Received: 31 December 2023 | Revised: 30 June 2024 | Accepted: 29 July 2024

DOI: 10.1002/psp4.13222

ARTICLE

A fully automatic tool for development of population pharmacokinetic models

Xiaomei Chen[1] | **Rikard Nordgren**[1] | **Stella Belin**[1] | **Alzahra Hamdan**[1] | **Shijun Wang**[1] | **Tianwu Yang**[1] | **Zhe Huang**[1] | **Simon J. Carter**[1] | **Simon Buatois**[2] | **João A. Abrantes**[2] | **Andrew C. Hooker**[1] | **Mats O. Karlsson**[1]

[1]Department of Pharmacy, Uppsala University, Uppsala, Sweden

[2]Roche Pharma Research and Early Development, Roche Innovation Center Basel, Basel, Switzerland

Correspondence
Mats O. Karlsson, Department of Pharmacy, Uppsala University, Box 580, 75123 Uppsala, Sweden.
Email: mats.karlsson@farmaci.uu.se

Abstract

Population pharmacokinetic (PK) models are widely used to inform drug development by pharmaceutical companies and facilitate drug evaluation by regulatory agencies. Developing a population PK model is a multi-step, challenging, and time-consuming process involving iterative manual model fitting and evaluation. A tool for fully automatic model development (AMD) of common population PK models is presented here. The AMD tool is implemented in Pharmpy, a versatile open-source library for pharmacometrics. It consists of different modules responsible for developing the different components of population PK models, including the structural model, the inter-individual variability (IIV) model, the inter-occasional variability (IOV) model, the residual unexplained variability (RUV) model, the covariate model, and the allometry model. The AMD tool was evaluated using 10 real PK datasets involving the structural, IIV, and RUV modules in three sequences. The different sequences yielded generally consistent structural models; however, there were variations in the results of the IIV and RUV models. The final models of the AMD tool showed lower Bayesian Information Criterion (BIC) values and similar visual predictive check plots compared with the available published models, indicating reasonable quality, in addition to reasonable run time. A similar conclusion was also drawn in a simulation study. The developed AMD tool serves as a promising tool for fast and fully automatic population PK model building with the potential to facilitate the use of modeling and simulation in drug development.

Study highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

Model-informed drug discovery and development (MID3) plays an important role in drug development and evaluation. Currently, population pharmacokinetic (PK) models are mostly developed manually and are time-consuming.

WHAT QUESTION DID THIS STUDY ADDRESS?

The study describes a fully automatic model development (AMD) tool that consists of a series of modules to develop all the components of a common population PK model. The modules for building structural, random effect and residual error models feature novel search algorithms. The AMD tool attempts to match approximately how a modeler might build a population PK model.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

The AMD tool is systematic, extensive, efficient, and flexible without input from any model file. The AMD tool was tested on 10 real PK datasets (intravenous or extravascular dosing) and in a simulation study. In both studies, the AMD tool provided final PK models with reasonable quality and running times. Different running sequences of the selected modules did not show a clear superiority of any sequence.

HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTICS?

The presented AMD tool is a promising and powerful tool that can facilitate and accelerate MID3.

INTRODUCTION

Model-informed drug discovery and development (MID3) has been widely used throughout the whole drug discovery and development process to facilitate decision-making, increase the success rate of later-phase clinical trials, and thus reduce development costs.[1,2] Regulatory agencies also apply MID3 to facilitate medicine evaluation, describe drug use in labeling, and inform policy.[3] Population pharmacokinetic (PK) modeling is an important part of MID3 in both the preclinical and clinical phases. Even after regulatory drug approvals, population PK modeling is an efficient approach for precision medicine and personalized dosing.[4]

The development of a population PK model is generally a challenging and time-consuming procedure that involves iterative manual model fitting by modelers. A population PK model applies nonlinear mixed-effect modeling (NLME), covering several different components, such as structural and covariate models, inter-individual variability (IIV) and inter-occasion variability (IOV) models (random effect), and residual unexplained variability (RUV) models.[5] For each component, there are a variety of different models to choose from, leading to numerous model possibilities. The common manual modeling approach consists of stepwise model construction, and it rarely assesses all potential models. In addition, the model development path and model selection are susceptible to the experience and preference of modelers and thus are subjective to a certain extent. Alternatively, a series of automatic model-building methods have been previously developed to facilitate the model construction process.[6-11] The stepwise covariate model (SCM) building procedure[6,7] is used to develop a robust covariate model. Another covariate modeling method is COnditional Sampling used for the Stepwise Approach based on Correlation tests (COSSAC),[8] which utilizes individual parameters sampled from conditional distributions. With a similar concept of using samples from conditional distributions, the stochastic approximation for model-building algorithm (SAMBA)[9] can automatically develop covariate models, correlation models, and error models. An automated search algorithm for population pharmacokinetic-pharmacodynamic (PKPD) model selection utilizing a genetic algorithm has also been proposed,[10,11] where a single genetic algorithm is used for all modeling components, and a large number of candidate models need to be estimated with the aim of finding the global minimum. This last method requires users to provide programming code to generate all tested models.

In this article, a fully automatic model development (AMD) tool is presented that covers common PK models and requires limited input from users (users do not need to provide any model code but only basic information about the studied drug and the dataset). The AMD tool consists of several search modules covering all the important elements of a population PK model. In each module, a model search begins from a starting model and applies a module-specific search algorithm based on the feature of the corresponding modeling component to automatically generate a series of candidate models, from which the optimal model is then selected based on a selection criterion. The modules operate sequentially in the whole AMD workflow, and a final population PK model is then given

by the tool. The AMD tool is implemented in Pharmpy,[12] an open-source software package for pharmacometric modeling, written in Python and readily executable in both Python and R.

The AMD tool is not only a tool that integrates the previous automatic model-building methods but also includes novel methods that automatically develop PK structural models, random-effect models, and RUV models. In this work, we describe in detail those modules with novel automatic model-building methods in addition to introducing the AMD tool. Moreover, three model development sequences were compared in terms of model selection and running time through the evaluation of the AMD tool using a series of 10 clinical PK datasets.

METHODS

Overview of the full AMD tool

The full AMD tool as a part of Pharmpy was developed using Python.[13] In addition to Python, Pharmpy can also be run using R[14] via the pharmr R package.[12] Pharmpy has versatile functionalities to facilitate pharmacometric modeling and aims to provide a platform to perform automatic model building. Those include the ability to read and manipulate model files and datasets, collect and present modeling results, modify models by adding or removing parameters or model components, etc. One of the main features of Pharmpy is the Pharmpy model object that is abstracted from model files of specific population PK software, enabling model modification independent of the original model language. Pharmpy is able to call external modeling software to perform model estimation based on model files that are automatically generated from Pharmpy model objects. The current version of the AMD tool requires NONMEM[15] for model estimation.

The AMD tool can automatically build a population pharmacokinetic model with the overall workflow shown in Figure 1. Users only need to provide basic information that includes: a PK dataset formatted for NONMEM, an optional data description file, administration route of the PK data (intravenous and/or extravascular administration), and initial values of PK parameters in the starting model. Users can also stipulate settings for the automatic search away from the default, such as the search space, or with a different starting model. During the model development, the AMD tool automatically generates a starting model, which uses the first-order conditional estimation (FOCE) with interaction as the default parameter estimation method. Users have the option to provide a starting model that applies other estimation methods or settings. From the starting model, the AMD tool then runs different modeling modules sequentially and selects the best model from each module to pass to the next module. The default module sequence is structural, IIV, RUV, IOV, allometry, and covariate modeling, though users can manually select which modules to use and the sequence of those modules. During the search, the AMD tool will evaluate candidate models in parallel, when possible, using available cores to accelerate the search process, which is especially efficient when there are many candidate models with long run times. By default, to be chosen as the final model, a candidate model needs to be free of minimization errors (except for rounding errors) and meet the model selection criteria of corresponding steps.

