Mathematics & Statistics | Lancaster University

Adaptive designs: why to use them and some experiences

Thomas Jaki

t.jaki@lancaster.ac.uk

Medical and Pharmaceutical Statistics Research Unit

Mathematics & Statistics | Lancaster University

MRC | Hubs for Trials Methodology Research

Probability of approval of new drug regimen (Wong et al., 2019)

- phase II: $<40\%$
- phase III: $\sim 60\%$

Probability of approval of new drug regimen (Wong et al., 2019)

- phase II: $<40\%$
- phase III: $\sim 60\%$

Likely causes for failure:

- taking forward futile treatments
- studying the wrong patient population
- poor precision (optimal dose, maximum tolerated dose, safety)

Can we do better?

Adaptive Designs

Idea

Modify an ongoing trial

Idea

Modify an ongoing trial

by design

based on reviewing accrued data at interim

to enhance flexibility

Idea

Modify an ongoing trial

by design

based on reviewing accrued data at interim

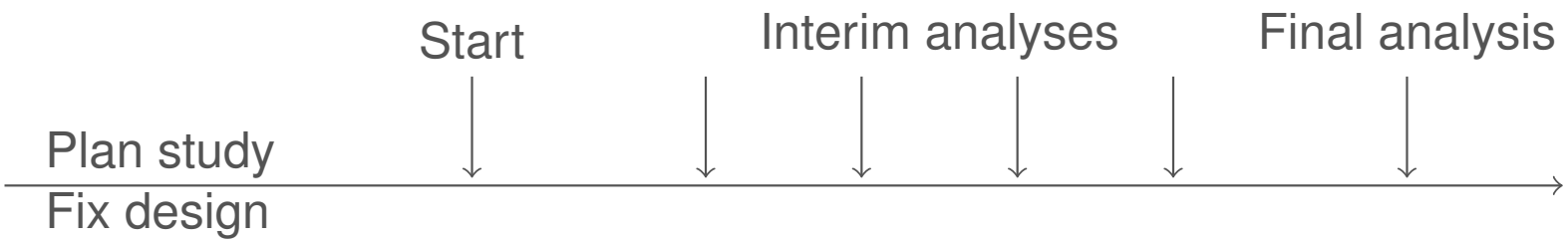
to enhance flexibility

without undermining the study's integrity and validity.

(Chow et al. 2005)

Fixed sample design

Adaptive design



At each interim:

- decide whether or not to stop
- change sample size
- change allocation ratio
- drop or add a dose
- . . .

The TAILoR study

Mathematics & Statistics | Lancaster University

TAILoR: Telmisartin And Insulin Resistance in HIV.

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

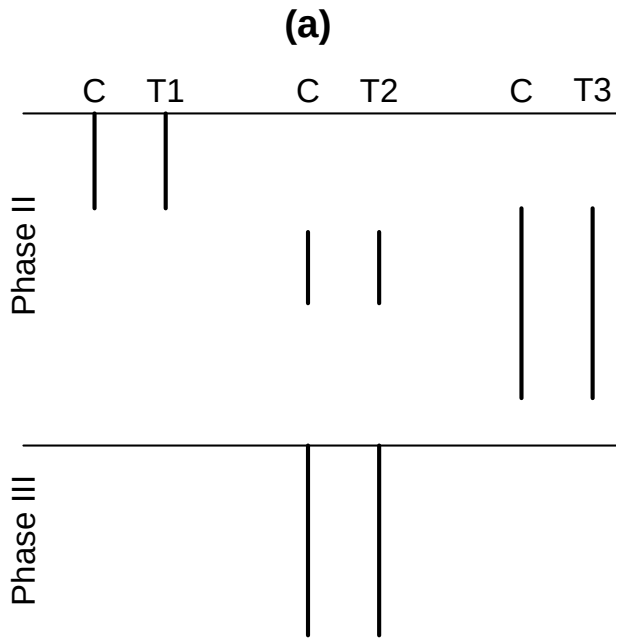
Treatment: Different doses of a licensed drug (in a different therapeutic area). Inappropriate to assume a monotone dose-response relationship.

