Statistical methods for reliably updating meta-analyses

Mark Simmonds

Centre for Reviews and Dissemination
University of York, UK

With:
Julian Elliott, Joanne McKenzie, Georgia Salanti,
Adriani Nikolakopoulou, Julian Higgins

THE UNIVERSITY of York
Centre for Reviews and Dissemination

Conflict of interest

This project was funded by:

Cochrane Methods Innovation Fund

I am the co-creator of one of the methods considered

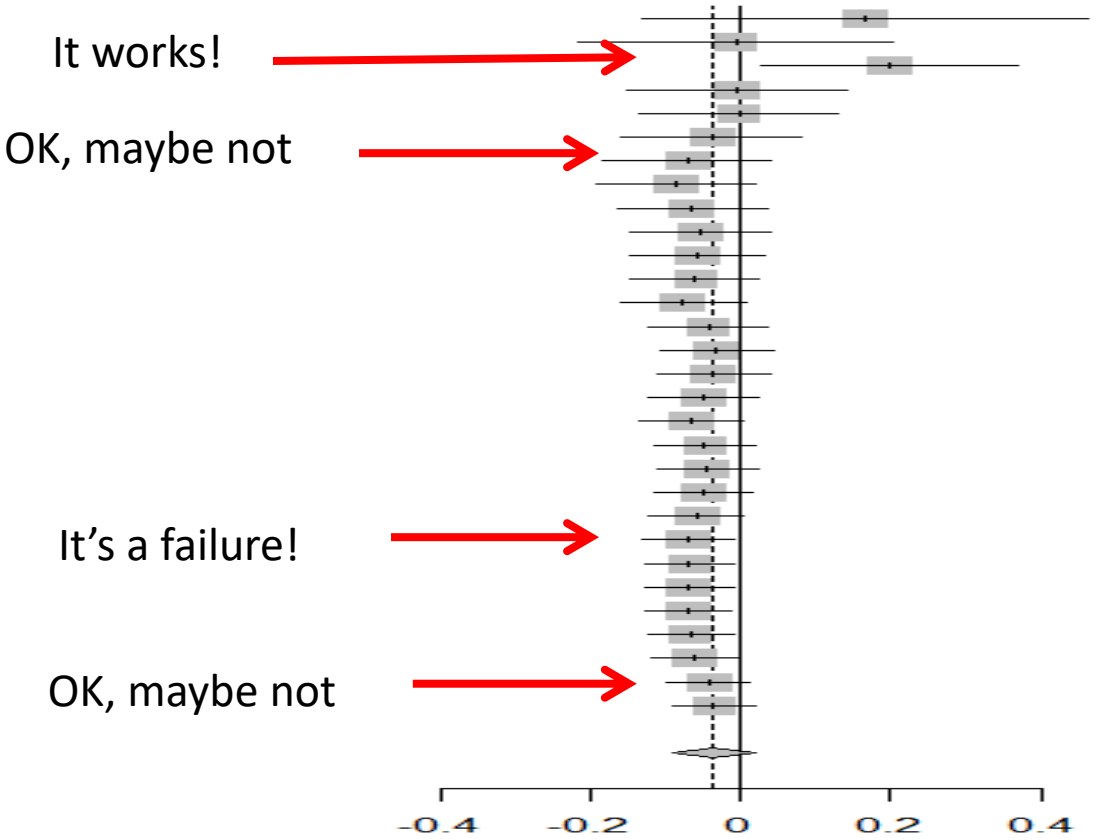
I have no financial or other conflicts of interest in relation to this presentation

