

De-regressing MACE versus domestic EBV for Genomics

P.G. Sullivan, J. Jamrozik & G.J. Kistemaker

Canadian Dairy Network

What is de-regression?

Reversing the BLUP regression of \hat{g} (EBV) towards parent average (PA)... How?

➤ **Scale** \hat{m} to increase $V(\hat{m})$

$$\hat{g} = PA + \hat{m}$$
$$DRP = PA + \frac{\hat{m}}{REL_{EDC}}$$
$$V(DRP) > V(\hat{g})$$
$$V(DRP) = V(y) ???$$

Did we scale properly?

- With our **approximations**, did we apply the **correct scaling** with $\frac{\hat{m}}{REL_{EDC}}$?

1. Were EDC approximated correctly?

$$REL_{EDC} = f(EDC_{bull})$$

2. What about EDC of parents?

$$REL_{\hat{m}} = f(EDC_{bull}, EDC_{sire}, EDC_{dam})$$

$$\hat{\sigma}_m^2 = \sum \hat{m}^2 / \sum REL_{\hat{m}}$$

Did we scale properly?

- With our **approximations**, did we apply the **correct scaling** with $\frac{\hat{m}}{REL_{EDC}}$?

3. Complications transforming MT to ST?

- MTEDC of **bull** versus **bull+sire+dam**?
- Is $REL_{\hat{m}}$ the same from correlated traits?

$REL_{\hat{m}}$ from correlated traits

Given a set of: REL_{bull} , REL_{sire} , REL_{dam}

Is $REL_{\hat{m}}$ the same for any data source?

Tested with a 2-trait x 3-animal MME:

Data Set	Domestic data			Foreign (corr'd) data		
	Bull	Sire	Dam	Bull	Sire	Dam
Domestic (D)	#	#	#	-	-	-
Foreign Low (L)	-	#	-	##	-	##
Foreign High (H)	-	-	-	##	##	##

$REL_{\hat{m}}$ by Data Set: **H > L > D**, esp. if low r_g

Did we scale properly?

- With our **approximations**, did we apply the **correct scaling** with $\frac{\hat{m}}{REL_{EDC}}$?

3. Complications transforming MT to ST?

- MTEDC of **bull** versus **bull+sire+dam**?
- Is $REL_{\hat{m}}$ the same from correlated traits?

NO!

Did we scale properly?

- With our **approximations**, did we apply the **correct scaling** with $\frac{\hat{m}}{REL_{EDC}}$?

