Roger Harbord

roger.harbord@bristol.ac.uk
www.epi.bris.ac.uk/staff/rharbord/

A modified test for small-study effects

in meta-analyses of controlled trials with binary endpoints

Acknowledgements

- Co-authors
  - Jonathan Sterne, University of Bristol
  - Matthias Egger, University of Bern

- Teachers
  - Anne & John Whitehead, University of Reading

Aim

- *Develop a modified version of the Egger test that:*
  - has better controlled false-positive rate

  while keeping:
  - reasonable statistical power
  - simplicity

Outline

- What are small-study effects?
- How can we detect them?
- Why are existing techniques questioned?
- Is there a better method?
- How does it compare in simulations?

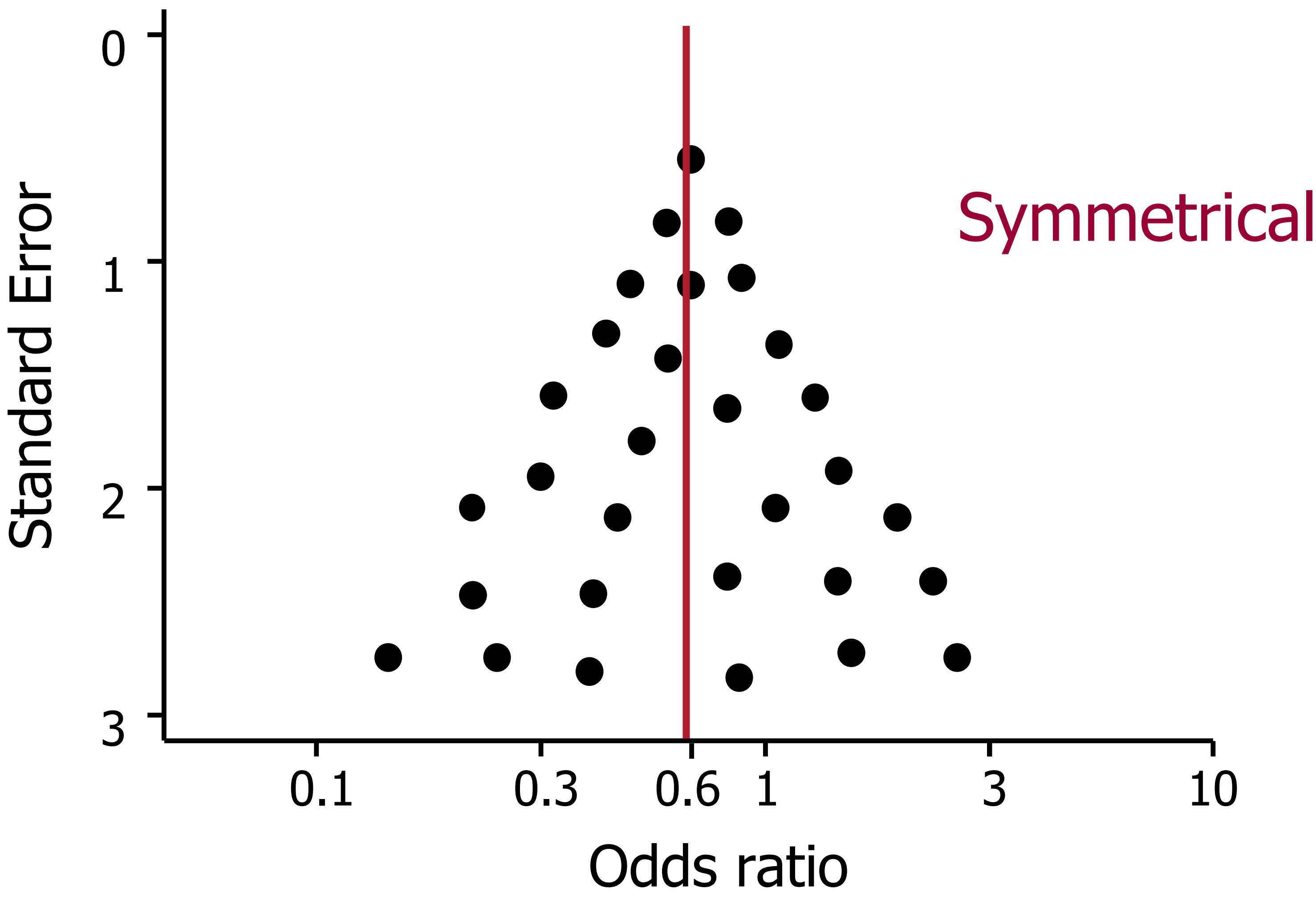
- Summary

Small-study effects

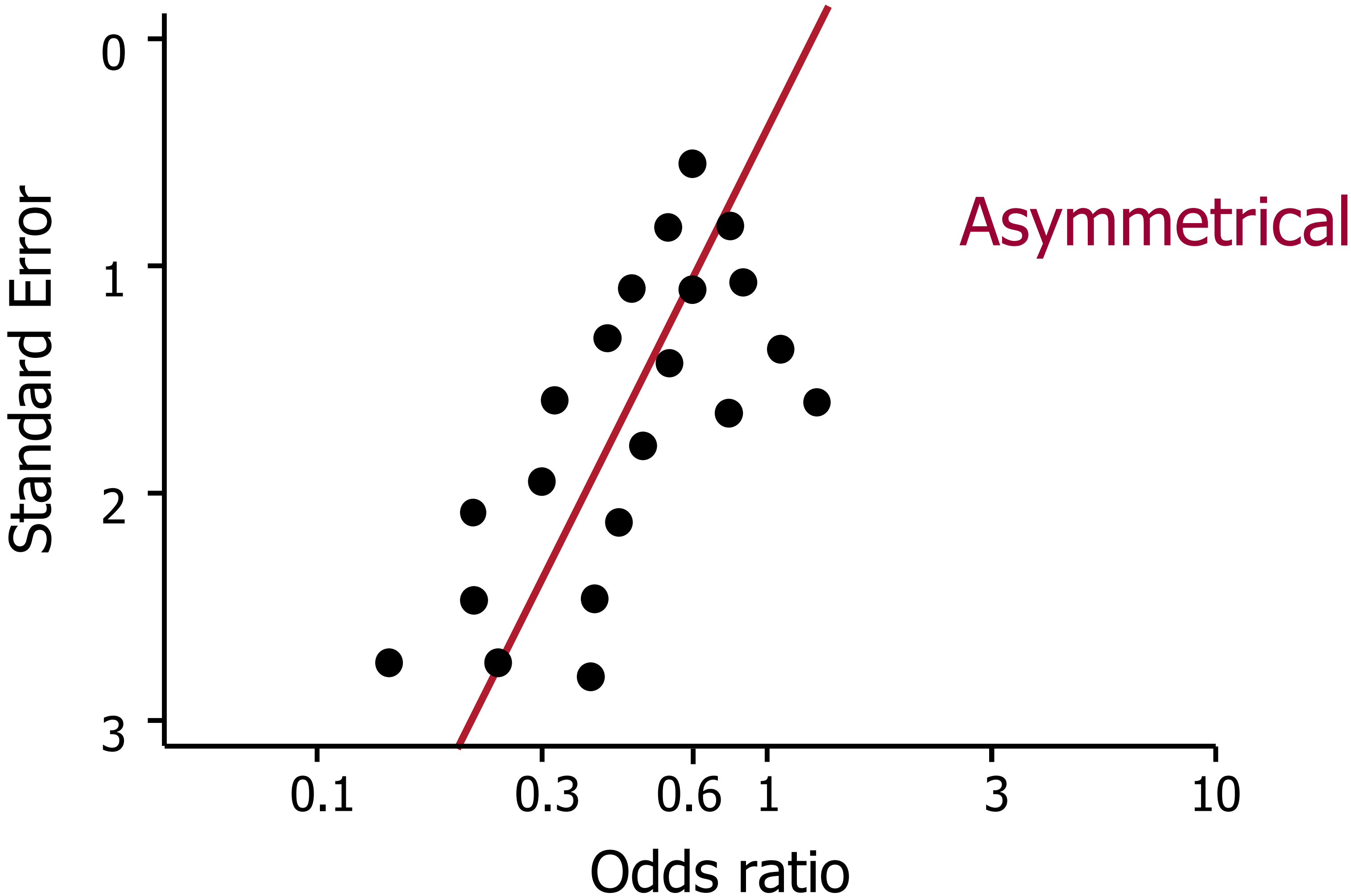
- **a tendency for smaller trials in a meta-analysis to show greater treatment effects than the larger trials**

- May be due to:
  - Publication bias
  - Smaller trials having poorer quality
  - Genuine differences in treatment effects

Funnel plot – no bias



Funnel plot – bias present



Egger test: definition

θ : treatment effect (e.g. log odds ratio)

- Regress θ on SE(θ) with weights 1/Var(θ)
- *t*-test of slope = 0

equivalently:

- Regress θ /SE(θ) on 1/SE(θ) without weights
- *t*-test of intercept = 0

(Egger *et al.* BMJ 1997, Sterne *et al.* J Clin Epi 2000)

