
University of Pittsburgh
Introduction to Systematic Review and Meta-Analysis: A Health Care Perspective
Sally C. Morton
Department of Biostatistics
University of Pittsburgh
Methods for Research Synthesis:
A Cross-Disciplinary Approach, October 2013

Cross-Disciplinary Communication From the Healthcare Systematic Review World

- Terminology
- Institute of Medicine (IOM) standards for systematic reviews
  - Quality of individual studies
  - Heterogeneity
  - Strength of the body of evidence

How Did We Get Here?

"The Evidence Paradox" (Sean Tunis):

- **18,000+ RCTs published each year**
- **Tens of thousands of other clinical studies**
- **Systematic reviews routinely conclude that:**

"The available evidence is of poor quality and therefore inadequate to inform decisions of the type we are interested in making."

Terminology

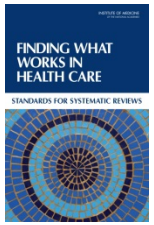
Systematic Review (SR): Review of a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant research, and to collect and analyze data from the studies that are included in the review

Meta-analysis (MA): Use of statistical techniques in an SR to integrate the results of included studies to conduct statistical inference

Adapted from the Cochrane Collaboration Glossary

Key Points

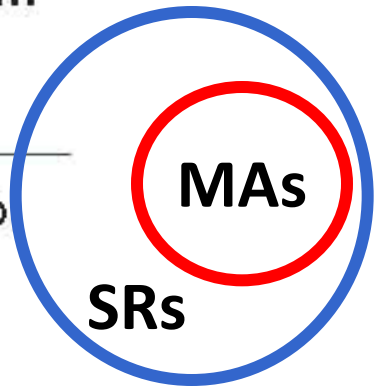
1. **MA should not be used as a synonym for SR**
2. **An MA should be done in the context of an SR**
3. **"An MA should not be assumed to always be an appropriate step in an SR. The decision to conduct an MA is neither purely analytical nor statistical in nature."**



STANDARD 4.3

Decide if, in addition to a qualitative analysis, the systematic review will include a quantitative analysis (meta-analysis)

