# Adaptive Methods in Clinical Research

*Lecture 1: Single arm binary outcome designs*

Martin Law

MRC Biostatistics Unit

October 30, 2023

# A dose of reality

Ben Goldacre, Guardian 1-09-08:

- **Before 1935**: doctors were basically useless

# A dose of reality

Ben Goldacre, Guardian 1-09-08:

- **Before 1935**: doctors were basically useless
- **1935-1995**: antibiotics, dialysis, transplants, intensive-care units, heart surgery, every drug you've ever heard of

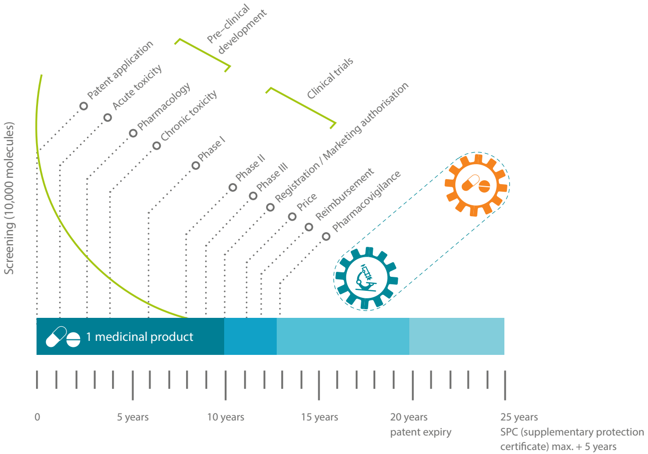
# A dose of reality

Ben Goldacre, Guardian 1-09-08:

- **Before 1935**: doctors were basically useless
- **1935-1995**: antibiotics, dialysis, transplants, intensive-care units, heart surgery, every drug you've ever heard of
- **1995-now**: the low-hanging fruit of medical research has all been harvested, and the industry is rapidly running out of new drugs

# The development process

PHASES OF THE RESEARCH AND DEVELOPMENT PROCESS



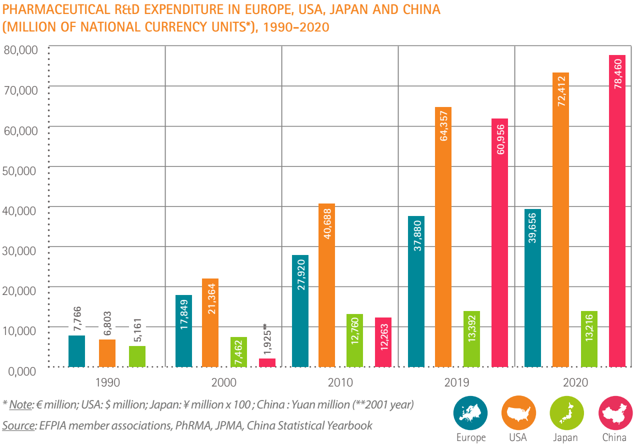

Source: European Federation of Pharmaceutical Industries and Associations (2022)

# Drug development

Development of a novel medicinal product

- takes 10-15 years
- costs several hundred million euros on average
  - largest contributors are confirmatory (phase III) trials
  - often involve thousands of patients with follow-up period frequently lasting years

# Cost on R&D in Pharmaceutical industry

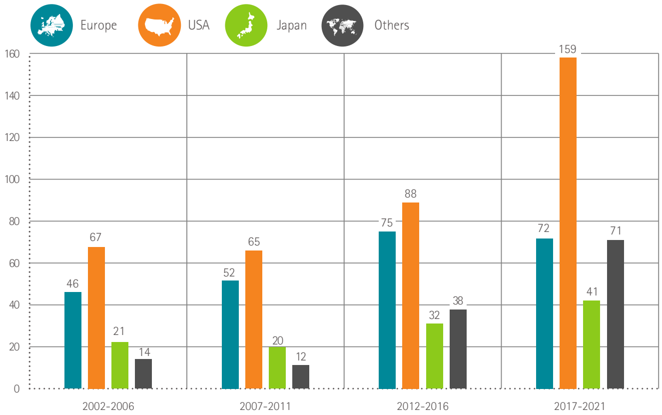



Source: European Federation of Pharmaceutical Industries and Associations (2022)

