance. As is the case for all part of this Guidelines (and especially the present chapter, "Evaluation Step") one must carefully create a balance among several criteria.

4.2.5 Genetic parameters

Genetic parameters are population and model dependent and differ from population to population. Considering the high rate of gene flow among all countries it is not unreasonable to assume that the pace change in the genetic constitution of every population is much faster than before.

> **Interbull Recommendation**
> Phenotypic and genetic parameters should be estimated as often as possible and definitely, at least, once per generation. It is also recommended that all aspects of estimation procedures for estimation of variance components (data structure, method and model of estimation, effects included in the model and so on) should be as similar as possible to the estimation procedures for breeding values.

4.2.6 Use of phantom parent groups

There are only a few countries that do not use phantom parent groups, and this seems to be a diminishing phenomenon. However, in many occasions formation of homogeneous, coherent phantom parent groups is problematic, because of the low number of (especially foreign) animals that can reasonably be attributed to be of (or coming from) the same origin.

Interbull Recommendation
The evaluation procedure should be certain to group unknown parents according to breed, country of origin, selection path and birth date or some other method to establish time trends. The procedures used for formation of phantom parent groups must give special attention to the imported animals in order to correctly evaluate these in the national GES. Phantom parent groups should have a minimum size of 10-20 animals.

For traits with low heritability there may be a need to have larger groups. If there is a need for merging different groups to attain reasonable size, priority should be given to merging those groups that contribute to the establishment of the time trend. However, caution must be taken when large discrepancies exist between means of groups with different origins.

4.3 Post-Evaluation Steps

Post-evaluation steps comprise all those decisions and activities that are related to the communication of evaluation results to farmers, owners and traders of genetic material, other genetic evaluation centers and researchers.

Post-evaluation steps are the means of interfacing with the end users, whether domestic or international. These steps are powerful tools for educational, extension work and marketing activities. No matter how good a job is done in previous steps, it is in the post-evaluation steps that one can harvest the fruits, whose value cannot be overestimated. These steps may not have direct impact on accuracy of evaluations, however, they are what makes the systems understandable and transparent to the users. For a general discussion on post evaluation steps see Jeffries (2000).

4.3.1 Criteria for official publication of evaluations

The criteria used for a bull to get an official EBV (proof) is usually based on bull’s number of daughters, number of herds the bull is represented in by his daughters, minimum reliability of the bull’s estimated breeding value, or combinations thereof. Based on the information contained in IBB 24 the minimum number of daughters in different genetic evaluation systems is as low as 5 and as high as 100, and minimum number of herds ranges between 1 and 20. Minimum reliability is usually 50-75%. There are obvious differences between countries in the levels of reliability or accuracy of animal evaluations required for official publication of these evaluations in the home country.

It must generally be emphasized that the daughters of each parent animal should be spread over many herds in order to get accurate EBVs. In particular, a sire evaluation based on a small number of herds (less than 10), with unusual distributions of daughters over these herds, may not be an accurate predictor of the bull's future progeny. While Interbull does not presume to dictate minimum levels of accuracy for individual countries, consensus seems to be that sires should not get official evaluations until they have daughters in a reasonable number of herds. Current practice at the Interbull Centre is to accept for evaluations those bulls that have daughters in at least 10 herds. For bulls with the second crop daughters (imports) the current practice is to consider a minimum of 75 daughters in 50 herds for Holstein, 15 daughters in 10 herds for Guernsey and 40 daughters in 20 herds for other breeds.

One problem in understanding and comparison of practices of different countries is associated with translation of number of daughters and herds to reliability values. However, recent developments and experiences with effective daughter contributions (EDC, visit www.interbull.org for more information) may help to alleviate the problems.

