Data exchanges for beef international evaluation

Thierry Pabiou, Eric Venot, Jean Guerrier

Data exchange for beef international evaluation.

Thierry Pabiou[1], Eric Venot[2], Jean Guerrier[3]

[1] *Irish Cattle Breeding Federation, Highfield house, Bandon, Co Cork, Ireland*
[2] *Institut National de la Recherche Agronomique, SGQA, F78352 Jouy en Josas, Cedex, France*
[3] *Institut de l'Elevage, Department Genetique, 149 rue de Bercy, 75595 Paris Cedex 12, France*

Abstract

When a country wishes to join the international genetic evaluation project for beef cattle, it will be ask to provide several specific files of data: 501 file for the identification of genetic links between countries, 502 files for the pedigree and several performance files depending on the type of evaluation. The essential of the files exchange will take place between the country and InterBeef who act as the manager and the regulator who set up the data validation process. Within InterBeef, two central files will be used during a genetic evaluation: the cross reference file (CRF) and the Ancestors File (AF) gathering respectively the 501 and 502 files from the member countries. The organisation of the system has been put together while running the first prototypes of international evaluation between France and Ireland and is now open to improvement.

1. Introduction

The development of an international genetic evaluation system of beef cattle was initiated in 2001 with a contracted project between INRA (France), the Irish Cattle Breeding Federation (Ireland), the Meat and Livestock Commission (United Kingdom), Institut de l'Elevage (France), and the Interbull sub-committee of the International Committee for Animal Recording (ICAR). A prototype has been developed for the joint genetic evaluation of weaning weight of Charolais and Limousine breeds. The treatment of other traits such as calving ease, weaning muscularity, weaning skeletal and slaughter weight will follow. Pedigree and performances obtained from several countries are to be treated together. In order to facilitate data validation and treatment, specific file formats were defined for file exchanges (following the ICAR recommendations outlined in the International Agreement of Recording Practices, 2001). Every country willing to participate to the joint genetic evaluation is asked to provide data in accordance with this standard files format.

This document describes the different files to be prepared by a foreign country who wants to participate to the international genetic evaluation (INTERBEEF) of beef cattle.

2. Breed

For a good understanding, in genetic evaluation terms, two aspects of 'breed' have to be developed:

- **The breed for identification** is defined by the breed code attributed by the country of first registration – in most cases, the country of birth.
- **The breed for evaluation** represents a group of breeds which undergo a common genetic evaluation. The breed for evaluation can therefore differ from the breed for identification and will be handled by InterBeef prior to each genetic evaluation.

3. The Animal International Identification

A unique international identification for each animal is used in the genetic evaluation system and will be referred to as the Animal International IDentification (AIID). The AIID is constructed as follow:

Breed + Country + Sex + ID number

- '*Breed*' refers to the code (char 3.) (Interbull,2007) for *the breed for identification of the animal* in the country of first registration,
- '*Country'* is the code (char 3.) (Interbull, 2007) for *the country of first registration* of the animal,
- '*Sex'* is the code (M=male, F=female) for *the sex* of the animal,
- '*ID number'* refers to *the identification number* of the animal (Char 12.) in the country of first registration.

Ex: CHLFRAM006327826864 is the AIID of

- *A Charolais* *CHL*
- *First registered in France* *FRA*
- *A male* *M*
- *With the ID number in France* *006327826864*

Ex: BAQIRLF125693456875 is the AIID of

- *A Blonde d'Aquitaine* *BAQ*
- *First registered in Ireland* *IRL*
- *A female* *F*
- *With the ID number in Ireland* *125693456875*

4. Cross identification of animals

The accuracy of a joint genetic evaluation is based on the quality of the genetic links between the participating countries through the AIID of common animals. The construction and the validation of the AIID are therefore crucial. The first step for a country that wishes to join the international evaluation system is to identify all the foreign animals presents in its database and create a cross identification file called 501 file.

41. The 501 file

The 501 a *cross identification file generated by a country* who wishes to include its animals in the international genetic evaluation process. This file gathers all foreign animals present in the country's database in a specific format (see Appendix 1) and *establish the link between the AIID of the animals and their identification tag(s) in the national database.*
Once this 501 file is created (Interbull, 2007), the country sends it to InterBeef.

