MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

Adaptive Methods in Clinical Research

Lecture 6: Estimation after Adaptive Designs

David Robertson

MRC Biostatistics Unit

1st November 2023

Overview

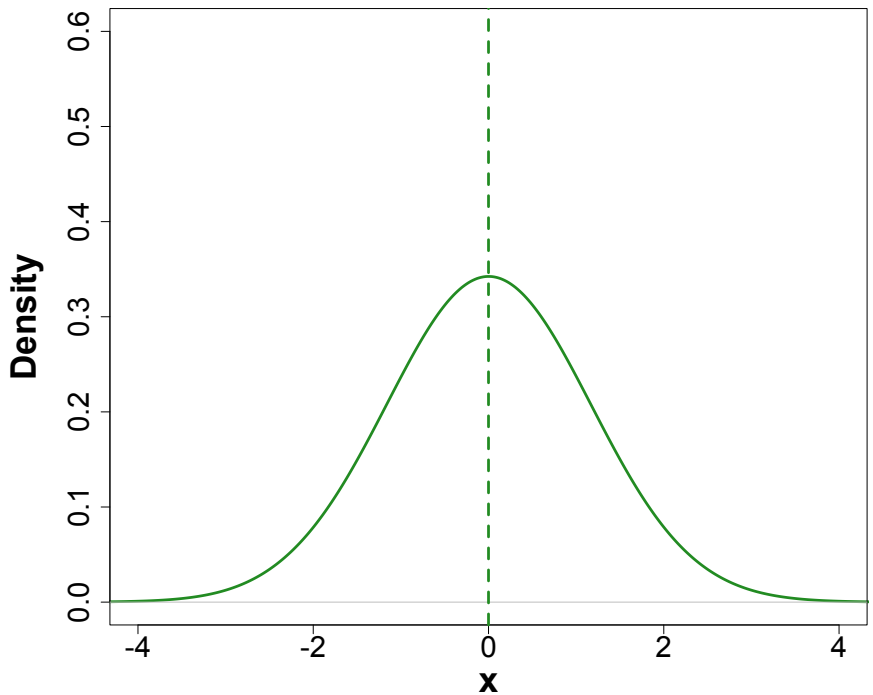
1. What are the issues with estimation after ADs?
2. Unbiased and bias-reduced point estimation
3. Adjusted confidence intervals
4. Practical aspects and guidelines

Motivation

- Appropriate estimation of treatment effects is an important part of trial validity
- Key issues for adaptive clinical trials is that the usual end-of-trial estimators can be prone to bias and confidence intervals (CIs) can have incorrect coverage (as well as other problems – see later)
  - Bias = *"A systematic tendency for the estimate of treatment effect to deviate from its true value" (FDA)*
- The question of estimation of treatment effects in an adaptive clinical trial has received comparatively less attention

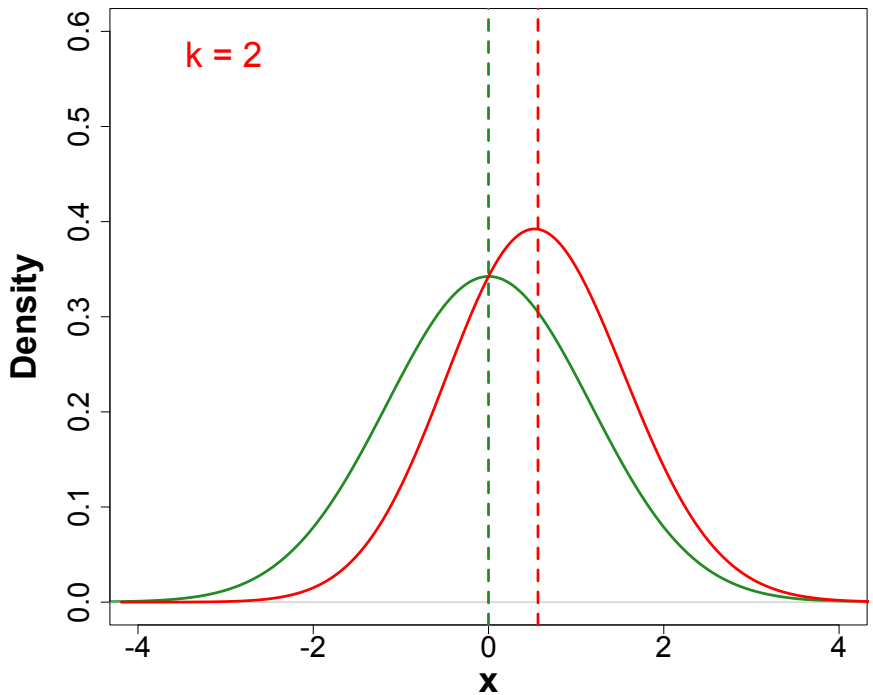
Motivation – point estimation

- Suppose outcome measure of experimental treatment follows a standard normal distribution
- What happens when we select the best-performing (i.e. treatment with highest observed mean) of *k* such treatments?



