Systematic Reviews (SRs)

Focusing on Therapy

Manit Srisurapanont, MD

Professor of Psychiatry

Department of Psychiatry

Faculty of Medicine, Chiang Mai University

(Email: msrisu@cmu.ac.th)

For Research Skill 2 (MSY3), November 3, 2020, Chiang Mai.

CHIANG MAI UNIVERSITY
FACULTY OF MEDICINE

Outline

- SRs: their roles in clinical practice
- Conducting SRs: steps and tasks
- Advance and practical issues of SRs

ข.ความรู้ความสามารถทางวิชาชีพและทักษะทางคลินิก

หมวดที่1. ภาวะปกติและหลักการดูแลทั่วไป

1.2 สามารถรวบรวมข้อมูล และประเมินปัญหาสุขภาพของบุคคล ครอบครัว และ ชุมชนในความรับผิดชอบได้ โดยใช้วิธีทางวิทยาการระบาดพื้นฐานดังนี้

1.2.5 Systematic reviews

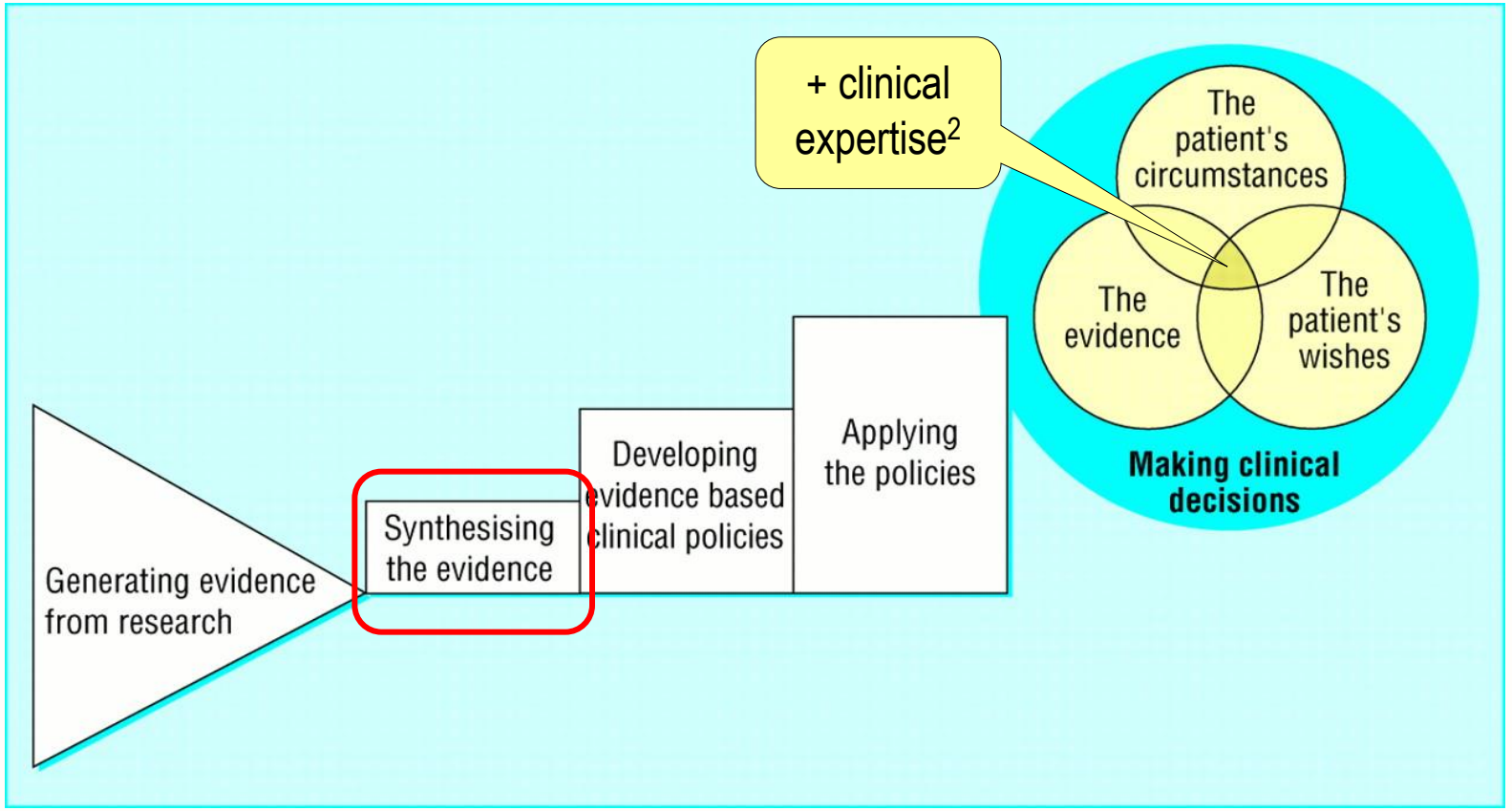
- 1.2.5.1 identifying and selecting studies
- 1.2.5.2 quality of evidence assessments
- 1.2.5.3 combining the findings of independent studies
- 1.2.5.4 variation between study findings
- 1.2.5.5 summarizing and interpreting results