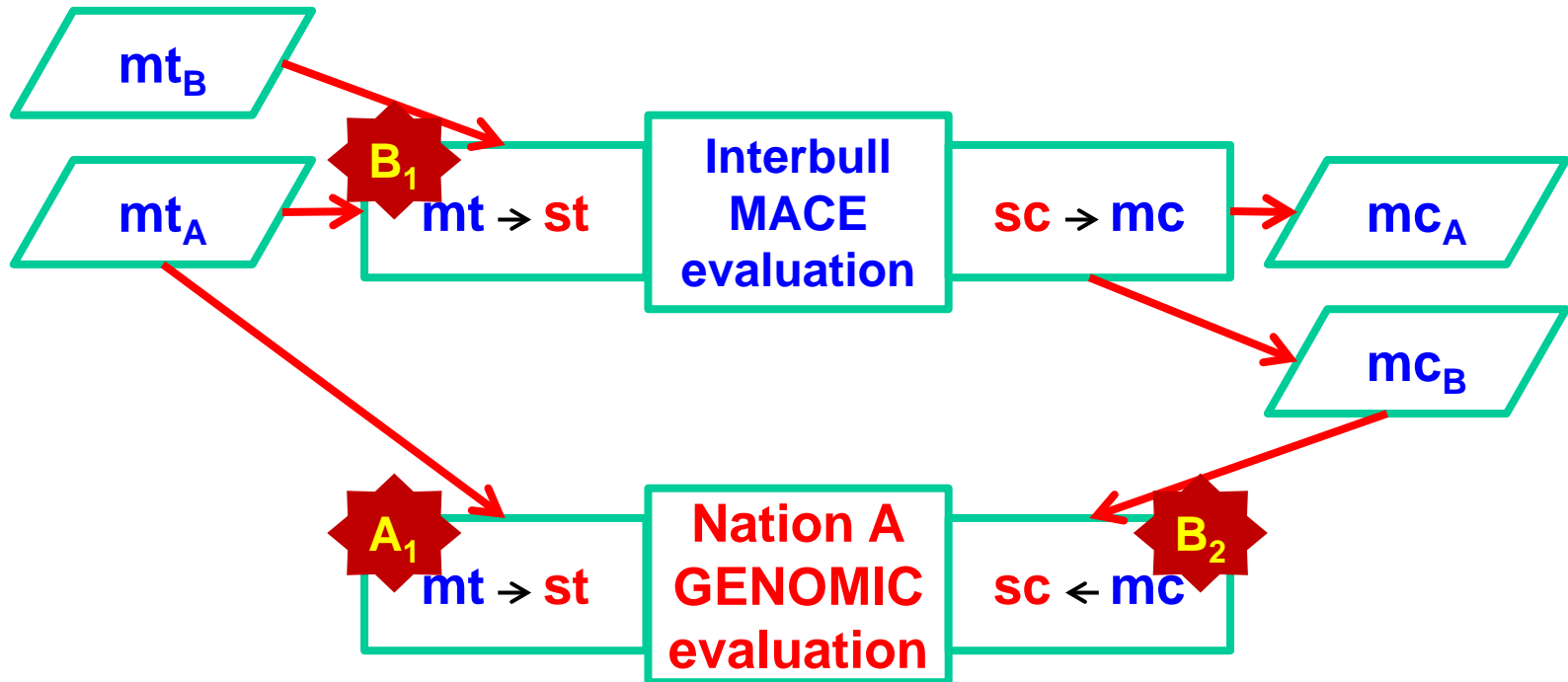
3. Complications transforming MT to ST?

- MTEDC of **bull** versus **bull+sire+dam**?
- Is $REL_{\hat{m}}$ the same from correlated traits?

NO!

4. Can we **transform** from multi-trait to single-trait information **TWICE** ? ...

Double Transformation



mt = multi-trait EBV
st = single-trait EBV

mc = multi-country EBV
sc = single-country EBV

Did we scale properly?

- **Probably not... $V(\text{SNP})$ usually too high**
 - MT → ST and REL_{EDC} vs $REL_{\hat{m}}$
- **Young bulls affected more than proven**
 - **Inaccurate** comparisons between top **young vs proven** bulls
- **Larger errors for Foreign bulls**
 - **Inaccurate** comparisons between top **foreign vs domestic** bulls, both young and proven

