MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

Adaptive Methods in Clinical Research

Lecture 1: Single arm binary outcome designs

Martin Law

MRC Biostatistics Unit

17th September 2024

Link to material

https://github.com/adaptive-designs/course-24

A dose of reality

Ben Goldacre, Guardian[1]:

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

[1] https://www.theguardian.com/business/2008/sep/01/pharmaceuticals.drugs

A dose of reality

Ben Goldacre, Guardian[1]:

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

[1] https://www.theguardian.com/business/2008/sep/01/pharmaceuticals.drugs

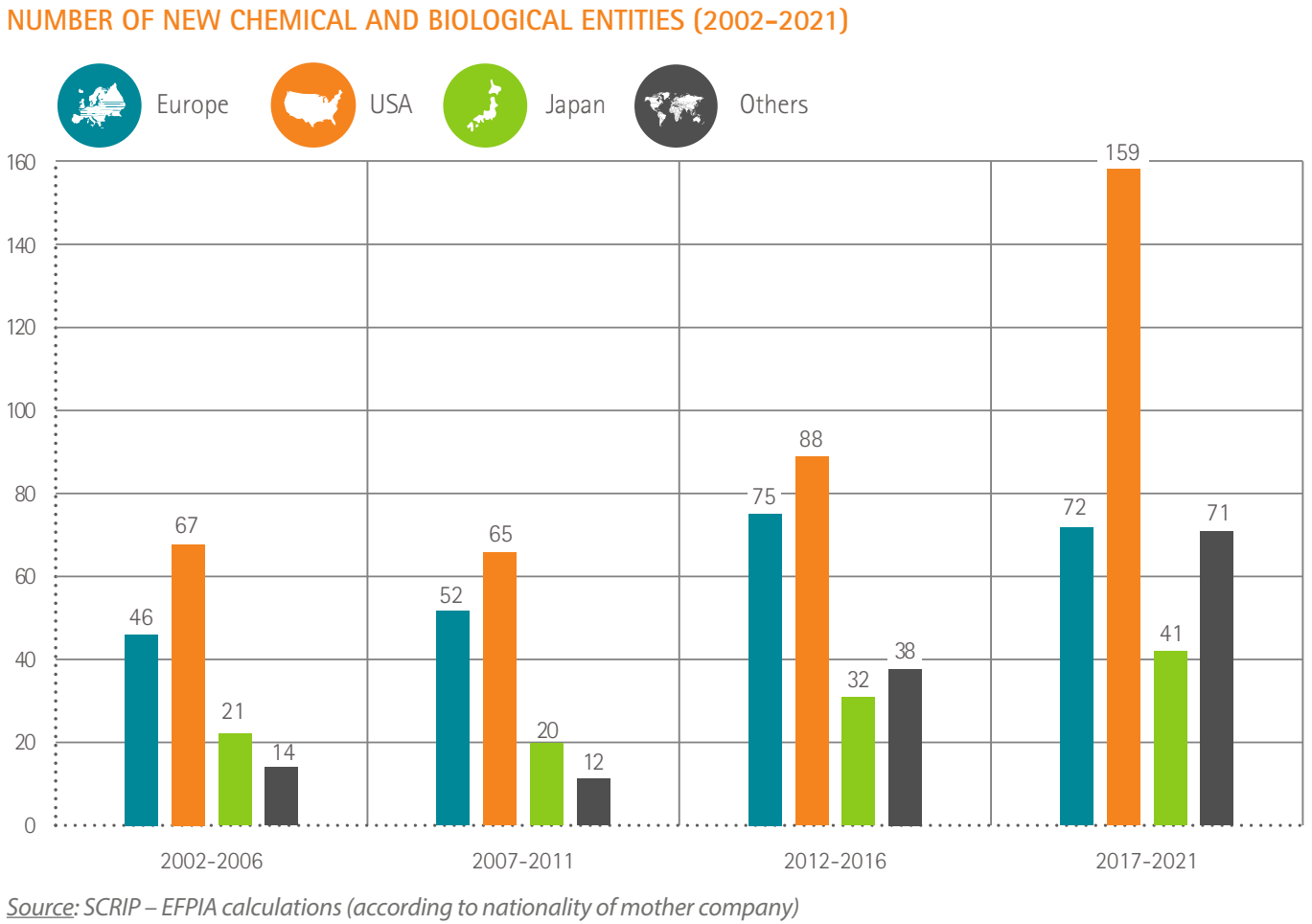
A dose of reality

Ben Goldacre, Guardian[1]:

- **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

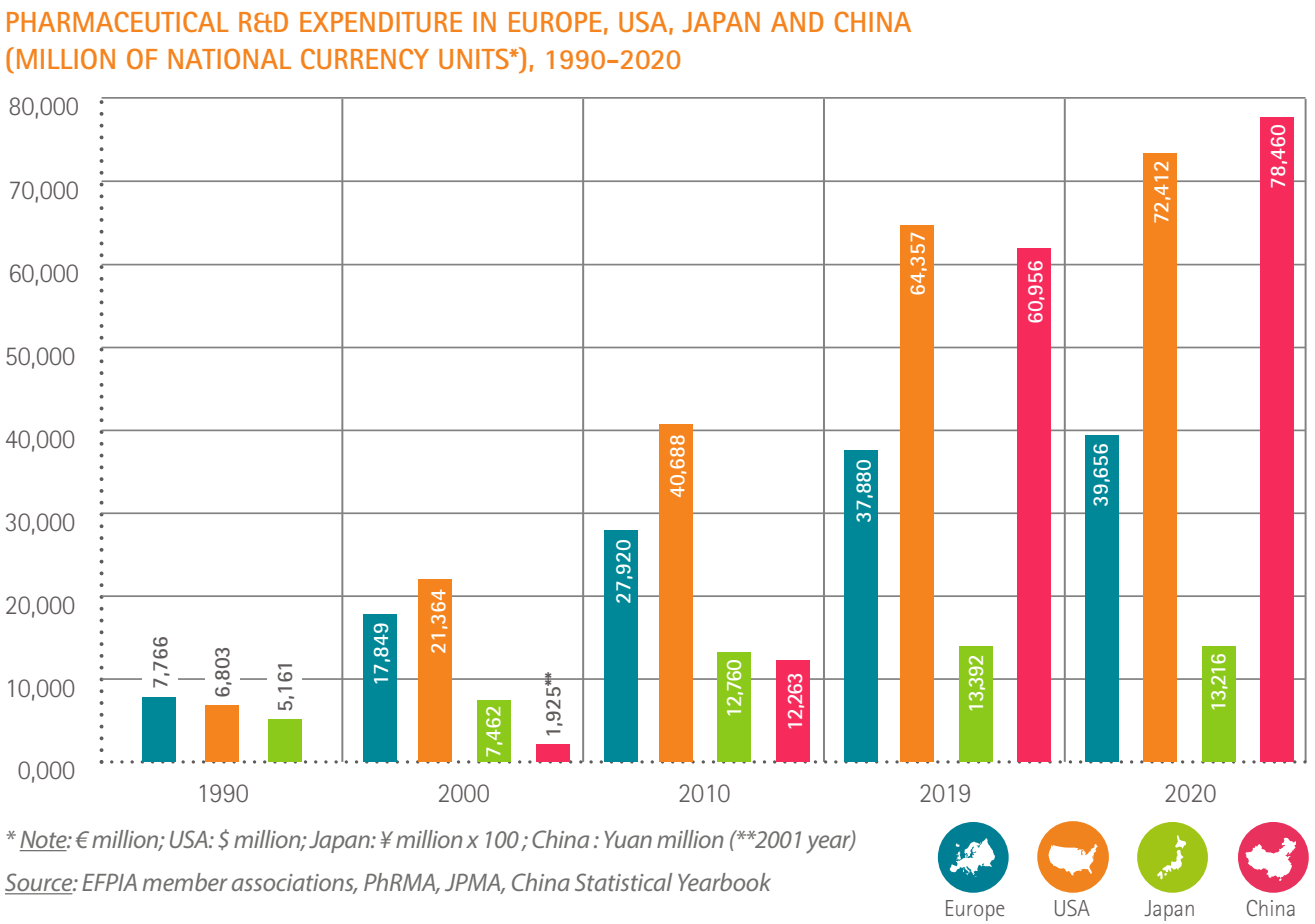
[1] `https://www.theguardian.com/business/2008/sep/01/pharmaceuticals.drugs`

