Innovative trial design and effect size estimation

Bias, de-biasing, and when it is considered to be important

Ursula Garczarek

Sepul Bio a Business Unit of Thea

Innovative designs for pivotal trials and effect size estimation

Bias, de-biasing, and ~~when~~ why it is considered to be important

Ursula Garczarek

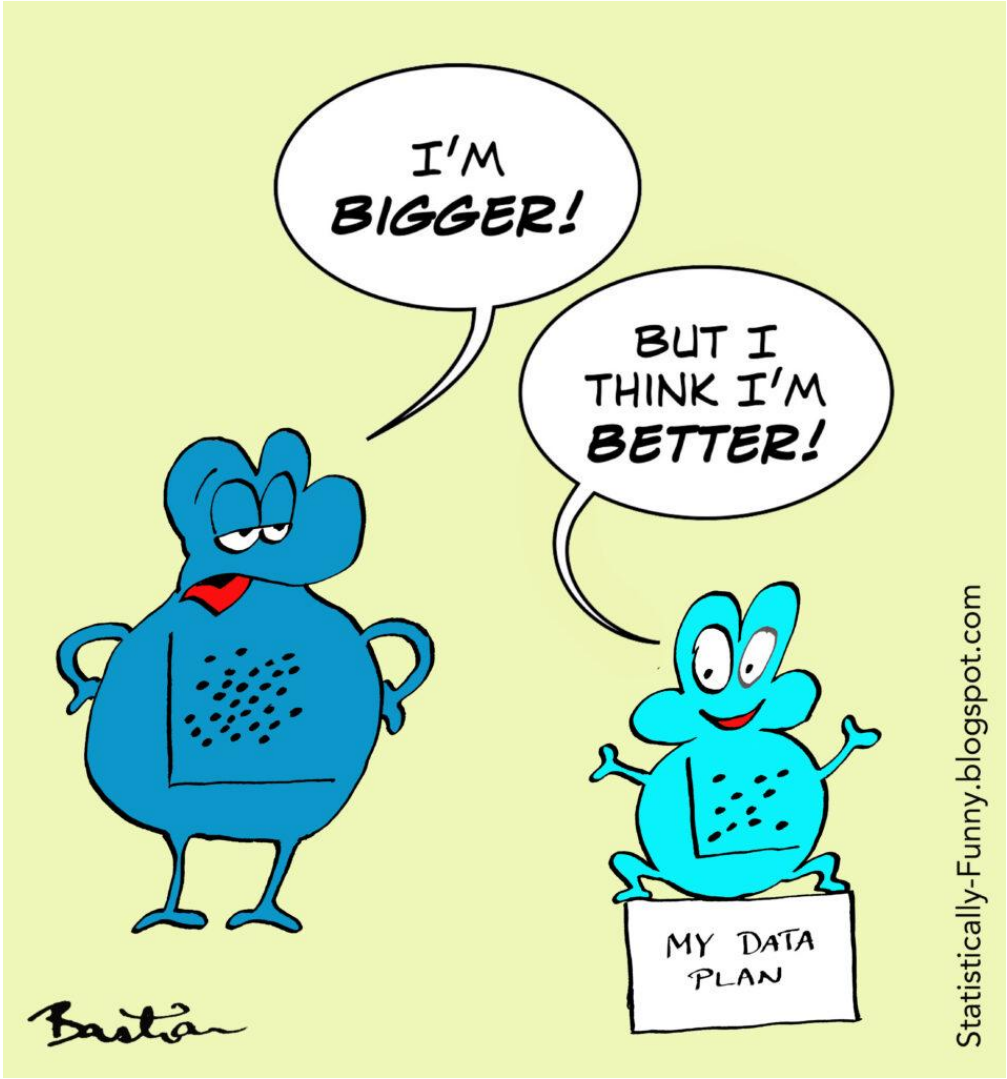
Sepul Bio a Business Unit of Thea

Disclaimer

The authors alone are responsible for the views expressed in this publication, and those views do not necessarily represent the decisions, policies or views of their respective institutions or companies.