MRC Biostatistics Unit
UNIVERSITY OF CAMBRIDGE

# New molecular entities

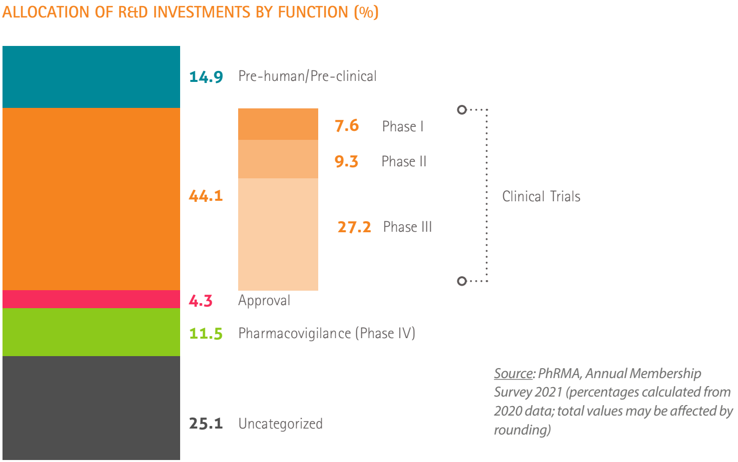
NUMBER OF NEW CHEMICAL AND BIOLOGICAL ENTITIES (2002-2021)




*Source: SCRIP – EFPIA calculations (according to nationality of mother company)*

Source: European Federation of Pharmaceutical Industries and Associations (2022)

# Cost on R&D in Pharmaceutical industry




Source: European Federation of Pharmaceutical Industries and Associations (2022)

# Success rates

According to a recent review (Wong, Siah & Lo, Biostatistics, 2019), between 2000 and 2015

- **41.0%** of confirmatory clinical trials overall and
- **64.5%** of confirmatory clinical trials in oncology

have been unsuccessful.

# Success rates

According to a recent review (Wong, Siah & Lo, Biostatistics, 2019), between 2000 and 2015

- **41.0%** of confirmatory clinical trials overall and
- **64.5%** of confirmatory clinical trials in oncology

have been unsuccessful.

- **13.4%** of treatments entering Phase I receive approval
- In oncology only **3.4%** of treatments entering Phase I receive approval

# Consequences

- Avoid going straight into large and expensive phase III trials
- Take more care during phases I and II
- Explore the potential of “new” statistical methods:
  - Sequential designs
  - Adaptive designs
  - Bayesian methods

What is an adaptive design?

What is an adaptive design?

A clinical trial design that can react (or adapt!) based on the trial data so far. This adaptation should (ideally) be specified in advance, both in terms of the changes we make and when we look at the incoming data.

What is an adaptive design?

A clinical trial design that can react (or adapt!) based on the trial data so far. This adaptation should (ideally) be specified in advance, both in terms of the changes we make and when we look at the incoming data.

Possible adaptations:

- Stop the trial early when we observe a very strong (or weak) treatment effect;
- Re-estimate the required sample size;
- Stop recruitment to a poorly-performing treatment.

# Why consider adaptive designs?

Different benefits for different forms of adaptation, including:

- Stopping early: fewer participants required on average
- Sample size re-estimation: more likely to reach required power
- Drop “loser” treatment: decrease the proportion of participants receiving poorly-performing treatment

In clinical trials, data accumulates steadily over time $\rightarrow$ natural to sequentially monitor results and perform interim analyses, for a number of reasons:

- **Ethical**: Minimise participants exposure to unsafe/ineffective treatments
- **Economic**: Allow *early stopping* $\rightarrow$ fewer patients needed on average.
- **Administrative**: Ensure trial is being run as planned.

# Interim analysis

In clinical trials, data accumulates steadily over time $\rightarrow$ natural to sequentially monitor results and perform interim analyses, for a number of reasons:

- **Ethical**: Minimise participants exposure to unsafe/ineffective treatments
- **Economic**: Allow *early stopping* $\rightarrow$ fewer patients needed on average.
- **Administrative**: Ensure trial is being run as planned.

