

Modeling unknown parent groups or metafounders in single step genomic BLUP – results of a simulation study

*Judith HIMMELBAUER, Hermann SCHWARZENBACHER,
Christian FUERST, Birgit FUERST-WALTL*

Interbull meeting, Bled, 2024-05-21

What is the aim of the study?

ssGBLUP for German-Austrian-Czech Fleckvieh population since April 2021

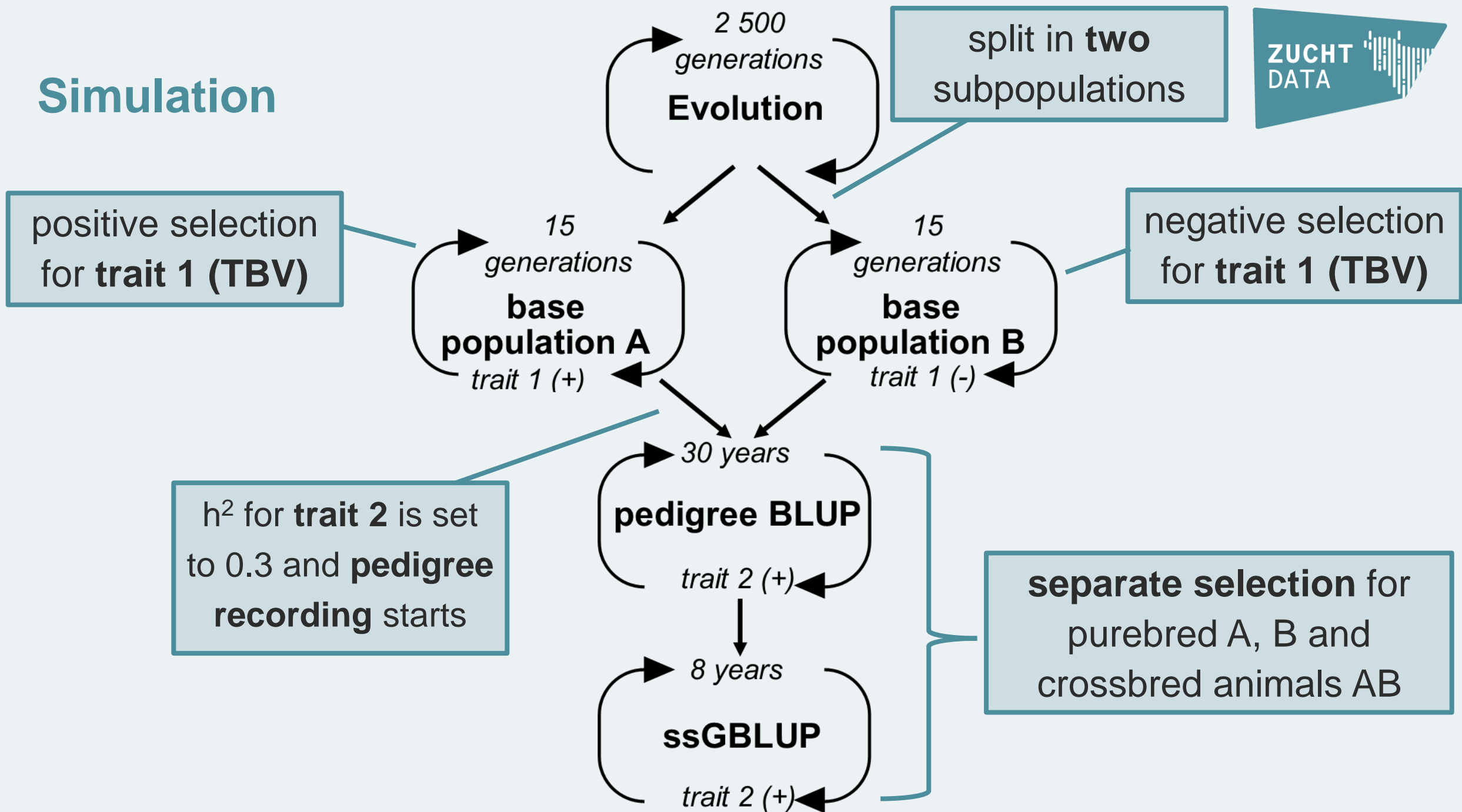
potential next step in the national evaluation: **metafounder (MF)**

