Model-based bioequivalence approach: robustness to model misspecification for sparse pharmacokinetic bioequivalence studies

Morgane Philipp[@1] Adrien Tessier[2] Mark Donnelly[3] Lucy Fang[3] Kevin Feng[3] Liang Zhao[3] Stella Grosser[4] Guoying Sun[4] Wanjie Sun[4] France Mentré[1] Julie Bertrand[1]

[@]morgane.philipp@inserm.fr, [1]Université Paris Cité, INSERM, IAME, UMR 1137, 75006 Paris, France, [2]Clinical Pharmacometrics, Quantitative Pharmacology, Servier, France,
[3]Division of Quantitative Methods and Modeling, Office of Research and Standards, Office of Generic Drugs, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring MD 20993, USA,
[4]Office of Biostatistics, Office of Translational Sciences, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring MD 20993, USA

Context

- Pharmacokinetic (PK) bioequivalence (BE) studies compare a reference (R) to a test (T) treatment in terms of exposure, i.e., the area under the curve (AUC) and the maximum plasma concentration (C_{max})
- Regulatory authorities generally recommend estimating AUC and C_{max} on data collected with a two-way, single dose, crossover study using a non-compartmental analysis (NCA) and using a two one-sided test (TOST)[1] on the treatment effect (β^T)[2,3]
- Due to challenges using NCA in PK studies with sparse samples, model-based BE (MBBE) has been proposed[4] however **the choice of the structural PK model remains a major challenge**[5]

Objectives

- **To evaluate the robustness of the MBBE approach to model misspecification in PK BE studies with sparse samples**
- **To investigate the impact of model selection (MS) and model averaging (MA)**

Methods

MBBE

- **Non linear mixed-effects model**
  - Fit of concentrations collected in R and T arms together
  - Treatment (β^T), sequence and period effects on all PK parameters μ
  - Population parameters obtained with the SAEM algorithm (Monolix 2020R1)
  - SE obtained by linearisation
  - Likelihood estimated by importance sampling
  - β^T_{AUC}, β^T_{Cmax} and their SE derived from the μ, β^T and their SE analytically or by Monte Carlo calculations (at least 800 samples)
- **MB-TOST**[4]
  - $H_{0,-\delta} : \{\beta^T_{AUC} \leq -\delta\}$ and $H_{0,\delta} : \{\beta^T_{AUC} \geq \delta\}$ and

$$W_{-\delta} = \frac{\hat{\beta}^T_{AUC} + \delta}{SE(\hat{\beta}^T_{AUC})} \geq z_{1-\alpha} \text{ and } W_{\delta} = \frac{\hat{\beta}^T_{AUC} - \delta}{SE(\hat{\beta}^T_{AUC})} \leq -z_{1-\alpha}$$

    with δ fixed to $\log(1.25)$ by the regulatory guidances[2,3], $SE(\beta^T)$ the standard error (SE) of β^T, $z_{1-\alpha}$ the quantile of a normal distribution at level $(1-\alpha)$ with α=5%
  - BE is concluded if the 90% confidence interval (CI) around the geometric mean ratio ($GMR = \exp(\beta^T)$) is within $[0.8; 1.25]$[2]
- **Extension to handle a set of M candidate models**

MS

  - Selection based on data from treatment R only according to the Akaike Information Criteria (AIC)
  - Fit of the selected model based on data from both treatment arms data and TOST

MA

  - The M models are fit based on data from both treatment arms, then weights are calculated[6]:

$$w_m = \frac{exp(-\Delta AIC_m/2)}{\sum_{m'=1}^{M} exp(-\Delta AIC_{m'}/2)}$$

    with AIC_m the AIC of model m and $\Delta AIC_m = AIC_m - min(AIC_{m=1,\dots,M})$
  - Models with weight under 0.5% are removed and weights are updated
  - Calculation of weighted $\hat{\beta}^T$ and SE($\hat{\beta}^T$)[6,7] and TOST

Real case study

Data

- Phase I, single-dose, two-way crossover BE study from Servier
- $N = 36$ healthy volunteers (i.e. 18 per arms) and $n = 22$ PK samples per period

Method

- Set of $M = 10$ candidate models with varying number of compartments for the distribution (1 or 2-COMPT) and/or absorption types with transit compartments (TRANSIT), delayed (LAG) and zero- or first-order (0/1-order)

