

# PARAMETER SIGNIFICANCE TEST USING MIXTURE MODELS (PaSTUM) ALLOWING TYPE-1 ERROR CONTROL FOR EXPOSURE RESPONSE MODELLING

Daniel Wojtyniak (1,2), Jinju Guk (2), Sebastian G. Wicha (1)

(1) Dept. of Clinical Pharmacy, Institute of Pharmacy, University of Hamburg, Germany

(2) TMCP Clinical Pharmacology, Boehringer Ingelheim Pharma GmbH & Co. KG

## Introduction

- Inflated Type I error (T1) rate in exposure-response (E-R) modelling can lead to a false-positive or ill-informed decision making in drug development.
- To mitigate the risk, individual model averaging (IMA) [1] was introduced to control T1 in treatment-response analysis (problems in E-R analysis).
- The aim of this study is to suggest a new method to extend IMA to E-R analysis by randomly imputing a missing variable (i.e. exposure).

## Methods

### Study-Design:

- The simulation study was conducted on a hypothetical antidiabetic drug with effect on fasting blood glucose.
- Three different direct (linear, log-linear and emax) and one indirect structural model (turnover with emax effect on Kout).
- Simulation and estimation was performed in NONMEM version 7.5.0.
- Each structural model consisted of 2- or 3- different dose groups (low, medium, high).
- with 3-, 10-, 30- or 100- patients per dose group as well as a placebo arm of equal size.
- All doses were given via 30 minute i.v. infusions every 8 hours for 3 days.
- Pharmacokinetic (PK) measurements were taken at: 0, 1, 3, 5, 7.9, 15.5, 23, 47.9, 66, 67, 68, 71.9 hours.
- Pharmacodynamic (PD) measurements were taken at: 0 and 73 hours.

### Parameter Significance Test Using Mixture Models (PaSTUM):

- In E-R the drug effect is dependent on an exposure metric like area under the concentration time curve (AUC).
- Applying IMA on an E-R analysis: Placebo patients cannot inform the mixture proportion parameter
- Placebo patients → both sub models are identical
- PaSTUM is based on the same idea as IMA: → using mixture models to test significance of arm information for model description → but extends IMA with randomly imputing exposure data for placebo patients.

### Models

- STA base model: DRUG = EFFECT · 0
- STA full model: DRUG = EFFECT · ARM

- PaSTUM base model: P(1) = 0.5

- PaSTUM full model:

$$P(1) = \theta_{mix} \cdot ARM + (1 - \theta_{mix}) \cdot (1 - ARM)$$

- PaSTUM sub model 1: DRUG = EFFECT } Identical in
- PaSTUM sub model 2: DRUG = 0 } base and full model

### Workflow

Each scenario analysis followed these workflows (Figure 1).

