Biases in Clinical Trials with Sequential Monitoring

Manjula Schou and Ian Marschner

Macquarie University, Sydney, Australia

MRC-HTMR Clinical Trials Methodology Conference, Bristol, October 2011

Truncated clinical trials

- A randomised controlled trial (RCT) that stops earlier than initially planned is called a **truncated RCT**
- There are various reasons why an RCT might be truncated:
  - Benefit: The experimental intervention apparently has superior efficacy to the control intervention
  - Safety: The experimental intervention apparently has unacceptable adverse effects
  - Futility: There is apparently no prospect of this study showing superior efficacy to the experimental intervention
- **In this talk we will only discuss truncation due to benefit**

Bias in truncated RCTs

- Early stopping requires a sequential monitoring rule that appropriately controls the type I error for multiple testing e.g. O'Brien-Fleming, Haybittle-Peto, Pocock *etc.*

- A valid sequential stopping rule will appropriately control the type I error but theoretically:
  - the observed treatment difference in a truncated RCT will tend to overestimate the actual treatment difference
  - the nominal p-value in a truncated RCT will be more extreme than is warranted

- This theoretical bias arises because random fluctuations in the direction favouring the experimental treatment can lead to truncation of the RCT

Debate over consequences

- **Truncated RCTs and systematic reviews are biased:**
  - *"Clinicians should view results of RCTs stopped early for benefit with skepticism"* (Montori *et al.*, *JAMA* 2005)
  - *"Most systematic reviews and meta-analyses including truncated RCTs fail to consider the possible overestimates of effect that may result from early stopping for benefit"* (Bassler *et al.*, *JClinEpi* 2007)
  - *"If reviewers do not note truncation and do not consider early stopping for benefit, meta-analyses will report overestimates of effects"* (Bassler *et al.*, *JAMA* 2010)

- **Truncated RCTs and systematic reviews are not biased:**
  - *"Systematic reviews are not biased by results from trials stopped early for benefit"* (Goodman, *JClinEpi* 2008)
  - *"Stopping at 50% or greater information has a negligible impact on estimation"* (Freidlin and Korn, *Clinical Trials* 2009)
  - *"Early termination . . . does not lead to substantive bias in the estimation of treatment effects"* (Berry *et al.*, *JAMA* 2010)
  - *"Estimates from trials stopped early for efficacy have negligible bias"* (Goodman, *JAMA* 2010)

