Introduction to systematic reviews

Dr Sally Hopewell
Oxford Clinical Trials Research Unit
NDORMS, University of Oxford

Outline of session

- What’s a systematic review
  - why do we need them?
  - What’s the process of conducting a systematic review?
- What’s a meta-analysis?
  - when can you do one?
  - how are the results displayed and interpreted?
Why do we need systematic reviews?

- Need information to make the right decisions
- But….too much information
- And…not enough time

- Individual trials may be biased or results presented out of context

Narrative review

- Conventional “narrative” literature review
  
  “Summary of the information available to the author from the point of view of the author”

- Can be very misleading as a summary from which to draw conclusions on overall evidence

- Reliable reviews must be systematic!
A systematic review collates all empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question.

- Clearly stated set of objectives with pre-defined criteria for studies
- Explicit reproducible methodology
- Systematic search to identify all studies meeting eligibility criteria
- Assessment of the validity of the findings of the included studies
- Systematic presentation and synthesis of the studies
Importance of systematic reviews

Decisions about health care require high quality information based on objective standards.
Results of a single trial are rarely sufficient to answer questions of best practices in clinical settings.
Much of the clinical research available is of relatively poor quality.
Resources are wasted each year on ineffective or harmful health care practices.

An example

Ian Roberts and his colleagues did the CRASH trial to address uncertainty about the effects of giving systemic steroids for people with acute traumatic brain injury, a treatment that had been in use for over three decades.
Systematic review of existing knowledge

Corticosteroids in acute traumatic brain injury: systematic review of randomised controlled trials

Philip Alderson, Ian Roberts


The review revealed important uncertainty about whether systemic steroids did more good than harm.

<table>
<thead>
<tr>
<th>Steroid</th>
<th>Control</th>
<th>Weight (%)</th>
<th>Odds ratio</th>
<th>Mantel-Haenszel odds ratio (95% confidence interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ransohoff 1972</td>
<td>9/17</td>
<td>3.1</td>
<td>0.43 (0.11 to 1.76)</td>
<td></td>
</tr>
<tr>
<td>Alexander 1972</td>
<td>16/55</td>
<td>8.0</td>
<td>0.62 (0.28 to 1.36)</td>
<td></td>
</tr>
<tr>
<td>Fauvel 1978</td>
<td>16/67</td>
<td>8.9</td>
<td>0.24 (0.09 to 0.60)</td>
<td></td>
</tr>
<tr>
<td>Cooper 1979</td>
<td>26/49</td>
<td>4.1</td>
<td>1.22 (0.46 to 3.12)</td>
<td></td>
</tr>
<tr>
<td>Hennesnemi 1979</td>
<td>35/81</td>
<td>10.4</td>
<td>0.99 (0.54 to 1.84)</td>
<td></td>
</tr>
<tr>
<td>Pitts 1980</td>
<td>114/201</td>
<td>12.4</td>
<td>1.24 (0.73 to 2.12)</td>
<td></td>
</tr>
<tr>
<td>Saul 1981</td>
<td>8/50</td>
<td>3.9</td>
<td>0.87 (0.31 to 2.47)</td>
<td></td>
</tr>
<tr>
<td>Braakman 1983</td>
<td>44/81</td>
<td>11.1</td>
<td>0.83 (0.45 to 1.56)</td>
<td></td>
</tr>
<tr>
<td>Giannotta 1984</td>
<td>34/72</td>
<td>3.1</td>
<td>1.15 (0.39 to 3.42)</td>
<td></td>
</tr>
<tr>
<td>Dearden 1986</td>
<td>33/68</td>
<td>5.8</td>
<td>1.84 (0.91 to 3.74)</td>
<td></td>
</tr>
<tr>
<td>Zagara 1987</td>
<td>4/12</td>
<td>1.4</td>
<td>1.00 (0.18 to 5.46)</td>
<td></td>
</tr>
<tr>
<td>Gaab 1994</td>
<td>19/133</td>
<td>9.2</td>
<td>0.91 (0.47 to 1.79)</td>
<td></td>
</tr>
<tr>
<td>Grumme 1995</td>
<td>38/175</td>
<td>18.7</td>
<td>0.83 (0.51 to 1.34)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>396/1061</td>
<td>100</td>
<td>0.91 (0.74 to 1.12)</td>
<td></td>
</tr>
</tbody>
</table>

Fig 1 Summary odds ratio for death at end of study
Addressing an important uncertainty

- Because the systematic review and a survey of clinical practice had revealed important uncertainty,
  - a large, publicly-funded, multicentre randomized trial was organised
  - the trial was registered prospectively
  - the protocol for the trial was published

Effect of intravenous corticosteroids on death within 14 days in 10008 adults with clinically significant head injury (MRC CRASH trial): randomised placebo-controlled trial

