MONASH University

Medicine, Nursing and Health Sciences

Baseline imbalance in trials included in systematic reviews: A cause for concern?

Joanne McKenzie, School of Public Health and Preventive Medicine
Cochrane Methods Symposium, 24th September 2013

Acknowledgements

- Jon Deeks, University of Birmingham, United Kingdom
- Andrew Forbes, Monash University, Australia
- Peter Herbison, University of Otago, New Zealand

Baseline imbalance

Some review authors are concerned about baseline imbalance.

- "… Posttest scores were used. In cases of statistically significant baseline difference, we used change scores. …"

- "An important discrepancy is the use of "post-treatment" data instead of "difference in change data." Effect is a measure of change caused by treatment, and thus should be based on the difference in change between groups, and not the "post-treatment data." In large trials or massive numbers of RCTs, this may be correct, but in reviews of few and smaller trials, this approach may bias the results due to imbalances caused by the randomization procedure."

Questions

- Do we need to be concerned about baseline imbalance when meta-analysing intervention effects from randomised trials?
- What is the impact of baseline imbalance on the pooled intervention effect estimate?
- Is the impact of baseline imbalance on the pooled intervention effect modified by the method of analysis in the trial?

Outline

- Background
  - Baseline imbalance
  - Common approaches to analysing continuous outcomes in trials
    - Examples using observed and generated data
  - Consideration of baseline imbalance in meta-analysis
- Simulation study
- Implications for review authors

Background

Baseline imbalance

- Considering:
  - continuous outcomes measured on the same scale
  - randomised trials
  - imbalance in the baseline measure of the outcome
  - baseline imbalance arising solely from chance
- Not considering:
  - baseline imbalance arising from mechanisms other than chance
    - inadequate sequence generation or allocation concealment
    - exclusions of participants, causing imbalance

Analysing continuous outcomes in a trial

Three common approaches:

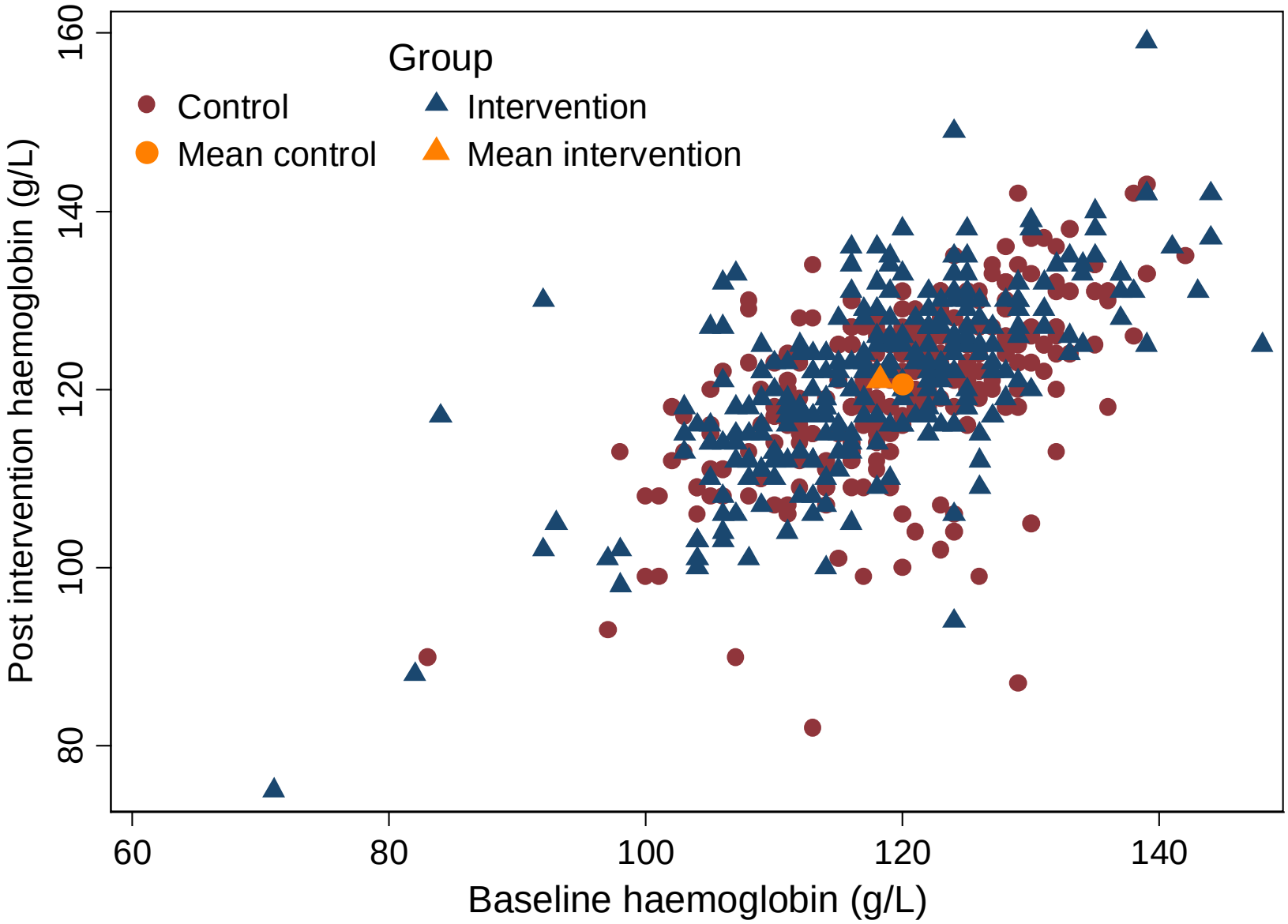
$$\hat{\theta}_{SAFV} = \bar{y}_{\text{int}} - \bar{y}_{ctrl}$$

$$\hat{\theta}_{SACS} = (\bar{y}_{\text{int}} - \bar{y}_{ctrl}) - (\bar{x}_{\text{int}} - \bar{x}_{ctrl})$$

$$\hat{\theta}_{ANCOVA} = (\bar{y}_{\text{int}} - \bar{y}_{ctrl}) - \beta(\bar{x}_{\text{int}} - \bar{x}_{ctrl})$$

where $\beta = \rho \dfrac{\sigma_y}{\sigma_x}$

Post intervention haemoglobin vs baseline haemoglobin



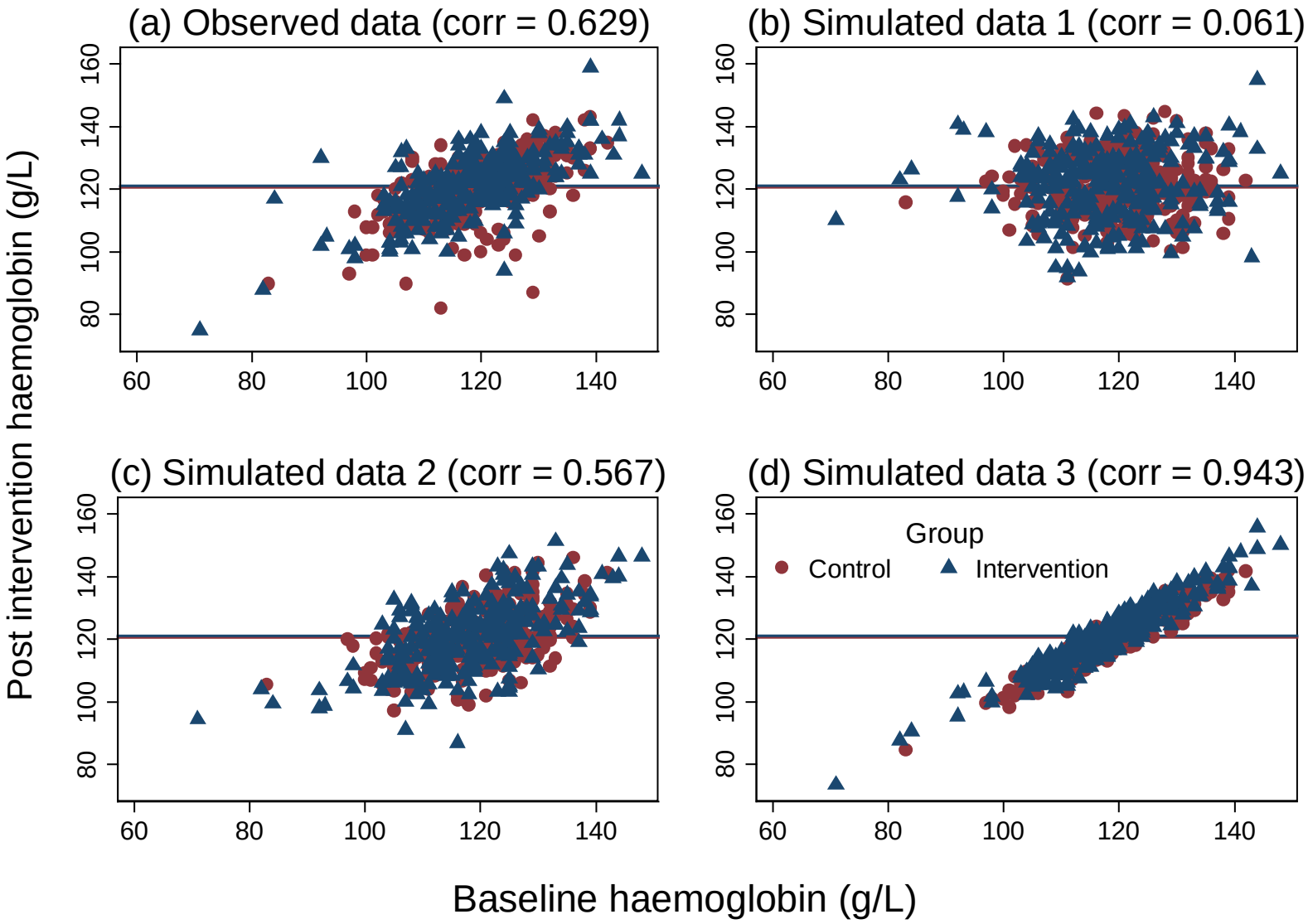
Data sets used to illustrate the impact of correlation on intervention effects