→ simulation study to test and investigate multiple aspects

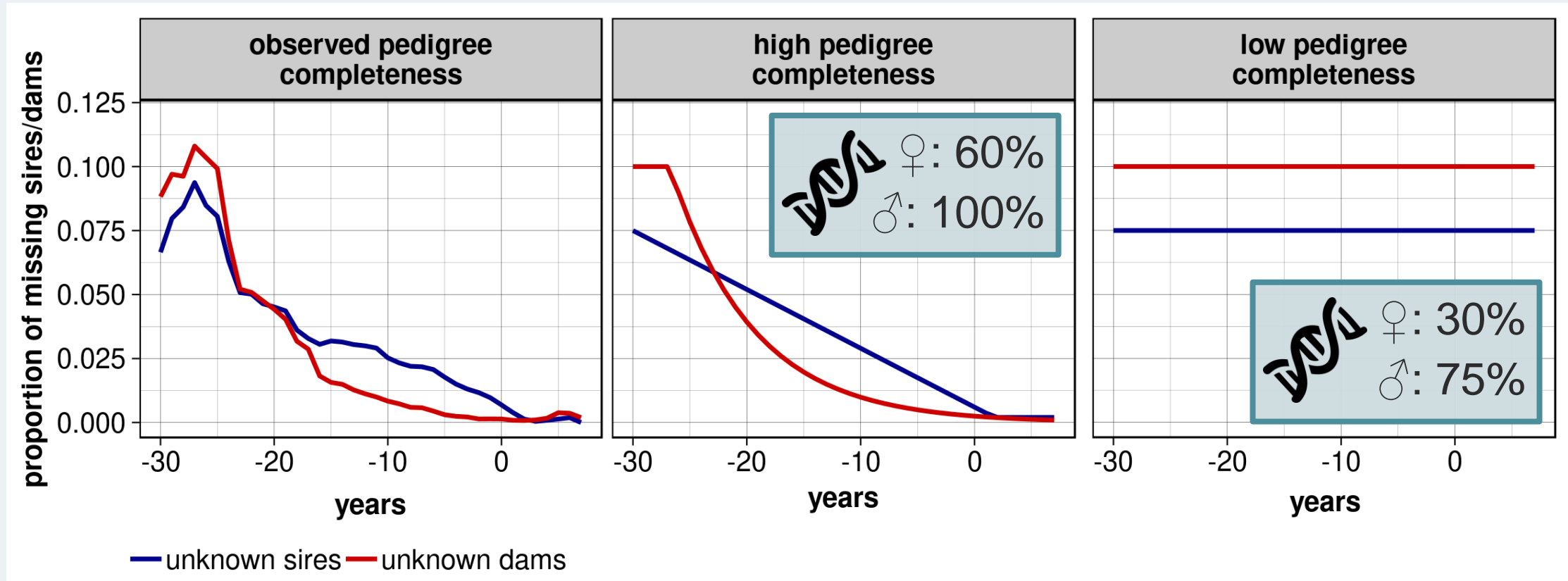
Content of the presentation:

- validation statistics for genetic evaluations with and without unknown parent groups (UPG) and MF
- effect of scaled variance components on GEBVs
- expected effects of MF in routine validation (linear regression , LR)

Simulation



Pedigree completeness - scenarios



Classification of UPG/MF



- **Full pedigree:**
 - true full pedigree without missing parents
 - 2 UPG or MF for the two subpopulations in the pedigree basis