Interbull Recommendation
In general, evaluation results should be accompanied by reliabilities for EBVs and considered as official for all animals entering national GES. For randomly sampled young bulls a minimum EDC value of 10 is recommended. Official publications of individual EBV by national genetic evaluation centers should include the most recent figures or information on:

a) Effective daughters contribution or number of daughters and their distribution over herds (*e.g.* number of daughters and herds, highest percentage of daughters in a single herd, etc);
b) Number or percentage of freshened daughters being excluded from the evaluations and also the number or percentage of evaluated daughters being culled before 305 days in the first lactation or alternatively before the second lactation. When lactations in progress are extended and used, the percentage of records in progress (RIP) should be given. For national GES practicing a test-day model average number of days in milk (DIM) for daughters of a bull is considered to be equivalent to %RIP in a lactation model;
c) The theoretically expected reliability of the evaluation;
d) The type of evaluation, *i.e.* whether the evaluation is a result of regular Artificial Insemination service (*i.e.* planned progeny testing program) or not. For AI proofs a distinction must be made between (1) those of domestic young sampling bulls; (2) those of simultaneously progeny tested young bulls, the so called "P" coded bulls; (3) those based on the second batch of daughters of already proven bulls, and (4) those resulting from use of imported semen (see also the section on Sire categories);
e) Breed and definition of the genetic base.

4.3.2 System validation

Results of the survey (IBB 24) show a wide range of measures adopted by different countries for the purpose of system validation. These can vary from ordinary data checks (which can be quite extensive in some countries), through ordinary checks of phenotypic values (means, ranges and variances), to comparison of breeding values at the individual level (for detection of sharp changes in the breeding values and examination of outliers) or at the population level (correlations with previous evaluations).

The three trend validation methods, originally proposed by Boichard *et al* (1995) and later adopted and recommended by Interbull, are also used extensively by many countries (for an operational description of these methods visit www.interbull.org). As it was mentioned in section on "Genetic evaluation model" it is also important that the assumptions, prerequisites and consequences of methods and models be checked regularly.

Interbull Recommendation
The three Interbull trend validation methods I, II and III should be used for validation of national evaluations. Monitoring and examination of Mendelian sampling and residuals could also be utilized.

4.3.3 Expression of genetic evaluations

Different countries have different traditions in expression of evaluations. Absolute EBVs and relative breeding values (RBV) and indexes (with or without constant mean and variance) all are common practice. One can argue for and against each of the above methods.

Interbull Recommendation
The use of absolute EBVs is recommended, though the use of RBVs for domestic use and composite traits or indices may continue. However, in order to facilitate the international use of domestically published breeding values, in addition to the domestically used method of expression, all traits should be expressed as absolute Estimated Breeding Values (EBV), in the metric system (if applicable). Such values relate directly to the additive genetic value of the animal itself as well as to actual amounts of products.

Evaluation centers should provide detailed information on the definition and statistical properties (including descriptive statistics) of EBVs and RBVs on their web sites.

4.3.4 Genetic base

Based on the survey (IBB 24) most countries/organizations (about ¾) use a stepwise change of base and in a larger proportion of cases the cows' birth year is used in defining the base.

Choice and definition of genetic base is both trait and model dependent. It is easy to see that a trait analyzed by a sire model is more likely to be associated with a base defined in terms of bulls, and a trait analyzed by animal model is more likely associated with a genetic base defined in terms of cows. Of course, females have a numerical advantage in giving a more stable genetic base. One point of difference between different GES is the time of the year that genetic base is changed. This latter decision is influenced by the breeding season and seasonal variation in workload and holiday seasons in different countries / production systems.

Interbull Recommendation
Interbull's recommendation for definition of genetic base at the national level for production traits is to utilize information of cows born at the onset of specific 5-year periods as is outlined below. Thus, member countries should endeavor to:

a) Use cows;
b) Use birth year;
c) Use all animals that entered national GES;
d) Use average genetic merit (EBV);
e) Use stepwise change of genetic base;
f) Change the base in the years ending with 0 or 5;
g) Use cows born 5 years before the onset of the new 5-year period;
h) Change the base in the first evaluation in the years ending with 0 or 5.

