

Performance evaluation of the Automatic Model Development (AMD) tool with real population pharmacokinetic datasets from clinical drug development projects



Franziska Schaedeli Stark, Sylvie Retout, Valérie Cosson, Sreenath Krishnan, Orwa Albitar, João A. Abrantes, Emilie Schindler

Roche Pharma Research and Early Development, Pharmaceutical Sciences, Roche Innovation Center Basel, Basel, Switzerland

Objectives

To evaluate the AMD tool in a real drug development context by assessing its performance in identifying models with appropriate descriptive and predictive properties, and identifying opportunities for its enhancements.

Conclusions

The AMD tool was able to generate reasonable descriptive and predictive models for 8/9 investigated real datasets, confirming results generated previously [1]. However, the AMD performance evaluation highlighted a trend to provide more complex models than the ones developed manually. The assessment allowed identifying improvements of the AMD tool and subtools that have been implemented in more recent versions. This work further showcases the usefulness of the AMD tool towards an automated modeling environment in drug development.

Introduction

- The AMD tool in PharmPy was designed for automatic, stepwise building of all components of common pharmacometric models using NONMEM [1,2,3].
- The potential of the AMD tool in rapidly producing reliable population PK (popPK) models in an early drug development setting has recently been demonstrated using simulated datasets [4].
- We evaluated the performance of the AMD tool with 9 real popPK datasets and compared the outcome to the models developed previously by pharmacometricians (“manual” model).

Methods

- Datasets from 9 previously developed popPK models were analyzed with the AMD tool, using the pharmpy R-wrapper pharmr v. 0.88.0 [5], generally with the default settings (Figure 1, Table 1).
- For datasets with covariate effects in the manual model, covariate search was performed subsequently using the COVsearch subtool.
- Final AMD models were compared to the manual ones in terms of model complexity, data description, parameter precision, predictive performance, covariate detection and impact on clinical decisions.
- All analyses were run in the model building environment Improve [6] using NONMEM 7.5.1.

Figure 1: Workflow for the AMD tool testing

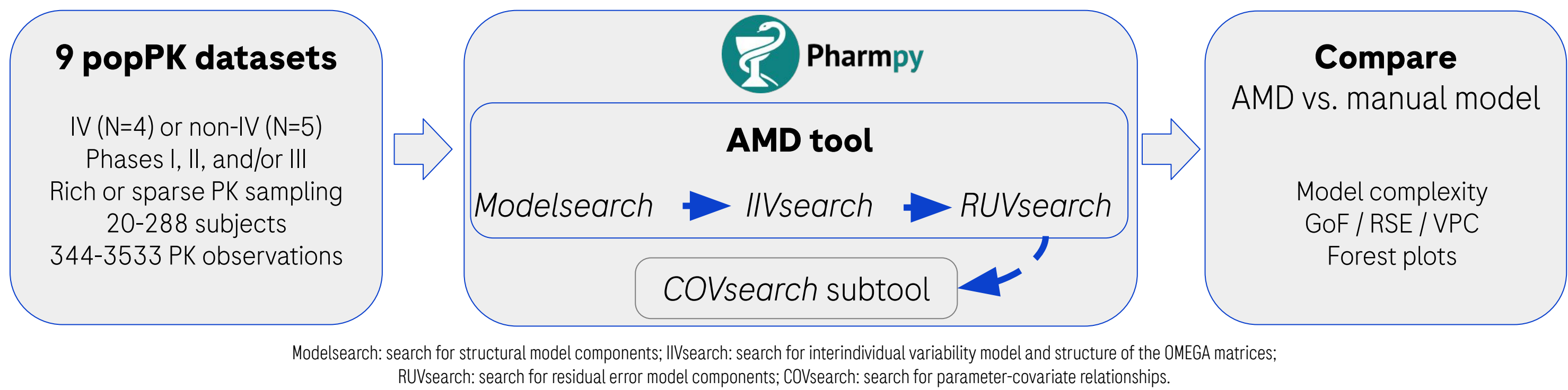


Table 1: AMD tool settings used for the model building with pharmpy/pharmr v.0.88