42. The Cross Reference File

The Cross Reference File (CRF) is the *multi-breed central file managed by InterBeef* and used in the international genetic evaluation process. The CRF is the accumulation of the validated lines *(see 43. Process of validation of AIID)* from 501 files generated by the member countries has the same format as the 501 files.

43. Process of validation of AIID

Before entering the genetic evaluation process, each AIID send to InterBeef *via* 501 files has to be validated. An AIID should be validated by the country it refers to in the field 'Country'. Thus, on reception of a 501 file, InterBeef have to dispatch the different lines to the relevant country for validation.
If the AIID is correct, the country for validation validates the lines by putting 'V' in the 'CHED' field of the 501 (see Appendix 1). If part of the AIID is incorrect, the field 'CHED' is marked 'D' and the line is non-validated. The country for validation can add a comment in the field 'CASE' to specify its reasons for non validation. A comment can also be added for validated lines regarding secondary information (date of birth and name).
If the country for validation is not a member of the joint beef genetic evaluation, InterBeef can choose to validate or not the lines: 'V' or 'D' in the 'CHED' field and 'III' in the 'CFC' field.

44. Update of the CRF

Each member country is individually responsible for the update of the InterBeef cross reference file.

5. Pedigree of animals

A country part of the joint genetic evaluation has to provide the full pedigree of the animals undergoing the genetic process. This pedigree file created by each country is called a 502 file.

51. The 502 file

The format of a 502 file is given in Appendix 2 and is *limited to the international identification of the animal and its parents.* This file should contain the ancestors of all the performance recording animals and of their parents: all animals known as sire or dam must also have an entry in this file as animal. In the end, the unknown sire or dam has to be indicated with 'SIIDC' or 'DIIDC' equal to "000" and 'SIID' or 'DIID' equal to "000000000000" (see Appendix 2).
Once this 502 file is created (Interbull, 2007), the country sends it to InterBeef.

52. The Ancestors File

The Ancestors File (AF) is the multi-breed central file managed by InterBeef and used in the international genetic evaluation process. The AF gathers, after validation steps, the different 502 sent by the member countries.

53. Process of validation of sire & dam

The following data from a 502 are INCLUDED in the central Ancestors File (the country providing the 502 is referred as 'the sending country'; refer to Annex 2 for the term explanation):

a) All the animals born in the sending country ('IIDC' = sending country) AND with sire (blank accepted) born in the sending country ('SIIDC' = sending country) AND with dam (blank accepted) born in the sending country ('DIIDC' = sending country).

b) Sire (blank non-accepted) of an animal born in the sending country ('IIDC' =

sending country) is born OUTSIDE the sending country ('SIIDC' ≠ sending country) IF that sire HAS an entry in the InterBeef cross reference file.

c) Dam (blank non-accepted) of an animal born in the sending country ('IIDC' = sending country) is born OUTSIDE the sending country ('DIIDC' ≠ sending country) IF that dam HAS an entry in the InterBeef cross reference file.

d) Animal born OUTSIDE the sending country ('IIDC' ≠ sending country) IF the country of origin IS NOT part of InterBeef AND IF that animal HAS an entry in the InterBeef cross reference file.

Data not included in either a), b) c) or d) cited above ARE NOT INCLUDED in the central ancestors file.

54. Update of the Ancestors File

Each member country is individually responsible for the update of the central ancestors file.

An update concerns the whole (replacement of all the ancestry lines for one country) pedigree. Every update is sent using a 502 file. An update can occur at any time but a deadline will be defined by InterBeef before the projected date of a genetic evaluation.

6. Other file for the management of the InterBeef CRF and AF

The 'DEL' file is a file accumulating the deleted lines ('CHED' = 'D') of the 501 files received by InterBeef. The format of the 'DEL' file is a 501/CRF format.

7. The performance files

All animals with a performance must have a line as animal in a corresponding 502 and/or the Ancestors File. It is advised that the performance files are created after the cross identification and pedigree steps are completed. The first columns of the different performance files consist of the international identification of the animal (see Annex 4 for an example of description).