For designation of genetic base the following convention should be followed:
1) A letter indicating breed of evaluation (*e.g.* A, B, G, H, J, or S for different breeds);
2) Two digits indicating the year of base established (*e.g.* 00 for year 2000);
3) A letter indicating type of animals included (*e.g.* C, or B, for cows or bulls);
4) A letter indicating the event used (*e.g.* B, or C, for birth or calving); and finally
5) Two digits to indicate the event's year (*e.g.* 95 for year 1995).

Therefore, all animals evaluated between, say, January 2000 and December 2005 in the Holstein breed evaluation will be compared with the genetic base established in the year 2000 called H00CB95, comprised of average EBV of all Holstein cows born 1995 that have entered the national GES.

If there is a need in traits other than the production traits for using bulls instead of cows, it is the Interbull recommendation that designation of the genetic base should follow the same convention as above, changing the letters or numbers as appropriate. For those countries that use a rolling base, and until they adopt the stepwise change of genetic base, the recommendation is to use information from cows born 7 years before the current evaluation. If the information from more than one year is used, the middle year should be used for the last two digits of the designation.

For the purpose of international comparisons, in addition to the nationally expressed genetic bases, one can also use Interbull evaluation results, if such evaluations exist for the desired country-breed-trait combination.

4.3.5 Number of evaluations per year

Number of official and unofficial evaluations per year is also very different in the various countries. The progress made in the field of computer science, with easier availability of high capacity computers, has made it easier for national evaluation centers to increase the number of evaluations per year. We can also observe a trend in that these two different evaluations, official and unofficial, are increasingly published on the Internet.

Another development in recent years is that an interdependence of national and international (Interbull) evaluations has developed that necessitates the coordination of these events.

Interbull Recommendation
It is recommended that national GES be scheduled to provide current and up-to-date inputs to the Interbull evaluations, which currently are performed four times per year (in February, May, August and November).

4.3.6 Use of indexes

Production and total merit indices are quite common among Interbull member countries. Number of traits included in the index and their relative weights is, obviously, dependent on the needs and wills of farmers, which in its turn is dictated by economical and political conditions prevailing in each country. From a domestic point of view, indices are very helpful, because they can provide an easy to understand way of translating biological characteristics of individual animals to ordinary, everyday economic values. From an international point of view, indices can be helpful in providing information on the traits of interest from different countries, even though some of these traits may not have an evaluation in all countries.

Interbull Recommendation
Countries are encouraged to have separate indices for different categories of traits, and for total economic merit.

4.3.7 Anticipated change

Changes of national evaluation systems, evidently influenced by theoretical and / or technical advances, seem to be taking place in phases or waves. First, a few countries adopt a change. After a while, when empirical confirmation of the theoretical expectations has been provided, other countries adopt the change. Simultaneous adoption of new methods in groups of countries is expected to provide an opportunity for genetic evaluation centers to properly utilize breeding values produced in other countries.

Interbull Recommendation
Genetic evaluation centers are encouraged to set up a long term, contingency timetable for possible future changes in all aspects of their GES. These timetables are expected to be announced world wide well in advance so that other genetic evaluation centers can accommodate to the changes.

4.3.8 Web site

Nowadays almost all countries have some information on their national genetic evaluation systems available on the Internet. A subjective and cursory examination of the various countries' information on the internet reveal large variation in the quality of information available, which is quite understandable considering the fact that this is a rather new phenomenon.

One obvious problem is the language. The main purpose of having the information on the Internet is the ease of conveying the information to the local users within each country, but the same information is invaluable to people in other countries as well.

Interbull Recommendation
National genetic evaluation centers and other relevant organizations should set up internet information sites that contain a complete documentation of the whole GES (including tables of overall statistics and EBVs of AI bulls). The information contents of these home pages are expected to be, at least, as detailed as the information published by Interbull in IBB 24 (visit www.interbull.org). Those parts of GES that are concerned with the processes (the way the data are treated) are recommended to be available in English in addition to the native language. National genetic evaluation centers should regularly update their links on the Interbull's home page.

