Adaptive designs - an overview and some examples

Thomas Jaki Philip Pallmann
Lancaster University

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

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

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

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

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

Reasons for failure (Arrowsmith & Miller 2013)

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)

Medical and Pharmaceutical
Statistics Research Unit

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)

Medical and Pharmaceutical
Statistics Research Unit

Likely causes for failure:

Likely causes for failure:

- taking forward futile treatments

Likely causes for failure:

- taking forward futile treatments
- studying the wrong patient population

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?

Medical and Pharmaceutical
Statistics Research Unit

Can we do better?

- avoid going straight into large and expensive phase III

Can we do better?

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

Medical and Pharmaceutical
Statistics Research Unit

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

Medical and Pharmaceutical
Statistics Research Unit

2. Adaptive Designs

Idea

Modify an ongoing trial

Idea

Modify an ongoing trial

by design or ad hoc

Idea

Modify an ongoing trial

by design or ad hoc

based on reviewing accrued data at interim

Idea

Modify an ongoing trial

by design or ad hoc

based on reviewing accrued data at interim

to enhance flexibility

Medical and Pharmaceutical
Statistics Research Unit

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)

Medical and Pharmaceutical
Statistics Research Unit

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
– simple estimates may be biased
– interim analyses may require unblinding

Medical and Pharmaceutical
Statistics Research Unit

Fixed sample design

Medical and Pharmaceutical
Statistics Research Unit

Fixed sample design

- **total** sample size known in advance

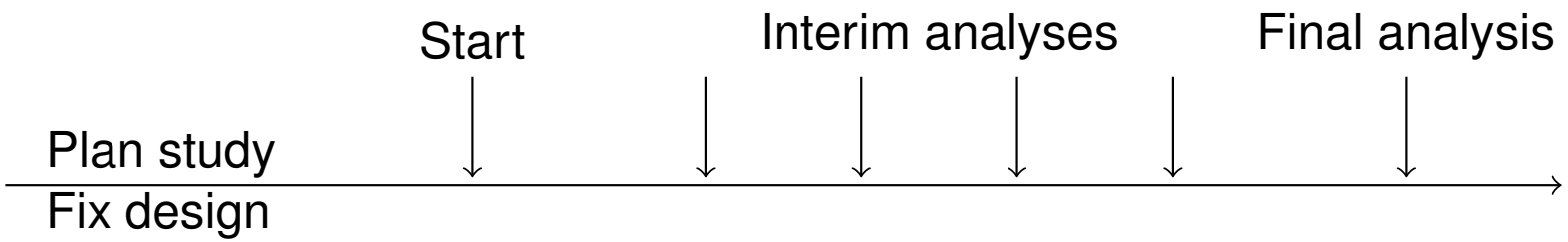
Medical and Pharmaceutical
Statistics Research Unit

Fixed sample design

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

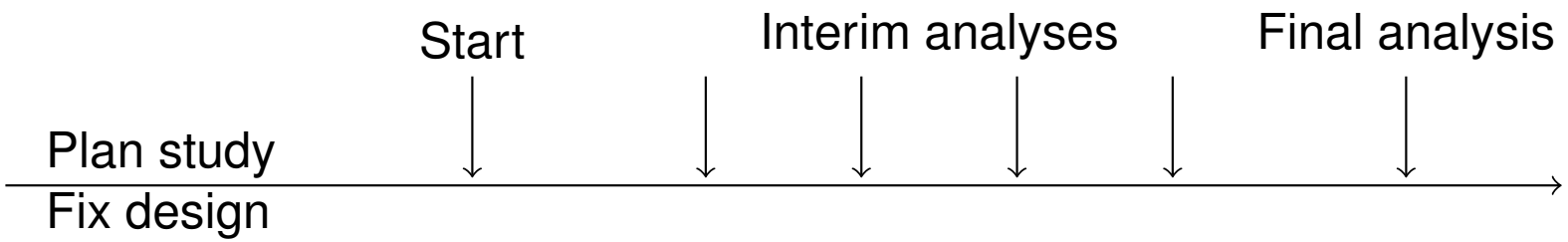
Medical and Pharmaceutical
Statistics Research Unit

(Group-)sequential design



Medical and Pharmaceutical
Statistics Research Unit

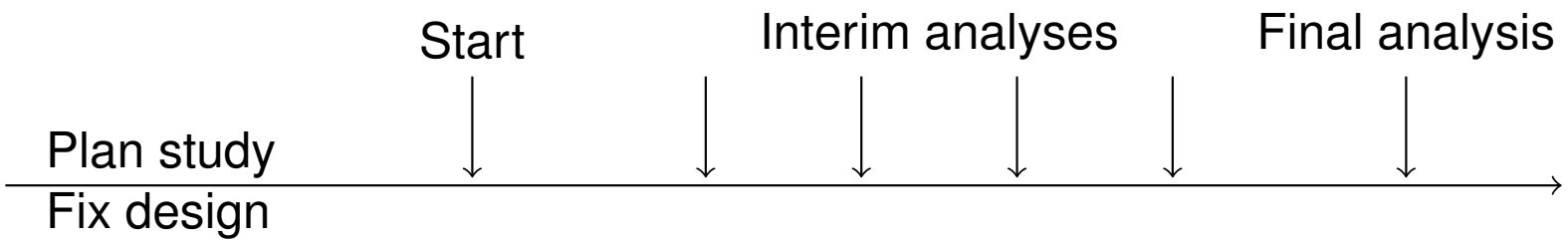
(Group-)sequential design



At each interim:

Medical and Pharmaceutical
Statistics Research Unit

(Group-)sequential design

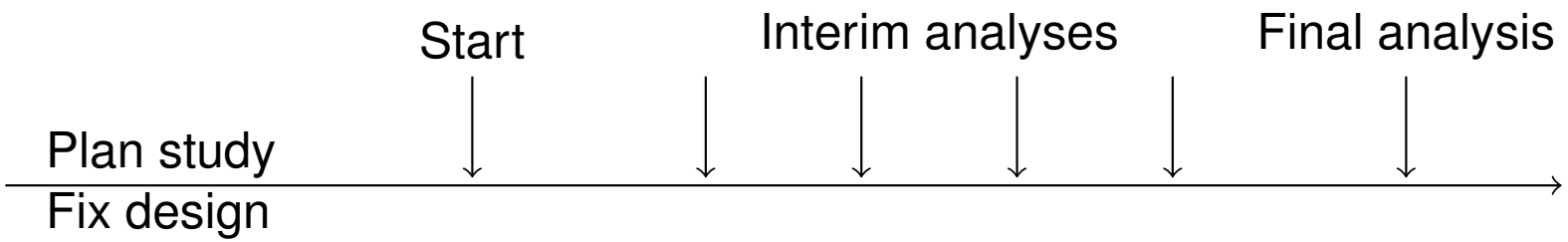


At each interim:

- decide whether or not to stop

Medical and Pharmaceutical
Statistics Research Unit

(Group-)sequential design

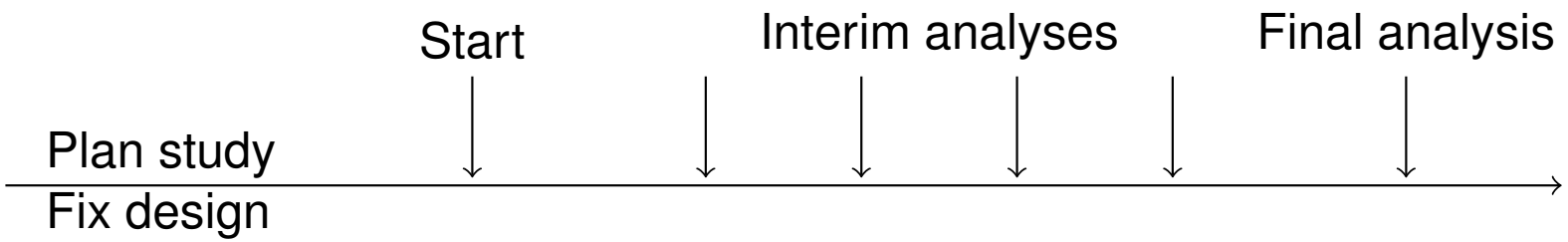


At each interim:

- decide whether or not to stop

Medical and Pharmaceutical
Statistics Research Unit

(Group-)sequential design

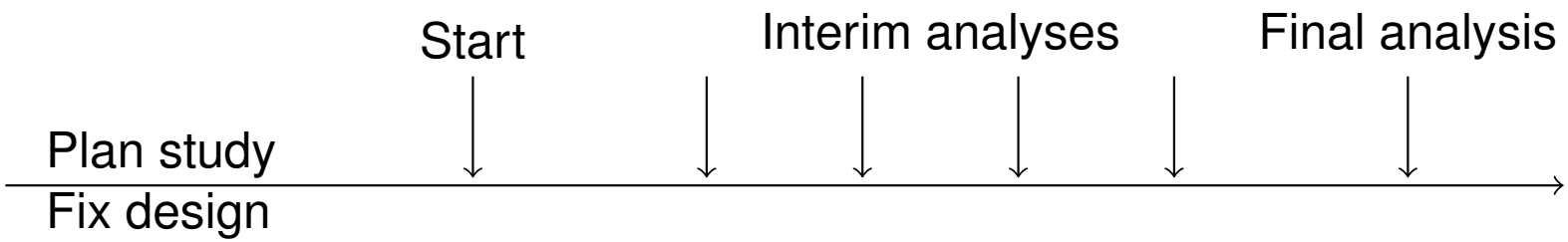


At each interim:

- decide whether or not to stop
- larger **maximum** sample size

Medical and Pharmaceutical
Statistics Research Unit

(Group-)sequential design



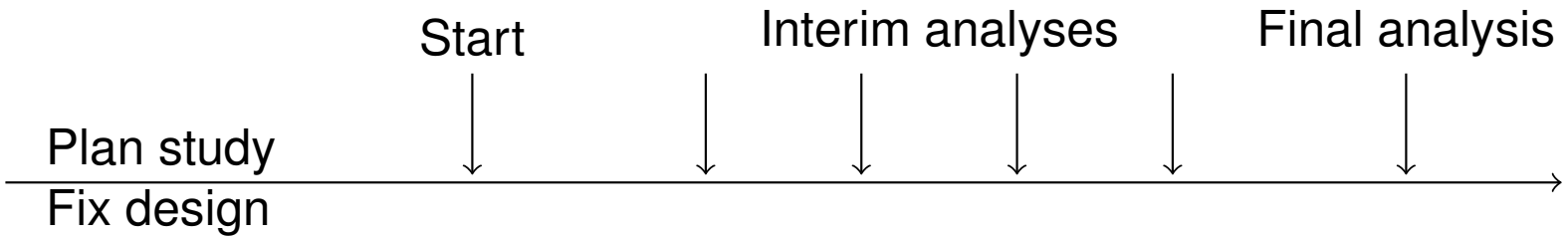
At each interim:

- decide whether or not to stop

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

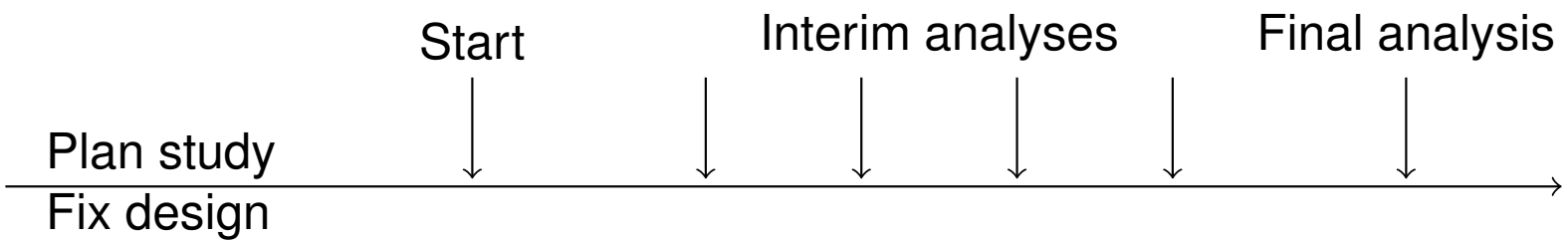
Medical and Pharmaceutical
Statistics Research Unit

Adaptive design



Medical and Pharmaceutical
Statistics Research Unit

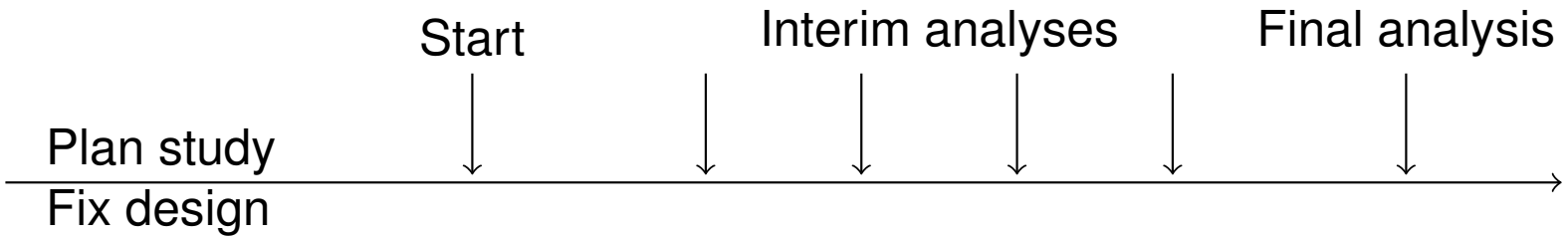
Adaptive design



At each interim:

Medical and Pharmaceutical
Statistics Research Unit

Adaptive design

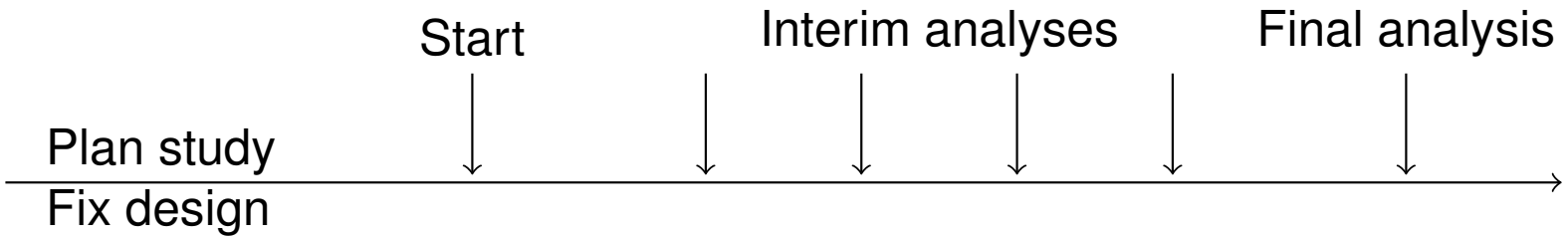


At each interim:

- decide whether or not to stop

Medical and Pharmaceutical
Statistics Research Unit

Adaptive design

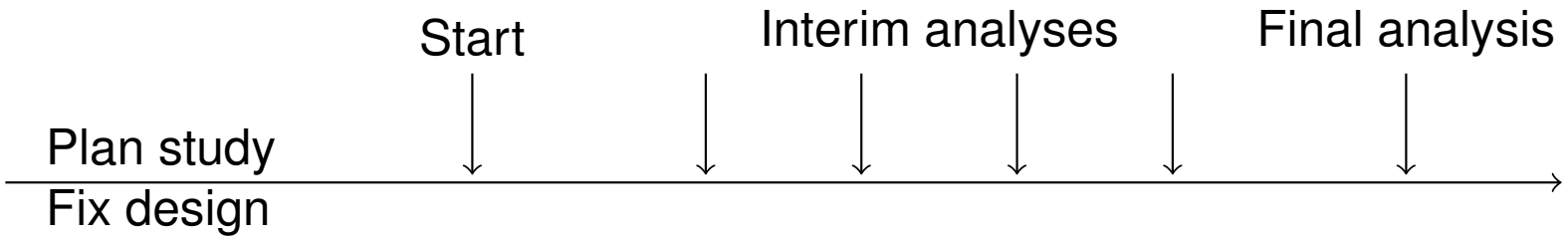


At each interim:

- decide whether or not to stop
- and many other options . . .