Subtool	Options / settings	Selection criteria
Modelsearch	Search space: Absorption: FO; FO with lag time; transit CMTs (n=1,3, or 10) with/without absorption CMT; sequential ZO-FO; Disposition: 1, 2, or 3 CMTs; Elimination: FO, MM, or combined FO-MM Algorithm: “reduced_stepwise” (i.e. stepwise evaluate and reduce the number of runs with the same model features.) IIV strategy: [CLVC]+[MAT]+[MDT]	BIC
IIVsearch	Algorithm: “brute force” (i.e. Step 1: Number of etas, Step 2: block structure)	BIC
RUVsearch	Search space: Prop, Prop+Add, IIV, RUV, time_var [4 bins], Power, p_value=0.05	OFV
COVsearch	Algorithm: SCM, p_forward=0.05, p_backward=0.01 - On parameters with IIV; exponential model for continuous covariates, multiplicative for categorical covariates	OFV

Results

- A final AMD model was obtained for 8 out of 9 datasets, with 100 to 920 model building steps. A comparison of AMD and manual models is shown in Table 2.
- Final AMD models had adequate descriptive and predictive performances.
- The selected models tended to be more complex and numerically less stable (e.g. high condition number and/or RSEs, rounding errors) than the manual models.
 - The Modelsearch subtool included an absorption delay in most models, and a 3-compartment disposition was selected in 7/9 AMD models, while most manual models had a 2-compartment disposition. Michaelis-Menten elimination was identified in 4 cases, while in manual models it was present in only two cases.
 - The RUVsearch subtool selected IIV on RUV in 6/8, and time-varying RUV in 5/8 models, while in the manual models these were not commonly present (not systematically tested).
 - The prediction-corrected VPCs indicated similar predictive performances, despite differences in structural, IIV and/or RUV models, compared to the respective manual models.

