

Impacts of genomic pre-selection on classical genetic evaluations

Vincent Ducrocq, Clotilde Patry



Overview

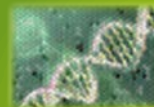
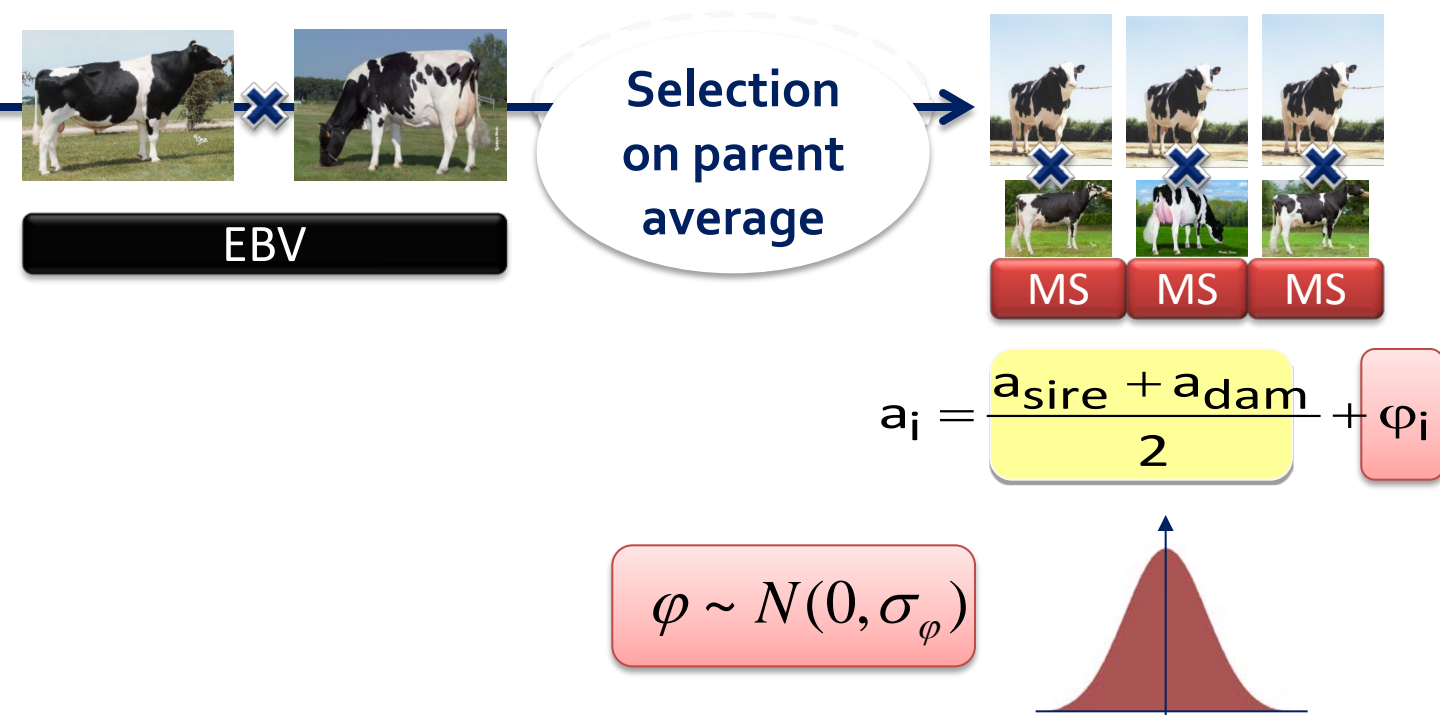
- ❑ Theoretical considerations suggest that genomic pre-selection of progeny-tested bulls leads to biases in “classical” genetic evaluations
- ❑ A first presentation in February 2009
- ❑ A study in three parts



