Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Boehringer Ingelheim

FAKULTÄT
FÜR MATHEMATIK, INFORMATIK
UND NATURWISSENSCHAFTEN

Evaluation of Machine-learning-informed Optimal Study Design in Preclinical Intravenous Rat Pharmacokinetic Studies

Keqi Shi (1,2), Jens M. Borghardt (2), Niklas Kroemer (2) & Sebastian G. Wicha (1)
(1) Dept. of Clinical Pharmacy, Institute of Pharmacy, University of Hamburg, Hamburg, Germany
(2) Global Drug Discovery Sciences, Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach, Germany

Introduction

Pharmacokinetic (PK) understanding in preclinical species is important for translation and prediction of human PK in drug discovery. Since **preclinical PK studies** of new drugs are mostly designed empirically and compound-agnostically, **optimal design (OD)** methodology incorporating compound information and therefore guiding the experimental design of these studies could represent a well-recognized tool to improve the quality of PK parameter estimates.

While OD is well recognized in clinical drug development [1], robust prior knowledge is not as accessible at the preclinical stage. Emerging **machine learning (ML)** predictions of PK parameters [2] in the preclinical field could overcome this limitation. Therefore, the aim of this study was evaluating the potential **benefit of ML-informed OD** on the identifiability of key PK parameters, i.e., clearance (CL) & volume of distribution at steady-state (V_{ss}), and recovery of PK profile in rat IV studies.

Methods

Data:

PK data after intravenous administration of over 2000 small molecules, alone or as a cocktail, in rats were included. These data comprised study designs, estimated compartmental PK parameters (hereafter referred as true parameters), and their corresponding ML predictions.

Study Design Optimization:

The optimized experimental designs were obtained by optimizing sampling times based on the original designs using R package PopED [3]. Five tested scenarios are listed in ***Table 1.***

Tab.1 Elaboration table of five design groups under different scenarios.

Design Scenarios	A priori used in OD		Optimality Criterion	
	ML-predicted	True*	D-opt	ED-opt
Reference standard design				
Realistic-case D-optimized design	√		√	
Realistic-case ED-optimized design	√			√
Best-case D-optimized design		√	√	
Best-case ED-optimized design		√		√

D/ED-opt=D/ED-optimality. The true parameter is marked with "*".

Study Design Evaluation:

PK study outcomes under different design scenarios were evaluated using stochastic simulation and re-estimation (SSE) (***Fig. 1***). Design performance was assessed based on recovery of key PK parameters and PK profile. The applied PK parameters recovery metric evaluates both parameter accuracy and precision, and is quantified as the proportion of designs for which the 95% confidence intervals (CI) of CL and V_{ss} fall within a predefined fold-error (FE) threshold, defined as:

$$P_{KeyPar}(FE) = \frac{1}{N}\sum_{i=1}^{N}(I_{i,j=CL}(FE) \cap I_{i,j=Vss}(FE))$$

Where: $$I_{i,j}(FE) = 1\left\{CI_{\theta_{i,j}} \in \left(\frac{\theta^*_{i,j}}{FE}, \theta^*_{i,j} \cdot FE\right)\right\}$$

$$CI_{i,j} = (\theta_{i,j} - 1.96 \cdot SE_{i,j}, \theta_{i,j} + 1.96 \cdot SE_{i,j})$$

PK profile recovery assessing accuracy of all PK parameters is quantified similarly but instead with Curve Similarity Index (CSI) within FE threshold, defined as:

$$P_{Shape}(FE) = \frac{1}{N}\sum_{i=1}^{N} 1\{CSI_i < FE\}$$

Where: $$CSI_i = 1 + \frac{AUC_{delta\_conc_i}}{AUC_{true_i}}$$

$$AUC_{delta\_conc_i} = \int_{t=0}^{t=24h} |C(t,\beta_i) - C(t,\beta_i^*)|\, dt$$

$$AUC_{true_i} = \int_{t=0}^{t=24h} C(t,\beta_i^*)\, dt$$

for parameter $j \in \{CL, V_{SS}\}$, i^{th} study design and parameter vector β

Sensitivity Analysis:

Sensitivity was tested against abundance of sample points and residual error size used in OD.