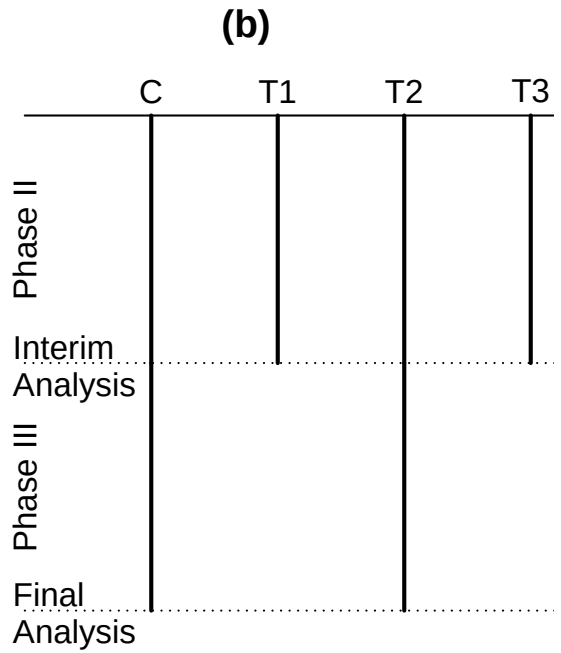
Endpoint: Change in insulin resistance as measured using HOMA-IR index (baseline - week 12).

Mathematics & Statistics | Lancaster University

Multi-arm multi-stage trials

- Compare several active treatments against common control
- Select one of more treatment at interim





Testing multiple hypothesis

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

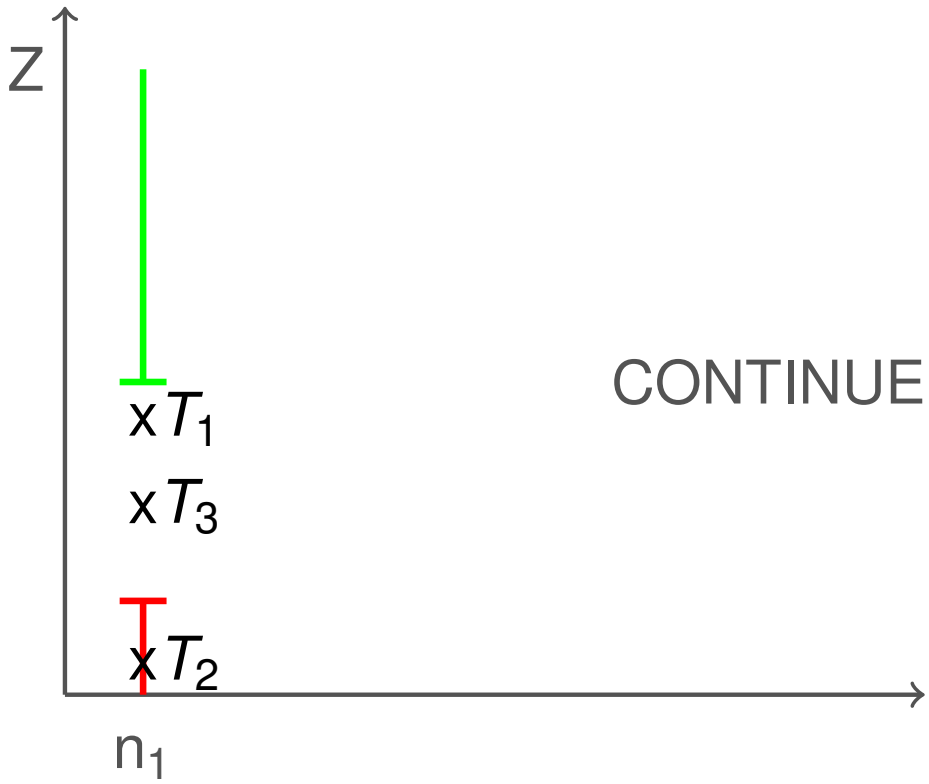
Individual null hypotheses:

$$\begin{aligned} H_1 : &\quad \mu_1 \leq \mu_0 \\ &\vdots \qquad\quad \vdots \\ H_K : &\quad \mu_K \leq \mu_0 \end{aligned}$$

