Group sequential designs with negative binomial data

Ekkehard Glimm[1] Tobias Mütze [2,3]

[1]Statistical Methodology, Novartis, Basel, Switzerland

[2]Department of Medical Statistics, University Medical Center Göttingen

[3]DZHK (German Centre for Cardiovascular Research), partner site Göttingen, Göttingen, Germany

Thanks to Tim Friede and Heinz Schmidli

PSI/BBS One-day meeting Basel

September 13, 2016

1. Motivating examples
2. Fixed design
3. Group sequential designs
4. Assessing operating characteristics
5. Discussion and outlook

Example 1: Clinical trials in heart failure

- Heart failure (HF) with preserved ejection fraction (HFpEF)
- Primary endpoint: Number of heart failure hospitalizations (HFH)
- HFH can be modeled with negative binomial distribution (Rogers et al., 2014)
- Example: the CHARM-Preserved trial (Yusuf et al., 2003)

Table: Heart failure hospitalizations in CHARM-preserved

	Placebo	**Candesartan**
Number of patients	1509	1514
Total follow-up years	4374.03	4424.62
Patients with ≥ 1 admission	278	230
Total admissions	547	392

- Rate ratio for recurrent heart failure hospitalizations according to negative binomial model $\theta = 0.71$

Example 2: Clinical trials in relapsing-remitting multiple sclerosis

- Primary endpoint: number of combined unique active lesions (CULAs)
- CULAs are modeled using the negative binomial distribution
- Example: Phase II study of Siponimod (Selmaj et al., 2013)
  - Placebo and five doses of Siponimod
  - Equal follow-up times (in general either 3 or 6 months)

Table: Monthly number of lesions (at 3 months)

	Placebo	**Siponimod 0.25 mg**	**Siponimod 0.5 mg**
Number of patients	61	51	43
Monthly CULAS	1.39	0.78	0.54

Statistical model

- Number of counts for patient $i = 1, \ldots, n_j$ receiving treatment $j = 1, 2$

$$Y_{ij} | \lambda_{ij} \sim \text{Pois}(t_{ij} \lambda_{ij})$$

- Follow-up per patient: t_{ij}
- Gamma-mixture for the rates

$$\lambda_{ij} \sim \Gamma\left(\frac{1}{\phi}, \frac{1}{\phi \mu_j}\right)$$

- Marginal distribution of counts

$$Y_{ij} \sim \text{NB}\left(t_{ij} \mu_j, \phi\right)$$

- Expected value and variance

$$\mathbb{E}\left[Y_{ij}\right] = t_{ij} \mu_j$$
$$\text{Var}\left[Y_{ij}\right] = t_{ij} \mu_j (1 + \phi t_{ij} \mu_j)$$

Hypothesis testing I

- Statistical hypothesis

$$H_0 : \frac{\mu_1}{\mu_2} \geq 1 \quad \text{vs.} \quad H_1 : \frac{\mu_1}{\mu_2} < 1.$$

- Hypothesis is tested using a Wald-type test of the maximum-likelihood estimators $\hat{\beta}_j$ of the log-rates $\beta_j = \log(\mu_j)$

- Wald-type test statistic

$$T = \frac{\hat{\beta}_1 - \hat{\beta}_2}{\sqrt{\frac{1}{\hat{I}_{\beta_1}} + \frac{1}{\hat{I}_{\beta_2}}}} \underset{\text{asymp.}}{\overset{H_0}{\sim}} \mathcal{N}(0, 1)$$

Hypothesis testing II

- Fisher information of log-rates β_j

$$I_{\beta_j} = \sum_{i=1}^{n_j} \frac{t_{ij} \exp(\beta_j)}{1 + \phi t_{ij} \exp(\beta_j)} = \sum_{i=1}^{n_j} \frac{t_{ij} \mu_j}{1 + \phi t_{ij} \mu_j}.$$

(Reminder: $I_{\beta_j}^{-1}$ is the asymptotic variance of the MLE $\hat{\beta}_j$.)

- Information level $\mathcal{I}_{fix}$ describes "knowledge" about unknown treatment effect

$$\mathcal{I}_{fix} = \frac{1}{\frac{1}{I_{\beta_1}} + \frac{1}{I_{\beta_2}}} = \frac{I_{\beta_1} I_{\beta_2}}{I_{\beta_1} + I_{\beta_2}}$$

- Sample size planning by solving equation

$$\mathcal{I}_{fix} \stackrel{!}{=} \frac{(q_{1-\beta} - q_{1-\alpha})^2}{(\beta_1 - \beta_2)^2}$$

Group sequential designs: Overview

- Test the hypothesis H_0 at several interim analyses and stop the trial if H_0 can be rejected (stop for efficacy)
- The interim analyses are performed with the Wald-type test using all data available up to that point in time
- Counts of patient i in treatment j at analysis k: $Y_{ijk} \sim NB(t_{ijk}\mu_j, \phi)$
- t_{ijk} is the follow-up time until analysis k
- The final analysis is performed when a prespecified information level $\mathcal{I}_{max}$ is attained (maximum information trial)

Type I error

- Critical values of the individual tests c_k must be chosen such that global type I error α, i.e.

$$\alpha \leq \mathbb{P}_{H_0}\left(T_k < c_k \quad \text{for at least one } k = 1, \ldots, K\right).$$

- Allocate global type I error $\alpha = \sum_{k=1}^{K} \pi_k$
- Type I error rate π_k for analysis k

$$\mathbb{P}_{H_0}\left(T_1 \geq c_1, \ldots, T_{k-1} \geq c_{k-1}, T_k < c_k\right) = \pi_k$$