4.3.1 Explain why a pooled estimate might be useful to decision makers

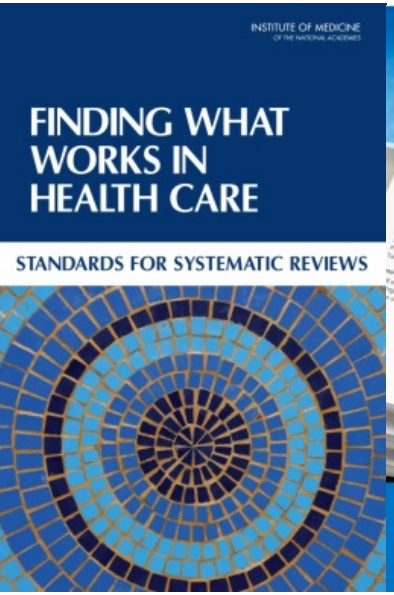


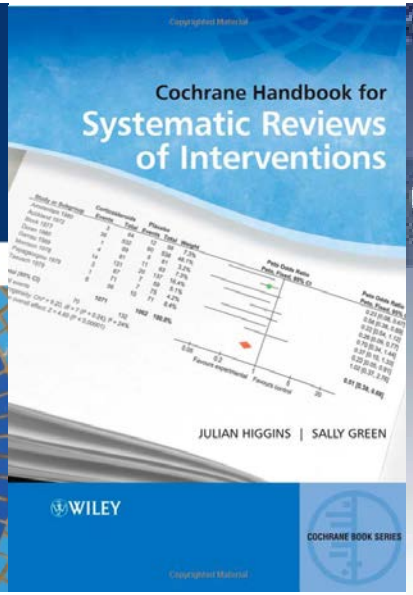
Steps of a Systematic Review

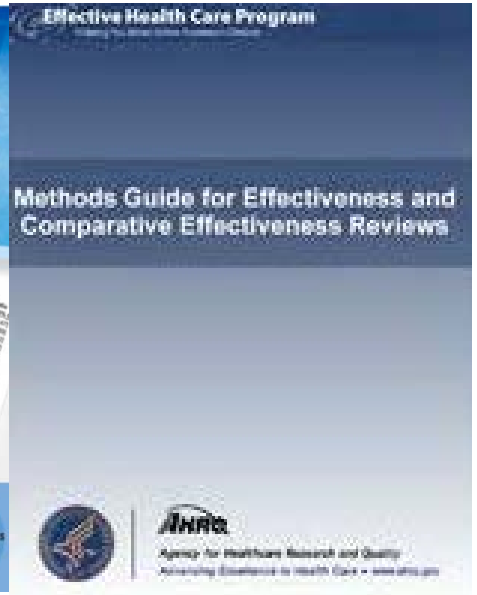
1. Develop a focused research question
2. Define inclusion/exclusion criteria
3. Select the outcomes for your review
4. Find the studies
5. Abstract the data
6. Assess quality of the data
7. Explore data (heterogeneity)
8. Synthesize the data descriptively and inferentially via meta-analysis if appropriate
9. Summarize the findings

Systematic Review Standards

A *standard* is a process, action, or procedure for performing SRs that is deemed essential to producing scientifically valid, transparent, and reproducible results








Annals of Internal Medicine | ACADEMIA AND CLINIC

The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration

Alessandro Liberati, MD, DrPH; Douglas G. Altman, DSc; Jennifer Tetzlaff, BSc; Cynthia Mulrow, MD, MSc; Peter C. Gøtzsche, MD, DrMedSci, MSc; John P.A. Ioannidis, MD; Mike Clarke, BA, DPhil; P.J. Devereaux, MD, BSc, PhD; Jos Kleijnen, MD, PhD; and David Moher, PhD


BMC Medical Research Methodology

BioMed Central

Research article

Open Access

Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews

Beverley J Shea*[1,5], Jeremy M Grimshaw[†2], George A Wells[3], Maarten Boers[†4], Neil Andersson[5], Candyce Hamel[†5], Ashley C Porter[5], Peter Tugwell[2], David Moher[6] and Lex M Bouter[†1]


Patient-Centered Outcomes Research Institute

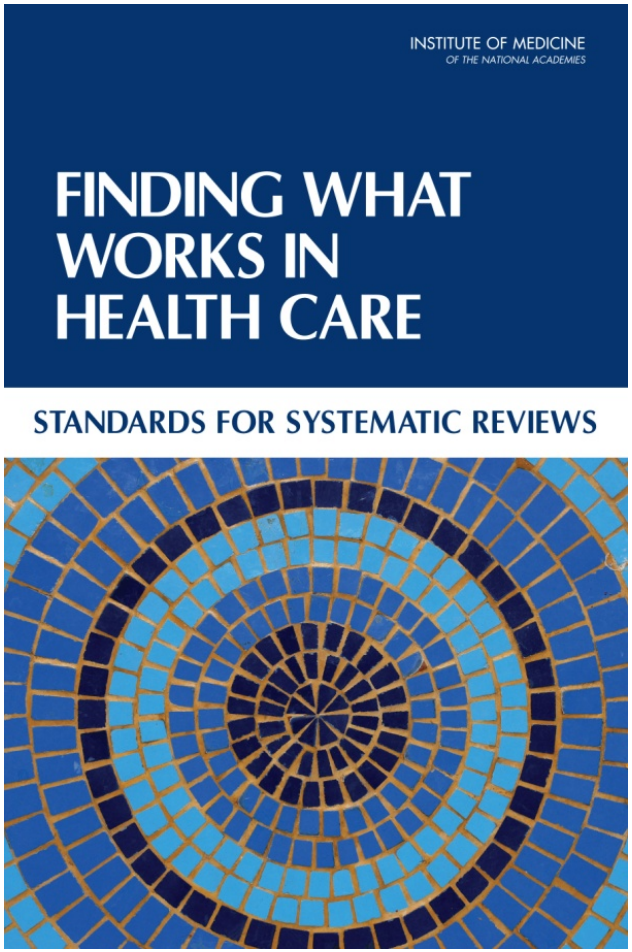
Draft Methodology Report:
"Our Questions, Our Decisions: Standards for Patient-centered Outcomes Research"