Dataset	Observed correlation	Follow-up haemoglobin (g/L)			
		Intervention group		Control group	
		Mean	SD	Mean	SD
Observed data	0.629	121.0	10.1	120.5	9.5

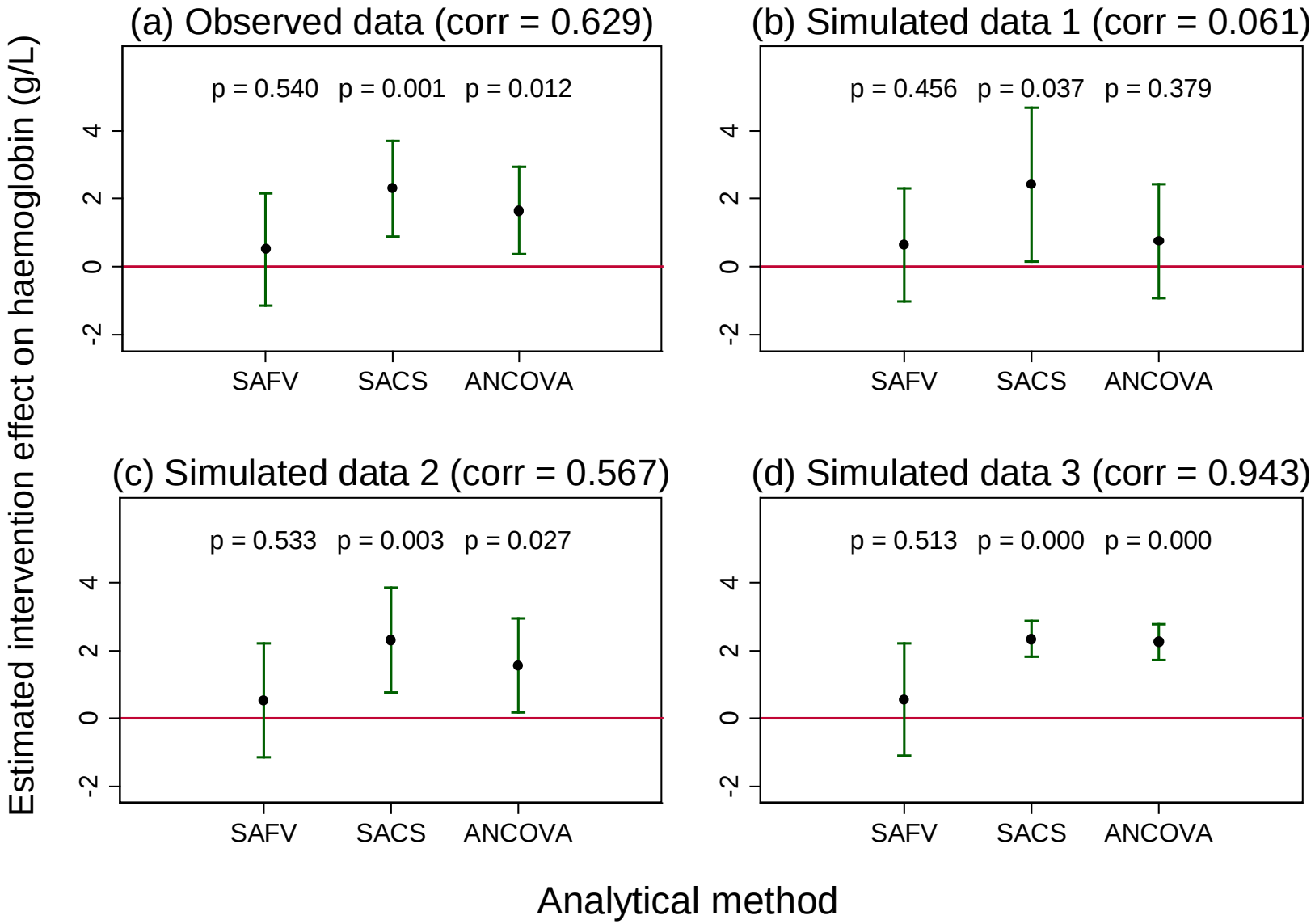
Data sets used to illustrate the impact of correlation on intervention effects

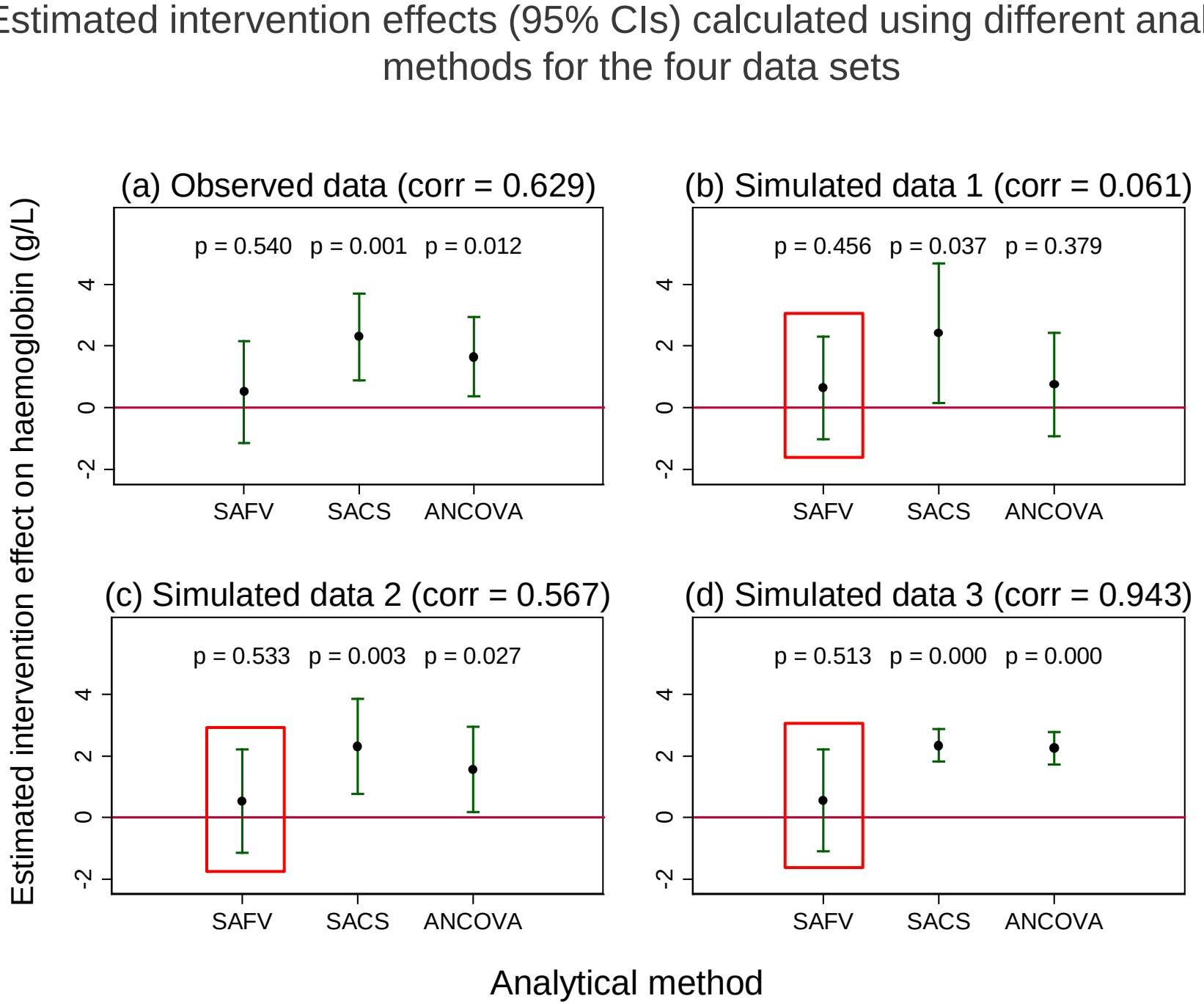
Dataset	Observed correlation	Follow-up haemoglobin (g/L)			
		Intervention group		Control group	
		Mean	SD	Mean	SD
Observed data	0.629	121.0	10.1	120.5	9.5
Simulated data 1	0.061	121.2	10.8	120.6	8.8
Simulated data 2	0.567	121.2	10.8	120.6	8.8
Simulated data 3	0.943	121.1	10.5	120.5	9.0

