Adaptive and Complex Innovative Designs across trial phases for accelerated approval

Ayon Mukherjee, David Robertson, Sofia S. Villar, Thomas Burnett

Agenda

8.30 – 8.40	Welcome and introductions
8.40 – 8.45	Learning objectives
8.45 – 9.15	Introduction to Adaptive Designs
9.15 – 10.00	Bayesian methods for clinical trials
10.00 – 10.15	Break
10.15 – 11.00	Project Optimus and the framework of CID
11.00 – 11.45	(Response) Adaptive Randomization
11.45 – 12.00	Recap of learning objectives and Q&A

Course learning objectives

- Understand the potential benefits and drawbacks of using adaptive designs.
- Learn about the breadth and depth of the toolbox offered by adaptive designs.
- Know where to search for more information on adaptive designs.
- Understand the difference between using Bayesian methods for design and/or inference.
- Learn about the current regulatory environment regarding Bayesian methods.
- See real-life examples of using Bayesian methods for clinical trials.
- Learn about key aspects of complex innovative designs and their practical implementation.
- Understand the operation of different dose-finding methodologies and how they differ.
- Learn about key aspects of Project Optimus and the importance of Bayesian adaptive designs in early phase dose escalation/optimization trials.
- Understand the relation between design, implementation and analysis aspects of complex adaptive designs (through an example using Bayesian response-adaptive randomization)

Lecture slides

`https://github.com/dsrobertson/RISW_SC01`

Q&As

https://app.sli.do/event/6u2ri9P3fbviNBTx84Utu8

Lecture 1: Introduction to Adaptive Designs

Thomas Burnett

tb292@bath.ac.uk

Department of Mathematical Sciences

Outline

Outline

Adaptive Designs

How can we use adaptive designs?

Error control and estimation

Master protocols

Bayesian methods

Motivation

Attrition rates for new developments (Arrowsmith 2011a, 2011b)

- phase II: $>80\%$
- phase III & submission: $\sim 50\%$

Reasons for failure (Arrowsmith & Miller 2013)

Phase II (2011–2012)

Phase III & submission (2011–2012)

Motivation

Attrition rates for new developments (Arrowsmith 2011a, 2011b)

- phase II: $>80\%$
- phase III & submission: $\sim 50\%$

Reasons for failure (Arrowsmith & Miller 2013)

Phase II (2011–2012)

Phase III & submission (2011–2012)

Motivation

Attrition rates for new developments (Arrowsmith 2011a, 2011b)

- phase II: $>80\%$
- phase III & submission: $\sim 50\%$

Reasons for failure (Arrowsmith & Miller 2013)

Phase II (2011–2012)

Phase III & submission (2011–2012)

Problems and solutions

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?

- avoid going straight into large and expensive phase III
- take more care during phases I and II
- consider adaptive and Bayesian designs

Idea

Modify an ongoing trial

by design or ad hoc

based on reviewing accrued data at interim

to enhance flexibility

without undermining the study's integrity and validity.

(Chow et al. 2005)

Idea

Modify an ongoing trial

by design or ad hoc

based on reviewing accrued data at interim

to enhance flexibility

without undermining the study's integrity and validity.

(Chow et al. 2005)

Idea

Modify an ongoing trial

by design or ad hoc

based on reviewing accrued data at interim

to enhance flexibility

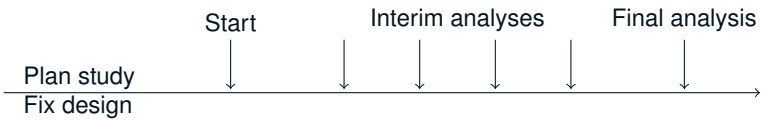
without undermining the study's integrity and validity.

