UniversitätsKlinikum Heidelberg

Timing of subgroup selection in adaptive enrichment designs

Laura Kohlhas, Meinhard Kieser

Institute of Medical Biometry and Informatics,

University of Heidelberg

Adaptive Designs and Multiple Testing Procedures Workshop, 29.04.2016

GEFÖRDERT VOM

Bundesministerium für Bildung und Forschung

Introduction

Background

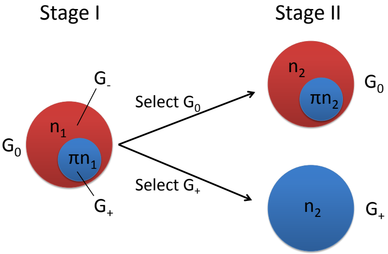
- a treatment may be more efficient in a subgroup compared to the total population
- example: subgroups identified by biomarker

Adaptive two-stage enrichment designs

- results of interim analysis are used to select the target population (subgroup and/or total population) / stop for futility
- only patients of the target population are enrolled in the second stage
- different classes of selection rules for selecting the target population

Notations

- G_0 = total population
- G_+ = subgroup
- G_- = $G_0 \setminus G_+$
- prevalence of subgroup G_+: π
- sample size:
  - stage I: n_1
  - stage II: n_2
  - overall: $N = n_1 + n_2$ **(fixed)**
- timing of interim analysis: $t = n_1 / N$



imbi

Notations and assumptions

- normally distributed outcome

Effect size

- in G_+: $\Delta_+ = \frac{\mu_{T+} - \mu_{C+}}{\sigma_+}$
- in G_-: $\Delta_- = \frac{\mu_{T-} - \mu_{C-}}{\sigma_-}$
- in G_0: $\Delta_0 = \pi\Delta_+ + (1 - \pi)\Delta_-$
- estimates in stage I: $\widehat{\Delta}_+, \widehat{\Delta}_-, \widehat{\Delta}_0$

Hypotheses and testing procedure

Hypotheses

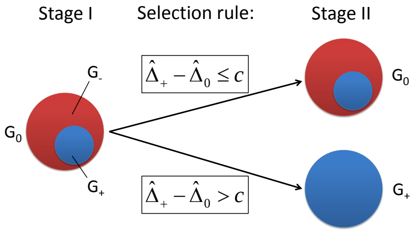
- $H_0^{(0)} : \Delta_0 \leq 0; \quad H_1^{(0)} : \Delta_0 > 0$
- $H_0^{(+)} : \Delta_+ \leq 0; \quad H_1^{(+)} : \Delta_+ > 0$
- $H_0^{(0+)} : \Delta_0 \leq 0 \cap \Delta_+ \leq 0; \quad H_1^{(0+)} : \Delta_0 > 0 \cup \Delta_+ > 0$

Combination of stage I and II: inverse normal combination test

Control of the FWER: closure principle (Simes procedure for $H_0^{(0+)}$)

Selection rule 1

Based on estimated difference in treatment effects



- Final analysis:
  - test $H_0^{(0)}$ if G_0 is selected
  - test $H_0^{(+)}$ if G_+ is selected

imbi

Selection rule 2

Based on estimated treatment effects (Jenkins et al. (2011))

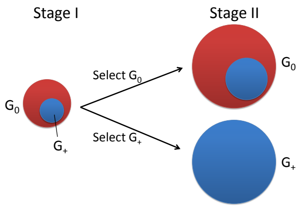
	$\widehat{\Delta}_0 > c_0$	$\widehat{\Delta}_0 \leq c_0$
$\widehat{\Delta}_+ > c_+$	continue with G_0, test $H_0^{(0)}$ and $H_0^{(+)}$	continue with G_+, test $H_0^{(+)}$
$\widehat{\Delta}_+ \leq c_+$	continue with G_0, test $H_0^{(0)}$	stop for futility

Jenkins M, Stone A, Jennison C (2011). An adaptive seamless phase II/III design for oncology trials with subpopulation selection using correlated survival endpoints. *Pharmaceutical Statistics* **10**: 347-356.

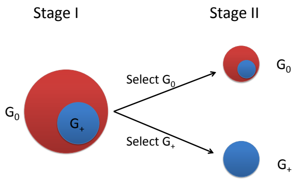
Timing of interim analysis?

- To what extent does timing affect the power (rejection of $H_0^{(0)}$ or $H_0^{(+)}$)?
- In which scenarios do we find large power differences?
- Does an early or a late interim analysis yield a higher power?