5 INTERNATIONAL EVALUATIONS

Although due to the existence of extensive international evaluations through Interbull's evaluation services and the expected expansion of these services to new traits and breeds, the need to produce recommendations for international genetic evaluations is not as acute as at the time of publication of IBB 4, however, the need for such recommendations is both real and continues to exist for the foreseeable future.

5.1 Problems in Across Country Comparisons

Interpretation of national genetic evaluation results and an across country evaluation need to address the problems associated with genetic correlations for the same trait between two countries being less than one. The following can be mentioned as possible causes for this.

5.1.1 Differences in climatic conditions

Individuals of the same dairy cattle breed are kept in different herds under a variety of environmental conditions, from arid to humid, from lowlands to high mountains, from tropical to temperate, from equatorial to almost polar. It is very easy to postulate that there are a large number of bulls with daughters in diverse climatic conditions and the breeding values estimated from different localities may reflect different sets of genes or the same set of genes reacting differently to the climatic conditions prevailing in different localities.

5.1.2 Differences in management practices

Management practices also show much variation between different countries and even between different regions of a single country. One example is the extensive grazing system practiced in some countries, which may also bring about some inevitable consequences like the need for synchronization of pregnancies to a larger extent than in any intensive year-round feeding system. Another example is the intensive, highly concentrate-dependent herds with limited access to outdoor space / grazing. Such herds show much variation with regard to their sizes. It is also easy to postulate that in the smaller herds there is a higher risk of preferential treatment of daughters of some bulls.

5.1.3 Differences in trait definitions

All of the economically interesting traits considered in different countries' national GES have a rather complicated operational definition. As an example, the trait commonly referred to as "milk yield" may involve quite different sets of genes depending on how we get from a single milking measurement on a test day to the final value of "milk yield" that is used as the observation in the evaluation model. Obviously the final value of "milk yield" depends, among others, on measurement method, milking frequency, extension method, pre-adjustment effects, number of lactations used (*e.g.* first vs. all), and so on.

5.1.4 Differences in genetic evaluation systems

Results of the surveys conducted by Interbull clearly show that the raw data obtained from single milk recording events go through a large number of steps until they are summarized in a few estimated genetic parameters for a whole population and one or few EBVs for each animal. Some examples of such steps are estimation of 24-hour yields from milk samples, extension of records, age adjustments, pre-adjustment of the records to different environmental factors and so on. While we may have some biological justifications for our interest in knowing how different animals perform under different environmental conditions or production systems, however, it is in our interest to know that the same set of raw data will lead to the reasonably similar result if subjected to different evaluation procedures.

5.1.5 Other complicating factors

There are also a number of other issues which we still do not have a full understanding of their impact. Error in recording the right Animal ID is an example (for an example see the study by Banos *et al.*, 2001). A mistaken identity may occur for a milk record or for an animal such as a bull or sire of a cow.

5.2 Comparison of Animal Evaluations

Number and proportion of the bulls that have progeny, and hence EBV, in more than one country is steadily increasing. Simultaneous publication of sire evaluations for the same bull in several countries, provided the bulls have been accurately evaluated in each country, is an important factor needed to convert breeding values from one country to another. It is therefore highly desirable that simultaneous and joint progeny testing of young bulls is promoted widely. To obtain an estimated (or converted) evaluation for the bulls in the countries where they have no evaluation several options are available, two of which are as follow:

5.2.1 Conversion equations

One procedure to obtain estimated breeding values in importing countries is to use conversion equations developed from simple regression analysis of bulls' progeny in the importing and exporting countries, that is a bull's performance in the importing country is predicted from its performance in the exporting country. In a simulation study (IBB 1, 1986) it was shown that the procedure earlier recommended by IDF (A-Doc 64, 1981), with modifications suggested by Goddard (1985) and Wilmink *et al*. (1986) would fairly accurately estimate the parameters of conversion equations (for more details and numerical examples see IBB 1 and IBB 4).