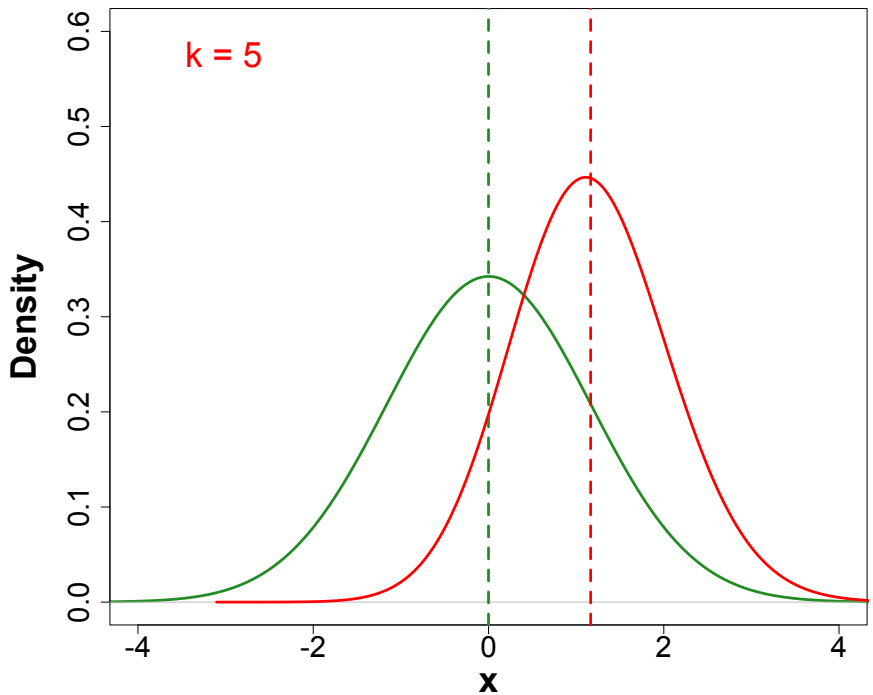
Motivation – point estimation

- Suppose outcome measure of experimental treatment follows a standard normal distribution
- What happens when we select the best-performing (i.e. treatment with highest observed mean) of *k* such treatments?



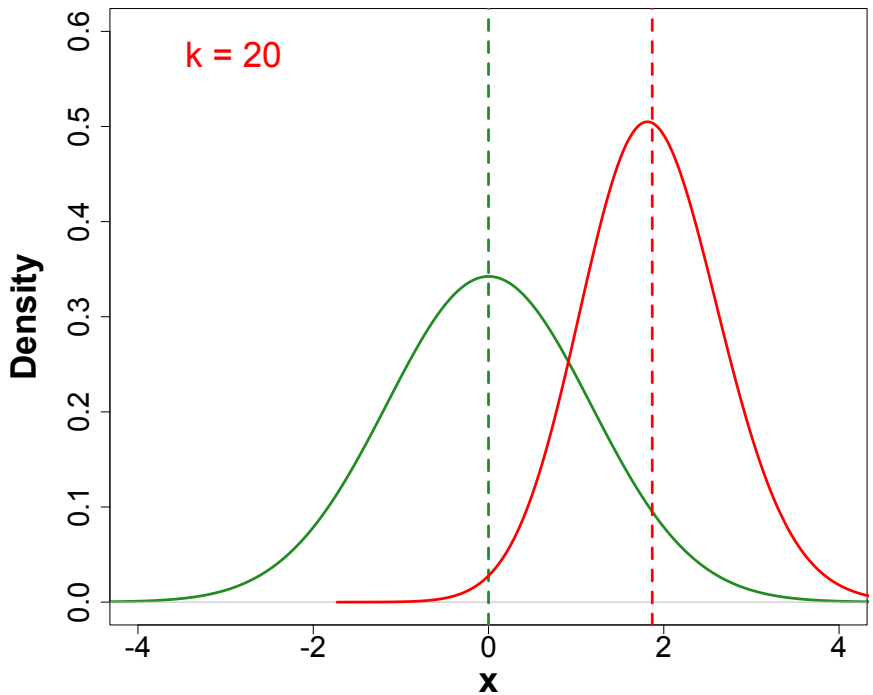
Motivation – point estimation

- Suppose outcome measure of experimental treatment follows a standard normal distribution
- What happens when we select the best-performing (i.e. treatment with highest observed mean) of *k* such treatments?