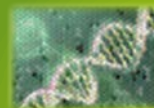
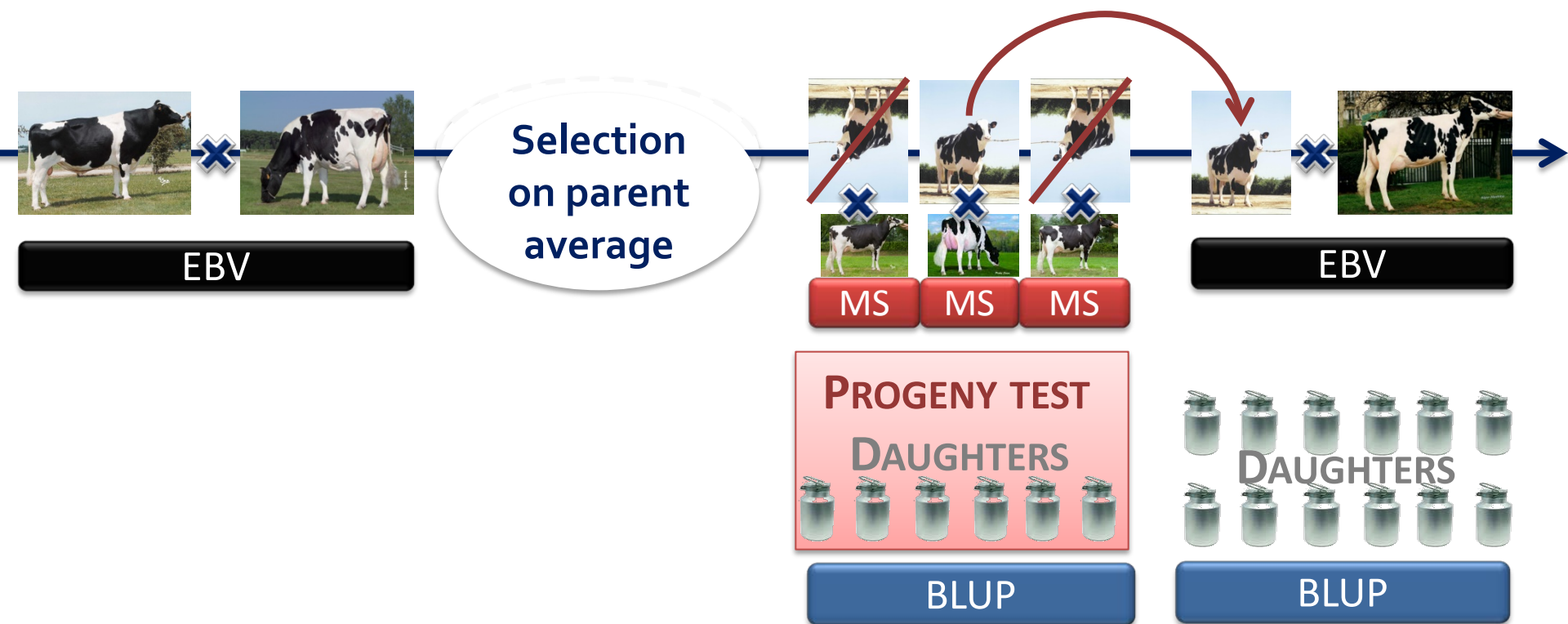
- ❑ Conclusions



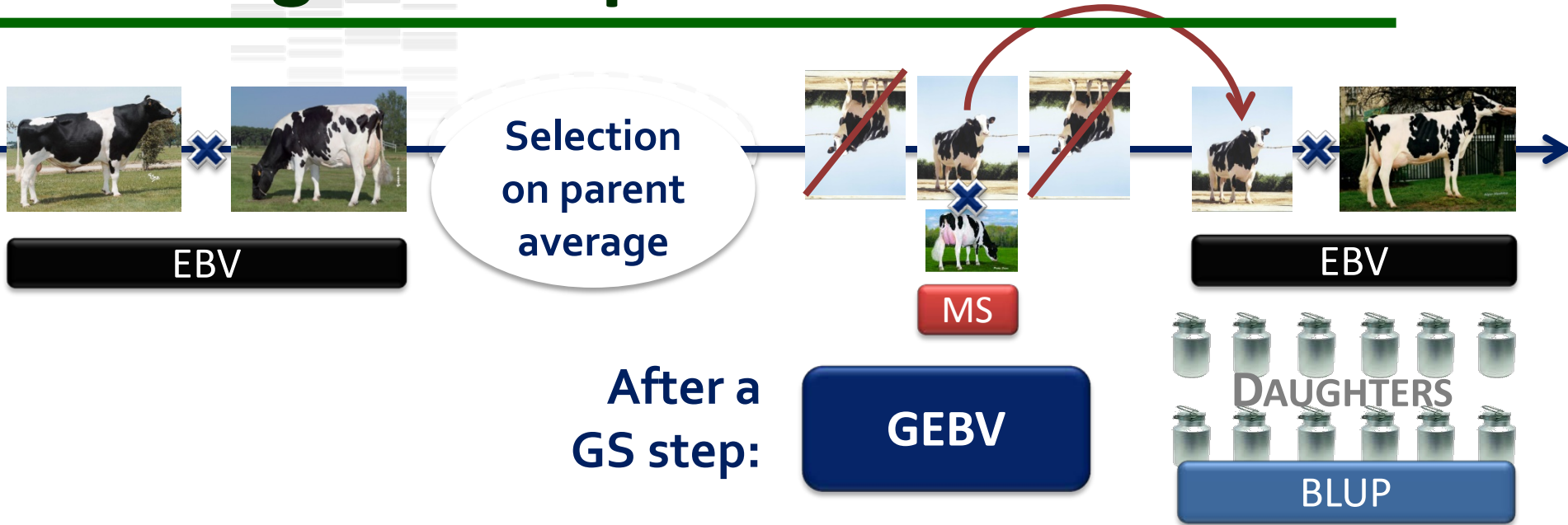
Mendelian sampling estimation: classical evaluation after progeny test



Mendelian sampling estimation: classical evaluation after progeny test



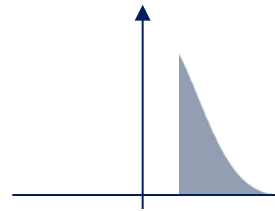
With genomic pre-selection



After a
GS step:

GEBV

**Violation of BLUP
assumption on MS
distribution**



~~$\varphi \sim N(0, \sigma_\varphi)$~~

?

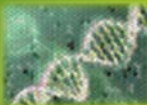


How to assess bias in national evaluations?