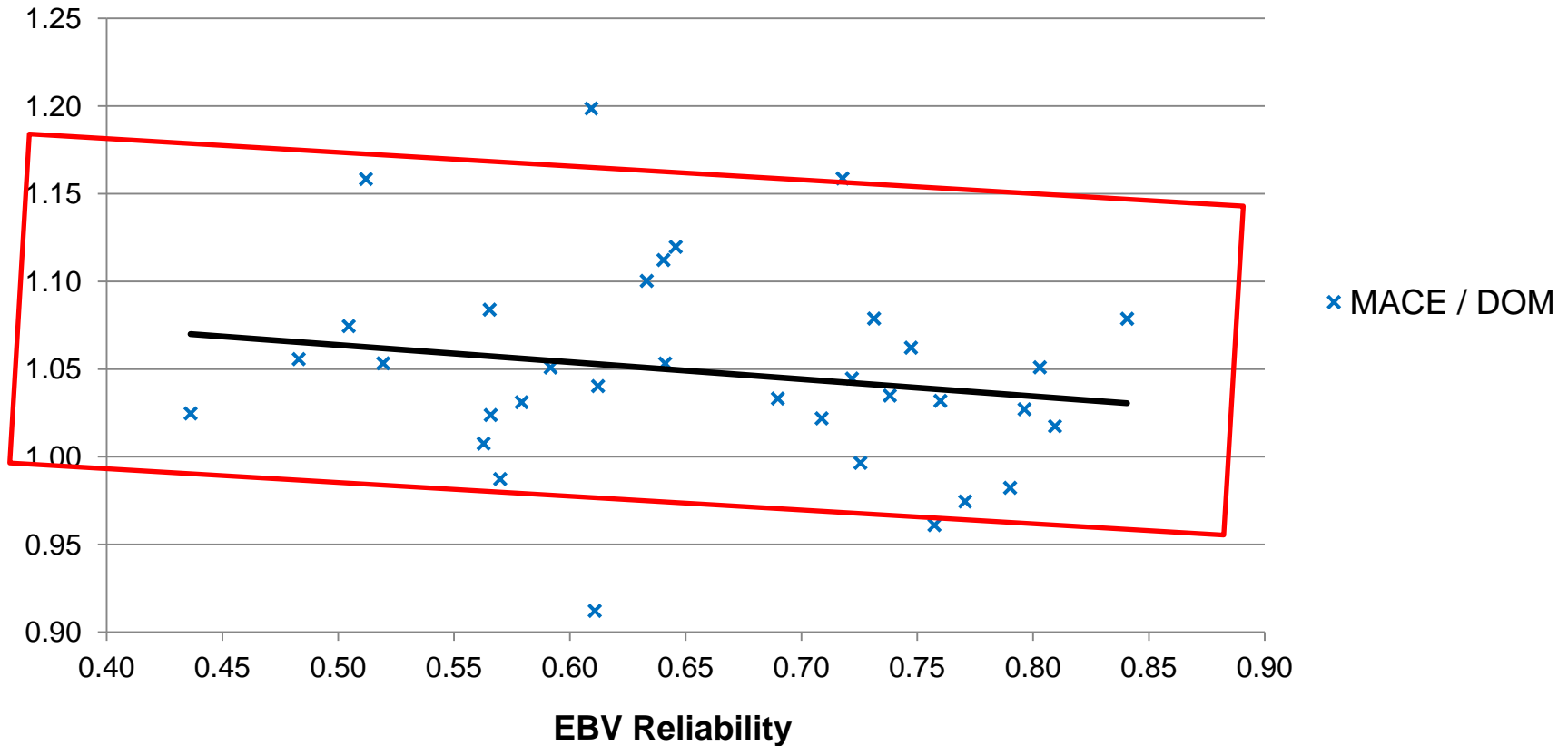
Heterogeneous V(DRP)

- **Genomic model does not distinguish between domestic versus foreign reference bulls**
 - i.e. model assumes MACE proof = Domestic proof
- **Estimate:** $\sigma_m^2 = f(\hat{m}, REL_{\hat{m}})$
 1. From DRP of MACE bulls (double-transform.)
 2. From DRP of DOMESTIC bulls (single-transform.)
- **Update scaling of \hat{m} for MACE bulls to match scaling of DOMESTIC bulls**
 - REML estimates account for Rel'y of bull and parents
 - Scaling more consistent for MACE vs DOMESTIC

MACE vs Domestic Data

(SD estimates Dec 2014)