New molecular entities



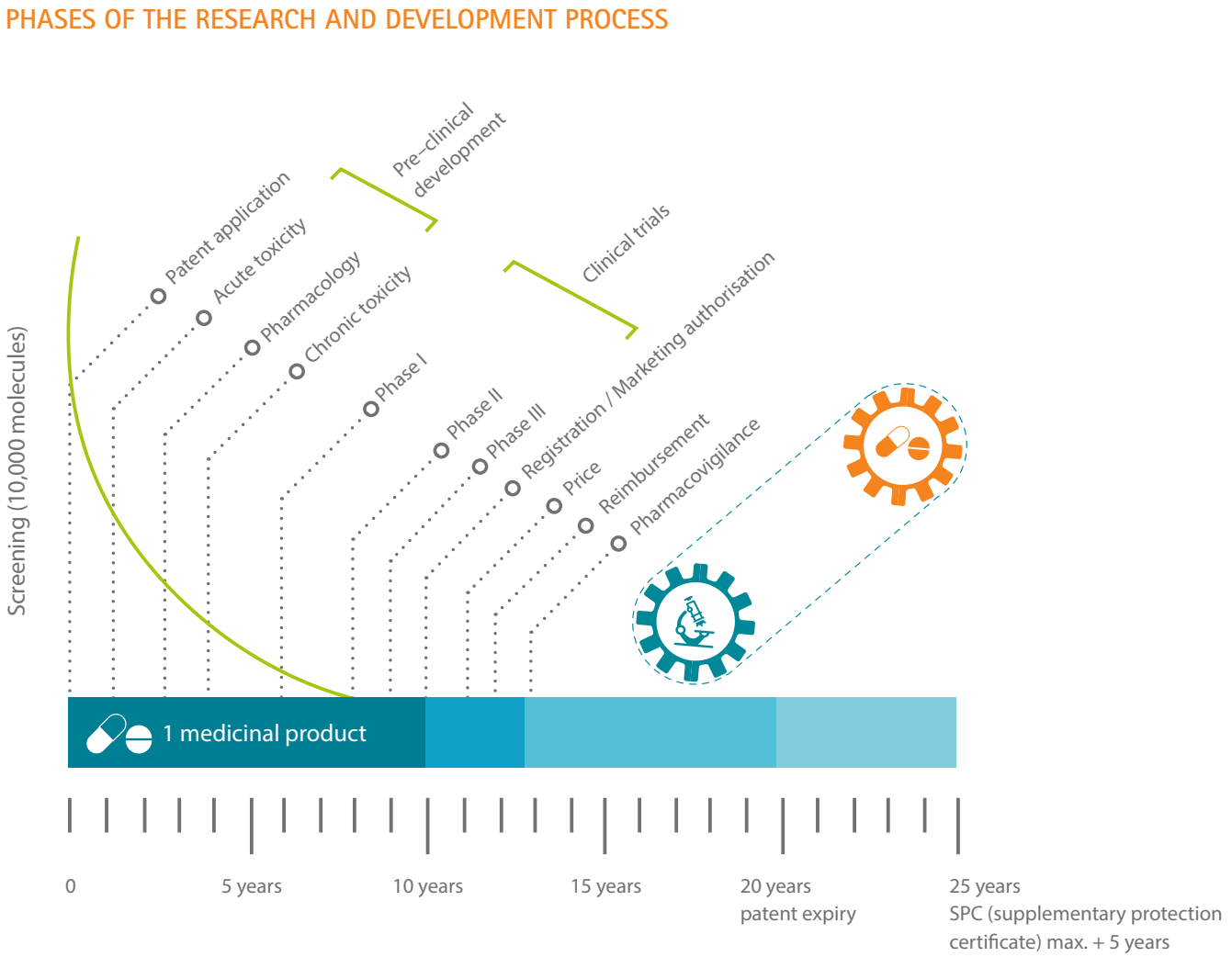
Source: European Federation of Pharmaceutical Industries and Associations (2022)

