KU LEUVEN

fwo

LPMX

Multiple imputation to handle missing covariate data in pharmacometrics modelling

My-Luong Vuong, Geert Verbeke, Erwin Dreesen

PKUK, Coppid Beech Hotel, Bracknell, UK
28th November 2024

Missing covariate data

- Common issue in pharmacometrics

Missing Data

- Various missing mechanisms

- Proper handling of missing covariate is crucial!

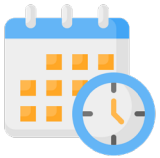
Handling missing covariate data in pharmacometrics

- **Single imputation (SI)** >>> high risk of bias 
- **Multiple imputation (MI)** >>> preferred but underexplored 
- No published study compared SI and MI under various degrees of missing covariate data 

Aim: Compare the accuracy of SI and MI for handling different levels of covariate missingness in pharmacometrics modeling

Data source[1]

- Complete dataset 

- Oral warfarin – single dose – 32 healthy volunteers 
- 251 warfarin concentrations – median of 6 per subject 
- Sampling times: t = {0.25, 0.5, 1.0, 2.0, 4.0, 6.0, 12.0, 24.0, 48.0, 72.0, 96.0} hours after dose 
- Patient characteristics: body weight, sex and age 

[1]O'Reilly RA, Aggeler PM. Studies on coumarin anticoagulant drugs. Initiation of warfarin therapy without a loading dose. Circulation. 1968;38:169-177.

Structural popPK model

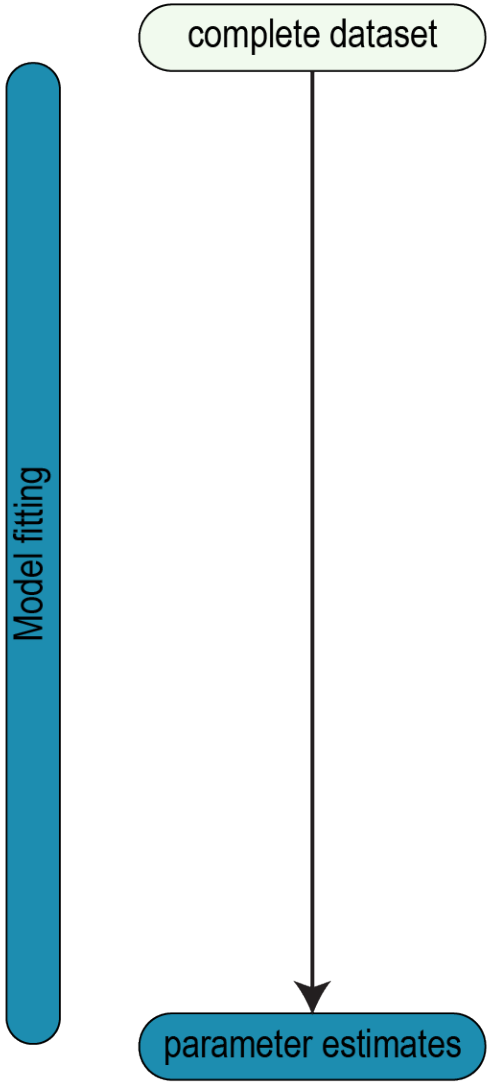
- A one-compartment model >>> Ka, Vd and CL
- Inter-individual variability >>> Ka, Vd and CL
- Residual variability >>> combined additive and proportional
- Covariate model:

$$TVCL = \theta_{1\text{-CL}} \times \left(\frac{\mathbf{WT}}{\mathbf{WT_{median}}}\right)^{\boldsymbol{\theta_{2-CL}}}$$

