

Interbull Scientific Advisory Committee (SAC)

Annual report (2005-2006) to the Interbull Steering Committee (SC)

Interbull SAC: Vincent Ducrocq, Mike Goddard, Larry Schaeffer, Georgios Banos (convener)

INTRODUCTION

Another year of considerable international activity has elapsed and SAC thanks Interbull and the Interbull Center for their good cooperation.

The objective of this report is to inform the Interbull SC of our activities during 2005-2006 and provide views on pertinent scientific issues, as identified and discussed during the 2005 Interbull meeting in Uppsala. Further e-mail communication also took place between SAC and the Interbull Center, and also among the SAC members. Focus was mainly placed on two areas: **(i)** Genetic evaluations for functional traits (in particular fertility) and **(ii)** Estimation of genetic correlation among countries in the Interbull MACE evaluation system.

1. GENETIC EVALUATIONS FOR FUNCTIONAL TRAITS (FERTILITY)

Although emphasis was chiefly placed on fertility traits, this section may also pertain to other functional traits that either have been included in the Interbull service portfolio or are being considered for future application.

Observations/Issues/Problems

1. Fertility has a moderate to high genetic correlation with milk yield, meaning that milk may be used as a predictor trait of fertility; it also means that milk should be considered in a genetic evaluation for fertility in order to account for selection bias by including data on which selection was based, thereby yielding unbiased parameter and breeding value estimates for fertility. Failure to account for milk may result in biased genetic evaluations.
2. Fertility traits are frequently binary (e.g. non-return rate) or categorical or censored (e.g. number of inseminations per conception), yet they are mostly analyzed as continuous variables with conventional linear models. This may also yield biased results.
3. Problems above notwithstanding, it is important to proceed with the calculation of reasonable (inter)national genetic evaluations and focus on continuous improvements of all systems.
4. Currently some countries calculate genetic evaluations for fertility using single- and others multi-trait models. Similar circumstances pertain to yield, conformation and mastitis/SCC evaluations. This heterogeneity of practice is not being duly accounted for in the system.
5. Some countries applying multi-trait evaluation models include milk yield in the analysis, thereby accounting for selection bias; others do not. The heterogeneity of this practice is even more important than the previous one since some national genetic evaluations are potentially biased while others are probably unbiased.

Applicability of current validation system

1. There is no point in performing the 3 trend validation tests on data that are potentially biased (e.g. national evaluations not accounting for milk yield). These methods are irrelevant/inadequate when input is biased; national genetic evaluations fraught with systematic bias due to selection may still pass the trend validation results.
2. Current trend validation methods are applicable to unbiased genetic evaluations for fertility traits (e.g. from multi-trait models including milk).
3. Mendelian sampling (MS) variance tests should be applicable in all cases given that MS is not affected by selection.
4. Checking the consistency of consecutive genetic evaluation runs may be applicable but it, too, suffers from the same the limitations imposed by biased input.

Alternative data quality assessment methods considering data subsets

1. In the 2004-2005 SAC report there was reference to alternative methods. In particular it was stated that: “The key to assessing the validity of any genetic evaluation would be to ensure that estimated breeding values (EBV) of bulls (especially of those whose semen will be broadly marketed) accurately predict future daughter performance. A useful test would be to re-compute today’s genetic evaluations after having excluded data from the last four years and compare them with adjusted daughter records (or daughter deviations) in the last four years or with currently official EBV. Regression of daughter deviations on reduced data EBV should be near unity. Differences between reduced data and official EBV should average zero, be independent of the former and have variance proportional to the change in reliability....”
2. Following up on the same topic, a complementary approach would be to pick out 1-2 cows from each contemporary (management) group and analyze the remaining data with the official national evaluation model. Records of the excluded cows may then be compared to their calculated pedigree indices from the test evaluation.
3. It is recommended that Interbull encourage participating countries to apply an explicit data quality assessment procedure based on cross-validation using independent subsets, as described above. Each country could prepare a report of the procedure applied. In all cases, tolerance criteria should observe statistical significance rules. This can first be done on a test basis with a view of implementing it at a later point as pre-requisite for participation in the international genetic evaluation service. The objective here would be to ensure the quality of national evaluation results.

Recommendations

1. Further research is needed to assess and quantify the actual impact of bias on fertility evaluations emanating from not accounting for milk yield.
2. Individual countries should consider appropriate models to ensure their genetic evaluations are unbiased (e.g. multi-trait models analyzing fertility and milk yield simultaneously or correcting for selection on milk yield in some other way).
3. At Interbull level, the challenge is to derive appropriate input to MACE. A few related points are outlined next:
 - a. In case of multi-trait national genetic evaluations (milk-fertility), single-trait de-regression (current system) may produce unbiased input to MACE, which, however,

would not model exclusively the functional trait (fertility), as it would carry a lot of weight on milk. The problem already exists with several conformation evaluations from multi-trait national models being dealt with in single-trait international analyses. The consequence may be more severe for the milk-fertility scenario because of their heritability difference placing more milk weight on the fertility proof. The impact of this should be investigated. Such an application may be useful but not optimal. It is important not to treat fertility evaluations that contain milk information as if they were “pure” fertility data.