Cost of R&D in Pharmaceutical industry



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

The development process



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

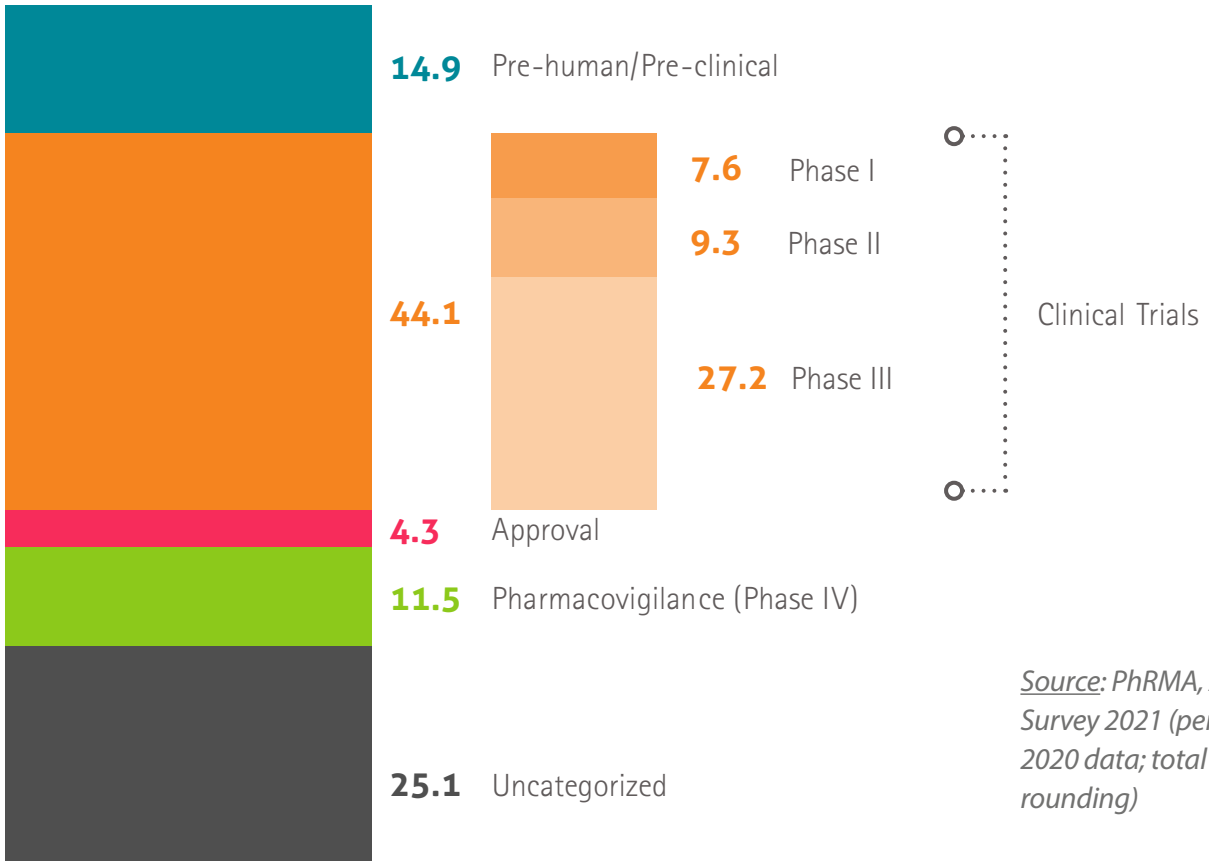
Drug development

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 with thousands of patients and follow-up period lasting years

Cost of R&D in Pharmaceutical industry

ALLOCATION OF R&D INVESTMENTS BY FUNCTION (%)



Source: PhRMA, Annual Membership Survey 2021 (percentages calculated from 2020 data; total values may be affected by rounding)

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

- **59%** of confirmatory clinical trials overall and
- **36%** of confirmatory clinical trials in oncology

were successful.

Success rates

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

- **59%** of confirmatory clinical trials overall and
- **36%** of confirmatory clinical trials in oncology

were successful.

- **14%** of treatments entering Phase I received approval
- In oncology only **3%** of treatments entering Phase I received 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:
  - Adaptive designs
  - Bayesian methods

Adaptive designs

What is an adaptive design?

Adaptive designs

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.

Adaptive designs

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

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.

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

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?

Single-arm binary outcome trials

We test $H_0 : p \leq p_0$ against $H_1 : p \geq p_1$. In general:

Single-arm binary outcome trials

We test $H_0 : p \leq p_0$ against $H_1 : p \geq p_1$. 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 α when $p = p_0$ and power our trial to a level $1 - \beta$ under $p = p_1$, where:

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.

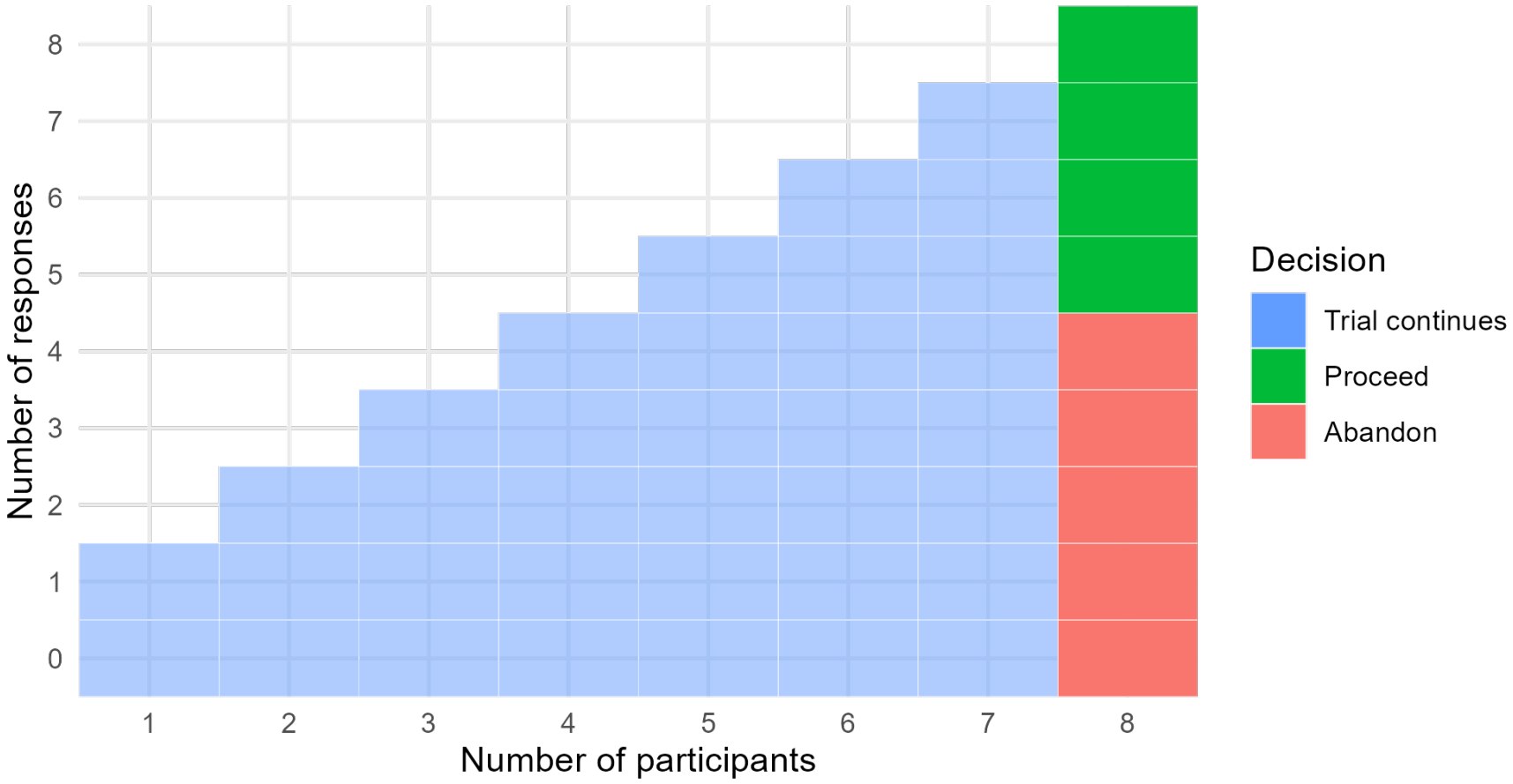
Single-stage design (A'Hern 2001)

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.