1.2.9 Evidence-based medicine

- 1.2.9.1 asking focused questions: translation of uncertainty to an answerable question
- 1.2.9.2 finding the evidence: systematic retrieval of best evidence available
- 1.2.9.3 critical appraisal: testing evidence for validity, clinical relevance, and applicability
- 1.2.9.4 making a decision: application of results in practice
- 1.2.9.5 evaluating performance: auditing evidence-based decisions

Evidence-based clinical practice

- The process of *systematically reviewing, appraising, and using* clinical research findings to aid the delivery of optimum clinical care to patients[1]



1. Haynes RB, Haines A. *BMJ.* 1998;317:273-276.; 2. Haynes RB, et al. *EBM.* 2002;7:36-38.

Hierarchy of strength of evidence[1] (for prevention and treatment decisions)

- (N-of-1 randomized trial)
- Systematic review (SR) of randomized trials
- Single randomized trials
- Systematic review of observational studies
- Single observational study
- Physiological study
- Unsystematic clinical observation

1. Guyatt G, et al. *Manual for Evidence-Based Clinical Practice*. 2008.

SR and meta-analysis[1]

- SRs - key characteristics include:
  - A clearly stated set of objectives with an explicit, reproducible methodology
  - A systematic search that attempts to identify all studies that would meet the eligibility criteria
  - An assessment of the validity of the findings of the included studies (e.g., assessment of risk of bias and confidence in cumulative estimates); and
  - Systematic presentation, and synthesis of the characteristics and findings of the included studies
- Meta-analysis: the use of statistical techniques to combine and summarize the results of multiple studies

Moher et al. *Syst Rev* 2015,4:1.

SR: a must for practice guideline development

- A key step for guideline development
  - 1991: Are systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances

    (Note: consensus statements was acceptable)
  - 2011: Are statements that included recommendations intended to optimize patient care that are informed by a SR of evidence and an assessment of the benefits and harms of alternative care options

    (Note: consensus statements is not acceptable)

IOM (Institute of Medicine). *Clinical Practice Guidelines We Can Trust*, 2011.

Conducting SRs: steps and tasks

Step 1: Protocol development and registration

- 1.1 Defining research question
- 1.2 Defining selection criteria
- 1.3 Designing the search strategies
- 1.4 Protocol development & registration

Step 2: Trial search, selection, assessment, and data extraction

- 2.1 Study searches & selection
- 2.2 Risk-of-bias assessment
- 2.3 Data extraction

Step 3: Data synthesis and report

- 3.1 Trial and trial-result description
- 3.2 Data synthesis (±)
- 3.3 Heterogeneity (±)
- 3.4 Publication bias (±)
- 3.5 Interpretation & report

NL Criteria 2012 vs. Tasks

NL Criteria 2010	Tasks
1.2.9.1 asking focused questions	1.1 Defining research question
1.2.5.1 identifying and selecting studies	1.2 Defining selection criteria
1.2.9.2 finding the evidence	1.3 Designing the search strategies
1.2.5.2 quality of evidence assessments	2.2 Risk-of-bias assessment
1.2.5.3 combining the findings of independent studies	3.2 Data synthesis
1.2.5.4 variation between study findings	3.3 Heterogeneity & publication bias
1.2.5.5 summarizing and interpreting results	3.2 Data synthesis 3.3 Heterogeneity & publication bias
1.2.9.4 making a decision: application of results in practice	(Advance and practical issues – GRADE approach)

Task 1.1: Defining a review question (objective)

- Example:
  - *"Can exercise or physical activity help improve postnatal depression and weight loss?"*

- Parts of a well-built clinical question (objective) (PICO)[1]:
  - Patients or problem: *postnatal depression and weight loss*
  - Intervention: *exercise or physical activity*
  - Comparison intervention (optional): *no*
  - Outcomes: *improve(ment) of postnatal depression and weight loss*

Saligheh M, et al. *Arch Womens Ment Health.* 2017;20:595-611

1.2 Defining selection criteria

- More details of the review question (objective)
  - Type of patients or problems, esp, (more specific) health problem/condition, population (eg, age, sex), comorbidity, setting (eg, hospitalization status)
  - Type of interventions, esp., drugs (w/ details), concomitant treatment
    - including comparators
  - Type of outcomes, esp., those interesting for cares, consumers, and policy makers
  - Type of studies, esp., RCTs only?, quality of studies

1.3 Designing the search strategies

- Defining the electronic databases to be searched, eg,
  - Bibliographic databases: eg, Pubmed (Medline)*, Scopus/EMBASE, Web of Science
    - Describing the search strategy of each database
  - Research registry databases, eg, Clinicaltrials.gov
- Reference lists from relevant primary and review articles
- Hand searching, grey literature, and conference proceedings
- Requesting trial and trial data from researchers and manufacturers

Pubmed (Medline) is a free bibliographic database in biomedical sciences.