• *Lancet* 2004;364:1321-28
Avoidable waste in deciding what research to do, Lancet series, 2014

- **Questions relevant to users of research?**
  - Low priority questions asked
  - Important outcomes not assessed
  - Over 50% studies designed without reference to systematic reviews of existing evidence

- **Appropriate research design, conduct and analysis?**
  - Over 50% of studies fail to take adequate steps to reduce biases
  - Studies with inadequate statistical power
  - Inadequate replication of initial observations

- **Efficient research regulation and delivery?**
  - Over 50% of studies never published in full
  - Biased under-reporting of studies with disappointing results
  - Biased reporting of data within studies

- **Accessible, full research reports?**
  - Over 30% of trial interventions not sufficiently described
  - Over 50% of planned study outcomes not reported
  - Most new research not interpreted in the context of systematic assessment of other relevant evidence

- **Unbiased and usable reports?**
  - Over 50% of trial interventions not sufficiently described
  - Over 50% of planned study outcomes not reported

**Research waste**

---

**Figure 5: Updated meta-analysis of effect of corticosteroids on death after head injury**

<table>
<thead>
<tr>
<th>Study</th>
<th>Adjusted control</th>
<th>Relative risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alexander 1972</td>
<td>16/55</td>
<td>22/55</td>
</tr>
<tr>
<td>Bassett 1972</td>
<td>9/57</td>
<td>12/13</td>
</tr>
<tr>
<td>Bax 1976</td>
<td>16/67</td>
<td>18/28*2</td>
</tr>
<tr>
<td>Cooper 1979</td>
<td>26/49</td>
<td>13/27*2</td>
</tr>
<tr>
<td>Hackett 1979</td>
<td>30/61</td>
<td>38/68</td>
</tr>
<tr>
<td>Potts 1980</td>
<td>11/281</td>
<td>38/4*3</td>
</tr>
<tr>
<td>Sutcliffe 1981</td>
<td>8/50</td>
<td>9/5</td>
</tr>
<tr>
<td>Brown 1993</td>
<td>6/58</td>
<td>4/78</td>
</tr>
<tr>
<td>Gennetta 1984</td>
<td>3/22</td>
<td>7/18*3</td>
</tr>
<tr>
<td>Dearden 1986</td>
<td>33/68</td>
<td>21/62</td>
</tr>
<tr>
<td>Chacon 1987</td>
<td>1/5</td>
<td>0/5</td>
</tr>
<tr>
<td>Zagon 1987</td>
<td>6/4</td>
<td>4/12</td>
</tr>
<tr>
<td>Styn 1987</td>
<td>13/98</td>
<td>5/14*2</td>
</tr>
<tr>
<td>Goob 1994</td>
<td>19/133</td>
<td>21/136</td>
</tr>
<tr>
<td>Gammie 1995</td>
<td>38/75</td>
<td>49/195</td>
</tr>
<tr>
<td>Zanatta 1995</td>
<td>0/30</td>
<td>0/30</td>
</tr>
<tr>
<td><strong>Overall</strong> (95% CI)</td>
<td>160/610 (25.3%)</td>
<td>152/610 (25.3%)</td>
</tr>
</tbody>
</table>

**VRC-004 trial**

- 450/2134 (21.1%)
- 450/2134 (21.1%)
- 0.96 (0.85-1.08)

- 1.18 (1.09-1.27)
- 1.12 (1.05-1.20)
Archie Cochrane

“It is surely a great criticism of our profession that we have not organised a critical summary, by specialty or subspecialty, adapted periodically, of all relevant randomised controlled trials.” (1979)

The Cochrane Collaboration

Preparing, maintaining and promoting the accessibility of systematic reviews of the effects of health care interventions
What’s the process of conducting a systematic review?

- Framework for the review
- Planning
- Reduce bias
- Access to peer review
- Avoid duplication of effort
Titles

- Titles should be succinct
- Format
  - ‘Intervention’ for ‘problem’ in ‘category’
- Include ‘a systematic review of?’
- Avoid abbreviations

Background

Contains:
- Description of the condition
- Description of the intervention
- How the intervention might work
- Why it is important to do this review
The review question

- The review question should specify the types of population (participants), types of interventions (and comparisons), and the types of outcomes that are of interest.

- The acronym PICO (Participants, Interventions, Comparisons and Outcomes) helps to serve as a reminder of these.