(Chow et al. 2005)

Fixed sample design

- **total** sample size known in advance
- no adjustment possible

Adaptive design



- larger **maximum** sample size
- lower **expected** sample size

At each interim:

- decide whether or not to stop
- change sample size
- drop or add a dose
- change the endpoint
- change the question

Pros and cons

+ highly flexible
+ very efficient
+ reflects medical practice
+ shorter trial and/or more accurate estimates
+ ethical

– highly flexible
– inefficient
– time-consuming to design
– post-trial estimation difficult
– simple estimates may be biased
– interim analyses may require unblinding

Outline

Adaptive Designs

How can we use adaptive designs?

Error control and estimation

Master protocols

Bayesian methods

What, which, who and does?

With so many different adaptive designs it is useful to categorise them by the questions they may help us answer.

Burnett, T., Mozgunov, P., Pallmann, P. et al. Adding flexibility to clinical trial designs: an example-based guide to the practical use of adaptive designs. BMC Med 18, 352 (2020). https://doi.org/10.1186/s12916-020-01808-2

PANDA (A Practical Adaptive & Novel Designs and Analysis toolkit)

`https://panda.shef.ac.uk/`

What is a safe dose?

Designs

Continual Reassessment Method (CRM)

An iterative process using all available trial data when choosing the dose for the next cohort of patients.

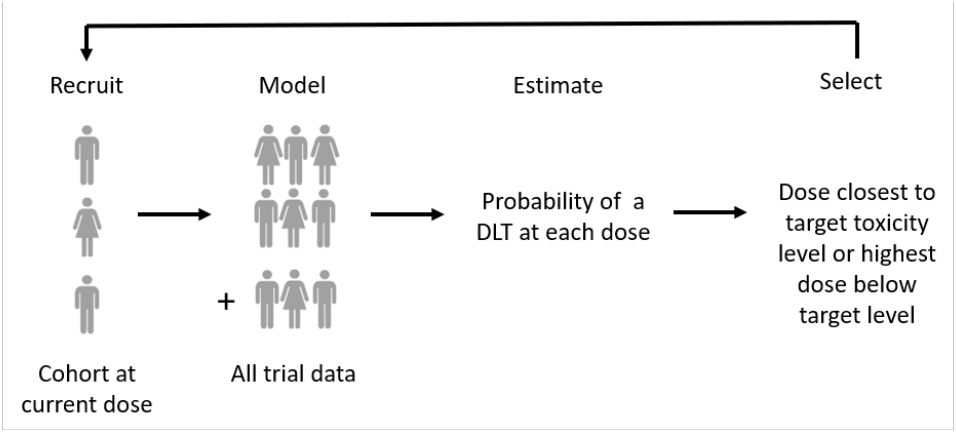
Escalation with overdose control (EWOC)

Skews the allocation criteria of CRM in favour of under-dosing patients rather than overdosing.

Bayesian Optimal Interval Design (BOIN)

Model assisted decision making for dose escalation.

CRM


Recruit
Model
Estimate
Select
Probability of a DLT at each dose
Dose closest to target toxicity level or highest dose below target level
+
Cohort at current dose
All trial data

The potential benefits

+ Decisions based on all trial data rather only the previous cohort.
+ Pre-specified target toxicity level.
+ Higher probability of selecting the dose nearest the maximum tolerated dose.
+ EWOC can reduce the chance of treating at excessively toxic doses.

– Additional complexity in designing the trial.
– Requires input from a suitably trained statistician.
– Can be seen as a black box.

Which is the best treatment?

Designs

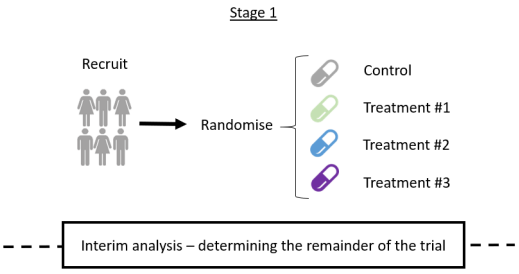
Multi-Arm Multi-Stage (MAMS)

Simultaneous comparison of multiple experimental treatment arms with a common control, allowing for early stopping for futility/efficacy.

Drop the loser (DTL)

Similar to MAMS with a fixed number of treatments dropped at each stage.