Updating meta-analyses

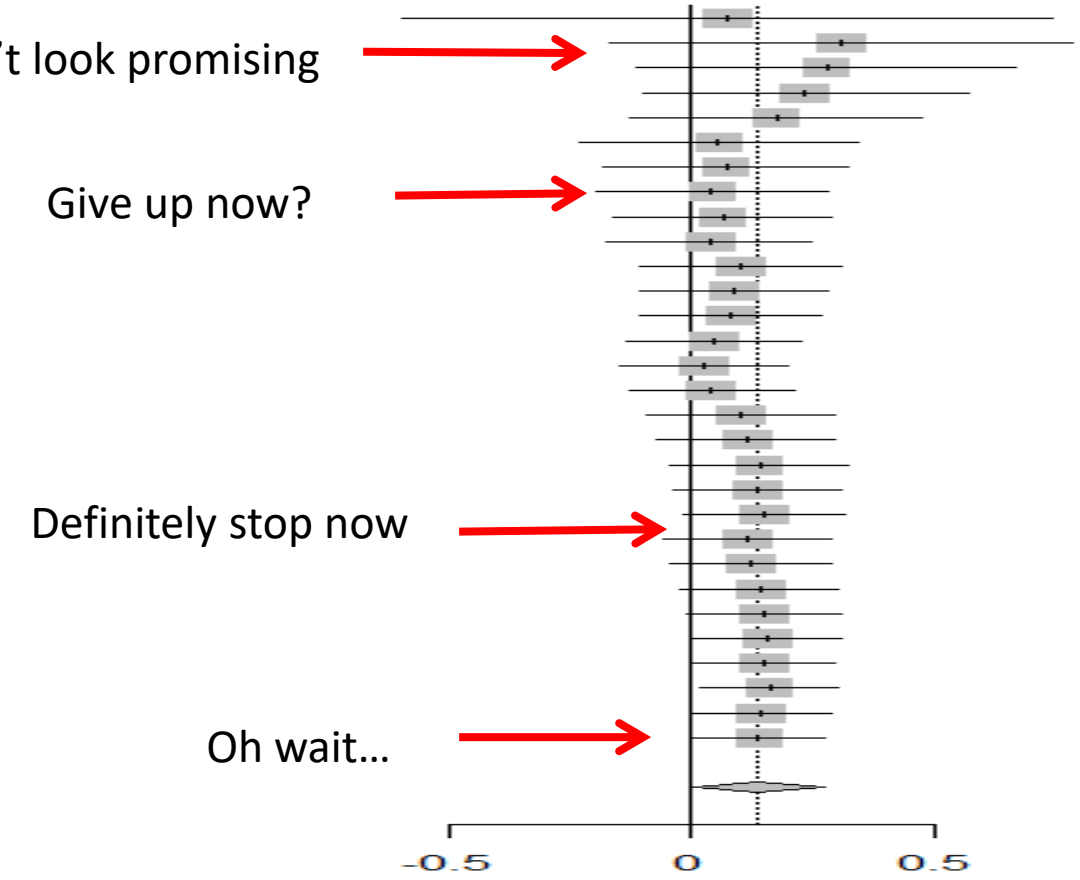
- When should we update a Cochrane review and meta-analysis?
  - Every two (or more/fewer) years?
  - As soon as new studies emerge?
  - When new data might alter our conclusions?
- Updating is time-consuming
- Conclusions can change over time
  - Risk of error if:
  - We make conclusions based on limited/poor data
  - We stop updating too soon
- Are the results robust?

Cumulative meta-analysis: false conclusions

Type I error



Type II error



Controlling error

- Control Type I and Type II error
  - **Sequential Meta-Analysis** *Higgins et al. Statistics in Medicine 2011: 30 903-921*
  - **Trial Sequential Analysis** *Wetterslev et al. J Clinical Epidemiology 2008: 61 64-75*
- Control Type I error
  - **Law of Iterated Logarithm** *Hu et al. Clinical Trials 2007: 4 329-340*
  - **"Shuster-Pocock" method** *Shuster and Neu. Research Synthesis Methods 2013: 4 269-279*
- Other methods
  - Fully Bayesian analysis
  - Consequences of adding new studies
  - Power gains from adding new studies