Implications of bias

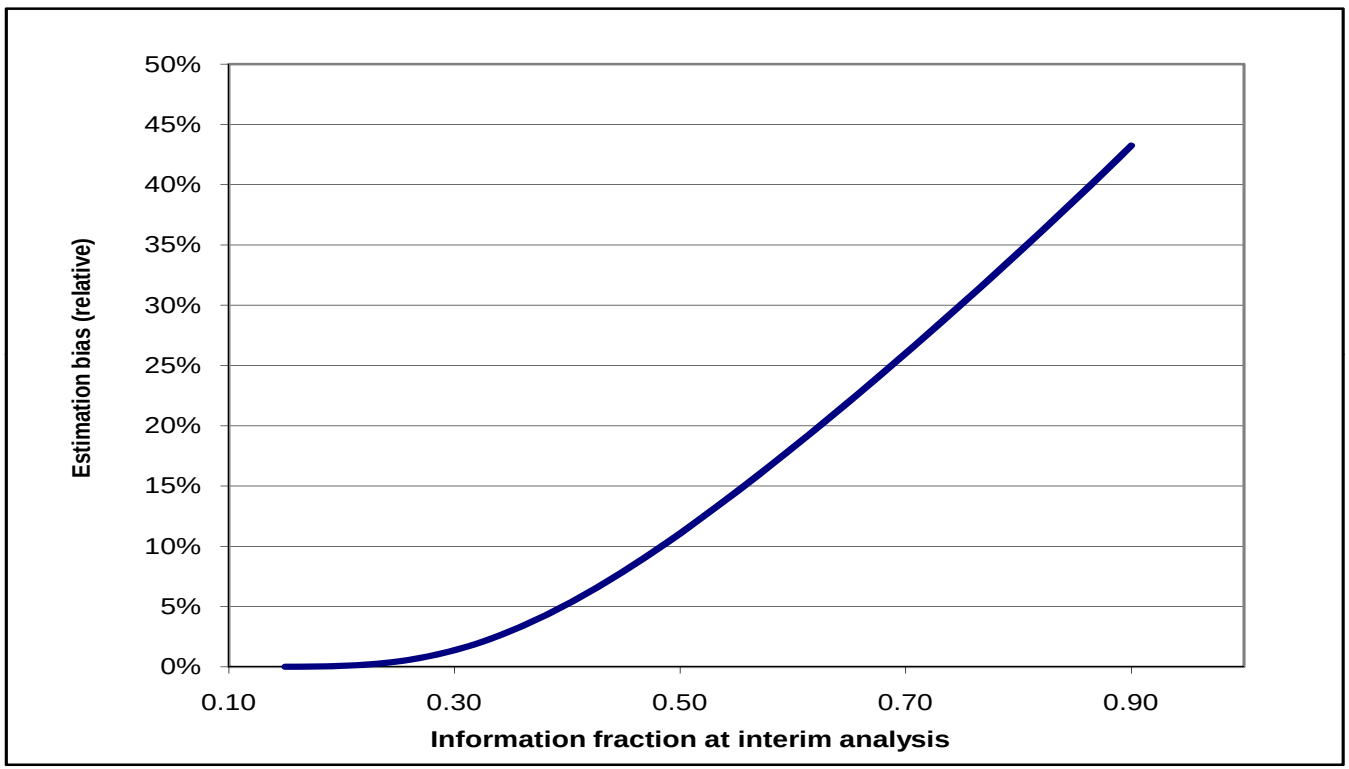
- Possible responses to the argument that truncated RCTs and systematic reviews are substantially biased include:
  - discontinuing the practice of truncating RCTs for benefit
  - excluding truncated RCTs from evidence synthesis
  - adjusting truncated RCTs for overestimation
- Discontinuing early stopping for benefit is currently not practical
- Excluding or adjusting truncated RCTs is practical
- **We examined the implications of excluding truncated RCTs from meta-analyses**

Conditioning on non-truncation

- Planned study:
  - Two groups each with sample size $2n$
  - Normal endpoint with unit variance and mean difference δ
  - Sample means $\overline{X}_1$ and $\overline{X}_2$ with observed difference $\hat{\delta}_{\text{crude}}$
- 1 interim analysis at information fraction $\tau \in [0, 1]$ with sample means $\overline{X}_1^{(\tau)}$ and $\overline{X}_2^{(\tau)}$ and observed difference $\hat{\delta}^{(\tau)}$
- Interim stopping rule (OBF): $\hat{\delta}^{(\tau)} > b_{\alpha,\tau}/\sqrt{\tau n}$
- Conditional on non-truncation:
  - $\hat{\delta}_{\text{crude}}$ is not normally distributed with mean δ
  - Distribution of $(\hat{\delta}_{\text{crude}}, \hat{\delta}^{(\tau)})$ is bivariate truncated normal
  - Adjusted estimate $\hat{\delta}$ and information $I(\delta)$ can be determined based on bivariate truncated normal and compared with $\hat{\delta}_{\text{crude}}$ and n