MAMS


Stage 1
Recruit
Randomise
Control
Treatment #1
Treatment #2
Treatment #3
Interim analysis – determining the remainder of the trial
Stage 2
Recruit
Randomise
Control
Treatment #1
Treatment #2
Treatment #3

Designs

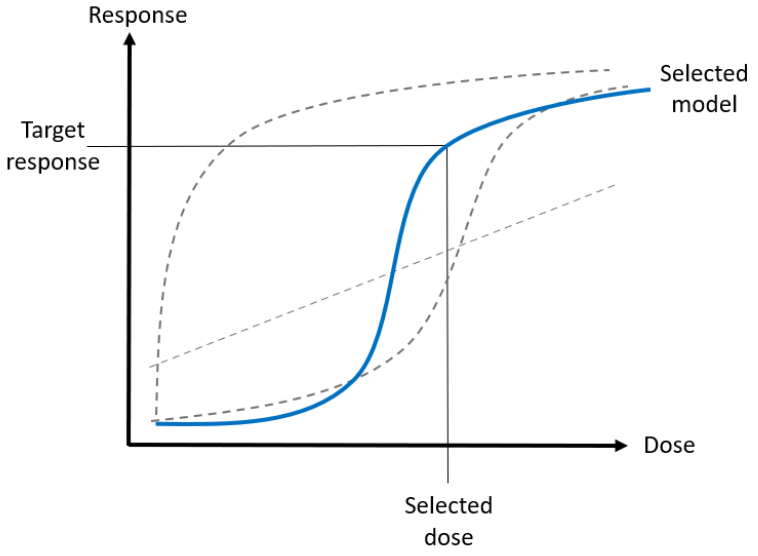
Response-adaptive randomization (RAR)

Modify the allocation probability based on accumulating data, for example based on observed efficacy of the treatment.

Multiple comparison procedures and modelling approaches (MCP-Mod)

Compares several dose-response models to detect whether there is a dose-response signal.

MCP-Mod


Response
Target response
Selected model
Dose
Selected dose

The potential benefits

+ Common control groups reduce the number of control patients.
+ Lower expected sample size.
+ Dropping of treatments/doses that do not benefit patients.
+ Increase the proportion of patient benefiting from treatment.
+ Efficient use of trial data through modelling.

– Additional complexity in designing the trial.
– Requires input from a suitably trained statistician.
– Uncertainty about the required number of patients and other logistical challenges.
– DTL can drop effective treatments.
– Require a quickly observed endpoint.
– MCP-Mod can be sensitive to model assumptions.

Who will benefit?

Designs

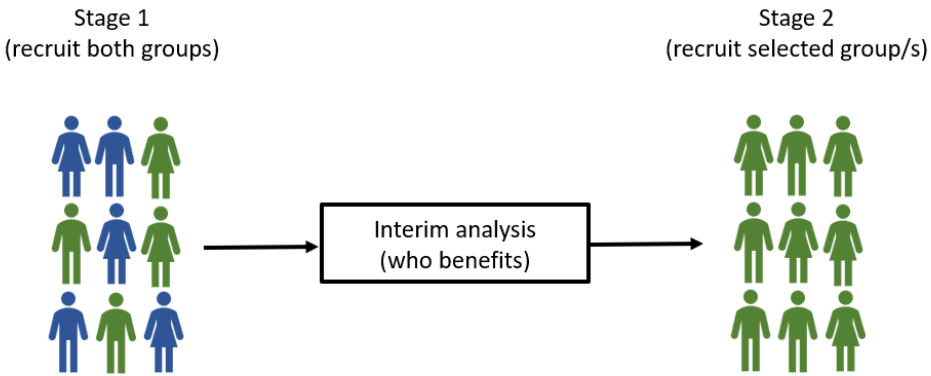
Covariate adjusted response adaptive

A form of RAR where randomization probabilities are skewed based on patients' observed bio-marker information.

Adaptive enrichment

Select bio-marker defined sub-groups to recruit from based on observed trial data at pre-planned interim analysis.