- **True missing pedigree:**
 - unknown pedigrees
 - classification based on true subpopulation, true age, and true sex

Genetic evaluations

for all ssGBLUP G was computed with APY

1. **ssGBLUP without UPG (no_UPG):** ssGBLUP without genetic groups in the pedigree
2. **ssGBLUP with UPG in A (UPG_alteredQP):** ssGBLUP with UPG in the pedigree (A and A_{22})
3. **ssGBLUP with UPG (UPG_fullQP):** ssGBLUP with UPG in A , A_{22} and G
4. **ssGBLUP with MF and true Γ (MF_true)**
5. **ssGBLUP with MF, true Γ and scaled variance components (MF_sc)**

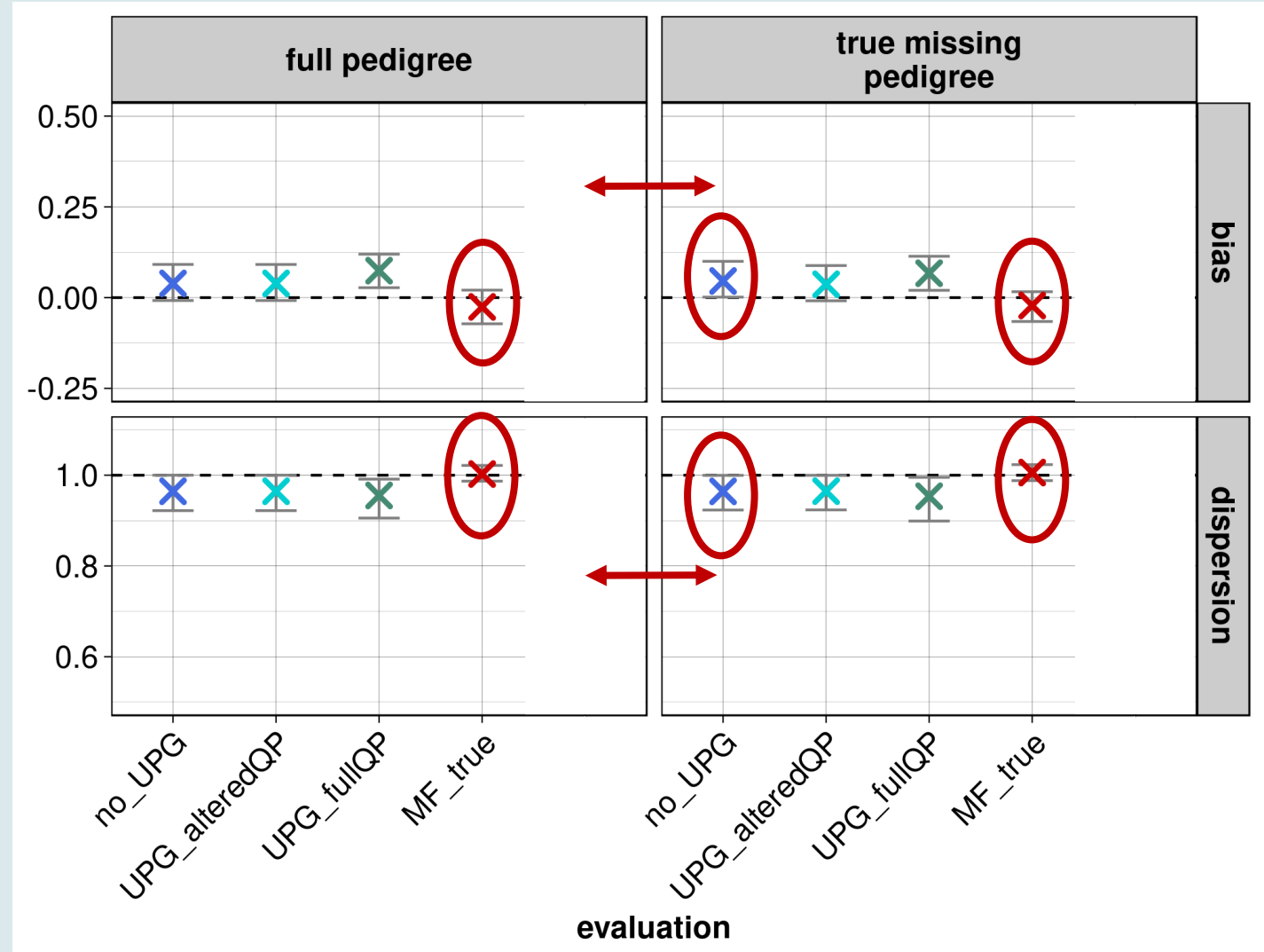
Comparison of evaluations: high pedigree completeness