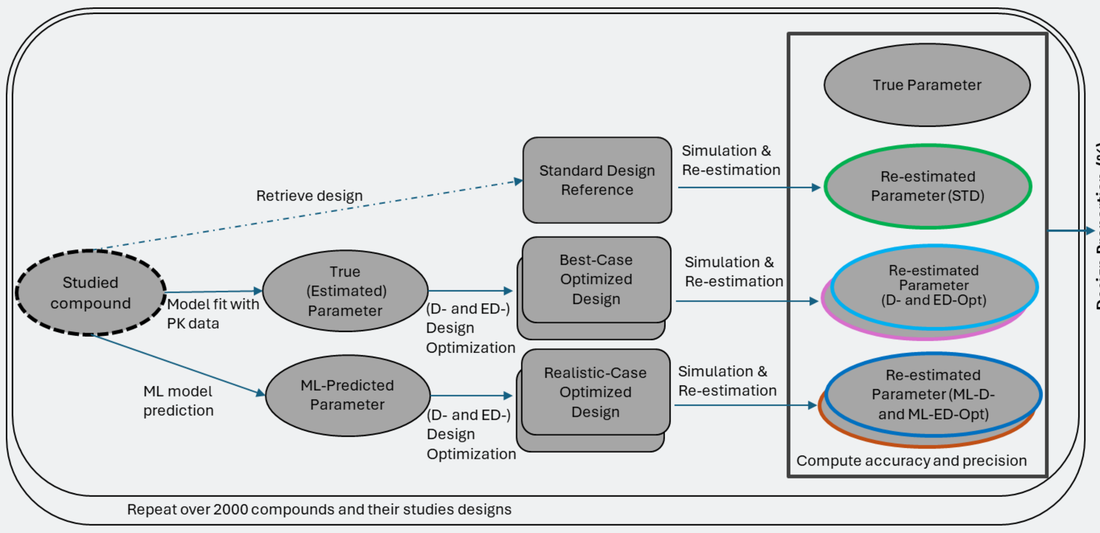

Fig. 1 Workflow and method diagram.
STD = standard design;
D-opt = design optimized with D-optimality with true parameters;
ED-opt = design optimized with ED-optimality with true parameters;
ML-D-opt = design optimized with D-optimality informed by machine learning;
ML-ED-opt = design optimized with ED-optimality informed by machine learning.

Results

- In best-case scenario, designs optimized based on true parameter and D-optimality showed overall superiority for all model types compared to the standard design, with improvement in qualified design proportion up to 21.9% (FE=1.4, P_{Std}=42.1%), and 13.7% (FE=1.8, P_{Std}=50.2%) for 1-cmt IV and 2-cmt IV models, respectively.
- In realistic scenario, ED-optimality was superior in compensating for ML prediction uncertainty (***Fig. 2***).
- In realistic scenario, ML-informed OD managed to outperform standard design in 1-cmt models (up to 11.1%, FE=1.5, P_{Std}=63.1%) but did not improve 2-cmt models (***Fig. 3***).

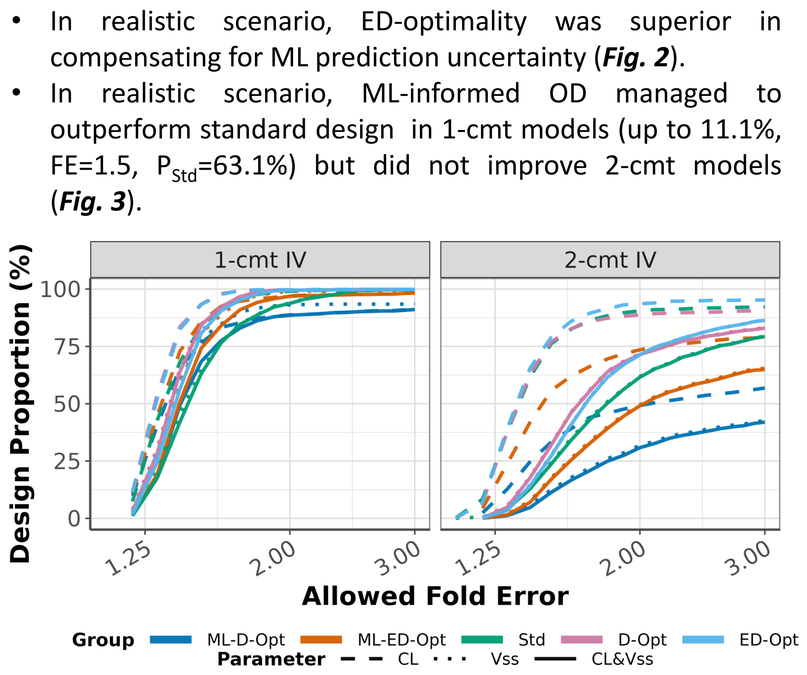

Fig. 2 Design performance across PK models for five design groups measured by key parameters recovery.

