Systematic review with meta-analysis: An overview

Professor Brigid Gillespie
Griffith University & Gold Coast Health, Queensland, Australia
Visiting Professor, Sahlgrenska Academy, University of Gothenburg, Sweden
ACCT Group Presentation, 15 December 2020

Acknowledgement: Some of the slides on meta-analysis have been adapted, based on Prof Joan Webster’s presentation
Overview

• What is a systematic review?
• What are the differences b/ a SR and other types of literature reviews?
• Components of a SR
• What is a meta-analysis?
• Components of a MA
• Summary of key points
What is a systematic review?

Focusses on a clearly formulated question that uses systematic and reproducible methods to identify, select and critically appraise all relevant research, and to collect and analyse data from the studies that are included in the review.

A systematic review
• Answers a focused research question.
• Employs a comprehensive, reproducible search strategy.
• Identifies ALL relevant studies (both published and unpublished).
• Assesses all results for inclusion/exclusion, and for quality.
• Presents an unbiased, balanced summary of findings.
• Involves a team of researchers looking at a complex research question.
• Takes months, or even years, to complete.
Types of systematic reviews

- A quantitative systematic review includes studies that have numerical data.
- A qualitative systematic review derives data from observation, interviews, or verbal interactions and focuses on the meanings and interpretations of the participants. It will include focus groups, interviews, observations and diaries.

Source: https://libguides.library.curtin.edu.au/systematic-reviews
## What are the differences between a SR and other types of literature reviews?

<table>
<thead>
<tr>
<th></th>
<th>Systematic review</th>
<th>Literature review</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Question</strong></td>
<td>Focused on a single question</td>
<td>Not necessarily focused on a single question, but may describe an overview</td>
</tr>
<tr>
<td><strong>Protocol</strong></td>
<td>Includes a peer review protocol or plan</td>
<td>No protocol is included</td>
</tr>
<tr>
<td><strong>Background</strong></td>
<td>Provides summaries of the available literature on a topic</td>
<td>Provides summaries of the available literature on a topic</td>
</tr>
<tr>
<td><strong>Objectives</strong></td>
<td>Clear objectives are identified</td>
<td>Objectives may or may not be identified</td>
</tr>
<tr>
<td><strong>Inclusion/exclusion criteria</strong></td>
<td>Criteria is stated before review is conducted</td>
<td>Criteria not specific or may not be specified</td>
</tr>
<tr>
<td><strong>Search strategy</strong></td>
<td>Comprehensive search conducted in a systematic way</td>
<td>Strategy not explicitly stated</td>
</tr>
<tr>
<td><strong>Article selection</strong></td>
<td>Process clear, explicit and replicable</td>
<td>Not always clearly described</td>
</tr>
<tr>
<td><strong>Process of evaluating articles</strong></td>
<td>Comprehensive evaluation of study quality</td>
<td>Evaluation of study quality may or may not be included</td>
</tr>
<tr>
<td><strong>Results and data synthesis</strong></td>
<td>Clear summaries based on high quality evidence</td>
<td>Summary based on studies where the quality of the articles may not be specified. May also be influenced by the reviewer's theories, needs and beliefs</td>
</tr>
<tr>
<td><strong>Discussion</strong></td>
<td>Written by an expert or group of experts with a detailed and well established knowledge of the issues</td>
<td>Written by an expert or group of experts with a well established knowledge of the issues</td>
</tr>
</tbody>
</table>

Adapted based on: https://libguides.library.curtin.edu.au/systematic-reviews
Steps in a systematic review

1. Check for existing reviews/protocols. If a systematic review answering question has been conducted, or if a specific research question that is clear and focused. Use the PICO tool.

2. Formulate a specific research question that is clear and focused. Use the PICO tool.

3. Develop and register your protocol, including the rationale for the review, and eligibility criteria.

4. Design a robust search strategy that is explicit and reproducible. Assistance from a health librarian with search terms and database searches is invaluable.

5. Conduct a comprehensive search of the literature by searching the relevant databases and other sources.


7. Extract relevant data from individual studies and use established methods to synthesise the data. If meta-analysis is appropriate, then include based on PICO question.