2×2 table

	Disease	Healthy	Total
Treatment	d_1	h_1	n_1
Control	d_0	h_0	n_0
Total	d	h	n

log odds ratio $\theta = \log\left(\frac{d_1 / h_1}{d_0 / h_0}\right)$

$$\mathrm{SE}(\theta) = \sqrt{\frac{1}{d_1} + \frac{1}{h_1} + \frac{1}{d_0} + \frac{1}{h_0}}$$

$\boldsymbol{\theta}$ and SE($\boldsymbol{\theta}$) are instrinsically correlated for binary endpoints

	Disease	Healthy
Treatment	~~18~~ 19	~~2~~ 1
Control	15	5

$$\theta = \log OR = \log\left(\frac{\overset{1/19}{\cancel{2/18}}}{5/15}\right) = \log(0.\overset{0.16}{\cancel{33}}) = \overset{-1.8}{\cancel{-1.10}}$$

$$SE(\theta) = \sqrt{\frac{1}{\underset{19}{\cancel{18}}} + \frac{1}{\underset{1}{\cancel{2}}} + \frac{1}{15} + \frac{1}{5}} = \underset{1.15}{\cancel{0.91}}$$

Macaskill test

d_1	h_1	n_1
d_0	h_0	n_0
$\boldsymbol{d}$	$\boldsymbol{h}$	$\boldsymbol{n}$

θ *:* log odds ratio

- Regress θ on n with weights dh / n
- t-test of slope = 0

- Properties
  - Better control of false-positive rate
  - Lower power

(Macaskill *et al.* Statist. Med. 2001)

A modified regression test

d_1	h_1	$\boldsymbol{n_1}$
d_0	h_0	$\boldsymbol{n_0}$
$\boldsymbol{d}$	$\boldsymbol{h}$	$\boldsymbol{n}$

- Define:
  - Efficient score $Z = "O - E" = d_1 - d\, n_1 / n$
  - Score variance (Fisher's information)

$$V = \frac{n_0\, n_1\, d\, h}{n^2 (n-1)}$$

- Regress Z / V on $\sqrt{V}$
- 2-sided *t*-test of intercept = 0

13

Z and *V* have much lower sampling correlation

	Disease	Healthy	
Treatment	~~18~~ 19	~~2~~ 1	20
Control	15	5	20
	~~33~~ 34	~~7~~ 6	40

$$Z = \overset{19}{\cancel{18}} - \frac{20 \times 33}{40} = \overset{19}{\cancel{18}} - 16.5 = \overset{2.5}{\cancel{1.5}}$$

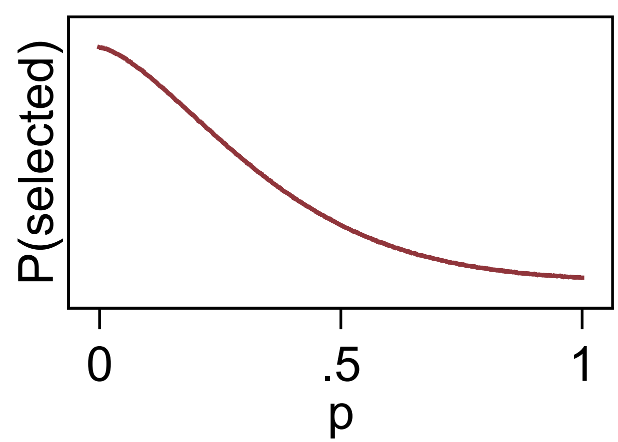
$$V = \frac{20 \times 20 \times \overset{34}{\cancel{33}} \times \overset{6}{\cancel{7}}}{40^2 \times (40 - 1)} = \overset{1.31}{\cancel{1.48}}$$

Design of simulations

based on those of Macaskill *et al.* Statist Med 2001,
Terrin *et al.* Statist Med 2003