Ease of application is the major advantage of this method, which presupposes existence of a sufficient number of bulls with national evaluations in both countries; alternatively, related animals (full-sibs, sire-son) may be used. The main problem is that such bulls are often a selected sample of the population, therefore, estimated regression coefficients are not representative of the population to which they will be applied and the estimated a and b values will be either over- or underestimated. On the other hand only two countries at a time may be compared by this method. One other shortcoming of this method is that re-ranking of animals is not possible (see also Weigel and Powell, 2000 and references therein). Numerical examples of Goddard (1985) and Wilmink *et al*. (1986) methods can be found in IBB 4.

5.2.2 Multiple-trait across country evaluation

Another procedure, currently practiced at the Interbull Centre and commonly known as Multiple-trait Across Country Evaluation (MACE) is due to Schaeffer (1994; see also Sigurdsson and Banos, 1995, Sigurdsson et al., 1996 and Rozzi and Schaeffer, 1996) which proposed a model to combine results of national GES from various countries in a joint analysis. The main model, as implemented at the Interbull Centre (Banos and Sigurdsson, 1996), is as follows:

$y_i = \mu_i 1 + Z_i Q g_i + Z_i s_i + e_i$

where

y_i = Observation vector associated with the national evaluation of a bull in country i, there may be multiple observations per bull (though only one observation from each country);

μ_i = Country of evaluation effect, reflecting the definition of the genetic base in country i;

g_i = Genetic group of bull effect (phantom group effect), defined by birth year, population of origin, and selection path;

s_i = Genetic effect of bull with variance $A\sigma^2_s$, A is the numerator relationship matrix of bulls;

e_i = Residual effect with variance $R\sigma^2_e$, R^{-1} is a diagonal matrix with diagonals reflecting the accuracy of daughter contribution in country i to the national evaluation of the bull;
Z_i = Incidence matrix relating observations to sires; and
Q_i = Incidence matrix relating sires to phantom groups.

The above model develops to a multiple-trait model where performance in each country is considered as a different trait, allowing for different genetic parameters in different countries and genetic correlation of less than unity among countries (Schaeffer, 1994). This implies that different country scales and GxE are taken into account. Another advantage of MACE is its capability to utilize all known male relationships between animals. This is especially important when sons of a bull sire are tested in different countries. This can partly compensate for the low number of multiple evaluations of the bulls coming from several countries. A more detailed description of operational steps that are performed at the Interbull Centre is available through Interbull's home page at www.interbull.org.

5.2.3 Choice of method for international comparisons

Even based on the cursory presentation and comparison of the conversion equations and MACE in the two previous sections it is evident that, in the absence of joint international evaluations based on performance records, even for bilateral comparisons MACE offers a better theoretical ground, has more advantages and suffers from lower number of shortcomings (for a comparison of these two methods see: Banos, 1998). The empirical evidence in support of the claim that MACE can be used successfully for all sorts of traits is overwhelming (see Interbull Bulletins for numerous examples).

Interbull Recommendations
For those country-breed-trait combinations that an Interbull evaluation exists, utilization of the Interbull results is recommended.

For those country-breed-trait combinations that an Interbull evaluation does not exist, utilization of the MACE methodology is recommended.

5.2.4 Minimum correlations and trait harmonization

An issue of importance in international comparisons is that how useful are international comparisons if the correlation among countries (e.g. correlation of bulls' breeding values) are very different from 1.0. Low correlation, among other things, contributes to re-ranking of bulls and low reliability of estimated breeding values for foreign bulls. Moreover, the value of such evaluation results in achieving genetic gain by importation of superior genetic material from other countries is questionable (Banos and Smith, 1991). Reasons for the correlation being different from 1.0 have been outlined in 5.1 above (Problems in Across Country Comparisons) and will not be elaborated here.