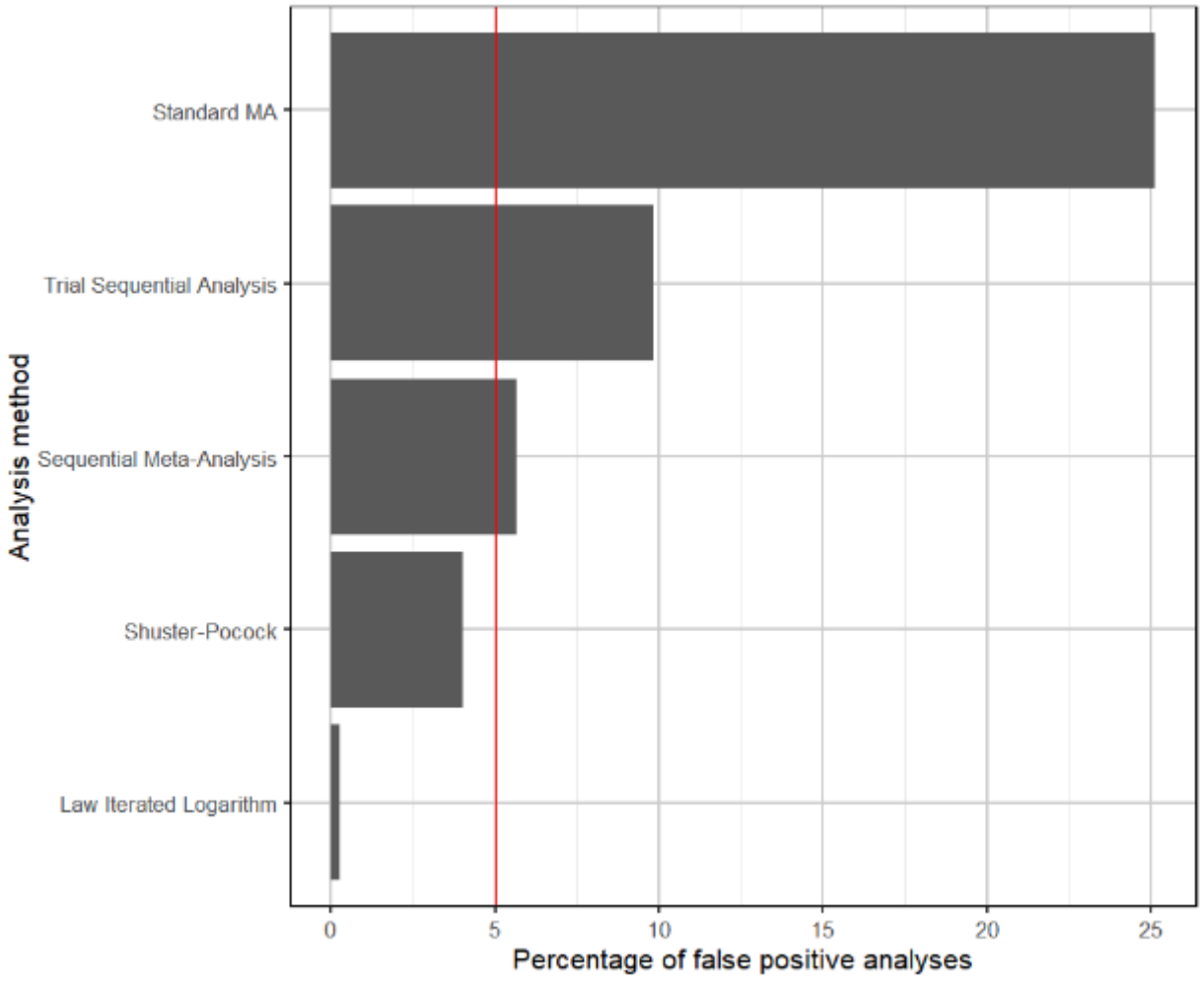
Simulation study

- Simulated meta-analyses varying:
  - True treatment effect: 0 or 0.1
  - Number of studies: 5 to 50
  - Heterogeneity: I^2 0 to 90%