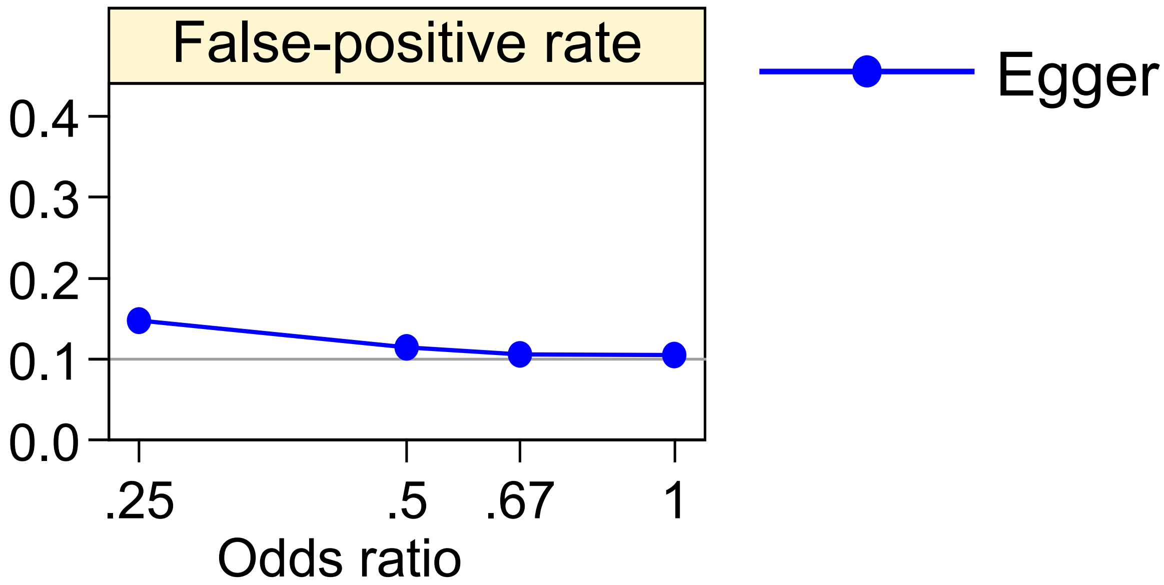
- 20 studies per meta-analysis
  - 11 × 100/group, 6 × 200/group, 4 × 300/group
- In control group, P(event) ~ U(0.1, 0.5)
- Set *OR* and between-study variance τ^2
- Simulate meta-analyses from binomials
  - 10 000 without selection
  - 10 000 'strong' selection: $P(\mathit{selected}) \propto \exp(-4p^{3/2})$

    (Begg & Mazumbdar Biometrics 1994)