71. Herd definition

As indicated by the ICAR recommendations, any group of animals kept for the same purpose and at the same location shall be regarded as a whole herd. For a performance record to be considered an official record, *the whole herd as defined above must be recorded*.

72. The different files

So far, InterBeef is working on the weaning weight trait – the corresponding performance file is called Calf adjusted weight and is coded 510. An example of that file can be found in Appendix 3.

Other files will be developed as the genetic evaluation expands.

8. Conclusion

The development of an international genetic evaluation system of beef cattle has increased the flow of data exchanged between countries. In order to manage these exchanges, InterBeef have called for a rigorous and efficient system based on international identification, common file format and rules to regulate the traffic. The current area of work is on means of integrating the cross bred animals in the evaluation.

This organisation has been put together while building of the first international evaluation (Venot, 2006); improvement can now be added.

References

Interbull, 2007: Guidelines to join InterBeef, version 1.

Venot *et al*, 2006: First steps towards a European joint genetic evaluation of the Limousine breed, Interbull meeting, Kuopio, June 2006.

Acknowledgement

Thanks to the following authors for their work in this field of data exchanges:

***G. Renand**, **2002**, Guidelines to enter into the International Beef Genetic Evaluations, December 2002, 17 pages.*

***L Journaux**, **2005**, Evolution du contenu des fichiers 501 et 502, Cahier des charges, June 2005, 25 pages.*

***L Journaux**, **2005**, Participation d'un pays aux evaluations InterBeef, Cahier des charges Version 1.00, December 2005, 22 pages.*

***E. Venot**, **2005**, Files descriptions for exchanging data, December 2005, 8 pages.*

Appendix 1 : Description of the 501 file

Field description	Label	Format	Start	Length	Status	Note	Comments
File code	**FCODE**	**Char**	**1**	**3**	**M**		**= 501**
Breed	**BREED**	**Char**	**4**	**3**	**M**	**I**	**International**
Country of birth	**IIDC**	**Char**	**7**	**3**	**M**	**II**	**identification**
Sex	**SEX**	**Char**	**10**	**1**	**M**	**III**	**(Interbull format)**
ID code of the animal	**IID**	**Char**	**11**	**12**	**M**	**IV**	
Name of the animal	NAME	Char	23	30	O		Complementary
Birth date of the animal	BDATE	Int	53	8	O	V	information
Breed	**BREEDC**	**Char**	**61**	**3**	**M**	**I**	**National**
Country	**NIDC**	**Char**	**64**	**3**	**M**	**II**	**identification**
Sex	**NSEX**	**Char**	**67**	**1**	**M**	**III**	**(Interbull format)**
ID code of animal	**NID**	**Char**	**68**	**12**	**M**	**IV**	**/ Alias**
Country of national ID	C_NID	Char	80	3	M	II	Country providing the national identification
Name of the animal	NNAME	Char	83	50	O		
Other ID1	OID1	Char	133	20	O	VI	Other
Other ID2	OID2	Char	153	20	O	VII	national
Other ID3	OID3	Char	173	20	O	VIII	identifications
Other ID4	OID4	Char	193	20	O	IX	
Other ID5	OID5	Char	213	20	O	X	
Check digit	**CHED**	**Char**	**233**	**1**	**M**	**XI**	**Validation / country**
Cause of rejection	**CASE**	**Char**	**234**	**10**	**M**	**XII**	**of international ID**
Country for creation	CFC	Char	244	3	M	II	
Date of creation	DOC	Int	247	8	M	V	
Country last modification	CFLM	Char	255	3	M	II	
Date of modification	DOLM	Int	258	8	M	V	

Status	M ⇔ Mandatory, O ⇔ Optional
Note	
I	List of Breed code available at Interbull
II	List of Country code available at Interbull
III	Sex : M ⇔ Male, F ⇔ Female only
IV	Identification Number in Interbull Format : Alpha-numeric codes only, Right justified, Left blanks being filled with Zero (“0”)
V	Numeric : YYYYMMDD (YYYY : year, MM : Month, DD : day)
VI	Other form of identifications used in the country (herd book number, tattoo…)
VII	V : validated ; D : one or several errors have been detected by birth country ; line yet to be validated
XII	To be filled if the ‘CHED’ field equals ‘D’ (identification non valid) ; the cause of rejection can be coded as follow: • B : problem / birth date • N : problem / name • S : problem / sex • I : problem / ID structure • R : breed code error