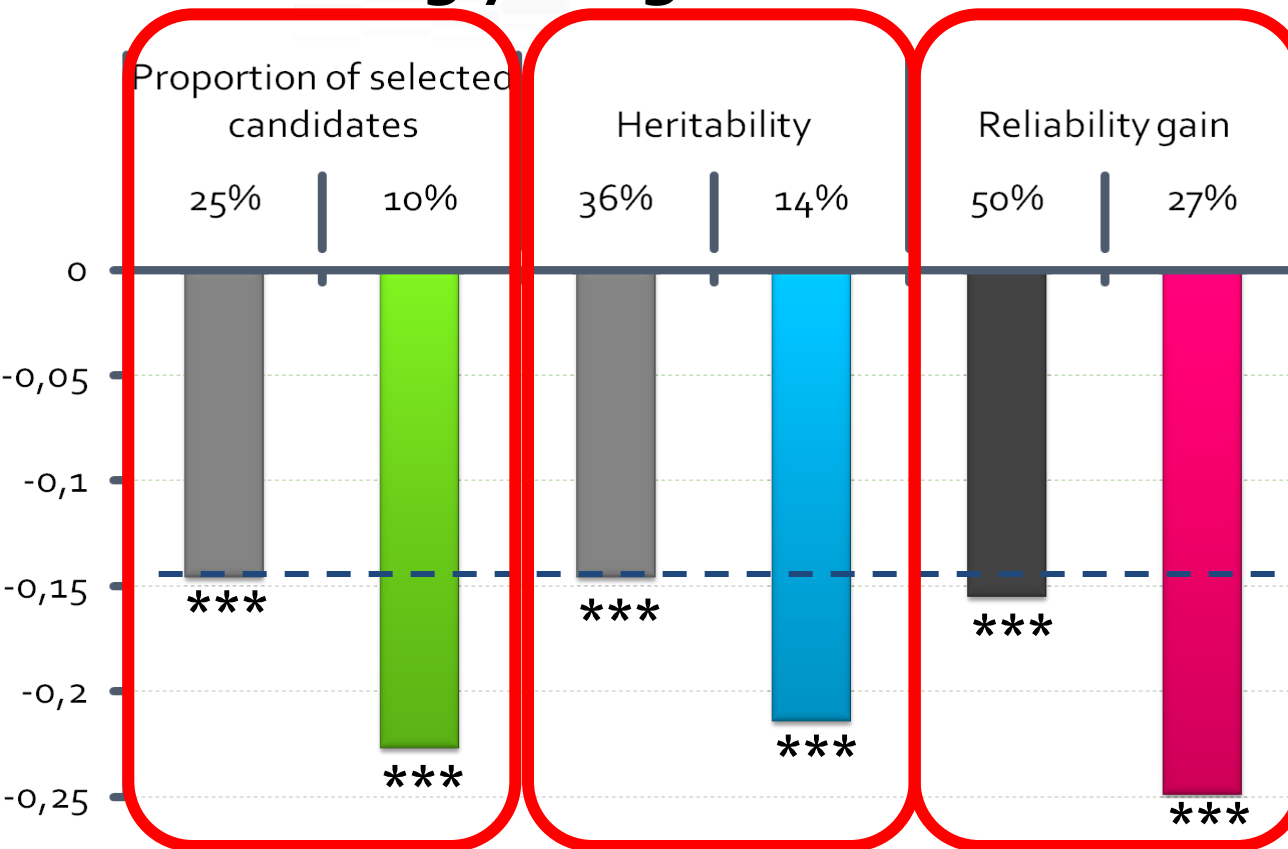
**Patry, C. and V. Ducrocq. 2011.
Evidence of biases in genetic evaluations due to genomic
preselection in dairy cattle.
J Dairy Sci 94:1011-1020**



- Holstein breed – data for French genetic evaluation
 - Joint simulation of TBV and GEBV
 - Influence of various factors on bias:
 - **Proportion of selected candidates:** 10% or 25%
 - 2 type traits => 2 levels of **heritability:** 14% or 36%
 - **genomic gain in reliability** : 50% or 27%
- ⇒ 5 scenarios of interest
- 50 replicates / scenario
 - **Bias = E(EBV - TBV)**



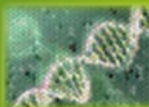
Bias among young sires after a GS step (in σ_a)



Generally, bias magnitude ranges from:



- 15 to 25% σ_a among YS
- 4 to 11% σ_a among their daughters

$E(EBV-TBV) < 0$ among young sires and their daughters



Accuracy of BLUP evaluations after PT or GS among YS

Scenario	Theoretical R^2	Observed $\rho^2(\text{TBV}, \text{EBV})$	MSE =var (EBV - TBV) + bias ²
After Progeny Testing	81.5%	75.6%	0.183
After a Genomic Selection step	81.5%	72.7%	0.188

From the simulations: biased EBV + a reduced accuracy

→ Need to account for a GS step in national evaluation model



How to reduce / eliminate the bias in national evaluations?

Patry, C. and V. Ducrocq, 2011.