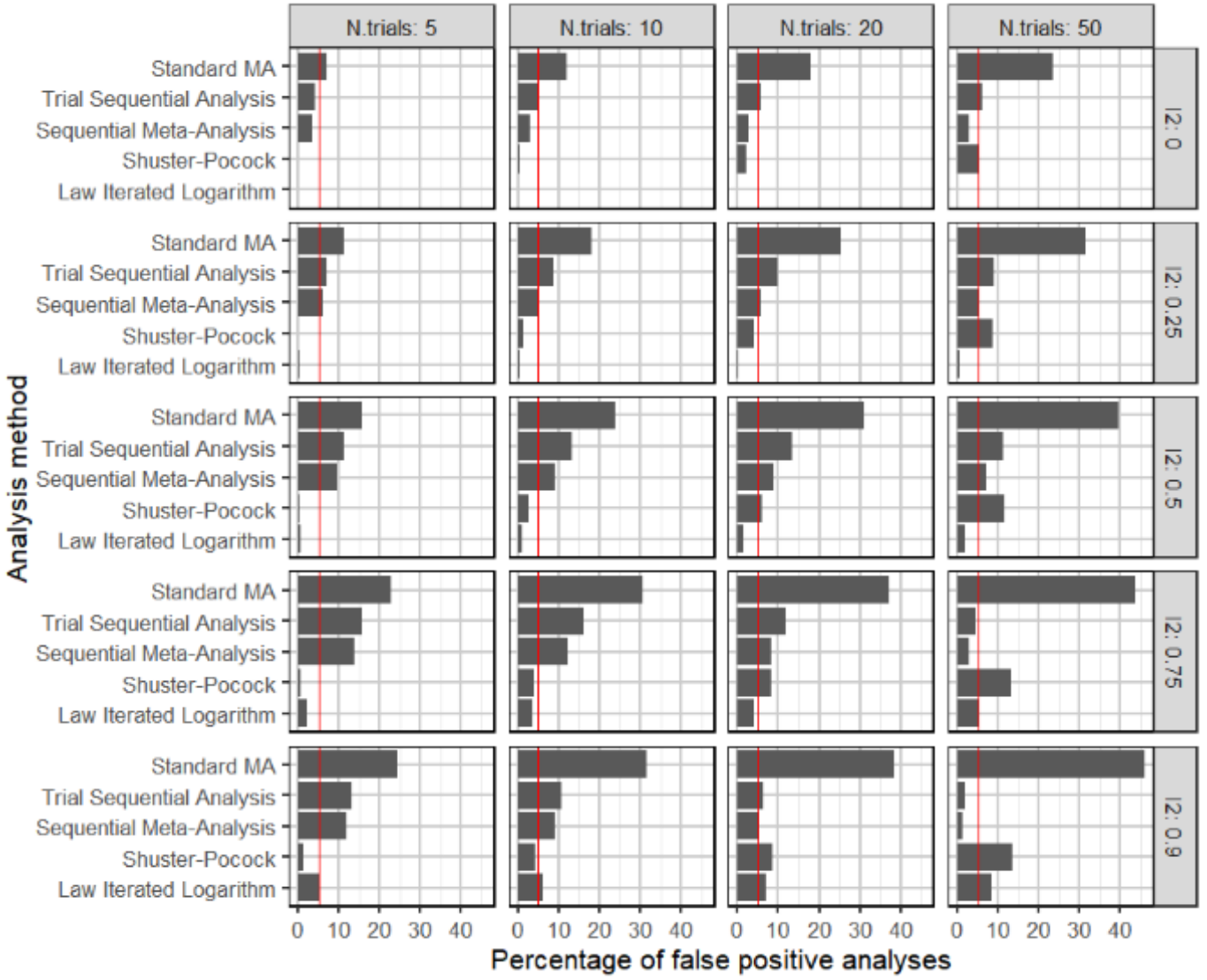
- Fixed total sample size of 9000
  - 90% power to detect effect of 0.1 if $I^2 = 50\%$