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

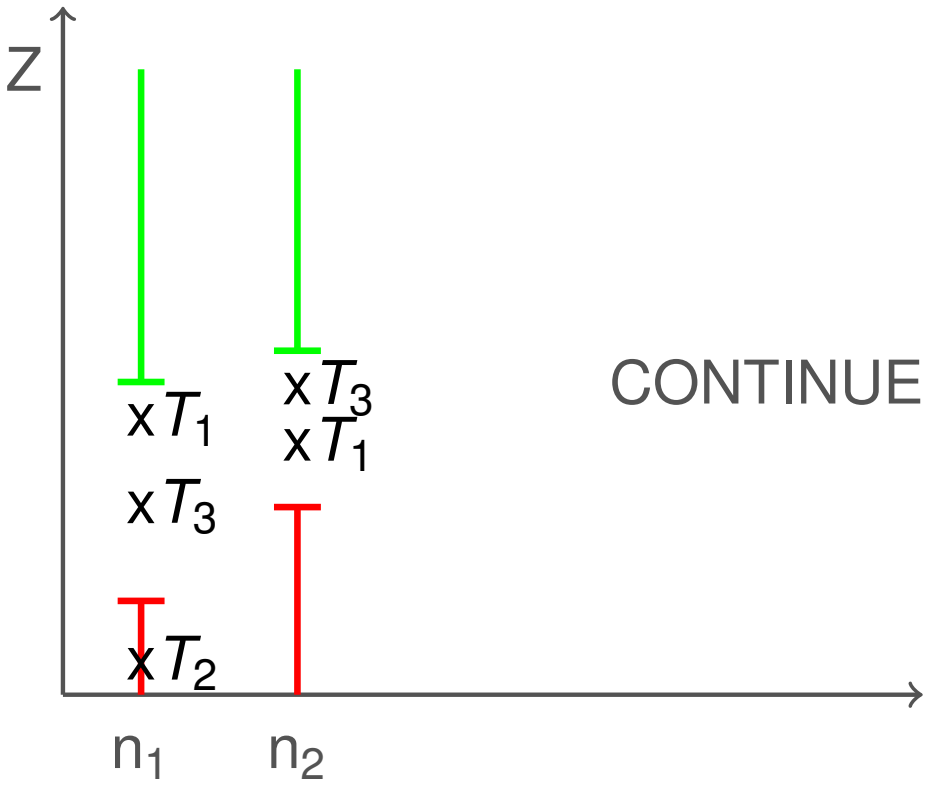
Familywise error rate (FWER): $P(\text{reject at least one true } H_k) \leq \alpha$

A multi-arm multi-stage design



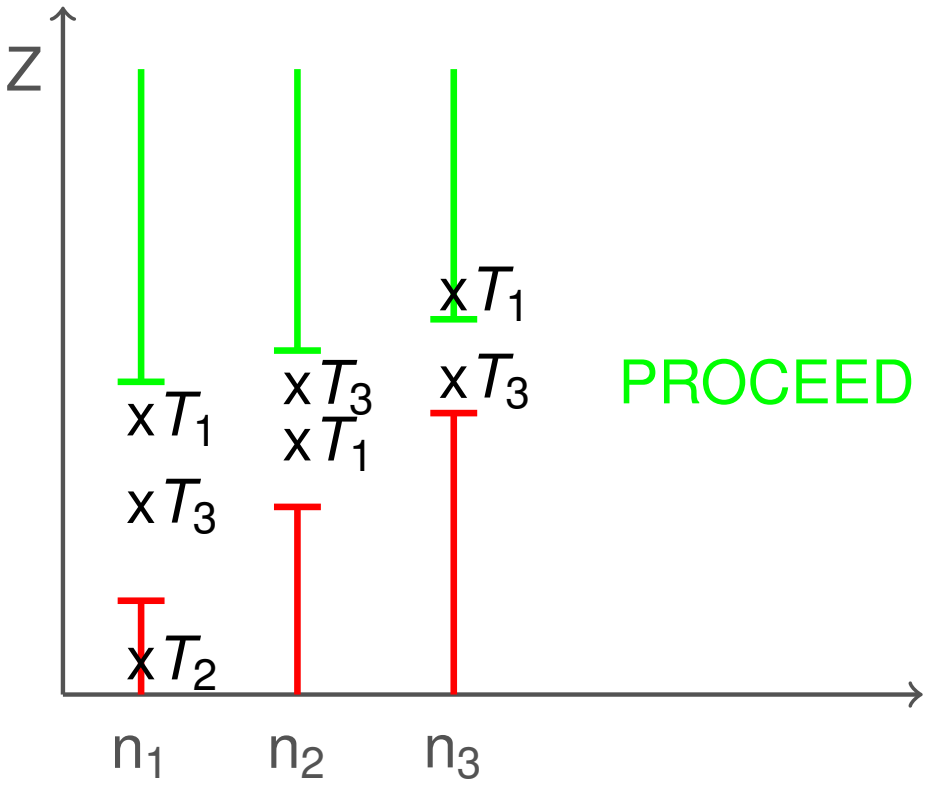
(Magirr et al, 2012)

