

Retrospective design analysis to evaluate power for covariate selection from observational data

Ron J. Keizer¹, Paola Mian², Michael McCarthy¹, Jordan Brooks¹

¹InsightRX, San Francisco, CA, USA. ² University Medical Center Groningen, the Netherlands

Background

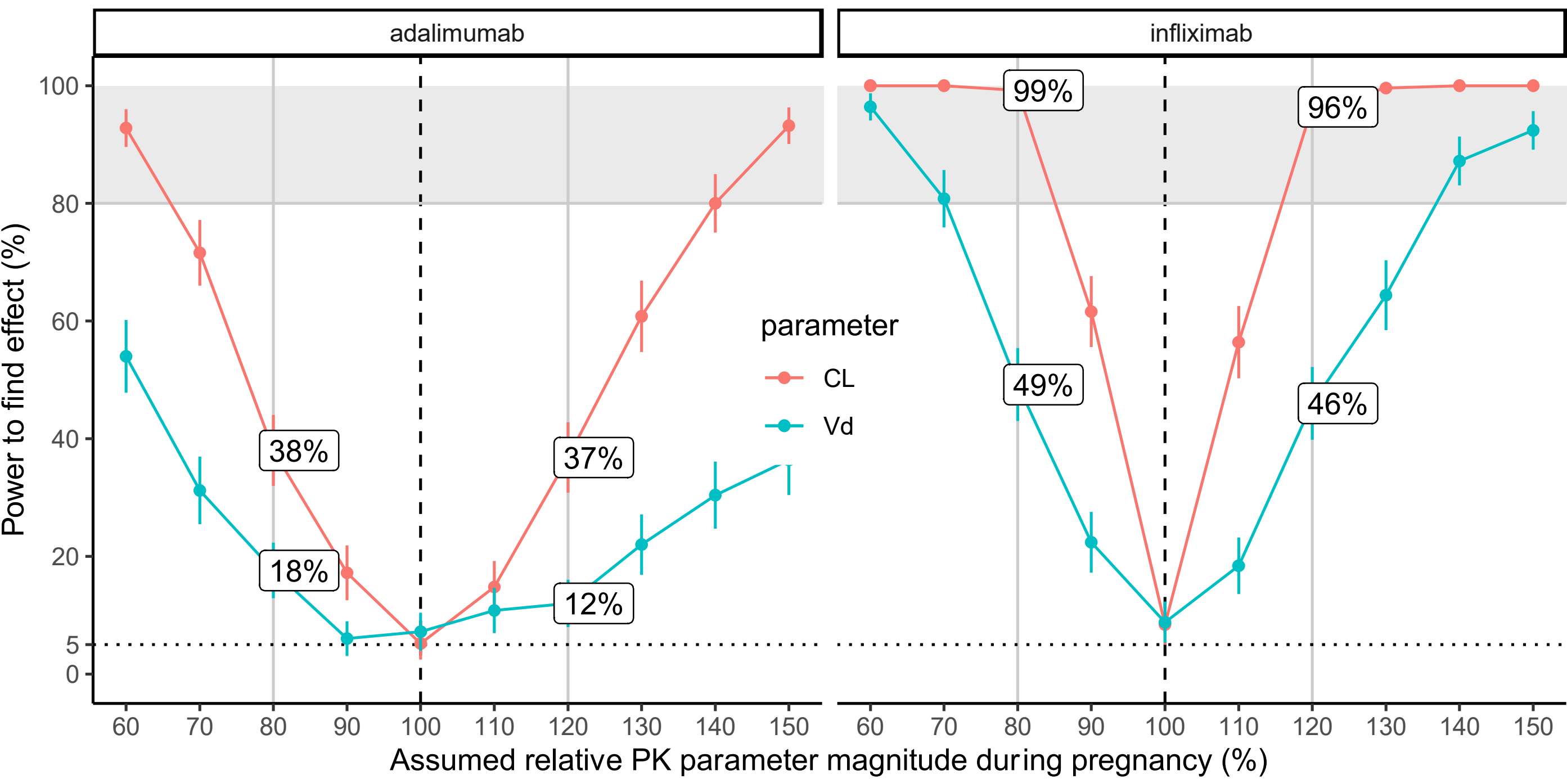
Population PK models are often developed from routinely collected, observational data, such as from **therapeutic drug monitoring**. However, in this setting it is often unclear for the covariate effects:

- Whether failure to reject the null hypothesis was due to an effect being truly absent, or that the dataset and analysis were **underpowered**.
- When the null hypothesis is rejected, it is often unclear whether the effect is a **spurious finding** or not.

Design analyses are used extensively within pharmacometrics in prospective trial design (e.g. PsN *sse*, *mcmp*).¹ However, design calculations can also be performed **retrospectively**, to contextualize the results of a covariate analysis. They have been in the past but under various nomenclature (see box "Nomenclature").

Importantly, "design analysis" as introduced by Gelman² differs from "posthoc power" analyses in which the value of the analysis' test statistic is in some way back-calculated to a power. This latter approach is a fallacy, as demonstrated by many (e.g. ³), as power does not depend on observed results.

Results – analysis 1 (pregnancy on mAb CL)



Plots of retrospective power vs assumed effects were useful to understand the ability of the dataset to identify covariate effects. Especially the calculation of the power at predefined clinical relevance cutoffs (e.g. P_{20} , at 20% effect) was useful to evaluate the study design's power.

The RDA for **analysis 1** revealed that the dataset for adalimumab was not sufficiently powered to identify clinically relevant effects for *CL* and *Vd* at $p<0.05$ (since P_{20} was lower than 80%) but that the dataset for infliximab was able to identify effects on *CL*.