Motivation – point estimation

- Suppose outcome measure of experimental treatment follows a standard normal distribution
- What happens when we select the best-performing (i.e. treatment with highest observed mean) of *k* such treatments?



Motivation – CIs

- Consider a group sequential trial with $J = 5$ equally-spaced analyses
- Coverage probabilities for standard/naive 90% CIs for Pocock and O'Brien-Fleming boundaries:

θ	Pocock	O'Brien-Fleming
0.0	0.881	0.907
0.5	0.846	0.878
1.0	0.869	0.883
1.5	0.891	0.892
2.0	0.918	0.914

Motivation

What are the issues with biased estimation and incorrect CIs?

- Reporting substantially biased estimates for a primary outcome measure following an adaptive design can result in poor decisions
- Concern about over- or under-estimation of treatment effects affecting further research
- Impact on health economic analyses

Estimation after adaptive designs

- Question: how to best construct point estimates and CIs for the treatment effect in adaptive trials
- Lack of established methods – still an active area of research

Unbiased and bias-reduced point estimation

- Lots of methods to remove or reduce the bias in the MLE
- Two broad classes:
  1. Unbiased estimators
  2. Bias-reduced estimators
- Also distinguish been *conditionally* and *unconditionally* unbiased estimators:
  - *Unconditionally* unbiased if it is unbiased when averaged across all possible realizations of an adaptive trial
  - *Conditionally* unbiased if it is unbiased only conditional on the occurrence of a subset of trial realizations
  - E.g. one might be interested in an estimator only conditional on a particular arm being selected at an interim analysis

Bias and MSE

- Mean bias of an estimator $\hat{\theta}$ of the parameter θ is

$$\text{bias}(\hat{\theta}) = \text{E}(\hat{\theta}) - \theta$$

- An estimator that has mean bias identically equal to 0 (for all possible values of $\theta \in \Theta$) is said to be *mean-unbiased*

- The *mean squared error* (MSE) is

$$\text{mse}(\hat{\theta}) = \text{E}(\hat{\theta} - \theta)^2$$

UMVUE

- Sensible way of choosing which unbiased estimator to use is to find the unbiased estimator with the smallest possible variance
  - Equivalently, this results in an unbiased estimator with the smallest possible MSE
- Such an estimator is called the *uniformly minimum variance unbiased estimator* (UMVUE)
- Use the Lehmann-Scheffé theorem:
  - Suppose S is a minimal sufficient and complete statistic, and U is an unbiased estimator
  - Then the estimator $\hat{U} = E(U|S)$ is the (unique) UMVUE

UMVCUE

- Consider a multi-arm trial which incorporates *treatment selection*
  - e.g. Treatments are ranked according to some rule and the best treatment(s) are taken forward to the next stage
- In this setting, it is more appropriate to additionally *condition* on the selection used
- This gives the *uniformly minimum variance conditionally unbiased estimator* (UMVCUE)
- Note that the UMVCUE is unbiased unconditionally as well

Median unbiased estimator

- So far, have been considering *mean* unbiasedness (the usual unbiased property)
- A reasonable alternative is to consider *median* unbiasedness
- An estimator θ_{MU} is median unbiased if $Pr(\theta_{MU} \leq \theta) = Pr(\theta_{MU} \geq \theta)$
  - i.e. the estimator underestimates just as often as it overestimates
- Median-unbiased estimators may be easier to construct than mean-unbiased estimators
- Median-unbiased estimators are invariant under one-to-one transformations (unlike mean-unbiased estimators)

Unbiased vs bias-reduced estimators

Why not always use an unbiased estimator?