- Choose π_k through error spending function $f : [0, \infty) \to [0, \alpha]$ with $f(0) = 0$ and $f(t) = \alpha, t \geq 1$:

$$\begin{aligned}
\pi_1 &= f\left(\mathcal{I}_1/\mathcal{I}_{max}\right), \\
\pi_k &= f\left(\mathcal{I}_k/\mathcal{I}_{max}\right) - f\left(\mathcal{I}_{k-1}/\mathcal{I}_{max}\right) \quad k = 2, 3, \ldots.
\end{aligned}$$

Critical values

- First critical value is the normal quantile $c_1 = q_{\pi_1}$
- Joint distribution $(T_1, \ldots, T_k)$ required to calculate critical value c_k
- Asymptotic normality of joint distribution has canonical form [Scharfstein et al., 1997]

$$(T_1, \ldots, T_k)' \to \mathcal{N}(0, \Sigma_k)$$

with

$$(\Sigma_k)_{(k_1,k_2)} = (\Sigma_k)_{(k_2,k_1)} = \sqrt{\frac{\mathcal{I}_{k_1}}{\mathcal{I}_{k_2}}}, \qquad 1 \leq k_1 \leq k_2 \leq k.$$

Practical considerations

- Information level depends on rates μ_j, shape parameter ϕ, follow-up times t_{ijk}, and sample size n_j
- At analysis k, $\mathcal{I}_k$ not known and is estimated by plugging in the rate and shape maximum-likelihood estimators
- Critical value c_k is not determined prior to the trial but at the time of analysis k
- $\hat{\mathcal{I}}_k$ is the estimated information level of stage k obtained with the data available at interim k

Practical considerations continued

- In practice the following estimators are considered

$$\hat{\pi}_1 = f\left(\hat{\mathcal{I}}_1/\mathcal{I}_{max}\right)$$

$$\hat{\pi}_k = f\left(\hat{\mathcal{I}}_k/\mathcal{I}_{max}\right) - f\left(\hat{\mathcal{I}}_{k-1}/\mathcal{I}_{max}\right) \quad k = 2, 3, \ldots$$

$$\left(\hat{\Sigma}_k\right)_{(k_1,k_2)} = \sqrt{\frac{\hat{\mathcal{I}}_{k_1}}{\hat{\mathcal{I}}_{k_2}}}$$

- Estimated information might decrease if sample sizes or time between analyses is small, i.e. $\hat{\mathcal{I}}_k < \hat{\mathcal{I}}_{k-1}$
  - then analysis is skipped $\Leftrightarrow$ critical value $c_k = \infty$
- "Locally" allocated type I error preserves the global type I error

$$\sum_{i=1}^{K} \hat{\pi}_k = \alpha$$

Planning of group sequential trials

- Power for given set of critical values $c_1, \ldots, c_K$

$$\text{Power} = 1 - \mathbb{P}_{H_1}\left(T_1 \geq c_1, \ldots, T_K \geq c_K\right)$$

- For rate ratio θ^* in alternative, joint distribution $(T_1, \ldots, T_K)$ approximately normal with mean vector $\log(\theta^*)(\sqrt{\mathcal{I}_1}, \ldots, \sqrt{\mathcal{I}_K})'$
- For planning purposes, we write

$$\mathcal{I}_k = w_k \mathcal{I}_{max}, \quad k = 1, \ldots, K, \quad w_k \in (0, 1]$$

- Calculate maximum information $\mathcal{I}_{max}$ required to obtain power of $1 - \beta$ by solving

$$1 - \mathbb{P}_{\theta^*}\left(T_1 \geq c_1, \ldots, T_K \geq c_K\right) = \beta$$

- Sample size, study duration, etc must be selected such that the maximum information is obtained

Simulation study - preface

- In the simulation, interim analysis time points are determined by theoretical information levels $\mathcal{I}_k$. The actual estimated information levels $\hat{\mathcal{I}}_k$ differ from this.
- Use of spending functions which imitate critical values of Pocock's test and O'Brien & Fleming's test
- Recruitment times uniform in fixed accrual period