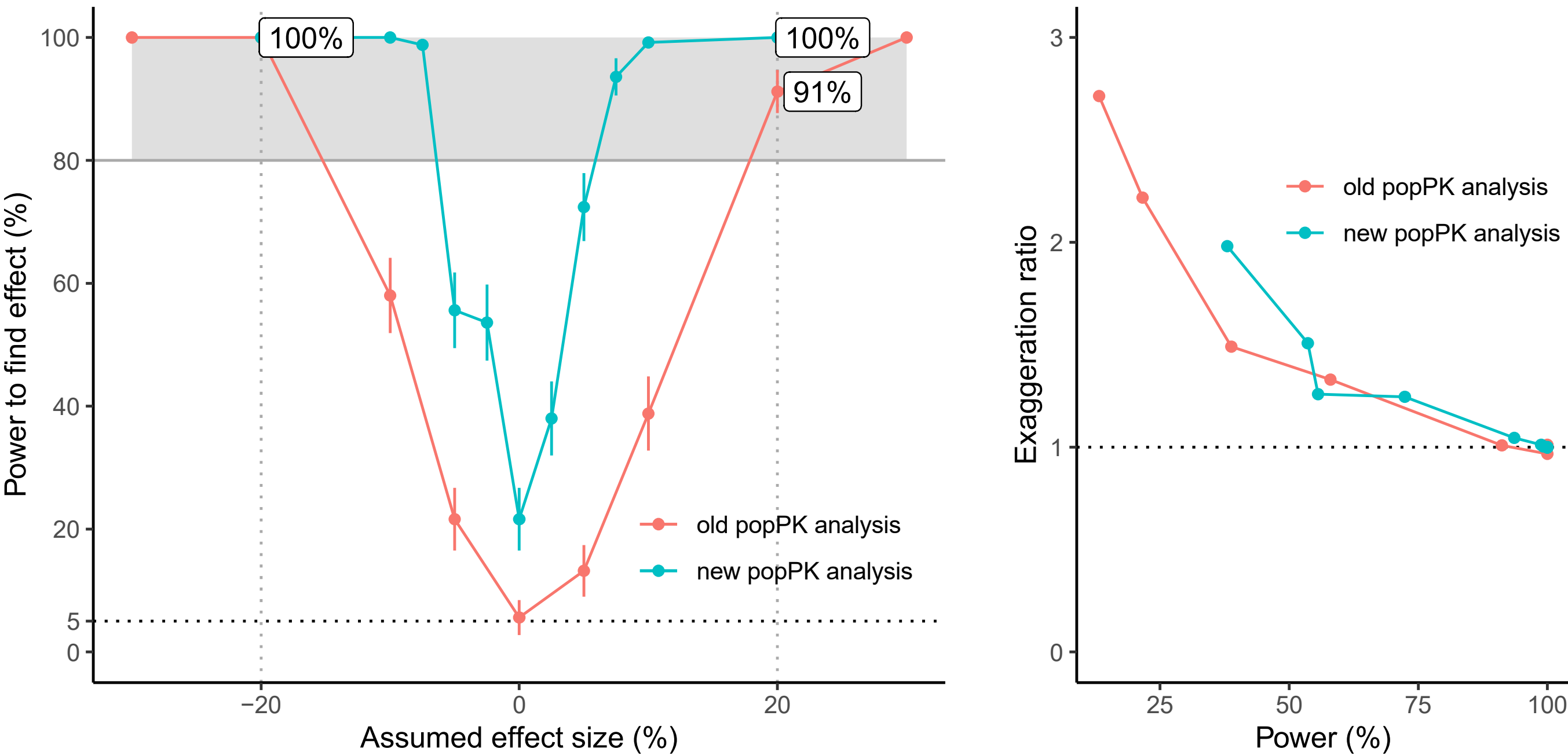
Objectives

- Explore retrospective design analysis (RDA) as an approach for observational pharmacokinetic analyses to provide more detailed information regarding covariate in-/exclusion.
- Propose updated nomenclature and useful plots and statistics

Methods – Datasets, prior analyses, tools

- Analysis 1:** a popPK analysis studying effect of **pregnancy on PK parameters** for two monoclonal antibodies in IBD. No significant effect was identified for either mAb, but it was unclear whether the datasets were sufficiently informative to find a clinically relevant effect if present.
- Analysis 2:** a popPK analysis for busulfan[4]. An earlier popPK analysis identified **co-medication** as a significant covariate on CL. However, a more recent analysis also found significance, but the sign of the covariate had flipped. We implemented an rDA to investigate the parameter-covariate relationship further.
- Both RDAs were implemented using the **sse** tool in PsN[5].

