Simulation-based evaluation of combined Likelihood Ratio Test (cLRT) and Multiple Comparison Procedure (MCP) approaches to identify a dose-response relationship in the oncology setting under different study design variables

Yomna M Nassar (1,2,3*), Jonas Schick (4), Charlotte Kloft (1), Sebastian Bossert (4), Alejandro Pérez Pitarch (3*)

(1) Department of Clinical Pharmacy and Biochemistry, Institute of Pharmacy, Freie Universität Berlin, Germany; (2) Graduate Research Training Program PharMetrX, Berlin/Potsdam, Germany; (3) Translational Medicine & Clinical Pharmacology, Boehringer Ingelheim Pharma GmbH & Co. KG, Ingelheim am Rhein, Germany; (4) Global Biostatistics & Data Sciences, Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach, Germany.

**Affiliation at the time of research*

Introduction

- In oncology, the FDA requests for more rational derivation of optimal doses [1] through:
  - Better leveraging of data from dose-finding trials
  - Identification of a dose-response (D-R) signal/shape [2]

  For →broader understanding of the impact of different doses on efficacy and toxicity
  →identification of the optimal dose for further investigation in confirmatory trials

- **Multiple Comparison Procedure-Modelling (MCP-Mod)**
  - A model-based approach for the design and analysis of phase II dose-finding trials [3]
  - Gained the FDA's qualification [4-7]
  - Later, **combined Likelihood Ratio Test-Modelling (cLRT-Mod)** was proposed as an extension to MCP-Mod [8]

cLRT-Mod :
- Based on the principles of NLME modelling
- Leverages longitudinal phase II trial data.
- Higher power than MCP-Mod in detecting a D-R signal
- Has shown a controlled type I error with simulated data [8] but type I error was inflated with the use of real data [9].

Objectives

- To evaluate the performance in identifying a D-R signal/shape of cLRT when using sum of longest diameter (SLD) continuous data versus MCP when using the best change in SLD from baseline data as non-time-variant endpoint in the oncology setting through simulation-based explorations

a) Challenge the performance under different settings:
- Strong and weak drug effects
- Different D-R shapes

b) Evaluate influence of study design variables:
- Different number of patients
- Different number of dose levels

Methods

Simulated data

Data-generating true model [10]
- Longitudinal continuous tumour response model
- Described:
  - The time course of SLD as a function of drug dose, given a pre-specified D-R shape e.g. linear, E_{max} ,…(Fig. 2)
  - Disease progression, probability of dropouts, deaths

Study design (dose-finding phase II study)

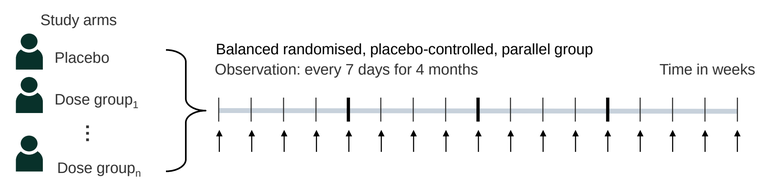

Figure 1: General design of the simulated studies to explore performance of cLRT and MCP

Simulation scenarios

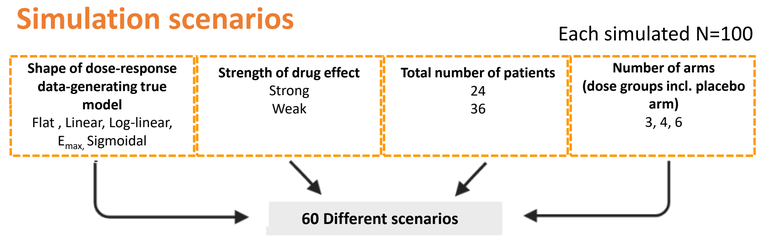

Figure 2: Variables defining the different simulation scenarios

Candidate models

- Tested against the simulated data
- Included same shapes of data-generating true model
- Flat shape (no D-R) was considered the reference

MCP analysis [3]

a) Pre-selected set of candidate models
↓
b) Multiple contrast tests
↓
c) Evidence of a D-R

DoseFinding R package.

cLRT analysis [8]

a) Fit each candidate model to each simulated dataset
↓
b) Calculate LRT between best-fitting candidate model (lowest AIC) and the no D-R (flat) model
→
c) Compare calculated LRT against the critical value approximated from data generated under the null hypothesis (H_0) (M=100 simulations), using the same procedure (a–b) but with placebo arm data
↓
d) Evidence of a D-R

NONMEM® and PsN were used for simulations and (re-)estimations

Evaluation

a) Power of detecting a D-R shape
b) Type I error
c) Selection of the best-fitting candidate model & identification of the true D-R shape of the data-generating true model