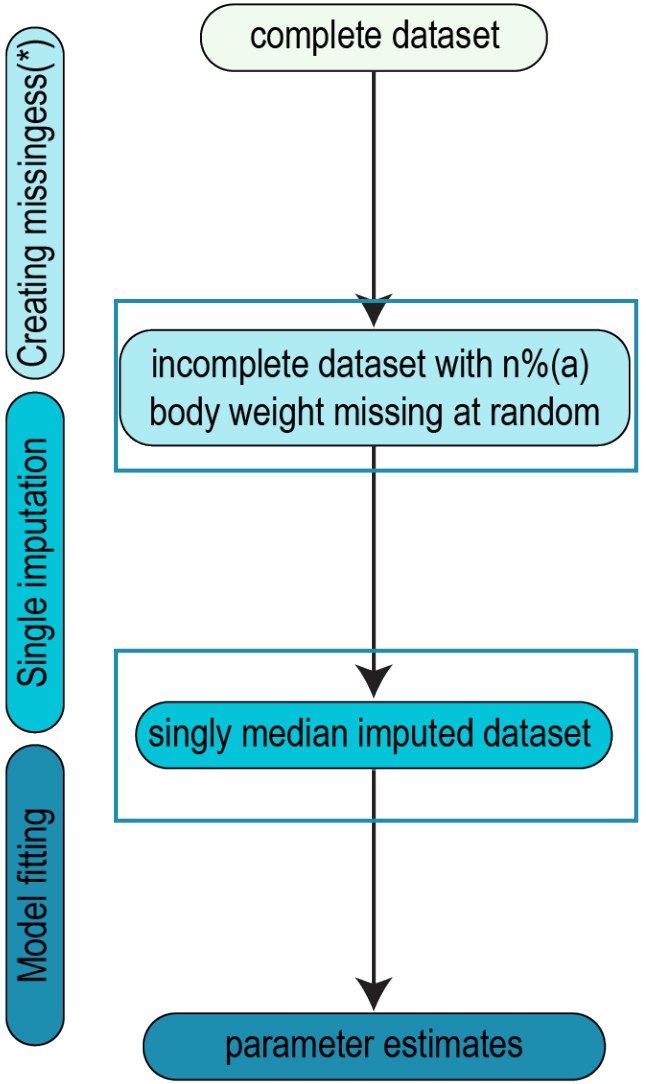
$$TVVd = \theta_{1\text{-Vd}} \times \left(\frac{\mathbf{WT}}{\mathbf{WT_{median}}}\right)^{\boldsymbol{\theta_{2-Vd}}}$$

Case study workflow

Complete dataset analysis

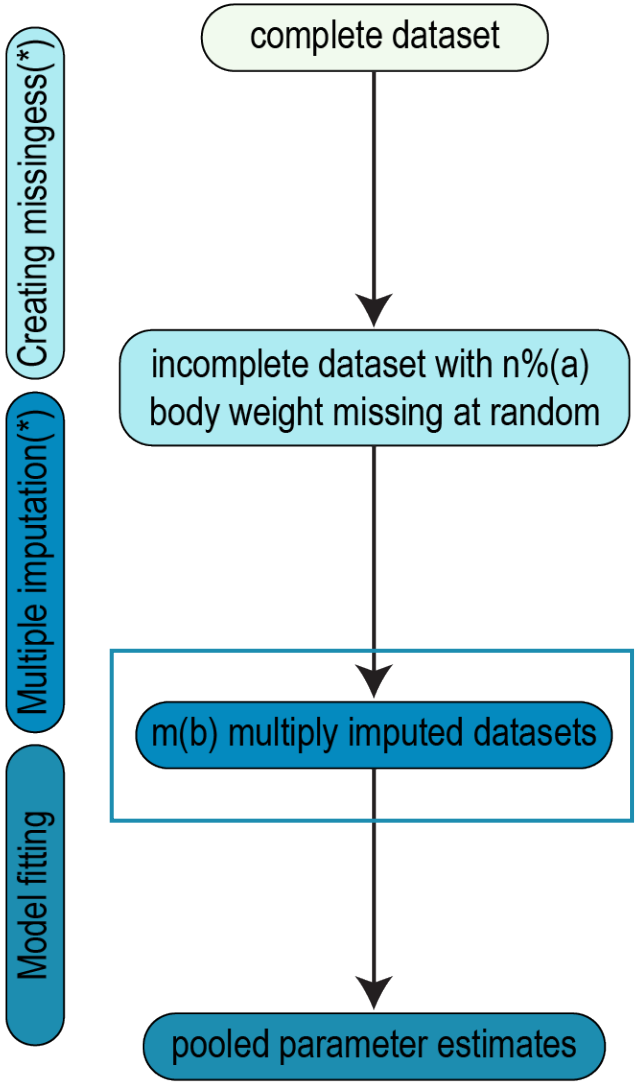


Singly imputed datasets analysis



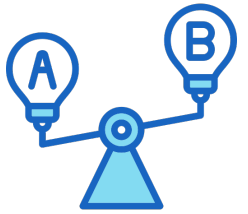
(a)n Є {6.25, 12.5, 25, 50, 75} (*)repeated 30 times

Multiply imputed datasets analysis



(a)n Є {6.25, 12.5, 25, 50, 75} (*)repeated 30 times
(b)m Є {20, 30, 45, 70, 105}

