Status on routine evaluation in NAV and Interbull

Gert Pedersen Aamand, Jørn Pedersen, Kevin Byskov, Anders Fogh and Ulrik Sander Nielsen

STØTTET AF
mælkeafgiftsfonden

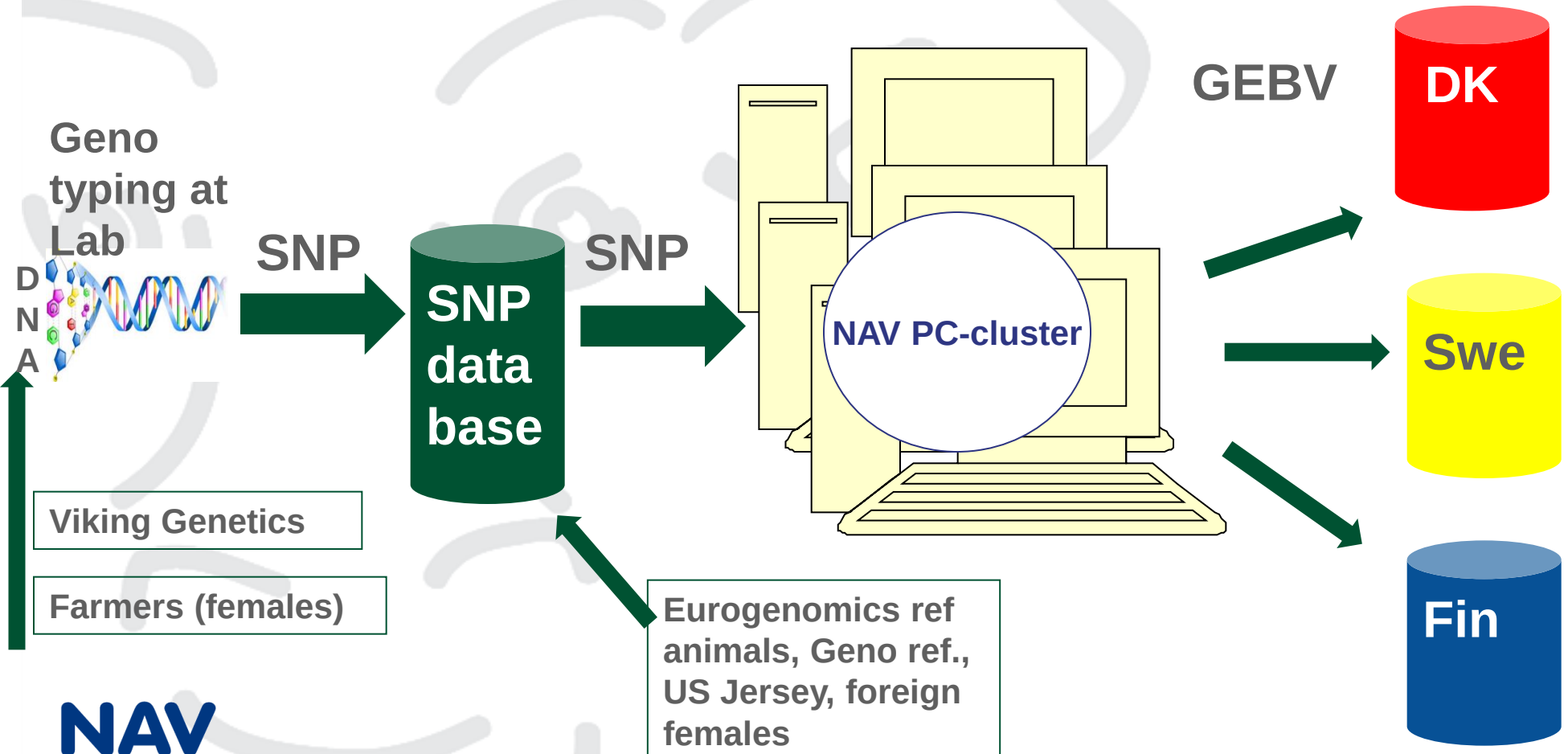
NAV routine

- **What are we doing today from DNA tissue to GEBV?**
- **Which steps have room for improvement in the short and longer run?**

Outline

NAV/national

1. Collection of DNA Tissue
2. Parentage verification
3. Exchange of genotypes
4. Imputation
5. DRP
6. Genomic prediction
7. Blending
8. Publication


Geno typing at Lab
DNA
SNP
SNP data base
SNP
NAV PC-cluster
GEBV
DK
Swe
Fin
Viking Genetics
Farmers (females)
Eurogenomics ref animals, Geno ref., US Jersey, foreign females

Collection of DNA

Collection of DNA - today

- Eartags
- Noose swap
- Blood
- Weak points
  - An extra operation
  - Correct link between tissue and animal id
  - An increasing challenge when number genotypes increase

Danish study

Future
Sample with minimal effort and maximal reliability!