Continuous monitoring is often impractical $\rightarrow$ examine the data at periodic intervals.

# Single-arm binary outcome trials

In oncology, a single-arm binary outcome trial (response/no response) often takes place after dose finding (Phase I).

Research question: Is the response rate $p$ of the selected dose large enough to continue development?

We test $H_0 : p \leq p_0$ against $H_1 : p > p_0$. In general:

We test $H_0 : p \leq p_0$ against $H_1 : p > p_0$. In general:

**Type-I error-rate**: $P(\text{reject } H_0 | H_0 \text{ is true})$

**Power**: $P(\text{reject } H_0 | H_1 \text{ is true})$

We wish to control the type-I error-rate to be $\alpha$ when $p = p_0$ and power our trial to a level $1 - \beta$ under $p = p_1$.

$p_0$: the greatest response rate that we deem typical for standard of care.

$p_1$: the smallest response rate that is large enough to warrant further study.

# Single-stage design (A'Hern 2001)

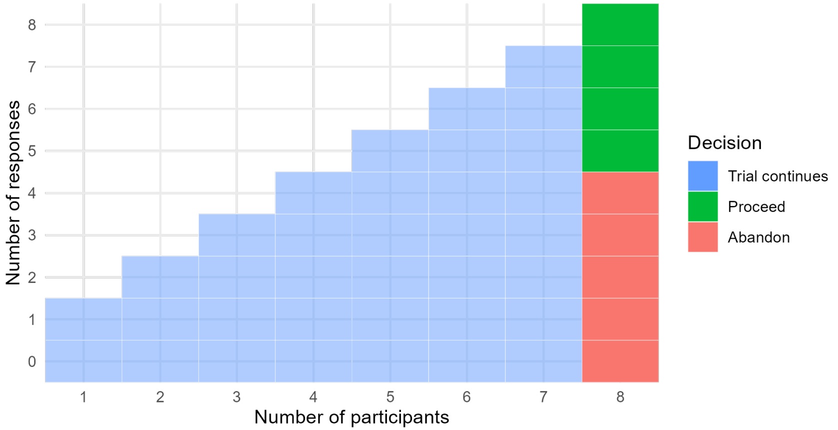
The most simple single-arm binary outcome design is the single-stage design:

Recruit $n_{max}$ participants. A go decision is made (i.e. the trial is deemed a success) if the final number of responses $S_{n_{max}}$ exceeds a specified boundary $r$ (i.e. $S_{n_{max}} > r$).

We choose a set of design parameters that satisfy specified type-I error-rate and power requirements for $p_0$ and $p_1$.

MRC Biostatistics Unit UNIVERSITY OF CAMBRIDGE

# Single-stage design (A'Hern 2001)




MRC Biostatistics Unit
UNIVERSITY OF CAMBRIDGE

## Simon design (Simon 1989)

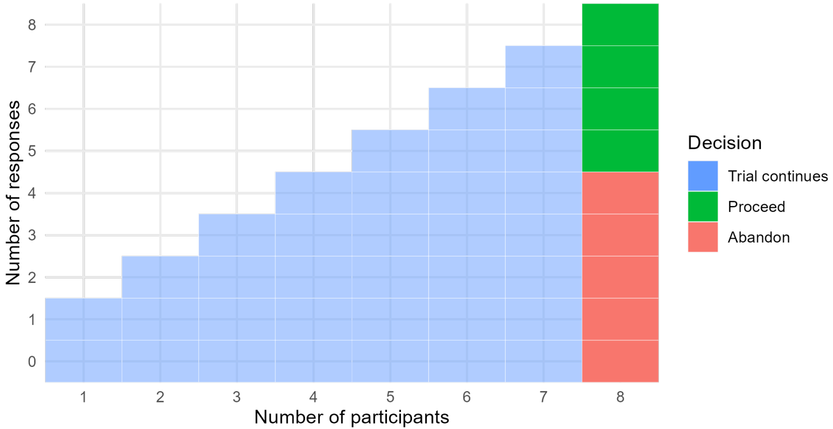
Simon design is a simple adaptation to the single-stage design:

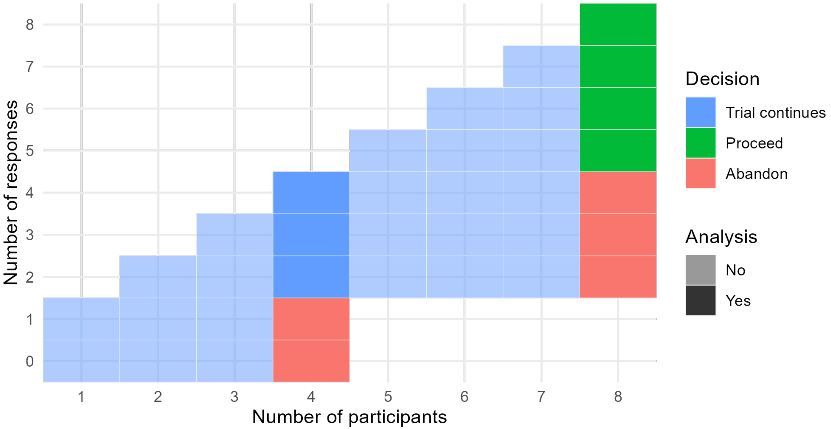
Include single interim analysis after $n_1$ participants.

If $S_{n_1} \leq r_1$, stop for lack of benefit.

Otherwise, recruit another $n_{max} - n_1$ participants and make a go/no-go (success/failure) decision at $n_{max}$.

# Single-stage design




Number of responses
8
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
8
Number of participants
Decision
Trial continues
Proceed
Abandon
Analysis
No
Yes

# Simon design

Simon designs can be obtained using the command `ph2simon` in the `clinfun` package. The command will return at least two designs:

- Minimax, which minimises the maximum sample size
- Optimal, which minimises the expected sample size under the null hypothesis ($ESS(H_0)$).

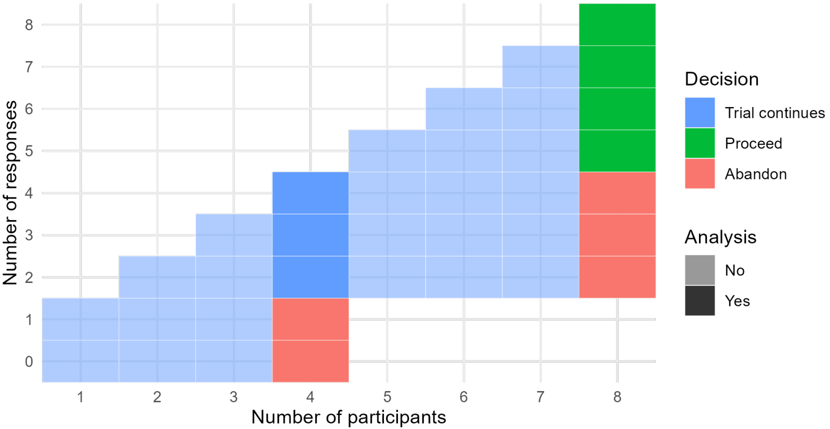
```
##
##  Simon 2-stage Phase II design
##
## Unacceptable response rate:  0.2
## Desirable response rate:  0.4
## Error rates: alpha =  0.05 ; beta =  0.1
##
##             r1 n1  r  n EN(p0) PET(p0)    qLo   qHi
## Minimax      5 24 13 45  31.23  0.6559 0.108 1.000
## Admissible   4 20 14 49  30.74  0.6296 0.058 0.108
## Optimal      4 19 15 54  30.43  0.6733 0.000 0.058
```