Performance assessment – SI and MI



- Model stability >>> model convergence rate

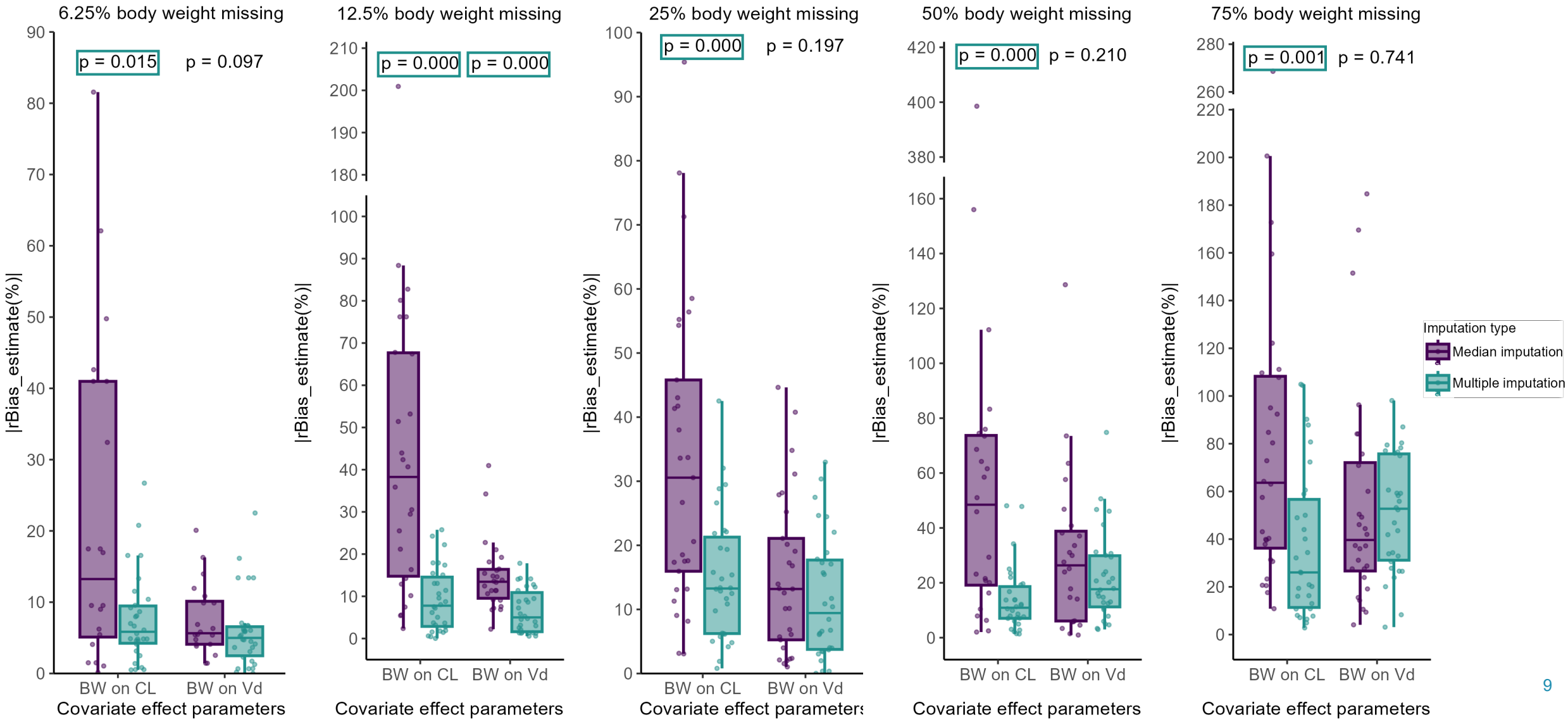
- Accuracy in estimating point estimates:
  - $|\text{rBias\_estimate}(\%)| = \left|\frac{\text{CovariateEstimate}_{\text{SI/MI}} - \text{CovariateEstimate}_{\text{complete}}}{\text{CovariateEstimate}_{\text{complete}}}\right| \times 100\%$
  - Compared |rBias_estimate(%)| >>> Wilcoxon rank-sum test

- Accuracy in reflecting uncertainty estimates:
  - $|\text{rBias\_standard-error}(\%)| = \left|\frac{\text{Standard-Error}_{\text{SI/MI}} - \text{Standard-Error}_{\text{complete}}}{\text{Standard-Error}_{\text{complete}}}\right| \times 100\%$
  - Compared |rBias_standard–error(%)| >>> Wilcoxon rank-sum test
  - Compared variances of |rBias_standard–error(%)| >>> Fligner-Killeen test