Scatter plots of post intervention haemoglobin vs baseline haemoglobin for observed and simulated data sets

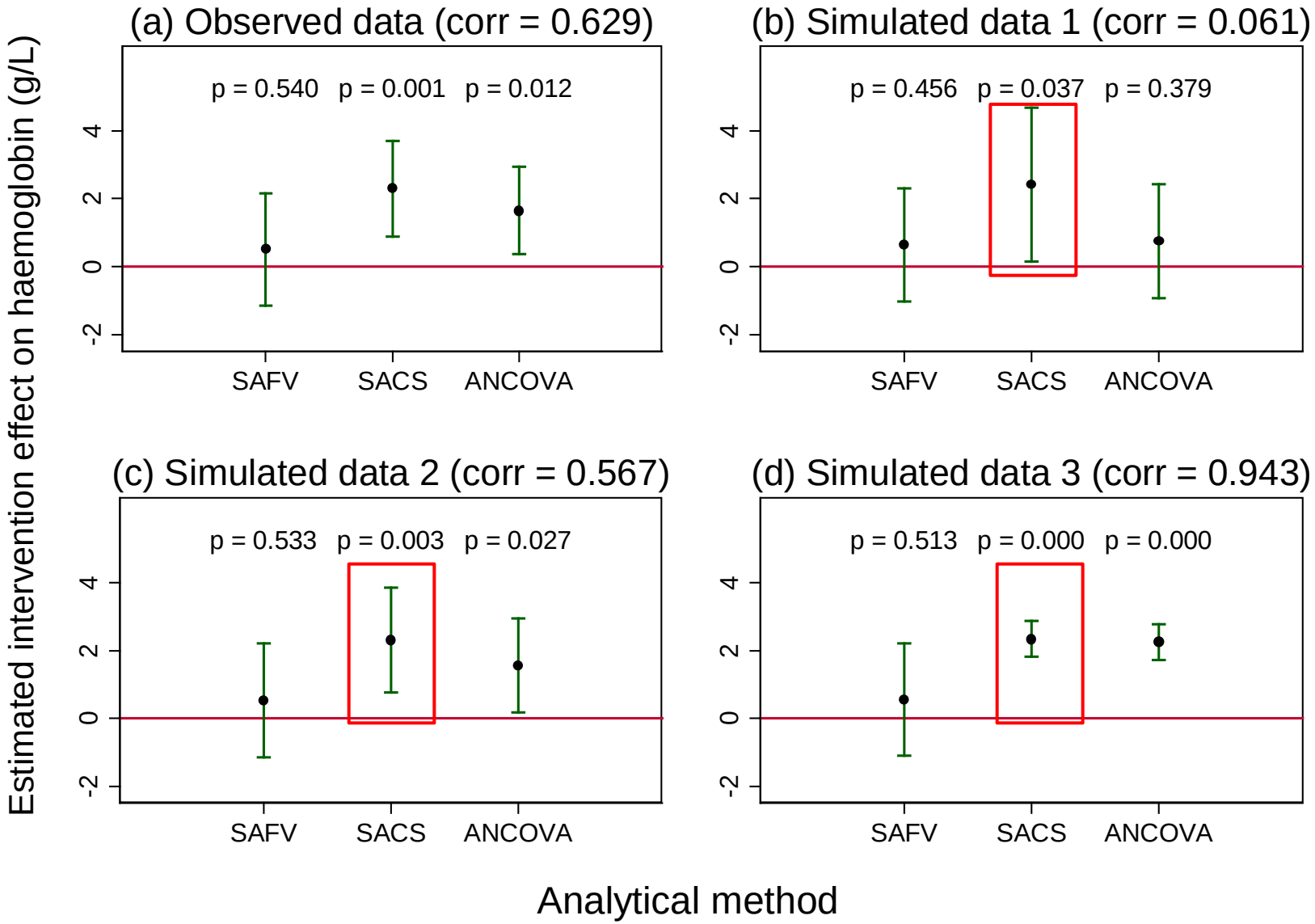


Estimated intervention effects (95% CIs) calculated using different analytical methods for the four data sets

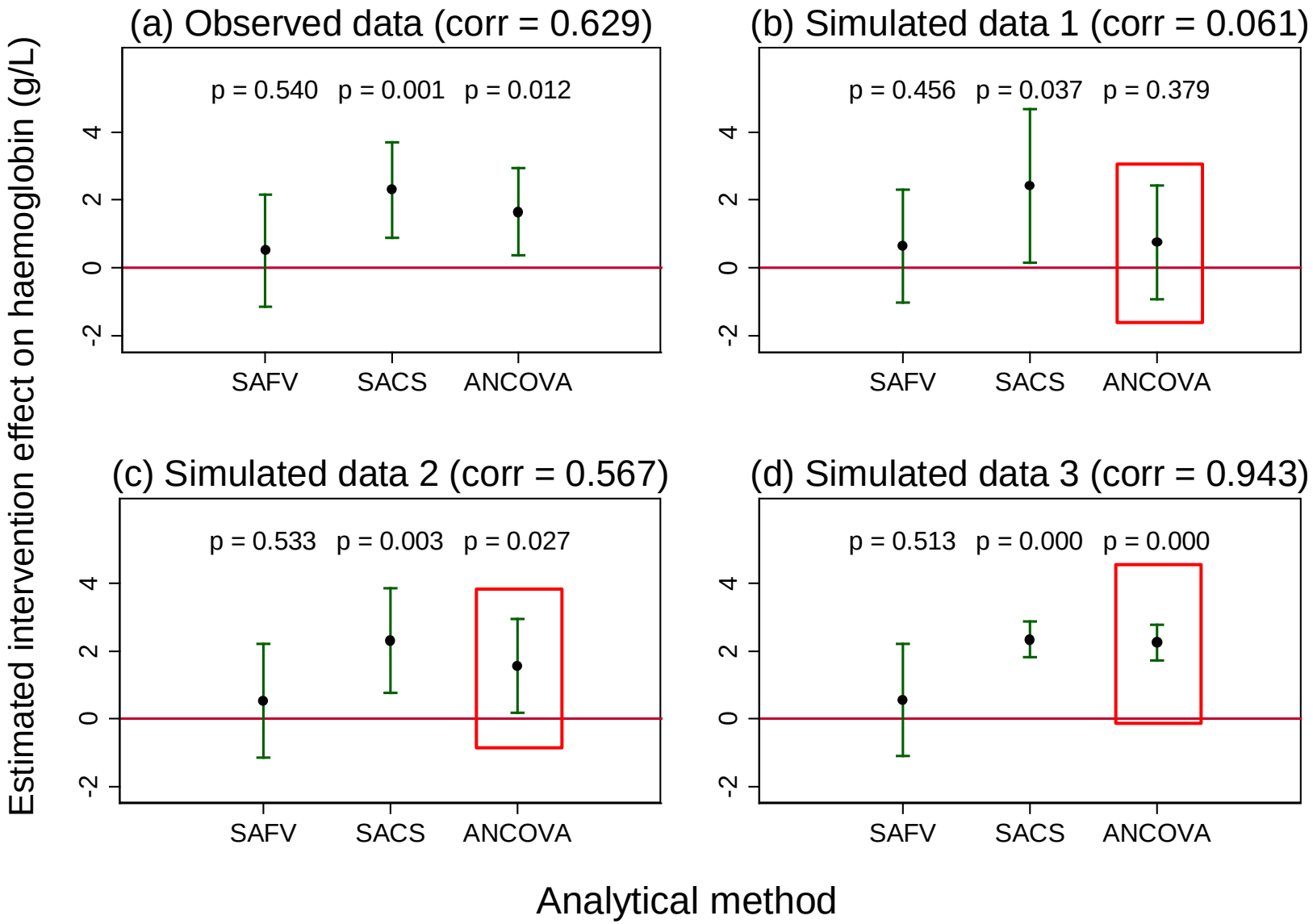



Estimated intervention effects (95% CIs) calculated using different analytical methods for the four data sets
(a) Observed data (corr = 0.629)
p = 0.540 p = 0.001 p = 0.012
(b) Simulated data 1 (corr = 0.061)
p = 0.456 p = 0.037 p = 0.379
(c) Simulated data 2 (corr = 0.567)
p = 0.533 p = 0.003 p = 0.027
(d) Simulated data 3 (corr = 0.943)
p = 0.513 p = 0.000 p = 0.000
SAFV
SACS
ANCOVA
-2
0
2
4
Estimated intervention effect on haemoglobin (g/L)
Analytical method