Early interim analysis:



Late interim analysis:



imbi

Simulation of power

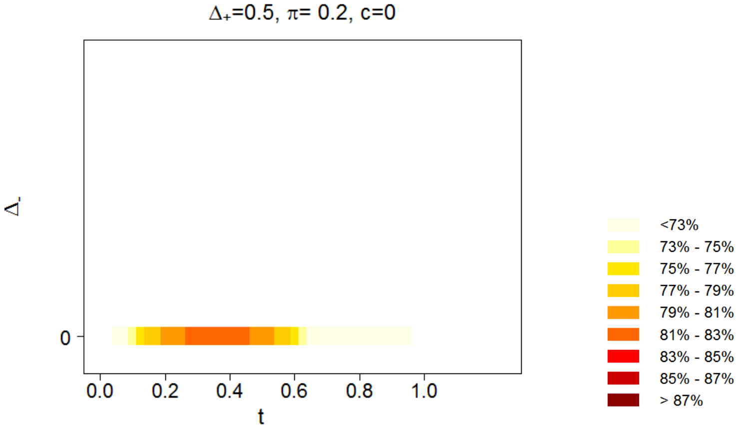
- simulation of 1,000,000 studies for 37 interim analysis times between 0.05 and 0.95
- $\Delta_+ = 0.5$
- $\Delta_- = 0, 0.05, 0.1, ..., 0.5$
- $\pi = 0.2, 0.7$
- N calculated for $t = 0.5$, and a power of 80% (probability to reject $H_0^{(0)}$ or $H_0^{(+)}$)
- selection rules 1 and 2

Results for selection rule based on estimated effect differences (selection rule 1)

Selection rule:

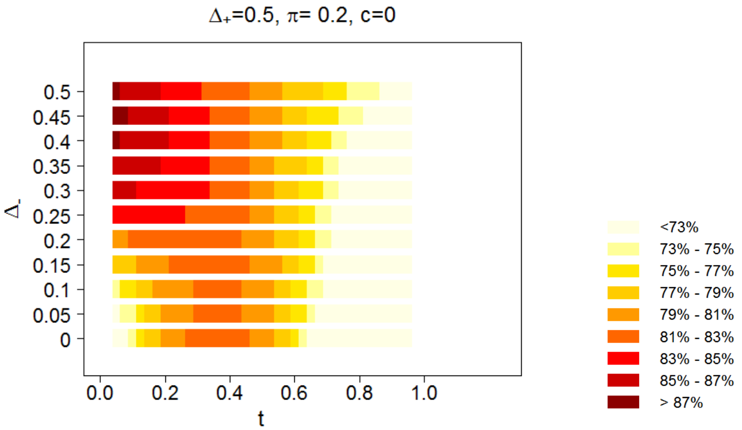
- $\widehat{\Delta}_{+} - \widehat{\Delta}_{0} \leq c \rightarrow$ continue with G_0 in stage II
- $\widehat{\Delta}_{+} - \widehat{\Delta}_{0} > c \rightarrow$ continue with G_+ in stage II

Power for selection rule 1 ($c = 0$, $\pi = 0.2$)



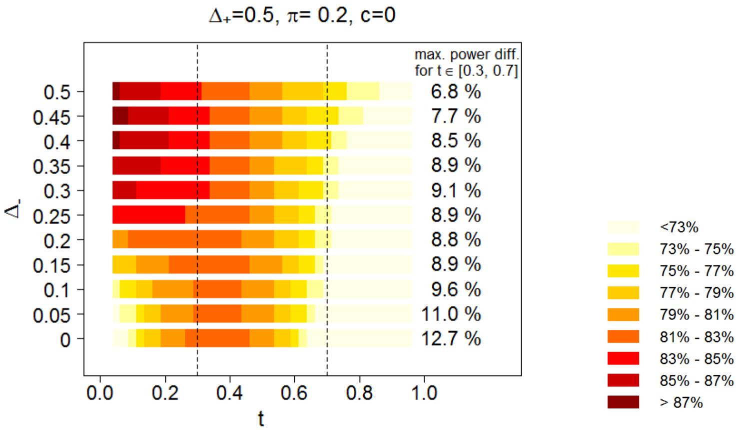
Probability to reject $H_0^{(0)}$ or $H_0^{(+)}$

Power for selection rule 1 ($c = 0$, $\pi = 0.2$)



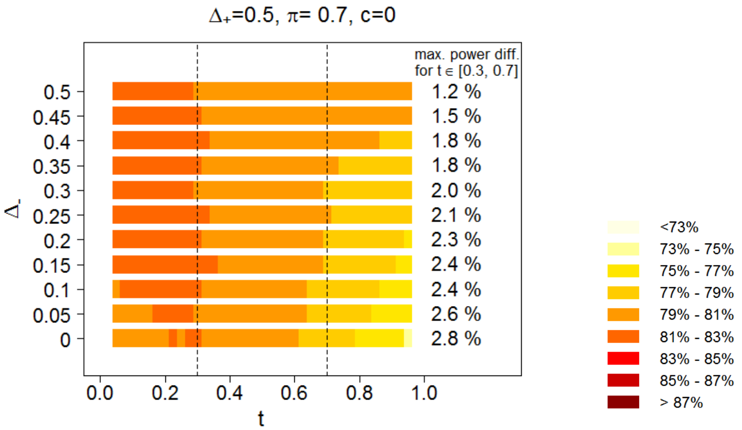
Probability to reject $H_0^{(0)}$ or $H_0^{(+)}$