$PET(p_0)$: Probability of Early Termination

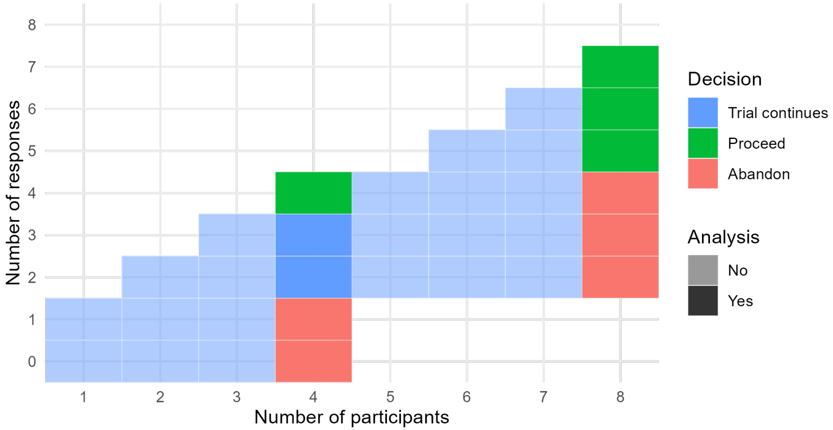
MRC Biostatistics Unit
UNIVERSITY OF CAMBRIDGE

# Mander and Thompson design (2010)

An extension to the Simon design, where the trial may additionally end early for a go decision after $n_1$ participants if $S_{n_1} \geq e_1$.


Number of responses
0
1
2
3
4
5
6
7
8
Number of participants
1
2
3
4
5
6
7
8
Decision
Trial continues
Proceed
Abandon
Analysis
No
Yes

# Mander and Thompson design




MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

# Obtaining design parameters

For these binary outcome designs, it is possible to calculate the probability of rejecting $H_0$ using the set of design parameters, e.g. $\{n_1, n_{max}, r_1, r\}$, conditional on a response rate.

As such, we do not directly calculate stopping boundaries. Instead, we can calculate type-I error-rate as $P(\text{reject } H_0 | p = p_0)$ and power as $P(\text{reject } H_0 | p = p_1)$.

With this in mind, we can calculate $\alpha$ and $1 - \beta$ for all possible combinations of $\{n_1, n_{max}, r_1, r\}$ to obtain a suitable set of design parameters for our choice of $p_0, p_1, \alpha$ and $1 - \beta$.

# Curtailment in trial design

## Non-stochastic curtailment

# Curtailment in trial design

### Non-stochastic curtailment

Stop when either success or failure is *certain*.

# Curtailment in trial design

### Non-stochastic curtailment

Stop when either success or failure is *certain*.

### Stochastic curtailment

MRC Biostatistics Unit
UNIVERSITY OF CAMBRIDGE

# Curtailment in trial design

## Non-stochastic curtailment

Stop when either success or failure is *certain*.

## Stochastic curtailment

Stop when either success or failure is *very likely*.

# “Very likely”?

# “Very likely”?

How do we define “very likely”?

How do we define “very likely”?

By using conditional power:

# Conditional power (CP)

# Conditional power (CP)

Conditional power is the current probability of trial success, conditional on some response rate $p$. This can be interpreted in different ways:

# Conditional power (CP)

Conditional power is the current probability of trial success, conditional on some response rate $p$. This can be interpreted in different ways:

- $P(\text{reject } H_0 \text{ at } n_{max} | p = p_1)$
- $P(\text{reject } H_0 \text{ at } n_{max} | p = \hat{p})$
- $P(\text{reject } H_0 \text{ at any point} | p = p_1)$

# Conditional power (CP)

Conditional power is the current probability of trial success, conditional on some response rate $p$. This can be interpreted in different ways:

- $P(\text{reject } H_0 \text{ at } n_{max} | p = p_1)$
- $P(\text{reject } H_0 \text{ at } n_{max} | p = \hat{p})$
- $P(\text{reject } H_0 \text{ at any point} | p = p_1)$

Stop if success is certain: CP=1 (NSC)

# Conditional power (CP)

Conditional power is the current probability of trial success, conditional on some response rate $p$. This can be interpreted in different ways:

- $P(\text{reject } H_0 \text{ at } n_{max} | p = p_1)$
- $P(\text{reject } H_0 \text{ at } n_{max} | p = \hat{p})$
- $P(\text{reject } H_0 \text{ at any point} | p = p_1)$

Stop if success is certain: CP=1 (NSC)

Stop if failure is certain: CP=0 (NSC)

# Conditional power (CP)

Conditional power is the current probability of trial success, conditional on some response rate $p$. This can be interpreted in different ways:

- $P(\text{reject } H_0 \text{ at } n_{max} | p = p_1)$
- $P(\text{reject } H_0 \text{ at } n_{max} | p = \hat{p})$
- $P(\text{reject } H_0 \text{ at any point} | p = p_1)$

Stop if success is certain: CP=1 (NSC)

Stop if failure is certain: CP=0 (NSC)

Stop if success is very likely: $CP > \theta_E$ (SC)

# Conditional power (CP)

Conditional power is the current probability of trial success, conditional on some response rate $p$. This can be interpreted in different ways:

- $P(\text{reject } H_0 \text{ at } n_{max} | p = p_1)$
- $P(\text{reject } H_0 \text{ at } n_{max} | p = \hat{p})$
- $P(\text{reject } H_0 \text{ at any point} | p = p_1)$

Stop if success is certain: CP=1 (NSC)

Stop if failure is certain: CP=0 (NSC)

Stop if success is very likely: $CP > \theta_E$ (SC)

Stop if failure is very likely: $CP < \theta_F$ (SC)

# Conditional power (CP)

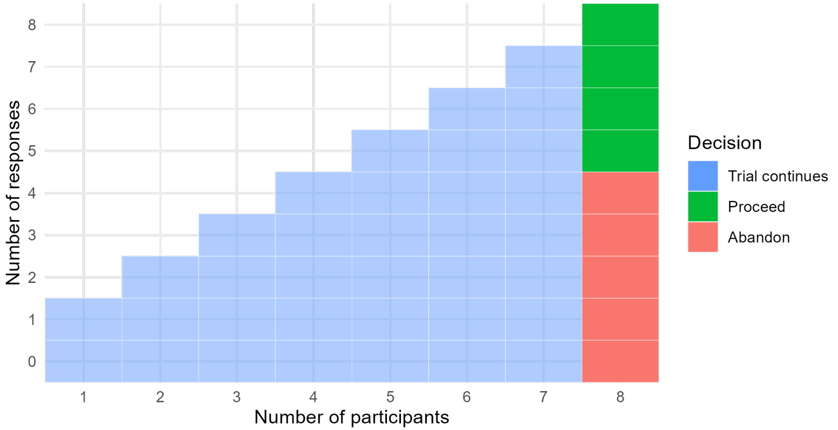
By planning to end a trial early due to a high or low CP, we can reduce that trial's expected sample size.

# Conditional power (CP)

By planning to end a trial early due to a high or low CP, we can reduce that trial's expected sample size.

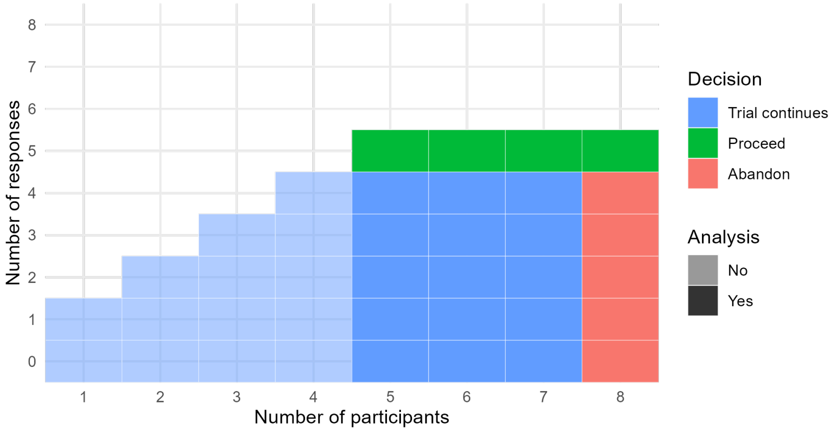
R package for Mander and Thompson design, non-stochastic curtailment and non-stochastic curtailment designs: `curtailment`. Latest version on github (though also on CRAN), link in References.