Estimated intervention effects (95% CIs) calculated using different analytical methods for the four data sets



Estimated intervention effects (95% CIs) calculated using different analytical methods for the four data sets



Comparing the trial analytical methods

- Estimates of intervention effect:
  - For a particular data set, the three analytical methods can produce different estimates of intervention effect.
  - Over the data sets (varying correlation), the ANCOVA estimate varies; SACS or SAFV does not.
- Standard errors:
  - The SE of the FV estimator is not affected by correlation.
  - Increasing correlation results a smaller SE for the CS estimator.
  - Correlation < 0.5, the SE of CS estimator is > SE of the FV estimator. This is reversed when the correlation is > 0.5.
  - For a particular correlation, the SE of the ANCOVA estimate is smaller compared with SEs of FV and CS.

Relationship between the methods

$$\hat{\theta}_{ANCOVA} = (\bar{y}_{\text{int}} - \bar{y}_{ctrl}) - \rho(\bar{x}_{\text{int}} - \bar{x}_{ctrl})$$

(assuming $\sigma_y^2 = \sigma_x^2$)

–When the correlation (ρ) is close to 0:

$$\hat{\theta}_{ANCOVA} \approx \hat{\theta}_{SAFV}$$

–When the correlation (ρ) is close to 1:

$$\hat{\theta}_{ANCOVA} \approx \hat{\theta}_{SACS}$$

When there is minimal baseline imbalance; the three methods produce similar estimates (since $(\bar{x}_{\text{int}} - \bar{x}_{ctrl}) = 0$)

Meta-analysis

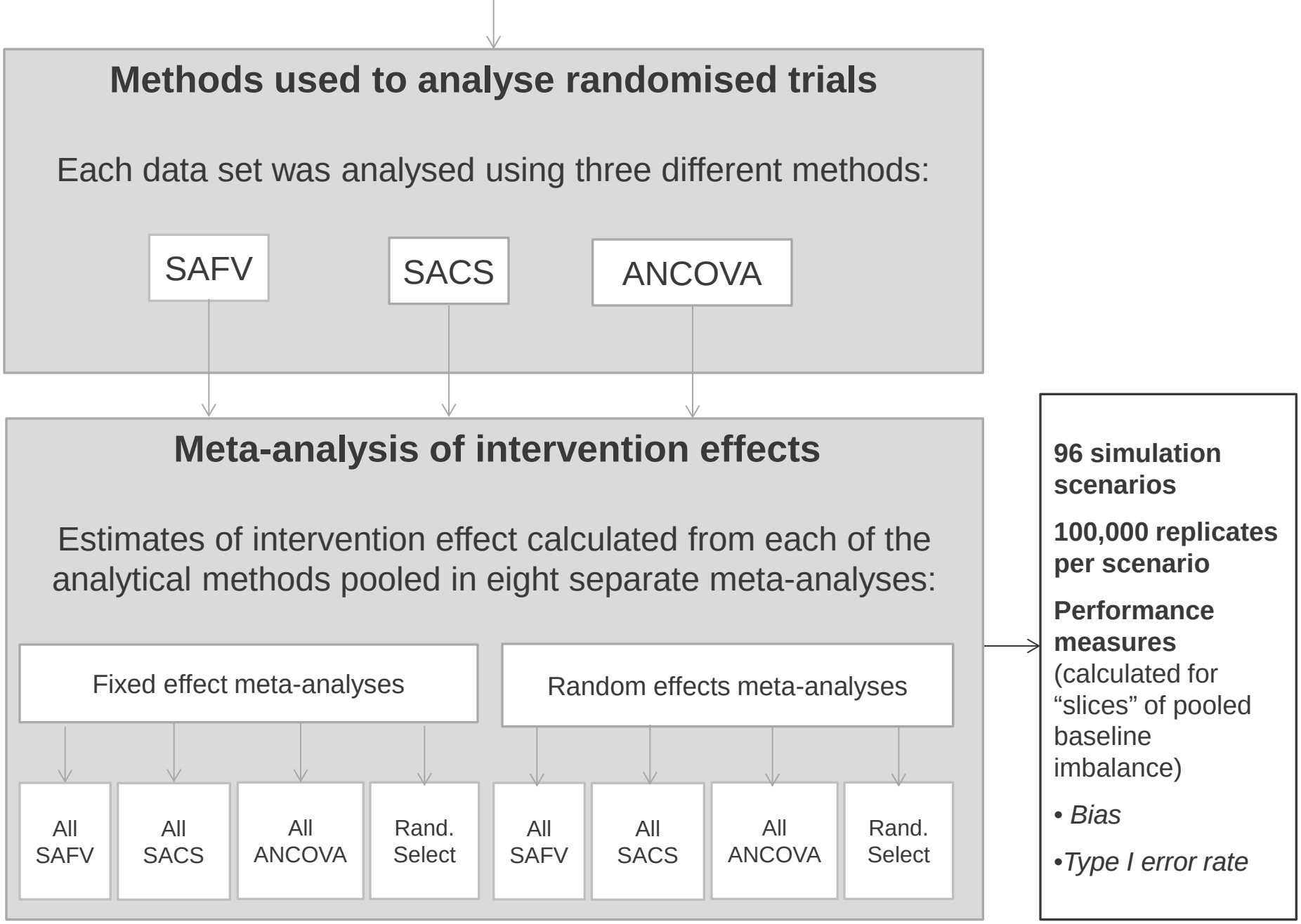
- When meta-analysing intervention effects, does it matter what analytical method was used to calculate the intervention effect in the trial?
- This question can be examined:
  - without consideration for pooled baseline imbalance, or
  - with consideration for pooled baseline imbalance.
- How do meta-analytical methods perform over hypothetical repetitions of meta-analyses:
  - where pooled baseline imbalance randomly varies? (unconditional inference)
  - with the same pooled baseline imbalance? (conditional inference)