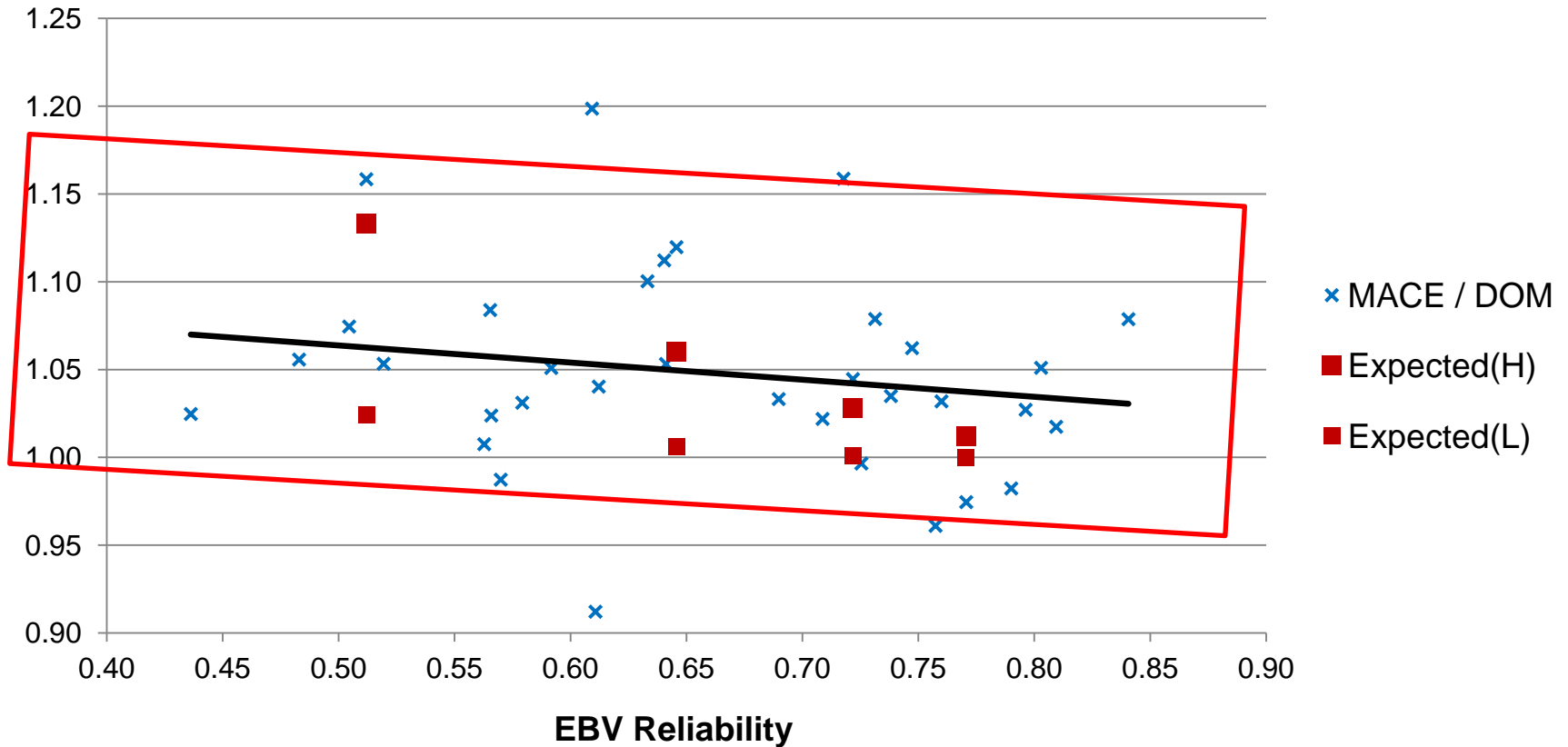
Ratio of Mendelian Sampling SD (MACE/DOM)
(higher if data are from the correlated country-trait)



Correlated vs Direct Trait

(Expected SD bias given $REL_{\hat{m}}$)

Ratio of Mendelian Sampling SD (MACE/DOM)
 (higher if data are from the correlated trait)



Genomic Validation

(Domestic Bulls with 2014 EBV)

	No Adjustment		Variance Adjustment	
	Slope of Prediction (2010 GPA → 2014 EBV)			
GLPI (~N(0,1))	0.93		0.97	
GLPI_PROD	0.96		1.00	
GLPI_DUR	0.94		0.99	
GLPI_HF	0.94		0.99	
	Average Difference (2010 GPA – 2014 EBV)			
	All bulls	Top 100 - All	All bulls	Top 100 - All
GLPI (~N(0,1))	0.37	0.08	0.31	-0.02
GLPI_PROD	0.23	0.12	0.21	0.04
GLPI_DUR	0.16	-0.01	0.11	-0.03
GLPI_HF	0.05	0.00	0.05	0.00