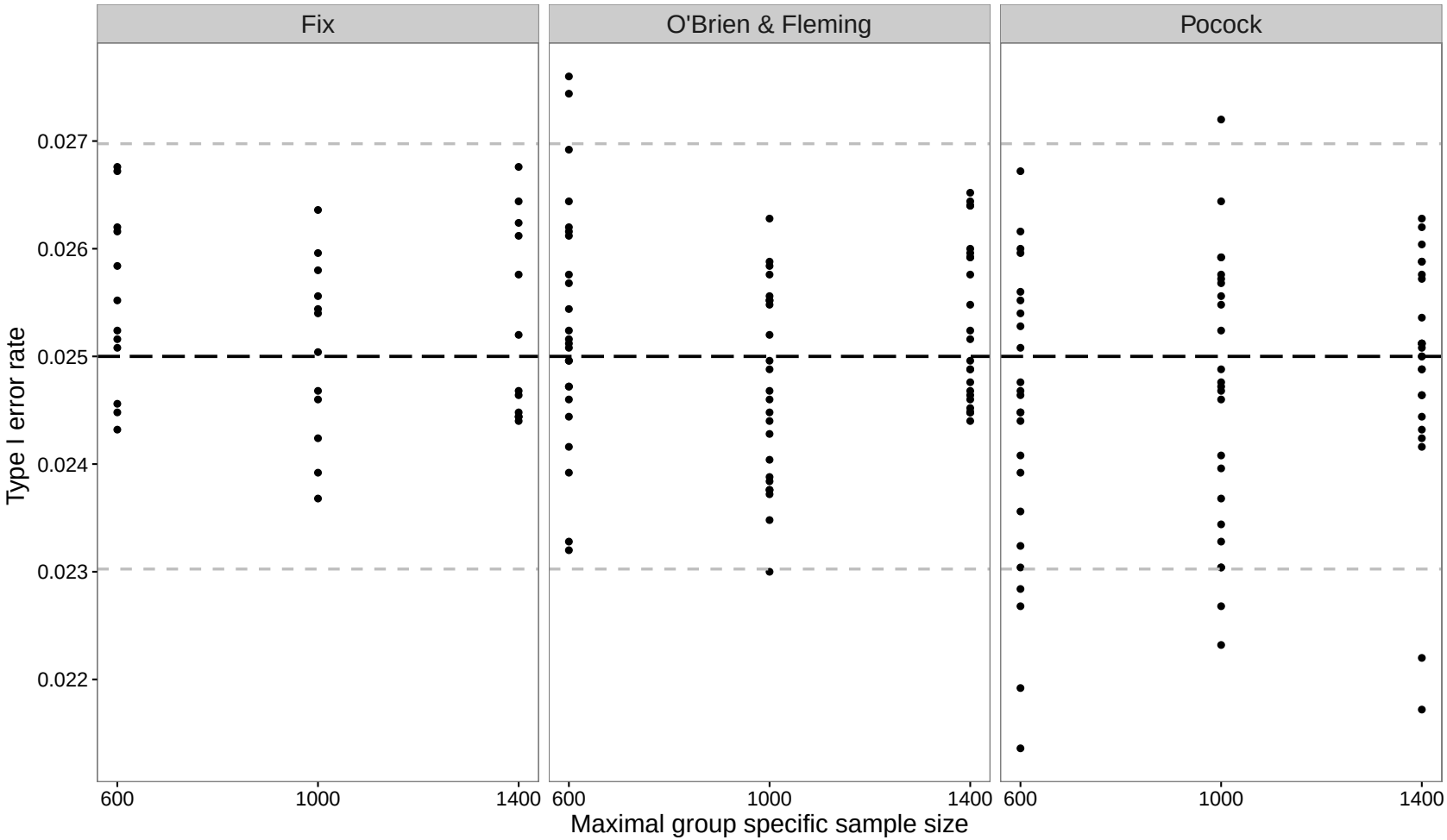
Simulation scenarios - type I error

- Simulation scenarios motivated by the number of hospitalizations from Example 1

Parameter	**Values**
Type I error rate α	0.025
Annual rates $\mu_1 = \mu_2$	0.08, 0.1, 0.12
Shape parameter ϕ	2, 3, 4, 5
Group sample size $n_1 = n_2$	600, 1000, 1400
Stages K	2, 5
Study duration	3.5 (years)
Recruitment period	1.25 (years)