Overview

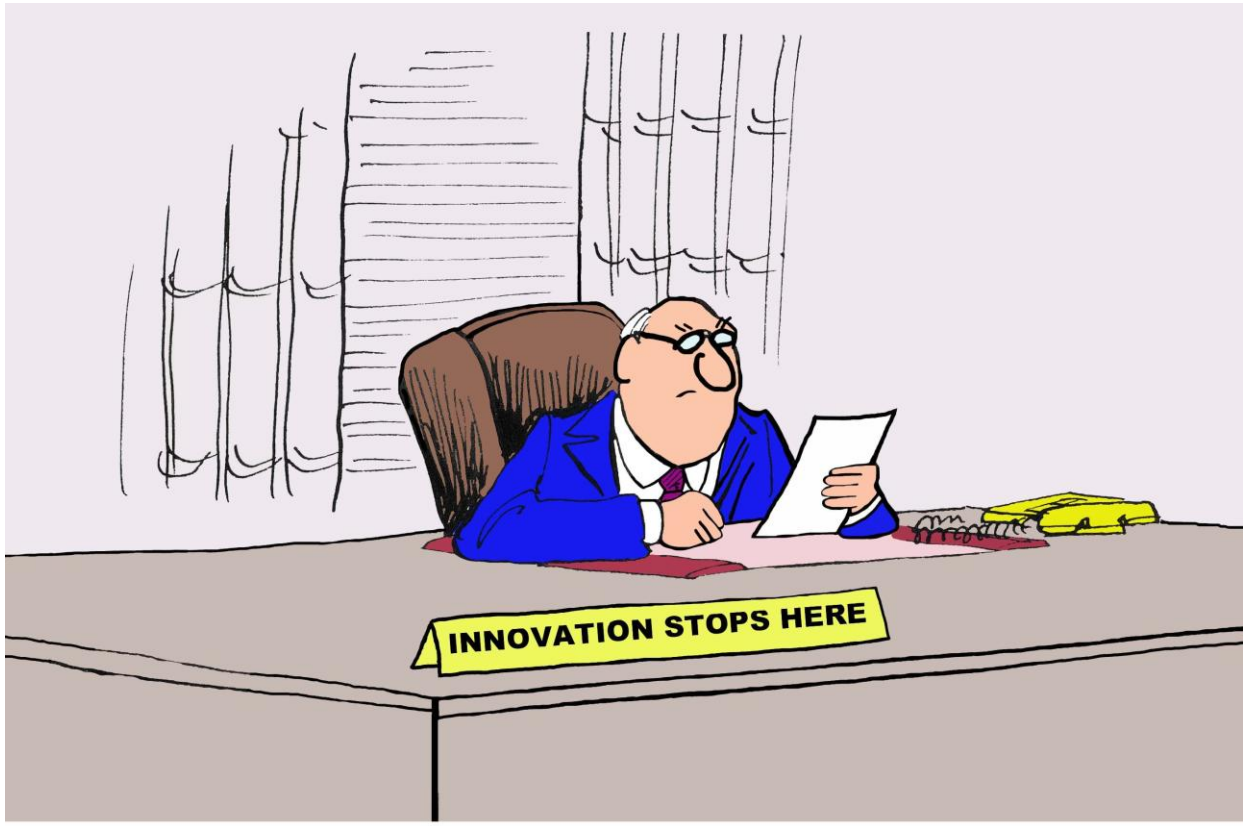
- Setting the frame
  - „Innovative“ Trial Designs for pivotal trials
  - „Bias“ in Effect Size Estimation
- Relevance
- Bias of the Estimator
  - Adaptive Clinical Trials
- Bias compared to Parallel Arm Design
  - Within-Subject Trials
- Conclusions



Hilda Bastian, Absolutely Maybe
http://creativecommons.org/licenses/by-nc-nd/4.0/

Innovative Trial Design

- Scope of this talk:
  - Established designs applied in pivotal trials
    - Trial designs with good type-1-error control
  - Innovative ~ not a 1:1 parallel design with well known clinical efficacy endpoint(s)
- Why so limited?
  - Without type-1-error control biased effect size estimation is to be expected