Sampling - part of normal work flow

Unique connection between tag and sample (also checked at lab)

Learnings from Dansih pilot study 2012/13

- Tags are easy to use
- Important to do tagging systematically
- Bit more time consuming than ordinary tags correct top/bottom important

Status!

- Allflex has develop a new part to existing applicator that fixate liquid container
- Tested in 3 herds in September/October 2013
- Tissue seems to have good quality no matter how it is conserved
- BUT we recommend that tissues is stored in deep freezer or refrigerator

Ready for use in large scale!

Procedure for large-scale collection

Evaluation of procedure

- Use in 2 - 3 months (start December 2013)
- 5 - 10 herds (presently 8 herds)
- Feedback from Farmers and Genoskan

Positiv respons – more herds

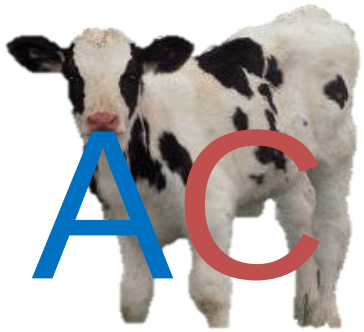
Negativ respons – revision of procedure

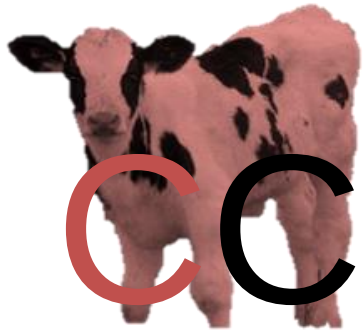
Parentage verification

Routine evalaution

- For most candidates we have a genotype of the sire, but not the dam
- Check for mendel errors
- Mendel errors indicate a disagreement between official pedigree and genotype

Principle behind Mendel Error Check





Mendel Error Check - Failed

- Rejected pedigree or sire is unknown / not genotyped
  - search for potential parents
- Technique similar as for ME-Check
  - Counting opposite homozygous loci
- Result is returned to VG or Växa
  - When errors are corrected genotype can be included again – farmers responsibility

How many genotypes are accepted?

- **% approved, rutine: 96-97 %**

Percent approved per breed, LD-project

- **JER: 94.5 %**
- **RDC: 97.9 %**

Errors are mainly ME

Percent approved per country, LD-project

- **DNK: 94.9 %**
- **SWE: 97.6 %**
- **FIN: 98.0 %**

% of errors correlated to herd size

Genotypes

Today

- In most cases we have a genotype of the sire, but not the dam – we do mendel error check
- Herds with many errors have to improve ear tagging/registration procedures

Genotypes

Near Future

- **MGS genotypes can be used to verify pedigree in some cases,**

BUT

- **It might create a puzzle (lots of manual work)**
- **Avoid an economic problem for a farmer if a calf suddenly has no dam (EU-support)**

Genotypes

Future

- Also a lot of (all) females genotyped
- Possible to assign parents automatically instead of traditional manual parentage verification
- Get everything running more smoothly, as automatic as possible, avoid bottlenecks

Reference group in routine evaluation Jan 2014

	HOL	RDC	JER
Bulls	25600	7800	1240+(1150 US)
Origin of ref animals	NLD, FRA, DEU, ESP	NOR	USA+CDN

Future reference group including cows?