Adaptive Enrichment



The potential benefits

+ Introduce balance of efficiency and ethical goals.
+ Recruit fewer patients from patient groups that do not benefit.
+ Offers efficient decision making when uncertainty about who will benefit.
+ Compromise between fixed sample alternatives while maintaining statistical operating characteristics.

– Additional complexity in designing the trial.
– Requires input from a suitably trained statistician.
– Changing eligibility criteria.
– Sensitive to bio-marker definition.
– Computationally intensive.
– Can be inefficient when unnecessarily.

Does the treatment work?

Designs

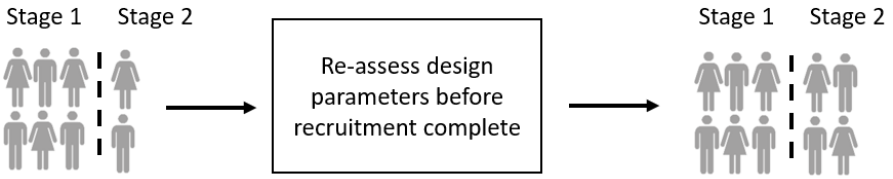
Group sequential designs

Enable early stopping of the trial while maintaining control of the type I error rate using pre-defined stopping rules.

Sample size re-assessment

Target an appropriate sample size for the trial by reassessing uncertainty about key design parameters.

Sample size re-assessment



The potential benefits

+ Reduce the expected sample size.
+ More common in practice.
+ Blinded and unblinded re-assessment possible.
+ Ensure a suitable number of patients are recruited.

– Additional complexity in designing the trial.
– Requires input from a suitably trained statistician.
– Potential increase the in the maximum sample size.
– Uncertainty about when the trial will end.
– Relatively few drawbacks to sample size re-assessment other than the practical issues.

Outline

Adaptive Designs

How can we use adaptive designs?

Error control and estimation

Master protocols

Bayesian methods

Motivation

In a late phase setting we may wish to formally test hypotheses formed about the treatments.

Through their use of interim analysis for decision making adaptive designs present several problems to consider when error control is desired.

Furthermore one may wish to consider this when performing inference, either through understanding the impact on the estimates or through adjusting the estimates accordingly.

Multiple analyses

What are combination tests?

Combination tests allow us to combine multiple trial stages.

They allow us to combine multiple stages of the trial under the assumption of conditional independence.

They yield a single test statistic.

Weighted Inverse Normal

The weighted inverse normal gives a combined p-value $P^{(c)}$ across the stages of the trial.

We find $P^{(c)}$ as follows.

$$Z^{(1)} = \Phi^{-1}(1 - P^{(1)}) \text{ and } Z^{(2)} = \Phi^{-1}(1 - P^{(2)}).$$

With pre-defined w_1 and w_2 such that $w_1^2 + w_2^2 = 1$,

$$Z^{(c)} = w_1 Z^{(1)} + w_2 Z^{(2)}$$

$$P^{(c)} = 1 - \Phi(Z^{(c)}).$$

The conditional error principle

An alternative construction to combination tests.

Given the observations from the Phase II stage of the trial, the conditional error is defined as

$$\alpha' = \mathbb{P}_{\theta=0}(\text{Reject } H_0 | Z^{(1)})$$

The probability of falsely rejecting the null hypothesis for the remainder of the trial must not exceed α'.

Discussion

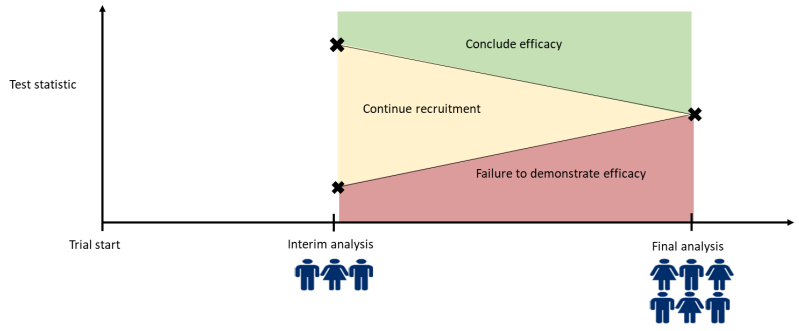
Combination tests and the conditional error principle are general approaches to combining data across a trial.