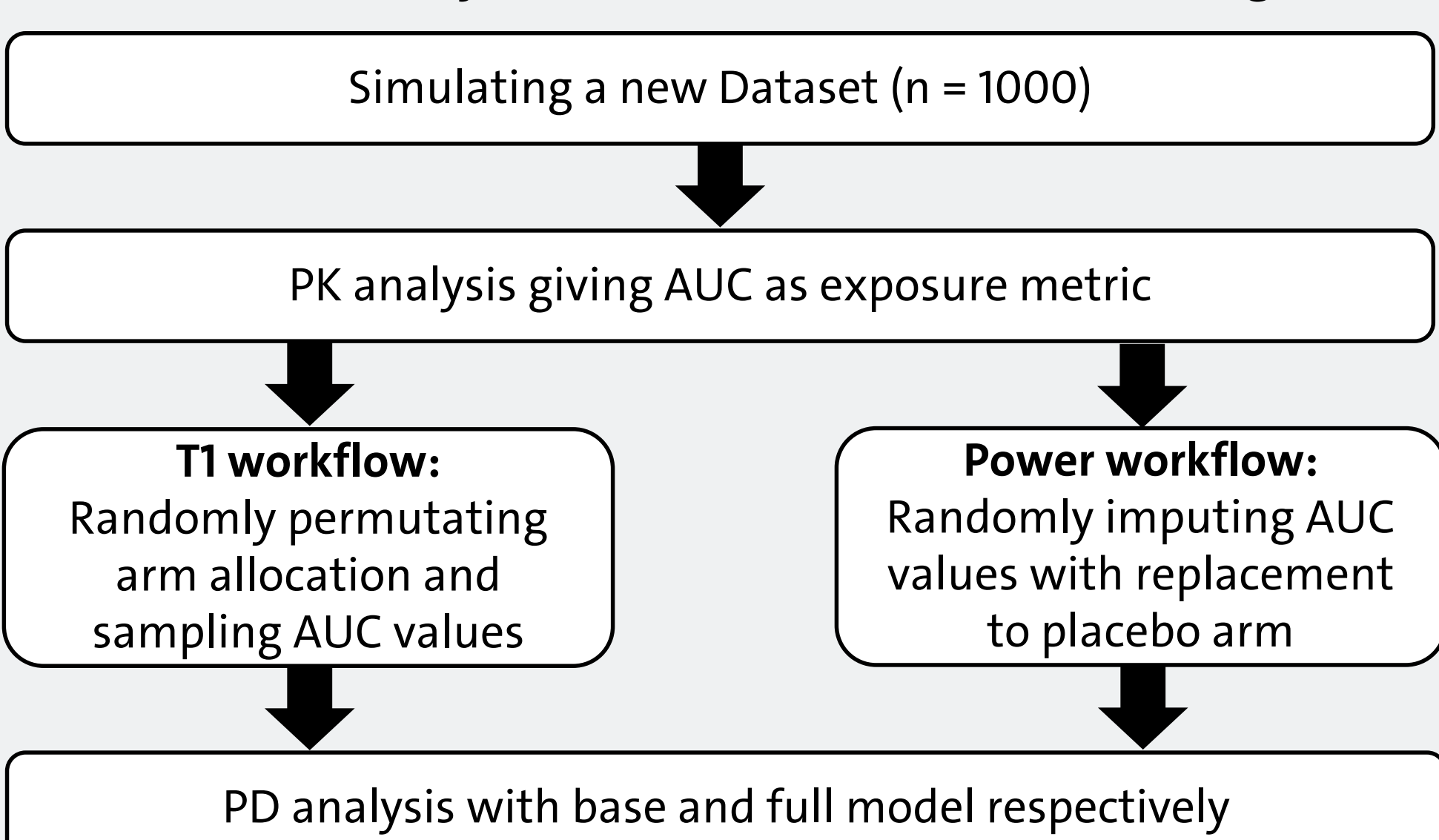


Figure 1: T1 rate and power setting for the simulation and estimation workflow

### Data Analysis:

- T1 rate and power calculation (Equation 1).

$$T1 \text{ rate/power} = \frac{|\Delta OFV|}{n} > \alpha_{0.5} \quad (1)$$

where [...] counts how often the condition inside is met,  $\Delta OFV$  is the difference in the objective function values between full and base model,  $\alpha_{0.5}$  is the critical chi-square value at the significance level  $\alpha = 0.05$  with degrees of freedom for PaSTUM being 1 and for STA being the number of parameters added for the drug effect,  $n$  is the number of simulations done.

- Predictive performance was evaluated by the relative root mean square error (rRMSE) (Equation 2) as well as relative bias (rBias) (Equation 3).

$$rRMSE = \sqrt{\frac{1}{n \cdot m} \cdot \sum_{i=1}^n \sum_{j=1}^m \frac{(\text{pred}_{i,j} - \text{obs}_{i,j})^2}{\text{obs}_{i,j}^2}} \cdot 100 \quad (2)$$

$$rBias = \frac{1}{n \cdot m} \cdot \sum_{i=1}^n \sum_{j=1}^m \frac{\text{pred}_{i,j} - \text{obs}_{i,j}}{\text{obs}_{i,j}} \cdot 100 \quad (3)$$

where  $n$  is the total number of individuals,  $m$  is the total number of observations per individual,  $\text{pred}_{i,j}$  is the prediction for FPG for the  $i$ th individual at the  $j$ th time point,  $\text{obs}_{i,j}$  is the observed FPG for the  $i$ th individual at the  $j$ th time point.