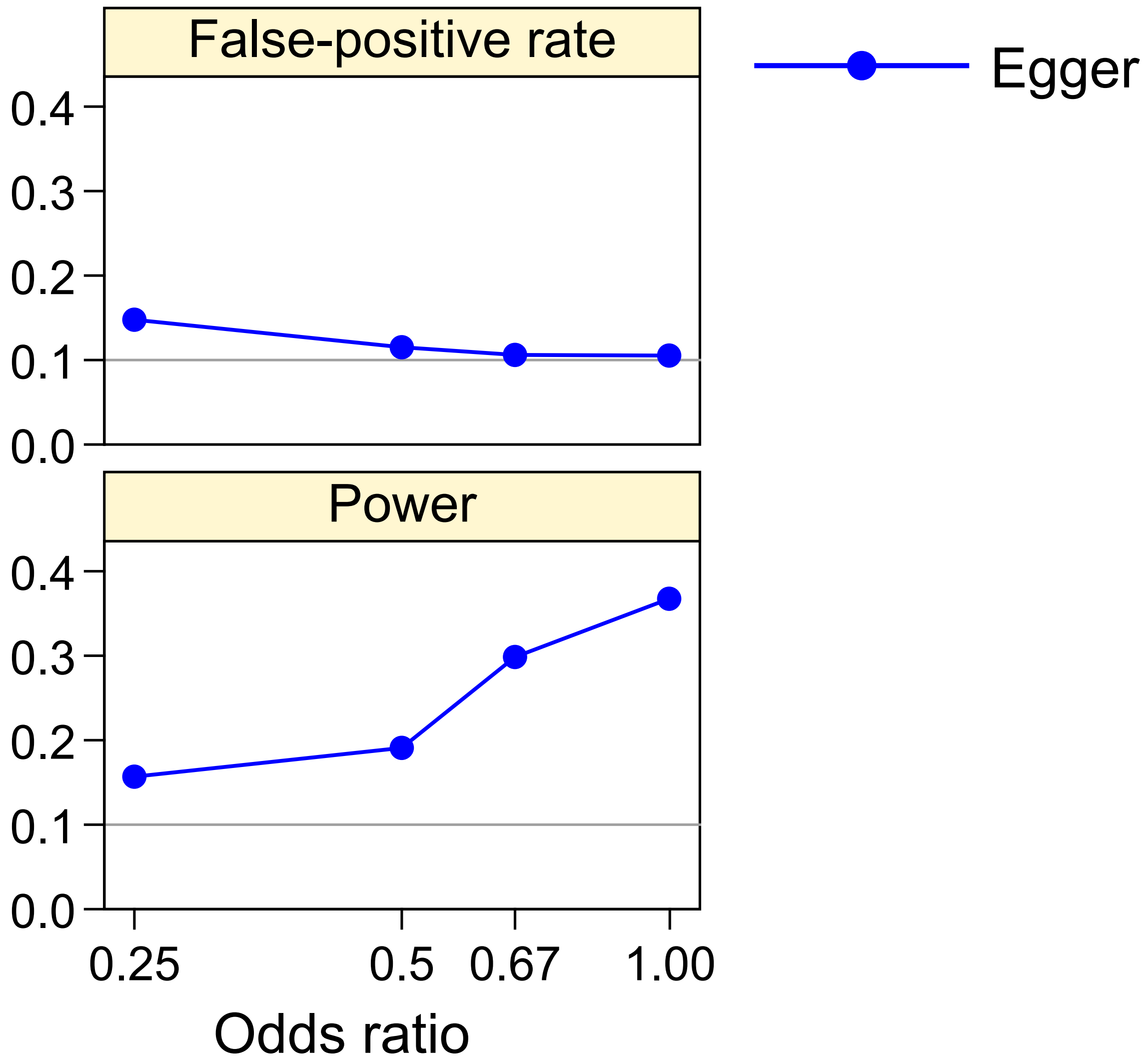


Results of simulations

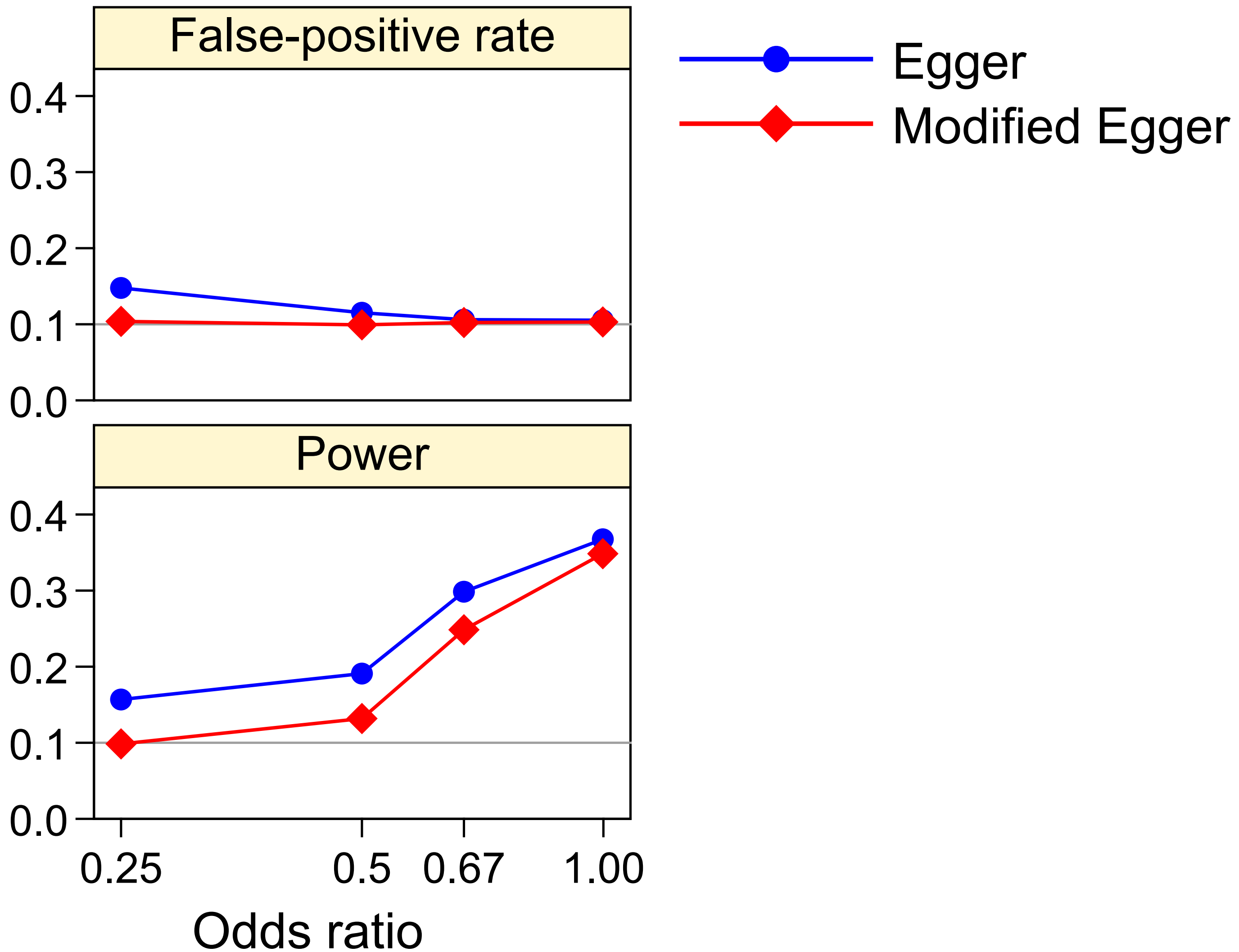