Key findings, Limitations & Conclusion

- Power was higher for cLRT vs MCP (Fig. 3) but at the expense of an inflated type I error (Fig. 4)
- Inflated cLRT type I error was in line with [9] but not [8]. This observation could be due to the:
  1. small sample size compared with [8, 9]
  2. placebo model misspecifications [9]
  3. selection bias [9]

  } impact H_0 distribution used to derive the critical value and calculate type I error
- At fixed sample size, cLRT power was not sensitive to different study designs (Fig. 3)

- In cLRT, a placebo group is needed to derive the critical value. Nevertheless, oncology trials generally do not provide these data.
- To overcome this limitation, we suggest to explore ways of incorporating external data (e.g. RWD), making assumptions when considered adequate (e.g. using data from very low dose levels gathered at earlier development stages and assuming irrelevant exposures), among other possibilities.

- Before cLRT can be applied to analyse dose-finding studies:
  1. Investigation of the robustness of cLRT to model complexities and study variables, e.g. sample size, is needed to justify its computational demands for simulations.
  2. Identify settings e.g. scenarios /data type associated with a controlled type I error

Results

a) Power

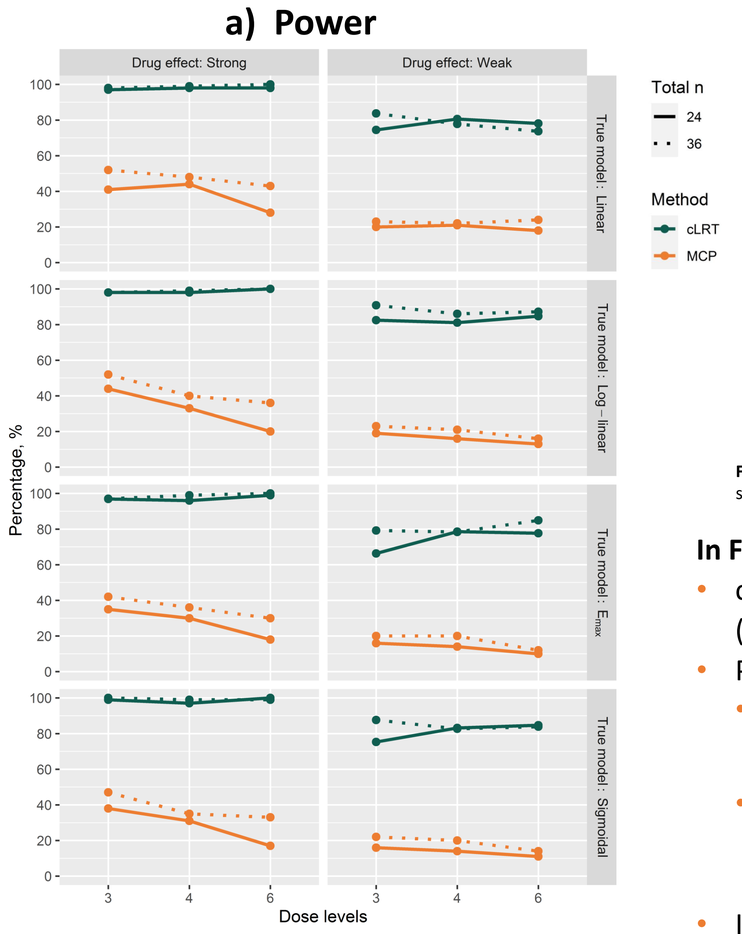

Figure 3: Power of cLRT and MCP across different study design variables stratified by strength of drug effect and shape of the data-generating D-R true model.

b) Type I error

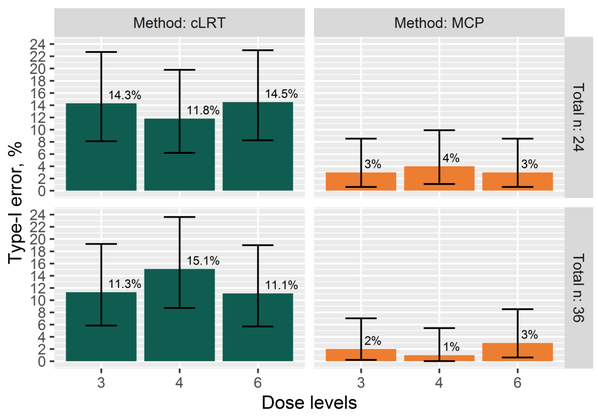

Figure 4: Type I error of cLRT and MCP per each group of scenarios sharing the same study variables and differing only in the shape of the true model and drug strength

In Fig. 3:
- cLRT power (89.8%) was higher than MCP power (27.0%)
- Power was also higher in:
  - Scenarios with a strong drug effect
    - $cLRT_{strong}$: 96–100%, MCP_{strong}: 17–52%
    - $cLRT_{weak}$: 73.7–90.8%, MCP_{weak}: 10–24%
  - Scenarios with a greater number of patients