- Precision and Accuracy for the E-R parameter estimates (power workstream) were compared against the estimates of the full STA model with 600 patients ("true" parameter) (Equation 4-5).

$$\bar{X} = \sum_{k=1}^s \frac{X_k}{s} \quad (4)$$

$$\sigma = \sqrt{\frac{1}{s-1} \cdot \sum_{k=1}^s (X_k - \bar{X})^2} \quad (5)$$

where  $\bar{X}$  is the mean of a predicted effect parameter for  $s$  simulations for a specific setting,  $X_k$  is the estimation for the  $k$ th simulation,  $s$  is the total number of simulations for each setting,  $\sigma$  is the standard deviation of a predicted effect parameter for  $S$  simulations for a specific setting.

### Success Criteria:

- T1 rate < 6.54% (upper bound of the 95% confidence interval (CI) of a 5% T1 rate)
- Statistical power > 80%
- rRMSE smaller than for STA model
- rBias 95% (CI) including 0
- Parameter estimates 95% CI including "true" parameter

## Results

- PaSTUM T1 rate, was controlled in most scenarios (48/64), while remaining sufficient power (64/64).
- STA showed no controlled T1 rate (0/64) (Table 1), while showing only slightly higher power (64/64) (Table 2).

Table 1: T1 rate for PaSTUM and STA

| Setting   | PaSTUM |           |       |          | STA    |           |       |          |
|-----------|--------|-----------|-------|----------|--------|-----------|-------|----------|
|           | linear | loglinear | emax  | turnover | linear | loglinear | emax  | turnover |
| L-H-3     | 9.97%  | 3.94%     | 8.64% | 9.65%    | 46.2%  | 12.5%     | 45.3% | 23.0%    |
| L-H-10    | 6.40%  | 4.46%     | 4.99% | 5.23%    | 64.5%  | 27.0%     | 73.2% | 53.5%    |
| L-H-30    | 6.74%  | 2.83%     | 4.21% | 5.50%    | 81.6%  | 80.2%     | 95.8% | 91.7%    |
| L-H-100   | 7.01%  | 2.40%     | 2.50% | 10.6%    | 95.3%  | 100%      | 99.7% | 99.1%    |
| L-M-3     | 9.27%  | 5.49%     | 7.30% | 8.43%    | 38.6%  | 16.0%     | 30.8% | 22.8%    |
| L-M-10    | 6.18%  | 5.48%     | 6.35% | 6.64%    | 60.9%  | 41.8%     | 57.8% | 54.7%    |
| L-M-30    | 5.90%  | 3.07%     | 4.70% | 5.58%    | 87.2%  | 92.2%     | 90.3% | 97.1%    |
| L-M-100   | 5.40%  | 4.00%     | 5.20% | 6.28%    | 99.5%  | 100%      | 100%  | 99.9%    |
| M-H-3     | 7.40%  | 4.42%     | 7.20% | 8.42%    | 37.9%  | 16.5%     | 37.7% | 18.4%    |
| M-H-10    | 4.51%  | 3.75%     | 5.71% | 4.47%    | 48.5%  | 48.7%     | 59.5% | 54.5%    |
| M-H-30    | 4.91%  | 4.20%     | 5.21% | 5.69%    | 42.6%  | 95.9%     | 80.1% | 96.3%    |
| M-H-100   | 5.11%  | 4.40%     | 5.50% | 6.11%    | 18.4%  | 100%      | 91.0% | 100%     |
| L-M-H-3   | 5.70%  | 4.31%     | 6.81% | 6.64%    | 47.8%  | 15.4%     | 52.6% | 28.3%    |
| L-M-H-10  | 5.48%  | 2.53%     | 5.65% | 6.11%    | 67.0%  | 49.3%     | 78.1% | 71.6%    |
| L-M-H-30  | 5.02%  | 1.80%     | 3.40% | 5.03%    | 76.7%  | 97.4%     | 95.9% | 98.8%    |
| L-M-H-100 | 5.81%  | 3.40%     | 5.40% | 7.07%    | 76.2%  | 100%      | 99.7% | 98.4%    |