1.4 Protocol development and registration

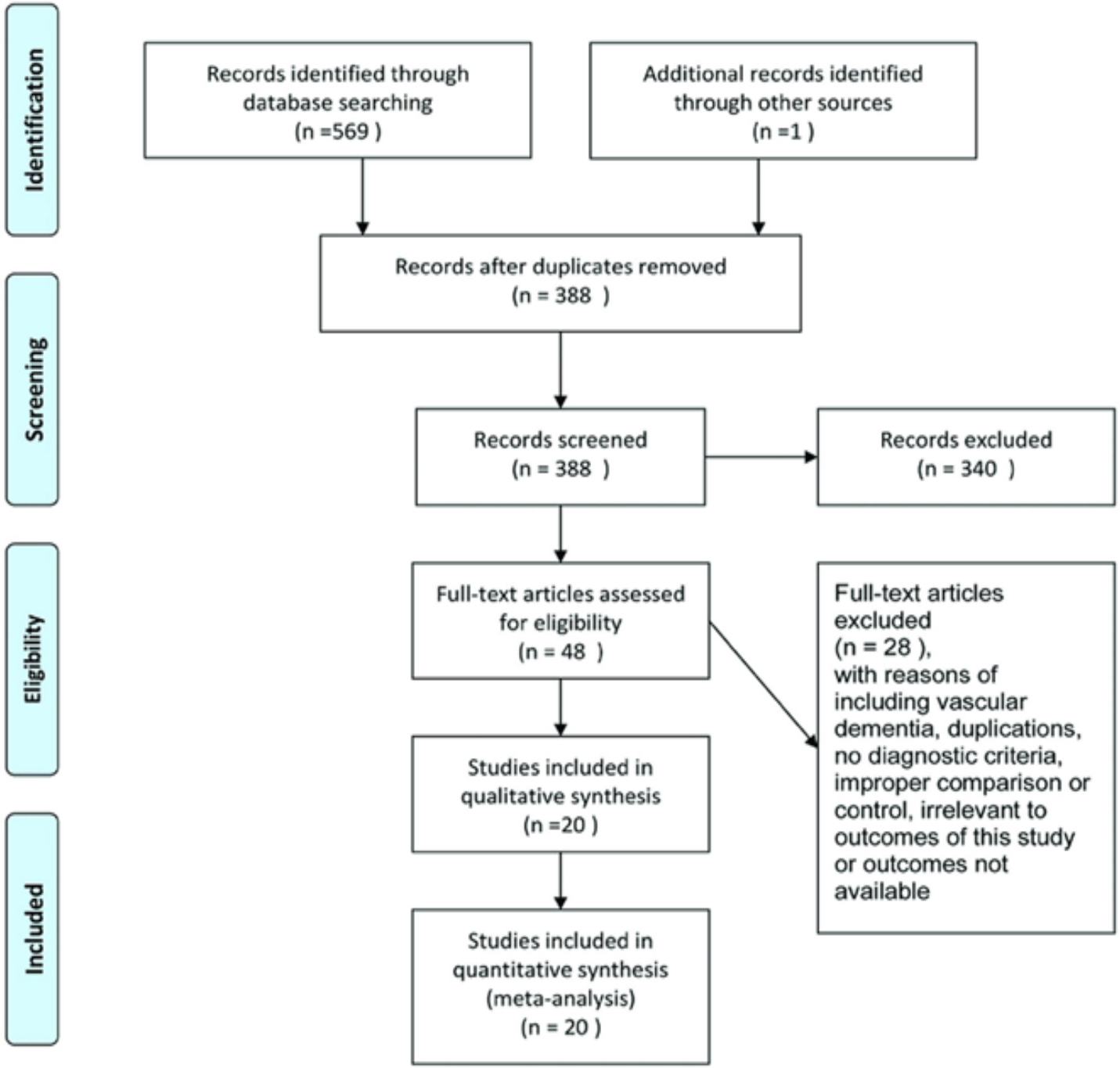
- Key items included
  - Review question (objective)
  - Selection criteria
  - Search strategy
  - Study selection
  - Outcomes and priority
  - Risk of bias of individual studies
  - Data items and extraction
  - Data synthesis
  - Heterogeneity
  - Publication bias

- Why and How?
  - Help reduce the bias possibly occurring during the conduct of SR
  - Describe the rationale, hypothesis, and planned methods of the review, eg. PRISMA-P[1]
  - Made publicly available, and registered in a registry such as PROSPERO, Open Science Framework (OSF; osf.io)