They can both guarantee error control at the pre-determined level whatever decision is made.

This allows complete freedom in how decisions are made.

Group sequential designs

Group-sequential stopping rules.



+ Stop early for efficacy
+ Drop ineffective treatments
+ Reduce the expected sample size

– higher maximum sample size
– more complex for design and analysis

Multiple hypotheses

Multiple treatments

Suppose in Phase II we wish to explore multiple experimental treatments.

For experimental treatment t we are interested in the treatment effect

$$\theta_t = \mu_t - \mu_0.$$

Hence the corresponding hypotheses are

$$H_{0,t} : \theta_t \leq 0 \text{ vs } H_{1,t} : \theta_t > 0.$$

Error control

There are several considerations when it comes to the error rate now we have multiple hypotheses.

Commonly in the confirmatory setting we require strong control of the FWER

$$\mathbb{P}_{\boldsymbol{\theta}}(\text{Reject any combination of true null hypotheses}) \leq \alpha \text{ for all } \boldsymbol{\theta}.$$

A Bonferroni correction is a well known method for adjusting for multiplicity.

Closed testing

Suppose we have two experimental treatments. We define a closed testing procedure as follows:

- $P_1^{(c)}$ for testing $H_{01} : \theta_1 \leq 0$,
- $P_2^{(c)}$ for testing $H_{02} : \theta_2 \leq 0$,
- $P_{12}^{(c)}$ for testing $H_{01} \cap H_{02} : \theta_1 \leq 0$ and $\theta_2 \leq 0$.

Then:

- Reject H_{01} globally if $P_1^{(c)} \leq \alpha$ and $P_{12}^{(c)} \leq \alpha$.
- Reject H_{02} globally if $P_2^{(c)} \leq \alpha$ and $P_{12}^{(c)} \leq \alpha$.

Estimation

Point estimates

One must consider the potential impacts of adaptation on estimation.

When adaptations are made one must consider the impact on estimation.

Robertson DS, Choodari-Oskooei B, Dimairo M, Flight L, Pallmann P, Jaki T. Point estimation for adaptive trial designs I: A methodological review. Statistics in Medicine. 2023; 42(2): 122-145. doi:10.1002/sim.9605

Robertson DS, Choodari-Oskooei B, Dimairo M, Flight L, Pallmann P, Jaki T. Point estimation for adaptive trial designs II: Practical considerations and guidance. Statistics in Medicine. 2023; 42(14): 2496-2520. doi: 10.1002/sim.9734

Confidence intervals

Similarly if confidence intervals are desired on must take care in their construction.

Does one desire repeated confidence intervals for each stage of the trial.

What other criteria are important?

Work in progress from the previously mentioned group will help answer these questions.

Regulatory guidance

U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER). Adaptive Designs for Clinical Trials of Drugs and Biologics Guidance for Industry. 2019.

ICH E20 is a work in progress that will further guide the implementation of adaptive designs.

Outline

What is a Master Protocol?

Platform

Highly adaptable trials to evaluate multiple therapies within a single design. A well-known example of this is the RECOVERY trial.

Umbrella

Multiple biomarker-defined patient subgroups paired with specific treatments.

Basket

Examine a single experimental treatment in multiple biomarker-defined subgroups

Outline

How are Bayesian methods used?

Bayesian methods represent a fundamental shift from the typically used Frequentist framework. The key difference is that the parameters such as the treatment effect are treated as random variables.

Analysis

One can plan a primary analysis using exclusively Bayesian methods.

Design

Alternatively one may construct a Frequentist analysis framework in such a way as to ensure control of the error rate of the trial while allowing complete freedom in how adaptive design decisions are made.