8. Interpret results, write a comprehensive report on all aspects of the systematic review. Present findings relative to their translation into clinical practice.
PICO question

- Population
- Intervention
- Comparator
- Outcome(s)
Example of a PICO question

Review objective: to assess the clinical effectiveness of repositioning regimens on the prevention of pressure injuries (PI) in adults, regardless of risk in any setting.

Population – any adult, without an existing PI, admitted to any healthcare setting.

Intervention(s) – comparisons b/ frequencies of repositioning, e.g., -2,-3,-4 hourly, different positions for repositioning, e.g., tilts.

Comparator – comparisons with standard practice, however defined by study authors.

Outcome – primary outcome cumulative incidence of PI
What is bias?

- A systematic error or deviation from the truth, in results or inferences.
- Biases can operate in either direction: different biases can lead to underestimation or overestimation of the true intervention effect.
- Biases can vary in magnitude: some are small (and trivial compared with the observed effect) and some are substantial.
- Even a particular source of bias may vary in direction: bias due to a particular design flaw (e.g. lack of allocation concealment), may lead to over estimation of effect.
- Because the results of a study may in fact be unbiased despite a methodological flaw, it is more appropriate to consider risk of bias.

https://handbook-5-1.cochrane.org/chapter_8/8_assessing_risk_of_bias_in_included_studies.htm
Assessing research quality in included papers

• Various tools for assessing risk of bias, e.g., Cochrane RoB tool for RCT (i.e., randomisation, allocation concealment, blinding, incomplete data, selective reporting, other bias)

• Risk of Bias (RoB) tools for non-randomised studies, e.g., ROBINS-I (“Risk Of Bias In Non-randomised Studies - of Interventions”)
Example of RoB assessment

Figure 4. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included clinical studies.

- Random sequence generation (selection bias)
- Allocation concealment (selection bias)
- Blinding of participants and personnel (performance bias): All outcomes
- Blinding of outcome assessment (detection bias): All outcomes
- Incomplete outcome data (attrition bias): All outcomes
- Selective reporting (reporting bias)

Legend:
- Green: Low risk of bias
- Yellow: Unclear risk of bias
- Red: High risk of bias

Studies included:
- Bergstrom 2013
- Defloor 2005
- Ghezeljeh 2017
- Manzano 2014
- Moore 2011
- Pickham 2018
- Young 2004
- Zhou 2014

Studies must be examining the same intervention against the same comparators, measuring the same outcome in the same population, using the same study design.
META-ANALYSIS

What is a meta-analysis?

• Combines the results of two or more studies
• Estimates a common or average treatment effect across studies
Why do a meta-analysis?

- Quantify treatment effects and their uncertainty
- Increase power
- Increase precision, larger sample size
- Explore differences between studies
- Settle controversies between studies
- Generate hypotheses
For example

- Eight trials studying the effect of different positioning regimens on the prevention of PI in adults.

- How can we summarise the effect of different positioning regimens across these trials?
When should you do a meta-analysis

- When more than one study has provided results about the same question
- When there are minimal differences across studies
- When the same outcome has been measured
- When data in each study are available
Some issues

- Only summary statistic available (e.g., $p = 0.01$) or OR 2.1 (95% CI 1.0 – 4.6).
- When more than one intervention has been used, e.g., turning regimen and support surfaces)
- When the outcome has been measured by different instruments, e.g., pain, QoL
When not to do a meta-analysis

Garbage in = garbage out!

• A meta-analysis is only as good as the studies in it.
• If included studies are biased
  ✓ Meta analysis results will be incorrect
  ✓ Will give more credibility and narrower confidence intervals
• If serious reporting biases are present
  ✓ Unrepresentative set of studies may give misleading result
Combining data

- Weighting studies
- More weight to studies that give more information
- More participants, more events, narrower confidence intervals
- Calculated using the estimated effect and its variance

**Comparison 2. 30° tilt 3-hourly overnight versus 90° tilt overnight**

<table>
<thead>
<tr>
<th>Outcome or subgroup title</th>
<th>No. of studies</th>
<th>No. of participants</th>
<th>Statistical method</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Pressure injury occurrence (stage 1 to 4)</td>
<td>2</td>
<td>252</td>
<td>Risk Ratio (M-H, Random, 95% CI)</td>
<td>0.62 [0.10, 3.97]</td>
</tr>
</tbody>
</table>