1. Shamseer L, et al. *BMJ.* 2015 ;349:g7647.

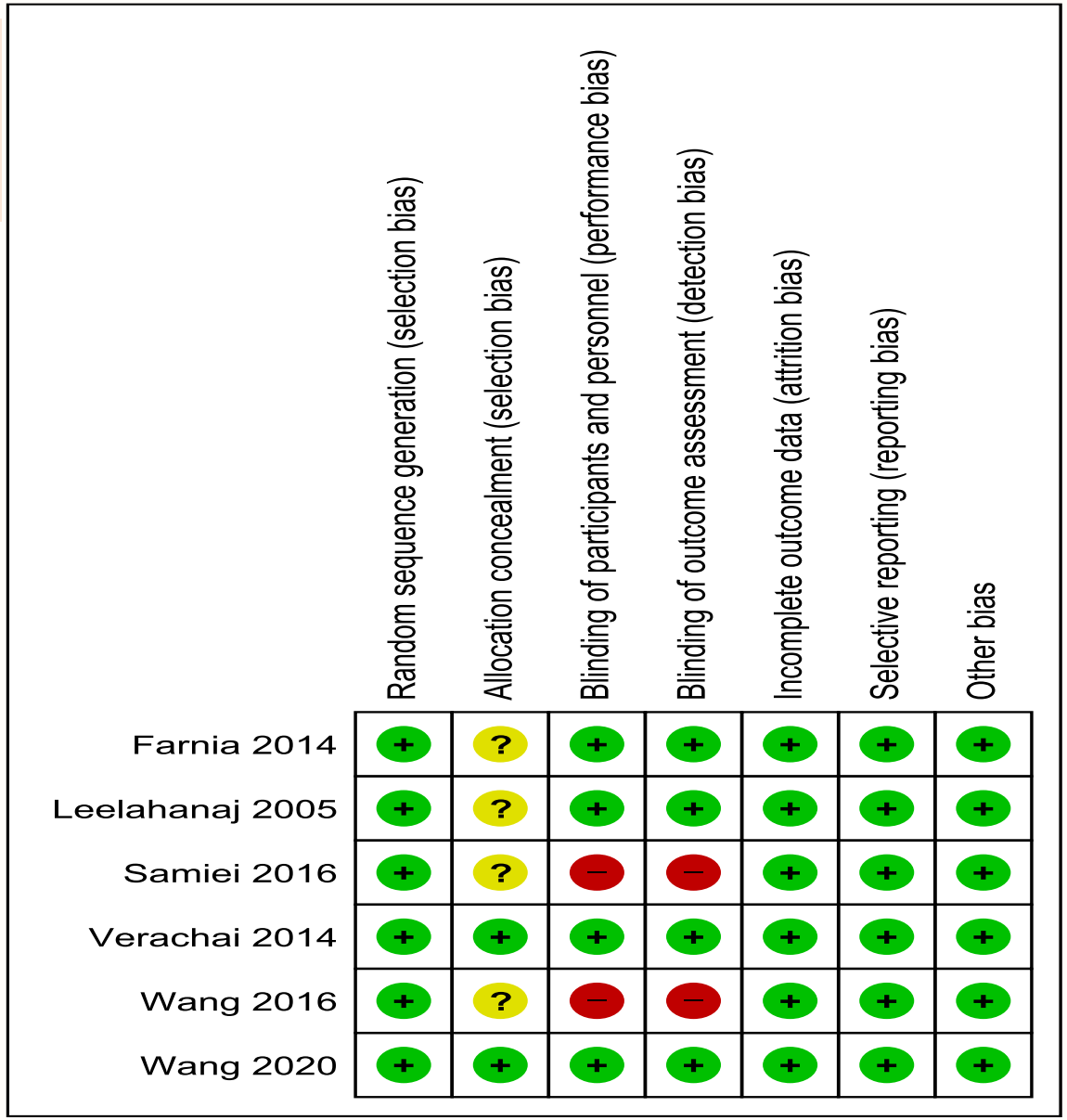
2.1 Study searches and selection

- Search multiple databases and select studies as described in the protocol
- Parallel independent selection (≥2 reviewers)
- Provide reasons for the exclusion of studies being assessed by reading their full-text articles
- Record items and studies included/excluded in all steps, as well as report using PRISMA Flow



2.2 Risk of bias assessment

- Help
  - understand the strength of evidence
  - exclude low-quality studies
- Examples
  - Cochrane criteria for assessing risk of bias in included studies (RoB1) (7 domains)[1]
  - Cochrane RoB2: Version 2 of the risk-of-bias tool for RCTs[2]
- Parallel independent assessment (≥2 reviewers)



1. Higgins J, et al. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*, 2011.
2. Sterne JAC, et al. *BMJ* 2019;366:l4898.