Single-stage design (A'Hern 2001)



Simon design (Simon 1989)

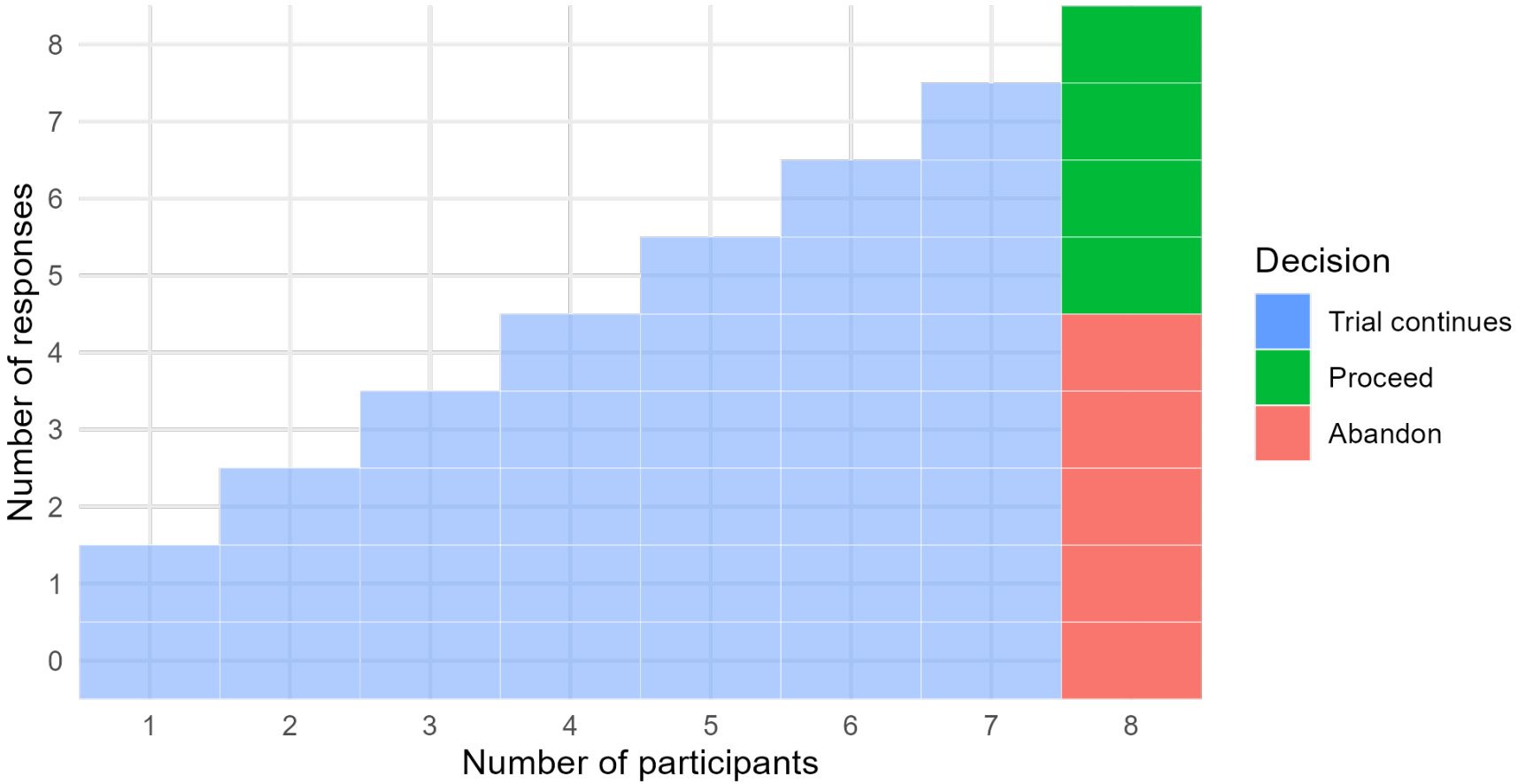
Simon design is a simple adaptation to the single-stage design:

Include single interim analysis after n_1 participants.

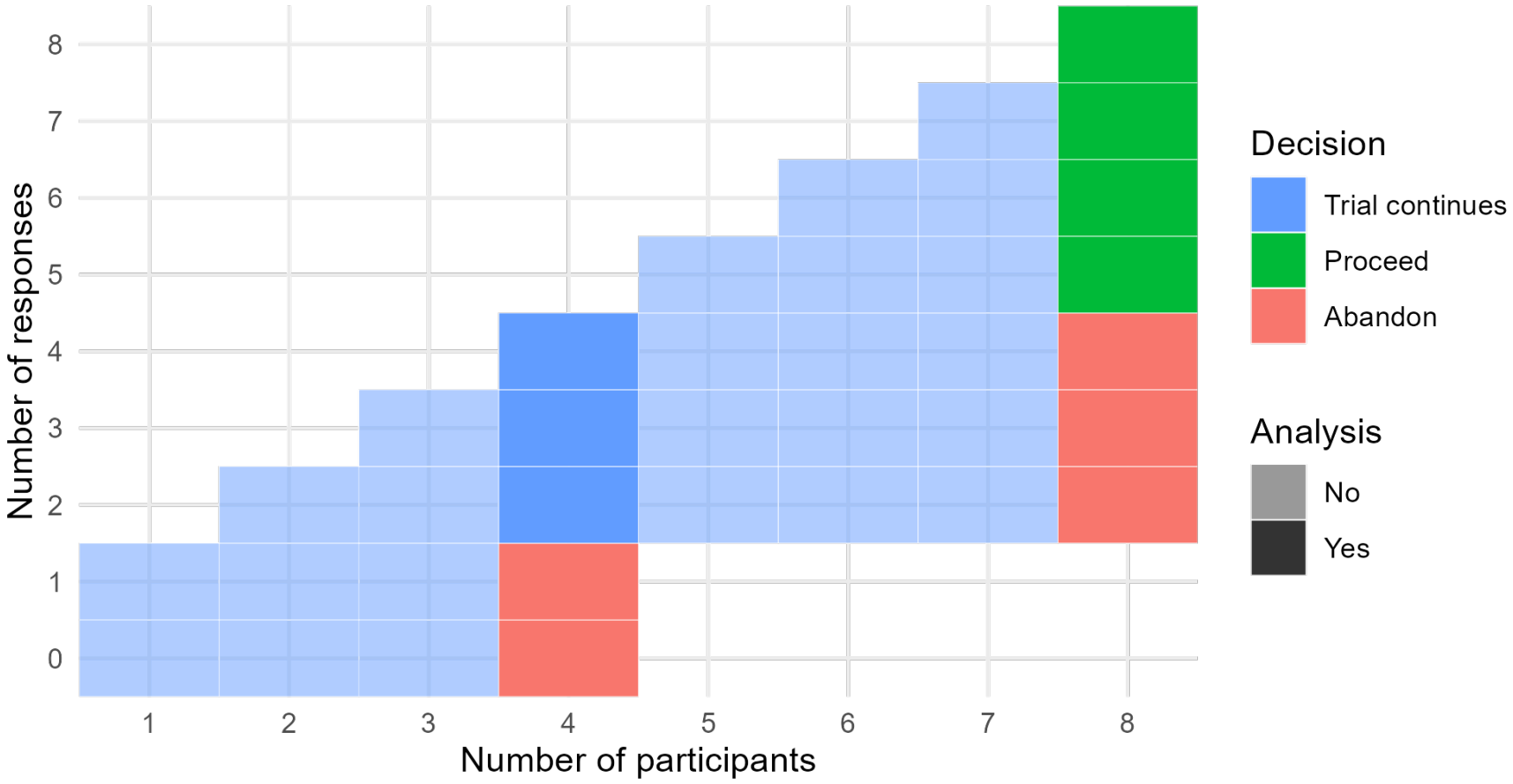
If $S_{n_1} \leq r_1$, stop for lack of benefit.

