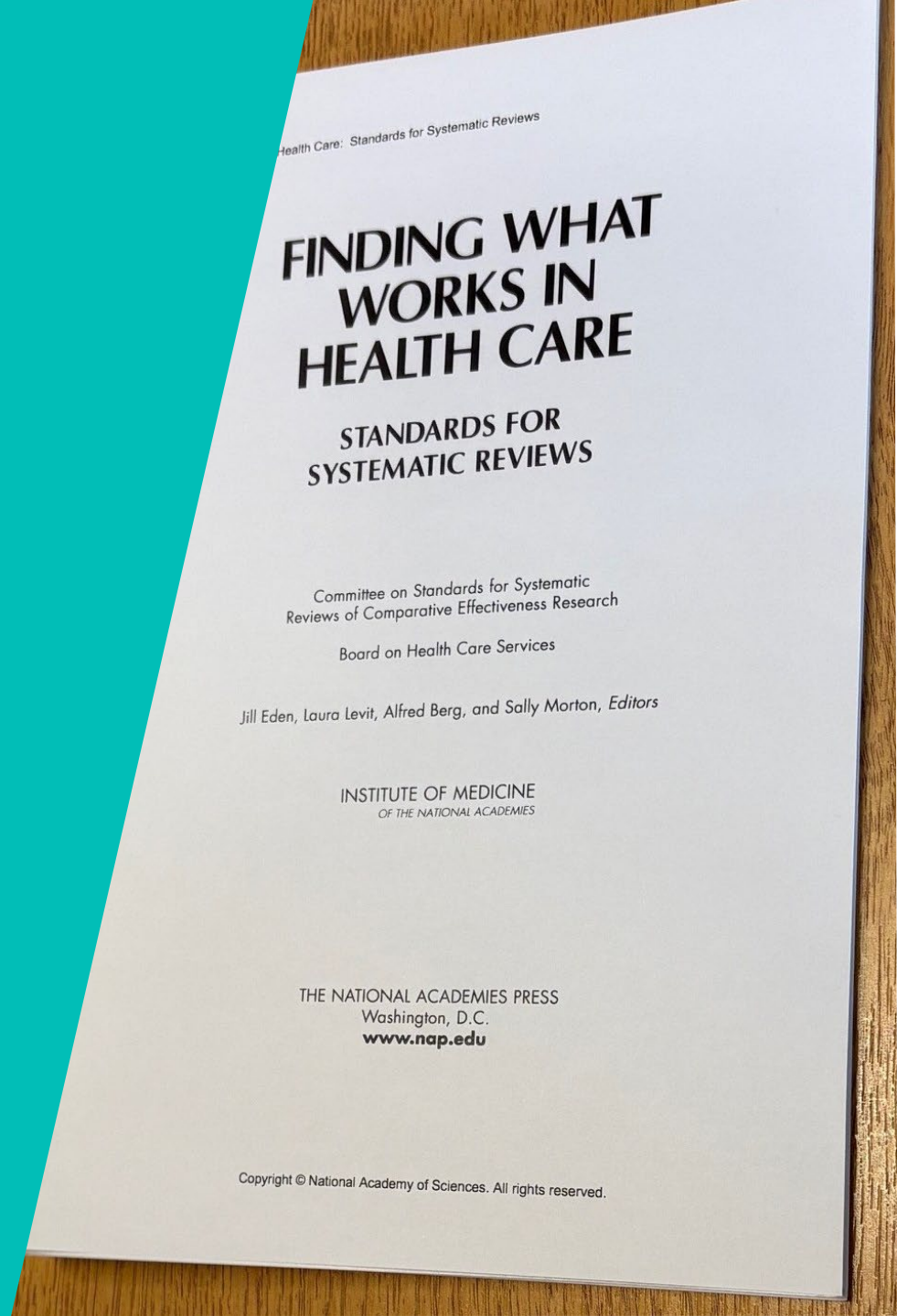
University of BRISTOL

NEW SYSTEMATIC REVIEW METHODS SINCE 2011

A personal perspective

Julian Higgins

University of Bristol, Bristol, UK



Disclosures

- Senior Editor, *Cochrane Handbook for Systematic Reviews of Interventions*
- Developer of methods and tools (including risk-of-bias assessment tools)
- User of methods and tools (>100 applied evidence synthesis papers)
- Commented on development documents for 2010 version of the standards

- I will restrict my remarks to systematic reviews on the *effects of interventions*

Some key methodological developments

- A general reflection is that methods have not advanced dramatically

- I shall offer a few remarks on
  - AI
  - Risk of bias and research integrity
  - The role of observational data
  - Meta-analysis methods