– no heterogeneity



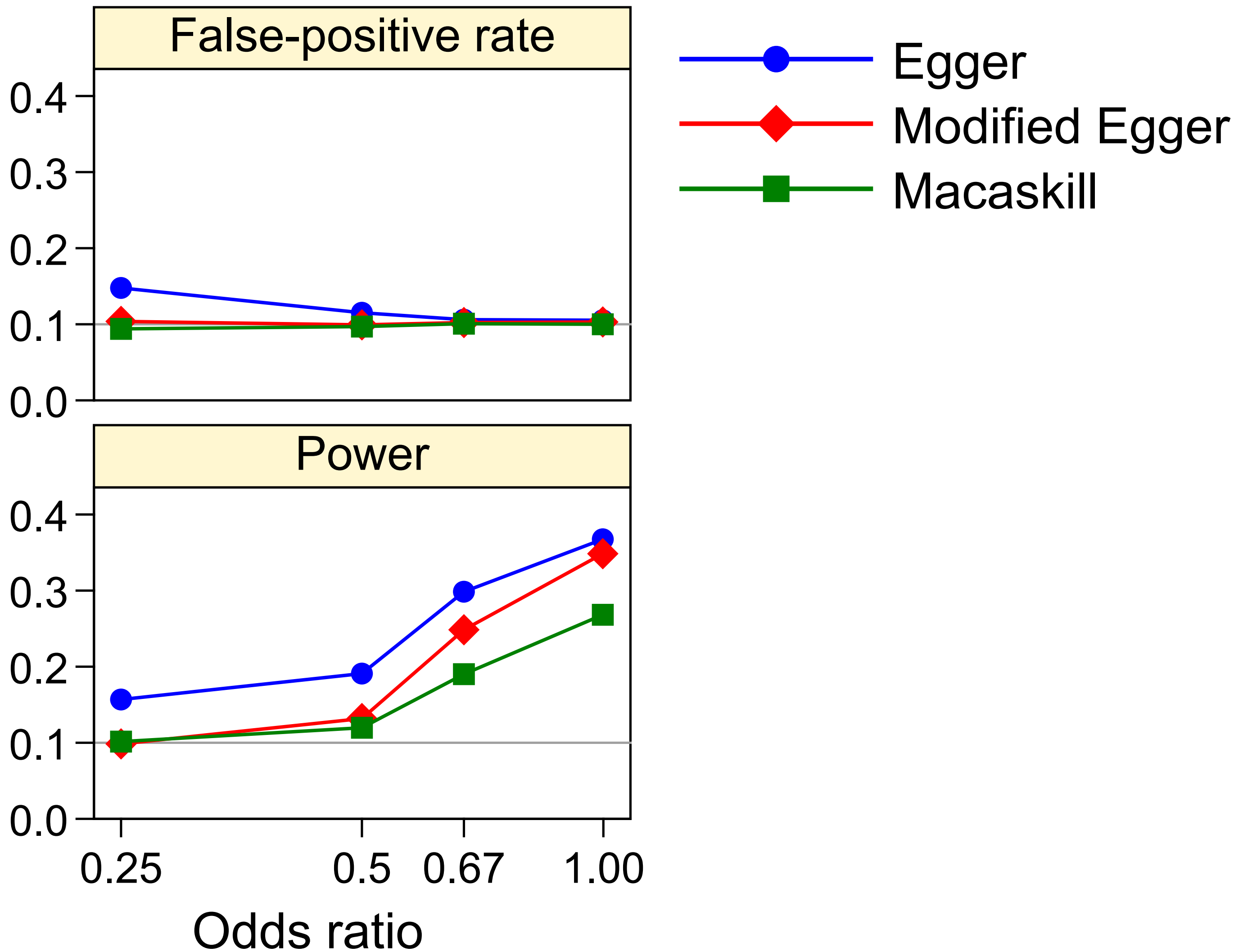
Results of simulations
– no heterogeneity



