# 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

# Some issues

- When can we stop updating a review?
- Conclusions can change over time
  - Risk of error if we stop too soon
- Type I error inflated by performing multiple analyses

# Controlling error

- Adapted from sequential clinical trial design
  - Sequential meta-analysis *(Higgins, Simmonds, Whitehead 2010)*
    - Includes Bayesian adjustment of heterogeneity
  - Trial sequential analysis *(Wetterslev, Thorlund, Brok, Gluud 2008)*
- Control Type I error
  - Law of Iterated Logarithm *(Lan, Hu, Cappelleri 2007)*
  - "Shuster-Pocock" method *(Shuster, Neu 2013)*
- Other methods
  - Fully Bayesian analysis
  - Robustness or stability of analysis
  - Consequences of adding new studies
  - Power gains from adding new studies



# Analyses of updated Cochrane reviews

- Searched for Cochrane reviews:
  - Updated in 2014-2015
  - At least one new trial added
  - At least one meta-analysis
    - That is statistically significant
    - At least 3 trials

- Included 76 reviews and 286 meta-analyses
  - 62% had statistically significant results
  - 44%  were of sufficient size to have 80% power to detect observed effect.

# Assumptions

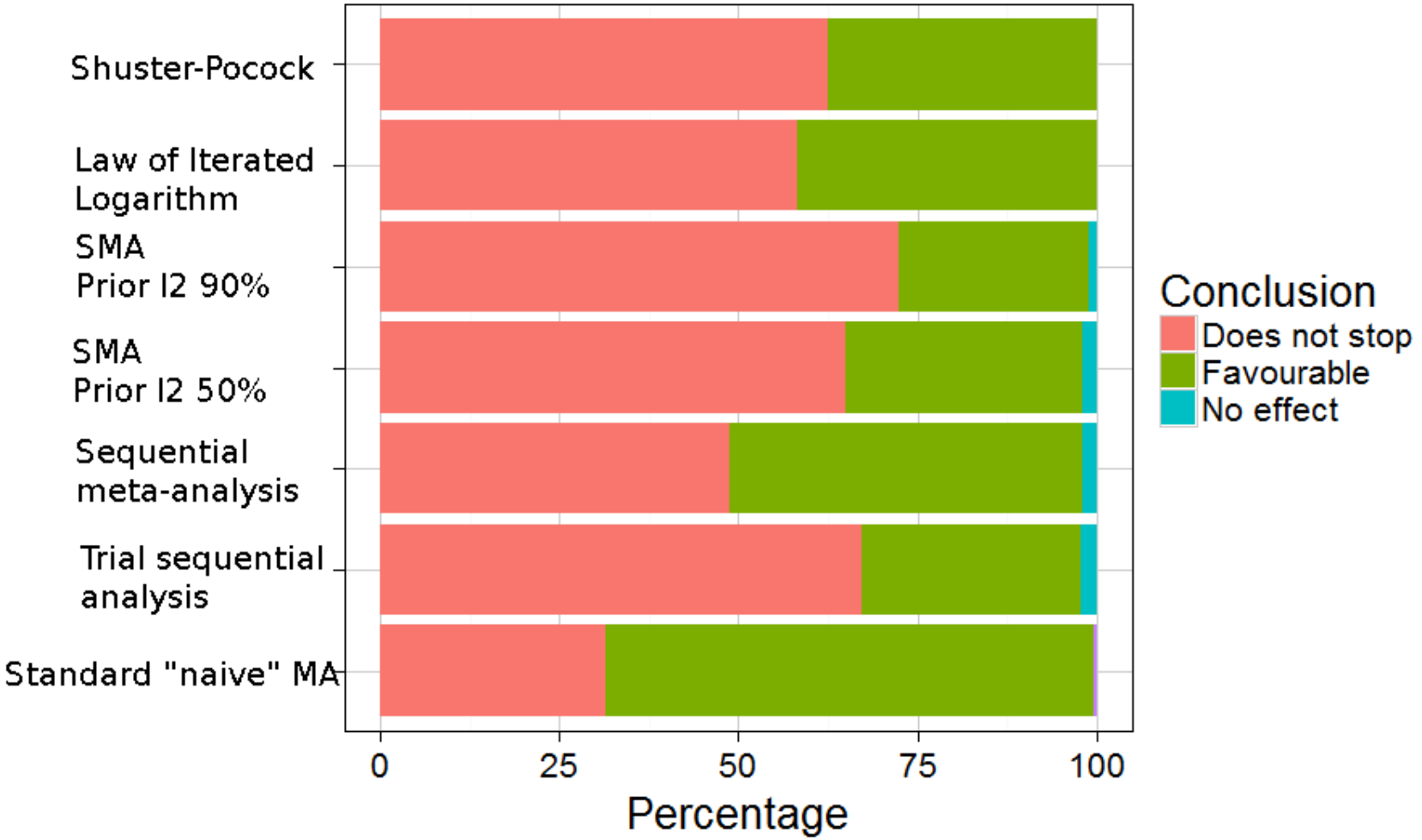
- Analysis using log odds ratio or SMD
- A new meta-analysis for each added trial