PCORI Methodology Committee

pcori

Mark Helfand, Alfred Berg, David Flum, Sherine Gabriel, and Sharon-Lise Normand, *Editors*

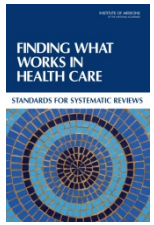
IOM Report on Standards for SRs (2011)



Committee Charge: Recommend methodological standards for SRs of comparative effectiveness research (CER) on health and health care

- **Assess potential methodological standards that would assure objective, transparent, and scientifically valid SRs of CER**
- **Recommend a set of methodological standards for developing and reporting such SRs**

Committee Methodology



- **Available research evidence**
- **Expert guidance from:**
  - **Agency for Healthcare Research and Quality (AHRQ) Effective Health Care Program**
  - **Centre for Reviews and Dissemination (CRD)**
  - **Cochrane Collaboration**
  - **Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group**
  - **Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA)**
- **Committee's assessment criteria:**
  - **Acceptability (credibility)**
  - **Applicability (generalizability)**
  - **Efficiency**
  - **Patient-centeredness**
  - **Scientific rigor**
  - **Timeliness**
  - **Transparency**

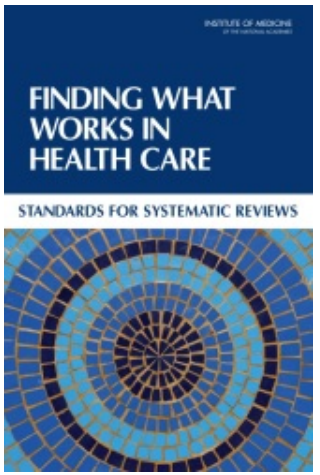
21 standards and 82 elements recommended

IOM standards are categorized into four subgroups:

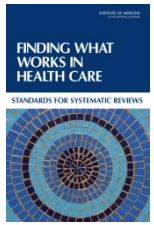
- **Initiating an SR**
- **Finding and assessing individual studies**
- **Synthesizing the body of evidence**
- **Reporting SRs**

STANDARD 3.3
Screen and select studies

3.3.1	Include or exclude studies based on the protocol's prespecified criteria
3.3.2	Use observational studies in addition to randomized clinical trials to evaluate harms of interventions
3.3.3	Use two or more members of the review team, working independently, to screen and select studies
3.3.4	Train screeners using written documentation; test and retest screeners to improve accuracy and consistency
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 text of articles identified in initial screening



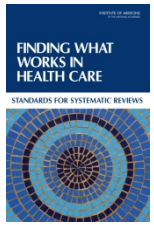
Buscemi et al., 2006 – "Single extraction was faster, but resulted in 21.7% more mistakes."



AHRQ – Ensure quality control mechanism; usually through use of independent researchers to assess studies for eligibility. Pilot testing is particularly important if there is not dual-review screening.

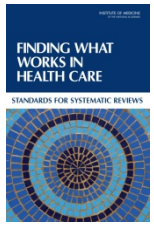
CRD – Good to have more than one researcher to help minimize bias and error at all stages of the review. Parallel independent assessments should be conducted to minimize the risk of errors.

Cochrane – At least two people, independently. Process must be transparent, and chosen to minimize biases and human error.



"Collectively the standards and elements present a daunting task. Few, if any, members of the committee have participated in an SR that fully meets all of them. Yet the evidence and experience are strong enough that it is impossible to ignore these standards or hope that one can safely cut corners. The standards will be especially valuable for SRs of high-stakes clinical questions with broad population impact, where the use of public funds to get the right answer justifies careful attention to the rigor with which the SR is conducted. Individuals involved in SRs should be thoughtful about all of the standards and elements, using their best judgment if resources are inadequate to implement all of them, or if some seem inappropriate for the particular task or question at hand. Transparency in reporting the methods actually used and the reasoning behind the choices are among the most important of the standards recommended by the committee."