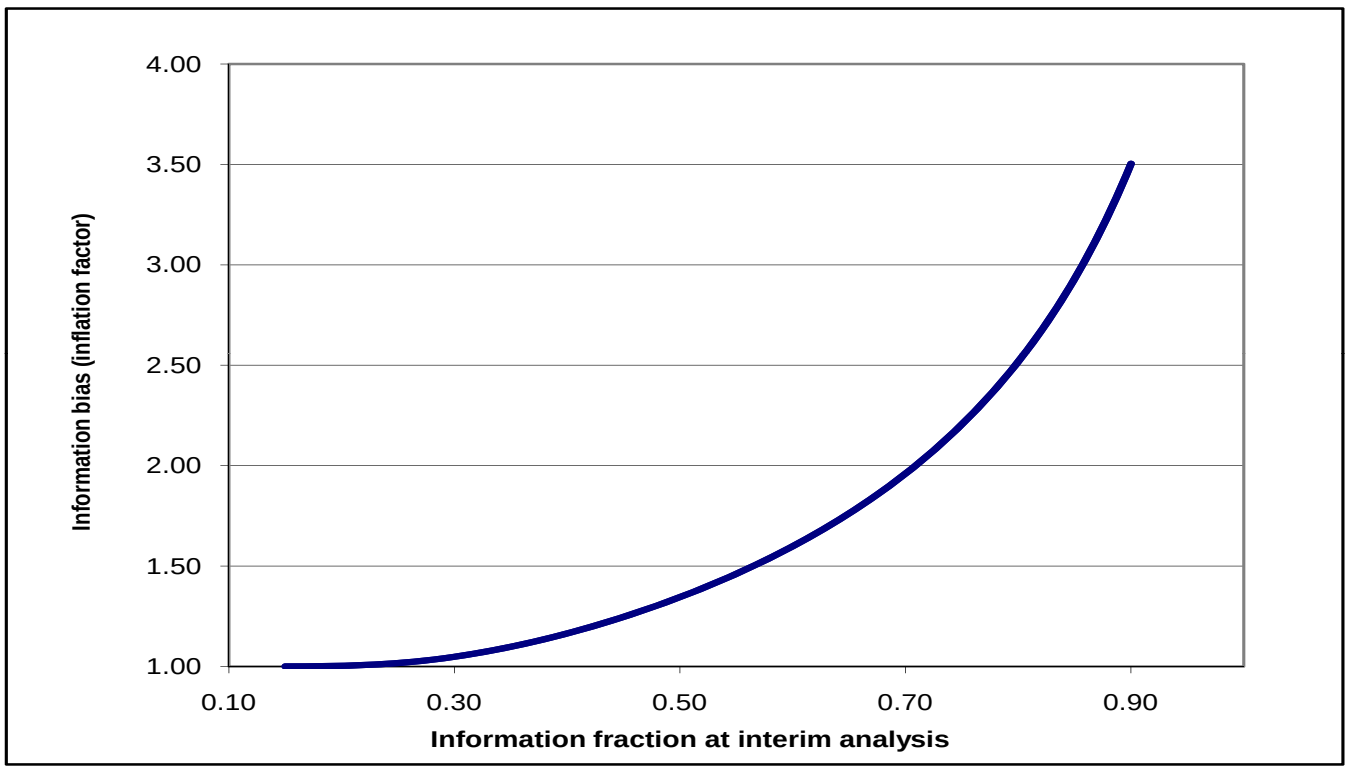
Estimation bias

$$\hat{\delta}_{\text{crude}} = \hat{\delta} - \frac{\sqrt{\tau}\phi(b_{\alpha,\tau} - \hat{\delta}\sqrt{\tau n})}{\sqrt{n}\Phi(b_{\alpha,\tau} - \hat{\delta}\sqrt{\tau n})} = \hat{\delta} - \mathrm{B}(\alpha, \tau, \hat{\delta}, n)$$