Bias of an estimator of effect size

Bias of an estimator:

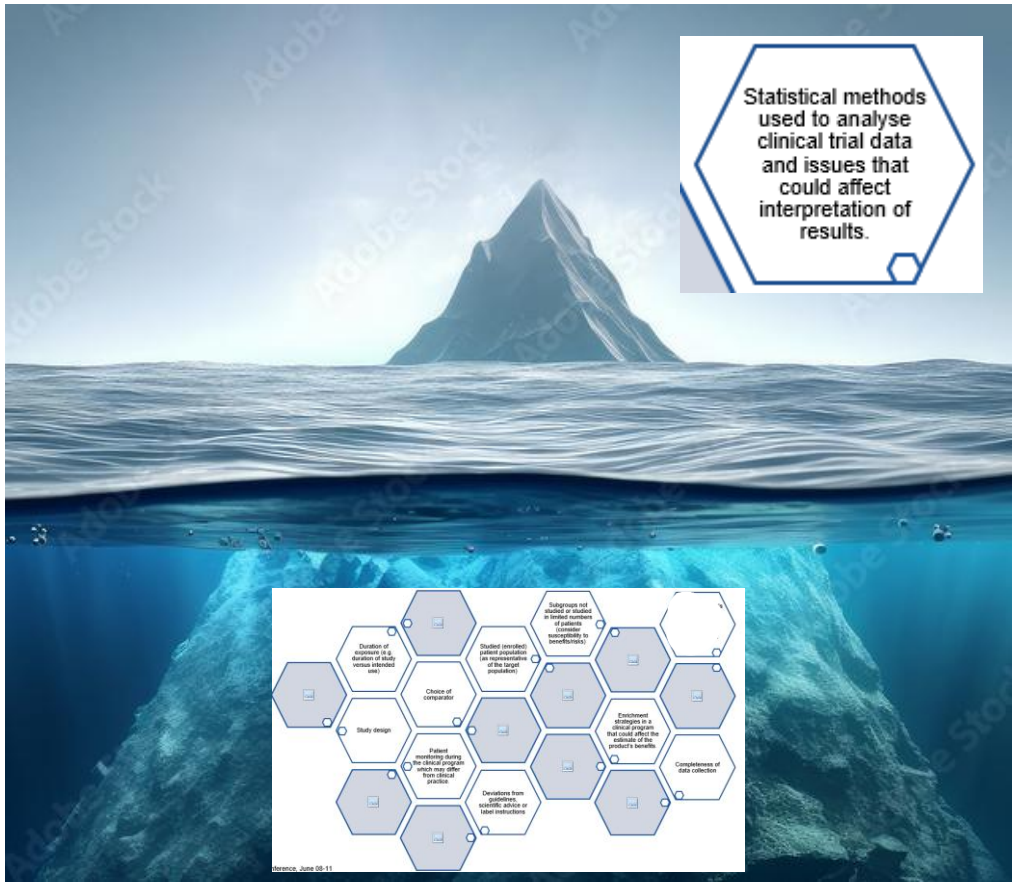
- Let *X:* Ω→ $\mathbb{E}$ be a random variable described by the probability space:

$$(\Omega,\Sigma,(P_\theta, \theta \in \Theta)),$$

- Let T = T(*X*) be a statistical point estimator of a function f(θ).
- Assume the expectation $E_\theta\{T\}$ exists.