After the end of the AMD search, summary tables are generated to show the details of the search process, including the information of all the candidate models (such as objective function values and number of parameters) in addition to the final selected model. Moreover, all the model and result files are available and organized per module. A summary table of potential influential individuals and outliers can also be obtained, which are predicted using a machine learning algorithm.[16] A user guide with detailed instructions for the AMD tool is available on the website https://pharmpy.github.io/latest/amd.html. In addition, Supplementary Material S1 provides example code and a simulated dataset to illustrate how to run the AMD tool in R.

The included modeling modules can also be run as stand-alone tools. They apply different search algorithms and appropriate selection criteria to build the specific model component. Some modules are based

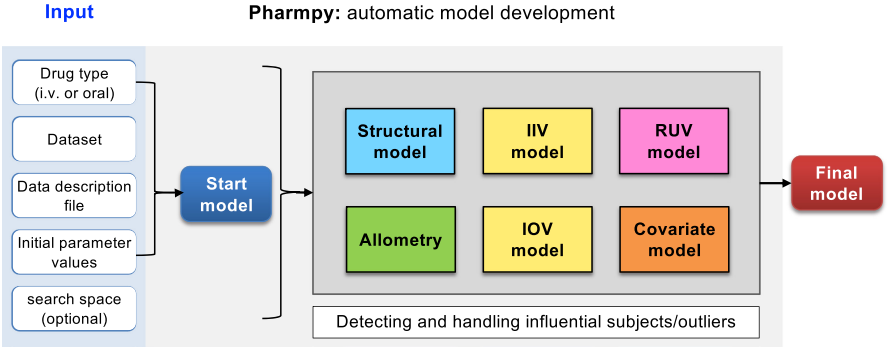

FIGURE 1 The automatic model development (AMD) tool workflow. IIV, inter-individual variability; IOV, inter-occasion variability; RUV, residual unexplained variability.

on an automatic screening algorithm that was developed previously. For example, the SCM building procedure[6] was re-implemented in the covariate modeling module (the covsearch tool). In the allometry module, allometric scaling is added to the model, which uses body weight for the scaling of clearance, compartment volumes, and inter-compartment clearance (see Supplementary Material S2 for more details). For the modules of PK structural modeling, random-effect modeling, and RUV modeling, novel search algorithms have been developed and are thus described in detail in the following sections.

Structural model building

A PK structural model uses a series of compartments to describe the change of drug amounts across various sites, reflecting the physiological processes in the body: absorption, distribution, and elimination.[17,18] The aim of structural model building is to identify the compartmental model that fits the PK data the best. The modelsearch tool in Pharmpy applies a novel search algorithm designed for an automatic search for the best PK structural model within a predefined model search space, given a dataset and a starting model (the simplest structural PK model within the defined search space). This is done through the creation of a series of candidate models, which are subsequently fitted and then ranked according to a pre-specified selection criterion.

Model search space

The modelsearch tool provides users with a wide range of models for defining a search space. The available models are categorized based on model features related to the absorption, absorption delay, distribution, and elimination processes (Figure 2a). Throughout the search procedure, all combinations among the four categories will be tested with certain exclusions, for example, a transit compartmental model with an absorption depot can only be coupled with a first-order absorption but not with the other two absorption models. A brief description of supported complex structural models is provided in Supplementary Material S3.

Reduced exhaustive stepwise algorithm

The reduced exhaustive stepwise algorithm evaluates all the models in the search space from a starting model in a stepwise manner (Figure 2b). In each step, a series of candidate models are created by changing one feature from the parent model (i.e., the starting model or models from the previous step). The initial values of the candidate models are taken from the estimated values of the parent

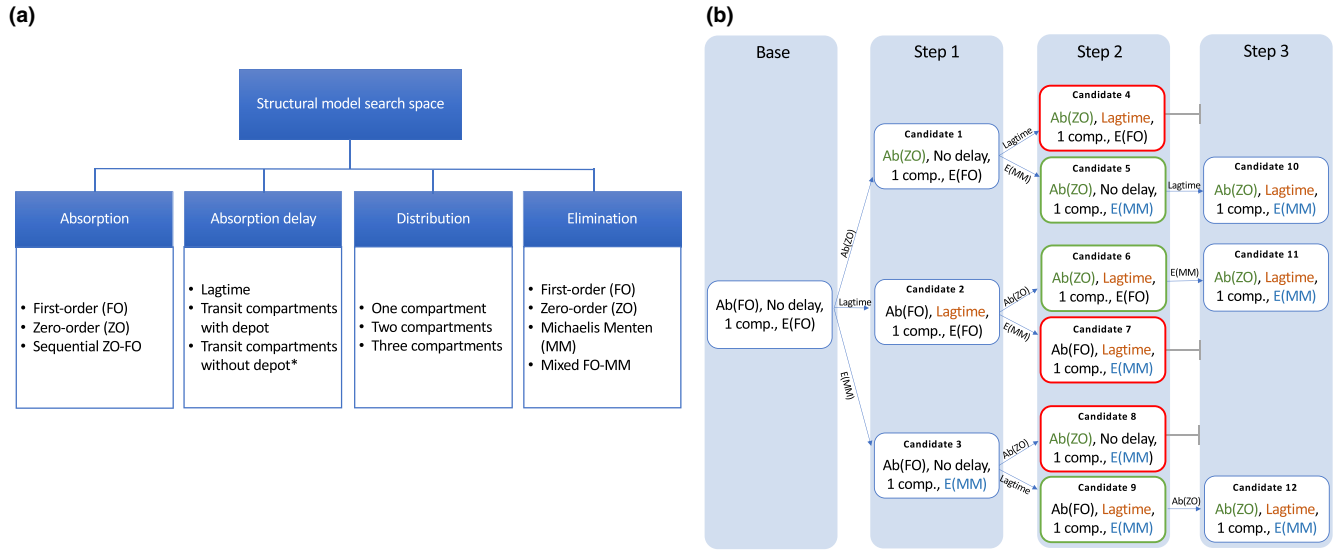

FIGURE 2 Illustrations of the automatic structural model development. (a) Structural models supported in the modelsearch tool. (b) An oral drug example to illustrate the reduced exhaustive stepwise algorithm. The search space of the example includes one-compartment model, first-order and ZO absorption, lag time absorption delay and no delay, as well as first-order and Michaelis-Menten elimination. The green-bordered candidate models are used to generate new models for the next step because they have the same or lower values of OFVs compared with their model replicates with different initial values (red-bordered candidate models). Ab: Absorption, comp.: Compartment, E: Elimination, FO: First-order, ZO: Zero-order, MM: Michaelis-Menten. *Note that there is one more parameter in the transit compartment model with depot compared with that without depot (see Supplementary Material S3 for more details).

models or algorithmically imputed for newly added parameters based on the existing parameter values (see Supplementary Material S3 for details). Using this method, replicated models are generated due to the different orders of feature modification and thus differ in the initial values of model parameters (e.g., candidate 4 and candidate 6 in Figure 2b). This duplication increases the robustness of the search process and the chance of reaching a global minimum. Among replicated models, the one with the lowest objective function value (OFV) proceeds to the next step. Options are provided regarding how to set IIV for newly added PK parameters (see Supplementary Material S3 for details).