Selection Criteria

- Type of studies
- Type of participants
- Type of interventions (and comparisons)
- Type of outcome measures
Levels of evidence

- Systematic reviews
- Randomised controlled trials
- Cohort studies
- Case-control studies
- Case series, case reports
- Editorials, expert opinion

Search Methods

- Show reader how studies were located
- Electronic searches:
  - Cochrane Central Register of Controlled Trials (CENTRAL)
  - Other electronic databases (e.g. Medline, Embase, PsycInfo, etc.)
- Searching other sources
  - Grey literature
  - Handsearching
  - Reference lists
  - Personal communication
  - Trial registers - ongoing studies
Data collection and analysis (1)

- Selection of studies
  - applying the selection criteria
    - Independently by more than one author
    - Identifying multiple reports of the same study
  - Should state how any disagreements will be resolved?
  - Selecting excluded studies
Data collection and analysis (2)

- Data extraction and management
  - which items?
  - how many authors?
  - format of data extraction sheet?
Data collection and analysis (3)

- **Assessment of risk of bias**
  - The Cochrane Collaboration has a recommended approach for randomized trials:
    - Risk of bias tool
      - Describe what was reported in the study
      - Assign a judgement relating to risk of bias
    - 6 parameters (the first 3 are most important)

Risk of bias: items to address

- Sequence generation (randomisation)
- Allocation concealment
- Blinding of participants, personnel and outcomes assessors
- Incomplete outcome data
- Selective outcome reporting
- Other (including topic-specific, design specific)
**Risk of bias summary**

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

<table>
<thead>
<tr>
<th>Study</th>
<th>Bias</th>
<th>Bias</th>
<th>Bias</th>
<th>Bias</th>
<th>Bias</th>
<th>Bias</th>
<th>Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detkicova 2004</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Kahe-Paradito 2002</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Mitocon 1988</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Mitocon 2007</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Narscote 1988</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Otte/Vada 1998</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

**Data synthesis**

- Analysis may include:
  - which comparisons?
  - to combine studies or not?
  - what statistical methods will be used?
  - subgroup analyses?
  - sensitivity analyses?
Summary of included studies

<table>
<thead>
<tr>
<th>Source</th>
<th>Setting</th>
<th>No. of Patients</th>
<th>Age Range</th>
<th>Inclusion Criteria</th>
<th>Antithrombotic Agent</th>
<th>Route</th>
<th>Follow-Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friedewald et al., 2005</td>
<td>ED</td>
<td>214</td>
<td>6 months–10 y</td>
<td>GE w/ mild to moderate deterioration and walking in the preceding 6 hours</td>
<td>Ondansetron</td>
<td>PO</td>
<td>1–2 weeks</td>
</tr>
<tr>
<td>Rinne et al., 2002</td>
<td>ED</td>
<td>187</td>
<td>1 month–22 y</td>
<td>GE and vomiting requiring IV rehydration</td>
<td>Ondansetron</td>
<td>IV</td>
<td>3–7 days</td>
</tr>
<tr>
<td>Ross and et al., 2007</td>
<td>ED</td>
<td>186</td>
<td>1–18 y</td>
<td>GE w/ actual or rehydration admitted in ED</td>
<td>Ondansetron</td>
<td>PO</td>
<td>1 week</td>
</tr>
<tr>
<td>Stark et al., 2006</td>
<td>ED</td>
<td>157</td>
<td>6 months–12 y</td>
<td>GE, recent anticoagulation, mild to moderate deterioration, and bilateral hypostasis</td>
<td>Ondansetron and desmopressin</td>
<td>IV</td>
<td>1 and 2 days</td>
</tr>
</tbody>
</table>

ED, emergency department; GE, gastrointestinal; IV, intravenous; PO, by mouth.

Optional part of a systematic review

- Systematic reviews
- Meta-analyses
What is a meta-analysis?

- Calculates a treatment effect based on pooled data from a group of studies
- Estimates a common treatment effect across studies
- Improves the precision of a point estimate by using all available data

When can/should you do a meta-analysis?

- When more than one study has estimated a treatment effect
- When there are minimal differences in characteristics across studies
- When the outcome has been measured in the same way
- When the data in each study are available
Performing a meta-analysis

- Calculate a single summary statistic to represent the effect found in each study
- Weighting each study gives us more information
  - More participants and more events combine to produce lower variance (e.g. narrower confidence interval) and more robust statistical results
- Display results graphically (forest plots)
  - Commonly used to assess heterogeneity
  - Provides a snapshot of statistical results

What does this forest plot tell us about the treatment?