$$b(\theta) = E_\theta\{T\} - f(\theta)$$

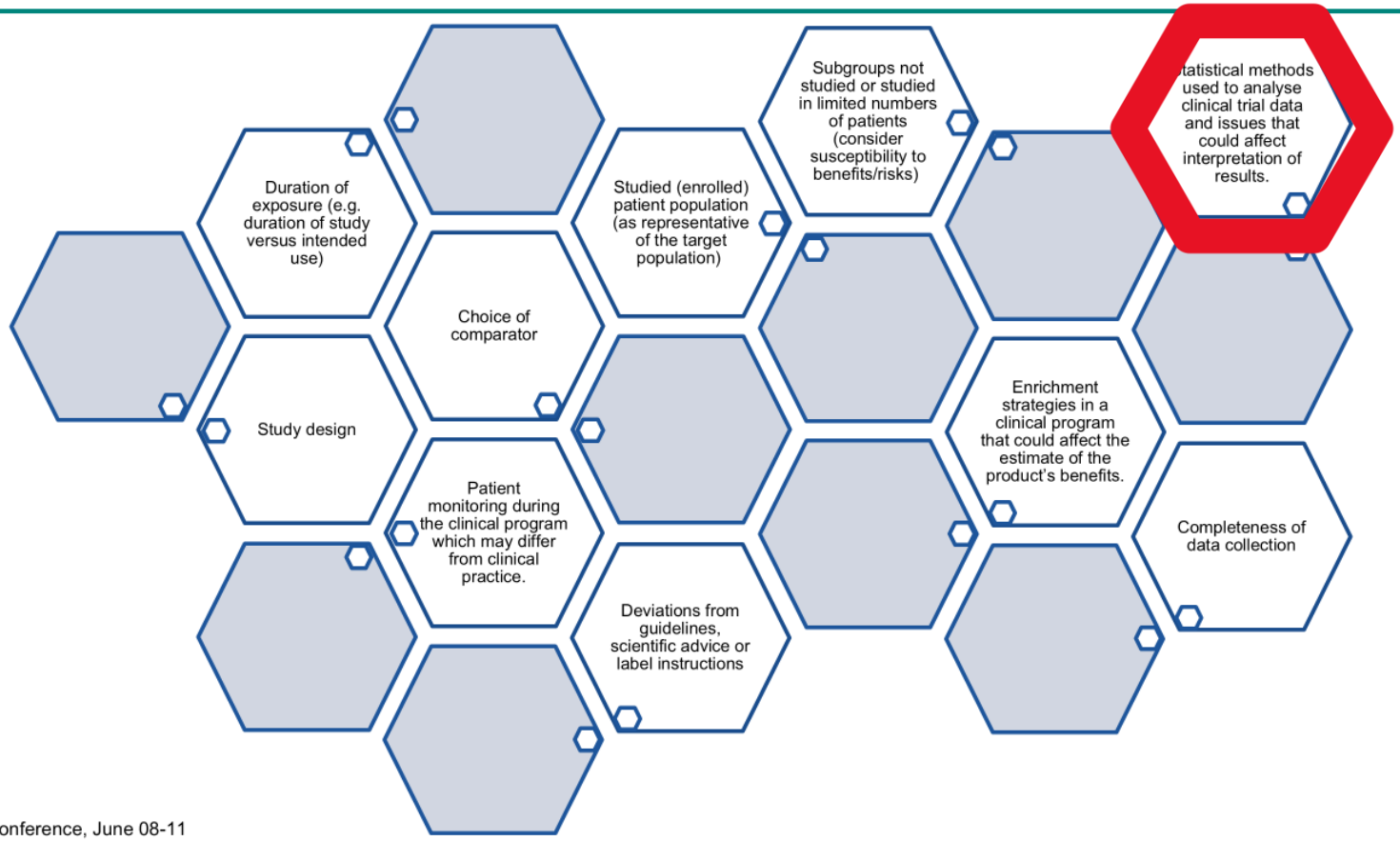
- If b(θ) ≠ 0 → T is a biased estimator of f(θ).