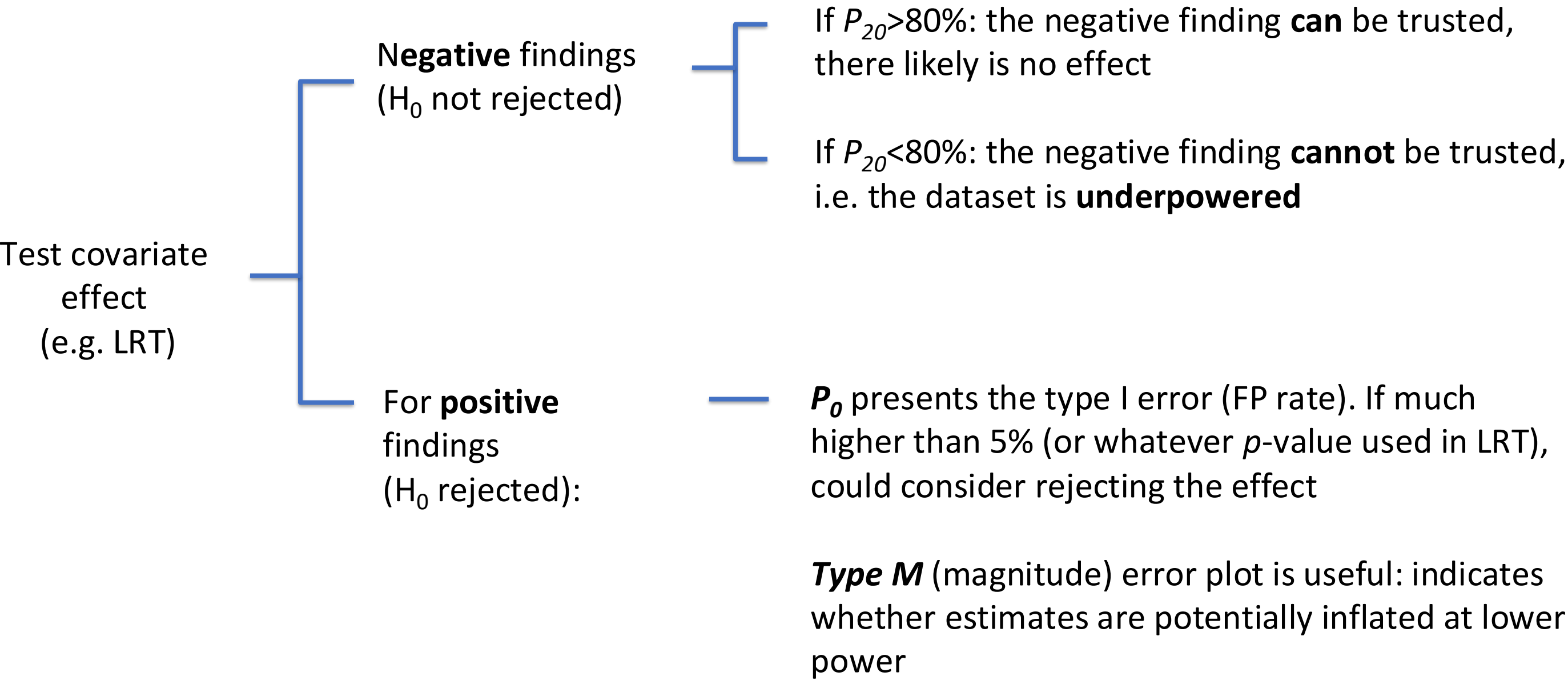
Results – analysis 2 (co-medication on busulfan CL)



For **analysis 2**, the RDA revealed that the previous analysis was in fact sufficiently powered and the effect only somewhat inflated (low **type M error**, or **exaggeration rate**), and that covariate inclusion into the final model was justified at the time.

The new population PK analysis on a larger dataset showed even higher power to identify effects, but did suffer from an **inflation of type 1 error** (false positive rate).

Interpreting Retrospective Design Analyses



References

1. Vong C et al.

2. Gelman A & Carlin J.

3. Hoenig JM & Heisey DM.

4. Shukla et al.

5. Perl-speaks-NONMEM.

6. M Tareq Abdelwahab (2020)

7. JH Hughes (2019)

8. C Emoto (2015)

9. N de Rouw (2021)

10. R Press (2009)

Nomenclature in prior literature

Reference	Nomenclature	Methodology
6	"A posteriori parametric power and design evaluation"	SSE
7	"Retrospective power analysis"	SSE
8	"Retrospective power analysis"	SSE
9	"Posthoc power analysis"	SSE
10	"Retrospective power analysis"	Bootstrap*

** Incorrect methodology, bootstrap should not be used to assess retrospective power*

Conclusions

- We propose to use the term **"retrospective design analysis"** (RDA) for analyses as presented here, to unify nomenclature in literature.
- RDA offers useful information in observational analyses where prospective design evaluation is not possible.
- While its use may benefit all observational analyses, RDA is especially useful when data is sparse and estimated effects are borderline significant.
- Useful for **negative findings**: power plots and computation of P_{20}
- Useful for **positive findings**: type M error plots and computation of P_0