MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

Adaptive Methods in Clinical Research

Lecture 9: How to design an Adaptive Trial

Pavel Mozgunov & Sofía S. Villar

MRC Biostatistics Unit

October-November 2023

How to use what you have learnt

- We have learnt about many advanced designs that can deal with many complex settings, but
- How do we know which to apply and when?
- When we know which, how do we make it work for **your particular trial**?

Planning for an Adaptive Trial

What is the research question?

- What is the objective of the trial?
- What is the patient population?
- What is the primary endpoint?
- What is after this trial?
- ...

Designs questions

- Distribution of the endpoint? [Binary/Ordinal/Continuous/Time-to-event...]
- Evaluation window and the recruitment rate?
- How many treatment arms? Why?
- Are treatments related to each other?
- Are we interested in all promising or just one?
- What control of type I error is of interest?
- What type of power is of interest?

Example: designing MAMS

Setting/Questions:

- Multi-arm study against the control;
- Two (or three?) experimental arms - a pilot study just completed
- Outcome the average of repeated measures over 12 weeks;

The **key goals** for the trial design:

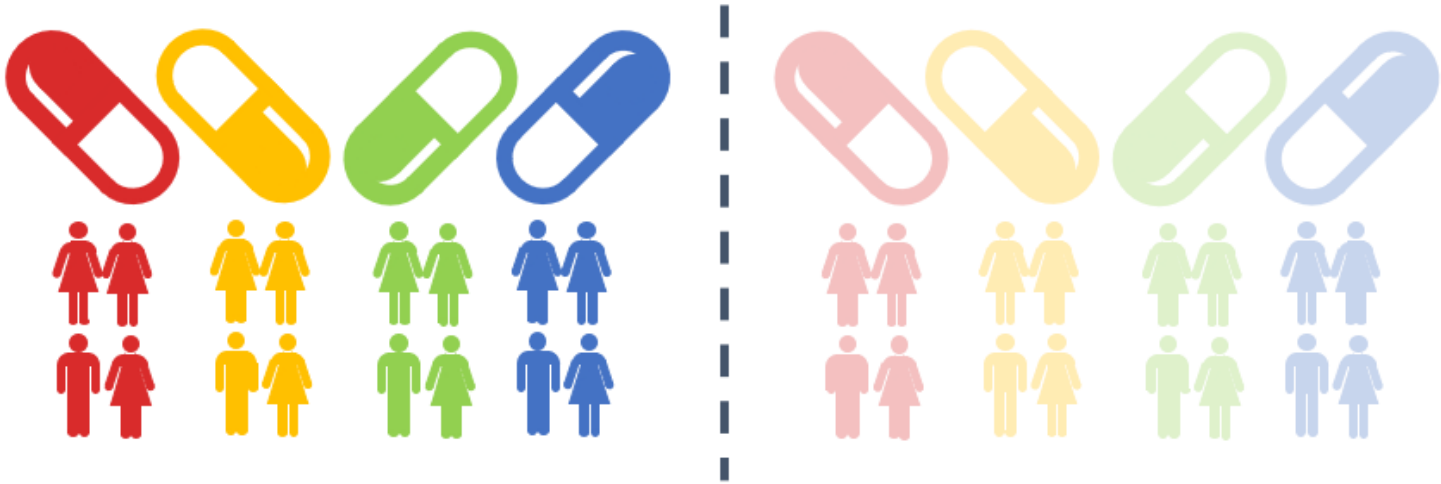
- Reduce the sample size,
- compare all treatments,
- removing ineffective treatments,
- making appropriate multiplicity adjustment.

The fixed sample design

+ Multiplicity adjustment possible
+ Reference design

– Cannot remove treatments during the trial
– Reference design

Multi-Arm Multi-Stage trial



Based on interim data possible decisions are:

- Futility - stop a treatment if it is not sufficiently effective,
- Efficacy - stop the trial if at least one treatment is effective.