Information bias

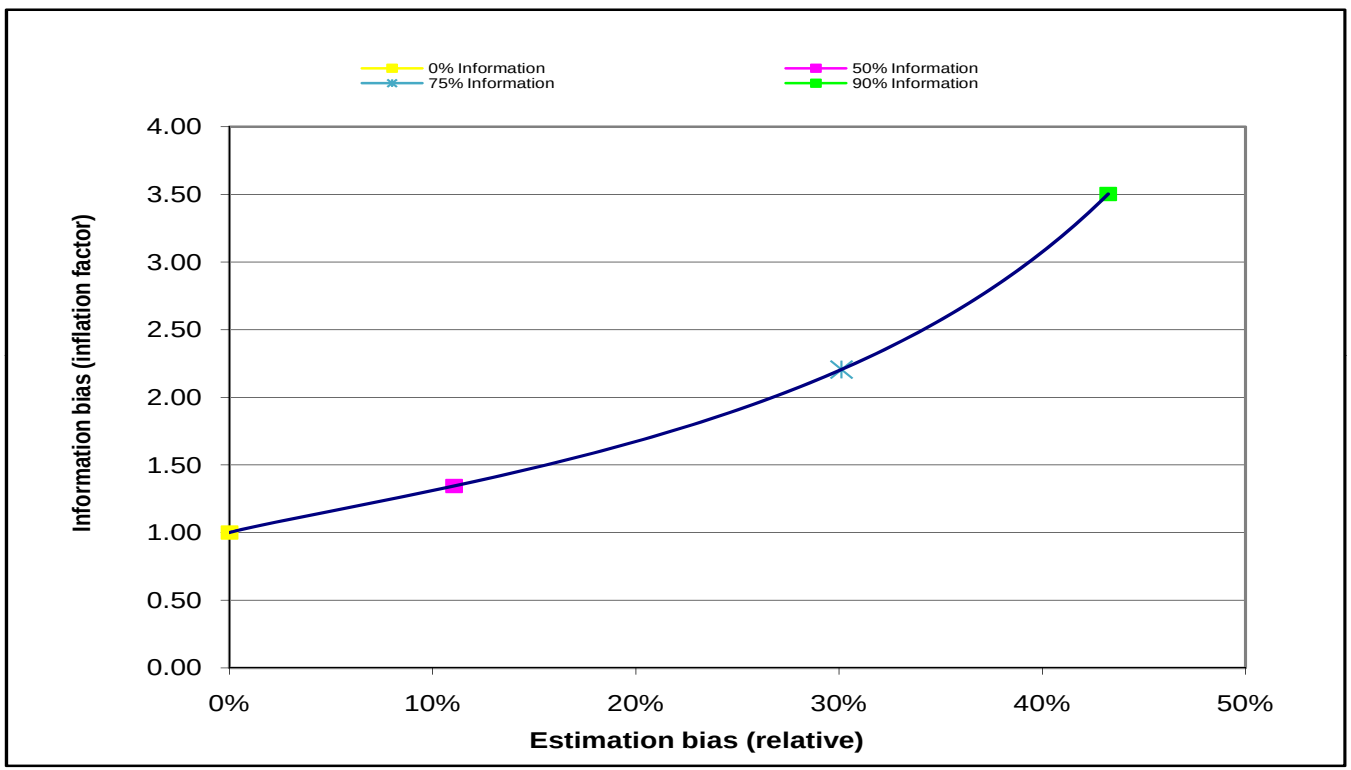
$$n = I(\delta) \Big/ \left\{ 1 - \mathsf{B}\sqrt{\tau n}\left(b_{\alpha,\tau} - \delta\sqrt{\tau n} + \mathsf{B}\sqrt{n/\tau} \right) \right\} = I(\delta)\mathsf{IF}(\alpha, \tau, \delta, n)$$



Double whammy

Conditioning on non-truncation:

- Information bias (IF) increases with estimation bias (B)
- In meta-analyses, both the estimates and the weights are biased
- Studies with the largest estimation bias are the most overweighted