**Analysis 2.1. Comparison 2: 30° tilt 3-hourly overnight versus 90° tilt overnight, Outcome 1: Pressure injury occurrence (stage 1 to 4)**

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>30° tilt 3-hourly overnight</th>
<th>90° tilt overnight</th>
<th>Weight</th>
<th>Risk Ratio (M-H, Random, 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moore 2011</td>
<td>3 Events 99 Total</td>
<td>13 Events  28 Total</td>
<td>54.7%</td>
<td>0.27 [0.08 , 0.91]</td>
</tr>
<tr>
<td>Young 2004</td>
<td>3 Events 18 Total</td>
<td>2 Events  13 Total</td>
<td>45.3%</td>
<td>1.75 [0.33 , 9.34]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>117</td>
<td>135</td>
<td>100.0%</td>
<td>0.62 [0.10, 3.97]</td>
</tr>
</tbody>
</table>

Total events: 15

Heterogeneity: \( \tau^2 = 1.24; \ \chi^2 = 3.21; \ \text{df} = 1 \ (P = 0.07); \ F = 69\% 

Test for overall effect: \( Z = 0.50 \ (P = 0.62) \)

Test for subgroup differences: Not applicable
Displaying results

- Forest plot
- Provides a ‘snapshot’ of statistical results
- Identifies heterogeneity
- Shows the effect of individual studies and the ‘summary’ effect across studies

**Analysis 1.1. Comparison 1: 2-hourly repositioning versus 4-hourly repositioning on any type of support surface, Outcome 1: Pressure injury occurrence (stage 1 to 4)**

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>2-hourly repositioning</th>
<th>4-hourly repositioning</th>
<th>Risk Ratio</th>
<th>Risk Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Events</td>
<td>Total</td>
<td>Events</td>
<td>Total</td>
</tr>
<tr>
<td>Bergstrom 2013</td>
<td>8</td>
<td>321</td>
<td>9</td>
<td>295</td>
</tr>
<tr>
<td>Defloor 2005</td>
<td>39</td>
<td>63</td>
<td>30</td>
<td>66</td>
</tr>
<tr>
<td>Manzano 2014</td>
<td>17</td>
<td>165</td>
<td>22</td>
<td>164</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td></td>
<td>549</td>
<td></td>
<td>525</td>
</tr>
<tr>
<td>Total events:</td>
<td>64</td>
<td></td>
<td>61</td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: Chi² = 3.65, df = 2 (P = 0.16), I² = 45%
Test for overall effect: Z = 0.41 (P = 0.68)
Test for subgroup differences: Not applicable

Review: Use of repositioning regimes and positions to prevent pressure injuries

Comparison: 2 hourly and 4 hourly

Outcome: pressure injuries of any stage

### Comparison 1. 2-hourly repositioning versus 4-hourly repositioning on any type of support surface

<table>
<thead>
<tr>
<th>Outcome or subgroup title</th>
<th>No. of studies</th>
<th>No. of participants</th>
<th>Statistical method</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Pressure injury occurrence (stage 1 to 4)</td>
<td>3</td>
<td>1074</td>
<td>Risk Ratio (M-H, Fixed, 95% CI)</td>
<td>1.06 [0.80, 1.41]</td>
</tr>
</tbody>
</table>

### Analysis 1.1. Comparison 1: 2-hourly repositioning versus 4-hourly repositioning on any type of support surface, Outcome 1: Pressure injury occurrence (stage 1 to 4)
The scale

The horizontal line at the bottom is the scale measuring the treatment effect.

The outcome is PI, the right side of the scale is greater than 1, BUT the diamond crosses the line of ‘no effect’.

<table>
<thead>
<tr>
<th>Outcome or subgroup title</th>
<th>No. of studies</th>
<th>No. of participants</th>
<th>Statistical method</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Pressure injury occurrence (stage 1 to 4)</td>
<td>3</td>
<td>1074</td>
<td>Risk Ratio (M-H, Fixed, 95% CI)</td>
<td>1.06 [0.80, 1.41]</td>
</tr>
</tbody>
</table>

**Comparison 1. 2-hourly repositioning versus 4-hourly repositioning on any type of support surface**

**Analysis 1.1. Comparison 1: 2-hourly repositioning versus 4-hourly repositioning on any type of support surface, Outcome 1: Pressure injury occurrence (stage 1 to 4)**