Simulation study

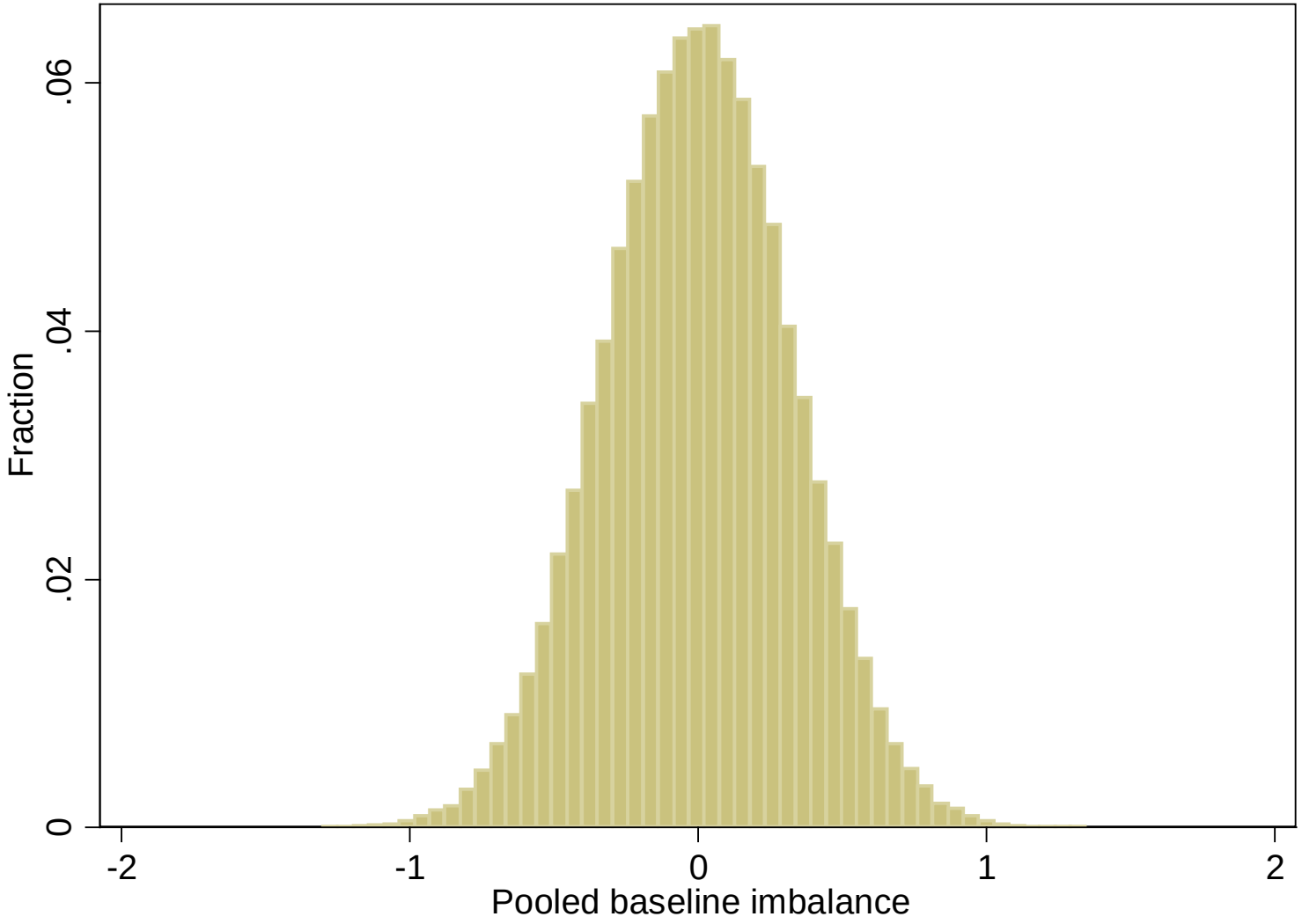
Generation of randomised trials

Randomised trial data sets randomly generated:

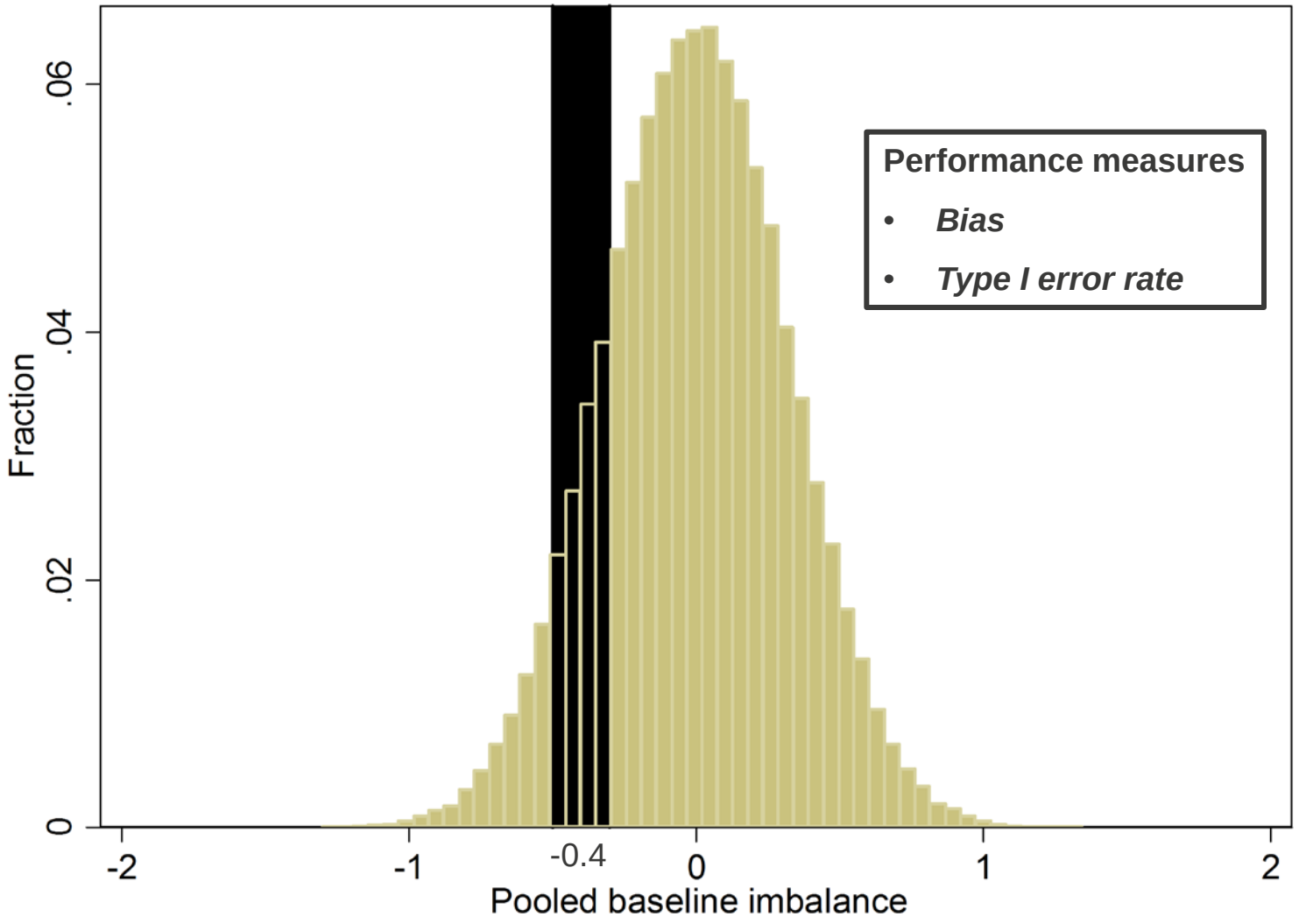
- Trial size (2 levels): 10/group (small) or 30/group (medium).
- Baseline and follow-up scores randomly sampled from a bivariate normal distribution:
  - Control group $\begin{pmatrix} X_{ctrl} \\ Y_{ctrl} \end{pmatrix} \sim BVN\left(\begin{pmatrix} 0 \\ 0.1 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$
  - Intervention group $\begin{pmatrix} X_{int} \\ Y_{int} \end{pmatrix} \sim BVN\left(\begin{pmatrix} 0 \\ \theta_{int} + \theta_{\tau^2} \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$
  - Four correlations: 0, 0.25, 0.5, 0.75
  - θ_{int} = 0.1 or 0.6
  - θ_{τ^2} was distributed $N(0, \tau^2)$, where τ^2 was either 0 or 0.6^2 representing no or moderate between trial heterogeneity.
- Number of trials per meta-analysis: 2, 3, or 5


Methods used to analyse randomised trials
Each data set was analysed using three different methods:
SAFV
SACS
ANCOVA
Meta-analysis of intervention effects
Estimates of intervention effect calculated from each of the analytical methods pooled in eight separate meta-analyses:
Fixed effect meta-analyses
Random effects meta-analyses
All SAFV
All SACS
All ANCOVA
Rand. Select
All SAFV
All SACS
All ANCOVA
Rand. Select
96 simulation scenarios
100,000 replicates per scenario
Performance measures (calculated for "slices" of pooled baseline imbalance)
• Bias
• Type I error rate