Table 2: Comparison of final models after manual and AMD model building process

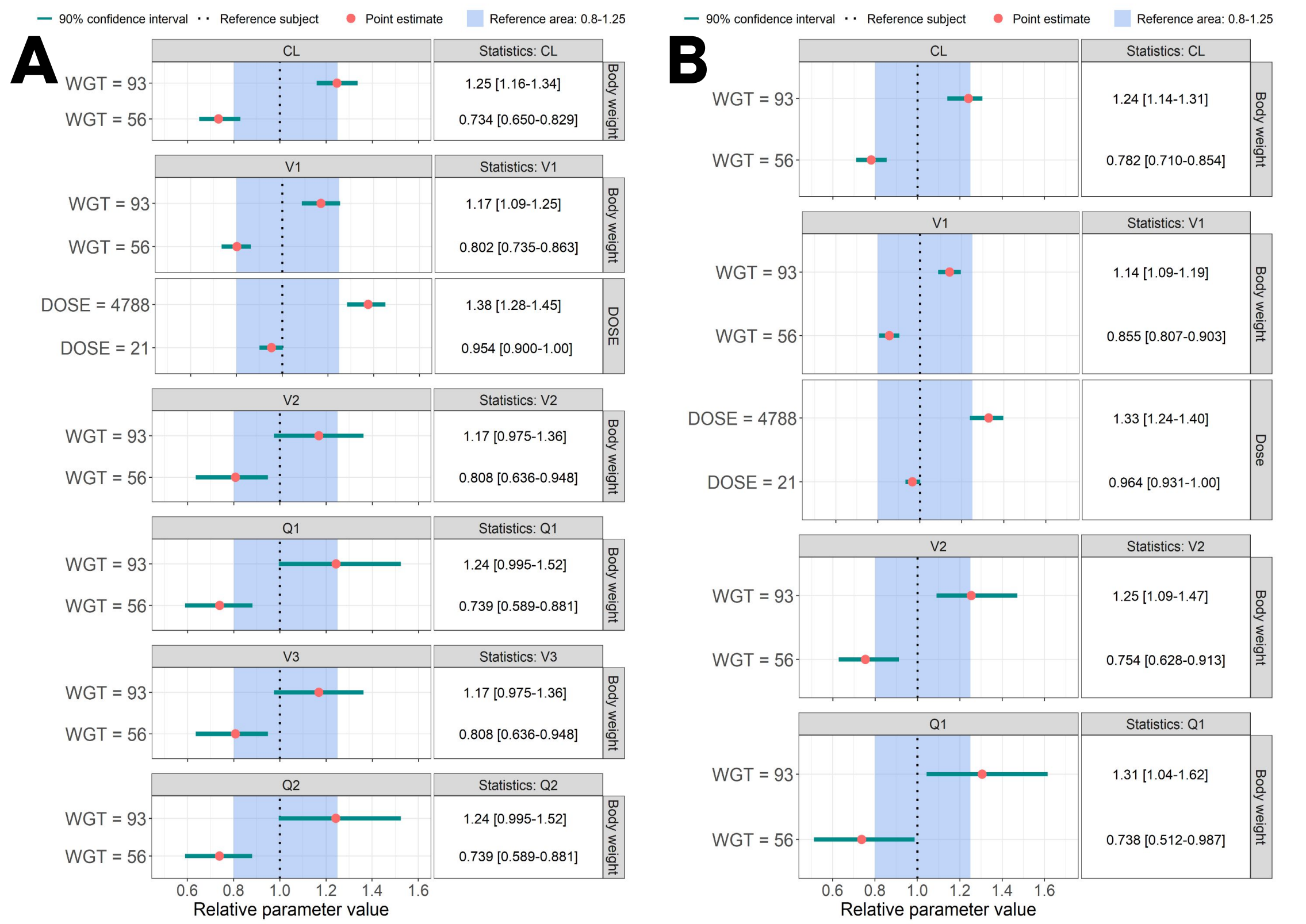
Dataset (N / n)		Absorption (nP)	Disposition, Elimination (nP)	IIV (nP)	RUV (nP)	Covariates (nP)	BIC
#1 (85/1792)	manual	i.v.	3-CMT, FO (6)	[CL]-[VC,OP1VP2] (7)	Prop (1)	[CL, OP1]-WT, [VC,VP1]-WT, VC-Dose (3)	34431
	AMD	i.v.	3-CMT, FO (6)	[CL,VC,OP1VP1] (10)	Prop+IIV, RUV (2)	CL-WT, VC-WT, VP1-WT, OP1-WT, VC-Dose (5)	34018
#2 (23/282)	manual	1-transit (1)	2-CMT, FO (4)	[CLVC]+[OP1VP1]+[KTR] (7)	Prop (1)	NA	4145
	AMD	Seq ZO-FO (2)	2-CMT, MM (5)	[MAT]+[MDT]+[CLMM,VC,OP1VP1] (12)	Prop (1)	NA	4056
#3 (266/1419)	manual	i.v.	2-CMT, FO+MM (6)	[CL,VCVP] (6)	Prop+IIV, RUV (2)	CL-[WT, GOMY, ALB], VC-WT, VP-[ETHN, ALK] (6)	8898
	AMD	i.v.	3-CMT, FO+MM (8)	[CL,VC,OP1] (6)	Prop+IIV, RUV+time_var (3)	CL-[WT, GOMY, EDOG, ALB], VC-[WT, ETHN], OP1-[ALK, RACE, BTMS2] (10)	8869
#4 (288/3533)	manual	Lag+2-transit (2)	2-CMT, FO (4)	[CL, V2, Q, V3]+[KA] (11)	Prop (1)	CL-[SMOK, SEX], VC-[SMOK, WT], VP-[SEX, WT], OP-SEX, KA-[FOOD, ETHN], BIO-ETHN, ALAG1+FOOD (11)	-2930
	AMD	3-transit+FO (2)	3-CMT, FO (6)	[CL,MDT]+[OP1VP1] (6)	Prop+IIV, RUV+time_var (3)	CL-[WT, SMOK, ETHN, SEX], VP1-[WT, SEX], OP1-[WT, SEX], MDT-FOOD (20)	-3573
#5 (41 / 964)	manual	FO (1)	2-CMT, FO (4)	[KA,CL, V2, Q, V3] (15)	Prop (1)	Allometric scaling with fixed exponents	2315
	AMD	Seq ZO-FO (2)	3-CMT, MM (7)	[CLMM]+[MAT,OP1]+[VC,OP2VP2] (10)	Prop+IIV, RUV+time_var (3)	ND	2148
#6 (20 / 344)	manual	i.v.	2-CMT, FO (4)	[CL,VC,OP1]-[VP1] (7)	Prop (1)	CL-WT, VC-WT (2)	155
	AMD	i.v.	3-CMT, FO (6)	[VC]+[CL,VP1,VP2] (7)	Prop+IIV, RUV+time_var (3)	ND	53
#7 (61/1908)	manual	n-transit (2)	2-CMT, FO+MM (6)	CL, V2, Q, V3, NTR, MIT, F, KM (8)	Prop (1)	Allometric scaling with fixed exponents	12391
	AMD	Lag + FO (2)	2-CMT, FO+MM (6)	NA (IIVsearch process stopped after 1500 steps)	NA	NA	NA
#8 (195/1371)	manual	FO (1)	2-CMT, FO (4)	[KA]+[CL]+[K32] (3)	Add+Prop+IIV, RUV (3)	VC-[WGT, SEX], K32-BMI; CL-[ETHN, ALT] (5)	11280
	AMD	10-transit (1)	3-CMT, FO (6)	[VC]+[CL,MDT,VP2] (7)	Prop (1)	VC-BMI; VP2-BMI; MDT-SMOKE (3)	11433
#9 (58/801)	manual	i.v.	2-CMT, FO (4)	[CL,VCVP1] (6)	Add+Prop (2)	CL-[STUD, SEX], VC-[STUD, SEX] (4)	6615
	AMD	i.v.	3-CMT, FO (6)	[OP1]-[CL,VC] (4)	Prop+IIV, RUV+time_var (3)	CL-[STUD, SEX], VC-[STUD, SEX] (4)	6504