# Add NSC to single-stage design

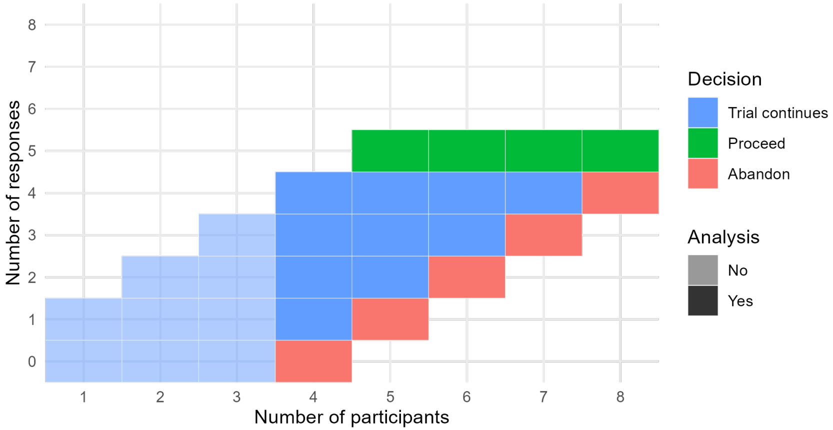



MRC Biostatistics Unit
UNIVERSITY OF CAMBRIDGE

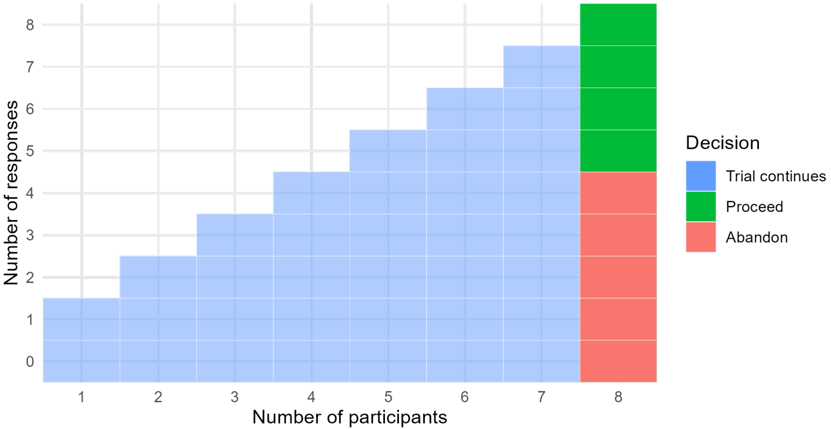
# Add NSC to single-stage design



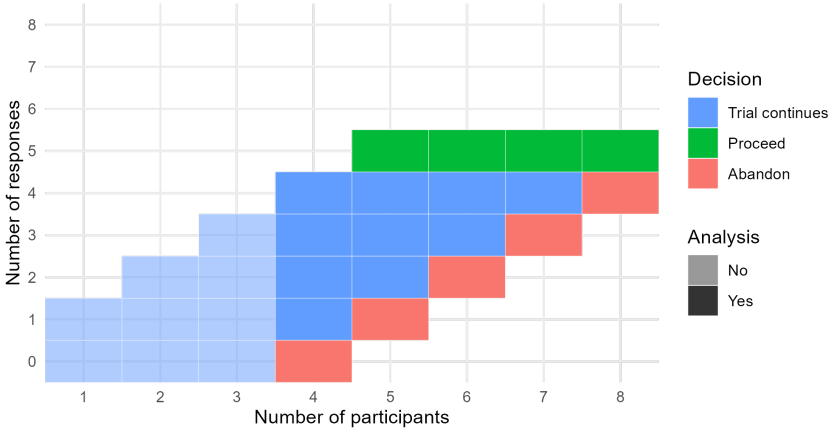
# Add NSC to single-stage design



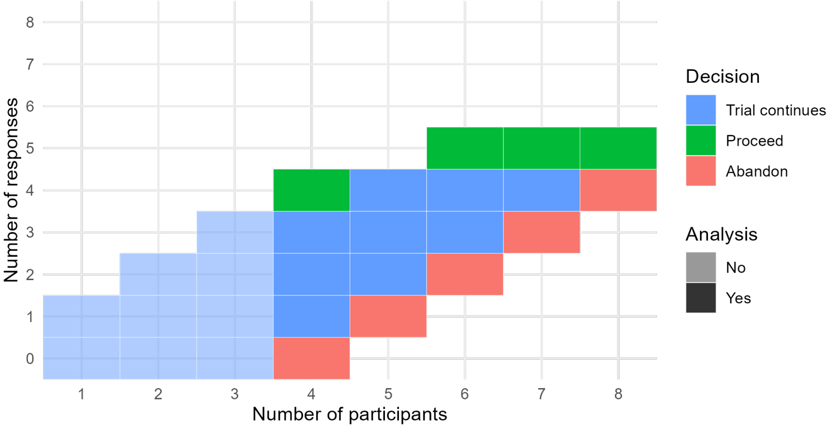
# Add SC to single-stage design



# Add SC to single-stage design

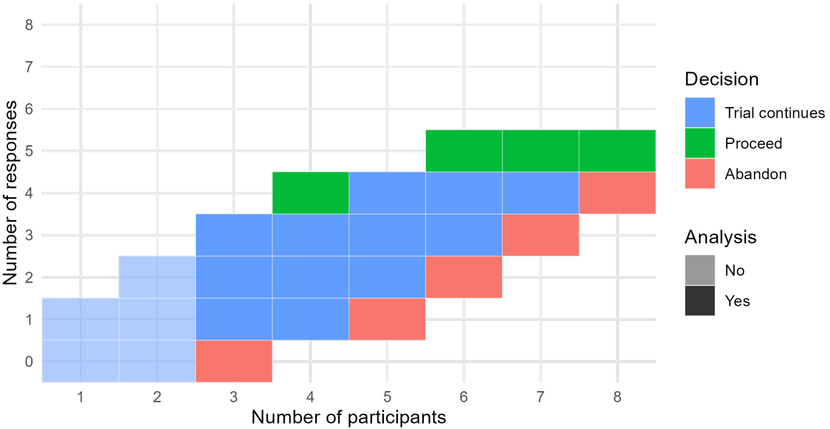


# Add SC to single-stage design




MRC Biostatistics Unit
UNIVERSITY OF CAMBRIDGE

# Add SC to single-stage design




MRC Biostatistics Unit
UNIVERSITY OF CAMBRIDGE

# Advantages, disadvantages: Simon design

Advantages:

- Simple, well-known to regulators
- Decreases number of participants receiving poorly-performing treatment

Disadvantages:

- When treatment works, trial will not end early (exception: you make a type II error!)
- At the points where early stopping takes place, what would be the probability of trial success if there was no stopping boundary? Not (typically) considered when choosing design parameters.

# Advantages, disadvantages: Mander and Thompson

Advantages:

- Simple
- Allows early stopping for promising treatment

Disadvantages:

- May want more information if trial going well

# Advantages, disadvantages: two-stage designs

Disadvantages:

- Final decision may be known with certainty between interim analyses