low proportions of unknown parents and high genotyping rate

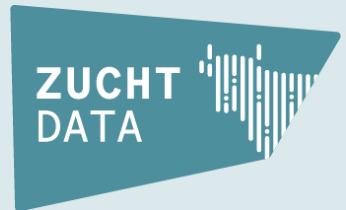
→ effects from MF/UPG on GEBVs are very small

less bias and dispersion with MF

→ positive effects from better alignment of A and G

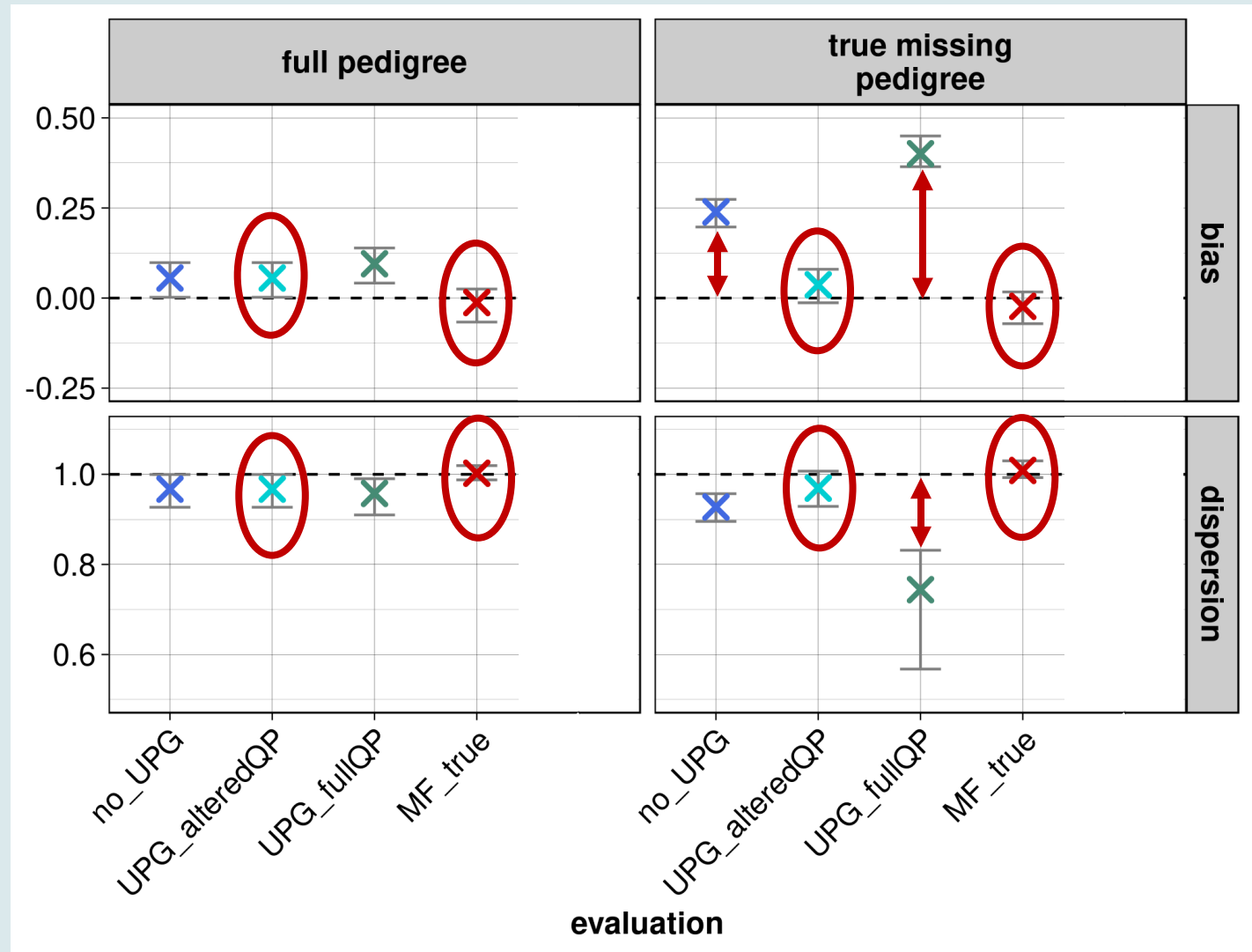


Comparison of evaluations: low pedigree completeness



clear differences for
no_UPG and
UPG_fullQP between
full and missing
pedigree

UPG_fullQP: double
counting (relationships
in G are already
complete)