Selection criterion

At the end of the model search process, all candidate models with successful minimization are ranked based on a predefined selection criterion. The modelsearch tool uses a search algorithm that combines stepwise and exhaustive approaches and thus involves comparisons among non-nested models. Therefore, likelihood-based information criteria are more suitable for those modules. The Bayesian Information Criterion (BIC) is used in the modelsearch tool for model selection instead of the Akaike information criterion (AIC) as the BIC tends to penalize against more complex models and thus select a simpler model. Specifically, the BIC is calculated according to Equation 1 for mixed-effects models[19]:

$$\mathrm{BIC} = -2\log L\left(\hat{\Psi}|\, y\right) + \dim\left(\Psi_R\right)\log N + \dim\left(\Psi_F\right)\log n_{\mathrm{tot}} \tag{1}$$

where $\hat{\Psi}$ is the vector of estimated model parameters, observations are expressed as y, N is the number of subjects, n_{tot} is the total number of observations, $\dim\left(\Psi_{\mathrm{R}}\right)$ is the number of parameters that are related to the random-effect components, and $\dim\left(\Psi_{\mathrm{F}}\right)$ is the number of parameters that are related to the fixed-effect components.

Random-effect model building

In NLME models, random effects describe the IIV and IOV, that is, the deviation of individual parameters from typical values of the population.[5,20] Generally, the IIV and IOV on PK parameters can be modeled as follows:

$$\begin{aligned} \theta_{i,j} &= g\left(\theta, \eta_i, \kappa_{i,j}\right) \\ \eta_i &\sim N\left(0, \Omega_P\right) \\ \kappa_{i,j} &\sim N\left(0, \Pi_P\right) \end{aligned} \tag{2}$$

where $\theta_{i,j}$ is the vector of parameter values of the ith individual at the jth occasion, g is a function to describe the relationship between typical values and individual parameters, θ is the vector of typical values of the PK parameters in the model, and the vectors η_i and $\kappa_{i,j}$ are random variables quantifying the IIV and the IOV of the parameter,[20] respectively. It is assumed that η_i and $\kappa_{i,j}$ are independently and normally distributed with a zero mean vector and variance–covariance matrices Ω_P and Π_P, respectively. The most common form of Equation 2 assumes a log-normal distribution and is used in the AMD tool by default:

$$\theta_{i,j} = \theta \bullet e^{\left(\eta_i + \kappa_{i,j}\right)} \tag{3}$$

IIV model search

Pharmpy provides a flexible module, the IIVsearch tool, that applies a novel search algorithm for the IIV model development. The IIVsearch tool offers two strategies to obtain the starting model: (1) to directly adopt an input model (e.g., the final selected model from the modelsearch tool) as the starting model; and (2) to modify the input model by adding IIV on all the PK parameters with either a diagonal or a full block variance–covariance matrix. Given a starting model, the IIVsearch tool implements a two-step exhaustive search process (Figure 3a): the first step is to identify the parameters incorporating IIVs; and the second step is to figure out the correlations among the IIVs, that is, the structure of the variance–covariance matrix of η_i. During each step, a series of candidate IIV models are generated by simplifying the starting model, whose parameter estimates are used to set the initial values of the candidate models. More specifically, in the first step different reduced models from the starting model are generated with one or more η_i removed and the original correlations unchanged. After estimating and ranking all the candidate models, the best model is selected and used as the starting model for the second step. In the second step, candidate models with one or more covariances fixed to zero are generated, among which the best final model is selected.

IOV model search

The IOVsearch tool in Pharmpy is a novel tool for automatically developing an IOV model. It consists of two steps (Figure 3b): (1) to select the parameters incorporating IOV; and (2) to further test if the previous IIV model can be reduced after adding an IOV model (i.e., remove one or more η_i for the parameters with existing $\kappa_{i,j}$). For each step, an exhaustive search similar to the IIV

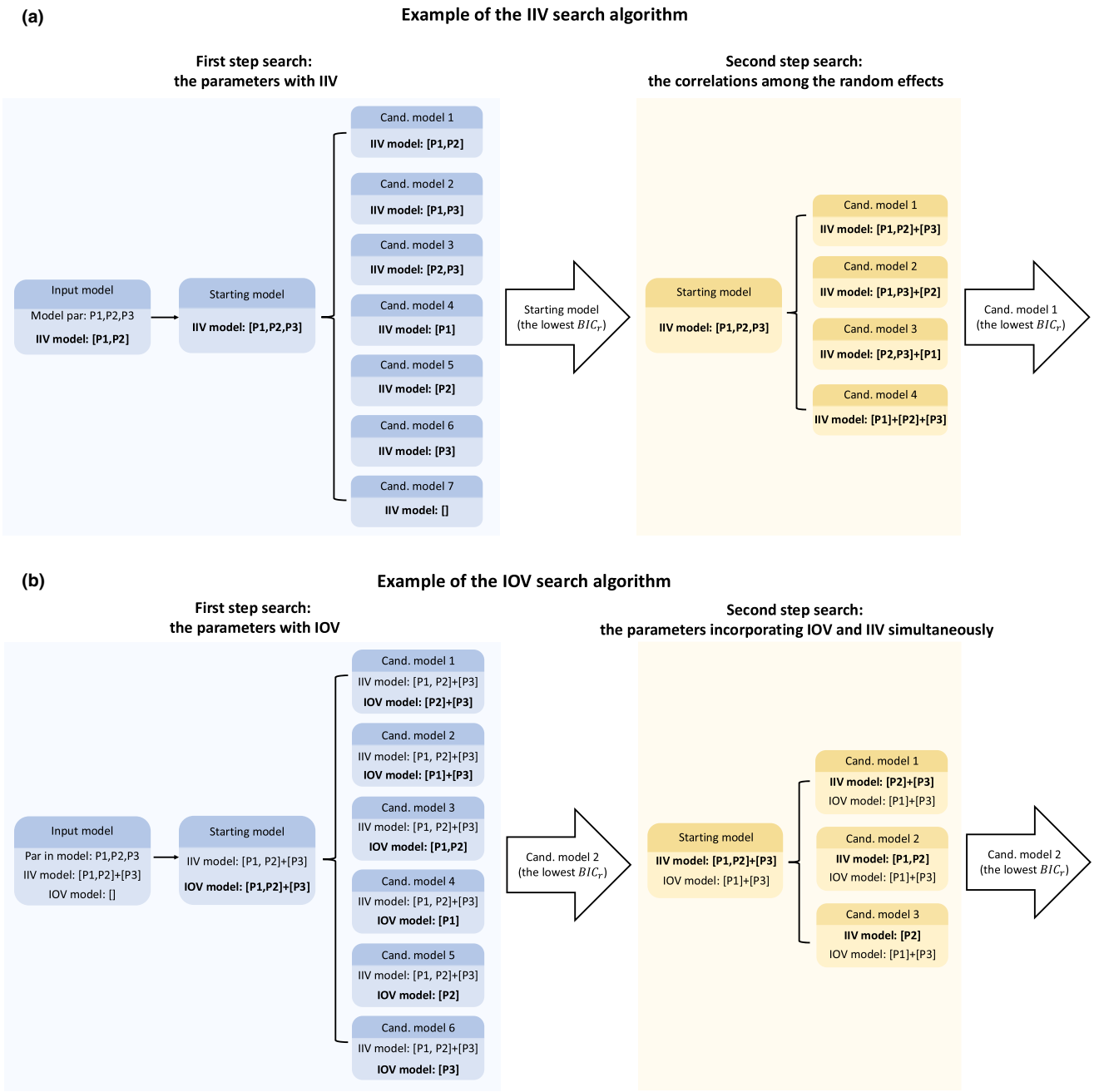

FIGURE 3 An illustration of the consecutive two-step search algorithm for IIV (a) or IOV search (b). The input model is a population PK model with three PK parameters as an example to demonstrate the algorithm. In the IIV/IOV model description, surrounding parameters with square brackets mean that random effects exist on the parameters, and the random effects are correlated. The plus sign is used to depict the mutually independent random effects. For example, [P1, P2] + [P3] indicates an IIV (or IOV) model where the PK parameters P1, P2, and P3 incorporate IIV (or IOV) and a correlation exists between the random effects on P1 and P2. The selection of the final model is conducted over all the models including the input model and the starting model. Cand.: Candidate.