2.2 Risk of bias assessment (for randomized trials) (RoB1)[1]

Bias	Rating: High, Low, or Unclear Risk	Eg, Low Risk
Selection bias: Random sequence generation	The method used to generate the allocation sequence: i) described in sufficient detail and ii) should produce comparable groups	Random codes were generated by computer
Selection bias: Allocation concealment	The method used to conceal the allocation sequence: i) described in sufficient detail and ii) could not been foreseen before or during enrollment	The randomization assignment was concealed in an envelope until the start of treatment
Performance bias: Blinding of participants & personnel	Methods used to blind study participants and personnel: i) described, especially, for the study that blindness is needed (eg, placebo use in the study)	Placebo used (participants and personnel were blinded to interventions)

1. Higgins J, Green S. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*, 2011

2.2 Risk of bias assessment (for randomized trials) (RoB1) (cont.)

Bias	Rating: High, Low, or Unclear Risk	Eg, Low Risk
Detection bias: Blinding of outcome assessment	Methods used to blind outcome assessors: i) described, especially, for the study that blindness is needed	Assessors were blind to intervention assignments
Attrition bias: Incomplete outcome data	Completeness of outcome data for each main outcome, including attrition and exclusions from the analysis.	Defining all participants included/excluded from each step of the study
Reporting bias: Selective reporting	No missing of an important outcome that had been defined	All defined outcomes were reported
Other bias	No other concerns of bias*	

2.3 Data extraction

- Data extraction is prone to human error, and can be minimized by
  - Designing and using a data extraction form and
  - Parallel independent extraction (≥2 reviewers)
- Missing data: contact the original authors

3.1 Trial and trial-result description

- Key elements of descriptive data synthesis (in text and tables)
  - PICO as described in the criteria for study selection
  - Factors possibly affecting outcomes, esp,
    - Patients' characteristics, eg, mean age, sex, illness severity/staging
    - Settings where the technology was applied
- Perform a narrative, descriptive (qualitative) summary with/without graphs and tables

3.2 Data synthesis

- AVOID meta-analysis if the data are too sparse, too low quality, or too heterogeneous to proceed with their statistical aggregation

3.2 Data synthesis – choices of outcomes and effect models

- Dichotomous data, eg,
  - Odds ratios (ORs)
  - Relative risks (RRs)
  - Risk differences (RDs)
- Continuous data, eg,
  - (Weighted) mean difference (WMD or MD): the same scale used for measuring an outcome (eg, weight)
  - Standardized mean difference (SMD): different scales used for measuring an outcome (eg, pain)

- Effect model[1]
  - Fixed: all studies share a common effect size
  - Random: there is *a distribution* of true effect sizes, and our goal is to estimate the mean of this distribution

[1] Borenstein M, et al. *Res Synth Method*. 2010:97-111.

3.2 Data synthesis – interpretation of SMDs

- SMDs can be computed by many methods
- Cohen's *d*:
  - *d* = (mt-mc)/(pooled sd)
  - mt = mean of treatment group
  - mc = mean of controlled group
- Hedges' *g*:
  - $g = d\left(1 - \frac{3}{4N - 9}\right)$
  - *N* = total sample size
  - a common method for computing SMDs
  - For large sample size, $g \cong d$

- Interpretation of Cohen's *d*[1]
  - 0.2 ≅ small effect size (ES)
  - 0.5 ≅ medium ES
  - 0.8 ≅ large ES
- Medium ES represent an effect likely to be visible to the naked eye of a careful observer
- Small ES to be noticeably smaller than medium but not so small as to be trivial
- Large ES to be the same distance above medium as small was below it.

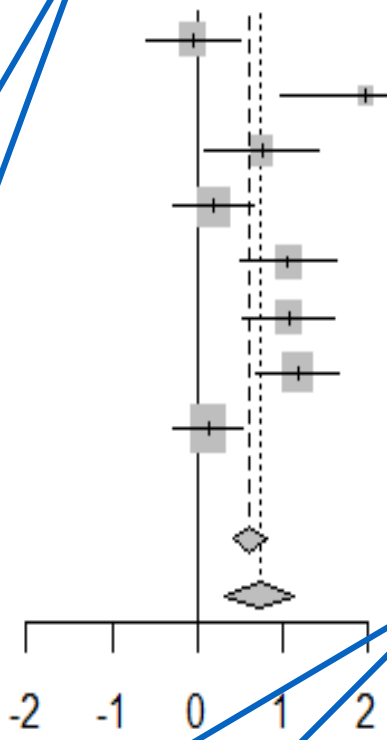
1. Cohen J. *Psychol Bull.* 1988;112(1):155-159.

3.2 Data synthesis – Forest plot

Study	Experimental Total	Experimental Mean	Experimental SD	Control Total	Control Mean	Control SD	SMD	95%-CI	Weight (fixed)	Weight (random)
Ikai (Yoga) 2014	25	0.04	0.5200	25	0.07	0.5800	-0.05	[-0.61; 0.50]	11.6%	12.8%
Lee (Yoga) 2014	14	7.42	0.7500	11	-6.53	10.3400	1.97	[0.98; 2.96]	3.6%	8.5%
Naveen (Yoga) 2016	16	1.95	5.8000	23	-2.40	5.6000	0.75	[0.09; 1.41]	8.1%	11.6%
Sungkarat (Tai Chi) 2018	33	0.17	0.1400	37	0.14	0.1700	0.19	[-0.28; 0.66]	16.1%	13.6%
Tolahunase (Yoga) 2018	29	5.57	6.0400	29	0.02	4.1100	1.06	[0.51; 1.61]	11.7%	12.8%
Gagrani (Meditation) 2019	30	11.10	11.6700	30	1.19	5.7400	1.06	[0.52; 1.61]	12.1%	12.9%
Ledreux (Meditation) 2019	39	-0.01	0.8100	39	-0.96	0.7900	1.18	[0.69; 1.66]	15.3%	13.5%
Nerry (Meditation) 2019	62	0.73	1.8900	37	-0.35	13.8300	0.13	[-0.28; 0.53]	21.5%	14.3%
Fixed effect model	**248**			**231**			**0.62**	**[0.43; 0.81]**	**100.0%**	--
Random effects model							**0.72**	**[0.31; 1.14]**	--	**100.0%**