Relevance: Benefit-Risk Framework

CIOMS Working Group XII

Sources of Uncertainties in SBRF



Relevance: FDA Guidance B/R

Examples of Important Considerations for FDA's Premarket Benefit-Risk Assessment of NDAs, BLAs, and Efficacy Supplements

The right "X: $\Omega \rightarrow \mathbb{E}$ " over $(\Omega, \Sigma, (P_\theta, \theta \in \Theta))$

- Clinical relevance of the study endpoints: importance to patients?
- Nature of the effect (e.g., survival, slowing disease progression, …)

The right "Ω" in $(\Omega, \Sigma, (P_\theta, \theta \in \Theta))$

- Defined subpopulations achieving a greater magnitude of benefit, having a greater need
- Generalizability of the demonstrated benefits to all populations likely to be prescribed the drug (e.g., older patients or patients with co-morbidities not extensively studied in the clinical trials)

The right "P_θ" in $(\Omega, \Sigma, (P_\theta, \theta \in \Theta))$

- The distribution of treatment effects in the clinical trial population
- Benefit attributed to the drug when studied in combination with other therapies

The right function $f(\theta)$

- Time course and durability of effect

The right function $T(X)$

- **Effect size and associated uncertainty**

Naive estimators in adaptive designs – conditional and unconditional bias

Unconditional bias

$$b(\theta) = E_\theta\{T\} - f(\theta) \neq 0$$

Conditional bias

$$b(\theta|STAGE) = E_\theta\{T|STAGE\} - f(\theta) \neq 0$$

Does it matter?

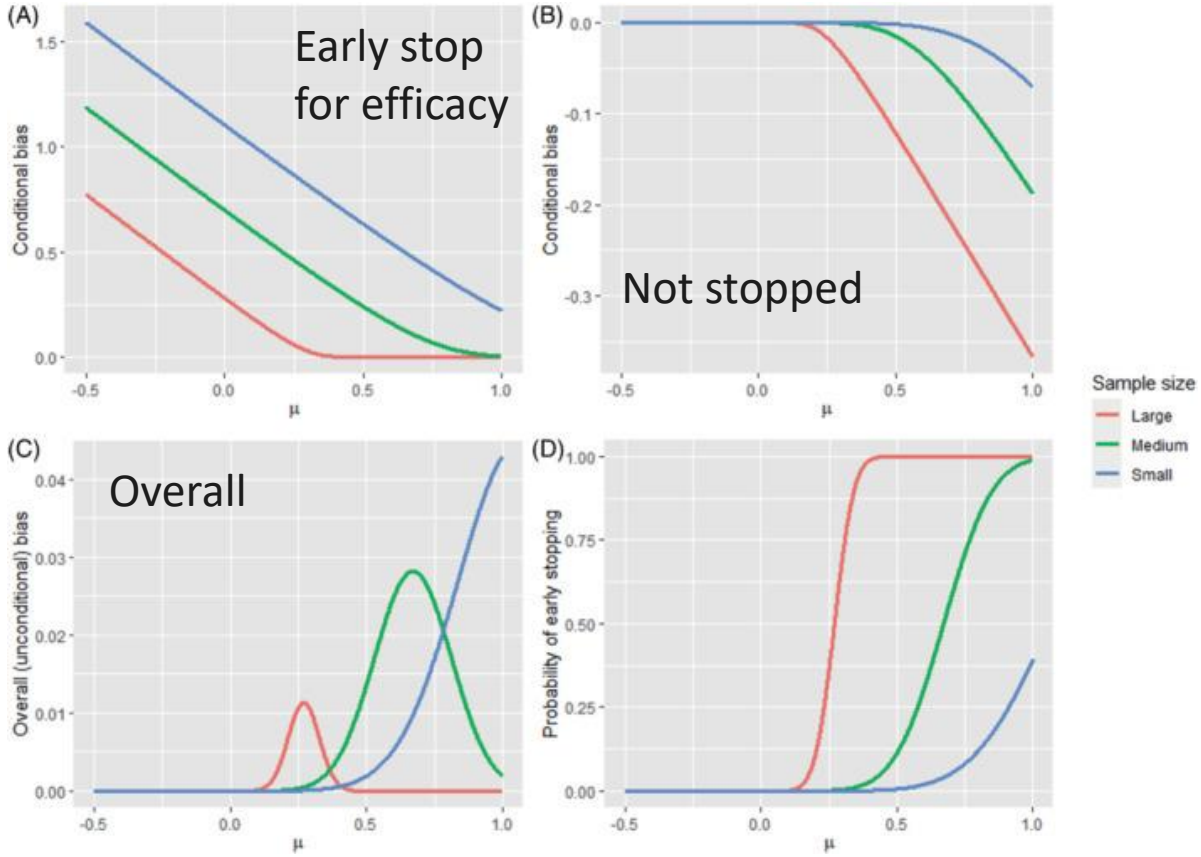
Evidence synthesis:

- No, when many trials
- Yes, drop the loser

Reimbursement:

- Underestimation may harm cost-benefit analyses for treatments with no early stop for efficacy

Two step group-sequential design, SD=1



Innovative trial designs and de-biased estimators (Robertson et al 2021 Part I)

- De-biased estimation often available but rarely used
- Beware of variance-bias tradeoff
- Resampling methods are well suited and flexible

Extract Table 3 for Group Sequential Designs

Method(s)	Pros and Cons	R packages or Literature
Mean-unbiased estimation	UMVUE has zero bias; higher MSE than naive or bias-adjusted estimators; may be conditionally biased.	R package OptGS
Median-unbiased estimation	Reduces bias; increased MSE; dependent on sample space ordering. Can be applied to adaptive group sequential designs	rpact, RCTdesign, AGSDesign, OptGS
Resampling	Bias substantially reduced with reasonable MSE	Whitehead et al 2020
Bias-reduced	Lower MSE; overcorrected bias possible; includes shrinkage methods.	RCTdesign, OptGS

Innovative trial designs and de-biased estimators (Robertson et al 2021 Part II)

Recommendations (Roberston et al 2021, Part II):

- Planning:
  - Conditional or unconditional?
  - Simulations following FDA guidance
  - Pre-specify
- At DMC
  - Provide both
- Final
  - Which estimator has been used for primary and secondary endpoints?

Example: Bias in Paired-Eye Designs?

Dream of the Statistician

Each subject acting as their own control

Huge sample size impact (~4 times)

BUT: it really changes the statistical model!
"X: $\Omega \rightarrow \mathbb{E}$ " over $(\Omega, \Sigma, (P_\theta, \theta \in \Theta))$

Potential Sources of Bias

Underestimation:

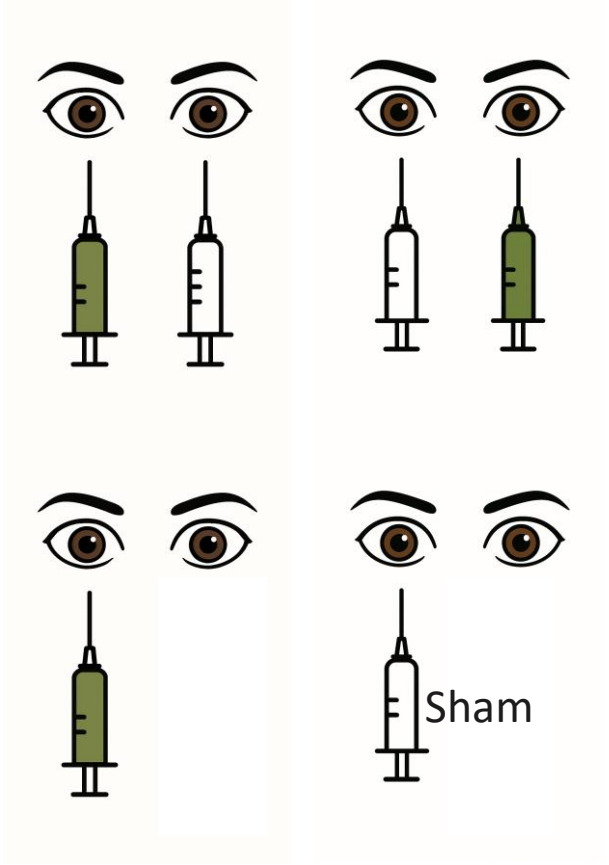
- Correlation on Perceptual Level
- Transfer of drug to the other eye
- Administration error

Overestimation:

- If a treatment effect exists, „Expectation" bias in psychophysical and PRO measurements

Innovations beyond Parallel Design: (Stated) Bias by Within-Subject Designs

- *„There is a bias **of the effect size** with a trial design, if a **true effect** between an active treatment and a control treatment is **systematically different** when assessed in a specific design compared to a **parallel arm design.“***
- **Rationale:** Parallel arm designs reflect treatment of the future which is per patient
- **Sceptisism:** *„Efficacy-Effectiveness Gap“* of parallel design.