- Bergstrom 2013
  - 2-hourly: 8 events, 321 total
  - 4-hourly: 9 events, 295 total, Weight: 15.4%
  - Risk Ratio: 0.82 [0.32, 2.09]

- DeGoor 2005
  - 2-hourly: 39 events, 63 total
  - 4-hourly: 30 events, 66 total, Weight: 48.2%
  - Risk Ratio: 1.36 [0.98, 1.89]

- Manzano 2014
  - 2-hourly: 17 events, 165 total
  - 4-hourly: 22 events, 164 total, Weight: 36.3%
  - Risk Ratio: 0.77 [0.42, 1.39]

- Total (95% CI): 64 total events, Weight: 100.0%
  - Risk Ratio: 1.06 [0.80, 1.41]
The line of no effect

- The vertical line in the middle is the line of no effect.
- Each horizontal line represents an individual study.
- If the horizontal line crosses the line of no effect, then there is no statistical difference between the treatment and control groups.
Individual trials

- Each study is given a square block representing the treatment effect.
- The size of the block is proportional to the weight given to that study.
- The horizontal line is the confidence interval (CI).
- The wider the confidence interval, the less likely the treatment effect is the true effect.
Each study

• For each study there is an ID (first author & year)
• Data for each trial are divided into experimental and control
• This is the % of weight given to each study in the pooled analysis
The summary statistic

- Data shown in the graph are also shown numerically
- The label above the graph indicates the summary statistic used

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>30° tilt 3-hourly overnight</th>
<th>90° tilt overnight</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Events</td>
<td>Total</td>
<td>Events</td>
</tr>
<tr>
<td>Moore 2011</td>
<td>3</td>
<td>99</td>
<td>13</td>
</tr>
<tr>
<td>Young 2004</td>
<td>3</td>
<td>18</td>
<td>2</td>
</tr>
</tbody>
</table>
The pooled result

- The diamond represents the treatment effect based on pooled results from the meta-analysis.
- The point estimate is represented by the vertical height of the diamond (0.62).
- The confidence interval is represented by the horizontal width of the diamond (0.10 to 3.97).
Interpreting confidence intervals

Always present the confidence interval with the treatment effect estimate

Precision
✓ The point estimate is the best guess of the effect of an intervention
✓ CI represents uncertainty – it is simply a range of values we can be reasonably sure contains the true effect

Significance
✓ If the confidence interval contains a null value
✓ It rarely means evidence of no effect
✓ It means effect cannot be confirmed or refuted by the available evidence

Consider what level of change is clinically important
Interpretation

- The heterogeneity between studies is represented by the $\chi^2$ & the $I^2$ (or can be assessed visually)
- The statistical difference between treatments is represented by the $Z$ score
Subgroup analyses

• Involve splitting all participant data into subgroups to make comparisons between them.

• Subgroup analyses for subsets of participants, e.g., males/females, or for subsets of studies, e.g., different geographical locations.

• Subgroup analyses to investigate heterogeneous results, or answer specific questions about particular patient groups, types of intervention or types of study.
Sensitivity analyses

• Analysis to determine how sensitive the results of a systematic review are to changes based on how it was done, e.g., one sensitivity analysis may explore the impact of using different meta-analysis models.

• Another sensitivity analysis may explore the impact of excluding or including studies in meta-analysis based on sample size, methodological quality, or variance. If results remain consistent across the different analyses, the results can be considered robust as even with different decisions they remain similar.

• Inclusion of studies based on quality or risk of bias can affect the pooled result of the meta-analysis

• E.g., sensitively analyses of studies at high risk of bias vs studies of low risk of bias
Other issues of interpretation

• Do the results make sense? i.e., biological plausibility.

• Do conclusions reflect findings? Avoid overstating inconclusive results.

• Applicability to clinical practice, the ‘so what’ question, external validity.
Summary of key points

• SR examine a focussed questions of clinical importance

• May or may not include meta-analysis, depending on how the outcomes were measured, levels of heterogeneity

• Require a research team where tasks can be allocated, +statistician if MA is included

• Decision to pool data influenced by degree of heterogeneity between studies.

• Conclusions must be supported by the results of the meta-analyses.

• Care in interpretation – sensitivity or sub-group analysis may be appropriate
THANK YOU
Prof. Brigid Gillespie
b.gillespie@griffith.edu.au