**Accounting for genomic pre-selection in national BLUP
evaluations in dairy cattle.**

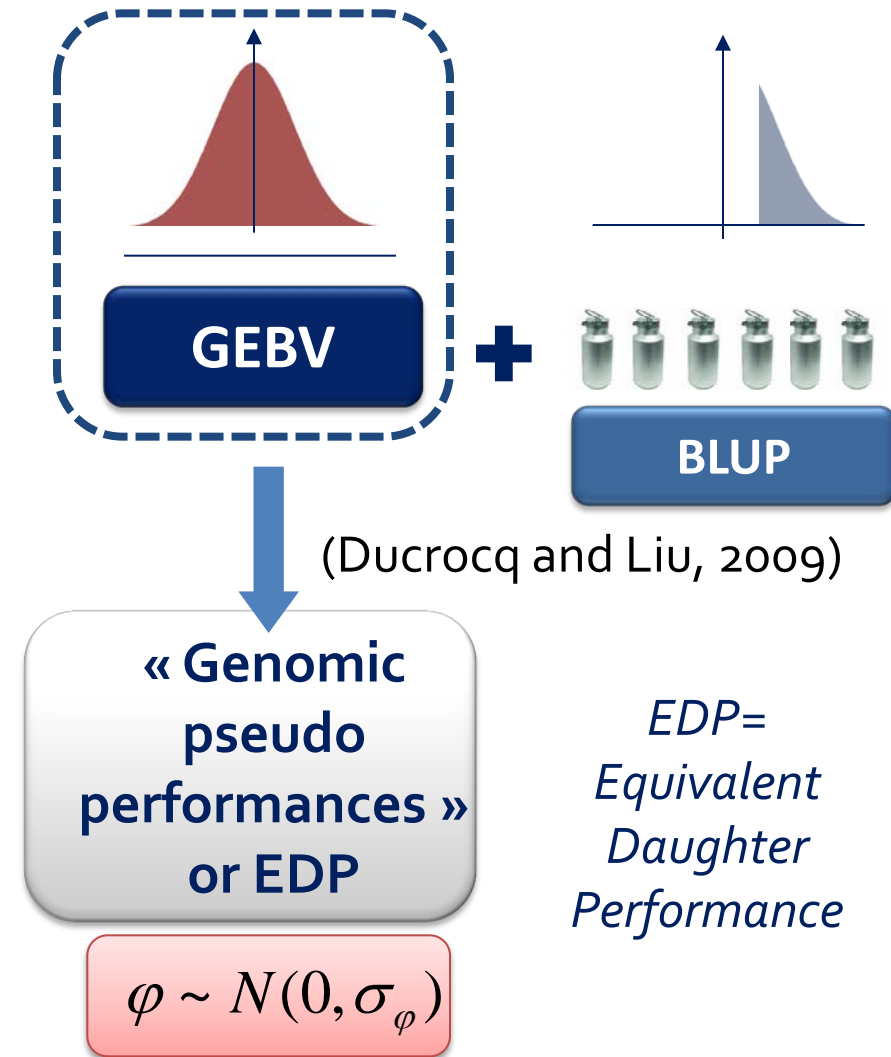
Genet Sel Evol 43:30



« All data on which selection is based should be included in the evaluation »

=

Data on culled and selected candidates



« All data on which selection is based should be included in the evaluation »

=

Data on culled and selected candidates

Mixed model equations :

$$(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \alpha\mathbf{A}^{-1})\hat{\mathbf{a}} = \mathbf{Z}'\mathbf{R}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$