A multi-stage design



(Magirr et al, 2012)

A multi-stage design



(Magirr et al, 2012)

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

Mathematics & Statistics | Lancaster University

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

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

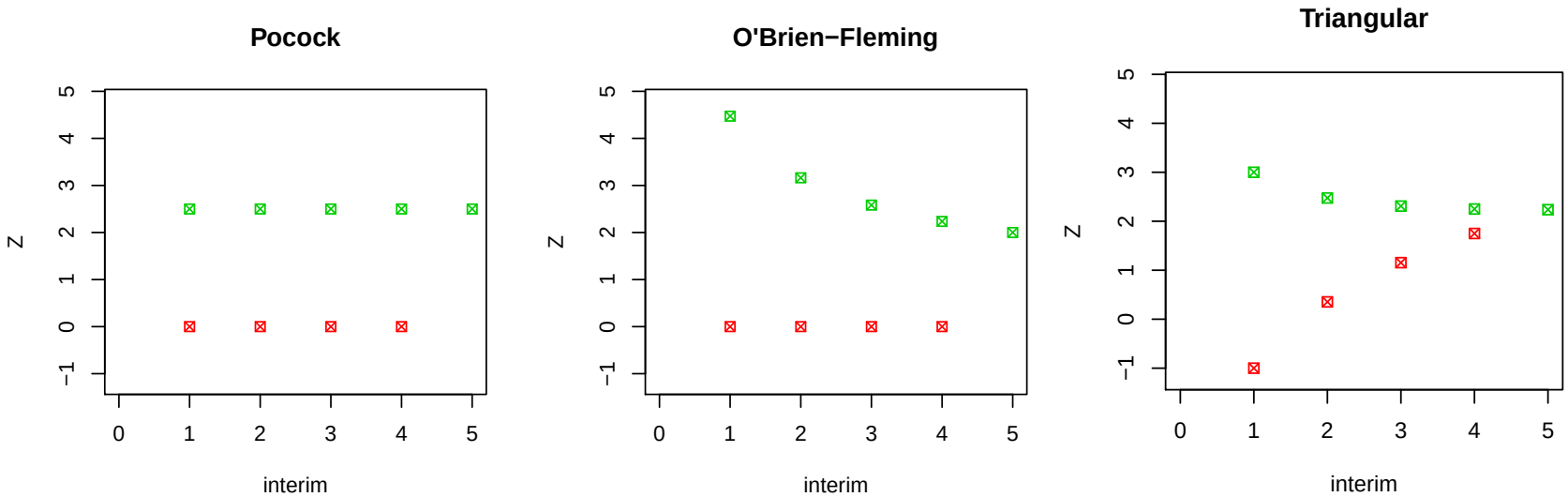
$$\alpha = 1 - \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{J \text{ times}} \left[\sum_{j=1}^{J} P\left\{ \left(\bigcap_{i=1}^{j-1} B_{1,i} \right) \cap A_{1,j} \middle| \bar{x}_0, H_G \right\} \right]^K dF(\bar{x}_0)$$

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

Mathematics & Statistics | Lancaster University

Boundary Constraints

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



Design considerations

Mathematics & Statistics | Lancaster University

Plan submitted for funding:

- 4 active doses (20, 40, 60 and 80mg)
- 2 interim analysis with O'Brien and Fleming type boundaries
- Method developed for this purpose (Magirr et al., 2012)

Mathematics & Statistics | Lancaster University

Design considerations

Mathematics & Statistics | Lancaster University

Plan submitted for funding:

- 4 active doses (20, 40, 60 and 80mg)
- 2 interim analysis with O'Brien and Fleming type boundaries
- Method developed for this purpose (Magirr et al., 2012)

Funded study:

- 3 active doses (20, 40 and 80mg)
- 1 interim analysis
- 370 patients to be recruited (336 evaluated needed)
- Funder was in general very happy with the design!

Mathematics & Statistics | Lancaster University

Design considerations

Mathematics & Statistics | Lancaster University

Plan submitted for funding:

- 4 active doses (20, 40, 60 and 80mg)
- 2 interim analysis with O'Brien and Fleming type boundaries
- Method developed for this purpose (Magirr et al., 2012)

Funded study:

- 3 active doses (20, 40 and 80mg)
- 1 interim analysis
- 370 patients to be recruited (336 evaluated needed)
- Funder was in general very happy with the design!