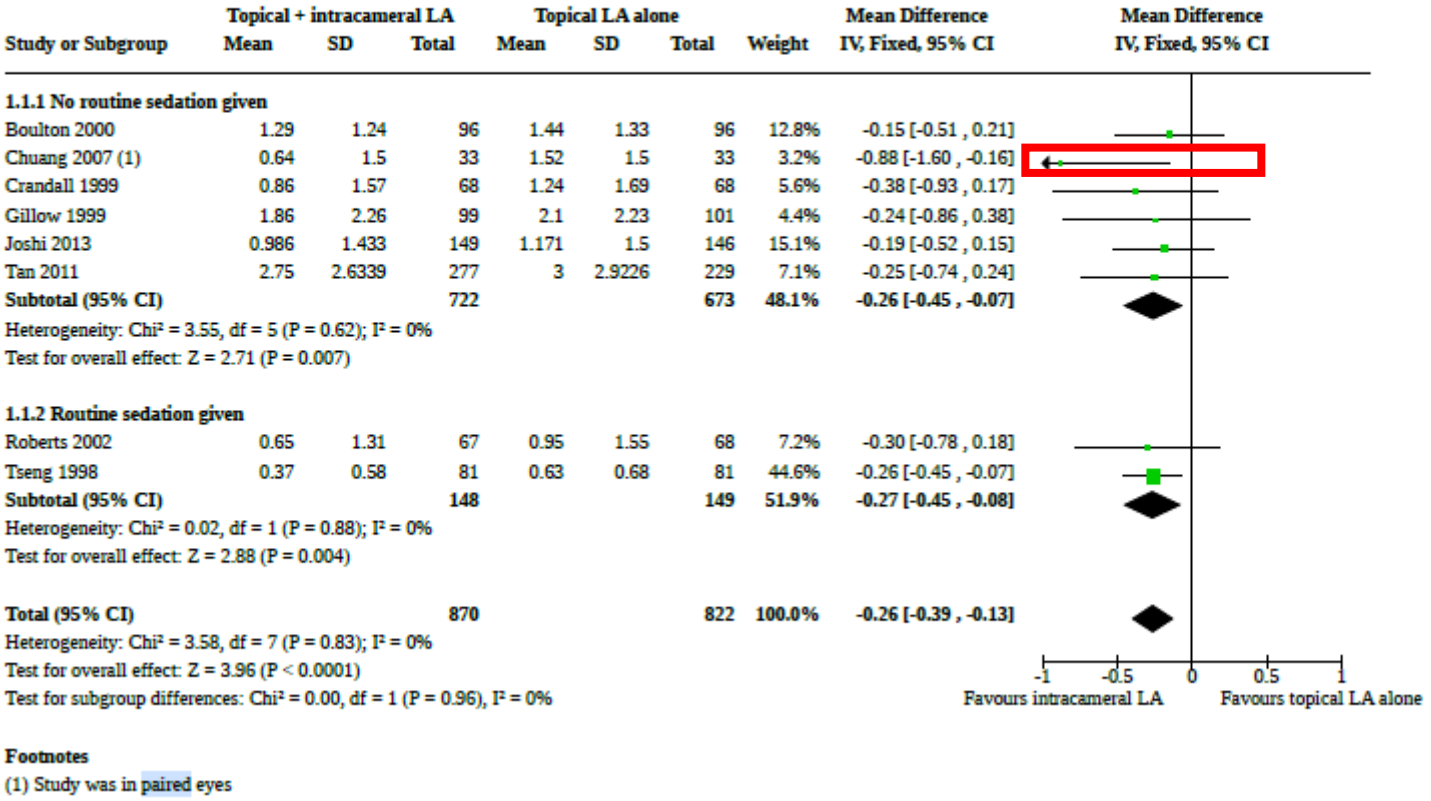


Learning from Meta-analyses
Cochrane reviews comparing paired vs parallel

Title	Paired/All	Paired vs Parallel
Laser-assisted subepithelial keratectomy (LASEK) versus photorefractive keratectomy (PRK) for correction of myopia	10/11	No effect, thus, no exaggeration effect
Intense pulsed light (IPL) therapy for the treatment of meibomian gland dysfunction	2/3	No exaggeration effect! No effect in paired-eye studies, some effect in parallel arm study, possibly due to poor masking.
Laser-assisted cataract surgery versus standard ultrasound phacoemulsification cataract surgery	13/42	No effect, thus, no exaggeration effect.
Laser-assisted subepithelial keratectomy (LASEK) versus laser-assisted in-situ keratomileusis (LASIK) for correcting myopia	2/4	No effect, thus, no exaggeration effect.
Topical anaesthesia plus intracameral lidocaine versus topical anaesthesia alone for phacoemulsification cataract surgery in adults	1/13	**The paired design shows a stronger effect on the pain scale.** **The difference is not explained well by general trial heterogeneity.**
Pharmacologic interventions for mydriasis in cataract surgery	1/14	The paired design shows a stronger effect on pupilarry diameter. The difference is well explained by general trial heterogeneity
Transepithelial versus epithelium-off corneal crosslinking for progressive keratoconus	1/13	No effect, thus, no exaggeration effect

Topical anaesthesia plus intracameral lidocaine versus topical anaesthesia alone for phacoemulsification cataract surgery in adults

Figure 4. Forest plot of comparison: 1 Topical anaesthesia alone versus topical anaesthesia with intracameral lidocaine for phacoemulsification cataract surgery in adults, outcome: 1.1 Intraoperative pain or discomfort (continuous 10-point scale).



Footnotes

(1) Study was in paired eyes

Chuang 2007: Treatments in both eyes were done 2 weeks apart.