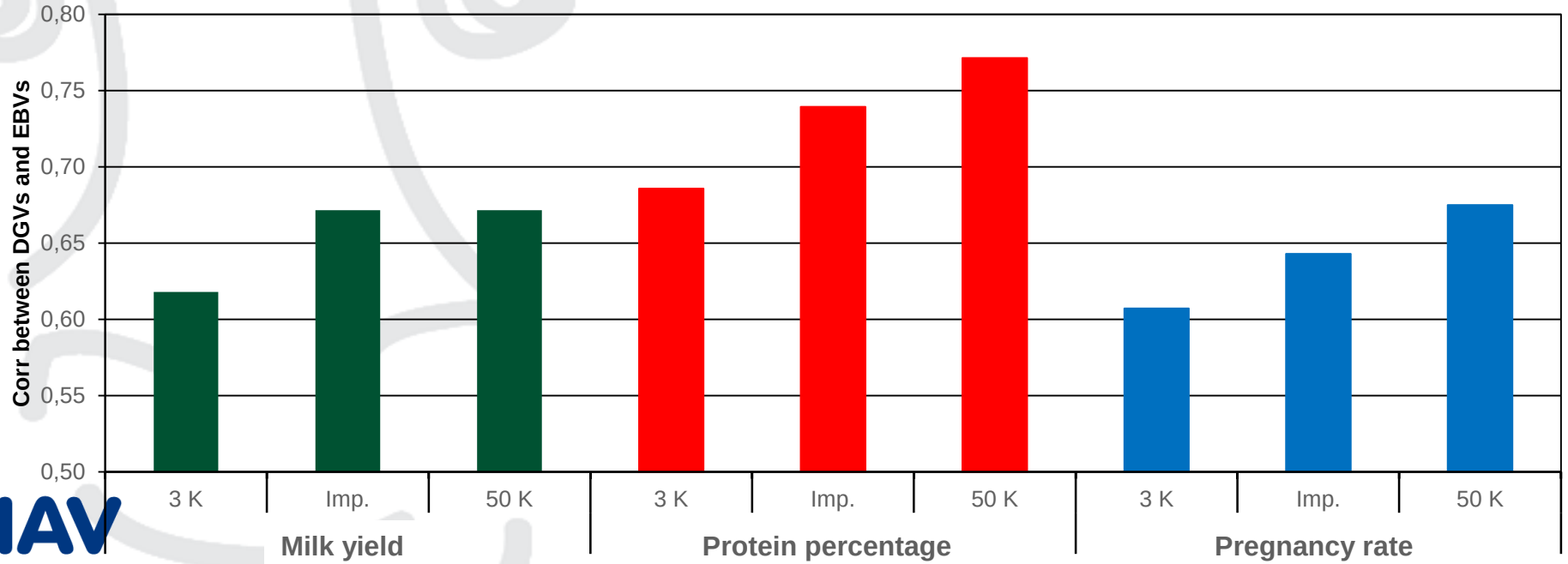
	HOL	RDC	JER
Bulls	25600	7800	1240+(1150 US)
Cows jan 2014 LD	0	3050	2300
Cows jan 2014 not LD project	5100	2350	760
Females total	11412	11938	5712

During 2014 amount of potential ref cows will increase significantly

Imputation

Why do we impute?

- Accuracy of GEBV
  - 54K > imputed LD > LD



American Jersey; modified after Weigel et al., 2010

Why do we impute?

- Cows (in referencegroup - RDC and Jersey) and heifers
  - Low density chip
- Illumina Bovine 54K chip
  - Version 1 and 2
- Eurogenomics (Holstein) og GENO (RDC)
  - NLD (HOL) and Geno (RDC) has previously used another chip
- Missing calls

Imputation

- When pedigree, sex, and ID is verified the genotype can be imputed
- HOL - FImpute
  - Quick & accurate in homogeneous populations
- RDC & JER - Beagle
  - Slow but accurate and robust

How accurate can we impute Genotype error rates

Animals genotyped in 2013

	00	01	10	11
HOL FImp LD	3.4 (67)	1.7 (14)	3.2 (1213)	0.7 (432)
RDC* Fimp LD	10.3 (9)	3.7 (3)	4.3 (1147)	0.7 (234)
RDC* Beagle LD	2.3 (9)	1.0 (3)	1.3 (1147)	1.0 (234)
RDC* Beagle LD Male/No Geno	3.0 (9)	3.7 (3)**	1.4 (1147)	1.1 (234)**
JER FImp LD	3.0 (5)	2.0 (3)	1.5 (110)	0.5 (75)
Jer Beagle LD	Not investigated, but we expect Beagle will perform better than Fimpute			

00 = No parents genotyped/parents unknown in pedigree
01 = Only dam genotyped
10 = Only sire genotyped
11 = Sire and dam genotyped

Imputation future

- **Impute all SNPs**
  - **Today 6K SNPs from LD is imputed to 54 K**
  - **4K from LD chip not available at 54K is not used**

Additional gain limited

Deregressed proofs

Used for Estimating direct genomic breeding values and Blending

Routine DRP

Sire MGS pedigree used

For reference animals

- **DRP for DFS bulls (NAV EBVs) and foreign ref bulls (Interbull EBVs)**
- **DRP for genotyped bulls having official EBVs based on a progeny test**