\+ Reduction in expected sample size

\+ Dropping ineffective treatments

\+ Multiplicity adjustment

– Increase in potential maximum sample size

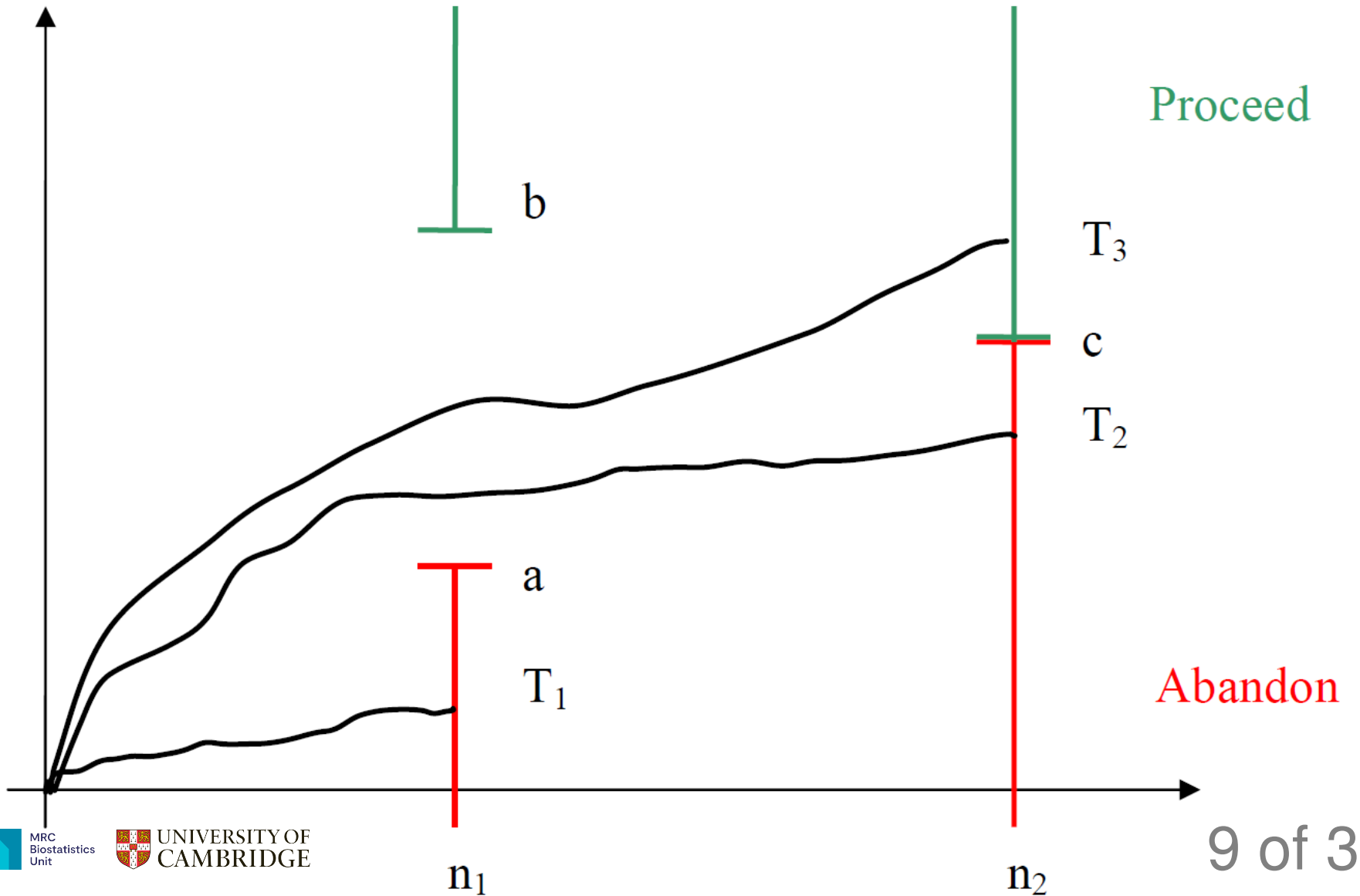
Drop the loser

Interim decision keep the best experimental treatment

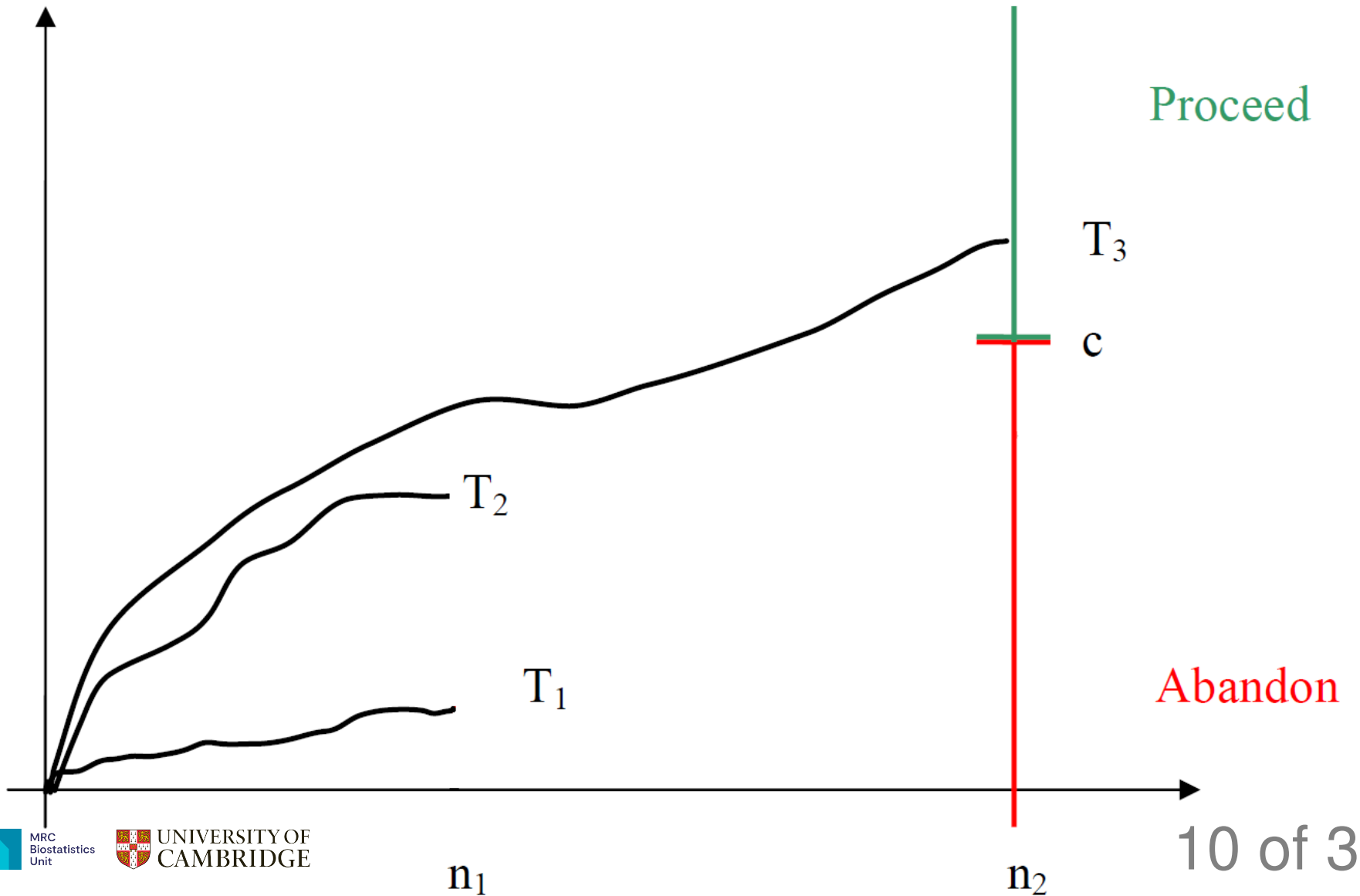
+ Guaranteed sample size
+ Reduction in sample size
+ Multiplicity adjustment

– Can only select one experimental treatment
– Treatments dropped regardless of efficacy
– Additional complexity

Comparing potential properties: Illustrating the difference in decision-making



Comparing potential properties: Illustrating the difference in decision-making



Comparing potential properties: Metrics

- Maximum sample size;
- Type I error and Power;
- Expected sample size (under different configurations);
- Possible sample sizes realisations;
- Probability to stop the trial earlier for futility/efficacy at the given stage;
- Expected duration of the trial;

Comparing potential properties

Under global null:

Distribution of sample sizes

Sample sizes	Frequency	Probability
204	274774	0.27
306	250122	0.25
357	246840	0.25
408	228264	0.23

Power to reject null for 1 treatments 0.01
Power to reject null for 2 treatments 0.00
Power to reject null for 3 treatments 0.00
Power to reject any null 0.05

Expected sample size 314

Under one good treatment:

Distribution of sample sizes

Sample sizes	Frequency	Probability
204	555102	0.56
306	229476	0.23
357	147084	0.15
408	68338	0.07

Power to reject null for 1 treatments 0.89
Power to reject null for 2 treatments 0.02
Power to reject null for 3 treatments 0.00
Power to reject any null 0.91

Expected sample size 264

Under two good treatments:

Distribution of sample sizes

Sample sizes	Frequency	Probability
204	715434	0.72
306	10247	0.01
357	209439	0.21
408	64880	0.06

Power to reject null for 1 treatments 0.41
Power to reject null for 2 treatments 0.55
Power to reject null for 3 treatments 0.01
Power to reject any null 0.97

Expected sample size 250

Under three good treatments:

Distribution of sample sizes

Sample sizes	Frequency	Probability
204	792119	0.79
306	1645	0.00
357	13488	0.01
408	192748	0.19

Power to reject null for 1 treatments 0.27
Power to reject null for 2 treatments 0.32
Power to reject null for 3 treatments 0.40
Power to reject any null 0.99

Expected sample size 246

The chosen design

Both adaptive designs offer:

- Reduction in sample size,
- incorporate all treatments,
- removing treatments during the trial,
- make appropriate multiplicity adjustment.

The choice between them came down to

MAMS

+ Dropping ineffective treatments.

DTL

+ Guaranteed sample size

The chosen design

MAMS was preferred and then the tweaks were explored:

- Number of interim analyses;
- Timing(s) of the interims;
- Futility only? Efficacy only? Both?
- Different futility/efficacy bounds.

Communicating Design to Funders/HAs

How do we fit something that complex in a little space provided?

Outline Application (aka Synopsis):

- Reference to the methodology;
- Maximum sample size;
- Range of expected sample sizes;
- Timing of the interim: after half of the maximum;
- Description of the decisions at the first stage;

161 words in total

Example: Feedback from the Panel

- The question on the primary endpoint - how does it link to the statistical analysis and sample size calculations;
- Please provide further details on the design;
- Timing of the interim and the impact of the overrun.

Full Application / Protocol

Title: Multi-Arm Multi-Stage Trial.

Summary: Brief rationale for an adaptive design.

"**Sample Size**" Section

- Maximum and minimum sample sizes & sample size required under various scenarios;
- Expected sample size (and number of simulations) accounting for waiting for the primary 12-weeks outcome to be evaluated;
- Projected timing of the interim (Month X of the trial);
- No pausing of the recruitment.

Full Application / Protocol

"**Analysis**" Section

- Adjusting covariates;
- Sample size per arm per stage;
- Critical boundaries that will be used to make the decision at the interim;
- Binding futility bound;
- Talking through the decisions that can be made at the interim.

617 words in total (for the adaptive features of the design)

Implementing an Adaptive Design

Conducting an Adaptive Trial

In theory, there is no difference between theory and practice. In practice, there is.

- Implementing (novel) adaptive designs requires close(r) collaborations between an interdisciplinary team.

  This includes: physicians, pharmacists, statisticians, data managers, health economits, sponsors and regulator.

- Close collaboration during trial conduct is a key component of any trial that includes adaptive decision rules.

- The process is involved but manageable and it is **essential** for implementing an efficient trial conduct.

Beyond the grant application and initial protocol

Your trial protocol has just been approved and funded, what next?

- The trial will be setting-up (almost always expect delays).
- Some aspects that will require additional statistical input are:
  (1) Trial's committees

  (2) Data management

  (3) Statistical Analysis Plan (SAP)

  (4) Randomization systems

  (5) Health Economics Analysis Plans

  (6) Consider publishing protocol and SAP Trials (journal link)

Data committees for ADs

- **Independent data monitoring boards** (iDMC) periodically review the interim data with an emphasis on safety, the overall conduct, progress toward the trial's goals, and efficacy.
- **Trial Sterering Commitee** (TSC) approves changes recommended by iDMC.
- **Trial Management Group** (TMG) represents others (sponsor, PI, wiedr team) should follow advice from TSC.
- The iDMC and TSC charters agreed upon between these committees and the trial team before the trial begins must clearly describe all pre-planned adaptations.
  Where blinding is possible, we avoid introducing bias by an unblinded iDMC making recommendations to a blinded TSC
- Input needed: recommend IDMC and TSC members with all the skills needed to implement the AD and revise charters

Example 1: role of committees

- NOTACS trial Dawson et al. [2022], Earwaker et al. [2022] preplanned SSR. At the start $n_{\text{min}} = 850$ and $n_{\text{max}} = 1152$
- After the SSR at interi, the iDMC recommended to the TSC a sample size increase to achieve 90%.