Heterogeneity: $I^2 = 78\%$, $\tau^2 = 0.2709$, $\chi^2_7 = 31.97$ ($p < 0.01$)

Standardised Mean Difference



Heterogeneity

Standardized Mean Difference for continuous data

Overall results with 95% CI

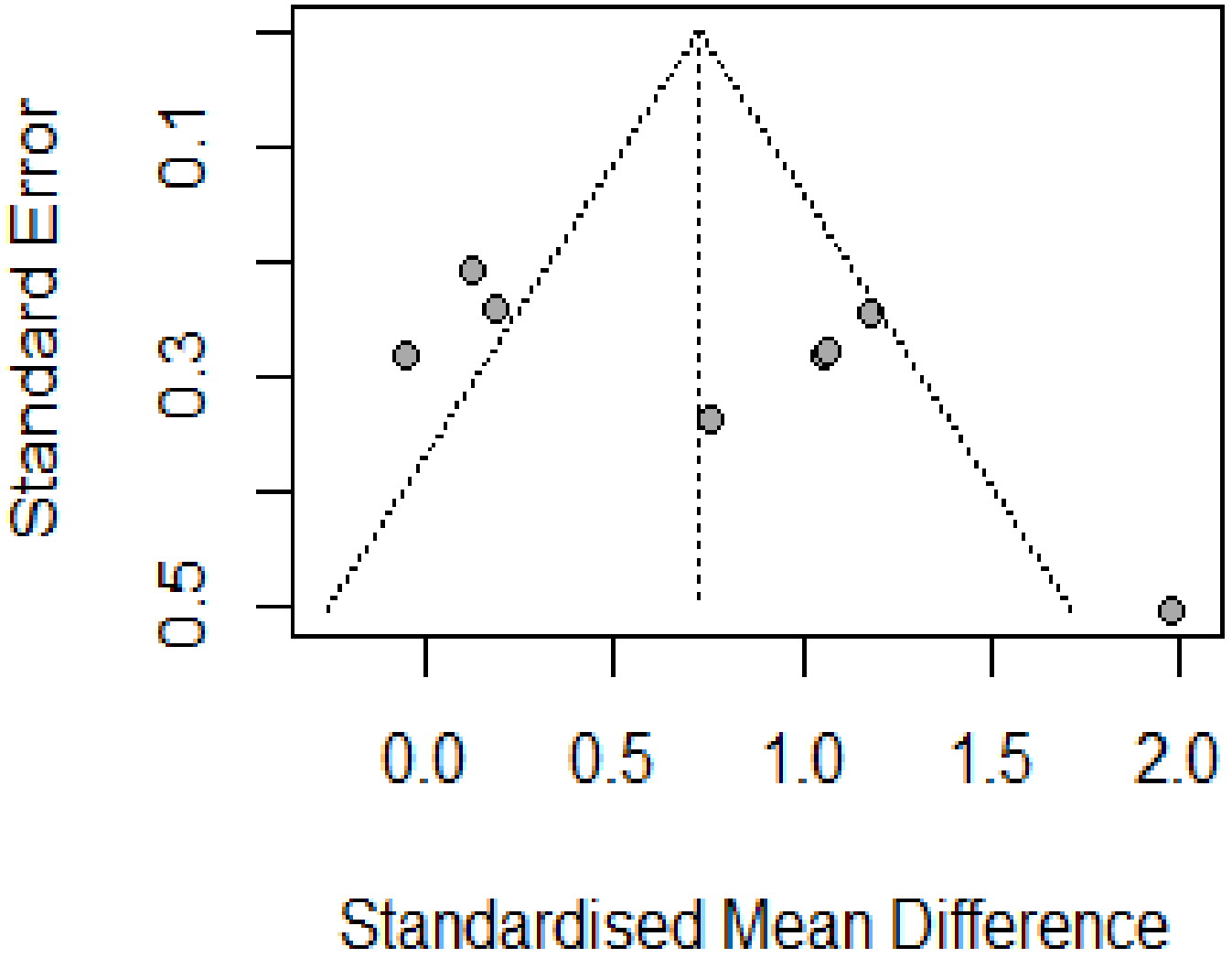
Data combined using a random- or fixed-effect model

3.3 Heterogeneity

- Heterogeneity assessment
  - Visualized using forest plots
  - Common statistical tests
    - Cochrane's (or Chi-square test of) Q test:
      - p (of $\chi 2$) < 0.1 = high heterogeneity
    - Higgin's & Thompson's I^2:
      - 25% = low, 50% = moderate, 75% = high
    - Between-study variance τ^2 (tau-squared):
      - many methods for calculation and difficult for interpretation
- Two approaches for the synthesis of data with high heterogeneity
  - Synthesizing data using a random-effect model
  - No data synthesis

3.4 Publication bias

- Studies showing the ineffectiveness of interventions may not be published
- Failing to include unpublished studies → overestimate the true effect of an intervention
- Tests for funnel plot asymmetry should be used only when ≥10 studies are included in the meta-analysis[1]



Egger's test (k=8): t = 1.7856, df = 6, p-value = 0.1244

1. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0; 2011.