- 25 000 Monte Carlo replications per scenario

Results - type I error



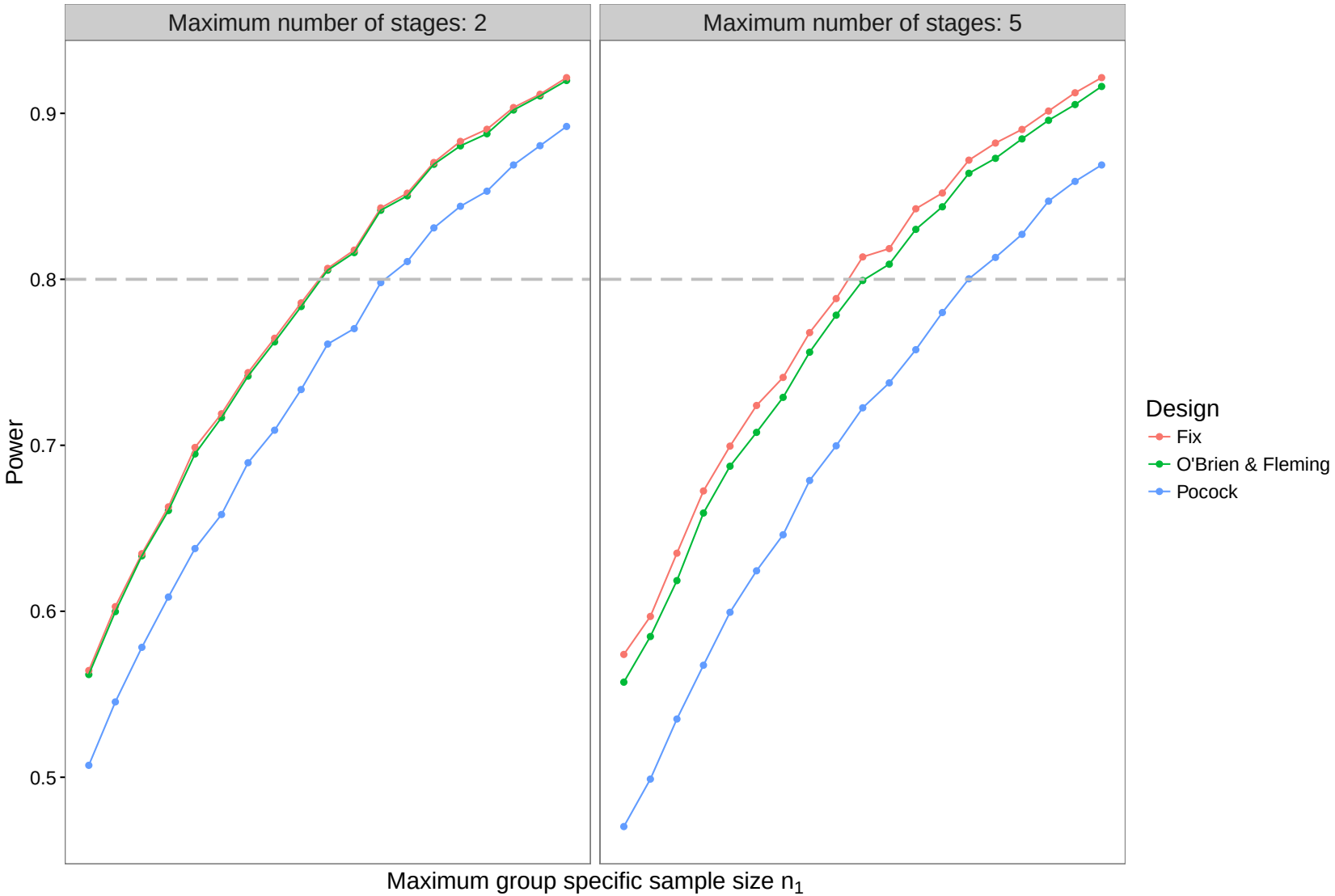
Simulation scenarios - power

- Parameters for the Monte Carlo simulation study of the power

Parameter	**Values**
Type I error rate α	0.025
Annual rate μ_1	0.0875
Annual rate μ_2	0.125
Rate ratio μ_1/μ_2	0.7
Group sample size $n_1 = n_2$	$600, 650, \ldots, 1500$
Shape parameter ϕ	5
Stages K	2, 5
Study duration	3.5 (years)
Recruitment period	1.25 (years)

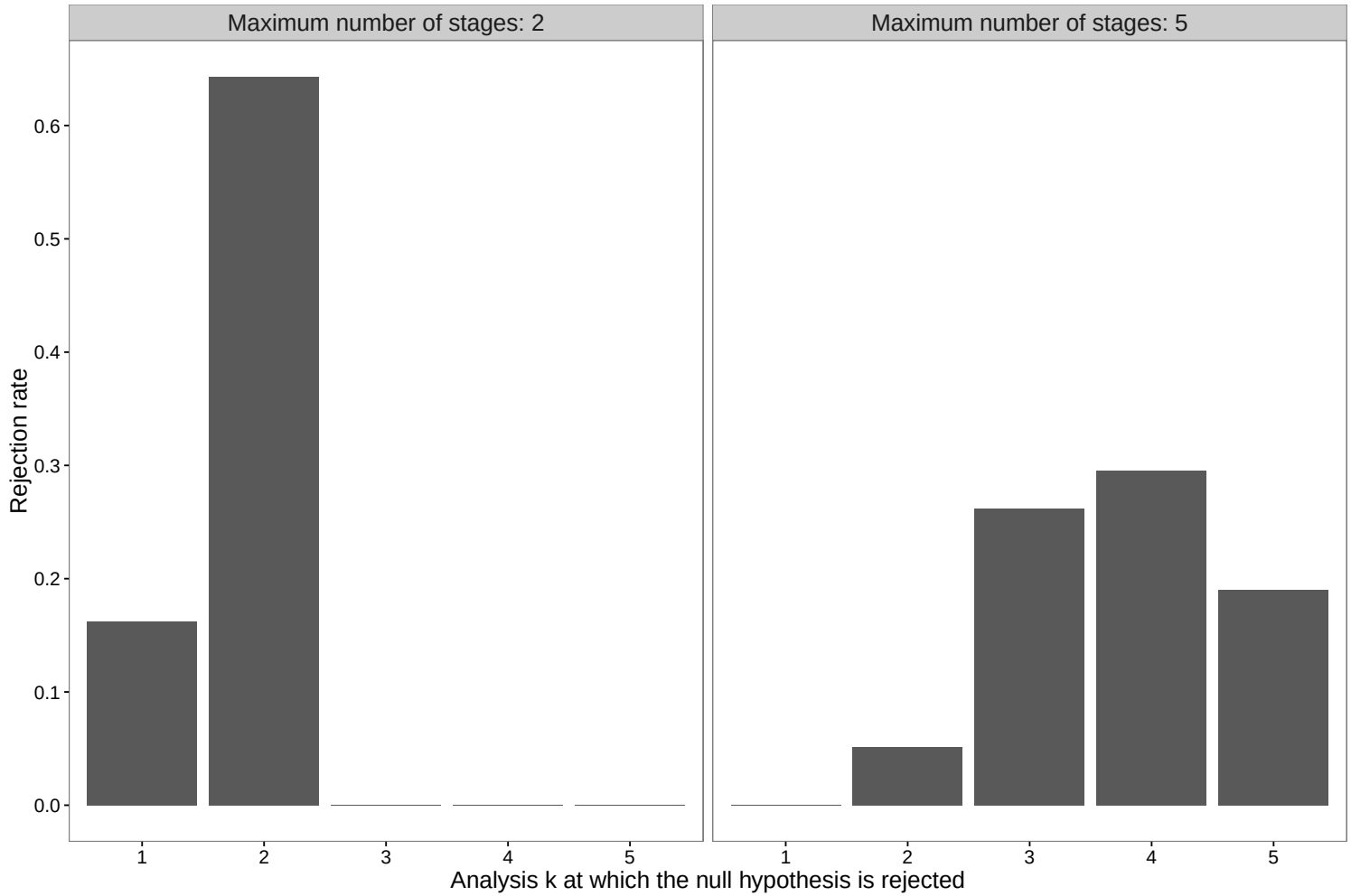
- 25 000 Monte Carlo replications per scenario