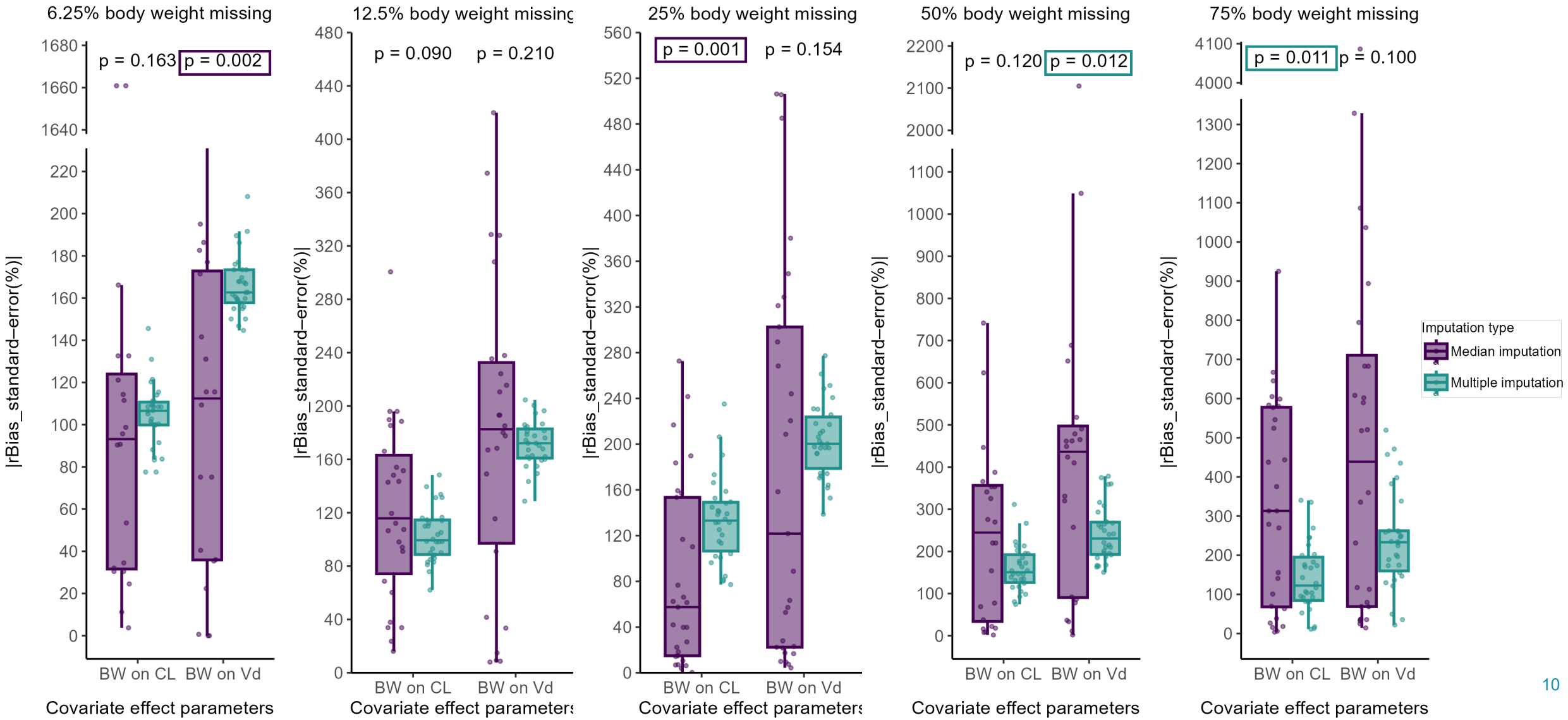
Model convergence rate is stable across 5 scenarios

	Missingness scenario	6.25% BW missing	12.5% BW missing	25% BW missing	50% BW missing	75% BW missing
Single imputation	Number of imputation	30	30	30	30	30
	Number of converged models (%)	20 (66.7)	26 (86.7)	29 (96.7)	24 (80.0)	28 (93.3)
Multiple imputation	Number of imputations	20	30	45	70	105
	Number of converged models (median [min – max])	17 (11 – 20)	26 (19 – 30)	38 (29 – 45)	62 (50 – 69)	95 (75 – 103)
	Percentage of converged models	**85.0**	**86.7**	**84.4**	**88.6**	**90.5**
	(median [min – max])	(55.0 – 100.0)	(63.3 – 100.0)	(64.4 – 100.0)	(71.4 – 98.6)	(71.4 – 98.1)

Multiple imputation yielded significantly more accurate covariate effect estimates



Multiple imputation produced more reliable uncertainty estimate at higher missing rates



Conclusions

- MI provides **more accurate** covariate effect **estimates** and **better uncertainty reflection**.

- We encourage **pharmacometricians** to **adopt MI over SI**

- **Get started**: mice, mimp

Research Foundation Flanders
Opening new horizons

myluong.vuong@kuleuven.be