Appendix 2 : Description of the 502 file

Field description	Label	Format	Start	length	status	Note	Comments
File code	**FCODE**	Char	1	3	M		**= 502**
Breed	**BREED**	Char	4	3	M	I	**International**
Country of birth	**IIDC**	Char	7	3	M	II	**identification**
Sex	**SEX**	Char	10	1	M	III	**(Interbull)**
ID code of the animal	**IID**	Char	11	12	M	IV	
Date of birth	**BDATE**	int	23	8	O	V	
Name	**NAME**	Char	31	50	O		
Breed of the sire	**SBREED**	Char	81	3	O	I	**International**
Country of birth	**SIIDC**	Char	84	3	O	II	**identification**
Sex	**SSEX**	Char	87	1	O	III	**(Interbull)**
ID code of the sire	**SIID**	Char	88	12	O	IV	
Breed of the dam	**DBREED**	Char	100	3	O	I	**International**
Country of birth	**DIIDC**	Char	103	3	O	II	**identification**
Sex	**DSEX**	Char	106	1	O	III	**(Interbull)**
ID code of the dam	**DIID**	Char	107	12	O	IV	

Statuts	M ⇔ Mandatory, O ⇔ Optional
Note	
I	List of Breed code available at Interbull
II	List of Country code available at Interbull
III	Sex : M ⇔ Male, F ⇔ Female only
IV	Identification Number in Interbull Format : Alpha-numeric codes only, Right justified, Left blanks being filled wit
V	Date YYYYMMDD format

Annex 3: Description of the510 file: calf adjusted weight file

Field description	Label	Format	Start	length	status	Note	*Example*
File code	**FCODE**	**Char**	**1**	**3**	**M**		*510*
Breed of the animal	**BREED**	**Char**	**4**	**3**	**M**	**I**	*LMS*
Country of birth	**IIDC**	**Char**	**7**	**3**	**M**	**II**	*FRA*
Sex	**SEX**	**Char**	**10**	**1**	**M**	**III**	*F*
ID code of the animal	**IID**	**Char**	**11**	**12**	**M**	**IV**	*008795005065*
Twin calving	**TWIN**	**Int**	**23**	**1**	**M**	**V**	*0*
Embryo transfer	**EMBT**	**Int**	**24**	**1**	**M**	**VI**	*0*
Herd	**HERD**	**Char**	**25**	**20**		**VII**	*FRA000123456789*
Reference age	**REFAG**	**int**	**45**	**3**	**M**	**VIII**	*200*
Adjusted weight	**ADWW**	**int**	**48**	**3**	**M**	**IX**	*256*
Contemporary group	CG	Char	51	30	M*		*11*
Environment effect1	ENV1	Char	81	30	M*		*2*
Environment effect2	ENV2	Char	111	30	M*		*356*
Environment effect3	ENV3	Char	141	30	M*		*1238*
Environment effect4	ENV4	Char	171	30	M*		*12*
Environment effect5	ENV5	Char	201	30	M*		*4*
Environment effect6	ENV6	Char	231	30	M*		
Environment effect7	ENV7	Char	261	30	M*		
Environment effect8	ENV8	Char	291	30	M*		

Statuts	M ⇔ Mandatory; O ⇔ Optional ; M* ⇔ Mandatory for the environmental effect included in the national evaluation model.
Note	
I	See appendix A for Breed code
II	See appendix B for Country code – ISO 3166 Alpha-3
III	Sex : M ⇔ Male, F ⇔ Female only
IV	Identification Number in Interbull Format : Alpha-numeric codes only, Right justified, Left blanks being filled with Zero ("0")
V	0 = single birth ; 1 = twin or more birth
VI	0 = no ET ; 1 = birth from ET
VII	Herd identification corresponding to the herd included in the contemporary group; format Char3 country code + Char12 number
VIII	Reference age in days
IX	Weight in **Kilogram only**