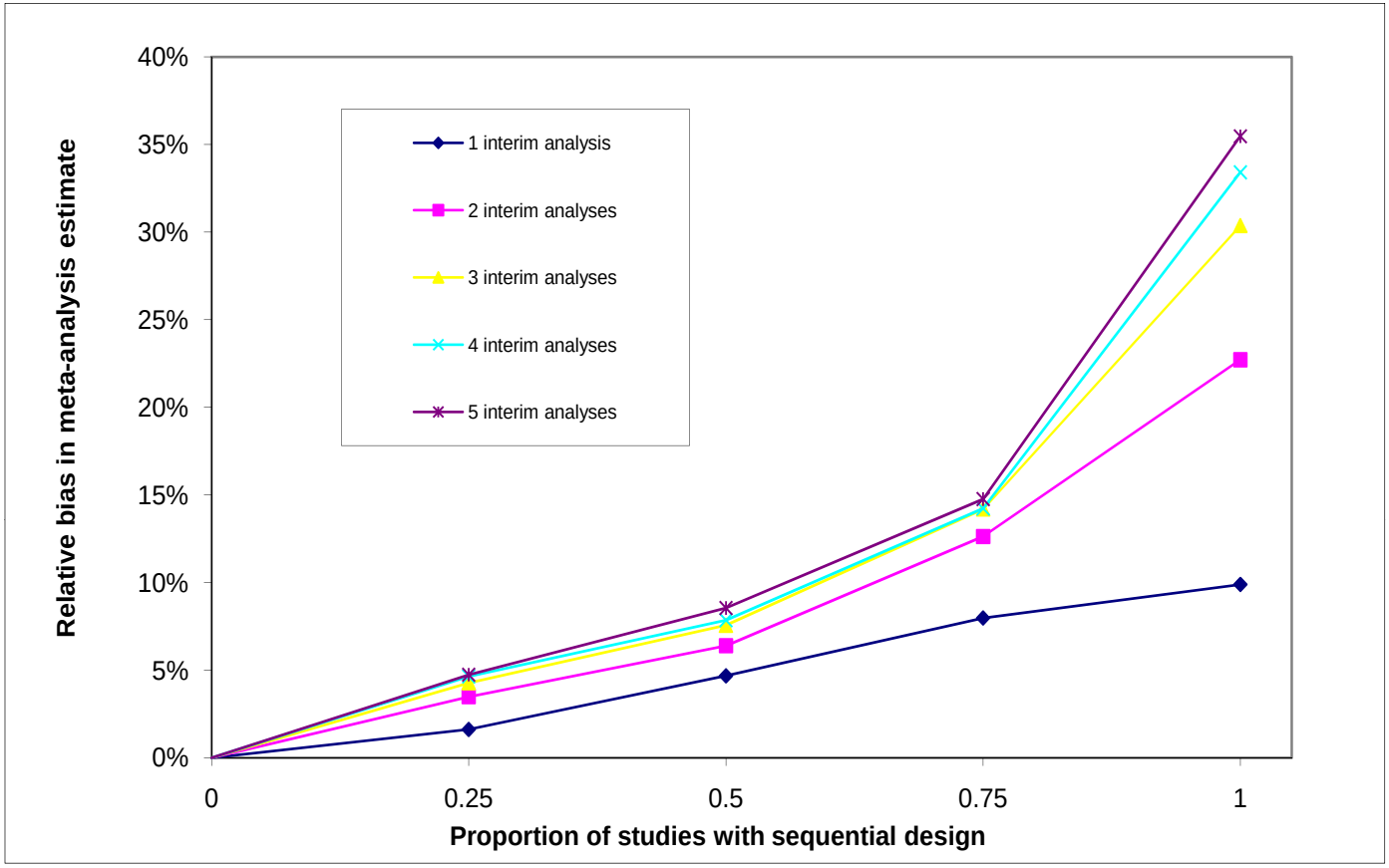


Meta-analysis simulations

- Systematic reviews involving 12 studies
- Normally distributed endpoint with treatment effect size 0.1–0.5
- Some studies (0% – 100%) are subject to possible truncation using OBF stopping boundary
- Remaining studies have a fixed (non-sequential) design
- Sequential studies have 1–5 equally spaced interim analyses
- All studies have 90% power with sample sizes 85–2100 per arm
- Fixed effects meta-analysis of treatment effect using:
  - non-truncated studies only
  - fixed design studies only
  - all studies

Non-truncated studies

- Under-estimation in meta-analyses restricted to non-truncated RCTs



All studies versus fixed design studies

- Relative bias and efficiency for treatment effect size 0.25

interim analyses	sequential	truncated	fixed bias	**all bias**	efficiency
1	25%	8%	0%	**1%**	79%
	50%	16%	0%	**1%**	55%
	75%	24%	0%	**1%**	29%
	100%	31%	–	**1%**	–
2	25%	12%	0%	**–1%**	79%
	50%	29%	0%	**0%**	56%
	75%	43%	0%	**–1%**	30%
	100%	58%	–	**0%**	–
3	25%	13%	0%	**–2%**	79%
	50%	33%	0%	**–1%**	57%
	75%	52%	0%	**0%**	31%
	100%	70%	–	**0%**	–

Conclusions

- Conditional on non-truncation:
  - treatment effects are underestimated and statistical information is overestimated
  - the most biased studies are the most overweighted studies in a meta-analysis
- Exclusion of truncated studies from meta-analyses leads to underestimation of treatment effects
- Inclusion of all studies in meta-analyses, both truncated and non-truncated, leads to unbiased estimation of treatment effects
- **Truncated clinical trials are <u>not</u> a source of bias in meta-analyses**
- **Exclusion of truncated clinical trials from meta-analyses <u>is</u> a source of bias**

Recommendations

- Truncated RCTs should not be excluded from meta-analyses
- All studies, both truncated and non-truncated, should be included in meta-analyses
- Adjusted treatment effects and p-values should be reported for truncated RCTs
- Further research is needed on the benefits of using adjusted treatment effects and information weights in meta-analyses