- b. Theoretically, a multi-trait de-regression would render unbiased de-regressed fertility proofs independent of milk. This is ideal input to MACE. Technically, the challenge would be to model the residual (within bull) covariance structure. A useful approximation might be to avoid to actually compute de-regressed milk proofs and use their values from single-trait milk de-regression instead. The relatively safe assumption that these are not very different is made here. Effectively, the multi-trait de-regression equations will collapse to a set of single-trait de-regression equations. The resulting fertility data are then de-regressed to the extent of the respective national genetic evaluation model, having removed the milk effect.
 - c. Alternatively, a trait decomposition (e.g. with canonical transformation) could take place at national level, resulting in unbiased fertility proofs uncorrelated to milk. The latter may also be technically challenging.
 - d. A different approach would be to use single-trait MACE to analyze milk and milk-corrected fertility. The two traits would be independent and render themselves suitable for a single-trait analysis. The key here would be to calculate appropriate milk-corrected fertility data.
 - e. In any case, it is important to ensure that unbiased, “pure” fertility proofs (that contain no milk data) are used as input to MACE.
4. When unbiased national de-regressed proofs are available that model fertility-only data i.e., they do not contain milk information, there may be no need to consider a multi-trait MACE. It is possible to combine single-trait MACE fertility and milk proofs subsequently, using selection index methodology.
 5. The focus should always be on improving the national (and international) genetic evaluation systems and make sure results are unbiased, rather than trying to pass the trend validation tests. Logically, these two should be consistent with each other. With traits like fertility (and survival), however, focus on passing these tests may not always lead to real model improvements.
 6. Appendix I illustrates various relevant scenarios.

2. ESTIMATION OF GENETIC CORRELATION AMONG COUNTRIES IN THE INTERBULL SYSTEM

There is a difference between these correlations and biological correlations. Interbull correlations are affected by:

1. True genotype-environment interaction (biological component).
2. Differences in trait definition between countries (recording/management component).
3. Differences in genetic evaluation models between countries (technical component).

It should be recognized that MACE is a meta-analysis based on processed data. The fact of the matter is that these parameters basically determine how much weight is placed on “foreign” proofs.

Regarding the question whether some Interbull correlation estimates should/could be set to unity, SAC comments as follows:

1. Although technically feasible, scientifically there isn't much evidence to support this. It would only make sense for countries that have merged their genetic evaluation operations and national databases.
2. One of the motivations might be lack of “believable” estimates between specific country pairs for some of the new traits. It is not obvious how setting correlations to 1 would solve the problem. Bayesian approaches and/or bending towards a common average or even the estimate for milk yield correlation would be preferable.

May 30, 2006

For the Interbull SAC

Georgios Banos, Convener

Appendix I

Scenarios for Interbull fertility evaluations

ST refers to single-trait analysis (fertility); MT refers to multi-trait analysis (fertility + milk); DYD refers to de-regressed proofs

Scenarios:	1	2	3	4	5	6	7	8
National evaluation	ST	MT	ST	MT	MT	MT	MT	ST
Interbull de-regression	ST	ST	ST	MT	MT	MT	ST	ST
Interbull MACE	ST	ST	ST	ST	ST	MT	MT	MT
Post-Processing	-	-	Selection index = ST(milk) + ST(fertility)	-	Selection index = ST(milk) + ST(fertility)	-	-	-
Comments	<ol style="list-style-type: none"> National proofs/DYD pertain to fertility only Potentially biased national proofs → possible bias in MACE. Computationally easy. 	<ol style="list-style-type: none"> National proofs/DYD unbiased. Fertility DYD contains milk info. Computationally easy. 	<ol style="list-style-type: none"> As scenario 1. Selection index methodology is used to combine ST MACE fertility and milk. 	<ol style="list-style-type: none"> National proofs/DYD unbiased. Fertility DYD independent of milk (appropriate input to MACE). Computationally more challenging than the first three. 	<ol style="list-style-type: none"> As scenario 4. Selection index methodology is used to combine ST MACE fertility and milk. 	<ol style="list-style-type: none"> National proofs/DYD unbiased. Fertility DYD independent of milk (appropriate input to MACE). Computationally more challenging than scenarios 4-5. 	<ol style="list-style-type: none"> National proofs/DYD unbiased. Fertility DYD contains milk info. Computationally as challenging as scenario 6. 	<ol style="list-style-type: none"> Potentially biased national proofs/DYD. Some (but not all) bias is accounted for by MT MACE. Computationally as challenging as scenario 6.

Current situation: mix of scenarios 1 and 2 (sub-optimal; possibly useful in the immediate run to gain experience with the trait).

Possible near-term improvement: mix of scenarios 3 and 5 (could be developed and implemented in the near future).

Ideal scenario: 6 (future application; needs considerable work).