Interbull Recommendation
If the correlation between two countries is lower than ≈ 0.70 the countries involved are recommended to investigate all possible causes of low correlation, especially to examine if trait definition, genetic evaluation model and problems associated with IDs are contributing to the low correlation. In such cases action to harmonize GES in the countries involved should be taken.

5.2.5 Validity of MACE results

Estimated breeding values for bulls by the MACE methodology obviously depend, among others things, on the national EBVs for bulls and genetic correlations between countries. Therefore, these should be re-estimated when ever new national evaluations are available.

Interbull Recommendation
Always the latest available national results should be used for the MACE analysis. New genetic correlations should be preferably estimated each time the breeding values are estimated, but certainly whenever:
- The change in sire variance in any of the countries involved is more than 5% compared to the previous evaluation;
- A change in methodology, base etc has occurred in either of the countries involved;
- There is a substantial increase / change in number of bulls with evaluations in either of the countries.

5.2.6 Publication of Interbull (MACE) evaluations

Status of the Interbull evaluations in each country, and whether they are considered official or not, is decided upon by national genetic evaluation centers. A summary of these policies can be found on the Interbull web site (www.interbull.org). However, publication and advertisement of Interbull evaluations is regulated by Interbull's "Code of Practice" and especially through the "Advertising Guidelines".

Interbull Recommendation
Publication of Interbull evaluation results, *i.e.* EBVs for all bulls (irrespective of their origin) in the domestic scale is the responsibility of the national genetic evaluation centers. These are expected to make the results available to all domestic and foreign interested parties in all countries participating in Interbull evaluations. As is the case for publication of national genetic evaluation results, it is Interbull recommendation that EBV's for all bulls be published together with the reliabilities for the estimates.

6 REFERENCES

Averdunk, G. & Dodenhoff. 2000. Data quality issues: from insemination to genetic analysis. pp 3-7 in Proceedings of the Interbull Technical Workshop, Verden, Germany, October 22-23, 2000. International Bull Evaluation Service. Bulletin No. 26, Uppsala, Sweden.

Banos, G. 1998. Review of international genetic evaluation procedures in dairy cattle. Animal Breeding Abstracts, 66, 585-592.

Banos, G. & Sigurdsson, A. 1996. Application of contemporary methods for the use of international data in national genetic evaluations. J. Dairy Sci., 79, 1117-1125.

Banos, G. & Smith, C. 1991. Selecting bulls across countries to maximize genetic improvement in dairy cattle. J. Anim. Breed. Genet. 108, 174-181.

Banos, G., Wiggans, G.R. & Powell, R.L. 2001. Impact of paternity errors in cow identification on genetic evaluations and international comparisons. J. Dairy Sc. 84, 2523-2529.

Boichard, D., Bonaiti, B., Barbat, A. & Mattalia, S. 1995. Three methods to validate the estimation of genetic trend for dairy cattle. J. Dairy Science. 78, 431-437.

Emanuelson, U., Fikse, F. & Banos, G. 1999. Impact of national evaluation models on international comparisons. Proceedings of the computational cattle breeding, '99 Workshop, Tuusula, Finland, International Bull Evaluation Service, Bulletin No. 20, Uppsala, Sweden.

Fikse, W.F. & Sullivan, P. 2000. Estimation of sire variance in international genetic evaluation models with genetic groups. J. Dairy Sci., 83, Suppl. 1, pp 60.

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

Groen, A. F., Steine, T., Colleau, J.J., Pedersen, J., Pribyl, J. & Reinsch, N. 1997. Economic values in dairy cattle breeding, with special reference to functional traits. Report of an EAAP-working group. Livest. Prod. Sci. 49, 1-21.

IDF. 1981. IDF recommended procedure for international comparison of genetic merit of dairy cattle. International Dairy Federation, Doc A-64.

Interbull, 1986. Procedures for international comparisons of dairy sires- current practice and evaluation of methods. International Bull Evaluation Service, Bulletin No. 1, Uppsala, Sweden.