3.4 Interpretation & report

- Interpret results within the context of current health care
- State the methodological limitations of both the primary studies (risk of bias) and the SR
- Make clinical recommendations practical and explicit
- Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)[1]
  - Including 27 checklist items
  - Aims to help authors improve the reporting of SRs and meta-analyses

Note: Cochrane Review has its own style.

1. http://www.prisma-statement.org/

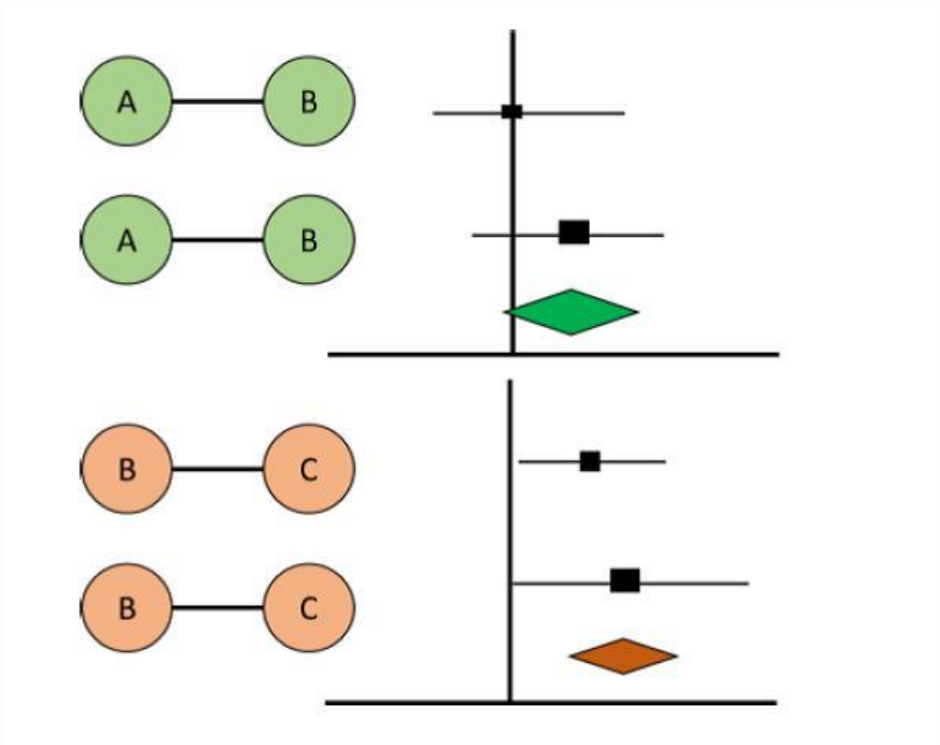
PRISMA 2009 Checklist

PRISMA: 8 of 27 items

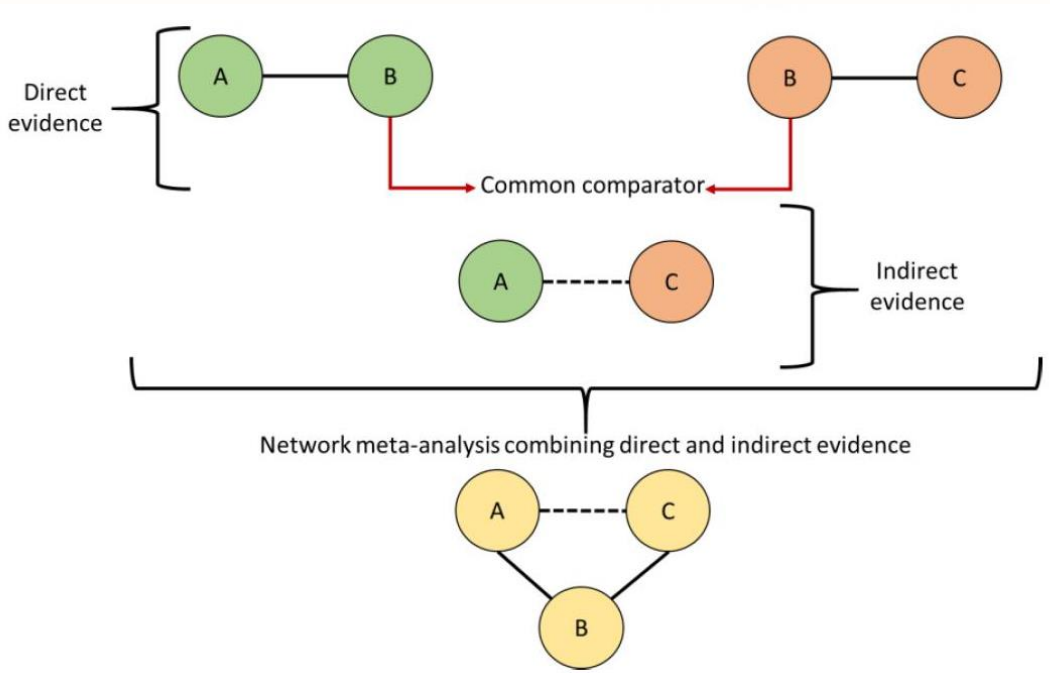
Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, 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	