Genomic Validation (Domestic vs Foreign Bulls)

Domestic	No Adjustment	Variance Adjustment
	Slope of Prediction (2010 GPA → 2014 EBV)	
GLPI (~N(0,1))	0.93	0.97
GLPI_PROD	0.96	1.00
GLPI_DUR	0.94	0.99
GLPI_HF	0.94	0.99

Variance adjustment to DRP

Genomic Validation

(Domestic vs Foreign Bulls)

Domestic	No Adjustment	Variance Adjustment
	Slope of Prediction (2010 GPA → 2014 EBV)	
GLPI (~N(0,1))	0.93	0.97
GLPI_PROD	0.96	1.00
GLPI_DUR	0.94	0.99
GLPI_HF	0.94	0.99
Foreign	No Adjustment	Variance Adjustment
	Slope of Prediction (2010 GPA → 2014 MACE)	
GLPI (~N(0,1))	0.86	0.90
GLPI_PROD	0.96	1.00
GLPI_DUR	0.83	0.88
GLPI_HF	0.91	0.97

Variance adjustment to DRP

Genomic Validation

(Domestic 2014 by proof type in 2010)

LPI ~N(0,1)	No Adjustment		Variance Adjustment	
	All bulls	Top 100	All bulls	Top 100
2010 DGV	Average Difference (2010 DGV – 2014 EBV)			
Domestic (DGV)	-0.02	0.12	0.00	0.06
Foreign (DGV)	0.23	0.30	0.10	0.15
Young (DGV)	0.29	0.68	0.21	0.50
2010 GEBV	Average Difference (2010 GEBV – 2014 EBV)			
Domestic (GEBV)	0.00	0.09	0.00	0.03
Foreign (GMACE)	0.23	0.32	0.21	0.27
Young (GPA)	0.37	0.45	0.31	0.29

Genomic Validation

(Domestic 2014 by proof type in 2010)