Interbull. 1986. Sire evaluation procedures for dairy production traits practiced in various countries 1985. International Bull Evaluation Service. Bulletin No. 2, Uppsala, Sweden.

Interbull. 1988. Sire evaluation procedures for dairy production traits practiced in various countries 1988. International Bull Evaluation Service. Bulletin No. 3, Uppsala, Sweden.

Interbull. 1990. Recommended procedures for international use of sire proofs. International Bull Evaluation Service. Bulletin No. 4, Uppsala, Sweden.

Interbull. 1992. Sire evaluation procedures for dairy production traits practiced in various countries 1992. International Bull Evaluation Service. Bulletin No. 5, Uppsala, Sweden.

Interbull. 1992. Sire evaluation procedures for non-production, growth and beef production traits practiced in various countries 1992. International Bull Evaluation Service. Bulletin No. 6, Uppsala, Sweden.

Interbull. 1996. Sire evaluation procedures for non-dairy-production and growth & beef 1996. International Bull Evaluation Service. Bulletin No. 13, Uppsala, Sweden.

Interbull. 2000. National genetic evaluation programmes for dairy production traits practiced in Interbull member countries 1999-2000. International Bull Evaluation Service. Bulletin No. 24, Uppsala, Sweden.

Jeffries, M. 2000. Comparing apples with apples – Best practice for presenting genetic evaluation results. pp 13-17 in Proceedings of the Interbull Technical Workshop, Verden, Germany, October 22-23, 2000. International Bull Evaluation Service. Bulletin No. 26, Uppsala, Sweden.

Jorjani, H. 1999. Connectedness in dairy cattle populations. pp 21-24 in Proceedings of the 1999 Interbull Meeting, Zurich, Switzerland, August 26-27, 1999. International Bull Evaluation Service. Bulletin No. 22, Uppsala, Sweden.

Jorjani, H. 2000. Impact of input data quality on national genetic evaluations. J. Dairy Sci., 83, Suppl. 1, pp 53.

Rekaya, R., Weigel, K.A. & Gianola, D. 1999. Bayesian estimation of parameters of a structural model for genetic covariances between milk yield in five regions of the USA. Proc. 50th Annu. Mtg. EAAP, Zurich, Switzerland. Wageningen Press, The Netherlands.

Rozzi, P. & Schaeffer, L.R. 1996. New deregression procedure used on type traits. Interbull technical workshop, Verden, Germany, 25-26 November, 1999.

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

Schaeffer, L. R. 1999. Animal Models, Course notes. University of Guelph, Guelph, Canada.

Sigurdsson, A. & Banos, G. 1995. Dependent variables in international sire evaluations. Acta Agric. Scand. Sect. A, Animal Sci. 45, 209-217.

Sigurdsson, A., Banos, G. & Philipsson, J. 1996. Estimation of genetic (co)variance components for international evaluation of dairy bulls. Acta Agric. Scand. Sect. A, Animal Sci. 46, 129-136.

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

Weigel, K. A. & Powel, R. 2000. Retrospective Analysis of the Accuracy of Conversion Equations and Multiple-Trait, Across-Country Evaluations of Holstein Bulls used Internationally J. Dairy Science 83, 1081-1088.

Wiggans, G.R. 2000. Issues in defining a genetic evaluation model. pp 8-12 in Proceedings of the Interbull Technical Workshop, Verden, Germany, October 22-23, 2000. International Bull Evaluation Service. Bulletin No. 26, Uppsala, Sweden.

Wilmink, J.B.M., Galesloot, P., Johnson, D., Ouweltjes, W., Rosati, A., Schaeffer, L., Steine, T. & VanRaden, P. 1998. Final report of the ICAR-Interbull working group on milk recording accuracy and its impact on genetic evaluations. 10 pages. Proceedings of the 31st ICAR Session, January 17-22, Rotorua, New Zealand.

Wilmink, J.B.M., Meijering, A. & Engel, B. 1986. Conversion of breeding values for milk from foreign populations. Livest. Prod. Sci. 14, 223-229.