- The PK profile shape recovery results confirmed the performance pattern across scenarios, showing benefit (4.7%) of ML-informed OD only in 1-cmt models (***Fig. 3***).

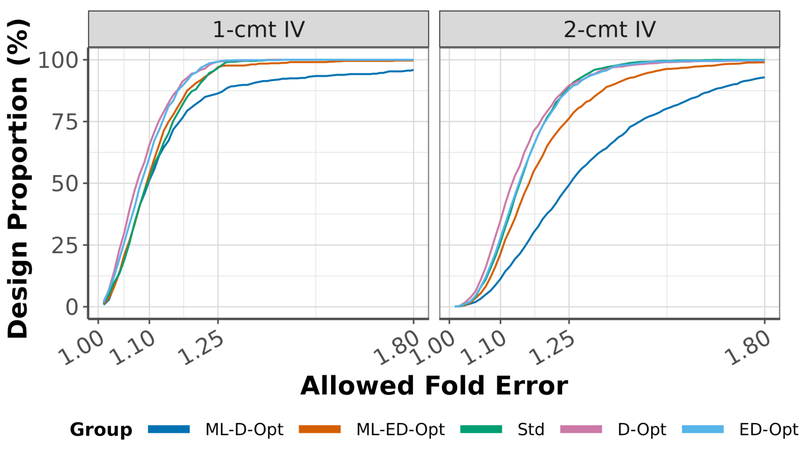

Fig. 3 Design performance across PK models for five design groups measured by PK profile recovery.

- The sensitivity analysis with varying numbers of sampling points (***Fig. 4***) and different levels of residual error (***Fig. 5***) corroborated abovementioned findings.
- Notably, the benefit of OD became more pronounced for reduced-sampling design in 2-cmt models, where optimized design outperforms the empirical standard design (***Fig. 4***)
- Lower residual error improved performance at stricter fold-error thresholds under the best-case scenario. However, when parameter priors were misspecified (realistic scenario), low residual error amplified the negative impact of this misspecification, resulting in a more pronounced degradation in design performance (***Fig. 5***).