1. Not (yet) available for all classes of adaptive designs
2. Increased MSE (bias-variance trade-off) leading to large SEs and wide CIs

Bias-adjusted MLE

- Whitehead (1986) proposed a bias-adjusted MLE, which adjusts the MLE $\hat{\theta}_{\text{MLE}}$ to remove bias
- Bias for the MLE as a function of the unknown parameter of interest θ is $b(\theta) = E(\hat{\theta}_{\text{MLE}}|\theta) - \theta$
- Bias-adjusted MLE $\hat{\theta}_{\text{BC}}$ is the numerical solution to

$$\hat{\theta}_{\text{BC}} = \hat{\theta}_{\text{MLE}} - b(\hat{\theta}_{\text{BC}})$$

Review of point estimation for adaptive trials

Received: 25 June 2021 | Revised: 21 October 2022 | Accepted: 1 November 2022
DOI: 10.1002/sim.9605

RESEARCH ARTICLE

Statistics in Medicine WILEY

Point estimation for adaptive trial designs I: A methodological review

David S. Robertson[1] | Babak Choodari-Oskooei[2] | Munya Dimairo[3] | Laura Flight[3] | Philip Pallmann[4] | Thomas Jaki[1,5]

Received: 25 June 2021 | Revised: 20 January 2023 | Accepted: 18 March 2023
DOI: 10.1002/sim.9734

RESEARCH ARTICLE

Statistics in Medicine WILEY

Point estimation for adaptive trial designs II: Practical considerations and guidance

David S. Robertson[1] | Babak Choodari-Oskooei[2] | Munya Dimairo[3] | Laura Flight[3] | Philip Pallmann[4] | Thomas Jaki[1,5]

- Provides a comprehensive overview of proposed approaches for different classes of adaptive designs, with available software
- Illustrates the computation of different estimators in practice using a real trial example
- Proposes a set of guidelines for researchers around the choice of estimators and the reporting of estimates following an adaptive design

Case study

- Two-arm RCT of oral cannabis extract vs placebo in adults with MS
- Primary endpoint: relief from muscle stiffness (yes/no) after 12 weeks
- Two-stage group-sequential design (maximum $n = 300$) with interim analysis (after $n = 200$)
  - O'Brien-Fleming efficacy stopping boundaries
  - Trial continued to the final analysis

	Interim data		Final data	
	Placebo	**CE arm**	**Placebo**	**CE arm**
Number of subjects with relief from muscle stiffness	12	27	21	42
Total number of subjects	97	101	134	143
Standardized test statistic	2.540		2.718	
OBF boundary	2.797		1.977	

Case study

Type of estimator	Estimator	Difference in proportions (SE)	Relative difference to overall MLE
MLE/naive	MLE (overall)	0.1370 (0.054)	–
Unconditionally unbiased/bias-adjusted	MLE (stage 1)	0.1436 (0.057)	+5%
	Median unbiased estimator (MUE)	0.1341 (0.054)	−2%
	UMVUE	0.1278 (0.054)	−7%
	Bias-corrected MLE (UBC-MLE)	0.1328 (0.055)	-3%
Conditionally unbiased/bias-adjusted	MLE (stage 2)	0.1139 (0.111)	−17%
	Conditional MUE (CMUE)	0.1851 (0.080)	+35%
	UMVCUE	0.1724 (0.071)	+26%
	Bias-corrected MLE (CBC-MLE)	0.1909 (0.073)	+39%

Desirable criteria for CIs

1. Correct coverage probability (validity) [Essential!]
2. Is an interval (i.e. not disjoint)
3. Narrower CIs are to be preferred
4. Consistent/compatible with the decision rule
5. Contains the point estimate of interest
6. Computational feasibility/simplicity

Classes of CIs

- Repeated CIs: can be calculated at any stage, not tied to a pre-specified stopping rule, but generally conservative (i.e. wide) and may not always contain the MLE
- Exact/monotone CIs: can only be calculated at the stage a trial stops according to a pre-specified stopping rule, narrower than repeated CIs
- Simultaneous CIs: ‘multiplicity-adjusted’ individual CIs

Methods to construct adjusted CIs

- Constructing an approximate pivotal quantity, e.g. using asymptotic normality
- Inverting a test statistic, i.e. exploiting the duality between confidence set and hypothesis test
- Bootstrap/resampling techniques
- Hybrid approaches, e.g. using different methods for lower and upper confidence bound

Example: Group sequential designs

Exact confidence intervals

1. Define an ordering of design space w.r.t evidence against $H_0 : \theta = 0$

2. Define a *p*-value function $P(\theta)$: gives probability that, at the stage the trial stopped, even more extreme evidence against H_0 could have been observed