search is applied. As for search space, users can choose to provide a list of parameters for testing in the IOV search. If not provided, the tool tests the parameters where there are already IIVs. For the variance–covariance matrix for IOV, the options are: (1) mutually correlated; (2) mutually independent; (3) the same correlation as the existing IIVs; and (4) a user-supplied structure.

Selection criterion

The selection criterion used in the IIV and IOV modeling procedures is the BIC specific for random effects selection.[21] The criterion, denoted as BIC_r, is calculated as:

$$\mathrm{BIC}_r = -2\log L\left(\hat{\Psi}|\, y\right) + N_{r\mathrm{par}}\log N \tag{4}$$

where N_{rpar} is the number of the random-effect parameters, that is, the number of non-zero elements in the matrices, Ω_P and Π_P. It should be noted that N_{rpar} is different from $\dim(\Psi_R)$ defined in Equation 1. During the model selection process, the model with the lowest BIC_r is selected as the best model for each step.

Residual error model building

The RUVsearch tool in Pharmpy is a novel stepwise automatic development tool for RUV models based on the method of resmod,[22] a fast postprocessing approach that evaluates RUV models by modeling the conditional weighted residuals (CWRES).[23] Briefly, CWRES are obtained from a pre-existing NLME model and subsequently treated as a dependent variable to build a base model. If the RUV model in the population PK model is appropriate, the CWRES exhibits a standard normal distribution irrespective of what the base model represents. The RUVsearch tool assesses a series of candidate RUV models for CWRES compared to the base model. Since the candidate models built based on the CWRES are simple, the estimation procedures are substantially faster than directly estimating those RUV models on the original data via NLME modeling.[22] The comprehensive description of the CWRES base model and the included RUV models used in the resmod procedure can be found in Supplementary Material S4. The evaluated models in the tool include (1) a proportional model; (2) a combined proportional and additive model; (3) a power model; (4) an IIV on RUV model; (5) a time-varying error model. The last three models, which relax certain assumptions about the residual model, have been described in detail previously.[22]

In the RUVsearch tool, users can specify which RUV models to test. To enable searches for combinations of different models, a search algorithm similar to the forward stepwise selection method was developed (Figure 4). The search procedures can be described as follows:

(1) *Input model*: an NLME model with proportional RUV is necessary to start the process (i.e., the search starts from the simplest model).
(2) *Resmod approach*: extended RUV models in the search space are evaluated by modeling the CWRES from the input model.
(3) *Likelihood ratio test (LRT)*: the LRT is performed between the CWRES base model and the extended models. If no significant extended model is found, the search is terminated and the tool returns the inputted NLME model as the final model. Otherwise, the model with the lowest p-value is selected as the best RUV model.
(4) *Confirmation with NLME modeling*: the selected model from Step 3, if any, is further verified on the original PK data using an NLME model. If the NLME model with the selected RUV model is also significantly improved compared with the input NLME model based on the LRT, the selection is confirmed. Otherwise, no change is made to the input model and the search is terminated.
(5) *Iteration*: Steps 2–4 are repeated using the selected RUV model from the previous round as the input NLME model. The search is carried out in the search space, excluding the selected models in the previous iteration(s). This process is repeated until the defined termination condition is met in Steps 3 and 4 or no model remains in the search space. It should be noted that the power model and the combined error model describe similar residual characteristics, thus the combination of the two models is excluded by skipping one model if the other one is added.
(6) *Final model*: The last iteration results in the final RUV model.

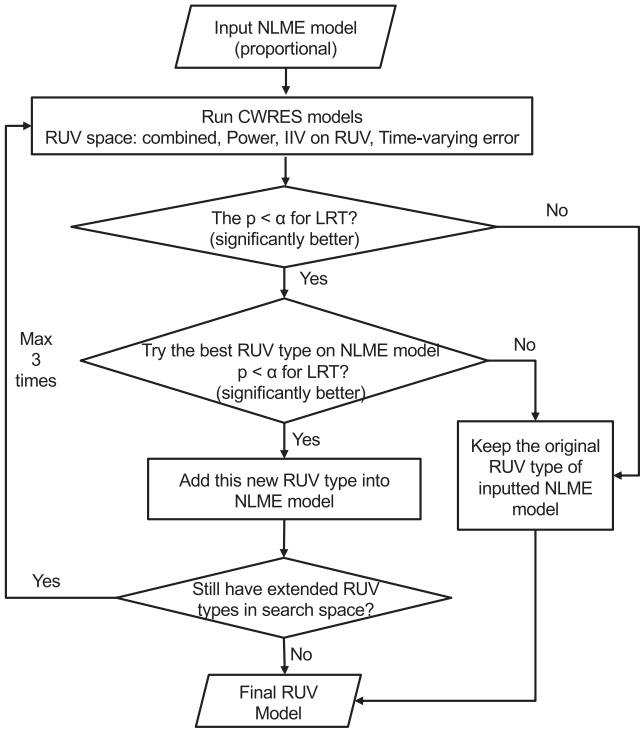

FIGURE 4 Flowchart of the RUVsearch tool. α: Significant level; LRT: likelihood ratio test; NLME: nonlinear mixed-effect modeling; CWRES: conditional weighted residuals.

The AMD tool evaluation

The developed full AMD tool was tested on a series of real datasets sourced from previous publications,[20,24-35] including five drugs for *i.v.* administration and five drugs for oral administration (Table 1). The evaluation focused

on the three modules using novel search algorithms, that is, structure model (S), IIV model (I), and RUV model (R), as well as the impact of the development sequence of those modules (SIR, SRI, or RSI) on the final result. It should be noted that the modules for developing the IOV model, allometric model, and covariate model were not evaluated here. All AMD procedures began from the default starting model, that is, a one-compartment model with first-order absorption (for oral drug) and first-order elimination, including IIV on all parameters and a correlation between clearance and volume of distribution. The search space for the structural modeling module comprised: (1) absorption: first-order (FO), zero-order (ZO), and sequential ZO and FO; (2) absorption delay: a lag time model and transit compartment models (1, 3, and 10 compartments) with a depot compartment; (3) distribution: 1- and 2-compartment models for oral drugs as well as 1-, 2-, and 3-compartment models for *i.v.* drugs; (4) elimination: FO. During the structural model search, the option to keep the IIV model in the starting model without adding IIV on any newly added parameters was used. The starting model of the IIV search step was obtained by adding IIV on all structural parameters in a full block. For the RUV search step, all optional residual error models were included for evaluation. The testing was run using Pharmpy/pharmr versions 0.38.0–0.58.4 with NONMEM 7.5.0 for model estimation. The model developments using the AMD tool were carried out using 20 cores on a cluster. To evaluate the quality of the final models given by the AMD tool, the models were compared through the BIC (Equation 1) and visual predictive checks (VPC) to the simplified published models without considering covariates and IOV if the published models were available for the tested datasets. In addition to real datasets, the AMD tool was further tested in a simulation study (see Supplementary Material S5 for more details).