Results

- MS: TRANSIT_2-COMPT model selected
- MA: TRANSIT_2-COMPT model with $w = 1$
- BE concluded with MB-TOST on AUC and C_{max}
  - $GMR_{AUC} = 1.00$ and $CI_{90_{AUC}} = [0.97; 1.04] \in [0.8, 1.25]$
  - $GMR_{C_{max}} = 1.04$ and $CI_{90_{C_{max}}} = [0.99; 1.10] \in [0.8, 1.25]$

[1] Schuirmann. *J. Pharmacokinet. Biopharm.* 1987; [2] FDA *Bioequivalence studies with pharmacokinetic endpoints for drugs submitted under an ANDA.* 2021; [3] EMA *Guideline on the Investigation of Bioequivalence.* 2010; [4] Dubois et al. *Stat. Med.* 2011; [5] Guhl et al. *PAGE 29 abstract 9719.* 2021; [6] Buatois et al. *AAPS J.* 2018; [7] Turek et al. *Comput. Stat. Data Anal.* 2012;

Simulation study

Settings

- Single-dose, two-way crossover study with $N = 20$ patients per sequence (within subject variability = 30%) and $n = 6$ PK samples per period {t=0.3, 3, 6, 12, 72, 144}
- $S = 200$ datasets simulated for each scenario

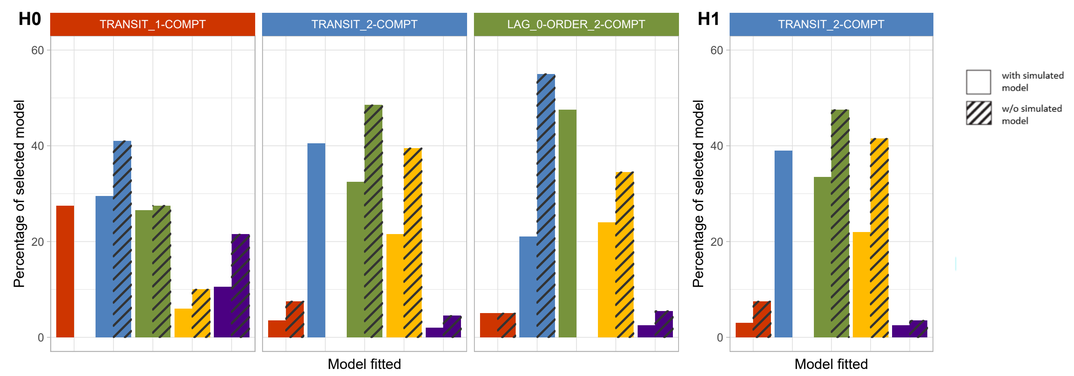
		Simulated model		
		TRANSIT_1-COMPT	TRANSIT_2-COMPT	LAG_0-ORDER_2-COMPT
Data fitted with	TRANSIT_1-COMPT TRANSIT_2-COMPT LAG_0-ORDER_2-COMPT LAG_1-ORDER_2-COMPT LAG_0-ORDER_1-COMPT MS with and w/o simulated model MA with and w/o simulated model	Under H_0 with β^T=log(1.25)	Under H_0 with β^T=log(1.25) and under H_1 β^T=log(0.95)	Under H_0 with β^T=log(1.25)