Results - power



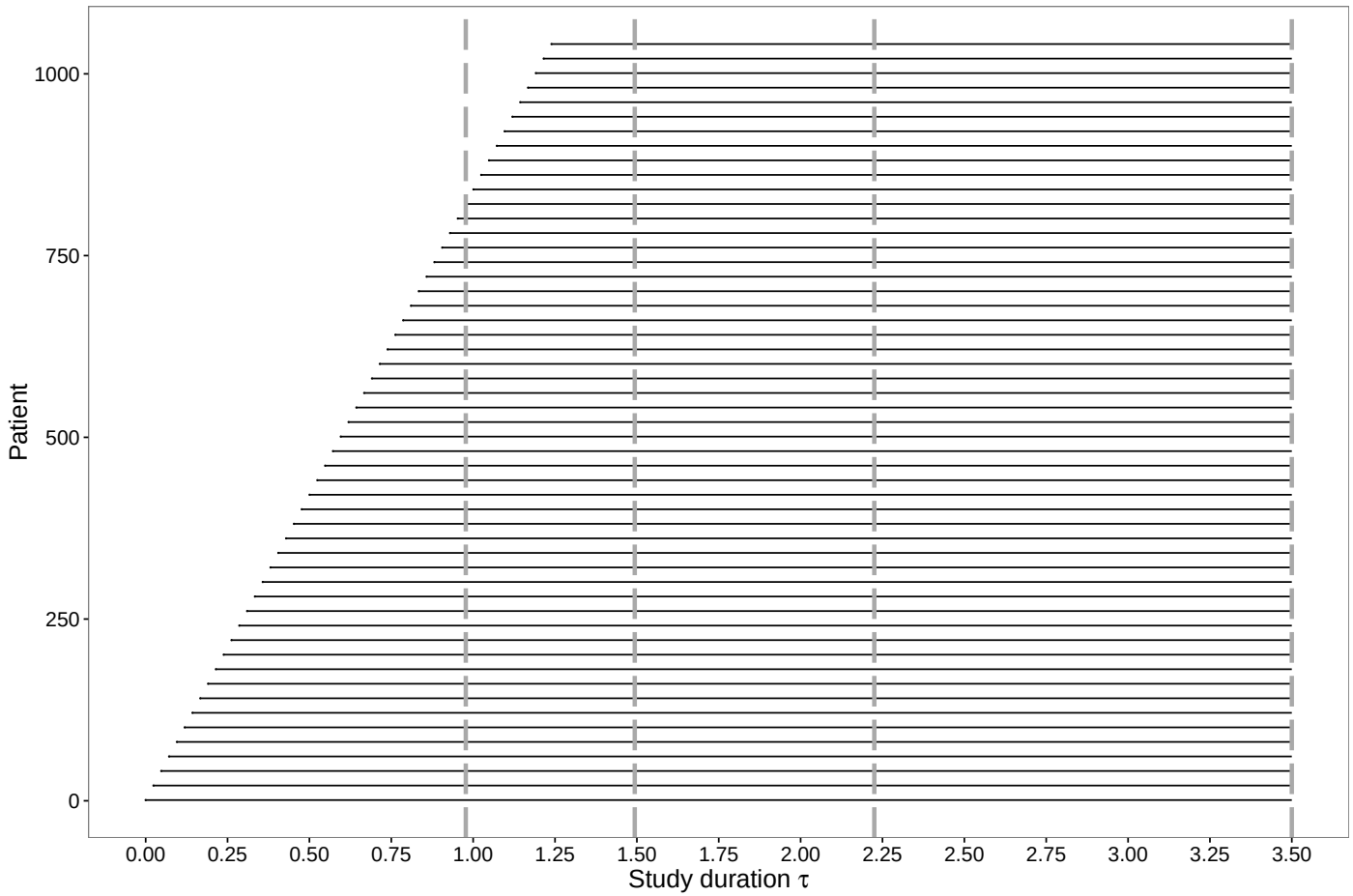
Results - stopping times

- Rejections by stage (O'Brien-Fleming, total power of 80%)



Results - gains from stopping early

- Study times at which in theory 25%, 50%, 75%, and 100% of the maximum information level $\mathcal{I}_{max}$ is attained