Medical and Pharmaceutical
Statistics Research Unit

Adaptive design

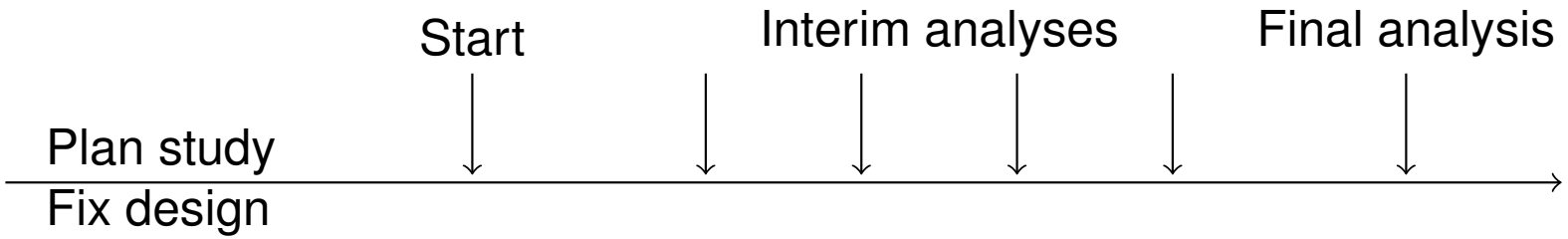


At each interim:

- decide whether or not to stop
- and many other options . . .

Medical and Pharmaceutical
Statistics Research Unit

Adaptive design



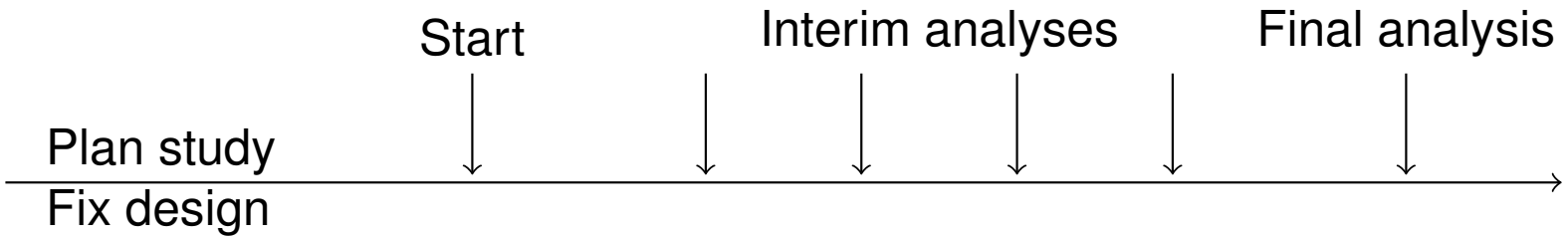
At each interim:

- decide whether or not to stop
- and many other options . . .

- larger **maximum** sample size

Medical and Pharmaceutical
Statistics Research Unit

Adaptive design



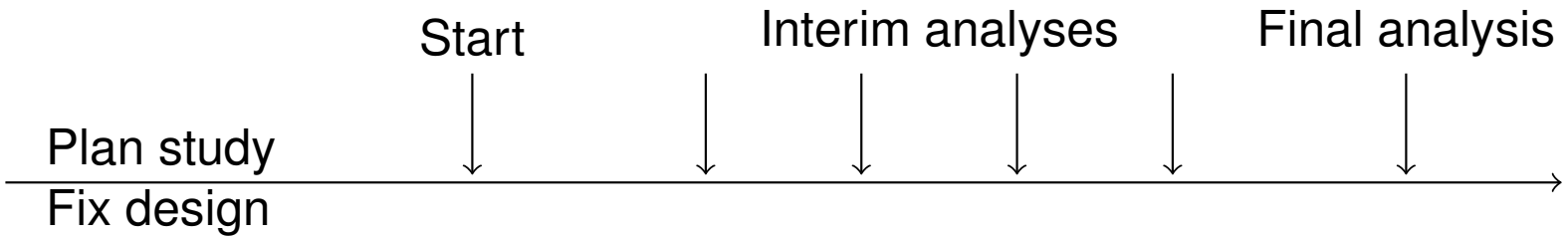
At each interim:

- decide whether or not to stop
- and many other options . . .

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

Medical and Pharmaceutical
Statistics Research Unit

Adaptive design



At each interim:

- decide whether or not to stop
- and many other options . . .

- larger **maximum** sample size
- lower **expected** sample size
- possibly more relevant answer

Medical and Pharmaceutical
Statistics Research Unit

How and what can we adapt?

Prospectively (“by design")

Medical and Pharmaceutical
Statistics Research Unit

How and what can we adapt?

Prospectively (“by design")

- adaptive randomization

How and what can we adapt?

Prospectively (“by design")

- adaptive randomization
- early stopping (safety, futility, efficacy)

Medical and Pharmaceutical
Statistics Research Unit

How and what can we adapt?

Prospectively (“by design")

- adaptive randomization
- early stopping (safety, futility, efficacy)
- drop-the-loser(s), pick-the-winner

How and what can we adapt?

Prospectively (“by design")

- adaptive randomization
- early stopping (safety, futility, efficacy)
- drop-the-loser(s), pick-the-winner
- sample size re-estimation (to achieve desired power)

Medical and Pharmaceutical
Statistics Research Unit

How and what can we adapt?

Prospectively (“by design")

- adaptive randomization
- early stopping (safety, futility, efficacy)
- drop-the-loser(s), pick-the-winner
- sample size re-estimation (to achieve desired power)

How and what can we adapt?

Prospectively (“by design")

- adaptive randomization
- early stopping (safety, futility, efficacy)
- drop-the-loser(s), pick-the-winner
- sample size re-estimation (to achieve desired power)

Concurrently (“ad hoc")

How and what can we adapt?

Prospectively (“by design")

- adaptive randomization
- early stopping (safety, futility, efficacy)
- drop-the-loser(s), pick-the-winner
- sample size re-estimation (to achieve desired power)

Concurrently (“ad hoc")

- modify inclusion/exclusion criteria

Medical and Pharmaceutical
Statistics Research Unit

How and what can we adapt?

Prospectively (“by design")

- adaptive randomization
- early stopping (safety, futility, efficacy)
- drop-the-loser(s), pick-the-winner
- sample size re-estimation (to achieve desired power)

Concurrently (“ad hoc")

- modify inclusion/exclusion criteria
- change primary endpoint

Medical and Pharmaceutical
Statistics Research Unit

How and what can we adapt?

Prospectively (“by design")

- adaptive randomization
- early stopping (safety, futility, efficacy)
- drop-the-loser(s), pick-the-winner
- sample size re-estimation (to achieve desired power)

Concurrently (“ad hoc")

- modify inclusion/exclusion criteria
- change primary endpoint
- modify doses

Medical and Pharmaceutical
Statistics Research Unit

How and what can we adapt?

Prospectively (“by design")

- adaptive randomization
- early stopping (safety, futility, efficacy)
- drop-the-loser(s), pick-the-winner
- sample size re-estimation (to achieve desired power)

Concurrently (“ad hoc")

- modify inclusion/exclusion criteria
- change primary endpoint
- modify doses
- extend treatment duration

Medical and Pharmaceutical
Statistics Research Unit

How and what can we adapt?