Scaling of variance components: low pedigree completeness

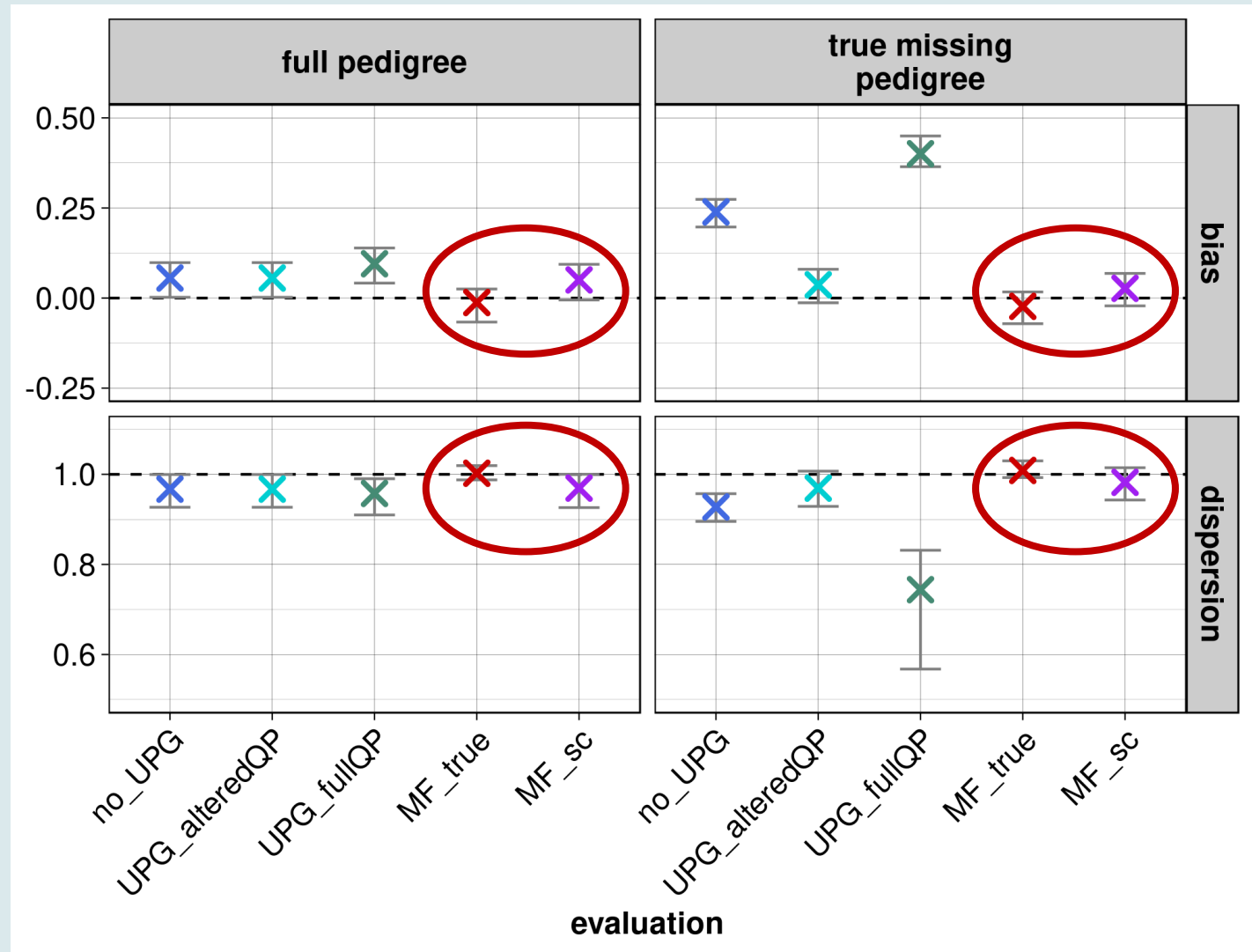


Scaling according to Legarra et al. 2015:

$$\sigma_{related}^2 \approx \frac{\sigma_{unrelated}^2}{1 + \frac{diag(\Gamma)}{2} - \bar{\Gamma}}$$

small negative effects on bias and dispersion

➔ maybe too high h²?



Estimation of variance components

h^2 was defined in the simulation after merging the populations

→ 0.3 is the „unrelated h^2 “

„related h^2 “ is higher

→ is in line with the theoretical derivations (Legarra et al. 2015)

without MF (Γ)			
	genetic variance	residual variance	h^2
mean	1,0663	2,3891	0,3088
min	0,9625	2,1878	0,2708
max	1,1142	2,6268	0,3224

What happens if Γ (MF) is used in the estimations?

with MF (Γ)			
	genetic variance	residual variance	h^2
mean	1,5212	2,3941	0,3887
min	1,3761	2,1918	0,3435
max	1,5941	2,6303	0,4045

Estimation of variance components

without MF (Γ)			
	genetic variance	residual variance	h^2
mean	1,0663	2,3891	0,3088
min	0,9625	2,1878	0,2708
max	1,1142	2,6268	0,3224

What happens if Γ (MF) is used in the estimations?

with MF (Γ)			
	genetic variance	residual variance	h^2
mean	1,5212	2,3941	0,3887
min	1,3761	2,1918	0,3435
max	1,5941	2,6303	0,4045

How well does the approximation using Legarra et al. 2015 fit the estimated h^2 ?

scaled			
	genetic variance	residual variance	h^2
mean	1,4970	2,3891	0,3854
min	1,3565	2,1878	0,3405
max	1,5644	2,6268	0,4006

scaling with the approximation in Legarra et al. (2015) leads to very similar results

➔ But why are the validation statistics (bias, dispersion) worse?