RESULTS

The AMD tool evaluation

The AMD tool successfully ran on the 10 real datasets. Figure 5a,b show that the BIC dropped compared with the starting model for most steps. The three sequential approaches did not always lead to the same final models (Table 2). The same structural models were given by the three approaches in nine out of 10 cases, with the warfarin example resulting in a different structural model (SIR and SRI: a two-compartment model with a three-transit compartment absorption; RSI: a 1-compartment model with a 10-transit compartment absorption). In contrast to the structural model, the IIV models and the RUV models were found to depend on the sequence of model development selection. Generally, more complex models were obtained when running the RUVsearch tool before the IIVsearch tool. SRI and/or RSI led to more parameters in the final models than SIR, except for desmopressin and warfarin (Table 2). Even in the cases of different final models, they tended to have similar BIC values with the differences in BIC <30 (except gentamicin that had a difference in BIC of 94). The final models with the lowest BIC were provided by SIR for three drugs, SRI for 6 drugs, and RSI for 4 drugs (Table 2). For the datasets with available published models, the final models obtained from the AMD tool exhibited a lower BIC than the modified published models, except for the SIR approach of gentamicin. In addition, the VPC plots of the AMD final models showed reasonable model quality and were comparable to the modified published models (Supplementary Material S6). In the simulation study with 10 simulated datasets, the AMD tool provided the same or more parsimonious models with the same or lower BICs compared with the model used for simulating the datasets. For more details, see Supplementary Material S5.

TABLE 1 The basic information of the 10 real datasets for the AMD tool evaluation.

Drug	Administration	Number of subjects	Number of observations
Daunorubicin[24]	*i.v.*	41	112
FactorVIII[25]	*i.v.*	34	714
Gentamicin[26,27]	*i.v.*	210	574
Pefloxacin[20]	*i.v.*	74	337
Tobramycin[28]	*i.v.*	155	388
Desmopressin[29]	Oral	28	373
Lopinavir[30]	Oral	30	315
Melagatran[31]	Oral	167	1177
Moxonidine[32,33]	Oral	73	1006
Warfarin[34,35]	Oral	32	246

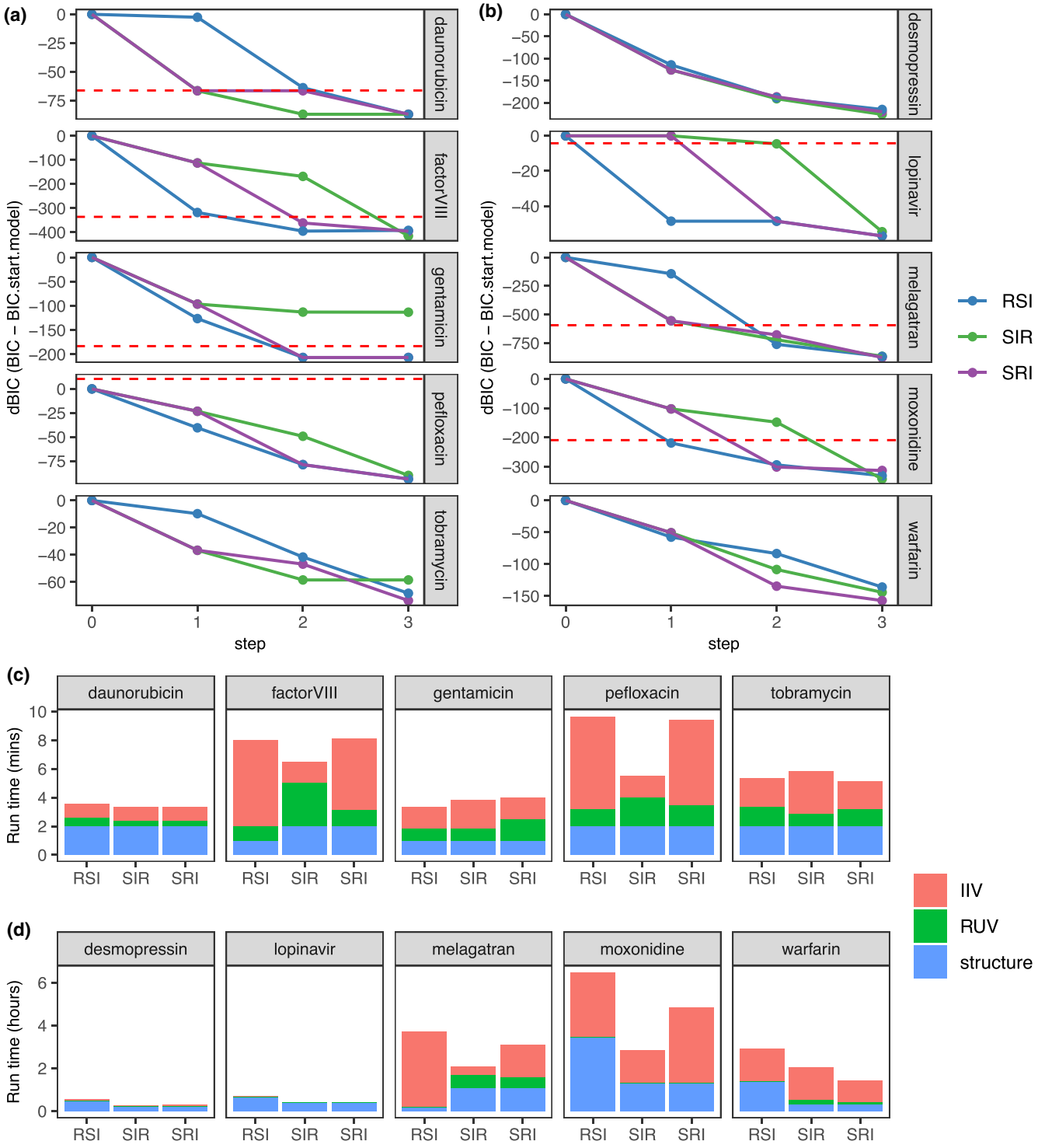

FIGURE 5 The evaluation results of the AMD tool on 10 real datasets including five *i.v.* drugs (a, c) and five oral drugs (b, d). (a) and (b) show the drop of BIC after each step compared with the starting models (Step 0) for three sequences of model selections (SIR, SRI, and RSI). Red dashed lines represent the BIC of the modified published models. Note that published models are not available for all the 10 datasets. (c) and (d) show the running time of the AMD tool. S: Structural model selection; I: Inter-individual variability (IIV) model selection; R: Residual unexplained variability (RUV) model selection.

The total AMD running times (Figure 5c,d) were shorter for *i.v.* drugs (3–10 min) than for oral drugs (17 min to 6.5 h). The selection for RUV models was generally the fastest process due to the high efficiency of the resmod approach and relatively fewer candidate models compared with the selections for structural models and IIV models. The IIV model selection tended to be slower with more parameters in the IIV model search,

TABLE 2 The final models of the AMD tool for three sequences of model selections (SIR, SRI, and RSI) as well as the modified published models and their BIC compared with the starting models (dBIC).

