Treatment selection in multi-arm, multi-stage clinical studies

Thomas Jaki[1], Dominic Magirr[1], Nigel Stallard[2]

[1] Department of Mathematics and Statistics
Lancaster University

[2] Warwick Medical School

10. September 2013

MRC Hubs for Trials Methodology Research
North West Hub

Medical and Pharmaceutical Statistics Research Unit

Outline

Medical and Pharmaceutical Statistics Research Unit

A multi-arm phase II trial

TAILoR: Telmisartin And Insulin Resistance in HIV.

Ambition: Reduce insulin resistance in HIV patients receiving antiretroviral therapy.

Treatment: 4 different doses of a licensed drug (in a different therapeutic area). Inappropriate to assume a monotone dose-response relationship.

Endpoint: Change in insulin resistance as measured using HOMA-IR index (baseline - week 12).

Testing multiple hypothesis

Responses: $X_{k,i} \sim N(\mu_k, \sigma^2), \quad i = 1, \ldots, n, k = 0, 1, \ldots, 4$

Individual null hypotheses:

$$\begin{array}{ll} H_1: & \mu_1 \leq \mu_0 \\ \vdots & \vdots \\ H_4: & \mu_4 \leq \mu_0 \end{array}$$

Teststatistics: $Z_k = \frac{\bar{X}_k - \bar{X}_0}{\sigma\sqrt{\frac{2}{n}}}$ for $k = 1, \ldots, 4$

Familywise error rate (FWER):

$P(\text{reject at least one true } H_k) \leq \alpha$

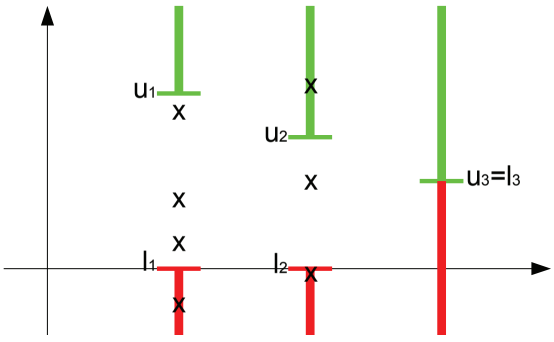
Analyses: J analysis planned

Design options

- Pre-planned adaptive designs (e.g. Stallard & Todd, 2003; Magirr et al, 2012)
  - \+ Sufficient statistics
  - \+ Analytic sample sizes
  - – Unexpected modifications difficult
- Flexible adaptive designs (e.g. Bretz et al, 2006)
  - \+ Very flexible
  - – Often not based on sufficient statistics

At interim analysis j

- if $Z_{k,j} < l_j$: treatment k is dropped from trial.
- if $Z_{k,j} > u_j$: can reject H_k and stop trial.



More Multiple Testing

- J-stage trial $\Rightarrow$ up to $4J$ hypothesis tests.

Strong control of FWER

$$P(\text{reject at least one true } H_k) \leq \alpha$$

Weak control of FWER

$$P(\text{reject at least one true } H_k \mid H_G) \leq \alpha$$

Fact: for this design, Strong control of FWER $\Leftrightarrow$ Weak control of FWER (Magirr et al, 2012).

Computing $P(\text{reject at least one true } H_k | H_G)$

Problem: Test statistics are correlated due to the common control.

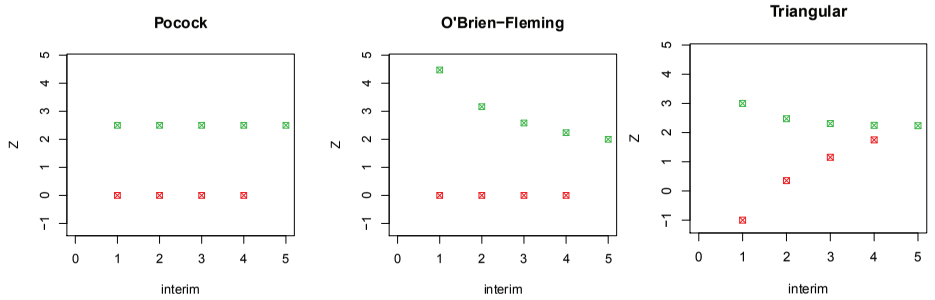
Solution: Condition on $\hat{\mu}_{0,J}$, the vector of sample means on control.

$$\alpha = 1 - \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{J \text{ times}} \left[\sum_{j=1}^{J} P\left\{ \left(\bigcap_{i=1}^{j-1} B_{1,i} \right) \cap A_{1,j} \middle| \hat{\mu}_0,\ H_G \right\} \right]^K dF(\hat{\mu}_0)$$

- $2J - 1$ unknowns $(l_1, \ldots, l_{J-1}, u_1, \ldots, u_J)$.

Boundary Constraints

For $J > 1$ set $l_h = g(u_J)$ and $u_h = f(u_J),\ h = 1, \ldots, J-1$.