Review: Therapeutic hypothermia for head injury
Comparison: Immediate hypothermia versus normothermia
Outcome: 01 Death at final follow-up

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Hypothermia Events</th>
<th>Control Events</th>
<th>Total</th>
<th>Weight</th>
<th>Odds Ratio M-H, Fixed, 95% CI</th>
<th>Odds Ratio M-H, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abidi 2006</td>
<td>1</td>
<td>15</td>
<td>3</td>
<td>11</td>
<td>2.7% 0.19 [0.02, 2.10]</td>
<td>2.7% 0.19 [0.02, 2.10]</td>
</tr>
<tr>
<td>Clifford 1992</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>0.7% 1.00 [0.04, 22.18]</td>
<td>0.7% 1.00 [0.04, 22.18]</td>
</tr>
<tr>
<td>Clifford 1993</td>
<td>8</td>
<td>23</td>
<td>8</td>
<td>22</td>
<td>4.0% 0.92 [0.28, 3.16]</td>
<td>4.0% 0.92 [0.28, 3.16]</td>
</tr>
<tr>
<td>Clifford 2001</td>
<td>83</td>
<td>190</td>
<td>48</td>
<td>178</td>
<td>30.0% 1.05 [0.68, 1.66]</td>
<td>30.0% 1.05 [0.68, 1.66]</td>
</tr>
<tr>
<td>Hirayama 1994</td>
<td>4</td>
<td>12</td>
<td>4</td>
<td>10</td>
<td>3.0% 0.90 [0.09, 8.17]</td>
<td>3.0% 0.90 [0.09, 8.17]</td>
</tr>
<tr>
<td>Jiang 2006</td>
<td>11</td>
<td>43</td>
<td>20</td>
<td>44</td>
<td>12.3% 0.41 [0.17, 1.02]</td>
<td>12.3% 0.41 [0.17, 1.02]</td>
</tr>
<tr>
<td>Moroni 1997</td>
<td>9</td>
<td>39</td>
<td>10</td>
<td>42</td>
<td>6.2% 0.96 [0.34, 2.69]</td>
<td>6.2% 0.96 [0.34, 2.69]</td>
</tr>
<tr>
<td>Meltzer 1998</td>
<td>3</td>
<td>12</td>
<td>3</td>
<td>13</td>
<td>1.8% 1.11 [0.10, 10.97]</td>
<td>1.8% 1.11 [0.10, 10.97]</td>
</tr>
<tr>
<td>Stickel 1999</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>8</td>
<td>Not estimable</td>
<td>Not estimable</td>
</tr>
<tr>
<td>Stickel 2001</td>
<td>8</td>
<td>45</td>
<td>6</td>
<td>51</td>
<td>4.1% 1.44 [0.40, 4.52]</td>
<td>4.1% 1.44 [0.40, 4.52]</td>
</tr>
<tr>
<td>Yan 2001</td>
<td>13</td>
<td>24</td>
<td>16</td>
<td>20</td>
<td>6.7% 0.30 [0.08, 1.10]</td>
<td>6.7% 0.30 [0.08, 1.10]</td>
</tr>
<tr>
<td>Zhang 2009</td>
<td>41</td>
<td>123</td>
<td>50</td>
<td>123</td>
<td>26.0% 0.72 [0.43, 1.20]</td>
<td>26.0% 0.72 [0.43, 1.20]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>539</td>
<td>522</td>
<td>338</td>
<td>100.0%</td>
<td>0.86 [0.61, 1.04]</td>
<td>0.86 [0.61, 1.04]</td>
</tr>
</tbody>
</table>

Total events 152 / 170
Heterogeneity: CH^2 = 8.53, df = 12 (p = 0.58); I^2 = 0%
Test for overall effect: Z = 1.66 (p = 0.10)
Subgroup analyses

Certain factors may produce misleading results of statistical analysis

- One way to assess the impact that differences in participant characteristics have on pooled results
- Differences across studies may lead to inaccurate measure of treatment effect
  - Example: participants with mild vs. severe level of disease, young vs. old
- Careful with interpretation
  - Subsequent studies often fail to confirm findings of subgroup results

Example of forest plot with subgroup analysis
Sensitivity analysis

- Investigates influence, bias, and robustness
- Variations in statistical methods, methodological quality, and degree of bias in each study can effect pooled result of meta-analysis
- Are the findings influenced by choice of statistical model?...
- Is bias in study methods (allocation concealment, blinding) affecting the outcome?
- Are the findings robust to different assumptions (intention to treat, missing data)?

A common sensitivity analysis is to repeat meta-analysis after removing trials at high risk of bias.
Other issues in interpretation

- Does the result make sense?
  - Biological plausibility

- Conclusions reflect findings
  - Don’t talk up inconclusive results

- Applicability to clinical practice
  - The ‘So what?’ question

*Cochrane Library* [www.cochrane.org](http://www.cochrane.org)
*Cochrane Collaboration* [www.cochrane.co.uk/en/index.htm](http://www.cochrane.co.uk/en/index.htm)
*Cochrane Handbook for Systematic Reviews of Interventions* [www.cochrane.dk/cochrane/handbook/hbook.htm](http://www.cochrane.dk/cochrane/handbook/hbook.htm)

Centre for Reviews and Dissemination (York)
* [http://www.york.ac.uk/inst/erd/index_guidance.htm](http://www.york.ac.uk/inst/erd/index_guidance.htm) Download Handbook for free !!