Drug	Selection sequence	Structural model	IIV model	RUV model	Parameter#	dBIC
Daunorubicin	SIR	2 comp	[CL,V2,Q]	Proportional	11	−87
	SRI	2 comp	[CL,V2,Q]	Proportional	11	−87
	RSI	2 comp	[V1,V2,Q]	Power	12	−87
	Published[23]	2 comp	[CL,V1]	Proportional	8	−66
FactorVIII	SIR	2 comp	[CL,V1,Q]+[V2]	Time; power; IIV	14	−415
	SRI	2 comp	[CL,V1,V2,RUV]	Power; IIV	16	−396
	RSI	2 comp	[CL,V1,V2,RUV]	Time; power; IIV	17	−392
	Published[24]	2 comp with baseline	[CL, V1, baseline]	Combined	13	−336
Gentamicin	SIR	2 comp	[CL]+[V2]	Proportional	7	−113
	SRI	2 comp	[CL,V1]	Power	9	−207
	RSI	2 comp	[CL,V1]	Power	9	−207
	Published[26]	2 comp	[CL]+[Q]	Combined	8	−183
Pefloxacin	SIR	2 comp	[CL,V1,V2]	Combined; IIV	13	−90
	SRI	2 comp	[CL,V1,RUV,V2]	Combined; IIV	16	−94
	RSI	2 comp	[CL,V1,V2,RUV]	Combined; IIV	16	−94
	Published[26]	1 comp	[CL]+[V]	Proportional	5	10
Tobramycin	SIR	2 comp	[CL]+[Q]	Proportional	7	−58
	SRI	2 comp	[CL]+[V1]+[Q]	Combined	9	−74
	RSI	2 comp	[CL]+[V1]+[Q]	Power	9	−68
Desmopressin	SIR	1 comp; SEQ-ZO-FO	[CL,V1,MAT,MDT]	Time; Combined; IIV	18	−226
	SRI	1 comp; SEQ-ZO-FO	[CL,V1,MDT]	Time; Combined; IIV	13	−220
	RSI	1 comp; SEQ-ZO-FO	[CL,V1,MDT]	Time; Combined; IIV	13	−215
Lopinavir	SIR	1 comp; FO	[CL]+[MAT]	Time; IIV	8	−55
	SRI	1 comp; FO	[CL,MAT,RUV]	Time; IIV	11	−57
	RSI	1 comp; FO	[CL,MAT,RUV]	Time; IIV	11	−57
	Published[29]	1 comp; FO	[CL]+[V]	Proportional	6	−4
Melagatran	SIR	1 comp; T3	[CL,MAT,MDT]	Time; IIV	13	−867
	SRI	1 comp; T3	[CL,MAT,MDT,RUV]	Time; IIV	16	−877
	RSI	1 comp; T3	[CL,V1,MDT,RUV]	Time; IIV	16	−867
	Published[30]	1 comp; LAG	[CL]+[V]+[MAT]+[LAG]+[F]	Proportional	10	−595
Moxonidine	SIR	2 comp; LAG	[V1]+[CL,V2,Q]	Time; IIV	16	−342
	SRI	2 comp; LAG	[CL,V1,MAT,V2,Q,RUV]	Time; IIV	29	−313
	RSI	2 comp; LAG	[CL,V1,MAT,V2,Q,RUV]	Time; power; IIV	30	−331
	Published[32]	1 comp; Tn	[CL,V]+[KA]+[N]+[MDT]	Proportional	12	−209
Warfarin	SIR	2 comp; T3	[CL,V1,Q]+[MDT]	Combined; Time24	16	−144
	SRI	2 comp; T3	[CL,V1,Q]+[MDT]	Combined; Time6	16	−157
	RSI	1 comp; T10	[CL,V1,RUV]+[MDT]	IIV	12	−136

Note that published models are not available for all 10 datasets.

S: structural model selection; I: inter-individual variability (IIV) model selection; R: residual unexplained variability (RUV) model selection.

Abbreviations for structural models: comp: compartment; SEQ-ZO-FO: sequential zero-order and first-order absorption; LAG: lag time model; T3: absorption model with three-transit compartments; T10: absorption model with 10 transit compartments; Tn: absorption model with the number of transit compartments estimated; FO: first-order absorption.

Abbreviations for IIV models: CL: clearance; MAT: mean absorption time; MDT: mean delay time; V1: central volume of distribution; V2: peripheral volume of distribution; Q: clearance between central and peripheral compartments; F: bioavailability; N: the number of transit compartments; [X,Y] represents there exists covariance parameter between X and Y in the IIV model; [X]+[Y] represents there is no covariance between X and Y in the model.

Abbreviations for RUV models: time: time-varying residual model with numbers representing cutoff time if the cutoff times are different; IIV: IIV on RUV model; combined: combined proportional and additive model; power; power residual model.

as seen for the drugs of factor VIII, pefloxacin, moxonidine, melagatran, and warfarin. Among the five oral drugs, desmopressin and lopinavir had the shortest total run times, which was probably due to a combination of small datasets (<400 observations) and simple models (the selected models had only one depot and one central compartment). The running times for the 10 simulated datasets (oral drug) were also acceptable, ranging from 98 to 125 min.

DISCUSSION

A novel tool for fully automatic model building has been presented here, which is efficient, flexible, and does not require users to provide extra model code. Model building of a population PK model is a complicated process of developing different model components. Those model components have different features, so different search algorithms were used for different model components in the AMD tool instead of using a universal search algorithm. During the development of the search methods, specific features and previous experience related to the corresponding model component were also considered. For example, initial values of population PK parameters have shown great influence on model estimation in the way that inappropriate initial values may lead to local minima or estimation failure.[15] As a result, the reduced stepwise exhaustive method used in the structural model search was designed to be robust against the setting of initial values: (1) a sequential model-building process generating reasonable initial values for downstream models and (2) replicated models with different initial values generated due to the different orders of feature modification. Another example is related to developing an IOV model, where the introduction of an IOV model may affect the existing IIV model.[20] As a result, the IIV on a parameter is re-evaluated after an IOV term is added to the same parameter.

Another important aspect of designing an automatic tool for model building is the selection criterion. In the AMD tool, the LRT and the BIC are used depending on whether the compared candidate models are nested in the respective module. In addition, a model needs to have successful minimization to be selected. Based on the simulation study (Supplementary Material S5), the final models selected by the AMD tool had the same or lower values for both BIC and the number of parameters compared with the simulation model, thus indicating a parsimonious tendency in model selection. This was also a finding made in another simulation-based assessment carried out by Duvnjak Z., et al.,[36] which showed accurate population and individual predicted PK profiles (in spite of more parsimonious models) as well as reasonable condition numbers and parameter uncertainty. In the evaluation study of the AMD tool on the real datasets presented in the current work, the AMD tool provided reasonable final models suggested by comparing them to the modified available published models in terms of BIC values and VPC. It should be noted that the published models may not be the absolute best models to describe the data. The comparison did not include imprecision of parameter estimates, condition numbers, or individual parameter values. More in-depth comparison is planned using the future version of Pharmpy, which may include more functions regarding model evaluation. Regarding the comparison among the three modeling sequences, they did not always give the same final model: the structural models were generally consistent, but the IIV and RUV models sometimes differed. SRI and RSI tended to give more complex final models than SIR. Despite the difference in the final models among the three sequences, their BIC values were the same or had modest differences, and were lower than the modified published models (except for gentamicin), which suggests similar performance. In practice, it is possible to run the AMD tool with different modeling sequence strategies and then choose the best final model.