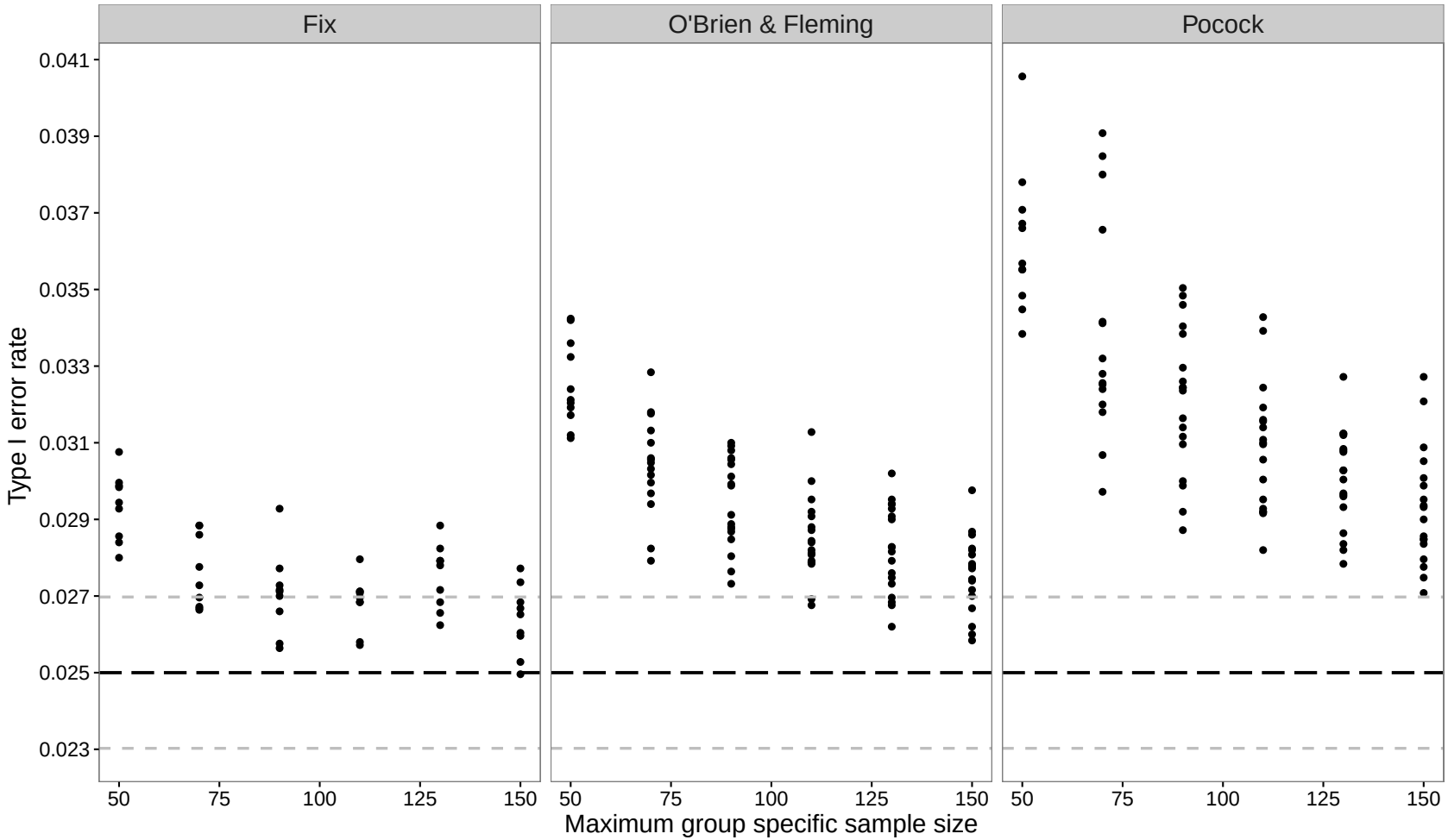
Simulation scenarios - type I error

- Simulation scenarios motivated by the CULAs from Example 2

Parameter	**Values**
Type I error rate α	0.025
6-month rates $\mu_1 = \mu_2$	6, 8, 10
Shape parameter ϕ	2, 3, 4
Group sample size $n_1 = n_2$	$50, 70, \ldots, 150$
Stages K	2, 3
Individual follow-up	0.5 (years)
Recruitment period	1.5 (years)

- 18 scenarios per group sample size for group sequential designs and 9 scenarios for the fixed design
- 25 000 Monte Carlo replications per scenario