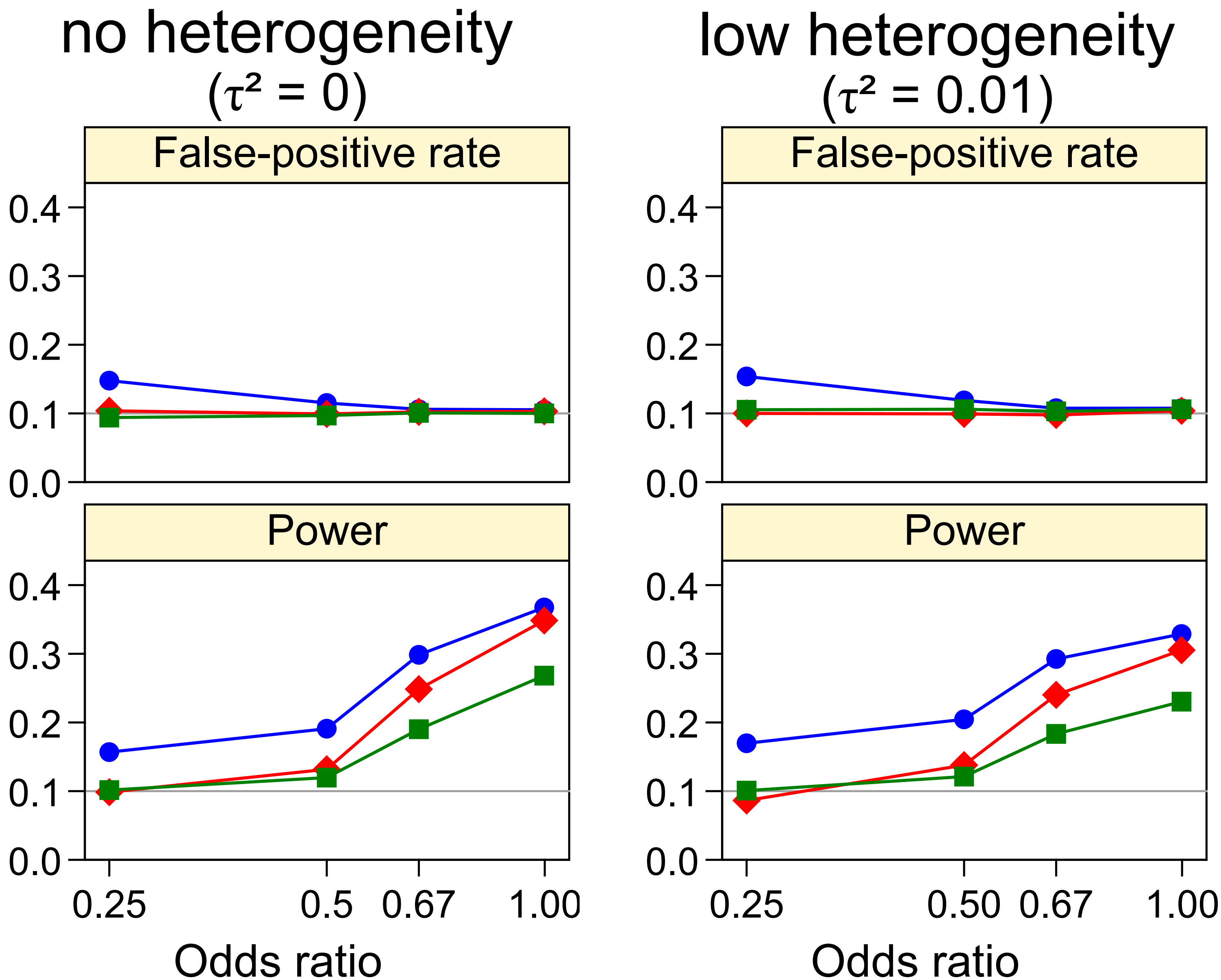
Results of simulations
– no heterogeneity