The AMD tool automates the part that can be automated to accelerate model-building process. It should be the responsibility of modelers to evaluate and interpret the final model from the AMD tool and make a judgment if the model is good enough for the modeling purpose and further application. It may be important to perform more extensive model evaluations on certain key models during the search, for example, the output model from each module. Sometimes, modelers may also want to check other models estimated during the search (especially those with similar BIC values to the final model) for further comparison and finally select one that is biologically and physiologically meaningful. The AMD tool outputs summary tables and all the related model files enabling users to easily check the model-building process.

The AMD tool is designed to be flexible to meet the different preferences and requirements of modelers. Users can define search space and decide which modules are included in the AMD tool and their sequence among other options. Users can provide their own starting model, and use different settings, for example, a different estimation method. Furthermore, users are able to repeat certain modules after the whole AMD process since each of the modeling modules in the AMD workflow can be used independently. For example, the IIV/IOV model may be reduced after adding covariate models, thus it would be worthy of re-evaluation. The AMD tool provides different modeling strategies for certain modules. There are different options for how IIVs are

added to new parameters during the structural model search (Figure S3-3). A comparison study showed that different IIV strategies may lead to different final models.[37] Using the strategies with more random-effect parameters to estimate, the running time was generally longer, and the probability of estimation errors was higher compared with the other strategies.

When designing an automatic model development tool, one important factor to consider is the efficiency of the search algorithm. During the development of the search algorithms of the AMD tool, the aim was to optimize efficiency with a balance between the final model quality and running times. In addition, the resmod method was used to increase the efficiency of the RUV model development. The running times in the evaluation study were deemed to be acceptable for most model development scenarios. It should be noted that the structural model search space set in the evaluation study did not include all the models listed in Figure 2a. Models with nonlinear elimination were not evaluated due to the lack of evidence of nonlinear PK for the 10 tested drugs. By doing so, the running time can be substantially reduced as the nonlinear elimination model uses a numerical solver instead of explicit solutions for the concentration calculation. The AMD tool uses analytical solutions for calculation and estimation whenever it is possible for the sake of high efficiency. In addition to the search space, the complexity of the evaluated models affected running times. For example, an absorption delay model with 10 transit compartments takes longer to run compared with that with three-transit compartments. Moreover, the two-step exhaustive method used in the IIV model search performs well when the number of IIV parameters is low. However, when the number is higher (≥6), many more candidate models will be generated and the running time can increase substantially. Other factors such as the dataset size (number of subjects or number of observations) may also affect the running times of the AMD tool. In practice, there may be different ways to improve the efficiency of using the AMD tool for model development. One way is to set an appropriate search space based on prior information. In the early phase of clinical trials, PK data usually consist of a small sample size but rich sampling data, and thus the AMD process with an extensive search space may not take a long time. On the other hand, for the later-phase clinical trials with a large pooled dataset, the search space may be reduced based on prior information to shorten running times.

The challenges of developing a fully automatic tool for population PK or PKPD model building are to make it versatile, robust, and efficient. That is also our goal for the developed AMD tool. The current version (v1.1.0) of the AMD tool has incorporated a series of new features. It can not only analyze PK datasets with either intravenous or extravascular administration, but also handle more complicated analyses, including multiple administration routes, metabolite data, data with below the limit of quantification (BLQ), PKPD modeling, and target-mediated drug disposition (TMDD) models. In addition, recent improvements include the selected model from each module being rerun with tweaked initial estimates to increase robustness against local minima. The new updates also include adding more options for model selection strictness to improve the quality of selected models, such as uncertainty evaluation. More flexibility in the AMD workflow is also available, such as carrying out certain modules twice (e.g., re-evaluation of IIV and RUV models after running covariate search or multiple sequential covariate searches). In the future, other search algorithms will be explored to increase the efficiency of building random-effect models and covariate models. A future version will be able to use nlmixr2 in addition to NONMEM for model estimation so that even the estimation engine used in the AMD tool may be freely available.

CONCLUSIONS

The AMD tool presented here is a systematic, extensive, efficient, flexible, and fully automatic tool for developing a population PK model. The AMD tool was successfully applied on a collection of datasets with acceptable running times and reasonable quality of the final models as demonstrated by the VPC and compared with the available published models. The comparison in different sequences of model component development did not show a clear superiority of any one sequence. In summary, the developed AMD tool serves as a promising tool to facilitate and accelerate the use of modeling and simulation in the MID3.

AUTHOR CONTRIBUTIONS

X.C., R.N., S.B., A.H., S.W., T.Y., Z.H., S.J.C., S.B., J.A.A., A.C.H., and M.O.K. wrote the manuscript. X.C., R.N., S.B., A.H., S.W., T.Y., S.B., J.A.A, A.C.H., and M.O.K. designed the research. X.C., R.N., S.B., A.H., S.W., and T.Y. performed the research. X.C., A.H., S.W., and T.Y. analyzed the data.

ACKNOWLEDGMENTS

The authors appreciate Emilie Schindler, Franziska Schaedeli Stark, Nicolas Frey, Sylvie Retout, and Valérie Cosson for testing the AMD tool and for their feedback during tool development. The application of the AMD tool on the real data and simulated data was enabled by resources in project NAISS 2022/5-584 provided by the National Academic Infrastructure for Supercomputing

in Sweden (NAISS) at UPPMAX, funded by the Swedish Research Council through the grant agreement no. 2022-06725.

FUNDING INFORMATION

This work was supported by F. Hoffmann-La Roche Ltd., Basel, Switzerland. Part of this work was produced during a scholarship period for master's studies (Alzahra Hamdan) funded by the Swedish Institute.

CONFLICT OF INTEREST STATEMENT

S.B. and J.A.A. are employees of F. Hoffmann-La Roche Ltd. S.W. is a former employee of Uppsala University and is currently employed with Simulations Plus. S.J.C. is a former employee of Uppsala University and is currently employed with AstraZeneca. A.C.H. and M.O.K. have received consultancy fees from, and own stock in Pharmetheus, all unrelated to this manuscript. All other authors declared no competing interests for this work. As an Associate Editor for *CPT: Pharmacometrics & Systems Pharmacology*, Andrew Hooker was not involved in the review or decision process for this paper.

ORCID

Xiaomei Chen https://orcid.org/0009-0002-0663-0532
Rikard Nordgren https://orcid.org/0009-0005-4428-6527
Stella Belin https://orcid.org/0009-0002-5275-7964
Alzahra Hamdan https://orcid.org/0009-0002-3102-2707
Shijun Wang https://orcid.org/0009-0004-9640-6001
Tianwu Yang https://orcid.org/0009-0005-0856-4497
Zhe Huang https://orcid.org/0009-0006-8674-5071
Simon J. Carter https://orcid.org/0000-0002-8220-2321
Simon Buatois https://orcid.org/0000-0002-1472-5366
João A. Abrantes https://orcid.org/0000-0002-7752-0229
Andrew C. Hooker https://orcid.org/0000-0002-2676-5912
Mats O. Karlsson https://orcid.org/0000-0003-1258-8297