H: high dose, L: low dose, M: medium dose, PaSTUM: Parameter significance test using mixture models, STA: standard approach, blue: T1 0-3.72%, green: T1 3.81-6.53%, red: T1 50-100%, yellow: T1 6.55-49.9%

Table 2: Statistical power for PaSTUM and STA

| Setting   | PaSTUM |           |       |          | STA    |           |      |          |
|-----------|--------|-----------|-------|----------|--------|-----------|------|----------|
|           | linear | loglinear | emax  | turnover | linear | loglinear | emax | turnover |
| L-H-3     | 98.5%  | 94.7%     | 95.6% | 97.8%    | 100%   | 100%      | 100% | 100%     |
| L-H-10    | 100%   | 100%      | 100%  | 100%     | 100%   | 100%      | 100% | 100%     |
| L-H-30    | 100%   | 100%      | 100%  | 100%     | 100%   | 100%      | 100% | 100%     |
| L-H-100   | 100%   | 100%      | 100%  | 100%     | 100%   | 100%      | 100% | 100%     |
| L-M-3     | 97.9%  | 96.7%     | 89.8% | 97.9%    | 100%   | 100%      | 100% | 99.8%    |
| L-M-10    | 100%   | 100%      | 100%  | 100%     | 100%   | 100%      | 100% | 100%     |
| L-M-30    | 100%   | 100%      | 100%  | 100%     | 100%   | 100%      | 100% | 100%     |
| L-M-100   | 100%   | 100%      | 100%  | 100%     | 100%   | 100%      | 100% | 100%     |
| M-H-3     | 100%   | 98.5%     | 100%  | 99.9%    | 100%   | 100%      | 100% | 100%     |
| M-H-10    | 100%   | 100%      | 100%  | 100%     | 100%   | 100%      | 100% | 100%     |
| M-H-30    | 100%   | 100%      | 100%  | 100%     | 100%   | 100%      | 100% | 100%     |
| M-H-100   | 100%   | 100%      | 100%  | 100%     | 100%   | 100%      | 100% | 100%     |
| L-M-H-3   | 100%   | 99.3%     | 100%  | 100%     | 100%   | 100%      | 100% | 100%     |
| L-M-H-10  | 100%   | 100%      | 100%  | 100%     | 100%   | 100%      | 100% | 100%     |
| L-M-H-30  | 100%   | 100%      | 100%  | 100%     | 100%   | 100%      | 100% | 100%     |
| L-M-H-100 | 100%   | 100%      | 100%  | 100%     | 100%   | 100%      | 100% | 100%     |

H: high dose, L: low dose, M: medium dose, PaSTUM: Parameter significance test using mixture models, STA: standard approach, green: Power: 80-100%, red: T1 0-79.9%

- PaSTUM predictive performance and bias in the power setting was similar to STA.
- PaSTUM outperformed STA regarding predictive performance and bias in the T1 setting (Figure 2).