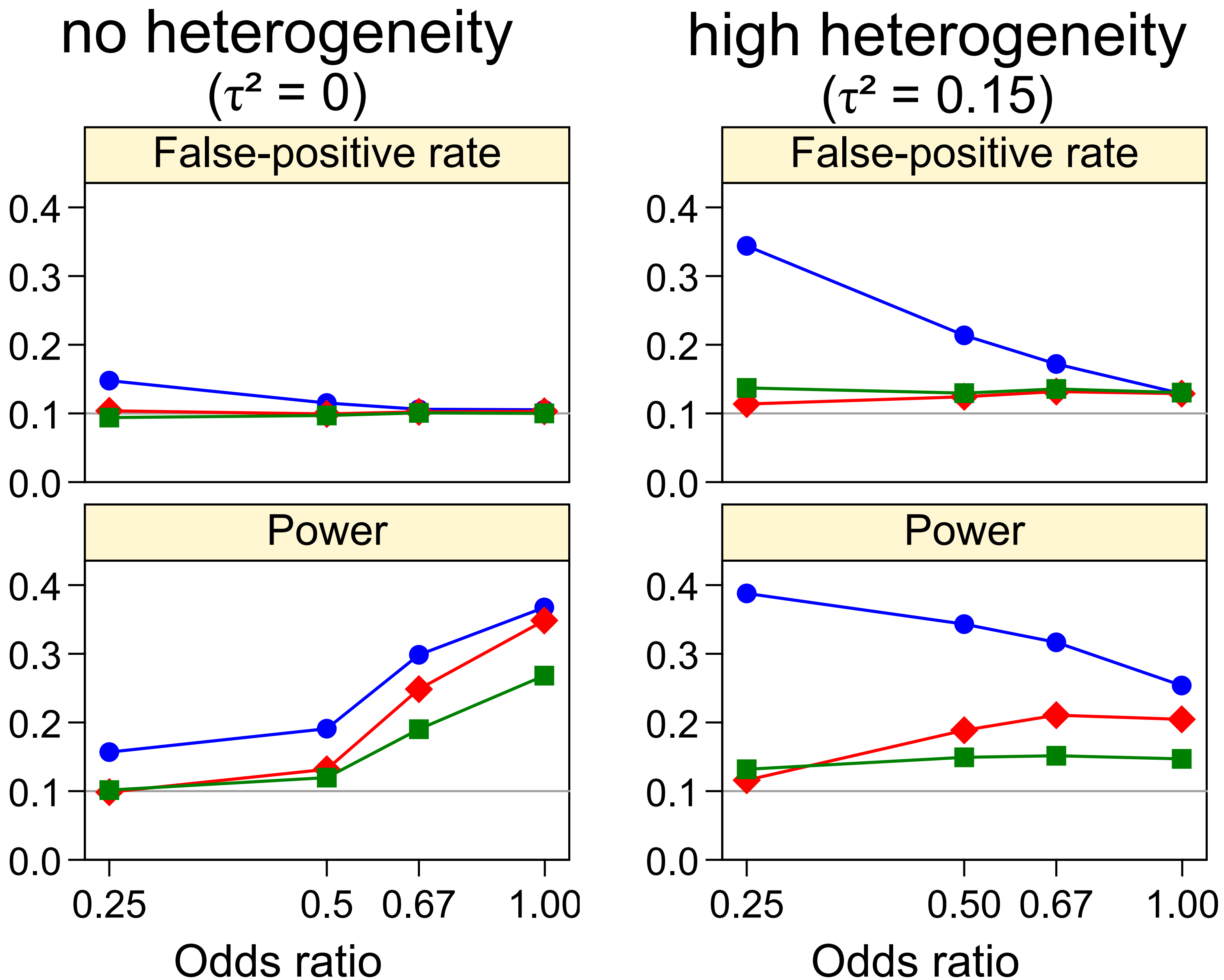
Results of simulations
– no heterogeneity



Results of simulations



Results of simulations:



Summary of further simulations

- Greater variation in study sizes:
  - Increases false-positive rates of all three tests
  - Macaskill test worse than Egger test with heterogeneity
  - Modified Egger still has lowest false-positive rate, but around 0.2 when $\tau^2 = 0.15$ so not acceptable
- Further simulations based on 78 published meta-analyses:
  - suggest modified test has acceptable false-positive rate for τ^2 less than around 0.04

Simulations – summary of results

- Modified test has:
  - ✓ lower false-positive rate than Egger test
  - ✓ similar power

× None of the tests work well with considerable heterogeneity
(τ^2 more than about 0.04)

Not assessed in simulations

- Other measure of treatment effect
  - Simulations only for log odds ratios
  - theory applies to other effect measures

- Properties in unbalanced trials
  - likely to be poor if imbalance high – e.g. diagnostic studies, cohort studies

Summary

- Funnel plots look at study-size effects
- Tests based on them can have problems
- New test greatly reduces one problem
- All tests poor if heterogeneity substantial

Harbord RM, Egger M, Sterne JAC. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Statistics in Medicine* (in press).

Preprint at www.epi.bris.ac.uk/staff/rharbord

Roger Harbord

roger.harbord@bristol.ac.uk
www.epi.bris.ac.uk/staff/rharbord/

A modified test for small-study effects

in meta-analyses of controlled trials with binary endpoints