IOM Standards Regarding Study Quality and Heterogeneity



STANDARD 3.6 **Critically appraise each study**	
3.6.1	Systematically assess the risk of bias, using predefined criteria
3.6.2	Assess the relevance of the study's populations, interventions, and outcome measures
3.6.3	Assess the fidelity of the implementation of interventions
STANDARD 4.2 **Conduct a qualitative synthesis**	
4.2.4	Describe the relationships between the characteristics of the individual studies and their reported findings and patterns across studies
STANDARD 4.4 **If conducting a meta-analysis, then do the following:**	
4.4.2	Address the heterogeneity among study effects

Why Assess the Quality of Individual Studies?

- **Combining poor quality studies may lead to biased, and therefore, misleading , pooled estimates**
- **Assessment of quality can be controversial and lead to its own form of bias**
- **Variety of methods exist including the Cochrane risk of bias tool**

- **Assessing quality of observational studies is very difficult**

Cochrane Risk of Bias Tool

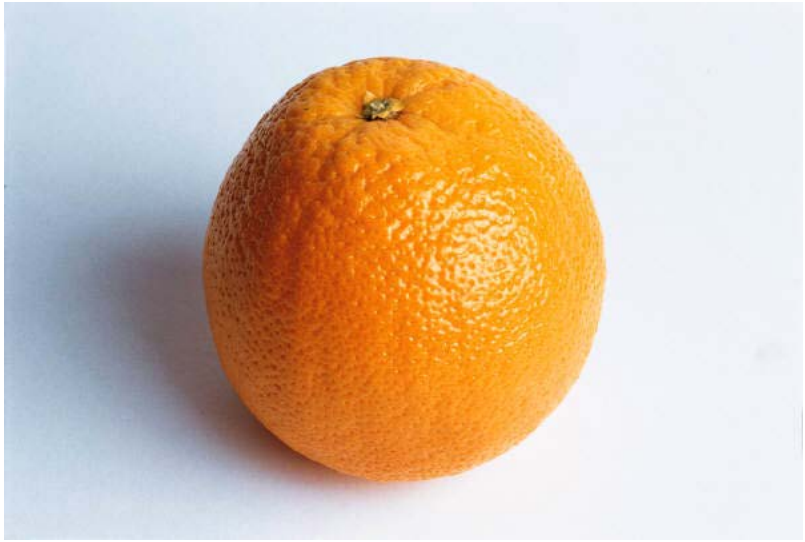
Domain	Description	Review authors' judgement
1. Sequence generation.	Describe the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups.	Was the allocation sequence adequately generated?
2. Allocation concealment.	Describe the method used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrolment.	Was allocation adequately concealed?
3. Blinding of participants, personnel and outcome assessors	Describe all measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.	Was knowledge of the allocated intervention adequately prevented during the study?

Domain	Description	Review authors' judgment
4. Incomplete outcome data	Describe the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. State whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomized participants), reasons for attrition/exclusions where reported, and any re-inclusions in analyses performed by the review authors.	Were incomplete outcome data adequately addressed?
5. Selective outcome reporting	State how the possibility of selective outcome reporting was examined by the review authors, and what was found.	Are reports of the study free of suggestion of selective outcome reporting?
6. Other sources of bias	State any important concerns about bias not addressed in the other domains in the tool. If particular questions/entries were pre-specified in the review's protocol, responses should be provided for each question/entry.	Was the study apparently free of other problems that could put it at a high risk of bias?

Figure 8.6.a: Example of a 'Risk of bias' table for a single study (fictional)

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	Low risk.	Quote: "patients were randomly allocated." Comment: Probably done, since earlier reports from the same investigators clearly describe use of random sequences (Cartwright 1980).
Allocation concealment (selection bias)	High risk.	Quote: "...using a table of random numbers." Comment: Probably not done.
Blinding of participants and personnel (performance bias)	Low risk.	Quote: "double blind, double dummy"; "High and low dose tablets or capsules were indistinguishable in all aspects of their outward appearance. For each drug an identically matched placebo was available (the success of blinding was evaluated by examining the drugs before distribution)." Comment: Probably done.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	Low risk.	Quote: "double blind". Comment: Probably done.
Blinding of outcome assessment (detection bias) (Mortality)	Low risk.	Obtained from medical records; review authors do not believe this will introduce bias.
Incomplete outcome data addressed (attrition bias) (Short-term outcomes (2-6 weeks))	High risk.	4 weeks: 17/110 missing from intervention group (9 due to 'lack of efficacy'); 7/113 missing from control group (2 due to 'lack of efficacy').
Incomplete outcome data addressed (attrition bias) (Longer-term outcomes (>6 weeks))	High risk.	12 weeks: 31/110 missing from intervention group; 18/113 missing from control group. Reasons differ across groups.
Selective reporting (reporting bias)	High risk.	Three rating scales for cognition listed in Methods, but only one (with statistically significant results) is reported.