AI: it will be (is?) better than humans

- Impact on methods:
  - potentially huge
  - in the short term:
    - recommendation for two independent humans?
    - teams might not screen all search results
    - search can be broader (it might not even be needed)

3.3.3 Use two or more members of the review team, working independently, to screen and select studies

3.5.1 At a minimum, use two or more researchers, working independently, to extract quantitative and other critical data from each study. For other types of data, one individual could extract the data while the second individual independently checks for accuracy and completeness. Establish a fair procedure for resolving discrepancies—do not simply give final decision-making power to the senior reviewer

- Need **evaluation data sets** and **evaluation systems**
- I think we need to keep humans in the loop (for now)

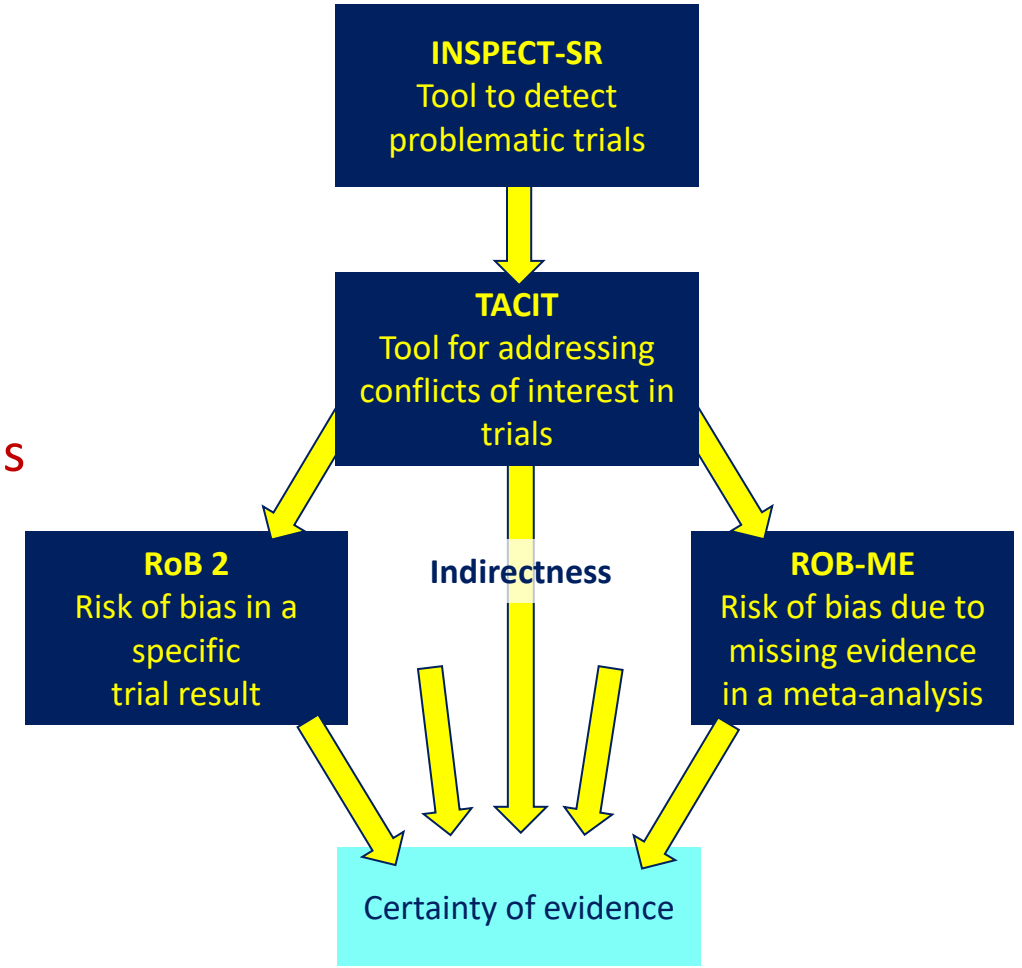
- Increase in *open access publication* also impacts on methods for searching, screening and study selection (often in conjunction with AI)

3.3.5 Use one of two strategies to select studies: (1) read all full-text articles identified in the search or (2) screen titles and abstracts of all articles and then read the full texts of articles identified in initial screening

Risk of bias and research integrity

3.6.1 Systematically assess the risk of bias, using predefined criteria

- Tools to examine **integrity** now available (vested interests, misconduct, fraud)
- We have a **new generation of tools** to assess risk of bias in estimates of intervention effect, rooted in improvements in understanding of causal inference
  - important focus on defining what the absence of bias looks like
  - defining the intervention effect of interest, or 'estimand'
- Beware 'simplification' of tools
  - Systematic reviews of intervention effects should be done to the *highest standards*, not the *easiest standards*