- Evaluation: proportion of model selected, type I error and power

Results

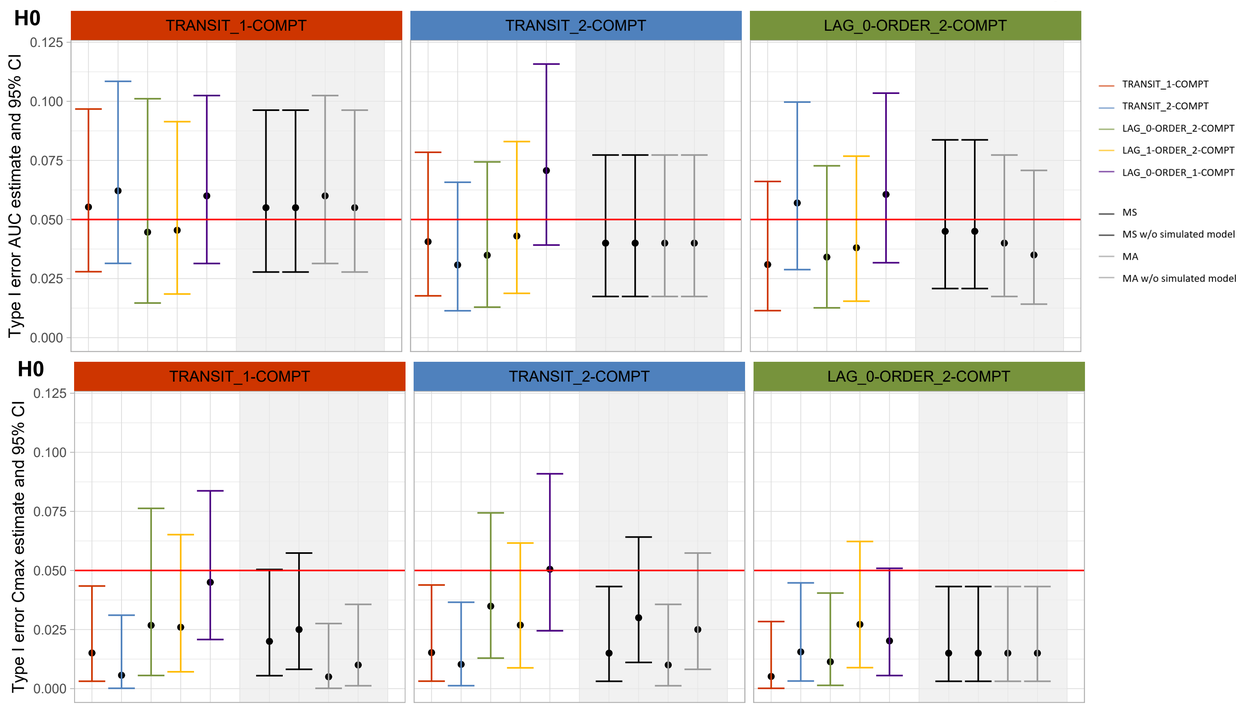
- **MS**

Candidate model selected (%) for each scenario when the simulated model was included or excluded from the candidate pool



  - with simulated model: right model selected in 28%, 41% and 48% cases under H_0 and 39% under H_1 → model not highly identifiable under simulation settings
  - w/o simulated model: selection driven by number of COMPT
- **MA** performs similarly to MS, with one candidate model having $w = 1$ in 68% and 73% of the cases under H_0 and H_1
- **Type I error**

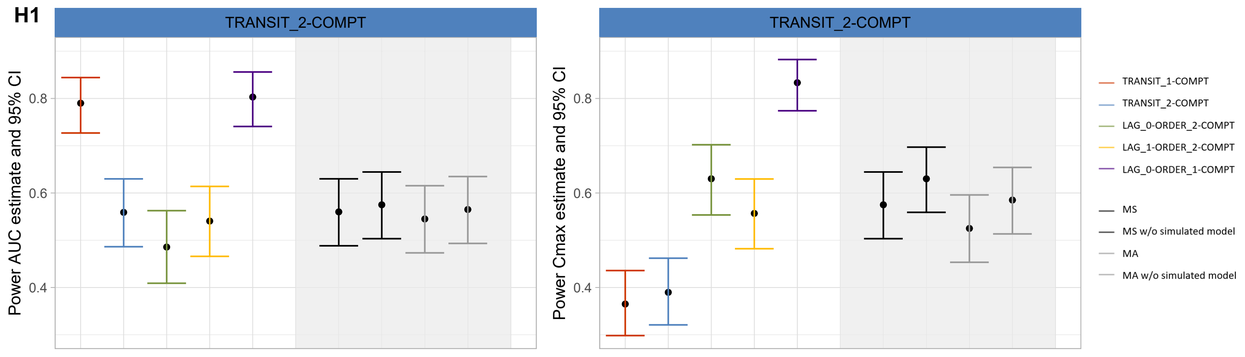
Estimate and 95% CI for AUC and C_{max} using M candidate models, MS and MA when the simulated model was included or excluded from the candidate pool



→ In general, the 95%CI around the type I error estimate includes the 5% target or is just below for C_{max}

- **Power**

Estimate and 95% CI for AUC and C_{max} using M candidate models, MS and MA when the simulated model was included or excluded from the candidate pool



→ MS and MA showed similar power for AUC and higher power for C_{max} when compared to the simulated model

Conclusion

The MBBE approach appears to be robust to model misspecification in our simulation study. MS and MA led to type I errors around or below 5% and ensured a reasonable power. No added value was observed with MA compared to MS. Further work is needed to assess the robustness when R and T have different PK models.