“Heterogeneity is Your Friend” (J. Berlin)



Fruit salad may, or may not, be tasty and interesting

Which are the apples and oranges, and how do they differ?

Definitions of Heterogeneity From a Health Care Perspective

Different types of heterogeneity:

- **Clinical heterogeneity (diversity): Variability in participants, interventions and outcomes**
- **Methodological heterogeneity (diversity): Variability in study design and risk of bias**
- **Statistical heterogeneity: Variability in treatment effects, resulting from clinical and/or methodological diversity**

- **Statistical heterogeneity is present if the observed treatment effects are more different from each other than would be expected due to chance alone**

Discuss Clinical/Methodological or "Substantive" Heterogeneity *Prior* To Analysis

- **Think first:** Are included studies similar with respect to treatment effect? Study design, subjects, treatments, etc. may affect results.
- **Include in protocol:** Sources of heterogeneity that you might stratify analysis on, or that you might include as independent variables in a meta-regression
- **Do statistics later:** Q statistic to test the hypothesis that the true (population) treatment effect is equal in all studies; and/or I-squared (I^2) statistic
- **Remember:** Tests for heterogeneity have low statistical power

Evaluating The Strength of The Body of Evidence

Annals of Internal Medicine

AHRQ Agency for Healthcare Research and Quality — Advancing Excellence in Health Care • www.ahrq.gov | USPSTF

SCREENING FOR BREAST CANCER USING FILM MAMMOGRAPHY
CLINICAL SUMMARY OF U.S. PREVENTIVE SERVICES TASK FORCE RECOMMENDATION

Population	Women Aged 40–49 Years	Women Aged 50–74 Years	Women Aged ≥75 Years
Recommendation	Do not screen routinely. Individualize decision to begin biennial screening according to the patient's context and values.	Screen every 2 years.	No recommendation.
	Grade: C	Grade: B	Grade: I (insufficient evidence)

Table 1. What the USPSTF Grades Mean and Suggestions for Practice

Grade	Definition	Suggestions for Practice
A	The USPSTF recommends the service. There is high certainty that the net benefit is substantial.	Offer/provide this service.
B	The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial.	Offer/provide this service.
C	The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is moderate or high certainty that the net benefit is small.	Offer/provide this service only if other considerations support offering or providing the service in an individual patient.

IOM Standard Regarding Strength of Evidence

STANDARD 4.1
Use a prespecified method to evaluate the body of evidence

4.1.1 For each outcome, systematically assess the following characteristics of the body of evidence:

- Risk of bias
- Consistency
- Precision
- Directness
- Reporting bias

4.1.2 For bodies of evidence that include observational research, also systematically assess the following characteristics for each outcome:

- Dose-response association
- Plausible confounding that would change the observed effect
- Strength of association

4.1.3 For each outcome specified in the protocol, use consistent language to characterize the level of confidence in the estimates of the effect of an intervention

- **Use consistent language to summarize the conclusions of individual studies as well as the body of evidence:**
  - **Presenting results not sufficient**
  - **Reviews often very long**

Evidence on assessment methods is elusive

- **Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach is becoming more popular**
- **Anecdotal evidence that**
  - **GRADE is difficult to apply**
  - **GRADE is being modified for specific situations**
- **GRADE starts by downgrading observational studies**

AHRQ Agency for Healthcare Research and Quality

Advancing Excellence in Health Care

Reliability Testing of the AHRQ EPC Approach to Grading the Strength of Evidence in Comparative Effectiveness Reviews (Berkman et al., RTI/UNC EPC)