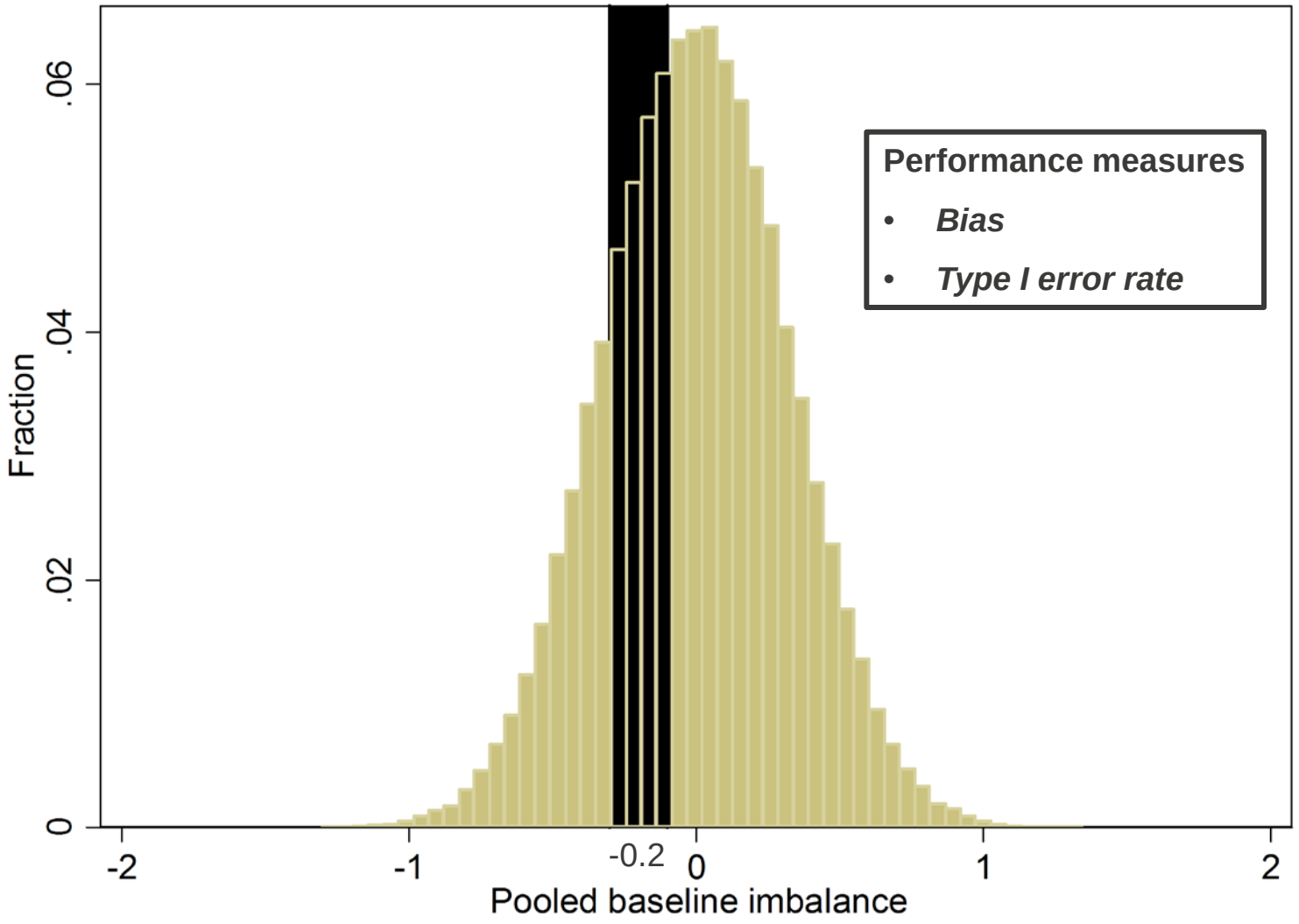
‘Slices’ of pooled baseline imbalance



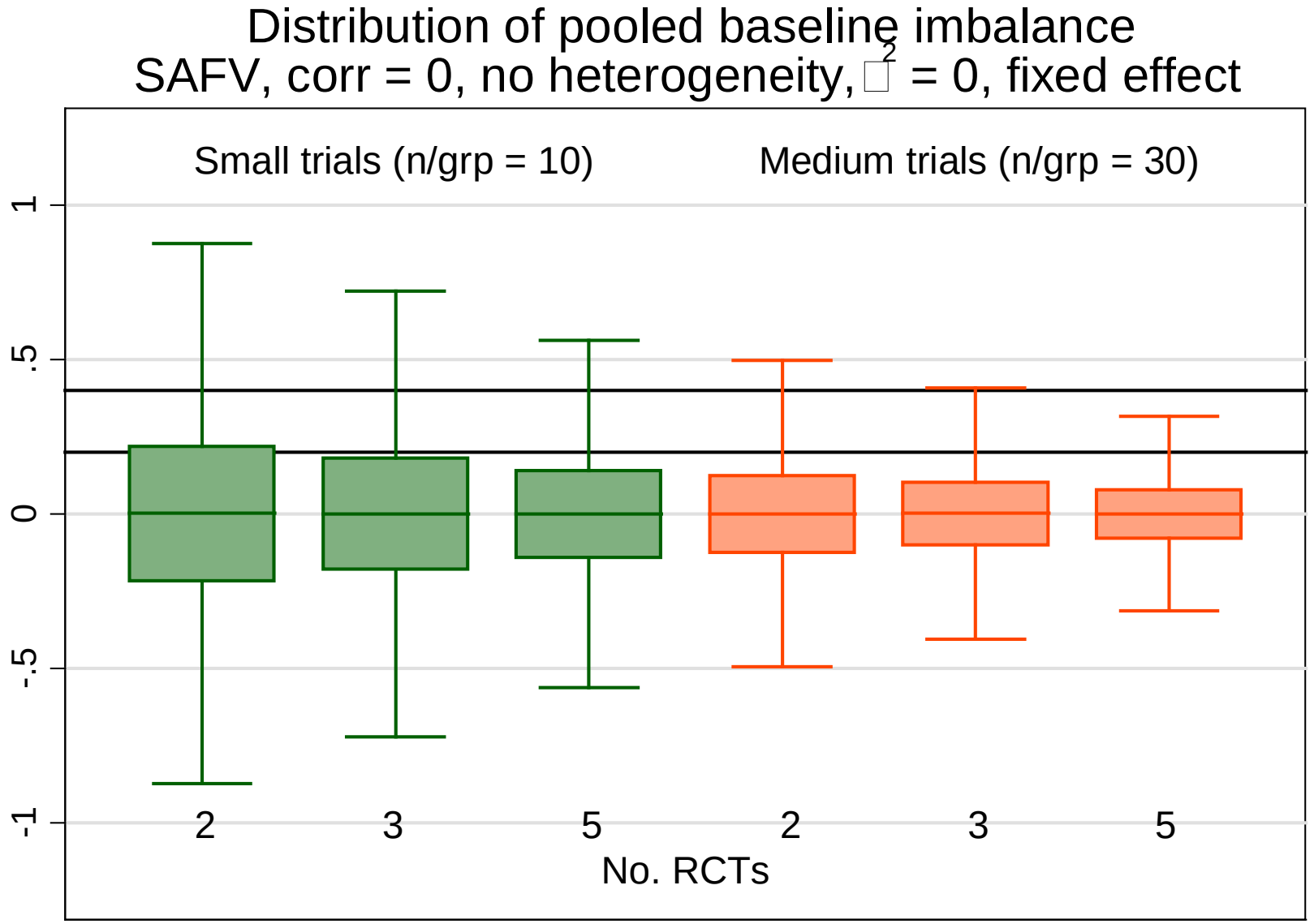
‘Slices’ of pooled baseline imbalance

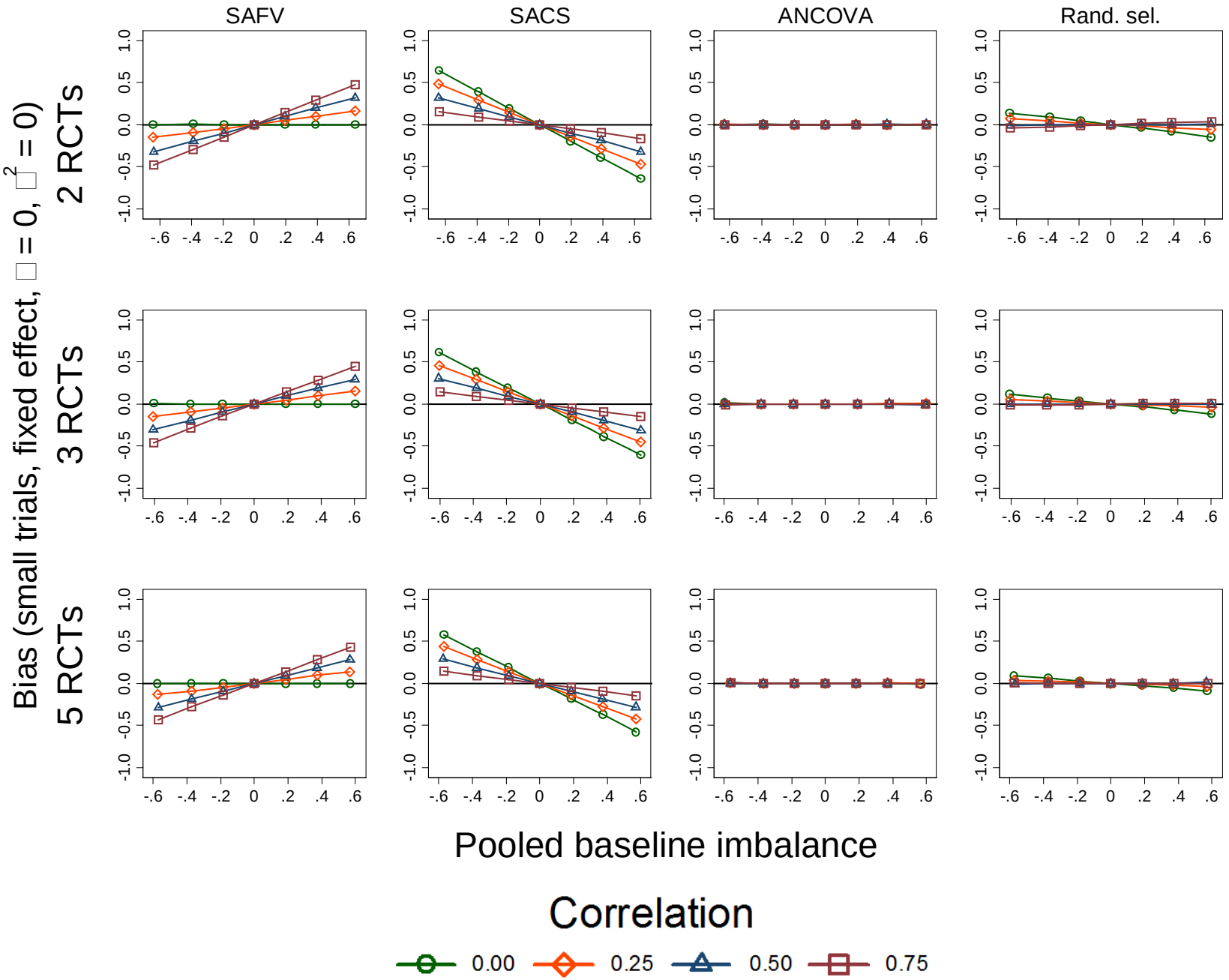


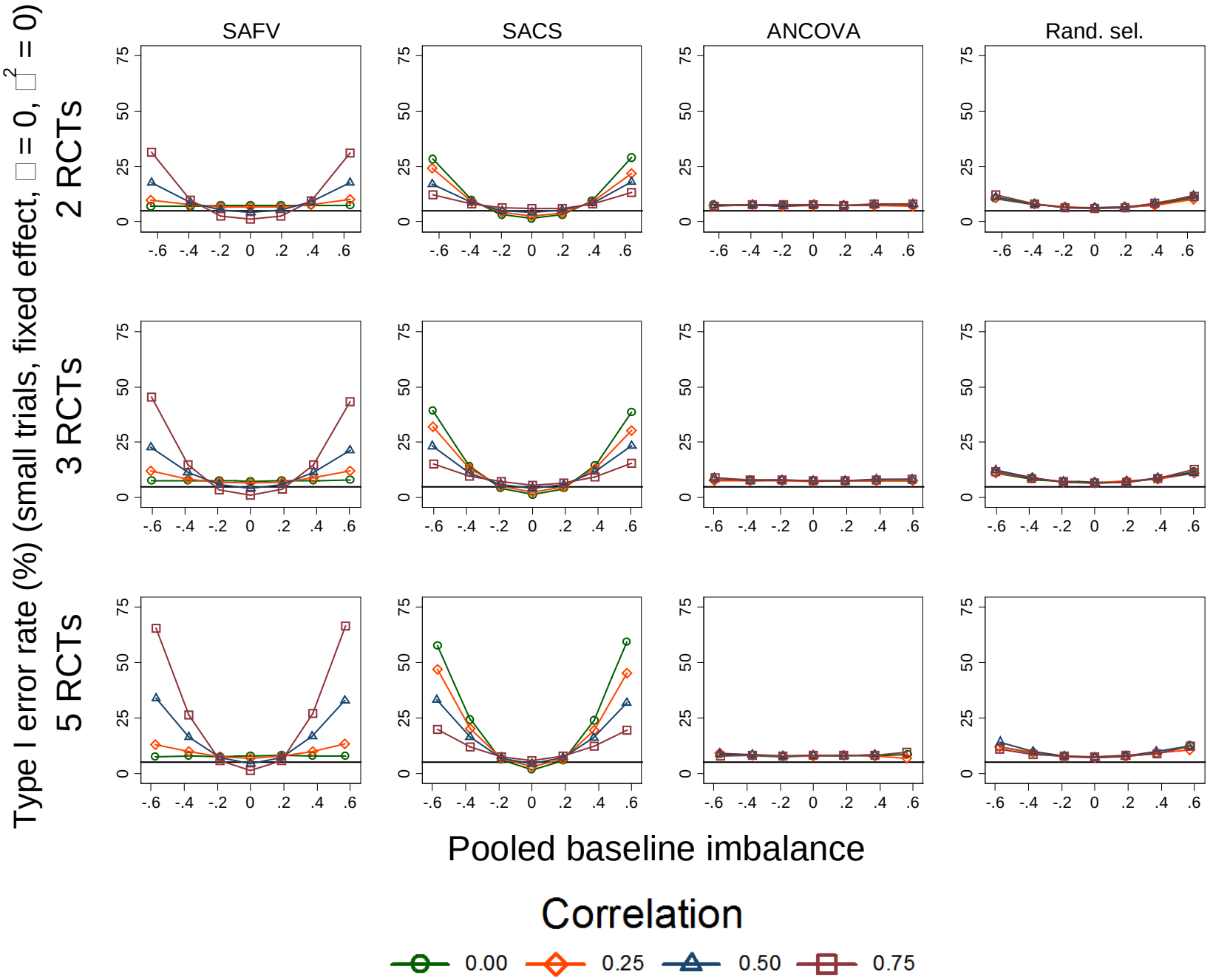
‘Slices’ of pooled baseline imbalance

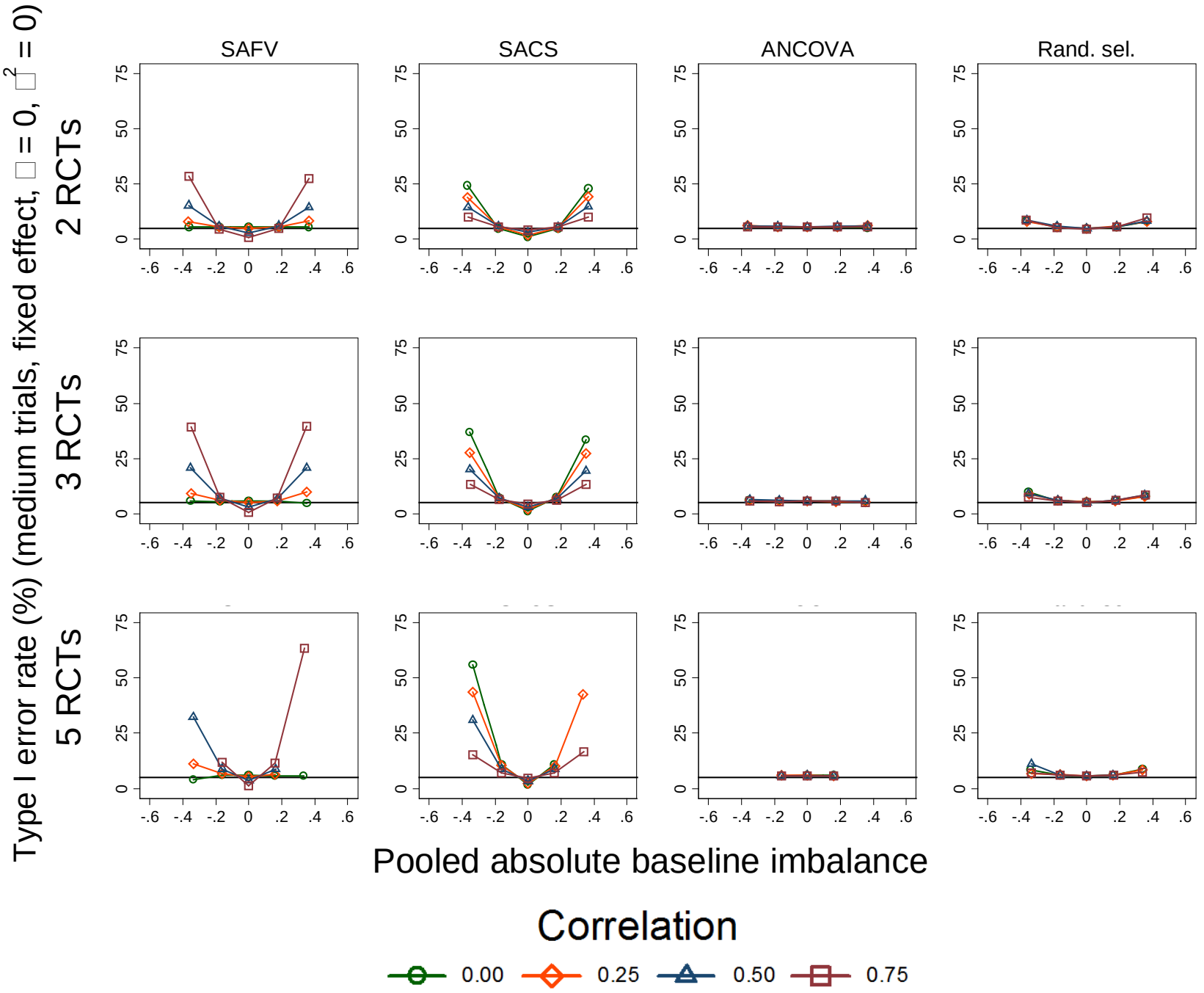


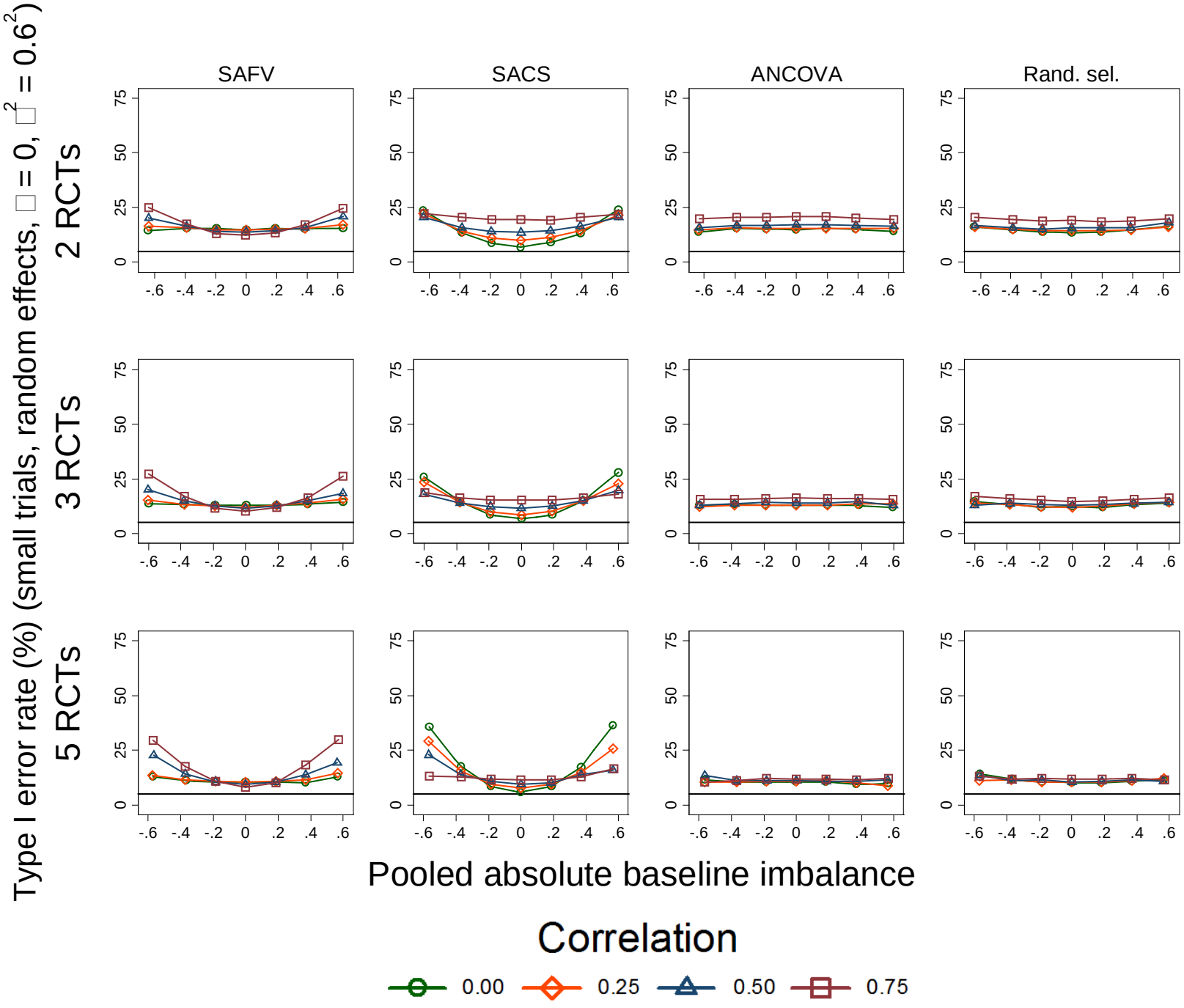
Results


Distribution of pooled baseline imbalance
SAFV, corr = 0, no heterogeneity, τ^2 = 0, fixed effect
Small trials (n/grp = 10)
Medium trials (n/grp = 30)
1
.5
0
-.5
-1
2
3
5
2
3
5
No. RCTs


SAFV
SACS
ANCOVA
Rand. sel.
2 RCTs
3 RCTs
5 RCTs
Bias (small trials, fixed effect, □ = 0, □² = 0)
Pooled baseline imbalance
Correlation
0.00
0.25
0.50
0.75


Type I error rate (%) (small trials, fixed effect, □ = 0, □² = 0)
2 RCTs
3 RCTs
5 RCTs
SAFV
SACS
ANCOVA
Rand. sel.
0
25
50
75
-.6 -.4 -.2 0 .2 .4 .6
Pooled baseline imbalance
Correlation
0.00
0.25
0.50
0.75


Type I error rate (%) (medium trials, fixed effect, □ = 0, □² = 0)
SAFV
SACS
ANCOVA
Rand. sel.
2 RCTs
3 RCTs
5 RCTs
0
25
50
75
-.6 -.4 -.2 0 .2 .4 .6
Pooled absolute baseline imbalance
Correlation