False positive rates – Type I error

- 20 trials / updates
- $I^2 = 25\%$

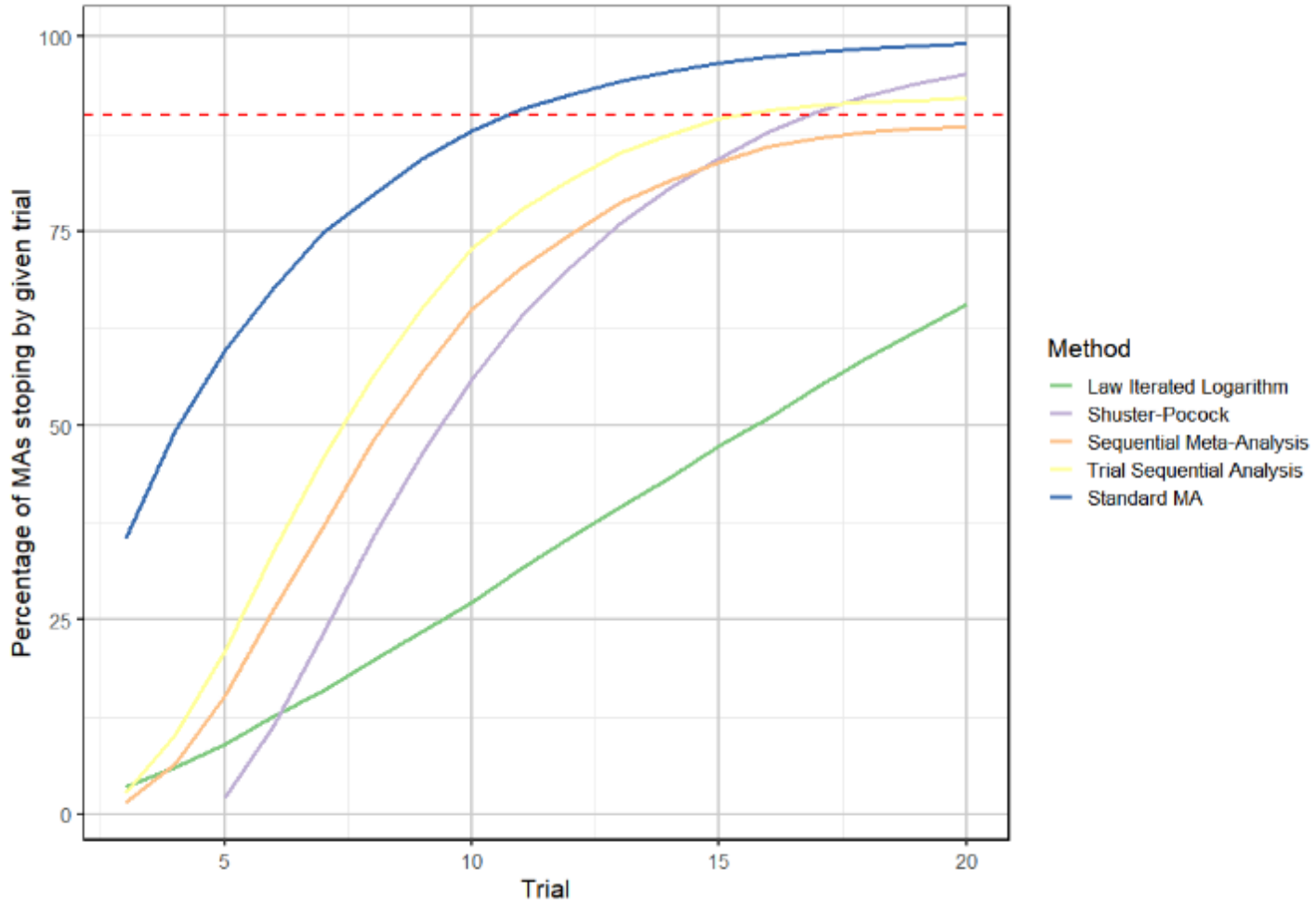


False positive rates – Type I error