Prospectively (“by design")

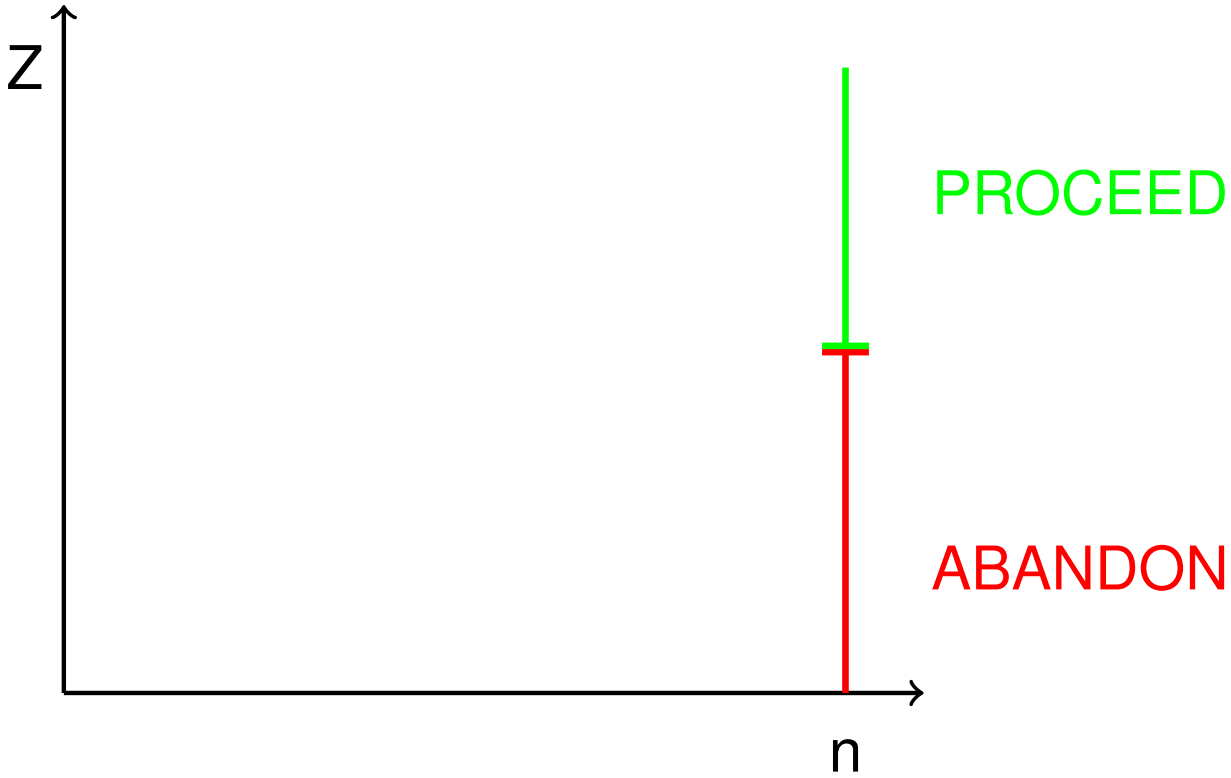
- adaptive randomization
- early stopping (safety, futility, efficacy)
- drop-the-loser(s), pick-the-winner
- sample size re-estimation (to achieve desired power)

Concurrently (“ad hoc")

- modify inclusion/exclusion criteria
- change primary endpoint
- modify doses
- extend treatment duration
- change trial objective (e.g., non-inferiority $\rightarrow$ superiority)

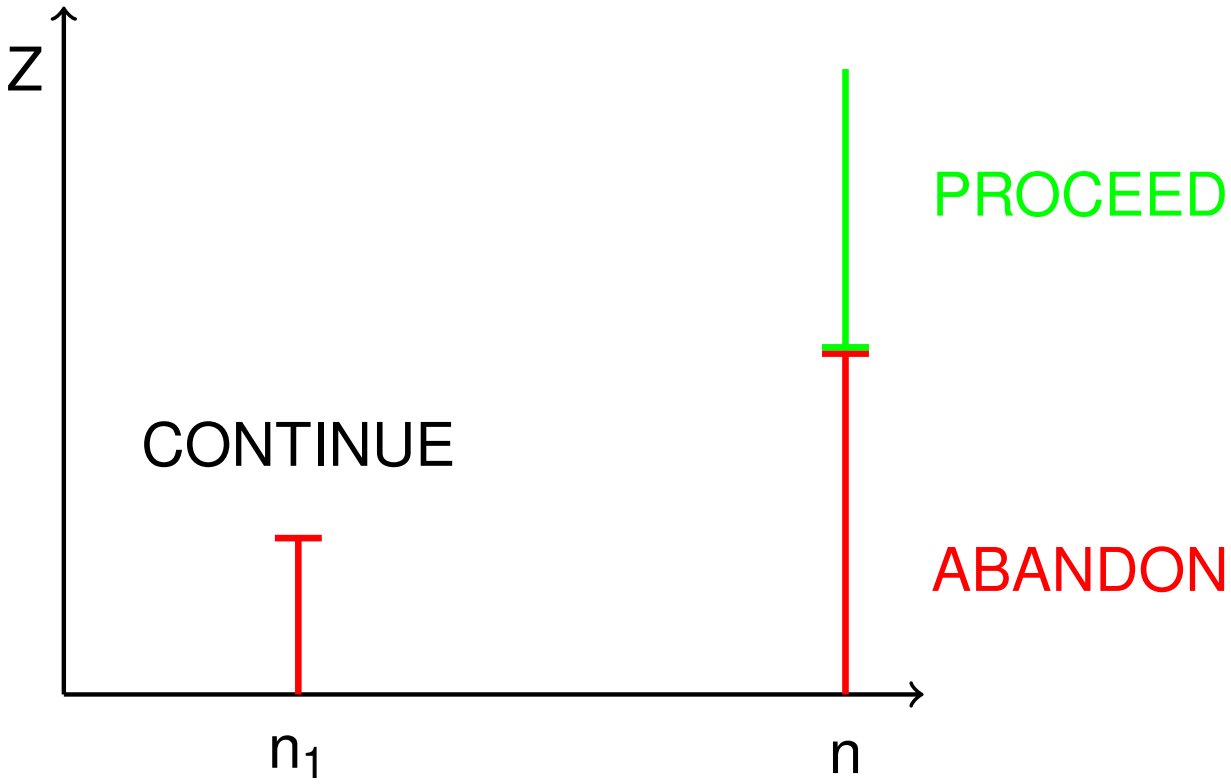
Medical and Pharmaceutical
Statistics Research Unit

A single-stage design



Medical and Pharmaceutical
Statistics Research Unit

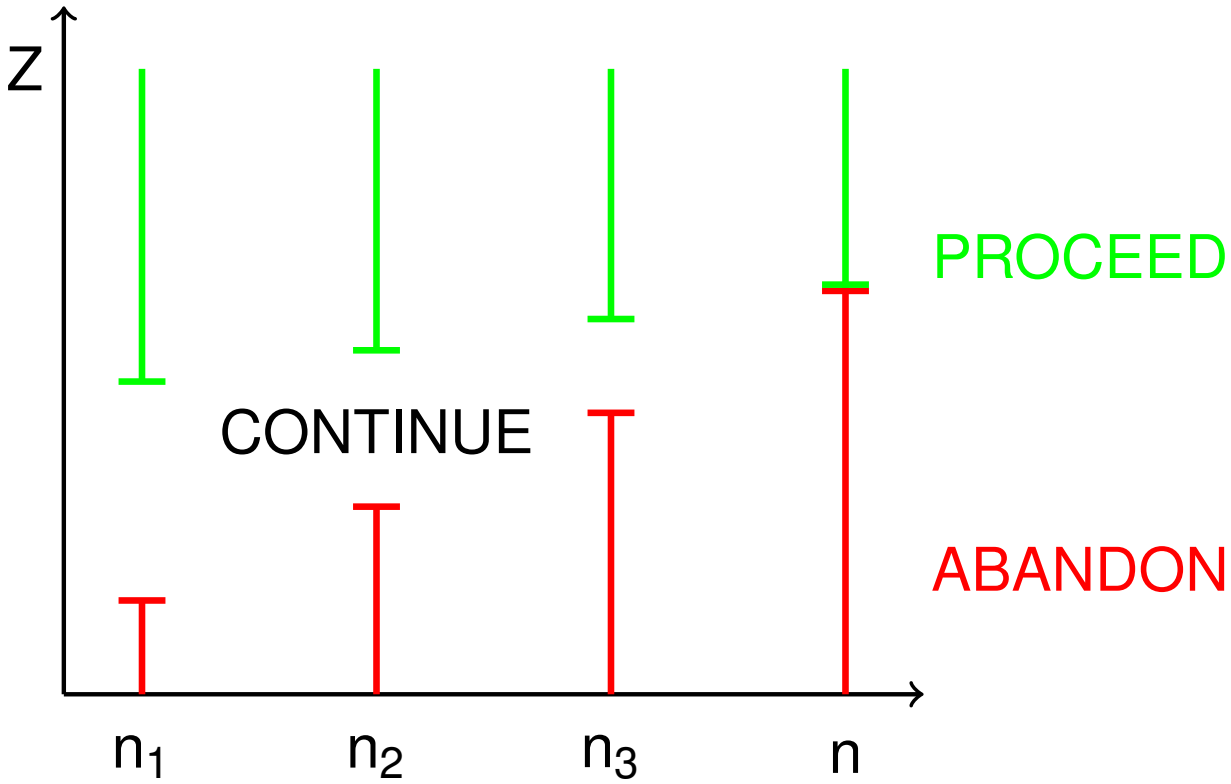
Simon's two-stage design



(Simon 1989)