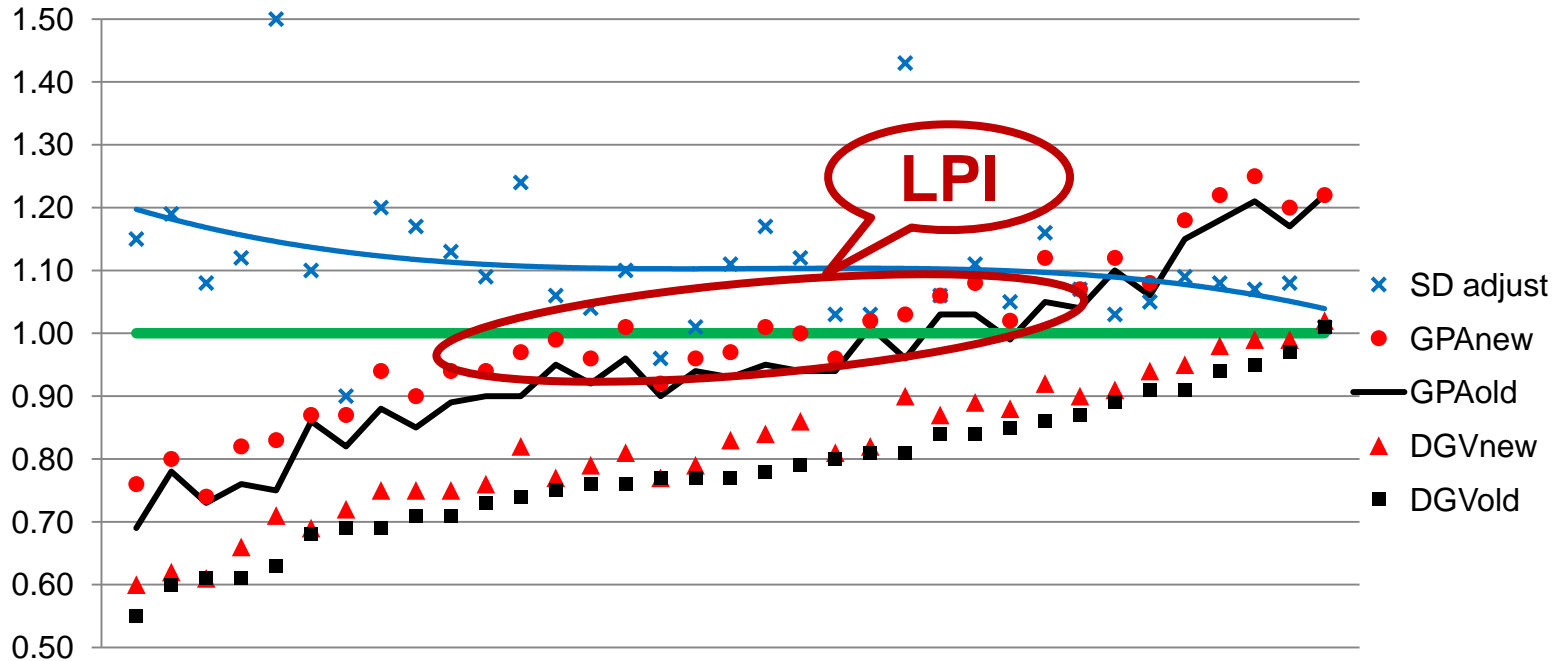
Blending: DGV+EBV

LPI ~N(0,1)	No Adjustment		Variance Adjustment	
	All bulls	Top 100	All bulls	Top 100
2010 DGV	Average Difference (2010 DGV – 2014 EBV)			
Domestic (DGV)	-0.02	0.12	0.00	0.06
Foreign (DGV)	0.23	0.30	0.10	0.15
Young (DGV)	0.29	0.68	0.21	0.50
2010 GEBV	Average Difference (2010 GEBV – 2014 EBV)			
Domestic (GEBV)	0.00	0.09	0.00	0.03
Foreign (GMACE)	0.23	0.32	0.21	0.27
Young (GPA)	0.37	0.45	0.31	0.29

Variance adjustment to DRP

Slopes of Prediction (2010 GPA → 2014 EBV)

Slopes of prediction
 (Dec 2014 EBV from 2010 genomic PA / DGV)



Summary

- **Variance adjustment to de-regressed MACE proofs reduced $V(\text{GEBV})$ for foreign bulls**
 - **Improved SNP estimates and GEBV for ALL bulls**
 - **Genomic validation results improved for Domestic and Foreign bulls, for both DGV and official GPA**
 - Slopes of predicting both EBV and MACE closer to 1
 - **Implemented in April 2015**
- **Slopes of prediction for DGV still below 1 for nearly all traits**
 - Lower slopes for low reliability traits from MT models

Summary

- Slopes of prediction much better for GEBV
 - CAN averaging of DGV and EBV is still helpful
- Addressed **multi-country** correlated data
- Can we also address correlated data from **multi-trait domestic** systems?
 - e.g. Adjust variance of DRP for DOMESTIC bulls, to account for higher $REL_{\hat{m}}$ when mainly correlated trait information?

Conclusions

- **De-regression is a variance-scaling of \hat{m}**
- **Single-trait genomic evaluation assumes single-trait EBV as input to de-regression**
- **Multi-trait EBV should have a different variance-scaling of \hat{m} especially if:**
 - EBV reliability is low and mainly from correlated traits
 - Genetic correlations in multi-trait model are low, e.g:
 - Time-dependent traits (e.g. Fertility and Survival)
 - Foreign bulls in the reference population

Final Note

- Truncated data were **official MACE** from 2010 vs. **domestic EBV by re-running 2010 data through current programs**
 - Can we do better with truncated MACE?
 - 4 years of changes to models and data (from any country) not in 2010 MACE ... **Lowers r_g**
 - Variance adjustment helps but cannot completely correct this problem
 - **Truncated MACE runs by Interbull could improve genomic validation test results**



=



?

Thank You!