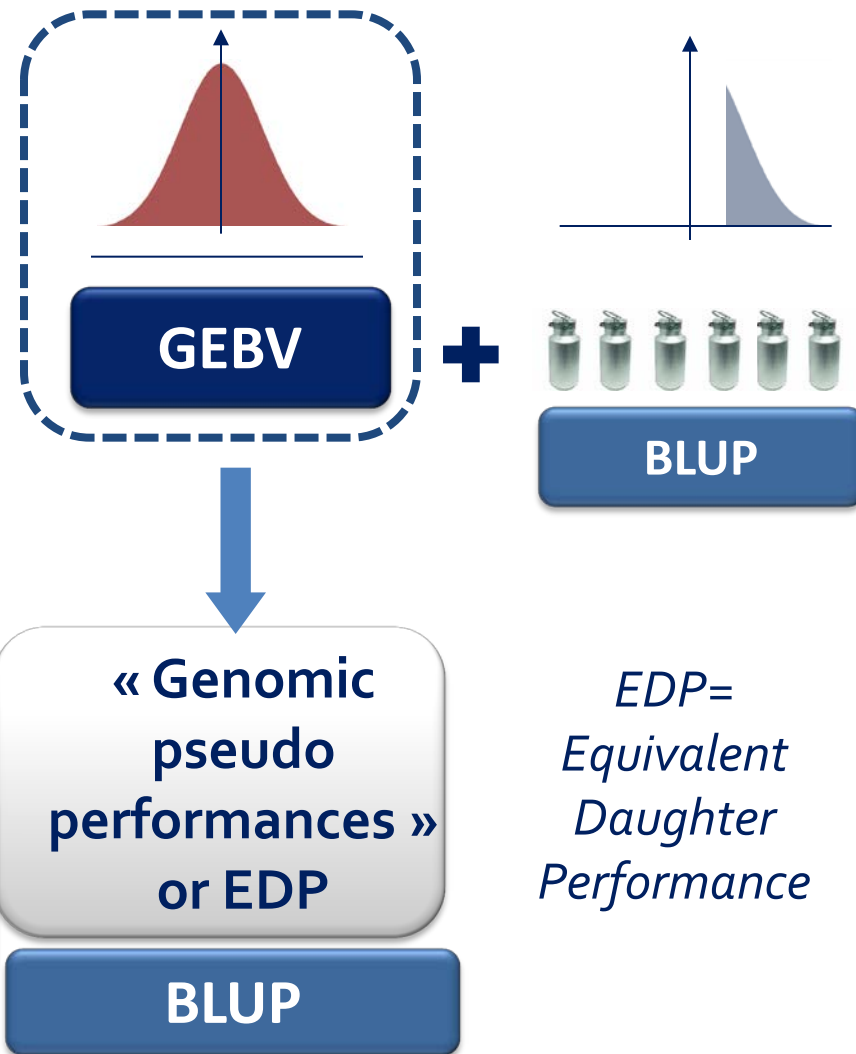
WEIGHT

1) gEDC
 $=f(h^2, \Delta R^2_{\text{genomic}})$

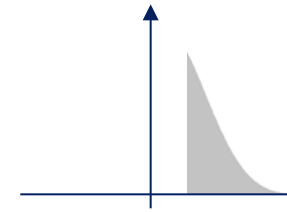
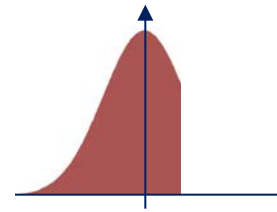
EDC = Equivalent Daughter Contribution

PERFORMANCES

2) EDP= de-regressed GEBV



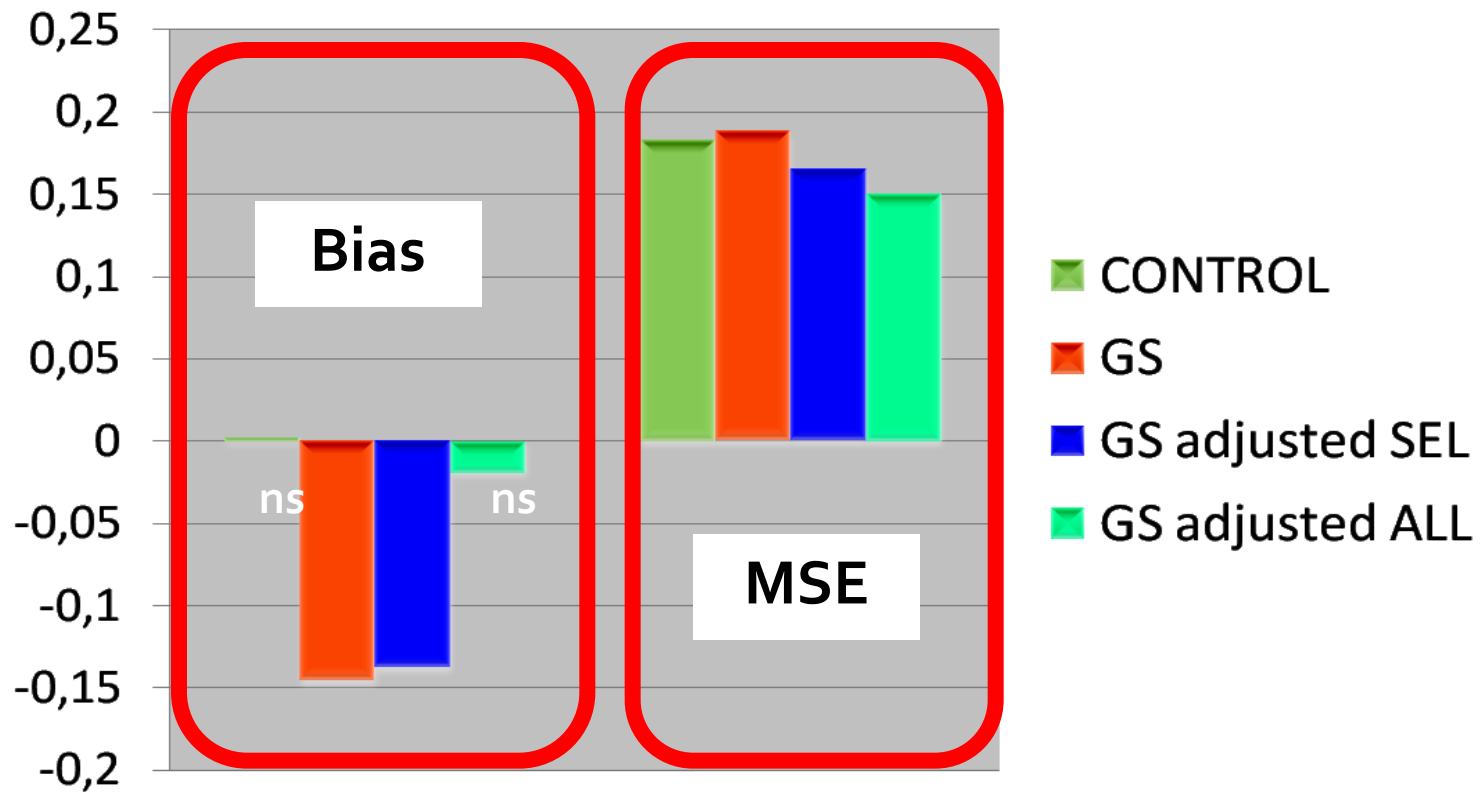
- On real data, same simulation framework
- 2 additional scenarios:



GS step	Scenario	EDP - culled YS -	EDP - selected YS -
No	Control (after PT)	No	No
Yes	Biased (after GS)	No	No
Yes	Adjusted for GS - SEL	No	Yes
Yes	Adjusted for GS - ALL	YES	Yes



Measures of **bias** = $E(TBV - EBV)$ and **MSE** = $\text{var}(TBV - EBV) + \text{bias}^2$
after standardization of EBV and TBV
for type trait ($h^2=36\%$) - selection rate = 25% - Young Sires



Bias adjustment

Including genomic pseudo-performances **corrects bias**

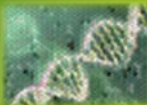
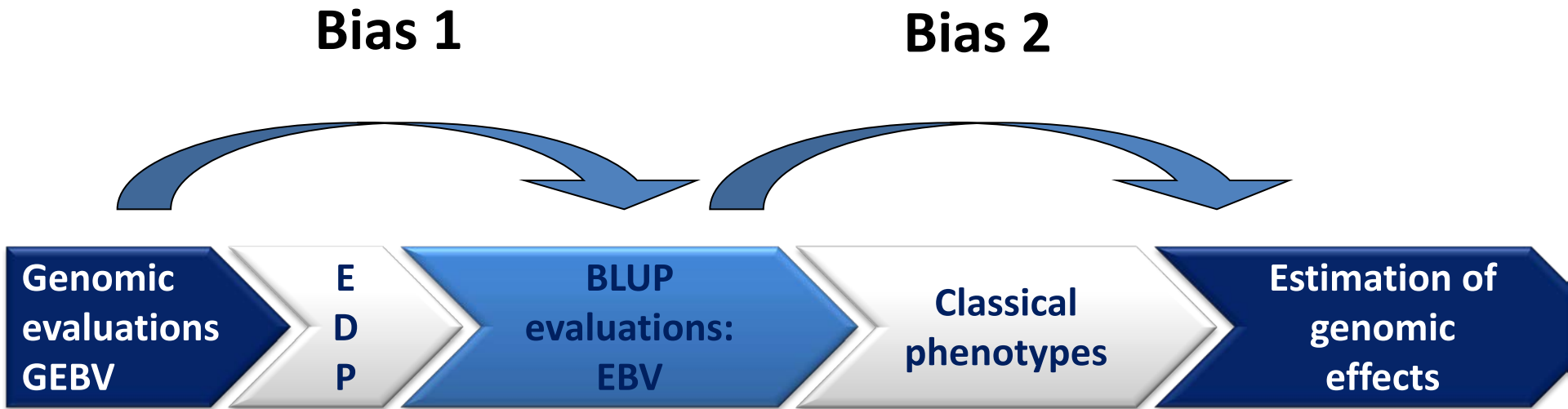
= one way to combine genomic, phenotypic and pedigree-based information BUT...

- 1) Weight given to genomic information** → *overestimated gEDC*
- 2) Double-counting** of genomic information in classical EBV: *once genotyped, the genotype of relatives does not add any information*
- 3) Dependency** between classical and genomic evaluations

Better alternatives exist (see next talks) !



Interdependency of evaluations





How genomic selection can affect international evaluations?

Patry, C., H. Jorjani and V. Ducrocq, 2013
Impact of pre-selected and biased national BLUP evaluations on international genetic evaluations, J Dairy Sci. 96, 3272-3284



4 situations at Interbull level for international evaluations (MACE):

Country strategies	Proofs for		Status	MACE issues
	Culled YS	Selected YS		
Progeny Test	EBV	EBV	Unbiased	Complete and correct
Genomic Selection	no	GEBV	Biased	Incomplete and incorrect
GS + adjustment	no	GEBV	Unbiased	Incomplete but correct
GS + adjustment	GEBV	GEBV	Unbiased	Complete and correct