N: number of subjects in the dataset; n: number of PK observations; nP: number of parameters in the model; CMT: compartment; FO: first order; ZO: zero-order; MM: Michaelis-Menten; transit: transit compartments; IIV: inter-individual variability model, parameters within brackets [] represent elements within each OMEGA block; RUV: residual unexplained variability; Add: additive; Prop: proportional; IIV, RUV: IIV on RUV, time_var: time-varying RUV allowing a higher RUV within a fixed time window early after dosing administration. Covariates: NA: not applicable; ND: covariate not tested; [CL, OP]-WT: [] indicate the same exponent estimated for both parameters; CL-[WT, GOMY,...]: [] indicate multiple covariate effects on one parameter. BIC: calculated according to the mixed (default) formula: $BIC = -2LL + n_{random_parameters} * \log(N) + n_{fixed_parameters} * \log(n)$ [2]

Color code for AMD vs manual model:	Same model	Higher complexity (>nP) and/or different components	Lower complexity (<nP)
-------------------------------------	------------	---	------------------------

- The covariates identified with the COVsearch subtool were generally aligned with the covariates in the manual models, and similar conclusions were made with respect to clinically relevant covariates for most models, as illustrated with the forest plots for a representative dataset in Fig. 2.
- The evaluation of COVsearch revealed some limitations of the tool due to limited flexibility, and more recent versions offer additional features, e.g. scm+ method [7], or prioritizing covariates based on their relevance in a specific drug development context.
- The results of this evaluation supported the implementation of updated strictness criteria to reduce the risk for selecting over-parameterized models, including stricter p-values during RUV search, the option to set a minimum number of significant digits, etc., that were implemented in newer versions of pharmpy/pharmr.
- AMD runs with more recent pharmr versions (0.106 or later) resulted in models that were generally less complex, without affecting the descriptive and predictive model performances.

Figure 2: Forest plots for the covariate effects in the manual model (A) and in the AMD model (B) for dataset #1



References:

- [1] Chen X, et al. PAGE 30 (2022) Abstr 10091 [www.page-meeting.org/?abstract=10091]
- [2] PharmPy/Pharmr AMD tool documentation [Internet]. <https://pharmpy.github.io>.
- [3] Abrantes J, et al. PAGE 30 (2022) Abstr 10051 [www.page-meeting.org/?abstract=10051]
- [4] Duvnjak et al. PAGE 31 (2023) Abstr 10529 [www.page-meeting.org/?abstract=10529]
- [5] Nordgren P, et al. PAGE 31 (2023) Abstr 10508 [<https://www.page-meeting.org/default.asp?abstract=10508>]
- [6] Hackl M et al. PAGE 31 (2023) Abstr 10586 [www.page-meeting.org/?abstract=10586]
- [7] Svensson RJ, Jonsson EN. CPT Pharmacometrics Syst Pharmacol. (2022) 11(9):1210-1222. doi: 10.1002/psp4.12838.

Acknowledgements

The authors would like to thank the Uppsala Pharmacometrics Research Group for their collaboration, Scinteco for their technical assistance, and Nicolas Frey for his valuable input.