Medical and Pharmaceutical
Statistics Research Unit

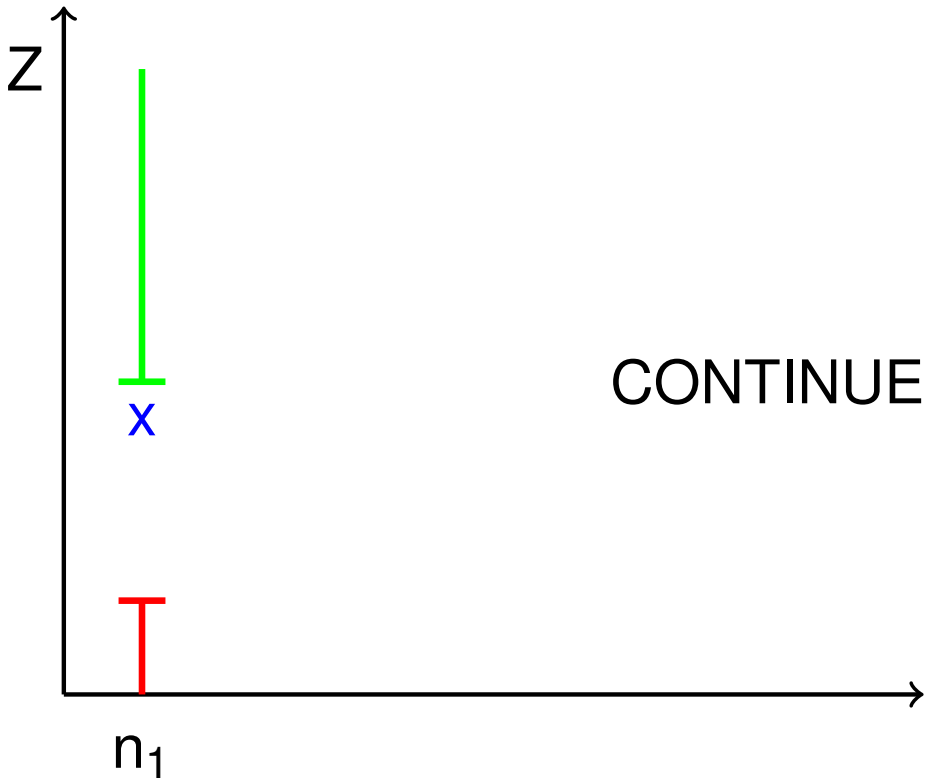
A multi-stage design



(Whitehead 1997)

Medical and Pharmaceutical
Statistics Research Unit

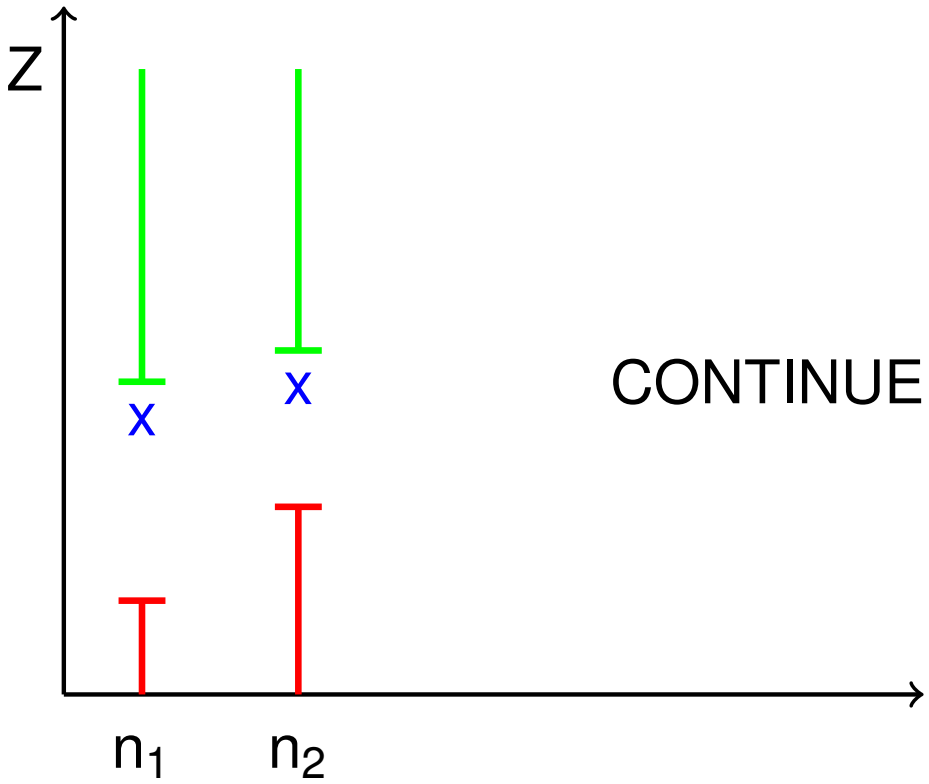
A multi-stage design



(Whitehead 1997)

Medical and Pharmaceutical
Statistics Research Unit

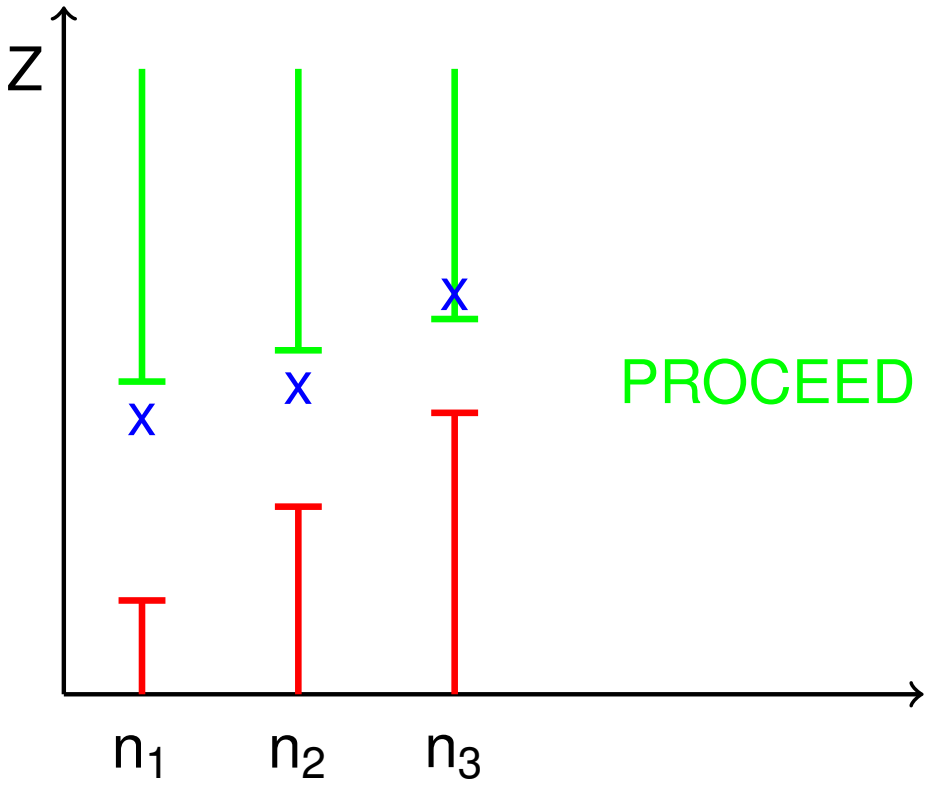
A multi-stage design



(Whitehead 1997)