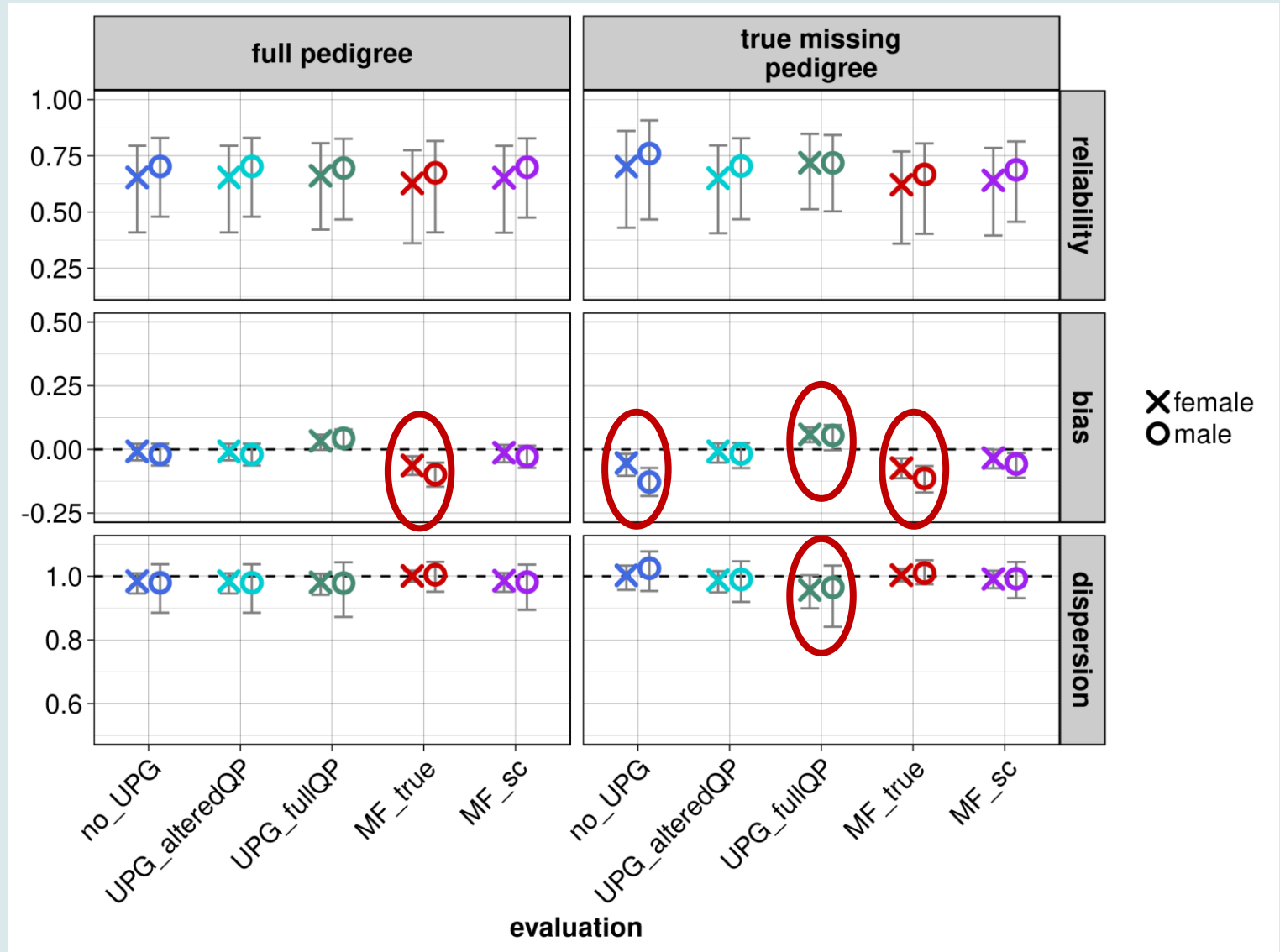
LR validation: low pedigree completeness



no big differences
between the evaluations

bias and dispersion of
no_UPG and
UPG_fullQP is not
detected

small „wrong“ bias for
MF_true



Conclusions

- MF have positive effects on bias and dispersion
- Wrong consideration of UPG can lead to extreme bias and dispersion (in (sub)populations with many unknown pedigrees)
- Scaling of variance components has no positive, but rather negative effects on validation statistics
- Validation with LR seems to be of limited use to assess the benefits of MF in this study