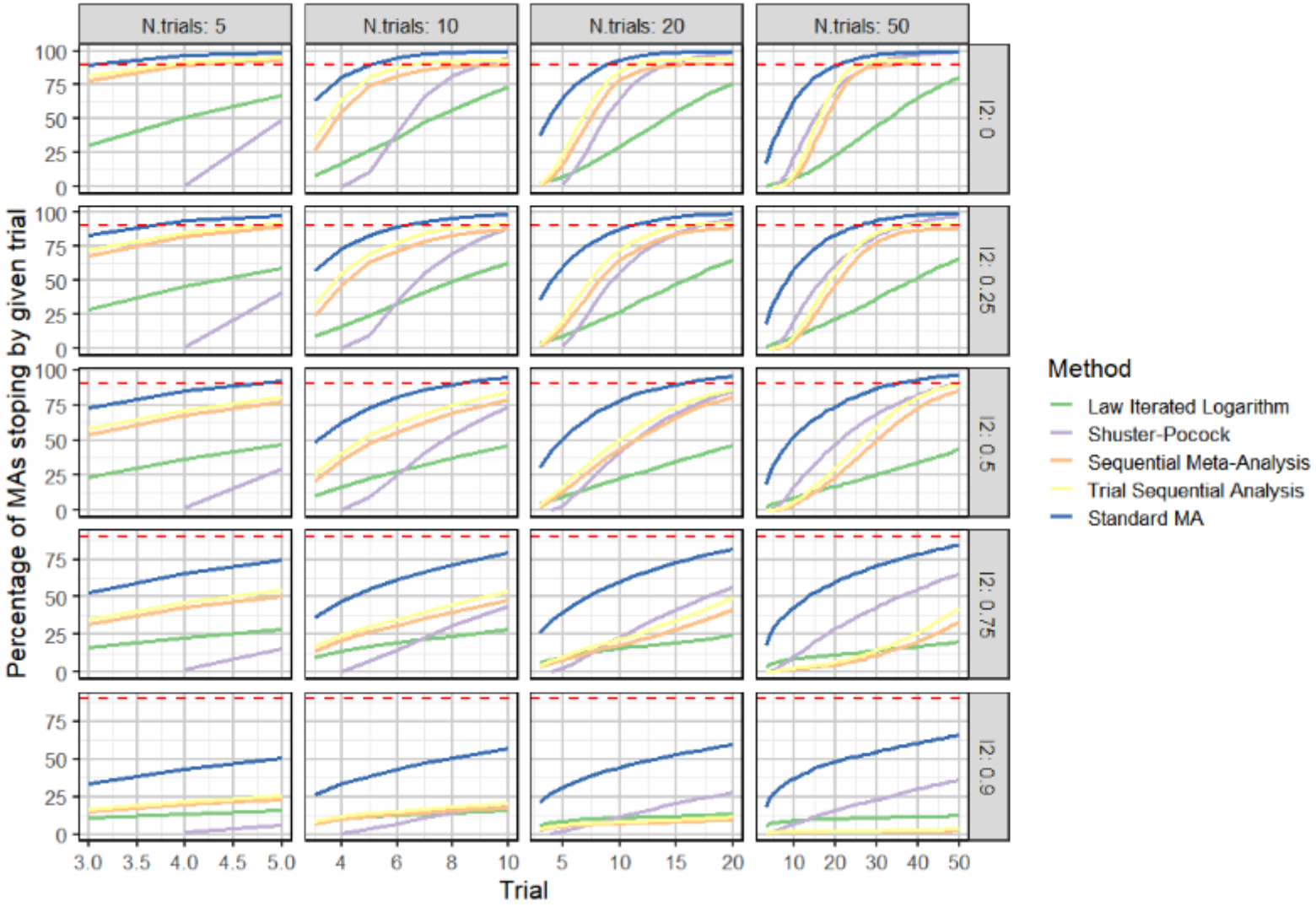


Cumulative power

- 20 trials / updates
- $I^2 = 25\%$



Cumulative power

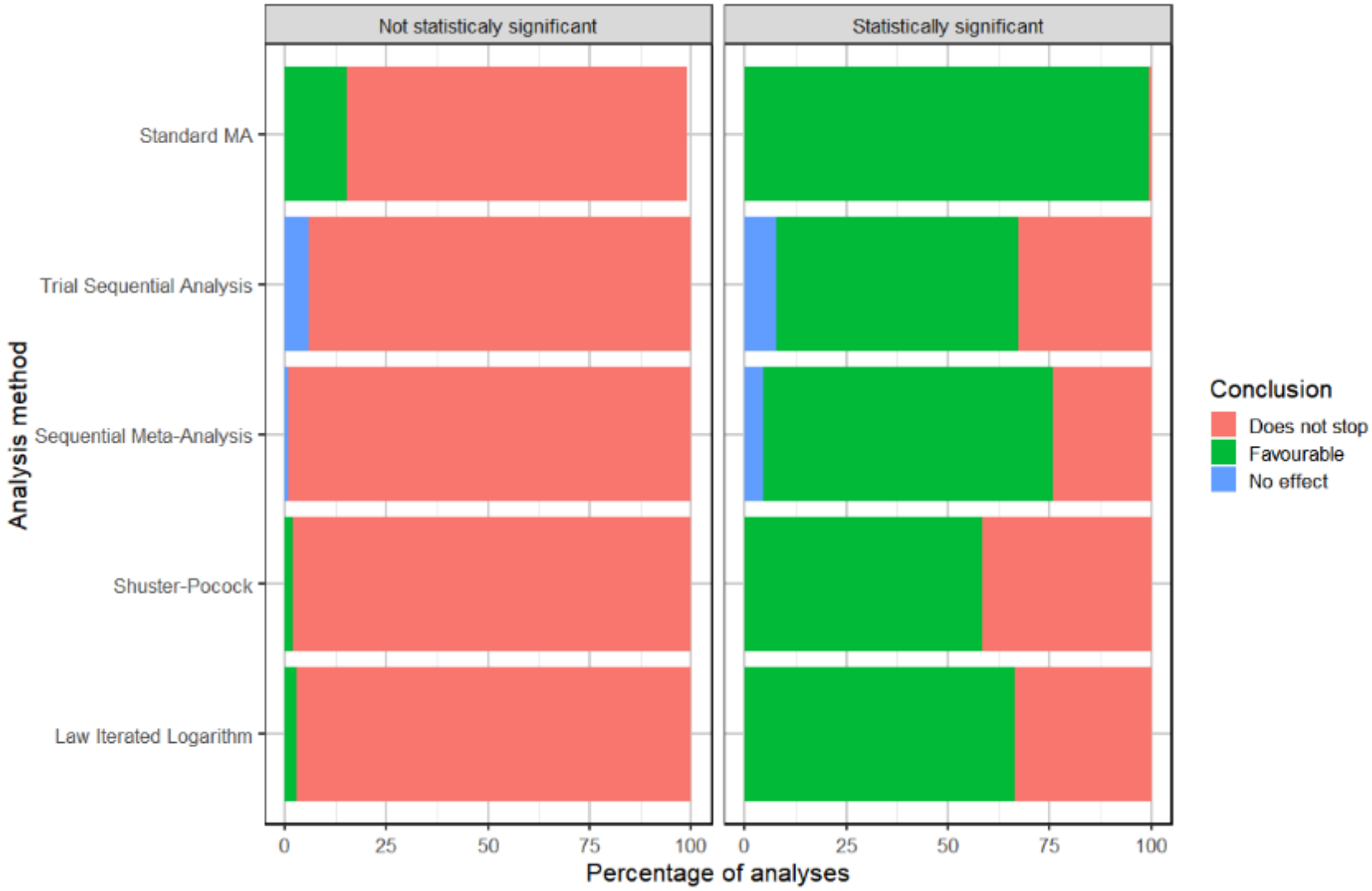