- Discrete data: may result in design with lower type-I error-rate or higher power than required.

# Advantages, disadvantages: curtailment

Advantage:

- Decreased expected sample size under both $p = p_0$ and $p = p_1$.

Disadvantages:

- More frequent monitoring required:
  - More work (both analysis and logistics)
  - Difficult if responses come quickly

- Curtailment can also be used in the two-arm setting
- Next lecture: multiple arms and multiple stages

# References


- R.P. A'Hern. Sample size tables for exact single-stage phase II designs. Statistics in Medicine, 20(6):859–866, 2001
- R. Simon. Optimal two-stage designs for phase II clinical trials. Controlled Clinical Trials, 10:1–10, 1989.
- A.P. Mander and S.G. Thompson. Two-stage designs optimal under the alternative hypothesis for phase II cancer clinical trials. Contemporary Clinical Trials, 31(6):572-578, 2010.
- M. Law, M.J. Grayling, and A.P. Mander. A stochastically curtailed single-arm phase II trial design for binary outcomes" Journal of biopharmaceutical statistics 32.5: 671-691, 2022.
- M. Law, M.J. Grayling, and A.P. Mander. A stochastically curtailed two-arm randomised phase II trial design for binary outcomes. Pharmaceutical Statistics, 20.2: 212-228, 2021.
- M. Law. Github repository. `https://github.com/martinlaw/curtailment`, 2023