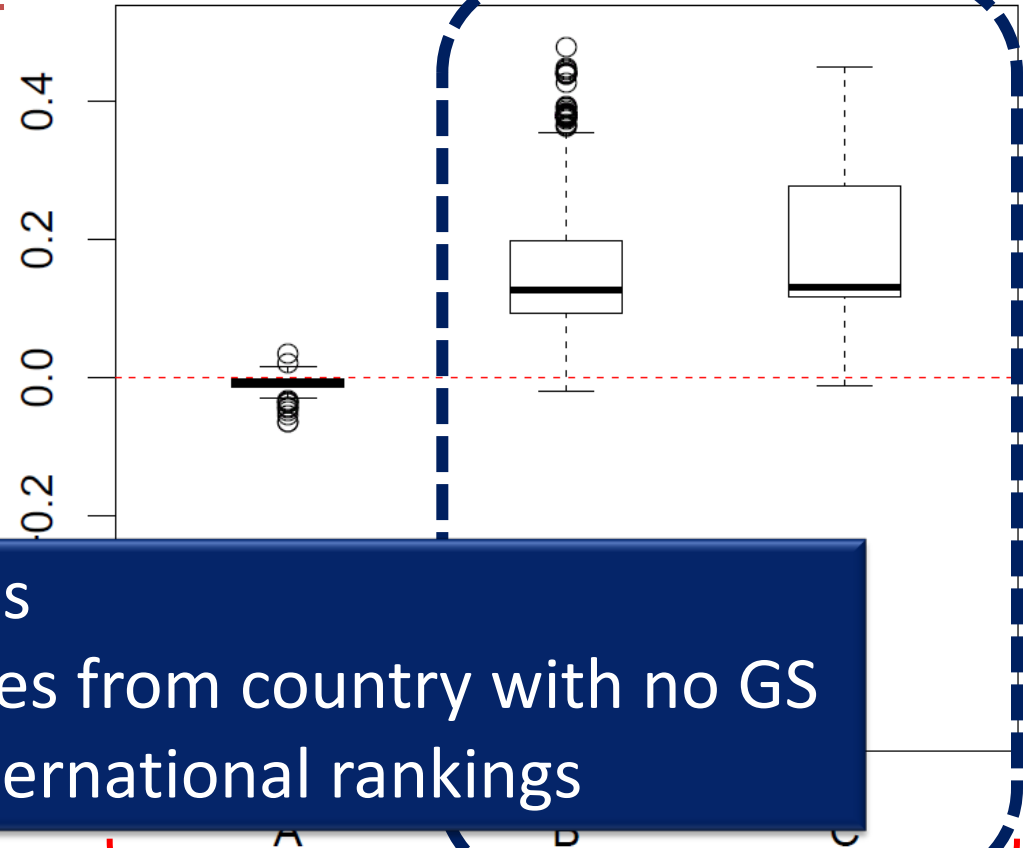


- Real data from 3 large countries: A, B, C
- Mimicking GS:
 - **No actual GEBV but EBV used as proxy of GEBV to compute MS estimates**
 - Selection based on MS estimates, within half-sib families
 - Genomic selection effects: **delete national proofs for “culled” young sires**
- Measure of bias among Young Sires
 - = E[MACE solution (PT) – MACE solution (GS)]**
 - By country of origin: young sires from A, B or C
 - On each scale: domestic versus foreign scales

Country A sends **pre-selected data** to Interbull

Effect of incomplete national data

**Bias (in σ_a)
On the domestic (A) scale**



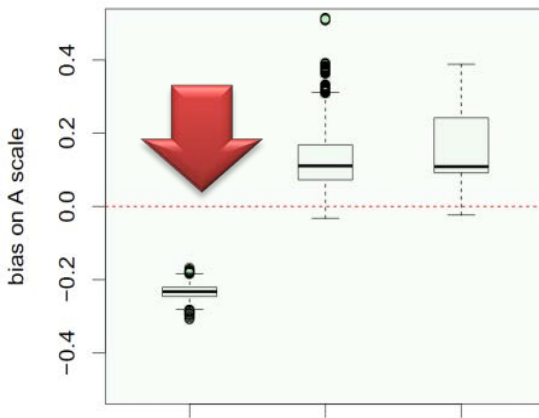
Among young sires by country of origin

- 1) Penalization of A sires
- ⇒ Consequences on sires from country with no GS
- ⇒ Consequences on international rankings

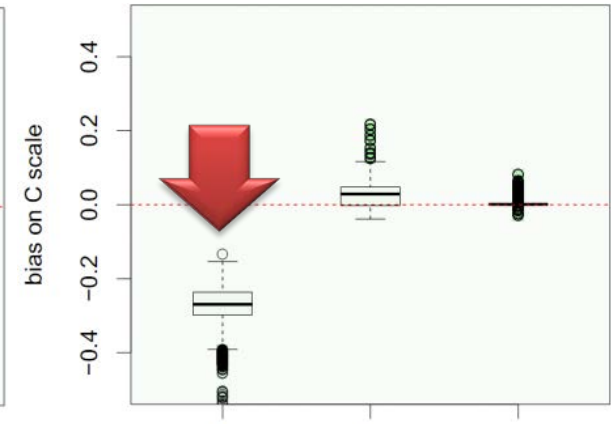
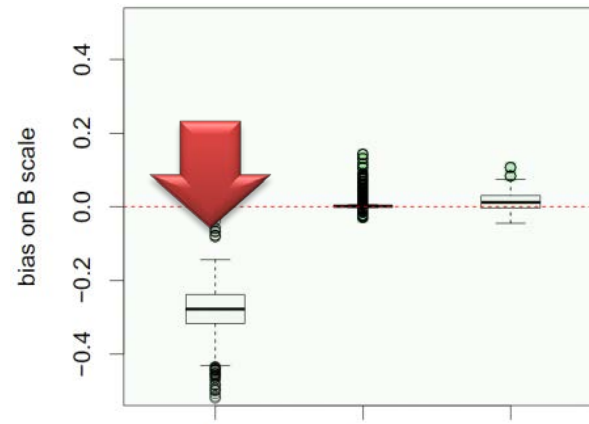
Effect of **incomplete and incorrect** national data

- Only 1 country implementing GS : A
- Without adjustment of national evaluations: A sends biased EBV