- 5% Type I error, 90% power
- “Desired” effect is same as observed

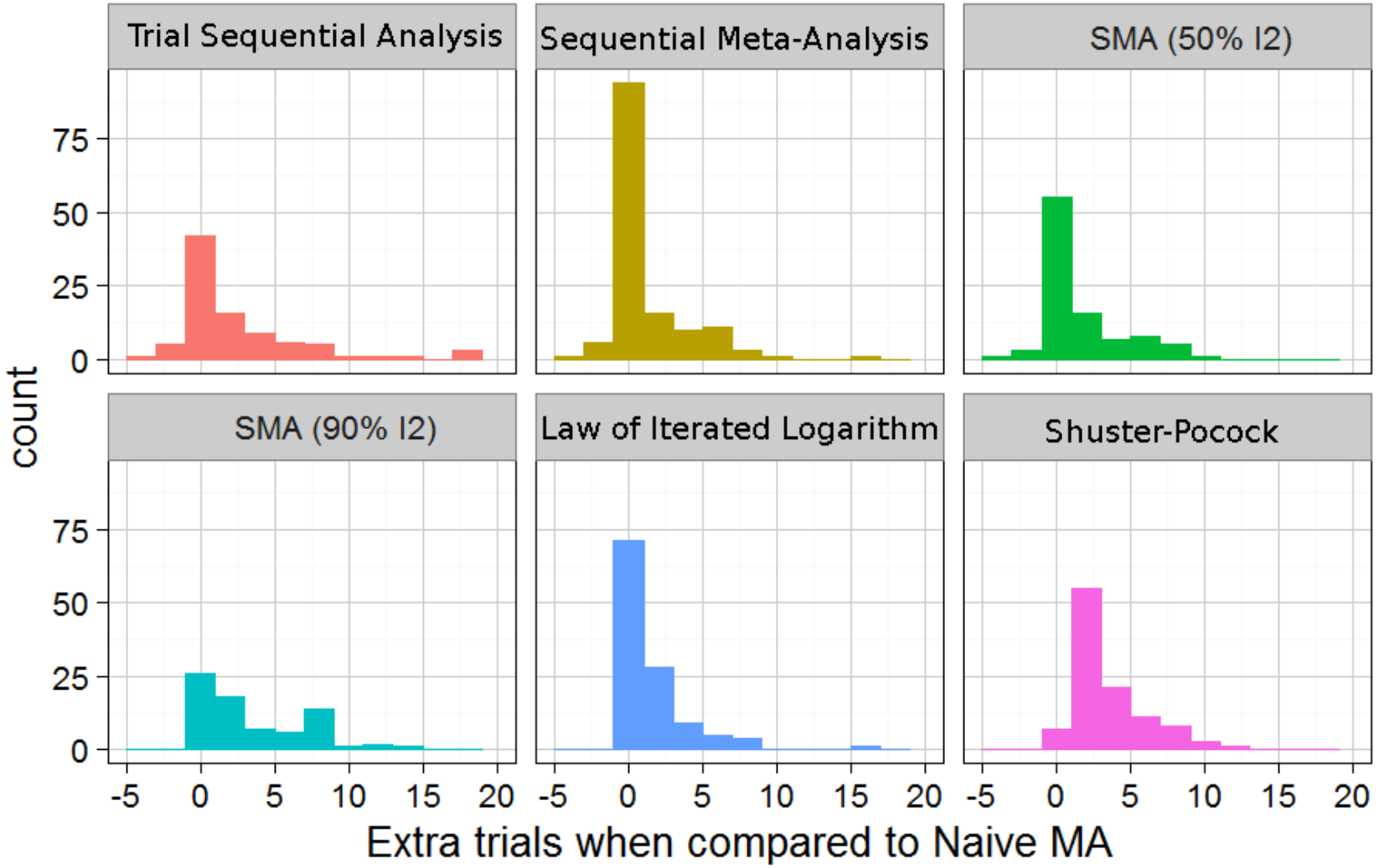
- Meta-analyses are uncorrelated



# Conclusions of analyses



# Additional trials to reach a conclusion



# “Inappropriate positives”

Conclusions of updated meta-analysis
where analysis with all trials is not statistically significant

| Method | Does not stop | Evidence of effect | No evidence of effect |
|---|---|---|---|
| Naïve MA | 83.8 | 15.2 | - |
| Trial Sequential Analysis | 99.0 | 0 | 1.0 |
| Sequential Meta-Analysis | 99.0 | 0 | 1.0 |
| SMA (50% I2) | 99.0 | 0 | 1.0 |
| SMA (90% I2) | 100 | 0 | 0 |
| Law of Iterated Logarithm | 98.1 | 1.9 | - |
| Shuster-Pocock | 98.1 | 1.9 | - |

# Conventional “Naïve” analysis

- Too many inappropriate positive conclusions
  - Elevated Type I error rate
  - But not vastly elevated for most updated reviews?
- Biased estimates of effect
- Significant results are often based on too little evidence?

# Controlling for error

- All methods appear to control for Type I error

- Increased complexity
- Need to select desired effect size, adjust for heterogeneity etc.

- May take longer before stopping

# Do we need these methods?

- Is the problem with “naïve” analysis serious enough in real Living Systematic Reviews?
- Do the methods needlessly delay a statistically significant result?
- When should they be implemented?
  - As part of protocol?
  - Only with statistically significant results?

# Implications for Living Systematic Reviews

- Reviews with many updates
  - Increased risk of type I error
  - Methods probably needed
- Starting with few trials
  - Need to identify required sample size
  - Methods needed as a caution if results statistically significant?
- Starting with many trials
  - Little new data expected, update for consistency
  - Methods not needed?