Results - type I error



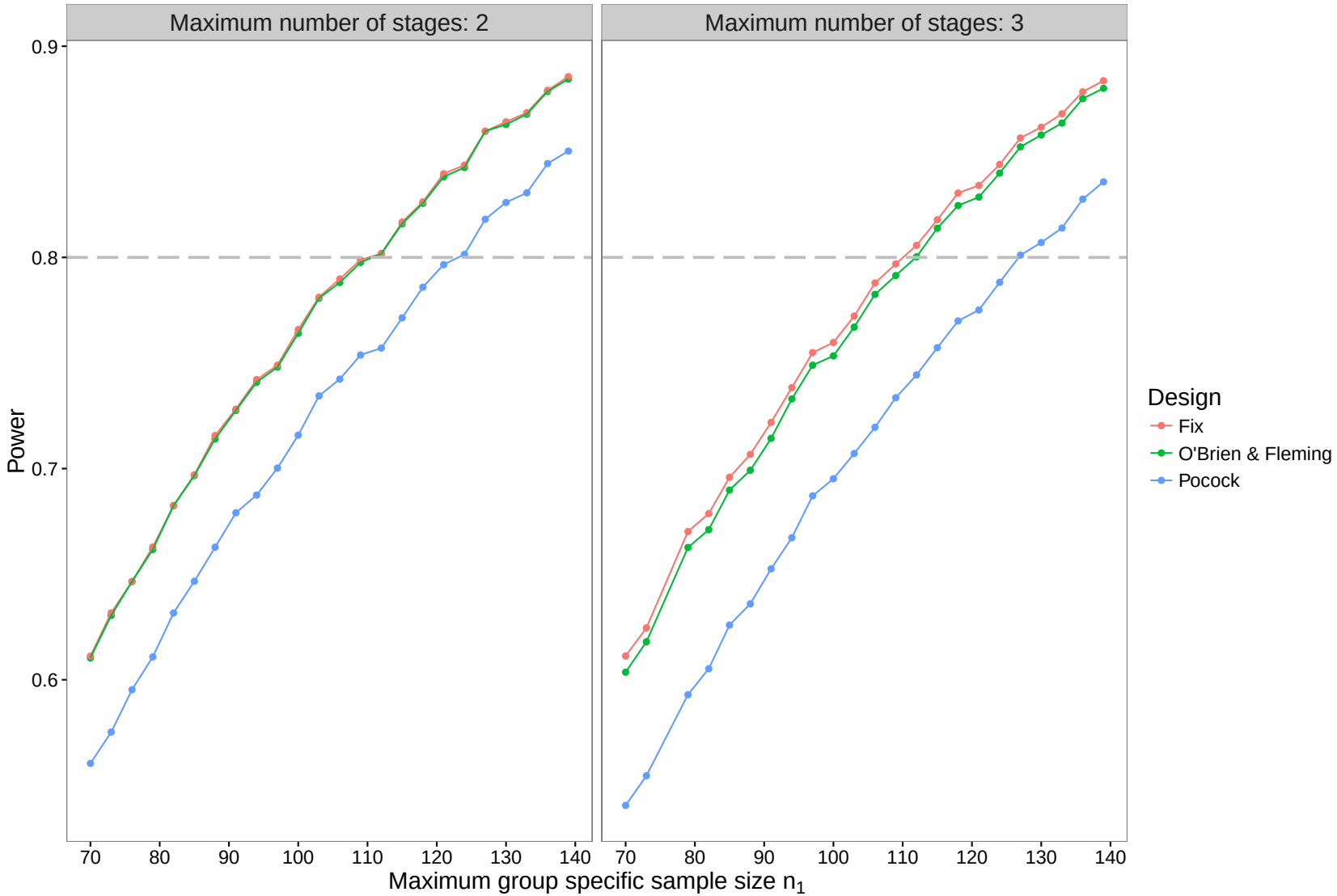
Simulation scenarios - power

- Parameters for the Monte Carlo simulation study of the power

Parameter	**Values**
Type I error rate α	0.025
6-month rate μ_1	4.2
6-month rate μ_2	8.4
Group sample size $n_1 = n_2$	$70, 75, \ldots, 140$
Shape parameter ϕ	3
Stages K	2
Individual follow-up	0.5 (years)
Recruitment period	1.5 (years)

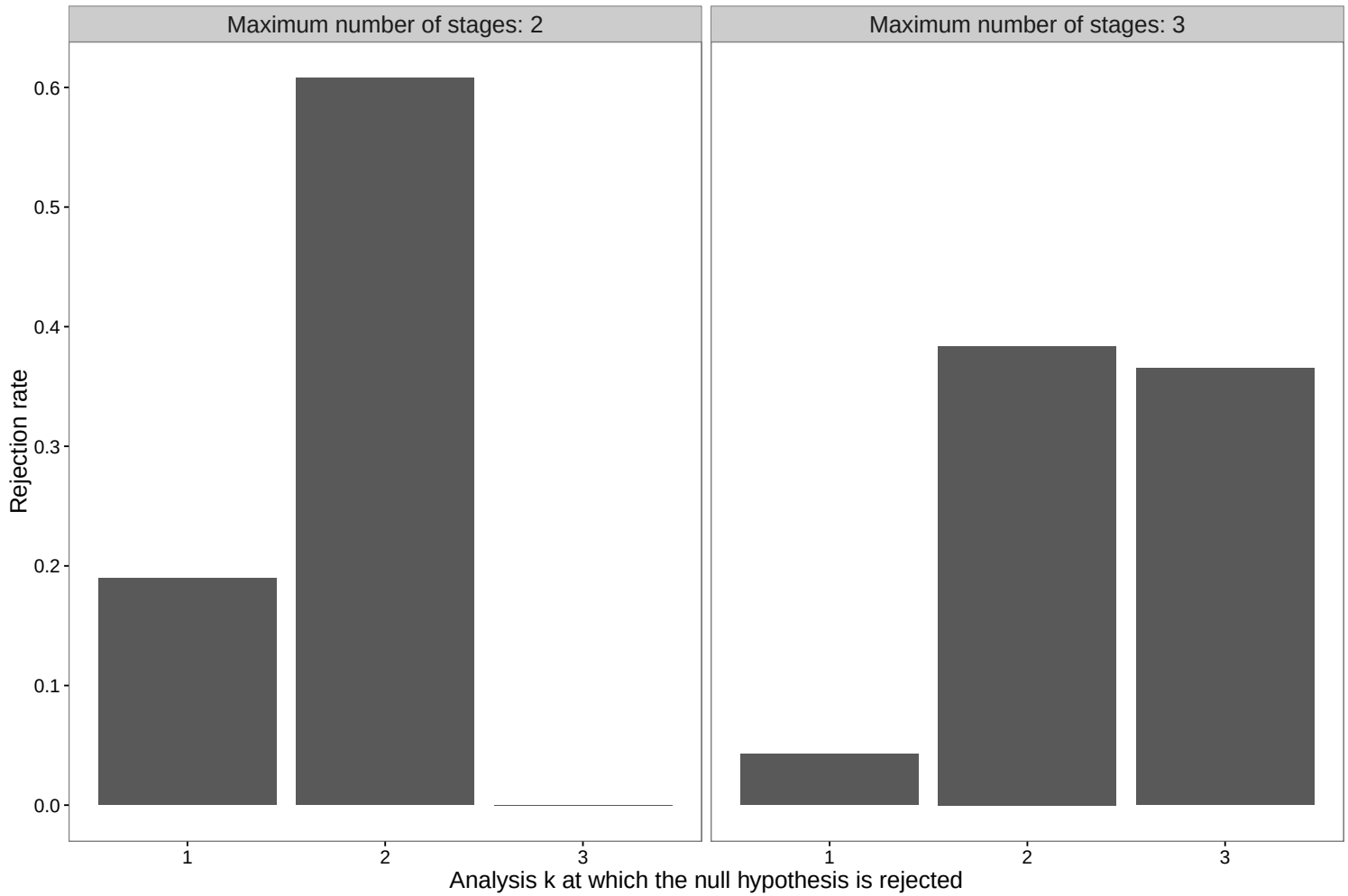
- 25 000 Monte Carlo replications per scenario