The role of observational data

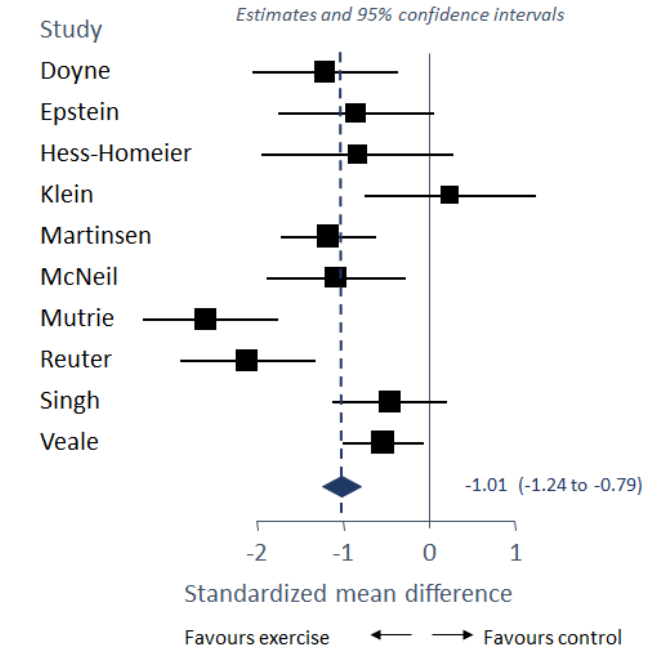
3.3.2 Use observational studies in addition to randomized clinical trials to evaluate harms of interventions

- Rebadging data as “real-world” does not make them more reliable
- But newer (and some older) methods can help us understand the limitations
  - e.g. target trial emulation (e.g. instrumental variables)
- Increasing focus on clarity of the question being asked (defining the estimand)

- This thinking is entering meta-analysis (which is welcome) – will lead to better synthesis of observational data (“causal meta-analysis”)

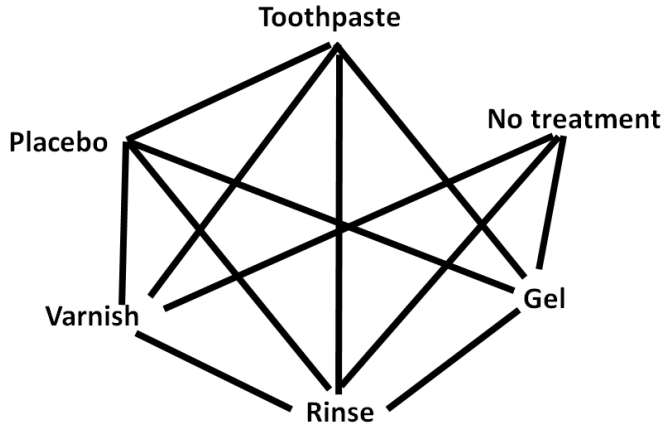
Meta-analysis methods

- Core ideas unchanged
  - though we now have better estimators for between-study variance in random-effects meta-analysis



- Major expansion of **network meta-analysis**
  - important developments particularly in **component network meta-analysis** and other methods to understand **complex interventions** and complex data sets

- We have better methods for **individual-level effect modifiers**, when data are available (and useful proposals for how to get hold of data quicker and cheaper)



Other major developments

- Increased emphasis on **reproducibility**
- Introduction of PROSPERO **registration** (2011)
- Increase in **co-production**
- Explosion in **types** of evidence synthesis
  - particularly scoping reviews, evidence (and gap) maps
- Living systematic reviews
  - important, but not really *new*
  - reinforce the need for **automation** and **reproducible pipelines**
- **SWARs** (studies within reviews)
  - encourage development and implementation of these

- A frightening number of systematic reviews are performed – many of them poorly

Help people get the old things right

- Too many massively-used methods and tools are **misused**
- e.g.
  - PRISMA is not a guideline for doing systematic reviews
  - Egger's test (and similar tests) is not a test for publication bias
  - I-squared does not measure the amount of heterogeneity
  - random-effects meta-analyses concern distributions of effects, not single effects
  - **studies shouldn't be ignored if they don't provide data in the format you wanted**
  - etc…

Concluding remarks

- Should change:
  - encourage people to examine all the evidence (not just that with convenient data)
  - impact of AI on methods (particularly searching, screening, data extraction)
  - methods to address research integrity

- It seems to me the standards already adequately address:
  - risk of bias (3.6.1)
  - observational data (3.3.2, 3.3.6)
  - meta-analysis methods (4.4) – *though guidance should say more on network meta-analysis*
  - living systematic reviews (3.1.7)

- Avoid being prescriptive about methods or tools