	24	36
cLRT	>66.3%	>73.7%
MCP	>10%	>12%

- Increased power with fewer dose levels and more patients per dose level→observed with MCP

c) Selection of the best-fitting candidate model and identification of the true shape of the D-R data-generating true model

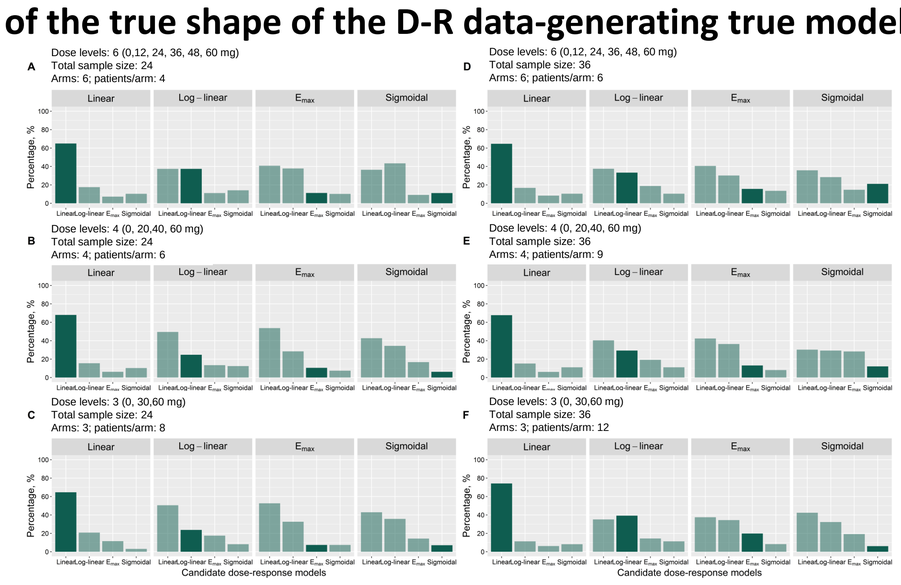

Figure 5: Proportion of the selected best-fitting candidate D-R models under the **strong** drug effect (**weak not shown**) highlighting the proportion of identified true shape of the D-R model for cLRT, per each group of scenarios, split by the underlying shape of the true data-generating dose-response model (facets). Each panel (**A-F**) represents results from the group of scenarios sharing the same set of study design variables, irrespective of the true shape of the data-generating dose-response model. **Bold columns:** proportion of the identified **true** underlying shape of the dose-response model.

In Fig. 5:
- Tendency towards selection of the *simplest* D-R shapes (linear, loglinear)
- Identification of true shape of D-R model was independent of study design
- More complex models (Emax, sigmoidal) had a weaker chance of being successfully detected compared with simpler models.

References

1. Food and Drug Administration. Project Optimus: Reforming the dose optimization and dose selection paradigm in oncology. https://www.fda.gov/about-fda/oncology-center-excellence/project-optimus (Accessed on: 08.12.2023).
2. European Medicines Agency. ICH Topic E 4: Dose Response Information to Support Drug Registration. https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-4-dose-response-information-support-drug-registration-step-5_en.pdf . (Accessed on: 08.12.2023).
3. F. Bretz *et al.* Biometrics 61: 738–748 (2005).
4. Food and Drug Administration. Drug Development Tools: Fit-for-Purpose Initiative. https://www.fda.gov/drugs/development-approval-process-drugs/drug-development-tools-fit-purpose-initiative . (Accessed on: 12.12.2023)
5. S. S. Shord *et al.* CPT: Pharmacometrics & Systems Pharmacology 12: 1573 (2023).
6. Food and Drug Administration. Statistical Review and Evaluation: Qualification of Statistical Approach. https://www.fda.gov/media/77169/download . (Accessed on: 08.12.2023).
7. Food and Drug Administration. FDA Qualification of MCP-Mod Method. https://www.fda.gov/media/99313/download . (Accessed on: 07.12.2023).
8. S. Buatois *et al.* Statistics in Medicine 40: 2435–2451 (2021).
9. E. Chasseloup et al. Pharmaceutics 15: 460 (2023).
10. S. M. Krishnan *et al.* 30th PAGE meeting, abstract 10071. (2022). www.page-meeting.org/?abstract=10071

Abbreviations

AIC:	Akaike information criterion
cLRT-Mod:	combined Likelihood Ratio Test-Modelling
D-R:	Dose-response
FDA:	Food and Drug Admisnitration
MCP-Mod:	Multiple Comparison Procedure- Modelling
NLME:	Nonlinear mixed-effects
RWD:	Real-world data
SLD:	Sum of longest diameter

Presented at the 32nd Population Approach Group Europe (PAGE) Meeting, Rome, Italy, 25-28 June 2024

This project was conducted during an internship sponsored by Boehringer Ingelheim.