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

Single-stage design



Simon design



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 design

```
##
##  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

Simon design: qLo, qHi?

Instead of minimising n_{max} *or* minimising $ESS(H_0)$, consider minimising the following weighted sum:

$$\omega n_{max} + (1 - \omega) ESS(H_0).$$

Each returned design minimises this sum for $\omega \in$ [qLo, qHi].

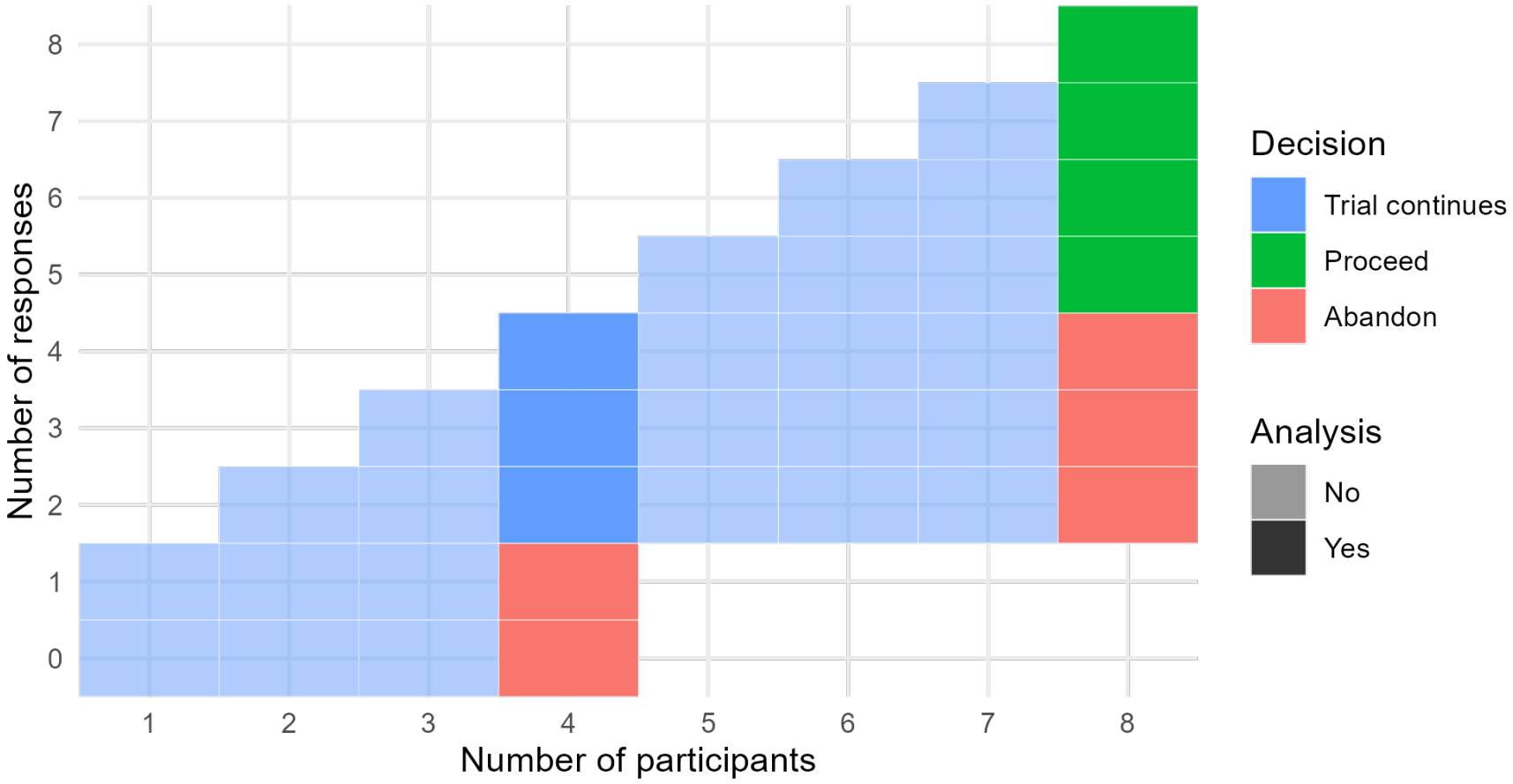
In other words, `qLo` and `qHi` denote the interval of weights for maximum sample size for which each design is best.