For blending:

- **All reference animals plus genotyped cows**

DRP - future reference

- **Single cows are only informative if EBV's are calculated by an Animal model, not a Sire model (Today SM for other disease, Calving traits and fertility)**
- **Apply AM pedigree instead of Sire-MGS pedigree in deregression**
- **Include all animals in degression**

DGV prediction

DGV prediction

- GBLUP
  - Same genetic parameters as in traditional model
  - No polygenetic effect
  - DGV's scaled according to validation results to get rid of inflation

DGV prediction

Observed:

- **Underprediction of genetic trend for candidates of Jersey and RDC – research focus**

Near future:

- **Get rid of underprediction**
- **Polygenetic effect might be considered again (but research have shown minimal effect) and substitute current scaling**

Future:

- **One step – all information simultaneously, selection (but today still bias problems)**

Preliminary results of inclusion of females in reference group

	RDC	JER
Production	X	XXX
Udder health	XX(X)	XXX
Body	XX	XXX
Feet and legs	-	X
Mammary	XXX	XXX
Milking speed	XXX	XXX
Temperament	-	-

XXX ≈ 5 pct. units

Note validation include daughters of proven sires only

Blending

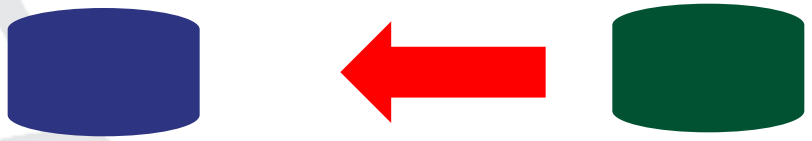
Blending – routine (GEBV) (Mantysaari, 2010)

Inputs:
DRPs genotyped animals

Inputs:
Scaled DGV + DGV reliabilities

Phenotypic information

Genomic information



Observation: Challenge to avoid double counting of information

Blending – future (GEBV)? (Taskinen & Mantysaari 2013)

Inputs:
DRPs genotyped animals

Inputs:
Scaled DGV + DGV reliabilities

Phenotypic information

Genomic information

Can we better avoid double counting of information?

One step (research focus)

All information utilised in the same step:

- **No risk of double counting**
- **Optimal use of information**
- **Handling selection**

Frequency and publication

Prediction

- 4 times a year based on ”new” phenotypic data and ”new” genotypes
- 8 time a year ”old” phenotypes and ”new” genotypes

Publication

- Monthly – females
- 4 times a year bull > 17 month (10 month in 2014?)

Summary

1. Collection of DNA Tissue (ear tagging)
2. Parentage verification (assigning of parents)
3. Exchange of genotypes (US jersey, candidates)
4. Imputation (limit gain in near future)
5. DRP (AM-pedigree in DRP step, all traditional models AM)
6. Genomic prediction (solve JER, RDC problem, polygenetic effect?, onestep solve bias)
7. Blending (new method?)
8. Publication (lower publication age bulls)

Summary

1. Collection of DNA Tissue (ear tagging)
2. Parentage verification (assigning of parents)
3. Exchange of genotypes (US jersey, candidates)
4. Imputation (limit gain in near future)
5. DRP
6. Genomic prediction
7. Blending