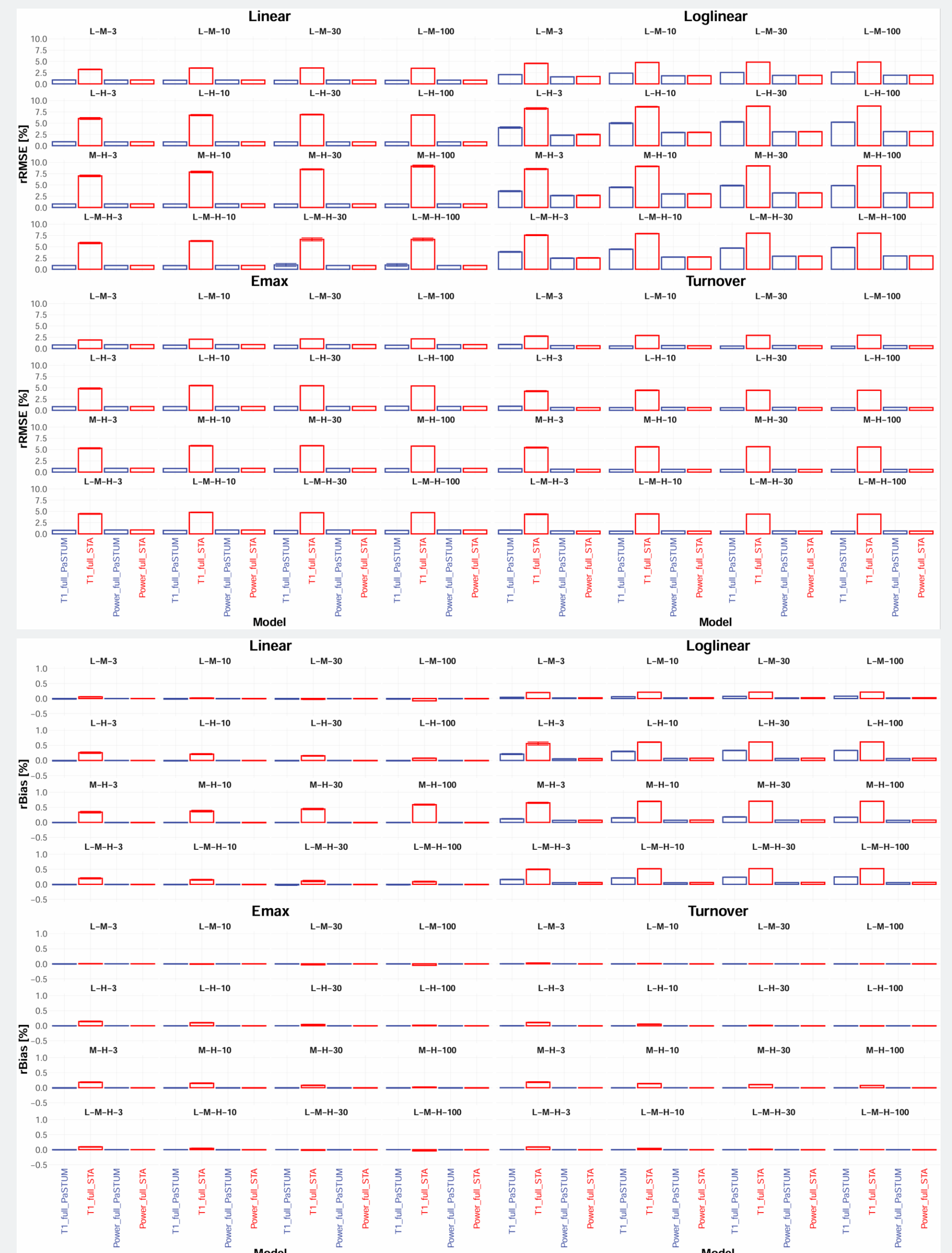


Figure 2: rRMSE (top) and rBias (bottom) mean as well as 95% CI of 1000 simulations for full PaSTUM and full STA models under all tested scenarios

- Precision and accuracy of drug effect parameter estimates for PaSTUM and STA models were similar.
- Parameter estimates in more complex model structures (direct emax and indirect) were worse for both methods (Figure 3).