ELSEVIER

Journal of Clinical Epidemiology 66 (2013) 1105e1117

Journal of Clinical Epidemiology

ORIGINAL ARTICLES

Interrater reliability of grading strength of evidence varies with the complexity of the evidence in systematic reviews

Nancy D. Berkman[a,*], Kathleen N. Lohr[a], Laura C. Morgan[a], Tzy-Mey Kuo[b], Sally C. Morton[c] [a]Division of Social Policy, Health, and Economics Research, RTI International (Research Triangle Institute), Research Triangle Park, NC 27709-2194, USA [b]Lineburger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, 101 E. Weaver Street, Carrboro, NC, 27599, USA
[c]Department of Biostatistics, University of Pittsburgh, 130 DeSoto Street, Pittsburgh, PA, 15261, USA

Accepted 5 June 2013

Study Design

- Inter-rater reliability of
  - 4 required domains (risk of bias; consistency; directness; and precision) for RCTs and observational studies separately
- 10 exercises from 2 published CER SRs on depression and rheumatoid arthritis
  - All exercises contained RCTs
  - 6 exercises included one or more observational studies
- Eleven pairs of reviewers (10 from 9 EPCs, 1 from AHRQ) participated in each exercise

Study/Design	N	Comparison	Quality	Results: Biologics vs. Oral DMARDs
TEMPO, 2005 RCT	451	ETN 25mg twice/wk vs. MTX	Fair	Remission at week 24: DAS < 1.6: 13.0% vs. 13.6% (P = NS) DAS28 < 2.6: 13.9% vs. 13.6% (P = NS) Remission at week 52: DAS <1.6: 17.5% vs. 14%, (P = NS) DAS28 < 2.6: 17.5% vs. 17.1%, (P = NS)
PREMIER, 2006 RCT	531	ADA 40 mg biweekly vs. MTX 20 mg/wk	Fair	Clinical remission (DAS28 < 2.6) at 1 year: 23% vs. 21%, (P = 0.582†)
Listing 2006 Prospective cohort study	1083	Biologics vs. conventional DMARDs	Fair	Odds of achieving remission (DAS28 < 2.6) at 12 months: Adjusted* OR, 1.95 (95% CI, 1.20-3.19); (P = 0.006) *Adjusted for age, sex, # of previous DMARDs, DAS28, ESR, FFbH, osteoporosis, previous txt with cyclosporine A. Matched pairs analysis DAS28 remission at 12 months: 24.9% vs. 12.4%, (P = 0.004)

Study Results

Domain or Strength of Evidence (SOE)	Independent Reviewer Agreement	Reconciled Reviewer Pair Agreement
Bias: RCTs	0.67 (0.61,0.73)	0.65 (0.56,0.73)
Bias: Observational studies	0.11 (0.05,0.18)	0.22 (0.13,0.32)
SOE: All studies	0.20 (0.16,0.25)	0.24 (0.14,0.34)
SOE: RCTs only	0.22 (0.17,0.28)	0.30 (0.17,0.43)

Study Conclusions

- **Inter-rater reliability is low**
- **Complex evidence bases, particularly those with a mix of randomized and observational studies, can be extremely difficult to grade**
- **Dual review with adjudication of differences improves reliability**
- **Additional methodological guidance needed for reviewers**
- **More research is needed on reliability, and the advantages and disadvantages of concrete rules for determining strength of evidence**

Final words

“To do a meta-analysis is easy,

to do one well is hard.”

- Ingram Olkin

Thank You

Sally C. Morton
Department of Biostatistics
Graduate School of Public Health
University of Pittsburgh

scmorton@pitt.edu

Are RCTs Enough?

Austin Bradford Hill Heberden Oration (1965):

"this leads directly to a related criticism of the present controlled trial – that it does not tell the doctor what he wants to know. It may be so constituted as to show without any doubt that treatment A is *on the average* better than treatment B. On the other hand, that result does not answer the practicing doctor's question – what is the most likely outcome when this drug is given to a particular patient?"

PRIMSA Flowchart and Checklist

Endorsed (required) by many top clinical journals

Figure 1. Flow of information through the different phases of a systematic review.