Results - power



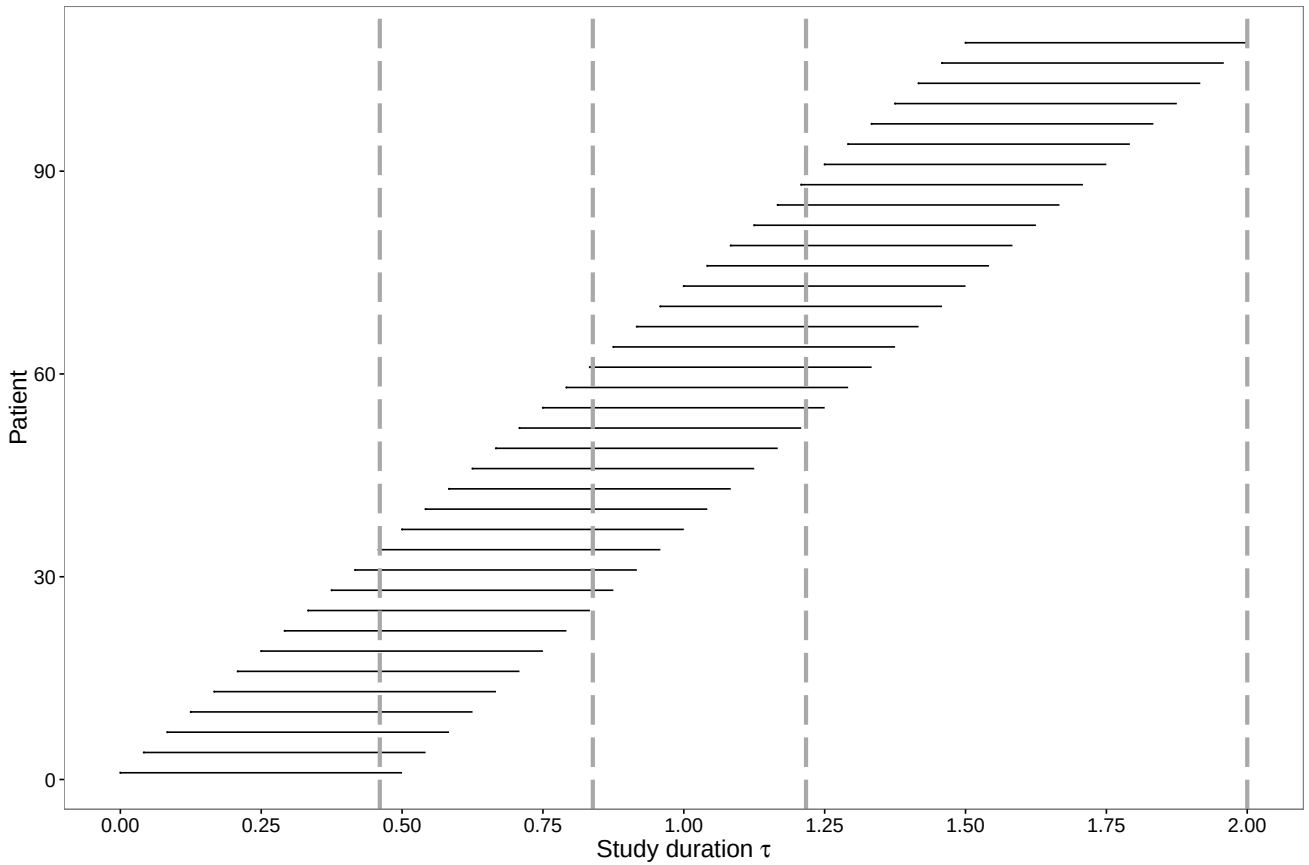
Results - analysis specific rejection rate

- Rate of stopping at a specific analysis at a power of 80%



Results - gains from stopping early

- Study times at which in theory 25%, 50%, 75%, and 100% of the maximum information level $\mathcal{I}_{max}$ is attained



Discussion and outlook

- Maximum-likelihood theory for negative binomial data results asymptotically in canonical form of joint distribution of test statistic
- Information level depends on rates, shape parameter, follow-up times, and sample size
- Future research on group sequential with negative binomial endpoints
  - Blinded information monitoring
  - Adaptive group sequential designs
  - Optimal designs
- Extend approach to quasi-Poisson models in the future

Bibliography

DO. Scharfstein, AA. Tsiatis, JM. Robins (1997).
Semiparametric efficiency and its implication on the design and analysis of group-sequential studies.
Journal of the American Statistical Association, **92**:1342–1350.

S. Yusuf et al. (2003)
Effects of candesartan in patients with chronic heart failure and preserved left-ventricular ejection fraction: the CHARM-Preserved Trial.
The Lancet, **362**:777–781.

JK. Rogers, SJ. Pocock, JJV. McMurray, CB. Granger, EL. Michelson, J. Östergren, MA. Pfeffer, SD. Solomon, K. Swedberg, S. Yusuf (2014)
Analysing recurrent hospitalizations in heart failure: a review of statistical methodology, with application to CHARM-Preserved.
European Journal of Heart Failure, **16**:33–40.

K. Selmaj, SKB. Li, HP. Hartung, B. Hemmer, L. Kappos, MS. Freedman, O. Stüve, P. Rieckmann, X. Montalban, T. Ziemssen (2013)
Siponimod for patients with relapsing-remitting multiple sclerosis (BOLD): an adaptive, dose-ranging, randomised, phase 2 study.
The Lancet Neurology, **12**:756–767.