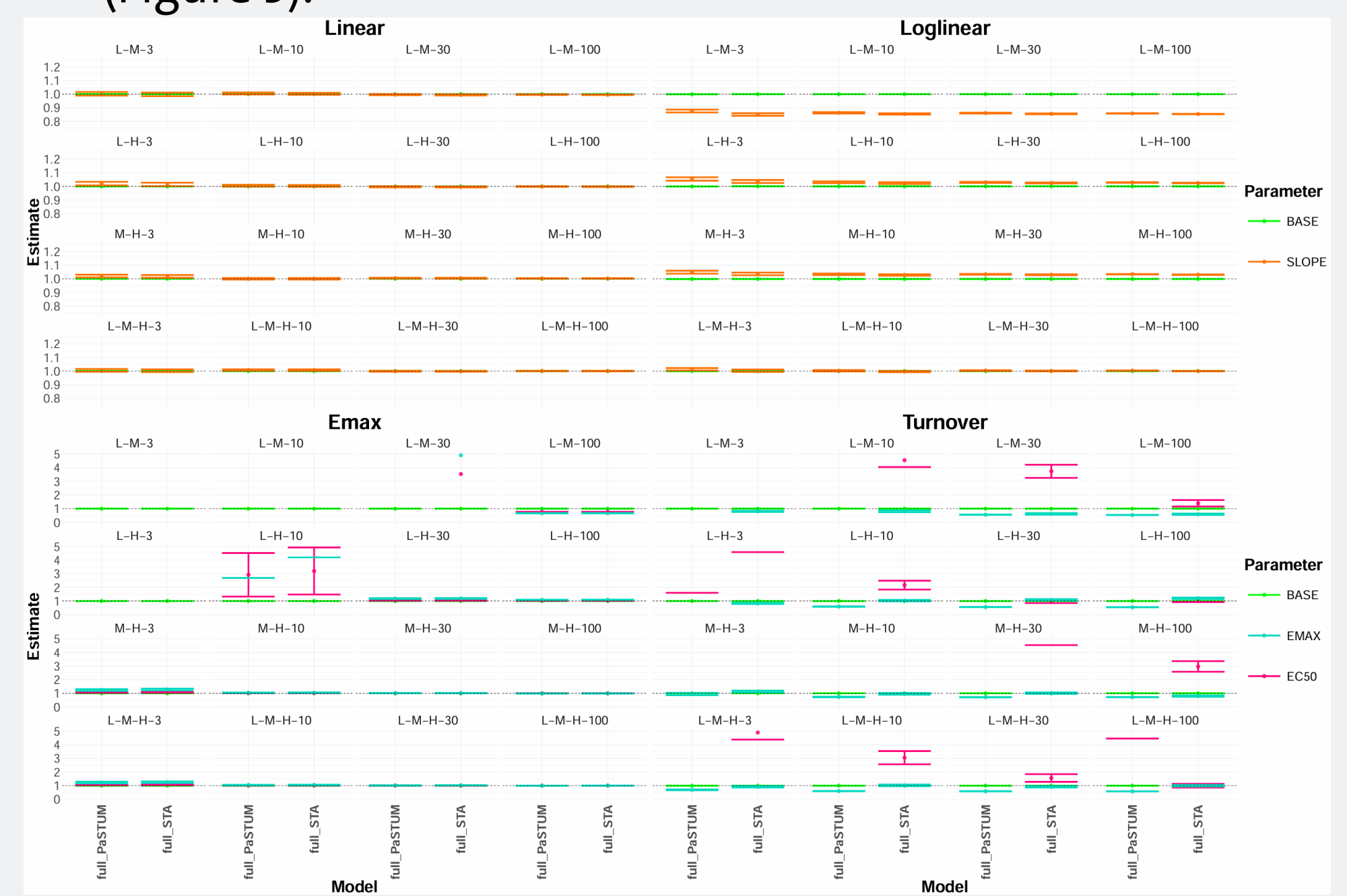


Figure 3: Precision and accuracy of drug effect parameter estimates for PaSTUM and STA full model for the power scenario

## Conclusion

PaSTUM showed well controlled T1 rate compared to STA, while still remaining high statistical power. The predictive performance as well as the parameter estimates are as good or better for PaSTUM than for STA.

### Acknowledgements:

This project has received funding from Boehringer Ingelheim Pharma GmbH & Co. KG.

### Literature

[1] Chasseloup E, Tessier A, Karlsson MO. Assessing Treatment Effects with Pharmacometric Models: A New Method that Addresses Problems with Standard Assessments. AAPS J. 2021;23:63.