0.00
0.25
0.50
0.75


Type I error rate (%) (small trials, random effects, □ = 0, □² = 0.6²)
SAFV
SACS
ANCOVA
Rand. sel.
2 RCTs
3 RCTs
5 RCTs
0
25
50
75
-.6 -.4 -.2 0 .2 .4 .6
Pooled absolute baseline imbalance
Correlation
0.00
0.25
0.50
0.75

Summary and implications

Summary

- For those concerned about baseline imbalance, concern should be about the pooled baseline imbalance, and not imbalance in the individual trials.
- Large pooled baseline imbalance (arising by chance) is not common.
    - Less likely as the number of trials increases, the sample size of the component trials increases, or both.
- When there is minimal baseline imbalance, all methods yield unbiased pooled intervention effects.

Summary

- When there is pooled baseline imbalance:
  - pooling ANCOVA estimates provides the best solution in terms of bias and type I error rate.
    - Problematic when there is moderate heterogeneity.
  - Pooling either all SAFV or SACS will yield biased estimates, with the bias dependent on the magnitude of baseline imbalance and correlation.

Implications for review authors

- Given large pooled (chance) baseline imbalance is not common, should we be concerned? Probably not.
- ANCOVA estimates should be used when available, however, they are not commonly reported.
- ANCOVA estimates can be recreated, but this is likely to involve algebraic manipulation and imputation of missing parameters.
  - Difficult to implement.
- An alternative to recreating ANCOVA estimates could be to pool either SAFV or SACS if we have knowledge of the 'typical' correlation.
  - E.g. SACS if correlation > 0.5 (less biased with baseline imbalance)

References

- Overall JE, Magee KN. Directional baseline differences and type I error probabilities in randomized clinical trials.[comment]. *Journal of Biopharmaceutical Statistics* 1992;2(2):189-203.
- Senn SJ. Covariate imbalance and random allocation in clinical trials. *Statistics in Medicine* 1989;8(4):467-75.
- Senn S. *Design and interpretation of clinical trials as seen by a statistician*. Chichester: Wiley & Sons, 1997.
- Wei L, Zhang J. Analysis of data with imbalance in the baseline outcome variable for randomized clinical trials. *Drug Information Journal* 2001;35:1201-1214.