Medical and Pharmaceutical
Statistics Research Unit

A multi-stage design



(Whitehead 1997)

Medical and Pharmaceutical
Statistics Research Unit

Example 1: Sample size re-estimation

Problem:

sample size formulae depend on nuisance parameters e.g., σ^2

Medical and Pharmaceutical
Statistics Research Unit

Example 1: Sample size re-estimation

Problem:

sample size formulae depend on nuisance parameters e.g., σ^2

Idea:

estimate the nuisance parameter(s) at interim and recalculate sample size

Example 1: Sample size re-estimation

Problem:

sample size formulae depend on nuisance parameters e.g., σ^2

Idea:

estimate the nuisance parameter(s) at interim and recalculate sample size

Methods:

Example 1: Sample size re-estimation

Problem:

sample size formulae depend on nuisance parameters e.g., σ^2

Idea:

estimate the nuisance parameter(s) at interim and recalculate sample size

Methods:

- for binary, Gaussian or survival outcomes

Medical and Pharmaceutical
Statistics Research Unit

Example 1: Sample size re-estimation

Problem:

sample size formulae depend on nuisance parameters e.g., σ^2

Idea:

estimate the nuisance parameter(s) at interim and recalculate sample size

Methods:

- for binary, Gaussian or survival outcomes
- exact (with unblinding) or approximate (without unblinding) (Gould 1995)

Medical and Pharmaceutical
Statistics Research Unit

Example 1: Sample size re-estimation

Problem:

sample size formulae depend on nuisance parameters e.g., σ^2

Idea:

estimate the nuisance parameter(s) at interim and recalculate sample size

Methods:

- for binary, Gaussian or survival outcomes
- exact (with unblinding) or approximate (without unblinding) (Gould 1995)
- negligible effect on type I error

Medical and Pharmaceutical
Statistics Research Unit

Sample size re-estimation

Example: DEVELOP-UK trial

- transplantation of reconditioned vs. standard donor lungs
- ex vivo lung perfusion (EVLP)
- phase III, multi-centre, unblinded, non-randomised, non-inferiority observational study
- primary endpoint: 12 months survival
- uncertainty in design parameters (only 50 transplants worldwide)

Medical and Pharmaceutical
Statistics Research Unit

Example 1: Sample size re-estimation
DEVELOP-UK trial

- 408 patients randomised to EVLP and standard
- 3:1 in favour of standard to ensure all available lungs are used
- interim analyses after 1/3 and 2/3 of total sample size
  - first: early stopping
  - second: early stopping, sample size re-assessment
- significance levels: 0.005 (first), 0.014 (second), 0.045 (final)

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

Dilemma:

low doses inefficient $\longleftrightarrow$ high doses toxic

Example 2: Phase I cancer trials

Dilemma:

low doses inefficient $\longleftrightarrow$ high doses toxic

Goal:

find a **maximum tolerated dose** (MTD)

where a prespecified proportion p_0 of patients

experience **dose-limiting toxicities** (DLTs)

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

Designs:

Example 2: Phase I cancer trials

Designs:

- 3 + 3 design (Carter 1973)
- . . . and variations (A + B, rolling 6, . . .)
- other 1- or 2-stage up-and-down designs
- accelerated titration design (Simon et al. 1999)
- continual reassessment method (O'Quigley et al. 1990)

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

Designs:

- 3 + 3 design (Carter 1973)
- . . . and variations (A + B, rolling 6, . . .)
- accelerated titration design (Simon et al. 1999)

} rule-based

- continual reassessment method (O'Quigley et al. 1990) — model-based

Medical and Pharmaceutical
Statistics Research Unit

Dose-finding / Phase I

Do NOT use rule based designs

Example 2: Phase I cancer trials

Continual reassessment method (CRM):

Example 2: Phase I cancer trials

Continual reassessment method (CRM):

- Bayesian adaptive design

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

Continual reassessment method (CRM):

- Bayesian adaptive design
- model-based

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

Continual reassessment method (CRM):

- Bayesian adaptive design
- model-based
- synthesizes prior belief about MTD and accumulating data

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

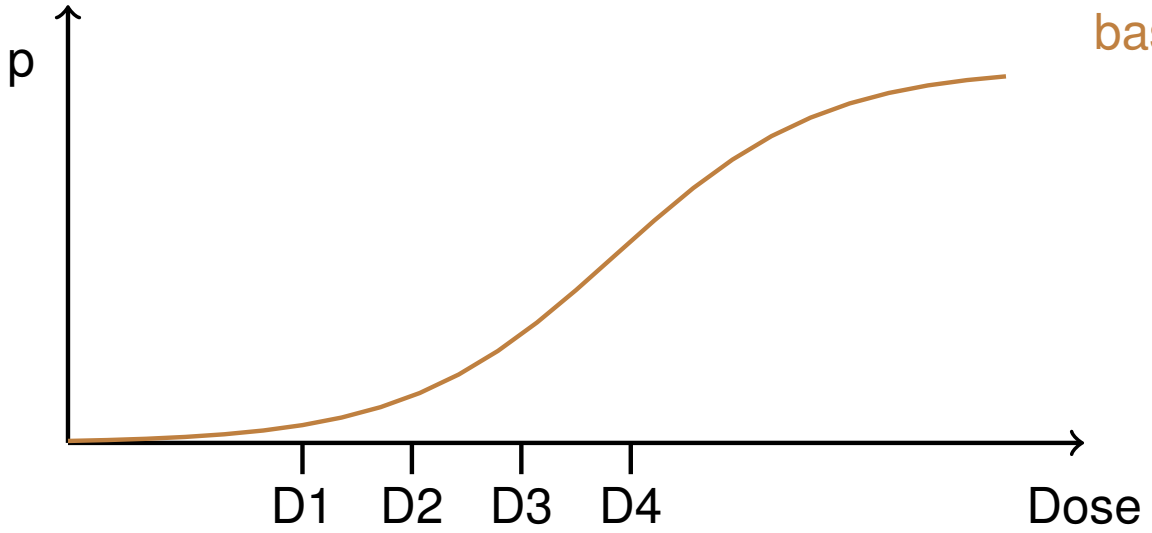
Continual reassessment method (CRM):

- Bayesian adaptive design
- model-based
- synthesizes prior belief about MTD and accumulating data
- uses information from **all** previous patients (“memory")

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

A CRM design:



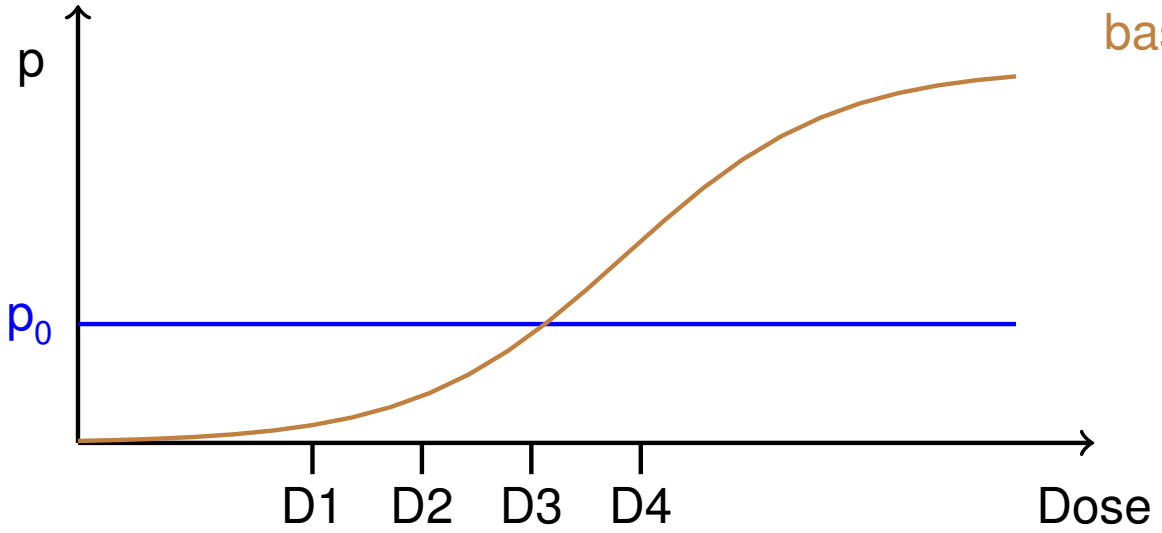
based on prior belief

p: proportion of DLTs

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

A CRM design:



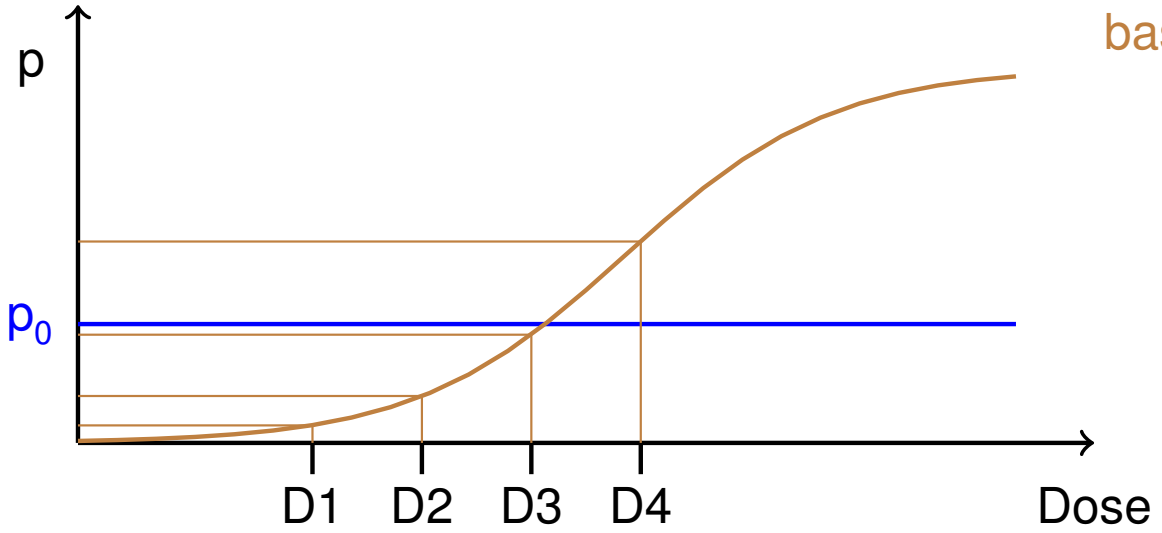
based on prior belief

p: proportion of DLTs

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

A CRM design:



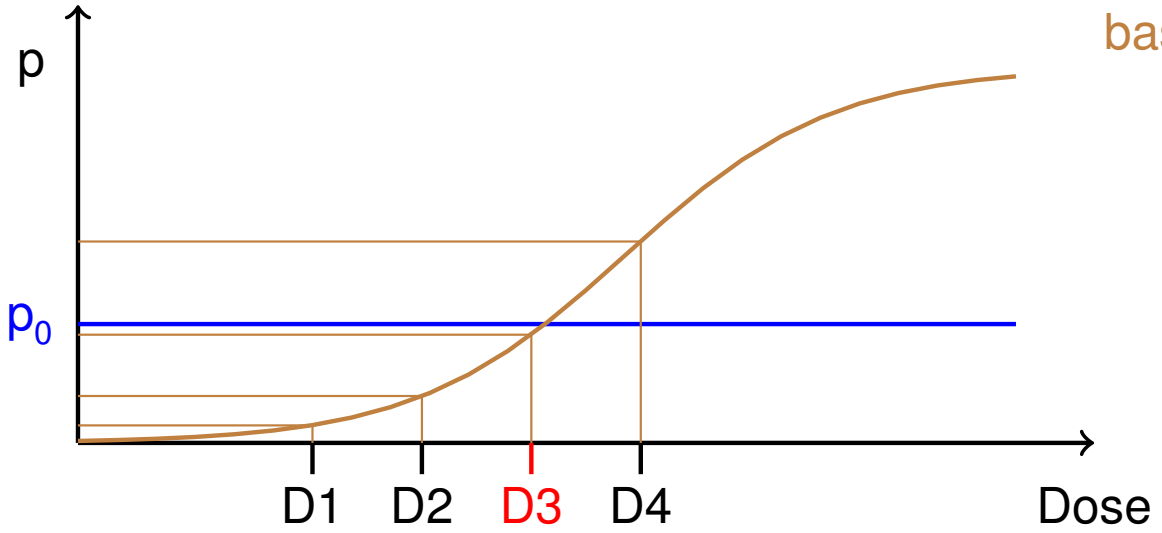
based on prior belief

p: proportion of DLTs

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

A CRM design:



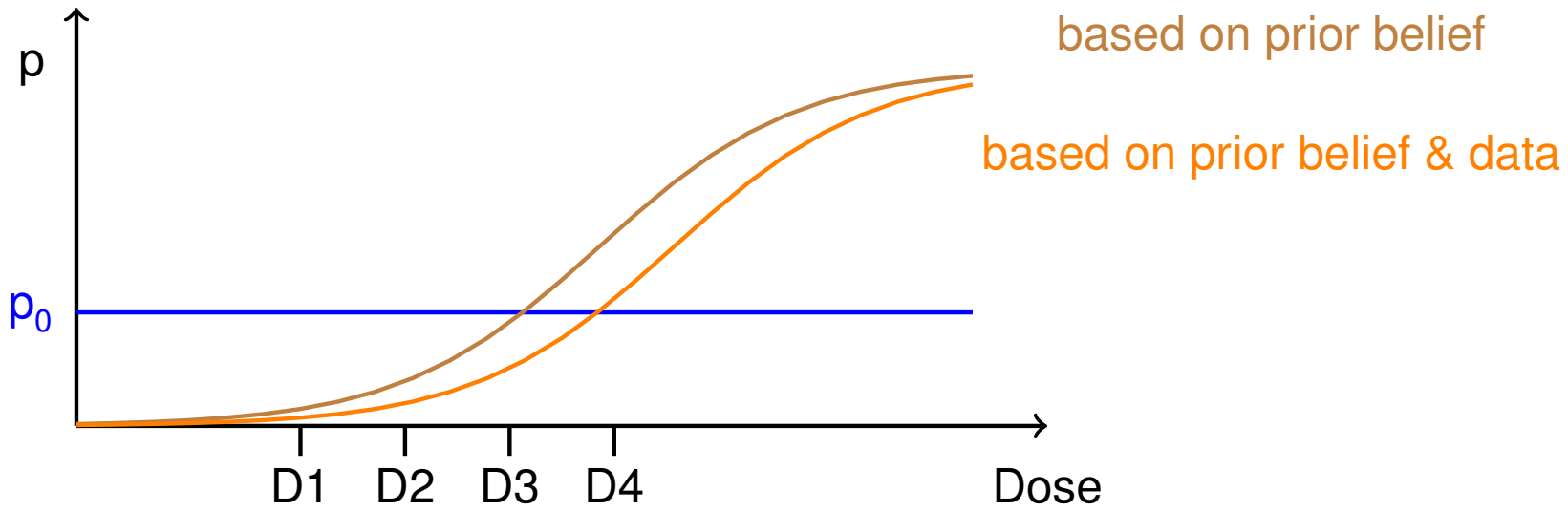
based on prior belief

p: proportion of DLTs

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

A CRM design:

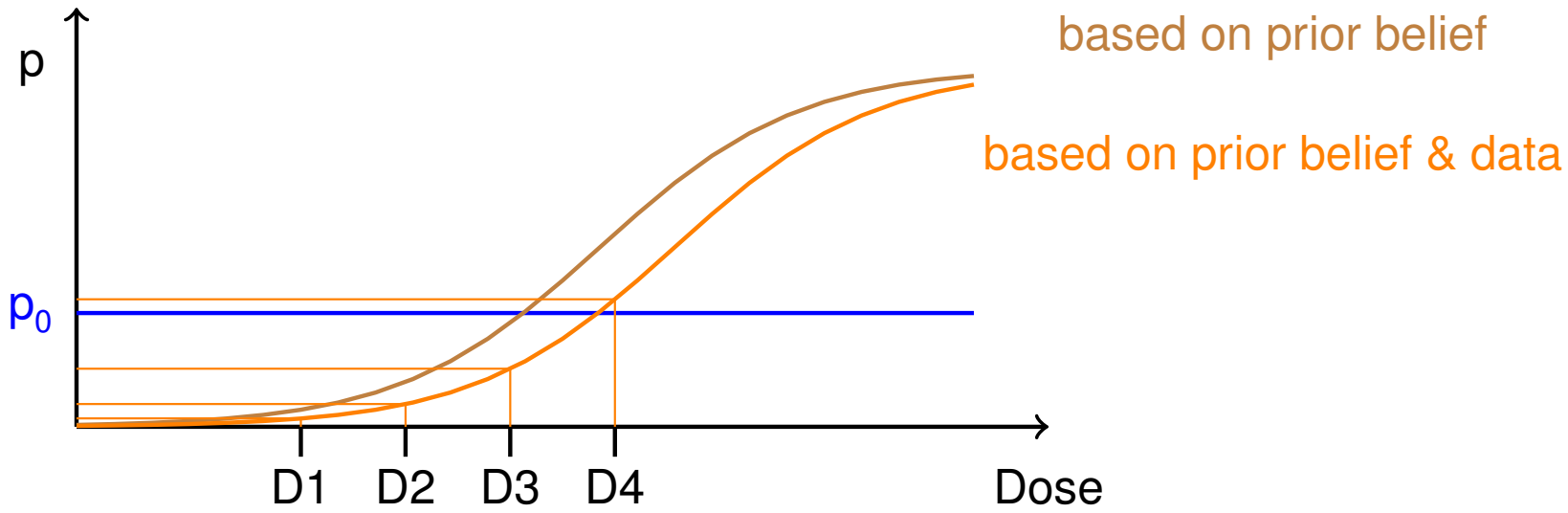


p: proportion of DLTs

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

A CRM design:

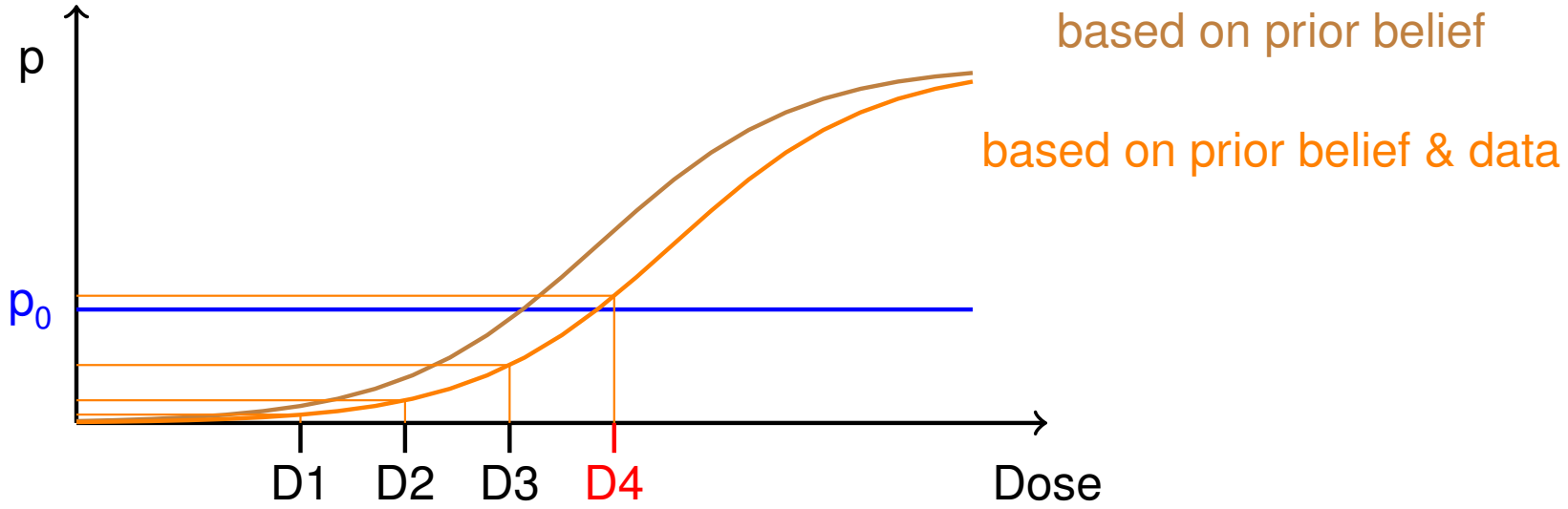


p: proportion of DLTs

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

A CRM design:



p: proportion of DLTs

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

CRM vs. 3 + 3:

- more likely to identify the correct dose
- more patients receive doses near MTD
- fewer overdoses
- clearly defined MTD
- MTD estimated with a measure of precision
- efficient use of all data
- flexible choice of toxicity levels
- easily extendable

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

Seamless phase I/II designs:

- phase I study “with a phase II flavour" (O’Quigley et al. 2001)

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

Seamless phase I/II designs:

- phase I study “with a phase II flavour" (O’Quigley et al. 2001)
- model toxicity and efficacy together

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

Seamless phase I/II designs:

- phase I study “with a phase II flavour" (O’Quigley et al. 2001)
- model toxicity and efficacy together
- identify the most successful dose (MSD) rather than MTD

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

Seamless phase I/II designs:

- phase I study “with a phase II flavour" (O’Quigley et al. 2001)
- model toxicity and efficacy together
- identify the most successful dose (MSD) rather than MTD
- efficient (patient population of interest in phase I)

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

Seamless phase I/II designs:

- phase I study “with a phase II flavour" (O’Quigley et al. 2001)
- model toxicity and efficacy together
- identify the most successful dose (MSD) rather than MTD
- efficient (patient population of interest in phase I)
- various methods using different models and stopping rules:

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

Seamless phase I/II designs:

- phase I study “with a phase II flavour" (O’Quigley et al. 2001)
- model toxicity and efficacy together
- identify the most successful dose (MSD) rather than MTD
- efficient (patient population of interest in phase I)
- various methods using different models and stopping rules:
  - efficacy-toxicity trade-off (Thall & Russell 1998, Thall & Cook 2004)

Medical and Pharmaceutical
Statistics Research Unit

Example 2: Phase I cancer trials

Seamless phase I/II designs:

- phase I study “with a phase II flavour" (O’Quigley et al. 2001)
- model toxicity and efficacy together
- identify the most successful dose (MSD) rather than MTD
- efficient (patient population of interest in phase I)
- various methods using different models and stopping rules:
  - efficacy-toxicity trade-off (Thall & Russell 1998, Thall & Cook 2004)
  - CRM-like (O’Quigley et al. 2001, Zohar & O’Quigley 2006)

Example 2: Phase I cancer trials

Seamless phase I/II designs:

- phase I study “with a phase II flavour" (O’Quigley et al. 2001)
- model toxicity and efficacy together
- identify the most successful dose (MSD) rather than MTD
- efficient (patient population of interest in phase I)
- various methods using different models and stopping rules:
  - efficacy-toxicity trade-off (Thall & Russell 1998, Thall & Cook 2004)
  - CRM-like (O’Quigley et al. 2001, Zohar & O’Quigley 2006)
  - decision procedure for continuous response (Zhou et al. 2006)

Medical and Pharmaceutical
Statistics Research Unit

Example 3: Multi-arm multi-stage trials (MAMS)

TAILoR trial (Pushpakom et al. 2015)

- Phase II study to evaluate treatment for side effect of HIV tri-regimen treatment (TAILoR)
- Superiority trial
- Several possible doses
- Normal distributed endpoint

Medical and Pharmaceutical
Statistics Research Unit

Example 3: Multi-arm multi-stage trials (MAMS)

Design (Magirr et al, 2012)

- 3 active arms compared to control
- Equal randomization between actives and control
- one interim analysis
  - Stop for superiortiy
  - Stop if no active arm appears promising
  - Drop any active arms that are not sufficiently promising
- In the trial we underestimated drop-out and could adjust at interim

Medical and Pharmaceutical
Statistics Research Unit

General considerations

- New methods either get results faster or are more accurate
- Tend to take more time to develop upfront
  - How can this cost be covered?
- Are the ideas sensible to implement?
  - Is an interim analysis feasible?
  - Can Bayesian methods be updated online?