Medical and Pharmaceutical Statistics Research Unit

Example

Table: TAILoR trial with $1-3$ stages. Boundaries and sample size for $\alpha = 0.05$ (one-sided) and $\beta = 0.1$ and equal sample size per arm and stage (n) is used. Total sample size is denoted $\mathcal{N}$.

J	Design	l	u	n	E($\mathcal{N}\|H_G$)	E($\mathcal{N}$|LFC)
1	Fixed	$-\infty$	2.16	84	420	420
2	OBF	$(0, 2.17)$	$(3.07, 2.17)$	44	342.3	346.9
	P	$(0, 2.38)$	$(2.38, 2.38)$	50	382.9	318.9
	T	$(0.81, 2.29)$	$(2.43, 2.29)$	50	309.4	302.0
3	OBF	$(0, 0, 2.18)$	$(3.78, 2.67, 2.18)$	31	307.6	317.3
	P	$(0, 0, 2.48)$	$(2.48, 2.48, 2.48)$	37	359.2	286.3
	T	$(0, 1.44, 2.34)$	$(2.71, 2.39, 2.34)$	36	292.5	285.2

Medical and Pharmaceutical Statistics Research Unit

Fully flexible designs

Use two fundamental concepts:

- p-value combination
- closed testing

Choices, choices, choices

Advantages of pre-planned adaptive tests

1. Sufficient statistics
2. Sample size calculations
3. Confidence intervals

BUT treatment selection is typically more complex than a pre-planned rule so do we need to use flexible adaptive designs?

Conditional error, König et al, 2008

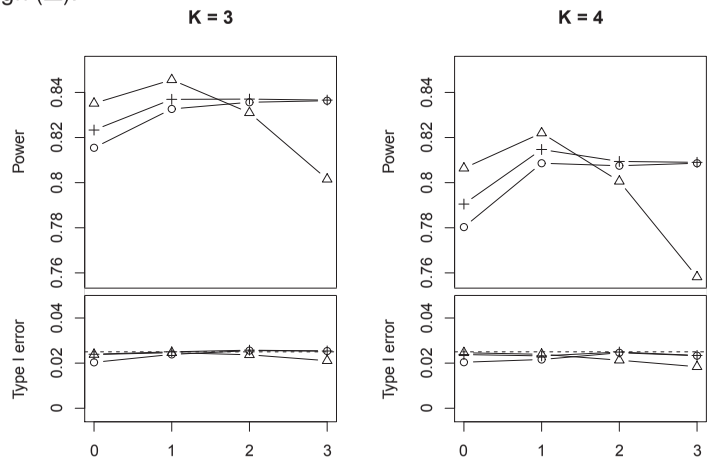
- The conditional error, $A(X)$, is the maximal probability under H of rejecting H with the original test, conditional on the interim data X
- $B(X)$ is the conditional error for a new test following an unplanned adaption.
- If $B(X) \leq A(X)$ the new test controls the FWER.
- Can be used to update boundaries to reflect a reduced number of treatments being continued

Simulation

- Compare pre-planned adaptive (without and with conditional error adjustment) and fully flexible design
- No early stopping for futility (no selection) in design
- Selection rule: All treatment that are within ϵ of best performing
- $K = 3$ or 4 and $J = 3$
- 100,000 simulation runs, $\alpha = 0.025$
- 1 treatment with an effect of 1 standard deviation

Simulation

Figure: Simulated type I error and power for the pre-planned adaptive method without (o) and with conditional error method (+) and the fully flexible design (Δ).



Discussion

- Different ideas to design multi-arm multi-stage trials discussed
- Introduced a strategy to add flexibility to pre-planned adaptive designs
- Conditional error used in unconventional way - traditionally used for data-dependent adaptations (e.g. sample-size re-estimation)
- Should be used as a back-up plan rather than the planned test
- Design of pre-planned adaptive designs available in `MAMS` package

References

- Bretz F, Schmidli H, König F,Racine A, Maurer W (2006) Confirmatory Seamless Phase II/III Clinical Trials with Hypotheses Selection at Interim: General Concepts. *Biometrical Journal*. 48(4):623â634.
- Dunnett CW (1984) Selection of the best treatment in comparison to a control with an application to a medical trial. In *Design of Experiments: Ranking and Selection*
- König F, Brannath W, Bretz F, Posch M (2008) Adaptive Dunnett Tests for Treatment Selection. *Statistics in Medicine* **27**, 1612-1625.
- Magirr D, Jaki T, Whitehead J (2012) A generalized Dunnett test for multi-arm, multi-stage clinical studies with treatment selection. *Biometrika*. 99(2):494-501
- Pocock, S. (1977) Group sequential methods in the design and analysis of clinical trials. *Biometrika* **64**, 191-199.
- O'Brien, P. and Fleming, T. (1979) A multiple testing procedure for clinical trials. *Biometrics* **35**, 549-556.
- Stallard N and Todd S (2003) Sequential Designs for Phase III Clinical Trials Incorporating Treatment Selection. *Statistics in Medicine*. 22:689-703.
- Whitehead, J. (1997) *The design and analysis of sequential clinical trials*. Wiley: Chichester