Lesson: Do not be afraid to propose an adaptive design to a funding agency

Mathematics & Statistics | Lancaster University

Interim analysis

Plan:

- Decision about stopping arms/study to be made by IDMC following pre-specified rules
- TMG to accept these recommendations

Interim analysis

Plan:

- Decision about stopping arms/study to be made by IDMC following pre-specified rules
- TMG to accept these recommendations

Reality:

- 2 arms recommended to be stopped
- TMG wanted to see unblinded data before confirming
- Lengthy discussions
  - Argued based on probability of success at study end for stopped arms is small

Interim analysis

Plan:

- Decision about stopping arms/study to be made by IDMC following pre-specified rules
- TMG to accept these recommendations

Reality:

- 2 arms recommended to be stopped
- TMG wanted to see unblinded data before confirming
- Lengthy discussions
  - Argued based on probability of success at study end for stopped arms is small

Lesson: Make sure TMG understands decision process and buys into the stopping rules.

A surprise

Plan:

- 10% patient drop-out

A surprise

Plan:

- 10% patient drop-out

Reality:

- 20% observed at interim (across all arms)
- Could adjust recruitment target from 370 to 377 patients

A surprise

Plan:

- 10% patient drop-out

Reality:

- 20% observed at interim (across all arms)
- Could adjust recruitment target from 370 to 377 patients

Lesson: Detailed data evaluation helped to prevent underpowered study.

A surprise

Mathematics & Statistics | Lancaster University

Plan:

- 10% patient drop-out

Reality:

- 20% observed at interim (across all arms)
- Could adjust recruitment target from 370 to 377 patients

Lesson: Detailed data evaluation helped to prevent underpowered study.

Lesson: An adaptive design does not always reduce sample size but here improved decision making.

Mathematics & Statistics | Lancaster University

A buddy system

- First multi-arm multi-stage design done by this CTU
- Worked closely with CTU statistician and provided oversight
  - e.g. CTU statistician drafted stat section for application, protocol, SAP. . . I commented/refined.
  - Strongly involved in communications around interim analysis
- In the meantime CTU has submitted at least 3 more multi-arm proposals for funding with limited involvment from us.

A buddy system

Mathematics & Statistics | Lancaster University

- First multi-arm multi-stage design done by this CTU
- Worked closely with CTU statistician and provided oversight
  - e.g. CTU statistician drafted stat section for application, protocol, SAP. . . I commented/refined.
  - Strongly involved in communications around interim analysis
- In the meantime CTU has submitted at least 3 more multi-arm proposals for funding with limited involvment from us.

Lesson: Buddy system can be very effective in training staff in novel methods.

Mathematics & Statistics | Lancaster University

A buddy system

Mathematics & Statistics | Lancaster University

- First multi-arm multi-stage design done by this CTU
- Worked closely with CTU statistician and provided oversight
  - e.g. CTU statistician drafted stat section for application, protocol, SAP. . . I commented/refined.
  - Strongly involved in communications around interim analysis
- In the meantime CTU has submitted at least 3 more multi-arm proposals for funding with limited involvment from us.

Lesson: Buddy system can be very effective in training staff in novel methods.

Mathematics & Statistics | Lancaster University

Status Quo of TAILoR

- Last patient last visit took place on Wednesday 29th June 2017
- no evidence for effect on primary endpoint on remaining dose.
- Some effect found in exploratory secondary analyses.

Status Quo of TAILoR

Mathematics & Statistics | Lancaster University

- Last patient last visit took place on Wednesday 29th June 2017
- no evidence for effect on primary endpoint on remaining dose.
- Some effect found in exploratory secondary analyses.

Lesson: An adaptive design does not prevent risk of over-interpretation of findings that have not been pre-specified.

Mathematics & Statistics | Lancaster University

Literature

Chow SC, Chang M, Pong, A (2005) Statistical consideration of adaptive methods in clinical development. *Journal of Biopharmaceutical Statistics*, **15**(4), 575–591.

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.

Pushpakom SP, Taylor C, Kolamunnage-Dona R, Spowart C, Vora J, Garcia-Finana M, Kemp GJ, Whitehead J, Jaki T, Khoo S, Williamson P. (2015) Telmisartan and insulin resistance in HIV (TAILoR): protocol for a dose-ranging phase II randomised open-labelled trial of telmisartan as a strategy for the reduction of insulin resistance in HIV-positive individuals on combination antiretroviral therapy. *BMJ open*. 5(10):e009566.

Wong CH, Kien WS, Andrew WL (2019) Estimation of clinical trial success rates and related parameters. *Biostatistics*, **20**(2):273–286.