76 Cochrane Reviews

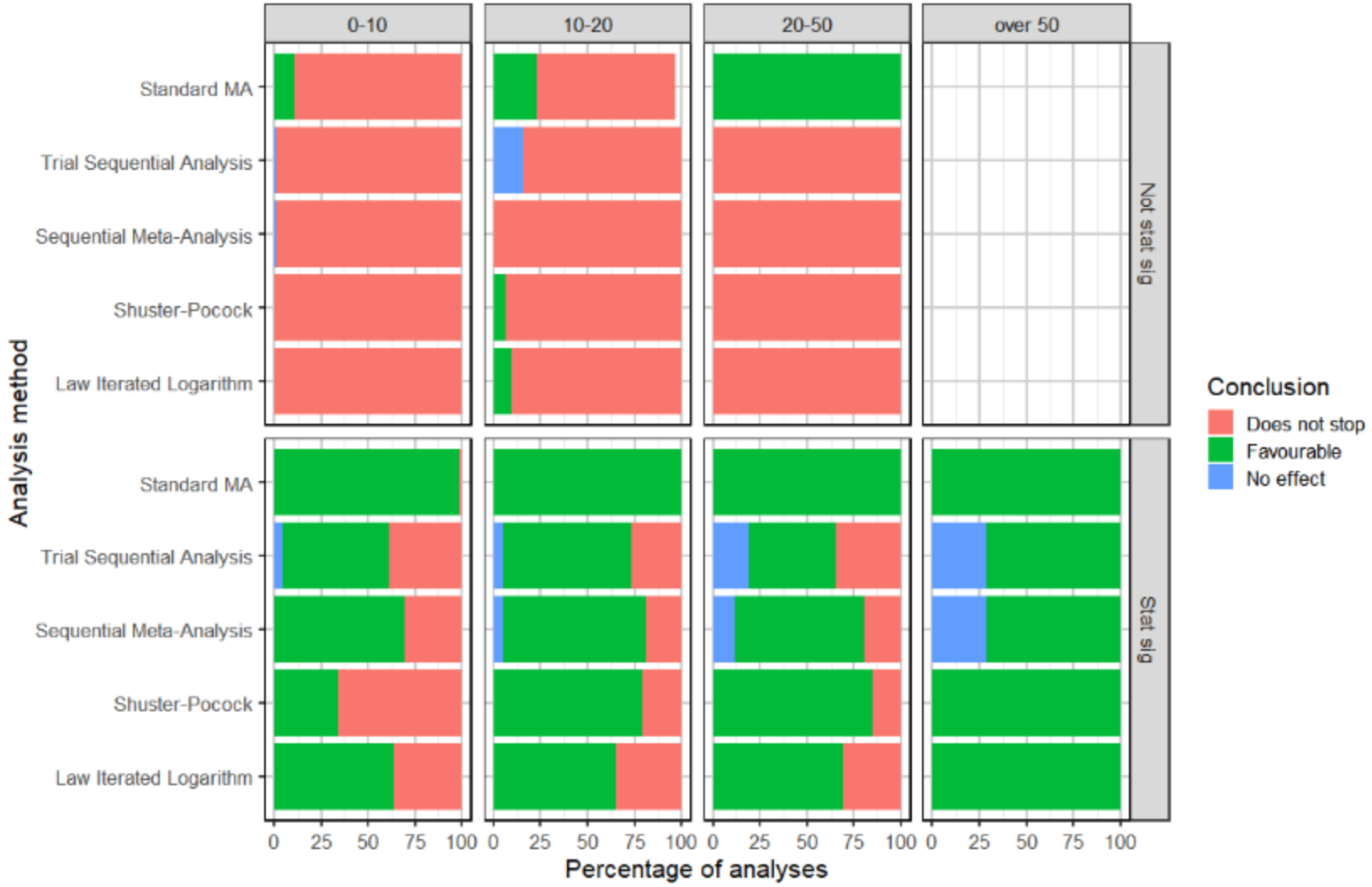
- 76 Reviews: 286 meta-analyses
  - 68% binary data
  - Median 9 trials (IQR 6 to 14)

