Advantages of Study Designs with Multiple Endpoints or Treatments

Gerd Rosenkranz
Center for Medical Statistics, Informatics and Intelligent Systems
Medical University of Vienna

Asterix End Symposium, Zaandam NL, 19th September 2017

Objectives of this talk

- To highlight new statistical methods for randomized studies with **multiple endpoints** or **multiple treatment arms** developed in the course of the Asterix project
- Talk will focus on relevance of approaches for rare diseases rather than on explanations why they work

- **Acknowledgements:**
  Susanne Urach, Robin Ristl, Martin Posch

Why multiple endpoints?

- To characterize the potential impact of treatment on different aspects of a disease
  - By **combining** multiple endpoints into **a single composite** endpoint
    - Example: treatment failure in transplantation studies consists of graft rejection, graft loss or death
  - By analysing multiple endpoints **separately**
    - Example: cognitive and functional scale in Alzheimer's disease

Multiple endpoints analysed separately

- For **separate** analyses it should be clear whether an impact on **any**, a **minimum number** or **all endpoints** is necessary to achieve the trial objectives
- The latter is important to account appropriately for **multiplicity** issues
- In the following we deal specifically with **co-primary** endpoints, i.e., an impact on all endpoints is required

Co-primary endpoints in rare diseases

- Eculizumab in paroxysmal nocturnal haemoglobinuria (PNH)
  - PNH rare blood disorder (0.1 in 10000) with high morbidity and mortality
  - Eculizumab approved based on the following results:

Endpoint	Placebo	Eculizumab	P-value
Patients with stabilized haemoglobin levels at study end	0/44	21/43	<0.001
#packed red blood cell units transfused during study	11.0(0.83)	3.0(0.67)	<0.001

Data from Eculizumab European Public Assessment Report, EMA 2007

Co-primary endpoints in rare diseases

- Treatment effect in all primary endpoints required for success
- Classical procedure:
  - Test each endpoint at a one-sided level α
  - Either you **win in all** endpoints – or you **loose** completely
- **Fall-back procedures** provide a **second chance** to win in at least one/some endpoint/s **without compromising the family-wise type I error rate**

Co-primary endpoints in rare diseases

- In the Asterix project, Ristl et al. (2016) developed fall-back procedures for 2 or 3 co-primary endpoints
- Fall-back procedure (for two co-primary endpoints):
  - Test each endpoint at one-sided level α
  - If you win in all, your trial reached its primary objective
  - Otherwise you may **win on the endpoint significant at level $\alpha/2$** (if any) in case the other one does not go "too far" in the wrong direction
- **Fall-back analysis does not affect the analysis of co-primary endpoints**

Co-primary endpoints in rare diseases

- Example
  - One sided $\alpha = 0.025$
  - Trial result: $p_1 = 0.01 < \alpha$, $p_2 = 0.03 > \alpha$
  - Fall-back analysis: $p_1 < \alpha/2 = 0.0125$ and $p_1 + p_2 \leq 1$
  - Significant impact of treatment on the first endpoint at a family-wise level $\alpha = 0.025$

Why multiple treatment arms?

- Comparing several treatments/doses with a control in one study requires less patients than conducting several trials
- More patients are randomized to a treatment arm than to a control
- Possibility of head-to-head comparison of several treatments

Multi-arm clinical trials in rare diseases

- Tadalafil in long-term treatment of pulmonary arterial hypertension (PAH) (prevalence: 1.8 in 10000)
- **742 subjects** from 151 centers in 39 countries randomized 1:1:1 to placebo (250), taldalafil 3mg (250) or 10mg (242)
- Primary endpoint: time to first morbidity/mortality event (defined by a long list not reproduced here)

Statistic	3mg	10mg
Hazard ratio (97.5% CI)	0.704 (0.516, 0.960)	0.547 (0.392, 0.762)
Logrank p-value	0.0108	< 0.0001

Data from Tadalafil European Public Assessment Report, EMA 2013

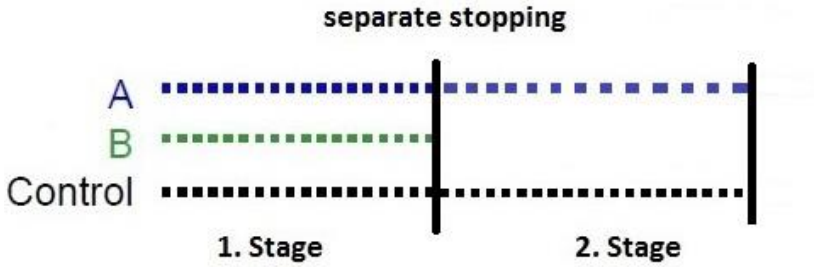
Multi-arm clinical trials in rare diseases

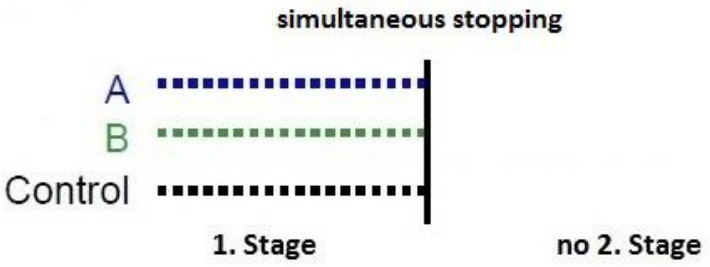
- Can we stop multi-arm clinical trials early, e.g., for futility or efficacy, to reduce sample size and trial duration?
- In the Asterix project, Urach & Posch (2016) developed multi-arm sequential designs with a **simultaneous stopping** rule to reduce the average sample size of a trial

Multi-arm sequential trials

- **Objective 1:**
  Detect **all** efficacious treatments/doses
- **Stop the treatment arm** where efficacy is shown and continue with the remaining treatment arms (separate stopping)

- **Objective 2**:
  Detect **at least one** efficacious treatment/dose
- **Stop the entire trial** if efficacy is shown for one treatment (simultaneous stopping)





Multi-arm sequential trials in rare diseases

Three arm trial in systemic sclerosis comparing relaxin 10ug/kg or 25ug/kg daily with placebo (relevant standardised difference = 0.4)

Stopping rule	ASM (0.4, 0.4)	ASM (0, 0)	Nmax (N=354)	Power to detect one dose	Power to detect both doses
Separate	295	235	390	0.91	0.70
Simultaneous	**272**	239	402	0.92	0.64

Average sample sizes for a relaxin effect of 0.4 for each dose, no relaxin effect, maximum sample size Nmax, and power to detect at least one or both efficacious treatments for separate and simultaneous stopping for efficacy or futility

Nmax and N for single stage design determined such that the efficacious treatment can be detected with 80% chance if one dose is efficacious (0.4) while the other is not.

Other project results on clinical trials with multiple endpoints/treatments/stages

- Optimal exact tests for multiple binary endpoints (plus software package that implements the proposed methods)
- Small sample simultaneous inference for multiple generalized linear models with dependent observations
- Analysis procedures for adaptive designs with a time to event endpoint
- Group sequential and adaptive designs with multiple endpoints
- Combined integrated protocol/basket trial design for a first in human study

Concluding remarks

- Sensible to consider for studies in rare diseases
  - Several endpoints instead of one
  - Several test treatments or doses instead of one
  - Sequential designs instead of single stage designs
- There is no holy grail – no approach is uniformly best but may be more appropriate than others in a specific context
- Trial objectives to be clearly spelled out
- Appropriate consideration of multiplicity is crucial
- Comforting to have alternative options to choose from