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.
Outline

- SRs: their roles in clinical practice
- Conducting SRs: steps and tasks
- Advance and practical issues of SRs
ข. ความรู้ความสามารถทางวิชาชีพและทักษะทางคลินิก หมวดที่ 1. 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

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

SR and meta-analysis

• SRs - key characteristics include:
  o A clearly stated set of objectives with an explicit, reproducible methodology
  o A systematic search that attempts to identify all studies that would meet the eligibility criteria
  o 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
  o 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

SR: a must for practice guideline development

• A key step for guideline development
  o 1991: Are systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances
    (Note: consensus statements was acceptable)
  o 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

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

1.2 Defining selection criteria

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

• Defining the electronic databases to be searched, eg,
  o Bibliographic databases: eg, Pubmed (Medline)*, Scopus/EMBASE, Web of Science
    ▪ Describing the search strategy of each database
  o 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
  o Review question (objective)
  o Selection criteria
  o Search strategy
  o Study selection
  o Outcomes and priority
  o Risk of bias of individual studies
  o Data items and extraction
  o Data synthesis
  o Heterogeneity
  o Publication bias

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

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)

---

### 2.2 Risk of bias assessment (for randomized trials) (RoB1)

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

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

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

3.2 Data synthesis – interpretation of SMDs

- SMDs can be computed by many methods
- Cohen’s $d$:
  - $d = (m_t - m_c)/(pooled\ sd)$
  - $m_t =$ mean of treatment group
  - $m_c =$ 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, $g \approx d$
- Interpretation of Cohen’s $d^1$
  - $0.2 \approx$ small effect size (ES)
  - $0.5 \approx$ medium ES
  - $0.8 \approx$ 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.

### 3.2 Data synthesis – Forest plot

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental Total Mean</th>
<th>Control Total Mean</th>
<th>Standardised Mean Difference</th>
<th>Overall results with 95% CI</th>
<th>Weight (fixed)</th>
<th>Weight (random)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ikai (Yoga) 2014</td>
<td>25 0.04 0.5200</td>
<td>25 0.07 0.5800</td>
<td>-0.05 [-0.61; 0.50]</td>
<td>11.6% 12.8%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee (Yoga) 2014</td>
<td>14 7.42 0.7500</td>
<td>11 -6.53 10.3400</td>
<td>1.97 [0.98; 2.96]</td>
<td>3.6% 8.5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Naveen (Yoga) 2016</td>
<td>16 1.95 5.8000</td>
<td>23 -2.40 5.6000</td>
<td>0.75 [0.09; 1.41]</td>
<td>8.1% 11.6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sun (Tai Chi) 2018</td>
<td>33 0.17 0.1400</td>
<td>37 0.14 0.1700</td>
<td>0.19 [-0.28; 0.66]</td>
<td>16.1% 13.6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tolahunase (Yoga) 2018</td>
<td>29 5.57 6.0400</td>
<td>29 0.02 4.1100</td>
<td>1.06 [0.51; 1.61]</td>
<td>11.7% 12.8%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gagrani (Meditation) 2019</td>
<td>30 11.10 11.6700</td>
<td>30 1.19 5.7400</td>
<td>1.06 [0.52; 1.61]</td>
<td>12.1% 12.9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ledreux (Meditation) 2019</td>
<td>39 -0.01 0.8100</td>
<td>39 -0.96 0.7900</td>
<td>1.18 [0.69; 1.66]</td>
<td>15.3% 13.5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nerry (Meditation) 2019</td>
<td>62 0.73 1.8900</td>
<td>37 -0.35 13.8300</td>
<td>0.13 [-0.28; 0.53]</td>
<td>21.5% 14.3%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Fixed effect model**

<table>
<thead>
<tr>
<th>Overall SMD</th>
<th>248</th>
<th>231</th>
</tr>
</thead>
</table>

**Random effects model**

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

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

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.*

# PRISMA: 8 of 27 items

<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TITLE</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>1</td>
<td>Identify the report as a systematic review, meta-analysis, or both.</td>
</tr>
<tr>
<td><strong>ABSTRACT</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structured summary</td>
<td>2</td>
<td>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.</td>
</tr>
<tr>
<td><strong>INTRODUCTION</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rationale</td>
<td>3</td>
<td>Describe the rationale for the review in the context of what is already known.</td>
</tr>
<tr>
<td>Objectives</td>
<td>4</td>
<td>Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).</td>
</tr>
<tr>
<td><strong>METHODS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protocol and registration</td>
<td>5</td>
<td>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.</td>
</tr>
<tr>
<td>Eligibility criteria</td>
<td>6</td>
<td>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.</td>
</tr>
<tr>
<td>Information sources</td>
<td>7</td>
<td>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.</td>
</tr>
<tr>
<td>Search</td>
<td>8</td>
<td>Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.</td>
</tr>
</tbody>
</table>
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\(^2\)

Advance and practical issues of SRs – GRADE (Grading of Recommendations, Assessment, Development and Evaluations) approach on rating the quality of evidence

<table>
<thead>
<tr>
<th>Study design</th>
<th>Initial quality of trial evidence</th>
<th>Lower if</th>
<th>Higher if</th>
<th>Final quality of SR evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomized trials</td>
<td>High</td>
<td>Risk of bias</td>
<td>Large effect</td>
<td>High</td>
</tr>
<tr>
<td>Observational studies</td>
<td>Low</td>
<td>Inconsistency</td>
<td>Dose response</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indirectness</td>
<td></td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Imprecision</td>
<td></td>
<td>Very low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Publication bias</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>