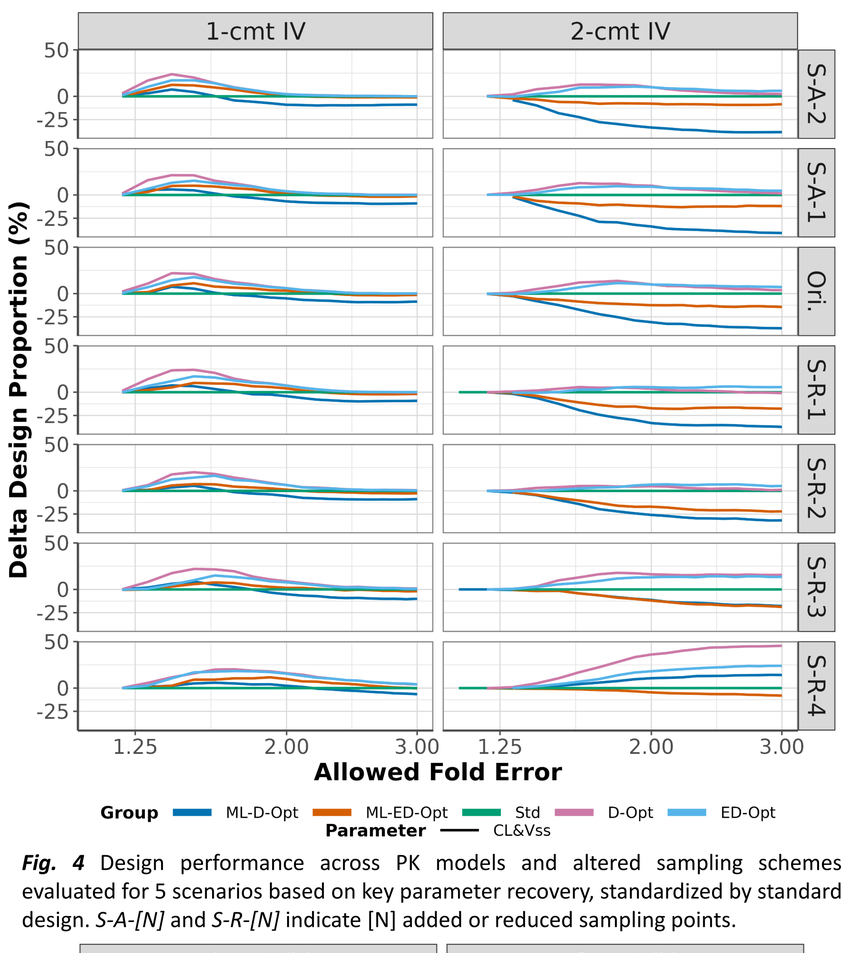

Fig. 4 Design performance across PK models and altered sampling schemes evaluated for 5 scenarios based on key parameter recovery, standardized by standard design. *S-A-[N]* and *S-R-[N]* indicate [N] added or reduced sampling points.

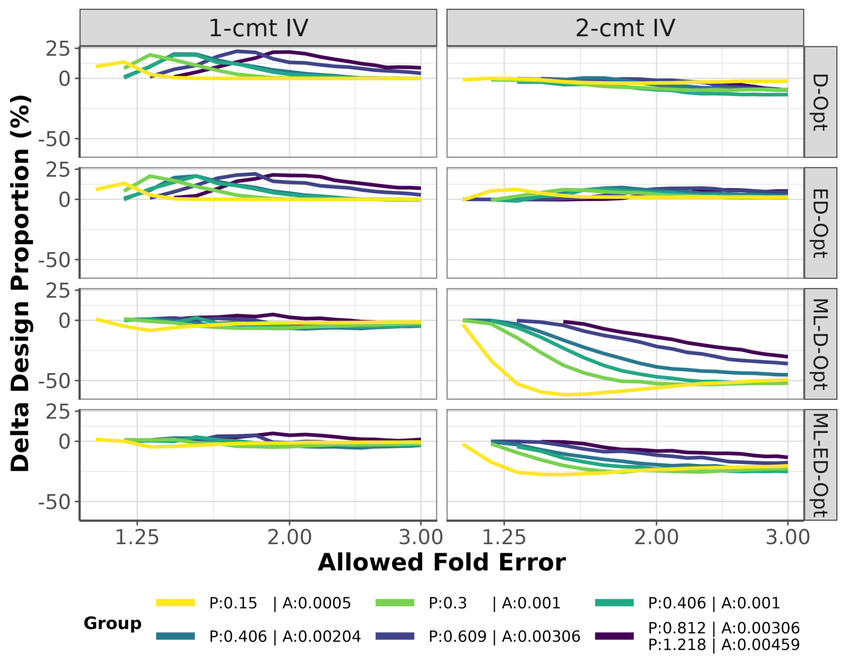

Fig. 5 Design performance across PK models and optimization groups with varying size of residual error, measured by key parameters recovery, standardized by standard design. P: proportional error; A: additive error (mM).

Conclusion

In realistic scenario, ML-informed OD provides marginal performance gains (~10%) for 1-cmt IV models under current ML prediction accuracy and with already adequate standard design. The main limitation remains the dependence on robust prior information. Further work is needed to define the gap, e.g., the level of ML improvement and technique required for ML-informed OD to provide consistent and practical benefits.

Acknowledgements

Boehringer Ingelheim Pharma GmbH & Co. KG provided funding.

Literature

[1] Mentré, F. *et al. CPT: Pharmacomet. Syst. Pharmacol.* **2**, 1–2 (2013).
[2] Walter, M. *et al. Clinical Translational Sci* **18**, e70150 (2025).
[3] Nyberg, J. *et al. Computer Methods and Programs in Biomedicine* **108**, 789–805 (2012).

Additional Acknowledgements

Moritz Walter, Miha Skalic, Lina Humbeck, Christofer Tautermann, and Cornelius Truenkle for their contribution to the ML predicted parameters and in vivo study insights.

visit us on
https://www.chemie.uni-hamburg.de