Advance and practical issues of SRs – network meta-analysis (NMA)

Pairwise meta-analysis[1]



Network meta-analysis[1]

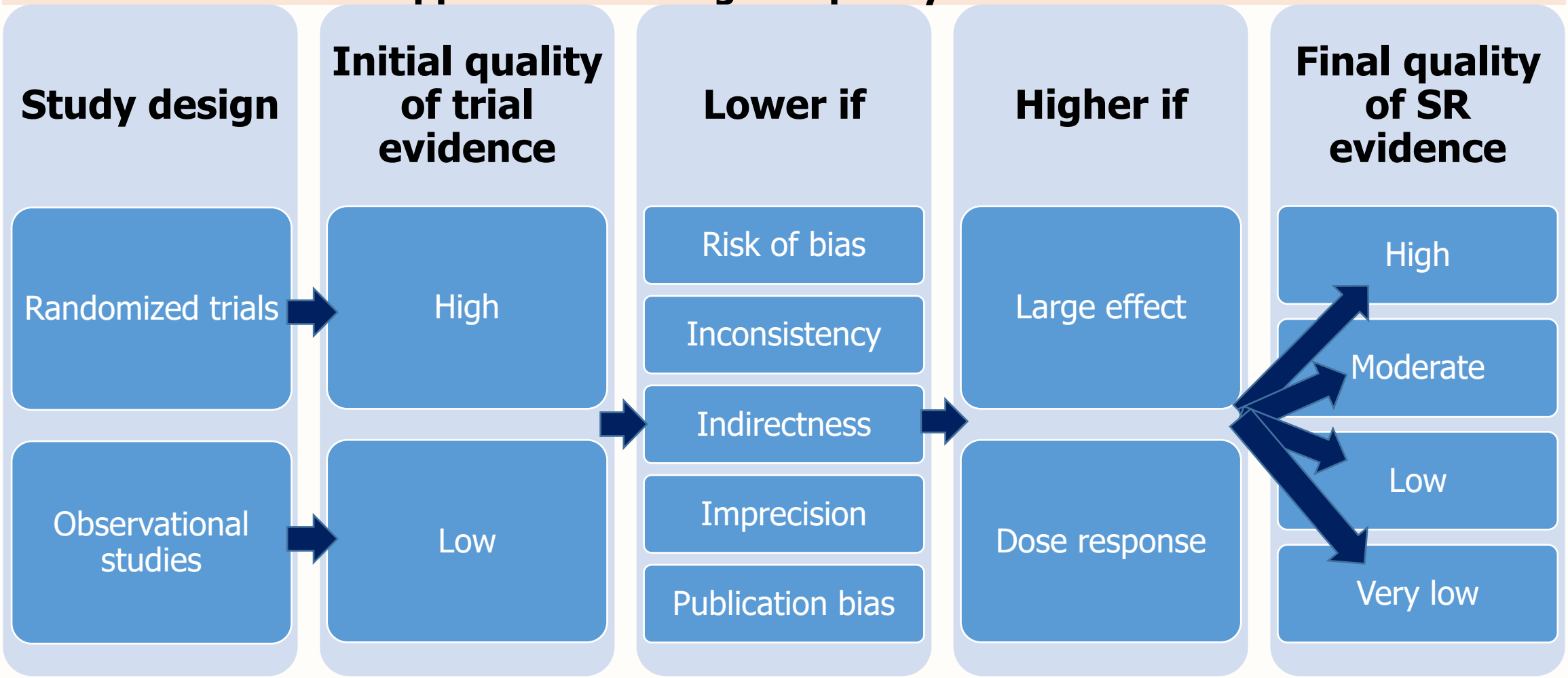


NMA should include only the studies for which the population, methodology and studied target condition is as similar as possible[2]

1. Tonin FS, et al. *Pharm Pract (Granada).* 2017;15(1):943.
2. Salanti G, et al. *PloS One*. 2014:9:e99682.

Advance and practical issues of SRs –

GRADE (Grading of Recommendations, Assessment, Development and Evaluations) approach on rating the quality of evidence[1]



1. Balshem H, et al. *J Clin Epidemiol.* 2011:64:401-406