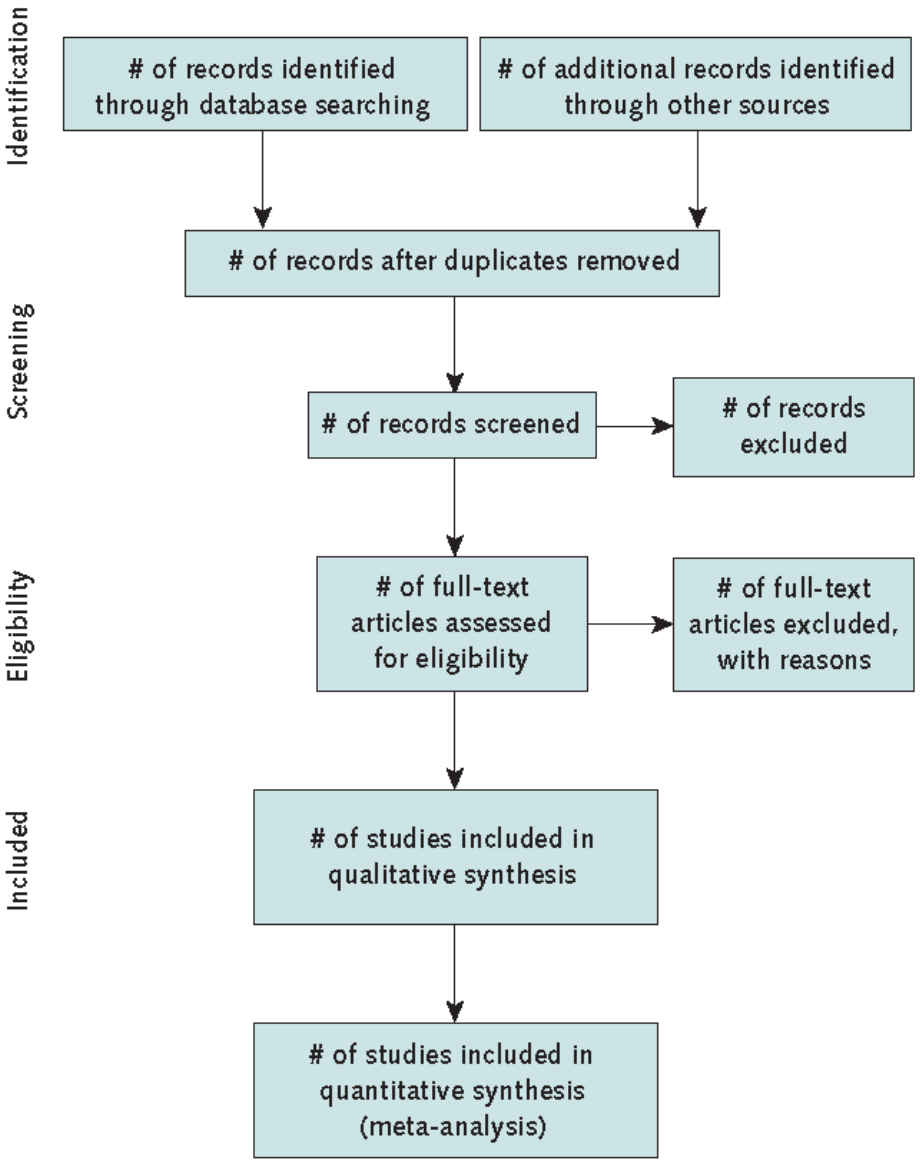


Table 1. **Checklist of Items to Include When Reporting a Systematic Review or Meta-Analysis**

Section/Topic	Item #	Checklist Item	Reported on Page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	

RESULTS		
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome-level assessment (see Item 12).
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group and (b) effect estimates and confidence intervals, ideally with a forest plot.
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).
DISCUSSION		
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., health care providers, users, and policy makers).
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review level (e.g., incomplete retrieval of identified research, reporting bias).
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.
FUNDING		
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.