  TSC then requested a “protocol amendment” to reflect the change in sample size.
- Discussion with TSC was needed to agree that a prospective sample size increase and this **NOT** a protocol change.

  Lesson Protocol should have discussed how to report the final sample size to participating centres.

Example 1: role of committees (ctd)

- During any trial (adaptive or not), design revisions may be needed to accommodate for practical issues (or issues not considered at design stage).
- Communication between those designing and those conducting must be open to all and frequent.
- Example: NOTACS SSR, iDMC recommended to go beyond initial sample size cap $n_{\max} = 1152$ to achieve 90% power

  Lesson Protocol should have discussed explicitly the $n_{\max} = 1152$ was derived on a pragmatic rationale of recruitment restrictions (rather than statistical considerations). Discussions were needed to get TSC to agree.

Example 2: Experiences with committees

Plan:

- Multi-Arm Multi-Stage Trial (3 experimental, 2 stages)
- Decision about stopping arms/study to be made by by iDMC following pre-specified rules
- TMG to accept these recommendations

Reality:

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

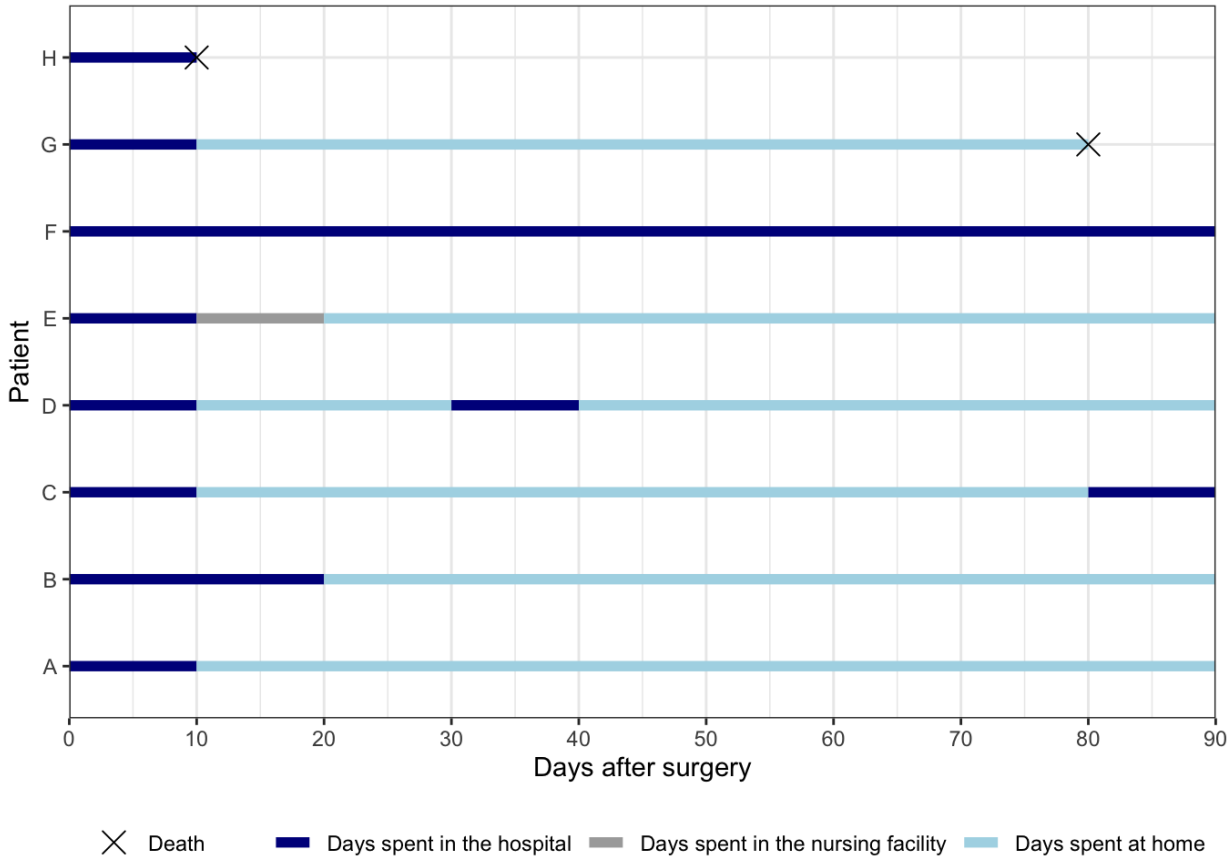
Lesson: Make sure not only iDMC and TSC but also TMG understands decision process and buys into the stopping rules.