- 62% had a statistically significant result using conventional analysis

Conclusions of analyses

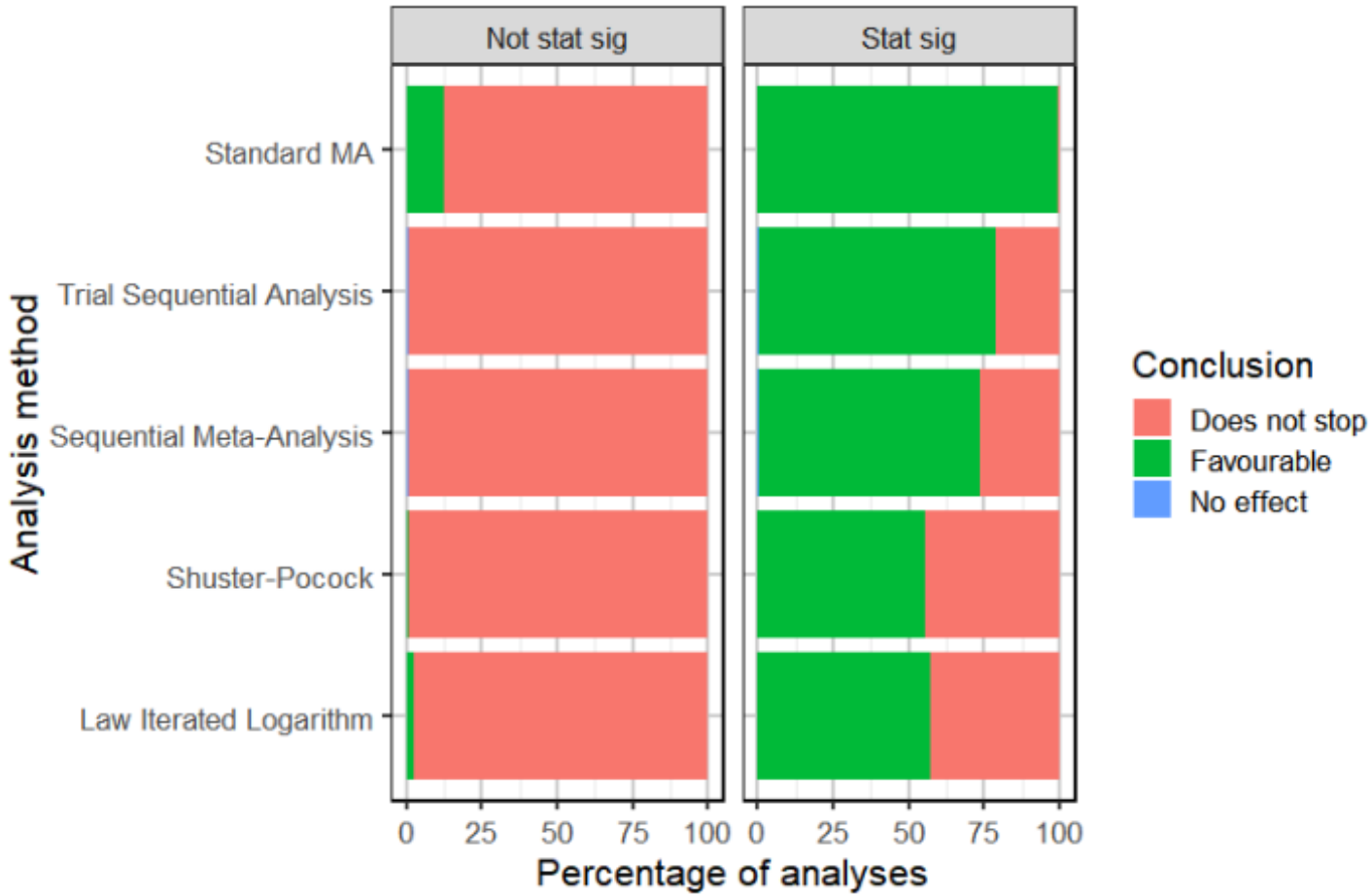


Conclusions of analyses



Realistic updating

- At most 4 updates
- After 50, 70, 90 and 100% of trials published



Conventional review updates

- Too many inappropriate positive conclusions
  - Elevated Type I error rate
  - But not vastly elevated for most real updated reviews?

- Many analyses showing “significant” results are based on too little evidence

Do we need sequential methods?

- Is the problem with standard reviews serious enough in real Cochrane reviews?
- Do the methods needlessly delay a statistically significant result?
- Should we avoid “statistical significance” altogether?

Practical conclusions for Cochrane reviews

- Remember that results may change over time
- Be cautious about interpreting “statistical significance”
  - Particularly with small sample sizes
- Consider the required sample size and statistical power
- Sequential methods may be useful in some reviews
  - Big effects but small sample sizes
  - Frequently updated or automated reviews