Mander and Thompson design (2010)

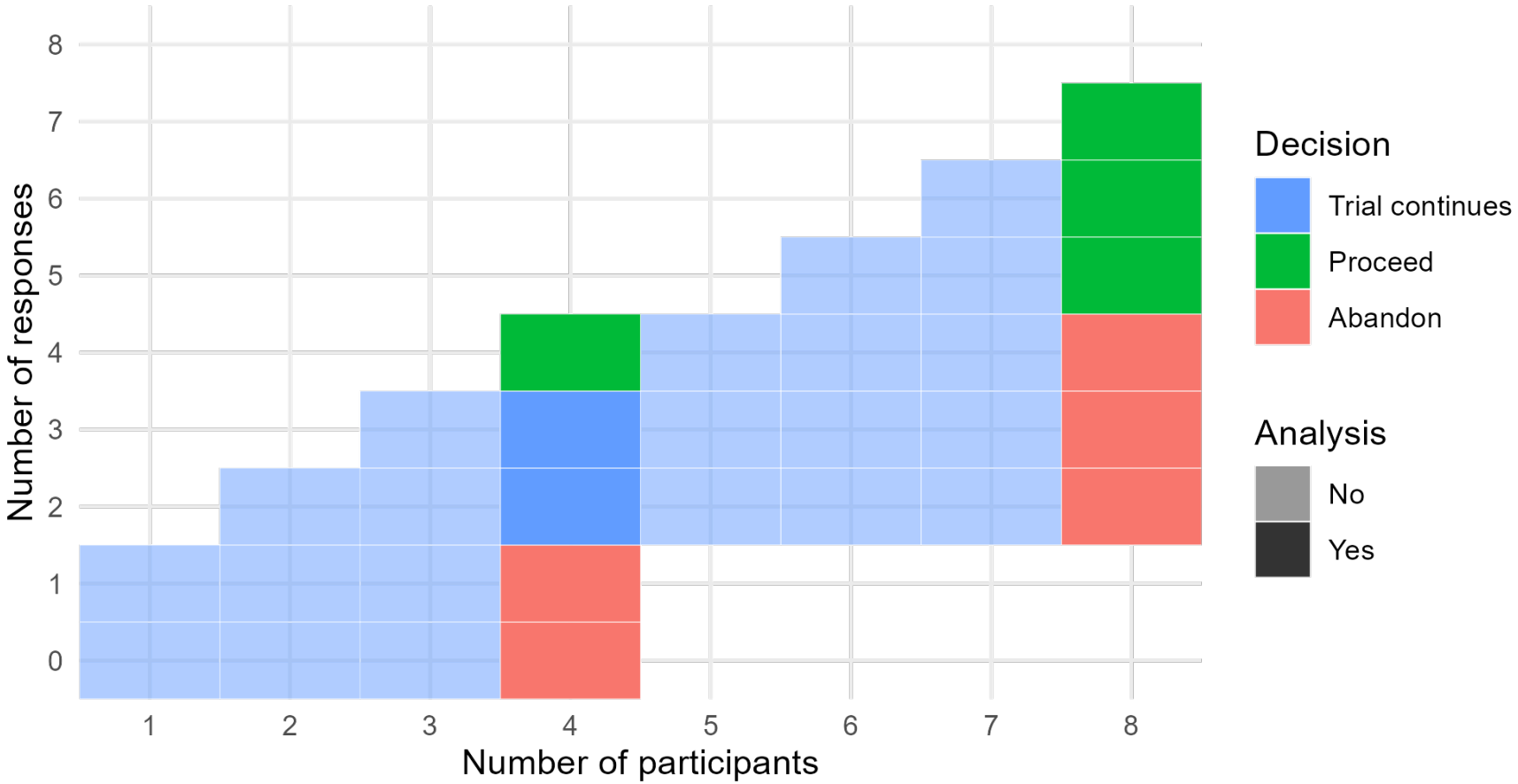
Mander and Thompson design (2010)

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

Simon design



Mander and Thompson design



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 α 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, α and $1 - \beta$.

Curtailment in trial design

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

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”?

“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)

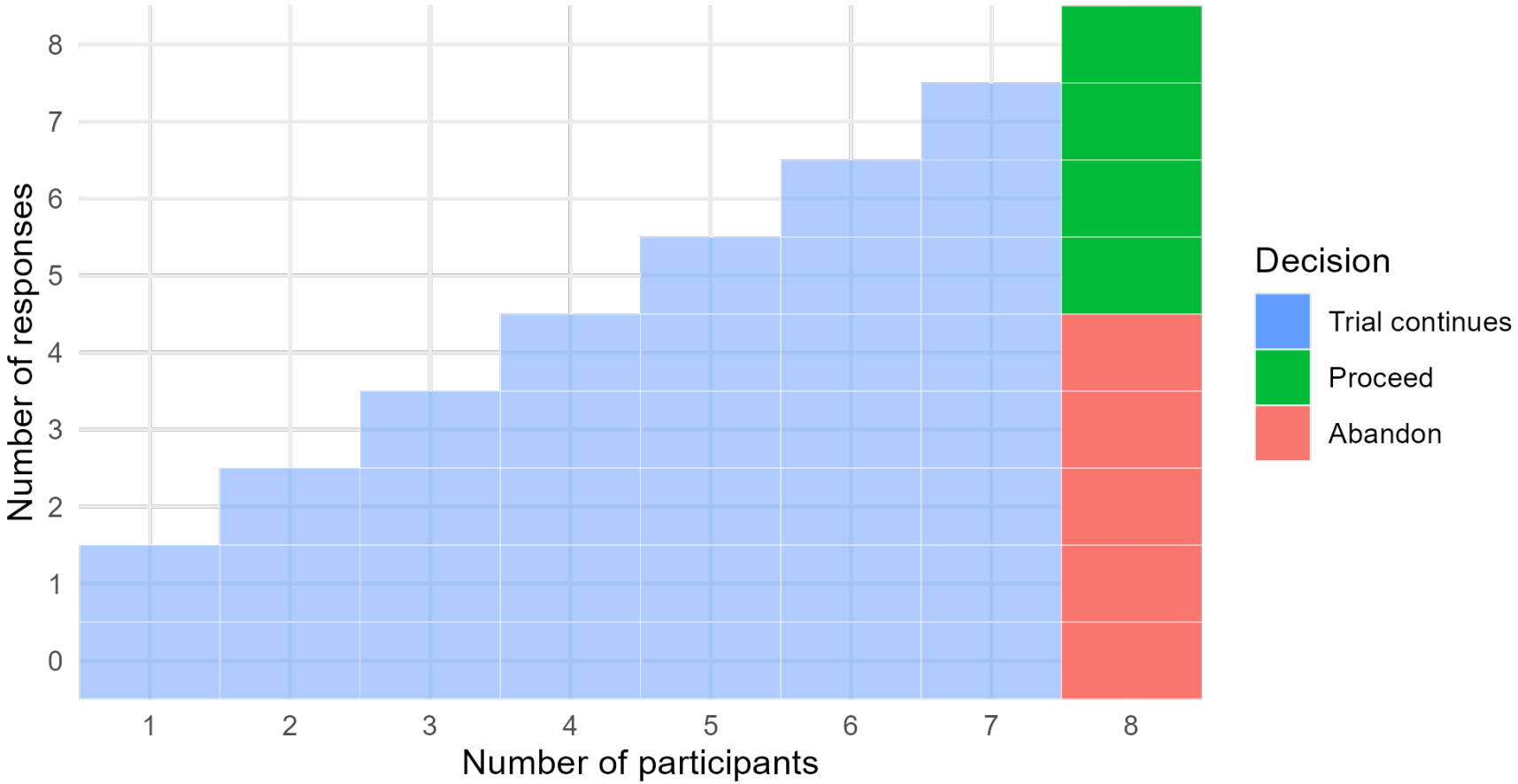