Data management

- Validity of interim (and final) analysis crucially depends on the quality of the data management process (data entry, storage, verification, correction and retrieval).
- The data team will need input to understand what “additional” variables (e.g., in addition to safety/completion rates of primary endpoint) are needed and when.
- Example: a trial with a planned change of the randomisation ratio based on a early efficacy readout would need to have the highest possible standard of verification and cleaning for the interim reports.

Example: data management for SSR

NOTACS primary endpoint “Days alive at home at 90 days”



On top of the above data (mortality and location for 90 days), the SSR required data on compliance treatment switch rates.

Writing a full SAP

- The trial's statistician may need input for any other planned analysis to be included in a full SAP.

  For example, should the SAP plan to report bias-adjusted estimation? which ones? for which analysis? [Lecture 8]

- Consider writing a section on the the implementation and reporting of interim analysis.

- Recommend reporting of trial enrollment, allocation, and follow-up according to the adaptive designs CONSORT extension statement Dimairo et al. [2020]

- Similar questions can arise for the team wirting a Health Economics Analysis Plan (HEAP)

Randomisation systems

- For non adaptive trials, the choice of randomisation already requires input from a trial statistician. Berger et al. [2021]

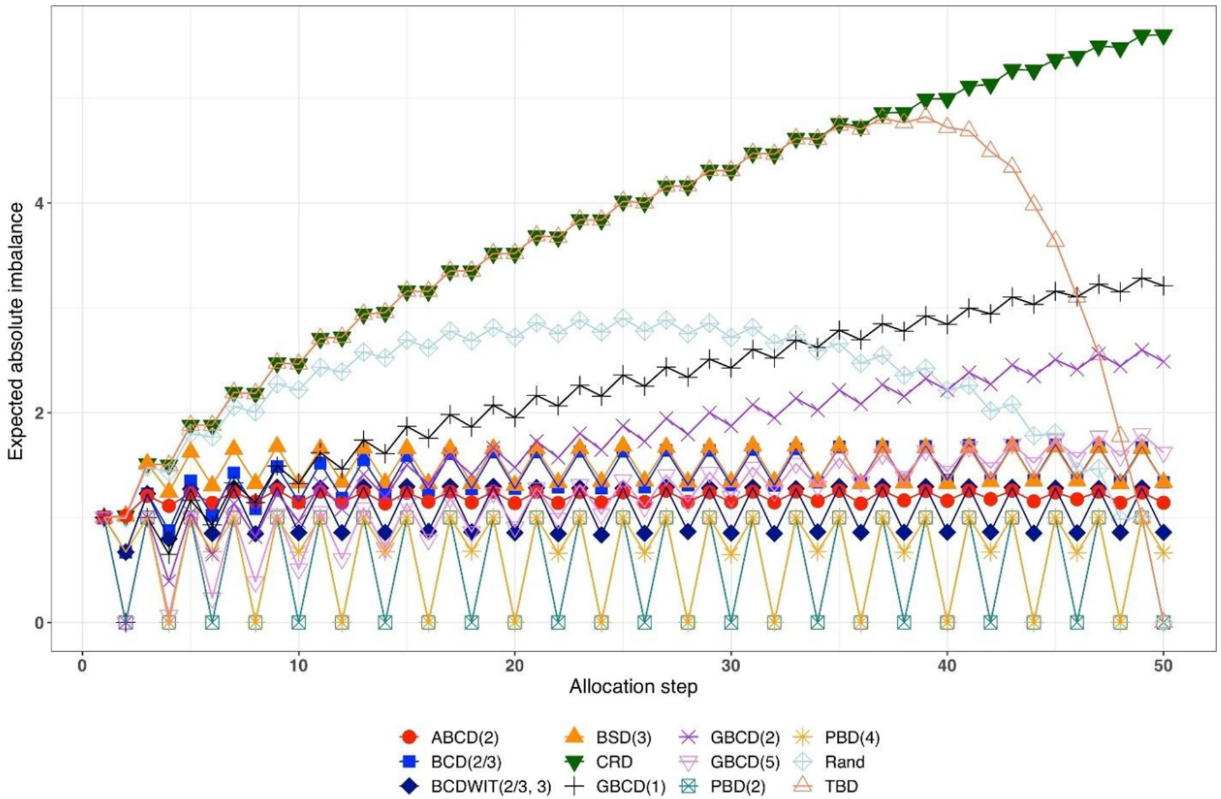
  Example: do we startify randomisation by centre? any other covariate? Do we use permuted block design?