Q Statistic – The Math

Let θ_i = true treatment effect for study i for $i = 1, ..., K$

Individual study i effect is estimated by T_i

Pooled treatment effect θ_P is estimated by $T_P = \dfrac{\sum_{i=1}^{K} w_i T_i}{\sum_{i=1}^{K} w_i}$

$H_o : \theta_1 = \theta_2 = ... = \theta_K = \theta_P$

Test statistic is $Q = \sum_{i=1}^{K} w_i (T_i - T_P)^2$

and $Q \sim_{H_0} \chi^2$ with $K - 1$ degrees of freedom

I-squared is a newer heterogeneity statistic that measures the percentage of variation across studies that cannot be explained by chance

Assessment of Heterogeneity – Proceed with Meta-analysis, or Not?

First, check your data!

Options:

- **Do not do meta-analysis**
- **Change treatment effect measure**
- **Explore heterogeneity via stratification or meta-regression**
- **Account for heterogeneity via random effects model (not advisable if heterogeneity large)**
- **Exclude studies (can do "leave-one-out" or jackknife test to determine individual study effect on heterogeneity)**

"Just a darn minute! — Yesterday you said that X equals **two**!"

Example Strength of Evidence Table

Self management for patients with chronic obstructive pulmonary disease

Patient or population: patients with chronic obstructive pulmonary disease
Settings: primary care, community, outpatient
Intervention: self management[1]
Comparison: usual care

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk **usual care**	Corresponding risk **self management**				
Quality of Life St George's Respiratory Questionnaire. Scale from: 0 to 100. (follow-up: 3 to 12 months)	The mean quality of life ranged across control groups from **38 to 60 points**	The mean quality of Life in the intervention groups was **2.58 lower** (5.14 to 0.02 lower)		698 (7)	⊕⊕⊕◯ **moderate**[2]	Lower score indicates better quality of life. A change of < 4 is not shown to be im-portant to patients.
Dyspnoea Borg Scale. Scale from: 0 to 10. (follow-up: 3 to 6 months)	The mean dyspnoea ranged across control groups from **1.2 to 4.1 points**	The mean dyspnoea in the intervention groups was **0.53 lower** (0.96 to 0.1 lower)		144 (2)	⊕⊕◯◯ **low**[3,4]	Lower score indicates improvement
Number and severity of exacerbations[5]	See comment	See comment	Not estimable[5]	591 (3)	See comment	Effect is uncertain
Respiratory-related hospital admissions (follow-up: 3 to 12 months)	**Low risk population**[6]		**OR 0.64** (0.47 to 0.89)	966 (8)	⊕⊕⊕◯ **moderate**[7]	
	10 per 100	**7 per 100** (5 to 9)				
	High risk population[6]					
	50 per 100	**39 per 100** (32 to 47)				

Example SOE Table, continued

Emergency department visits for lung diseases (follow-up: 6 to 12 months)	The mean visits ranged across control groups from **0.2 to 0.7 visits/person/year**	The mean visits in the intervention groups was **0.1 higher** (0.2 lower to 0.3 higher)		328 (4)	⊕⊕⊕◯ **moderate**[4]
Doctor and nurse visits (follow-up: 6 to 12 months)	The mean visits ranged across control groups from **1 to 5 visits/person /year**	The mean visits in the intervention groups was **0.02 higher** (1 lower to 1 higher)		629 (8)	⊕⊕⊕◯ **moderate**[8]

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; OR: Odds ratio;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Self-management is a term applied to any formalized patient education programme aimed at teaching skills needed to carry out medical regimens specific to the disease, guide health behaviour change, and provide emotional support for patients to control their disease and live functional lives. Of the 14 studies, there were four in which the education delivery mode consisted of group education; nine which were individual education and one which was written education material only. In six studies the use of an action plan for self-treatment of exacerbations was assessed.
[2] Seven other studies were not pooled and some showed non-significant effects.
[3] No allocation concealment in 1 study. Incomplete follow-up.
[4] Sparse data.
[5] Different definitions of exacerbations used and studies could not be pooled.
[6] The low and high risk values are two extreme numbers of admissions in the control groups from 2 studies (8 was rounded to 10% and 51 to 50%).
[7] 2 studies with very severe COPD patients weighed heavily in analysis. Therefore, there is some uncertainty with applicability of effect to all risk groups.
[8] Unexplained heterogeneity.