3. At the point the trial stops, a $100 \times (1 - \alpha)\%$ confidence interval for θ is given by $(\hat{\theta}_{lb}, \hat{\theta}_{ub})$ where:

$$P(\hat{\theta}_{lb}) = \frac{\alpha}{2} \quad \text{and} \quad P(\hat{\theta}_{ub}) = 1 - \frac{\alpha}{2}$$

Ordering the design space

- Stage-wise ordering most common
- For two test statistics (z_j, z_k), $P_{z_j}(\theta) \leq P_{z_k}(\theta)$ if:
  - $j < k$ and $z_j \geq e_j$
  - or if $j = k$ and $z_j \geq z_k$
  - or if $j > k$ and $z_k \leq f_k$

Example: Group sequential designs

Repeated confidence intervals

- Proposed by Jennison and Turnbull, 1989
- The RCIs for a parameter θ of interest are a sequence of confidence intervals C_j, $j = 1 \ldots, k$ where

$$P(\theta \in C_j \text{ for stages } j = 1, \ldots, k) \geq 1 - \alpha$$

  for all θ

- Can be computed at any stage, whether the trial stops or not, i.e. is valid regardless of how the decision to stop the study was reached

Proposed guidance

- The issue of estimation should be considered throughout the whole lifecycle of an adaptive trial
- The design and analysis of an adaptive trial are closely linked, and one should not be considered without the other
- Our main focus is on the confirmatory setting where analyses are fully pre-specified
- Builds on the relevant parts of the FDA and EMA guidance for adaptive designs and the adaptive designs CONSORT extension

Proposed guidance

Planning stage

- Context, aims and design of an adaptive trial should all inform the analysis strategy used, which includes the choice of estimators and CIs
- Decide on what exactly is to be estimated (i.e. the estimands of interest)
- Decide on the desired characteristics of estimator:
  - Conditional versus unconditional perspective
  - Bias-variance trade-off
- A review of the literature may be sufficient
- Otherwise, conduct simulations to explore

Proposed guidance

Pre-specification of analyses

- Statistical analysis plan (SAP) and health economic analysis plan (HEAP) should include a description and justification of the estimators and CIs used
- When available, unbiased or bias-reduced estimators should be used and reported alongside the standard MLE, as well as adjusted CIs
- If multiple adjusted estimators are available and are of interest, one adjusted estimator should be designated the‘primary’ adjusted estimator
  - Others included as sensitivity or supplementary analyses

Data Monitoring Committees

- When presenting interim results to DMCs, the issue of potential bias and incorrect CIs should also be considered
- The sensitivity of the standard MLE and CI (based on the interim data) to potential bias and incorrect coverage should be reported
- When unbiased or bias-reduced estimators are available, these should also be presented to the DMC

Proposed guidance

Reporting results for a completed trial

- There should be a clear description of the statistical methods used to estimate treatment effects
- Adjusted estimates taking the trial design into account are to be preferred
- FDA guidance: *"if naive estimates such as unadjusted sample means are used, the extent of bias should be evaluated, and estimates should be presented with appropriate cautions regarding their interpretation"*
- EMA guidance: *"methods to . . . provide confidence intervals with pre-specified coverage probability are required"*

Final thoughts

- Point estimates and CIs need adjusting after an adaptive design
- Estimation for flexible adaptive designs (e.g. with non-binding decision rules or unplanned adaptations) remains an open problem
- Methods for constructing CIs for some settings are limited and may not be particularly informative
- Further research is needed!

References

Dimairo M et al. (2020)
The Adaptive designs CONSORT Extension (ACE) statement: a checklist with explanation and elaboration guideline for reporting randomised trials that use an adaptive design.
BMJ, doi:10.1136/bmj.m115

European Medicines Agency (2007)
Reflection paper on methodological issues in confirmatory clinical trials planned with an adaptive design
https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-methodological-issues-confirmatory-clinical-trials-planned-adaptive-design_en.pdf

US FDA (2019)
Adaptive Designs for Clinical Trials of Drugs and Biologics: Guidance for Industry.
https://www.fda.gov/media/78495/download

Marschner IC (2021)
A General Framework for the Analysis of Adaptive Experiments
Statistical Science, 36(3):465–492.

Robertson DS et al. (2023a)
Point estimation for adaptive trial designs I: A methodological review
Statistics in Medicine, 42(2):122–145

Robertson DS et al. (2023b)
Point estimation for adaptive trial designs II: Practical considerations and guidance
Statistics in Medicine, 42(14):2496–2520