  What level of predictability of allocations is acceptable?

  What level of sample size imabalance is acceptable?

Randomisation systems: non-adaptive RCT

From: A roadmap to using randomization in clinical trials



Simulated expected absolute imbalance vs. allocation step for 12 restricted randomization procedures for $n = 50$. Note: PBD(2) and PBD(4) have forced periodicity absolute imbalance of 0, which distinguishes them from MTI procedures

Back to article page >

MAMS: Randomisation after the interim

Examples of additional questions an adaptive designs brinngs in.

- Multi-Arm Multi-Stage Trial
- 3 experimental, 1 control; 2 stages;
- Decision making at the interim:
  - Stop early (for efficacy or futility)
  - Drop one/two/three arms
  - Continue as is
- Equal randomisation for the first stage;

Q: What should be the randomisation for the second stage?

We do not know. One can include all the possible options (four in this case) from the onset and then choose the one that we need.

$1:1:1:1$ ($K=3$), $1:1:1$ (for one of the two $K=2$) and $1:1$ ($K=1$)

RAR: Randomisation systems

- If the trial involves RAR (response-adaptive randomisation):

  Do those conducting he trial have the resources to frequently change the randomization system according to the design?

  If yes, work with them to develop a randomisation system that fits design. If not, you have to advice those outsourcing it.

- Example: if you use BRAR in a blcoked implemetation, what to do with discretness? Say you have alloc. prob. 0.65 for control and blocks of 10 then permuted block would assign 6 patients to the control group.

From Design to Conduct: final thoughts

~~*Life*~~ **An adaptive design** *can only be understood by looking backward; but it must be ~~lived~~* **designed** *looking forward —* ~~*Kierkegaard*~~ **ESD course at BSU**

- Factor in the level of experience with adaptive designs of those implementing your design when comparing design options (favor simplicity if possible).

  When past experience is low, there can be a tendency to underestimate resources needed for interim analysis.

- If you are leading the first adoption of an adaptive trial, bear in mind substantial implementation time can be saved from lessons learnt along the way.

References

V. W. Berger, L. J. Bour, K. Carter, J. J. Chipman, C. C. Everett, N. Heussen, C. Hewitt, R.-D. Hilgers, Y. A. Luo, J. Renteria, et al. A roadmap to using randomization in clinical trials. *BMC Medical Research Methodology*, 21: 1–24, 2021.

S. N. Dawson, Y.-D. Chiu, A. A. Klein, M. Earwaker, and S. S. Villar. Effect of high-flow nasal therapy on patient-centred outcomes in patients at high risk of postoperative pulmonary complications after cardiac surgery: a statistical analysis plan for notacs, a multicentre adaptive randomised controlled trial. *Trials*, 23(1):699, 2022.

M. Dimairo, P. Pallmann, J. Wason, S. Todd, T. Jaki, S. A. Julious, A. P. Mander, C. J. Weir, F. Koenig, M. K. Walton, et al. The adaptive designs consort extension (ace) statement: a checklist with explanation and elaboration guideline for reporting randomised trials that use an adaptive design. *bmj*, 369, 2020.

M. Earwaker, S. Villar, J. Fox-Rushby, M. Duckworth, S. Dawson, J. Steele, Y.-d. Chiu, E. Litton, G. Kunst, G. Murphy, et al. Effect of high-flow nasal therapy on patient-centred outcomes in patients at high risk of postoperative pulmonary complications after cardiac surgery: a study protocol for a multicentre adaptive randomised controlled trial. *Trials*, 23(1):1–18, 2022.