Power for selection rule 1 ($c = 0$, $\pi = 0.2$)



Probability to reject $H_0^{(0)}$ or $H_0^{(+)}$

Power for selection rule 1 ($c = 0$, $\pi = 0.7$)



Probability to reject $H_0^{(0)}$ or $H_0^{(+)}$

Summary and further results (selection rule 1)

$c = 0$:

- smaller power for high t
- for small Δ_-: highest power between $t = 0.3$ and 0.4
- power varies more for smaller π

$c > 0$

- for higher c, power differences are smaller

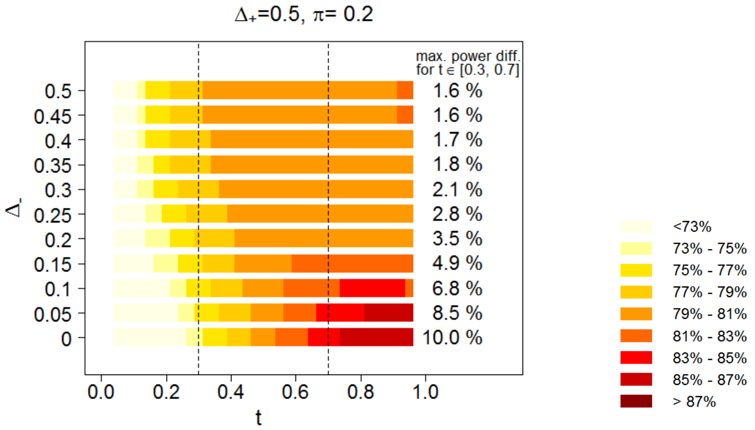
Results for selection rule based on estimated treatment effects (Jenkins et al. (2011))

	$\widehat{\Delta}_0 > c_0$	$\widehat{\Delta}_0 \leq c_0$
$\widehat{\Delta}_+ > c_+$	continue with G_0, test $H_0^{(0)}$ and $H_0^{(+)}$	continue with G_+, test $H_0^{(+)}$
$\widehat{\Delta}_+ \leq c_+$	continue with G_0, test $H_0^{(0)}$	stop for futility

Scenarios:

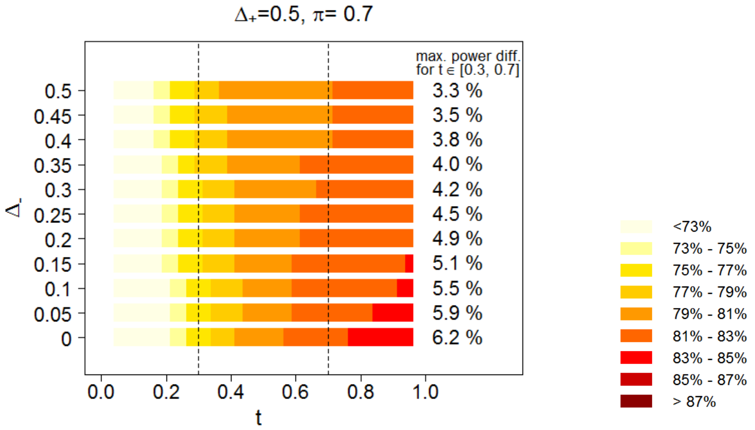
- $c_0 = 0.1, c_+ = 0.3$

Power for selection rule based on estimated treatment effects ($\pi = 0.2$)



Probability to reject at least one of the hypotheses $H_0^{(0)}$ or $H_0^{(+)}$

Power for selection rule based on estimated treatment effects ($\pi = 0.7$)



Probability to reject at least one of the hypotheses $H_0^{(0)}$ or $H_0^{(+)}$

Summary and further results (selection rule 2)

$c_0 = 0.1$, $c_+ = 0.3$

- small power for approximately $t < 0.4$ (the smaller t, the smaller the power)
- relative constant power function for approximately $t > 0.4$
  - power increases for high π and small Δ_-

different c_0, c_+

- power differences are smaller for smaller c_+ and smaller c_0

Conclusions

- power as a function of timing varies for different selection rules, effect sizes and prevalences
- in many scenarios, the power at $t = 0.5$ is not much smaller than the power maximum
- caution: power can be small for certain t:
  - selection rule based on estimated effect differences:
    - small power for high t (especially for small π, small Δ_- and small c)
  - selection rule based on estimated effects (Jenkins):
    - small power for small t

Conclusions

- in some scenarios, a higher power can be achieved by choosing a different interim analysis time to $t = 0.5$:
  - selection rule based on estimated effect differences:
    - power gain for smaller t (especially for small π)
  - selection rule based on estimated effects (Jenkins):
    - power gain for higher t if π is high, or π is small and Δ_- is small
- no general rule or recommendation that uniformly fits to all scenarios

Thank you for your attention!