One step

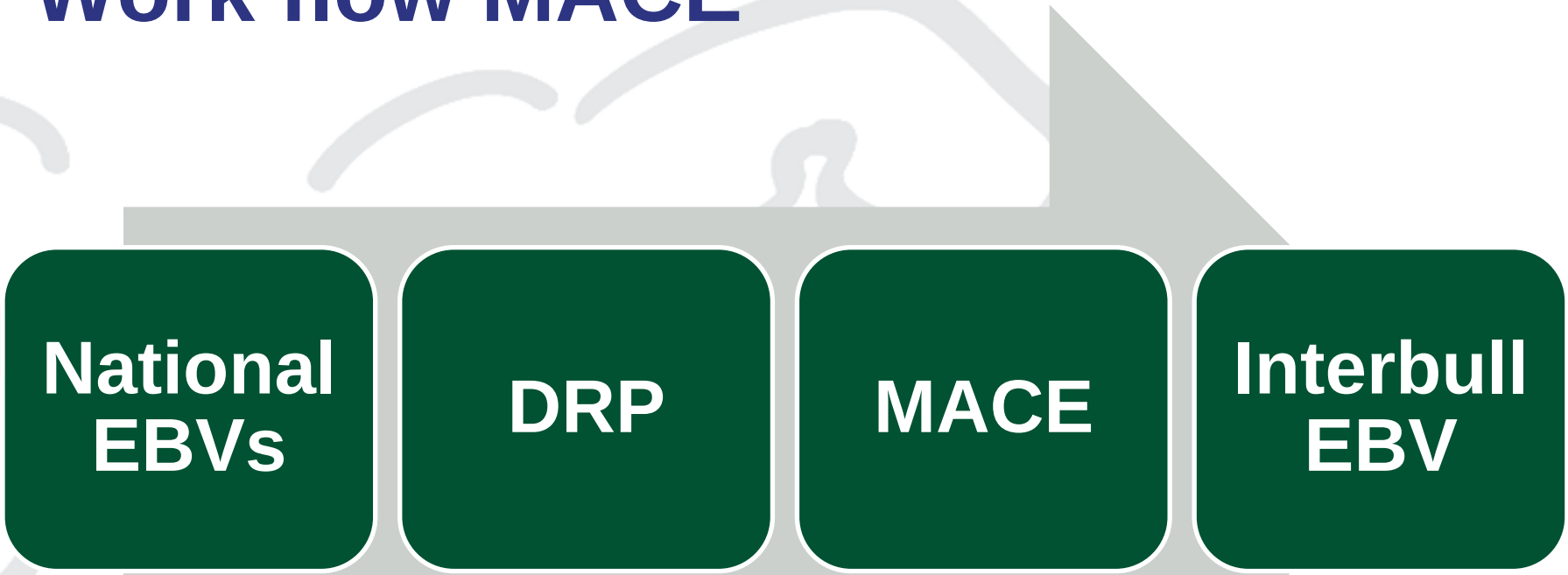
8. Publication (lower publication age bulls)

Outline

Interbull

1. GMACE
2. International reliabilities

Work flow MACE



National EBVs are subtracted the pedigree information and degressed – national reliabilities/EDCs used
Little variation in national EDCs

Work flow GMACE

National EBVs are subtracted the pedigree information and degressed – national reliabilities/EDCs used
National reliabilities have relatively large variation

GMACE status and plans

- February 2013 test run
- August 2013 (Implementation run =countries allowed to but not forced to publish)
- September 2013 test run
- December 2013 Implementation run
- End December 2013 – two new pilot runs (January 2104 national inspections ongoing)
- End January and early February IB ITC and SC meetings (decisions about routine)

GMACE status and plans

- End January and early February IB ITC and SC meetings (decisions about routine)
- February 2014 Test run
- April 2014 Routine or implementation run
- August 2014 Routine or implementation run

Interbull ongoing activities

- **International working group formed in August 2013 has to come up with a standard for calculating reliabilities before May 2014 (Martin Lidauer. MTT is in the group)**

General picture from GMACE results from test and implementation runs

- NAV finds unexpected high STD of IB GEBV on NAV scale for more traits
- The most extreme case is clinical mastitis
- NAV has communicated the findings to IB

Simple statistics. Young bulls birth year 2010 or later. Udder health (CM)

	Number	CM DFS	SD CM	REL CM
CAN	3017	97.2	7.7	56.8
DEU	6547	98.6	7.9	55.8
DFS	1601	101.2	7.6	57.6
FRA	7541	97.5	7.1	53.0
GBR	1004	98.2	7.8	56.1
ITA	1456	98.0	7.8	58.0
NLD	5093	98.8	6.9	51.2
US	28408	101.9	7.7	54.9

Very unexpected results in RED

Ratio MACE or GMACE results
STD (G)EBV national/STD (G)EBV DFS scale

Country	MACE EBV birth year 2006	GMACE GEBV birth year 2010
CAN	0,45	0,51
DEU	0.55	0.56
DFS	1.00	1.00
FRA	0.54	0.52
ITA	0.51	0.60
NLD	0.57	0.58
USA	0.49	0.58

We hope the IB pilot runs have solved the problems

Expected to be the same given progeny group size and reference group are about the same DFS versus foreign countries

Interbull ongoing activities

- Investigate effect on national assumption about GEBV reliabilities on GMACE results
- International working group has to come up with a standard for calculating reliabilities before May 2014

From country comparisons involving DFS, FRA, NLD, DEU

- Ranking of animals: NAV-DEU, NAV-NLD and DEU-NLD have same size of correlations
- NAV rank animals as efficient as other countries - IB validation reliabilities the same.

GMACE

- **Hope Interbull are able to estimate routine international genomic breeding values during 2014**