A larger effect size can be explained, if patients at the second treatment remember the pain level from the first treatment. For those with first placebo, and then active, they are relieved if pain is less than expected. For those that have first active and then placebo, they may be disappointed because the pain is larger than what they expected.

In that case it is not so straightforward which is the better repesentation of the „true“ effect size in real life: if people can compare or cannot compare to an anchor??

Conclusions

- Innovative trial designs can introduce additional sources of bias
- Discussed here:
  - Bias of the estimator
    - Cures are often possible but rarely applied
    - Not necessarily the most important bias for B/R?
  - Systematic difference to a result compared to a parallel trial design
    - As always, what is considered a bias depends on the definition of truth

https://sketchplanations.com/looking-under-the-lamppost

References

- Whitehead J, Desai Y, Jaki T. Estimation of treatment effects following a sequential trial of multiple treatments. Stat Med. 2020;39(11):1593-1609. doi:10.1002/sim.8497
- Robertson, D. S., Choodari-Oskooei, B., Dimairo, M., Flight, L., Pallmann, P., & Jaki, T. (2023). *Point estimation for adaptive trial designs I: A methodological review*. **Statistics in Medicine, 42**(2), 122–145. https://doi.org/10.1002/sim.9561
- Robertson, D. S., Choodari-Oskooei, B., Dimairo, M., Flight, L., Pallmann, P., & Jaki, T. (2023). *Point estimation for adaptive trial designs II: Practical considerations and guidance*. **Statistics in Medicine, 42**(13), 2496–2520. https://doi.org/10.1002/sim.9734
- Chuang, L.-H., Yeung, L., Ku, W.-C., Yang, K.-J., & Lai, C.-C. (2007). Safety and efficacy of topical anesthesia combined with a lower concentration of intracameral lidocaine in phacoemulsification: Paired human eye study. *Journal of Cataract & Refractive Surgery, 33*(2), 293–296. https://doi.org/10.1016/j.jcrs.2006.10.037