REFERENCES

1. Milligan PA, Brown MJ, Marchant B, et al. Model-based drug development: a rational approach to efficiently accelerate drug development. *Clin Pharmacol Ther*. 2013;93:502-514.
2. Madabushi R, Seo P, Zhao L, Tegenge M, Zhu H. Review: role of model-informed drug development approaches in the lifecycle of drug development and regulatory decision-making. *Pharm Res*. 2022;39:1669-1680.
3. Wang YN, Zhu H, Madabushi R, Liu Q, Huang SM, Zineh I. Model-informed drug development: current US regulatory practice and future considerations. *Clin Pharmacol Ther*. 2019;105:899-911.
4. de Velde F, Mouton JW, de Winter BCM, van Gelder T, Koch BCP. Clinical applications of population pharmacokinetic models of antibiotics: challenges and perspectives. *Pharmacol Res*. 2018;134:280-288.
5. Mould DR, Upton RN. Basic concepts in population modeling, simulation, and model-based drug development-part 2: introduction to pharmacokinetic modeling methods. *CPT Pharmacometrics Syst Pharmacol*. 2013;2:e38.
6. Jonsson EN, Karlsson MO. Automated covariate model building within NONMEM. *Pharm Res*. 1998;15:1463-1468.
7. Svensson RJ, Jonsson EN. Efficient and relevant stepwise covariate model building for pharmacometrics. *CPT Pharmacometrics Syst Pharmacol*. 2022;11:1210-1222.
8. Ayral G, Si Abdallah JF, Magnard C, Chauvin J. A novel method based on unbiased correlations tests for covariate selection in nonlinear mixed effects models: the COSSAC approach. *CPT Pharmacometrics Syst Pharmacol*. 2021;10:318-329.
9. Prague M, Lavielle M. SAMBA: a novel method for fast automatic model building in nonlinear mixed-effects models. *CPT Pharmacometrics Syst Pharmacol*. 2022;11:161-172.
10. Sale M, Sherer EA. A genetic algorithm based global search strategy for population pharmacokinetic/pharmacodynamic model selection. *Br J Clin Pharmacol*. 2015;79:28-39.
11. Ismail M, Sale M, Yu Y, et al. Development of a genetic algorithm and NONMEM workbench for automating and improving population pharmacokinetic/pharmacodynamic model selection. *J Pharmacokinet Pharmacodyn*. 2022;49:243-256.
12. Nordgren R, Belin S, Yngman G, et al. Pharmpy: a versatile open-source library for pharmacometrics. *PAGE 30* Abstr 10029 [www.page-meeting.org/?abstract=10029].
13. Van Rossum G, Drake FL. *Python 3 reference manual*. CreateSpace; 2009.
14. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2021.
15. Bauer RJ. NONMEM tutorial part I: description of commands and options, with simple examples of population analysis. *CPT Pharmacometrics Syst Pharmacol*. 2019;8:525-537.
16. Qutishat O, Carter SJ, Nordgren R, et al. The development of artificial neural networks for the prediction of influential individuals and outlying individuals and their application during the model building process. *PAGE 30* Abstr 10029 [www.page-meeting.org/?abstract=10029]
17. Lavielle M. *Mixed Effects Models for the Population Approach: Models, Tasks, Methods and Tools*. 1st ed. Chapman and Hall/CRC; 2014.
18. Rowland M, Tozer TN. *Clinical Pharmacokinetics and Pharmacodynamics: Concepts and Applications*. Wolters Kluwer Health/Lippincott William & Wilkins; 2011.
19. Delattre M, Lavielle M, Poursat MA. A note on BIC in mixed-effects models. *Electron J Stat*. 2014;8:456-475.
20. Karlsson MO, Sheiner LB. The importance of modeling interoccasion variability in population pharmacokinetic analyses. *J Pharmacokinet Biopharm*. 1993;21:735-750.
21. Delattre M, Poursat MA. An iterative algorithm for joint covariate and random effect selection in mixed effects models. *Int J Biostat*. 2016;16:75-83.
22. Ibrahim MMA, Nordgren R, Kjellsson MC, Karlsson MO. Model-based residual post-processing for residual model identification. *AAPS J*. 2018;20:1-9.
23. Hooker AC, Staatz CE, Karlsson MO. Conditional weighted residuals (CWRES): a model diagnostic for the FOCE method. *Pharm Res*. 2007;24:2187-2197.

24. Bogason A, Quartino AL, Lafolie P, et al. Inverse relationship between leukaemic cell burden and plasma concentrations of daunorubicin in patients with acute myeloid leukaemia. *Br J Clin Pharmacol*. 2011;71:514-521.
25. Björkman S, Folkesson A, Jönsson S. Pharmacokinetics and dose requirements of factor VIII over the age range 3-74 years: AA population analysis based on 50 patients with long-term prophylactic treatment for haemophilia a. *Eur J Clin Pharmacol*. 2009;65:989-998.
26. Rosario MC, Thomson AH, Jodrell D, Sharp CA, Elliott HL. Population pharmacokinetics of gentamicin in patients with cancer. *Br J Clin Pharmacol*. 1998;46:229-236.
27. Wählby U, Thomson AH, Milligan PA, Karlsson MO. Models for time-varying covariates in population pharmacokinetic-pharmacodynamic analysis. *Br J Clin Pharmacol*. 2004;58:367-377.
28. Aarons L, Vozeh S, Wenk M, Weiss P, Follath F. Population pharmacokinetics of tobramycin. *Br J Clin Pharmacol*. 1989;28:305-314.
29. Østerberg O, Savic RM, Karlsson MO, et al. Pharmacokinetics of desmopressin administrated as an oral lyophilisate dosage form in children with primary nocturnal enuresis and healthy adults. *J Clin Pharmacol*. 2006;46:1204-1211.
30. Elsherbiny D, Ren Y, McIlleron H, Maartens G, Simonsson USH. Population pharmacokinetics of lopinavir in combination with rifampicin-based antitubercular treatment in HIV-infected south African children. *Eur J Clin Pharmacol*. 2010;66:1017-1023.
31. Bååthe S, Hamrén B, Karlsson MO, Wollbratt M, Grind M, Eriksson UG. Population pharmacokinetics of melagatran, the active form of the oral direct thrombin inhibitor ximelagatran, in atrial fibrillation patients receiving long-term anticoagulation therapy. *Clin Pharmacokinet*. 2006;45:803-819.
32. Karlsson MO, Jonsson EN, Wiltse CG, Wade JR. Assumption testing in population pharmacokinetic models: illustrated with an analysis of moxonidine data from congestive heart failure patients. *J Pharmacokinet Biopharm*. 1998;26:207-246.
33. Savic RM, Jonker DM, Kerbusch T, Karlsson MO. Implementation of a transit compartment model for describing drug absorption in pharmacokinetic studies. *J Pharmacokinet Pharmacodyn*. 2007;34:711-726.
34. O'Reilly RA, Aggeler PM, Leong LS. Studies on the Coumarin anticoagulant drugs: the pharmacodynamics of warfarin in man. *J Clin Invest*. 1963;42:1542-1551.
35. O'Reilly RA, Aggeler PM, Services M, Francisco S, Hospital G. Studies on Coumarin anticoagulant drugs. *Circulation*. 1968;38:169-177.
36. Duvnjak Z, Schaedeli Stark F, Cosson V, Retout S, Schindler E, Abrantes JA. Simulation-based evaluation of the Pharmpy Automatic Model Development tool for population pharmacokinetic modeling in early clinical drug development. *CPT Pharmacometrics Syst Pharmacol*. 2024;1-15. doi:10.1002/psp4.13213
37. Hamdan A, Chen X, Belin S, et al. Development of pharmacokinetic structural models – Pharmpy model search tool. *PAGE 30* Abstr 10020 [www.page-meeting.org/?abstract=10020].

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Chen X, Nordgren R, Belin S, et al. A fully automatic tool for development of population pharmacokinetic models. *CPT Pharmacometrics Syst Pharmacol*. 2024;13:1784-1797. doi:10.1002/psp4.13222