By planning to end a trial early due to a high or low CP, we can reduce that trial's expected sample size.

Conditional power (CP)

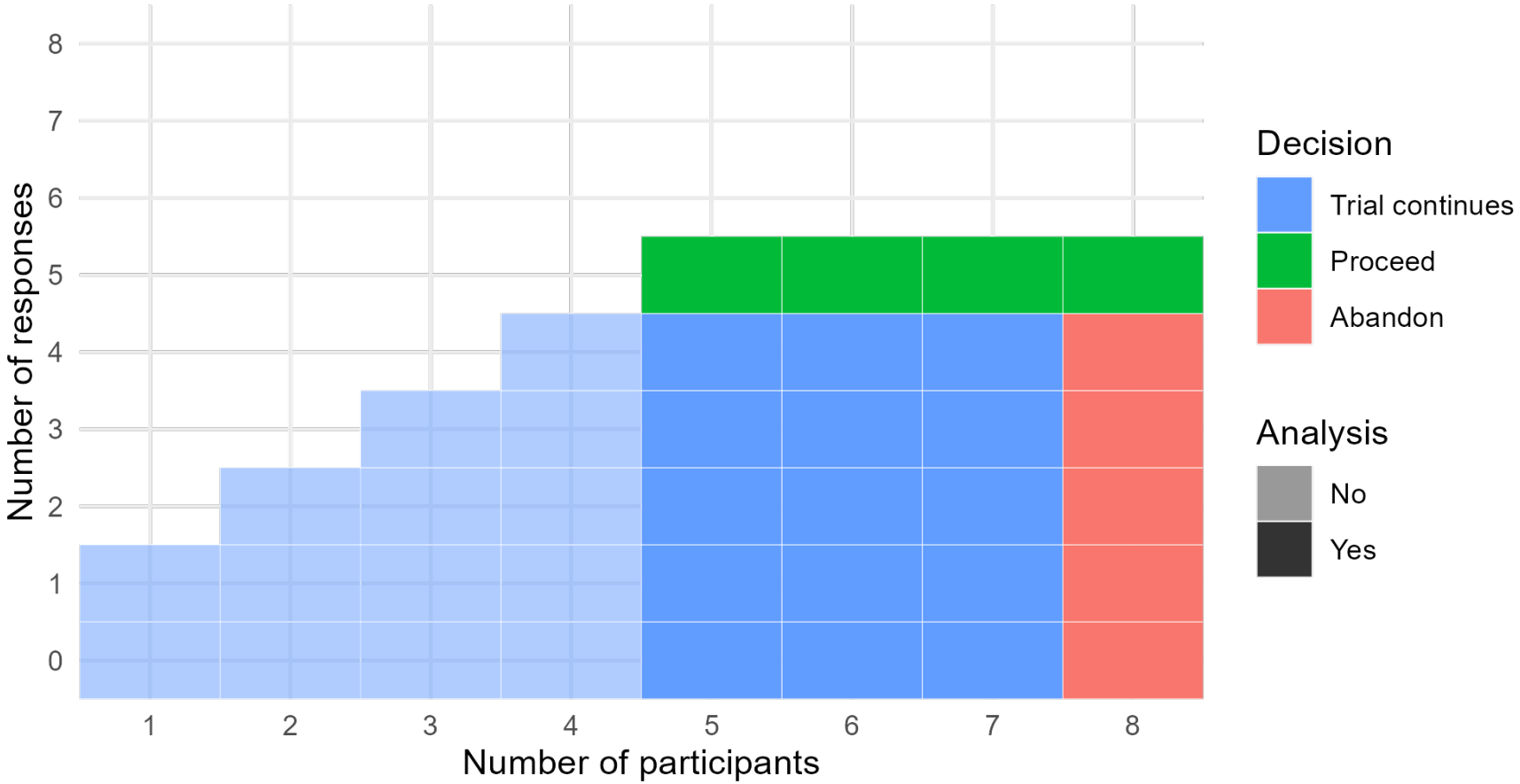
By planning to end a trial early due to a high or low CP, we can reduce that trial's expected sample size.

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

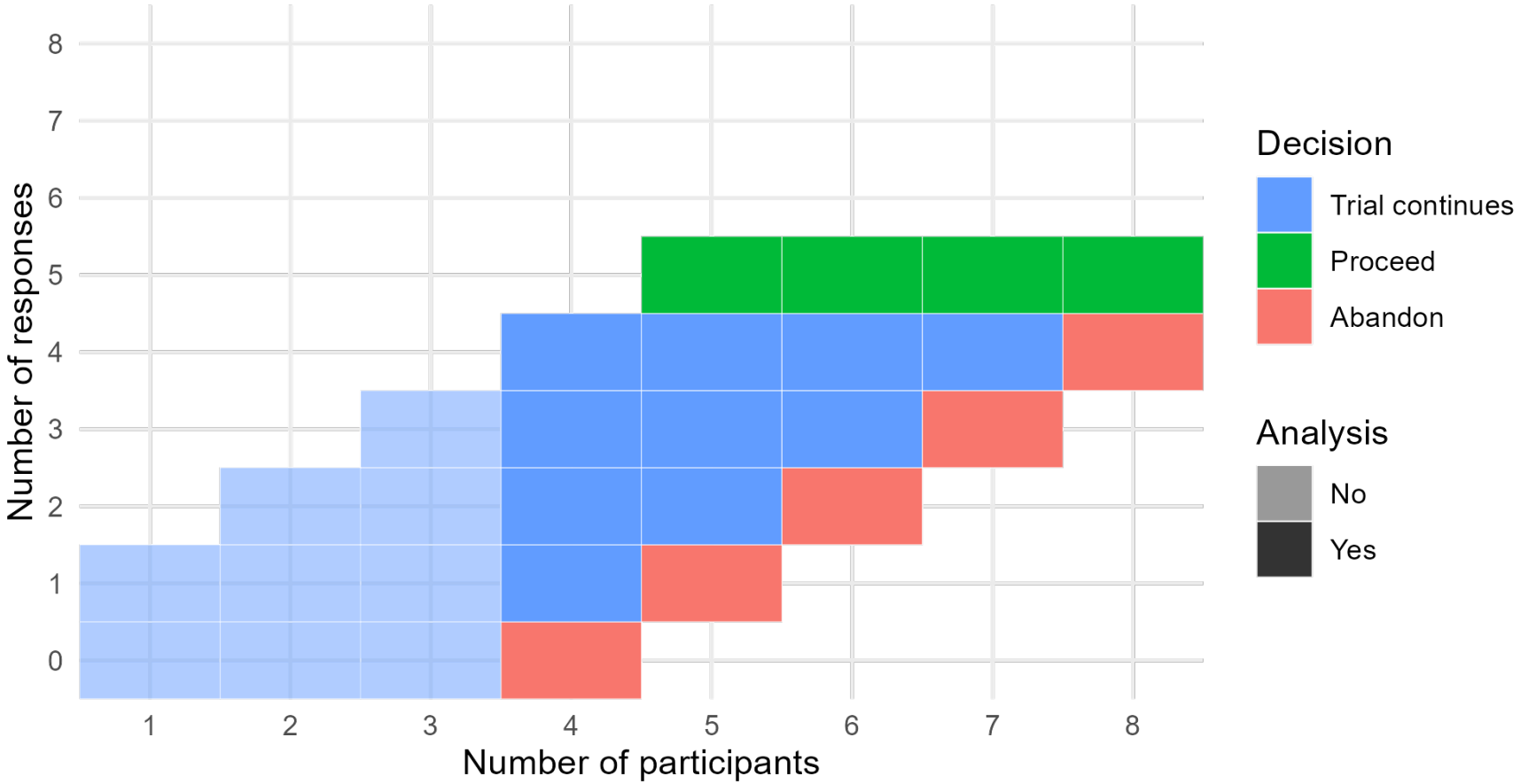
Add NSC to single-stage design



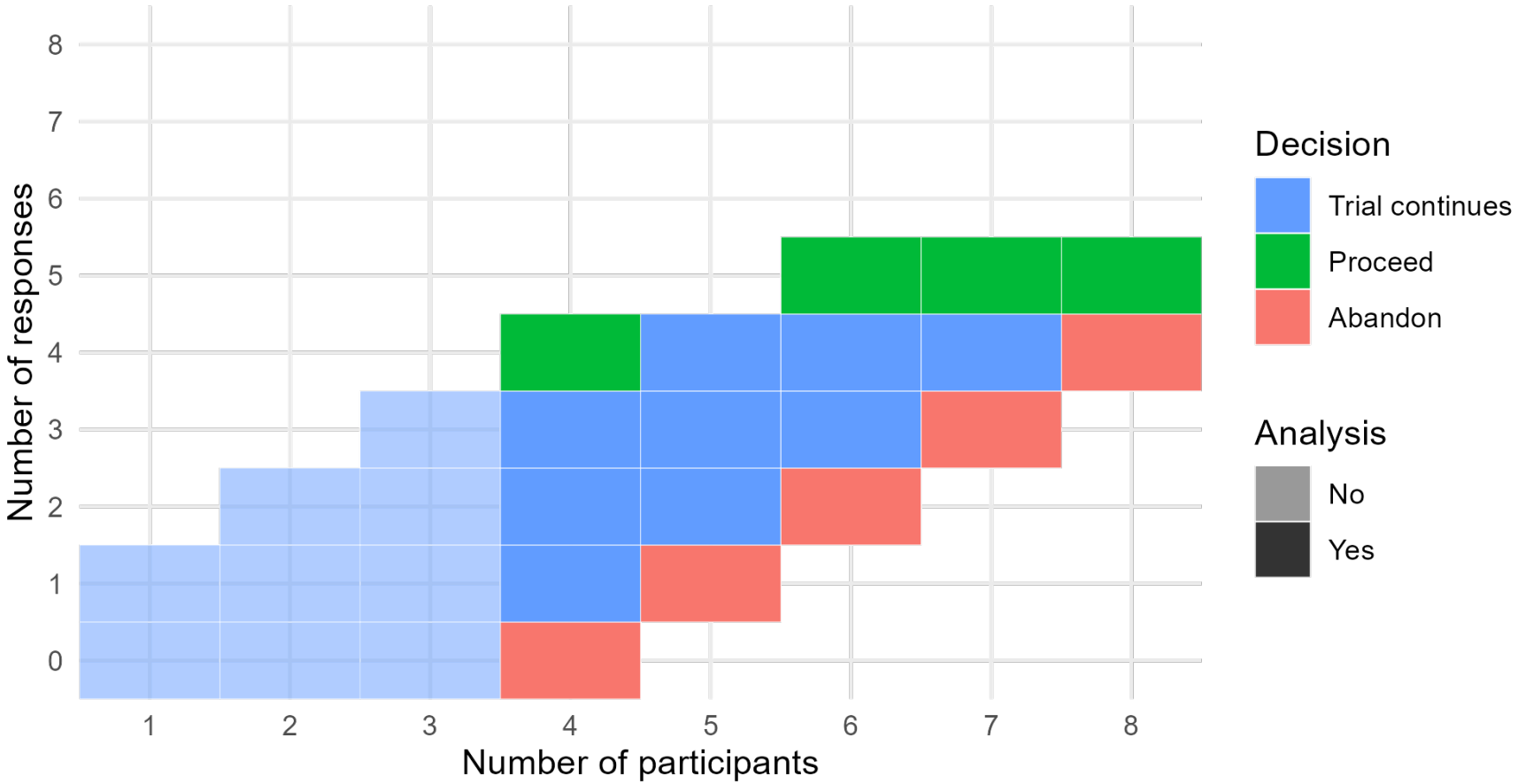
Add NSC to single-stage design



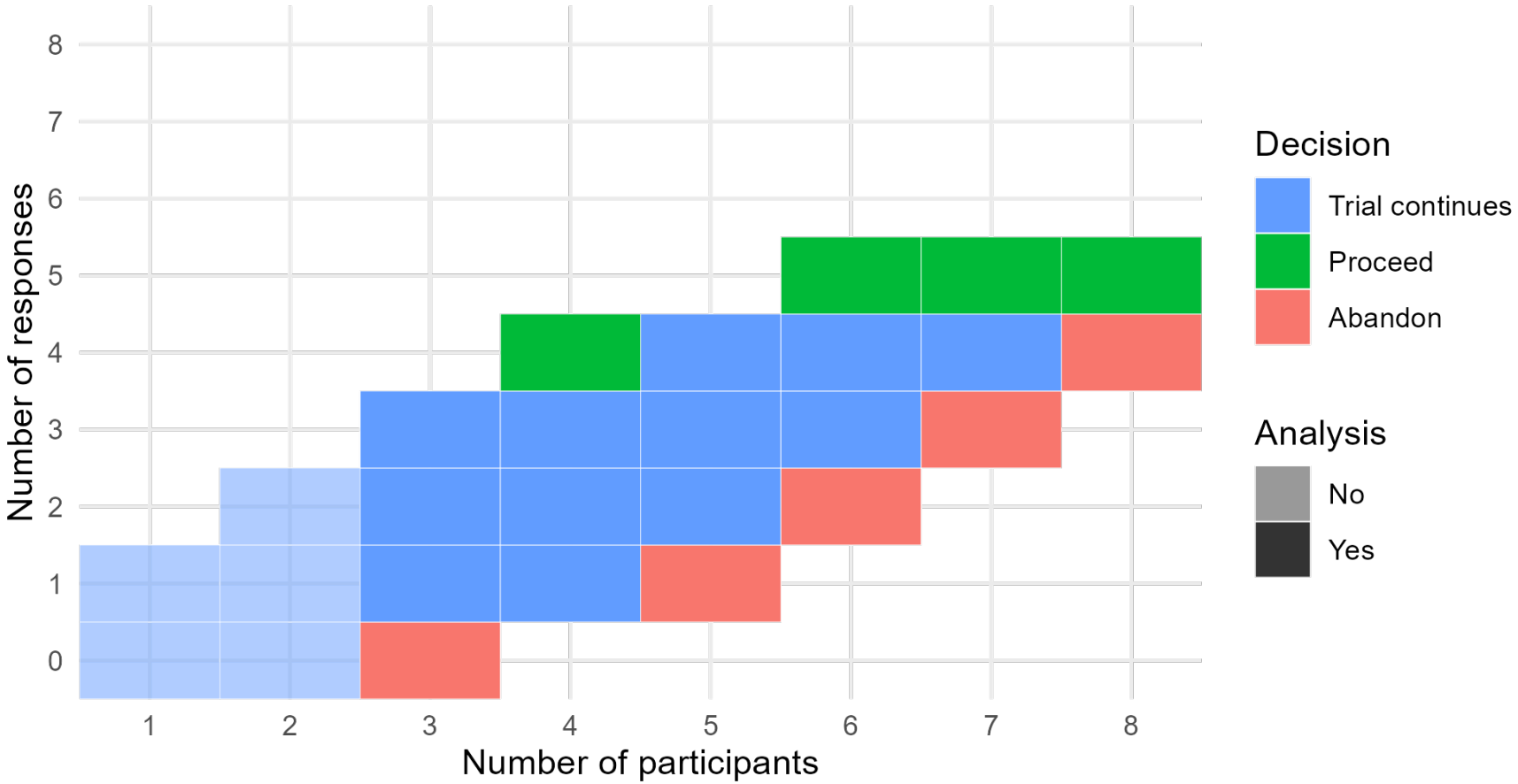
Add NSC to single-stage design



Add SC to single-stage design



Add SC to single-stage design



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 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
- Less information (as smaller ESS)

Beyond single-arm single-stage designs

- Curtailment can also be used in the two-arm setting

Beyond single-arm single-stage designs

- Curtailment can also be used in the two-arm setting
- Next: Practical session using some of these designs
- Next lecture: Group sequential designs

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. Github repository. `https://github.com/martinlaw/curtailment`, 2023