Bias on A scale



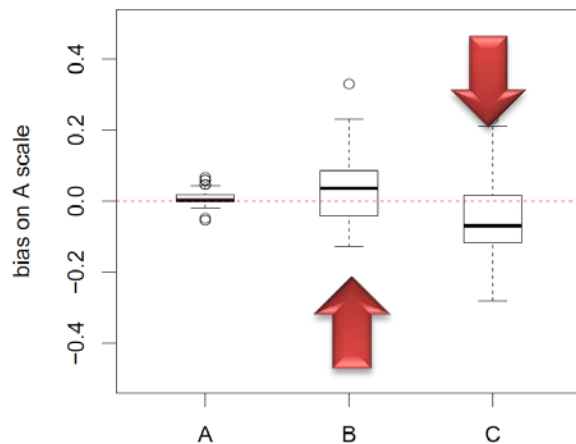
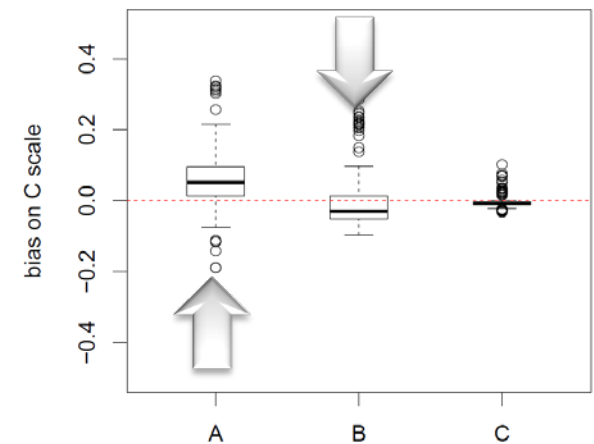
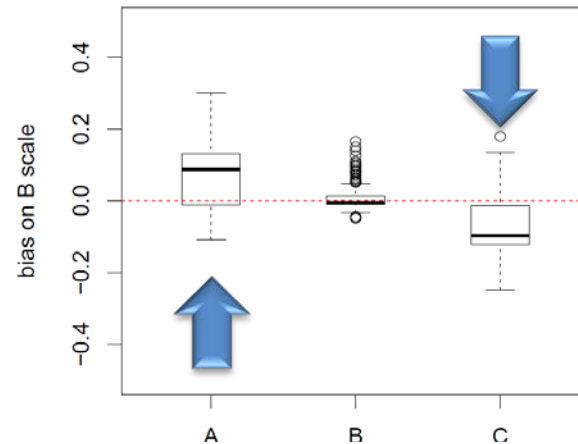
Bias on B and C scales



- 1) Penalization of A sires
- 2) Propagation of bias on all scales = impact on countries without own GS

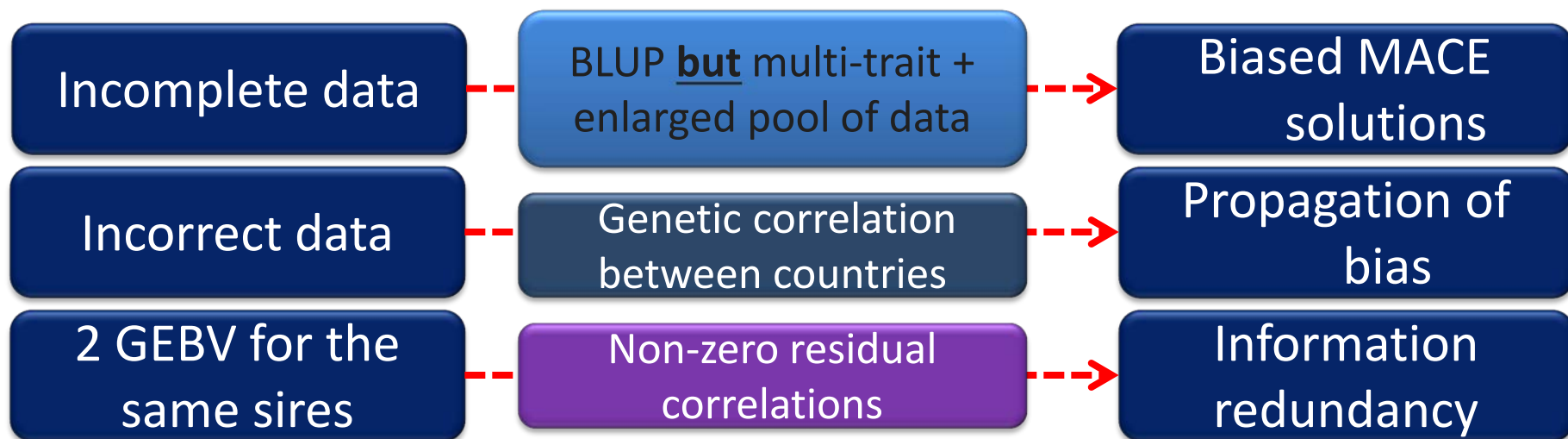


- A, B and C send incomplete data to Interbull

Bias on A scaleBias on B and C scales

- 1) Penalization of A sires
 - 2) Propagation of bias on all scales
 - 3) Difficult to predict the direction/magnitude of bias
- But international re-ranking is certain!**

- In summary:



Final remarks

- How to avoid such a large and widespread bias ?
- **Share all information on the selection process (!?)**
- At national **and** international levels
- Opportunities at national level
- Need to adapt routine evaluations ...**before daughters of genomically selected sires have records**
- At international level?
- **Develop tests to « validate » unbiasedness of national EBV ?**

All this is included in a more complex issue : properly combine genomic and classical information into GEBV => **not only maintain unbiased genetic evaluations but increase their accuracy**



Acknowledgments

