Introduction to systematic reviews

Dr Sally Hopewell
Centre for Statistics in Medicine
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!
Systematic review

- A systematic review collates all empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question.
Systematic review

Key characteristics included:

- 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

and Cochrane Database of Systematic Reviews.

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

- A systematic review collates all empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question.
Systematic review

Key characteristics included:

- 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
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>Steroid</th>
<th>Control</th>
<th>Weight (%)</th>
<th>Mantel-Haenszel odds ratio (95% confidence interval)</th>
<th>Odds ratio (95% confidence interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ransohoff 1972</td>
<td>9/17</td>
<td>13/18</td>
<td>3.1</td>
<td></td>
<td>0.43 (0.11 to 1.76)</td>
</tr>
<tr>
<td>Alexander 1972</td>
<td>16/55</td>
<td>22/55</td>
<td>8.0</td>
<td></td>
<td>0.62 (0.28 to 1.36)</td>
</tr>
<tr>
<td>Faupel 1976</td>
<td>16/67</td>
<td>16/28</td>
<td>8.9</td>
<td></td>
<td>0.24 (0.09 to 0.60)</td>
</tr>
<tr>
<td>Cooper 1979</td>
<td>26/49</td>
<td>13/27</td>
<td>4.1</td>
<td></td>
<td>1.22 (0.48 to 3.12)</td>
</tr>
<tr>
<td>Hernesniemi 1979</td>
<td>35/81</td>
<td>36/83</td>
<td>10.4</td>
<td></td>
<td>0.99 (0.54 to 1.84)</td>
</tr>
<tr>
<td>Pitts 1980</td>
<td>114/201</td>
<td>38/74</td>
<td>12.4</td>
<td></td>
<td>1.24 (0.73 to 2.12)</td>
</tr>
<tr>
<td>Saul 1981</td>
<td>8/50</td>
<td>9/50</td>
<td>3.9</td>
<td></td>
<td>0.87 (0.31 to 2.47)</td>
</tr>
<tr>
<td>Braakman 1983</td>
<td>44/81</td>
<td>47/80</td>
<td>11.1</td>
<td></td>
<td>0.83 (0.45 to 1.56)</td>
</tr>
<tr>
<td>Giannotta 1984</td>
<td>34/72</td>
<td>7/16</td>
<td>3.1</td>
<td></td>
<td>1.15 (0.39 to 3.42)</td>
</tr>
<tr>
<td>Dearden 1986</td>
<td>33/68</td>
<td>21/62</td>
<td>5.8</td>
<td></td>
<td>1.84 (0.91 to 3.74)</td>
</tr>
<tr>
<td>Zagara 1987</td>
<td>4/12</td>
<td>4/12</td>
<td>1.4</td>
<td></td>
<td>1.00 (0.18 to 5.46)</td>
</tr>
<tr>
<td>Gaab 1994</td>
<td>19/133</td>
<td>21/136</td>
<td>9.2</td>
<td></td>
<td>0.91 (0.47 to 1.79)</td>
</tr>
<tr>
<td>Grumme 1995</td>
<td>38/175</td>
<td>49/195</td>
<td>18.7</td>
<td></td>
<td>0.83 (0.51 to 1.34)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>396/1061</td>
<td>296/836</td>
<td>100</td>
<td></td>
<td>0.91 (0.74 to 1.12)</td>
</tr>
</tbody>
</table>

$\chi^2 = 15.99; \ df = 12; \ Z = 0.89$

**Fig 1** Summary odds ratio for death at end of study
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 10 008 adults with clinically significant head injury (MRC CRASH trial): randomised placebo-controlled trial

CRASH trial collaborators*

•Lancet 2004;364:1321-28
What’s the process of conducting a systematic review?
Why have a protocol?

- 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
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
ClinicalTrials.gov is a registry and results database of publicly and privately supported clinical studies of human participants conducted around the world. Learn more about clinical studies and about this site, including relevant history, policies, and laws.

ClinicalTrials.gov currently lists 158,418 studies with locations in all 50 states and in 185 countries.

Search for Studies
Example: "Heart attack" AND "Los Angeles"

Advanced Search | See Studies by Topic
See Studies on a Map

Search Help
- How to search
- How to find results of studies
- How to read a study record

Locations of Recruiting Studies
- Non-U.S. Only (50%)
- U.S. Only (44%)
- Both U.S. & Non-U.S. (6%)

Total N = 31,616 studies
Data as of January 02, 2014

International Clinical Trials Registry Platform (ICTRP)

Welcome to the WHO ICTRP
The mission of the WHO International Clinical Trials Registry Platform is to ensure that a complete view of research is accessible to all those involved in health care decision making. This will improve research transparency and will ultimately strengthen the validity and value of the scientific evidence base.

The registration of all interventional trials is a scientific, ethical and moral responsibility.
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
PRISMA Flow diagram

1. Identification
   - # of records identified through database searching
   - # of additional records identified through other sources
   - # of records after duplicates removed

2. Screening
   - # of records screened
   - # of records excluded

3. Eligibility
   - # of full-text articles assessed for eligibility
   - # of full-text articles excluded, with reasons

4. Included
   - # of studies included in qualitative synthesis
   - # of studies included in quantitative synthesis (meta-analysis)
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

<table>
<thead>
<tr>
<th>Study</th>
<th>Random sequence generation</th>
<th>Allocation concealment</th>
<th>Blinding of participants and personnel</th>
<th>Blinding of outcome assessment</th>
<th>Incomplete outcome data</th>
<th>Selective reporting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delicicza 2004</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
</tr>
<tr>
<td>Kahve-Paradiso 2002</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
</tr>
<tr>
<td>Morrocona 2007</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
</tr>
<tr>
<td>Norscane 1998</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>?</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>Antiemetic Agent</th>
<th>Route</th>
<th>Follow-Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freedman et al., 2006</td>
<td>ED</td>
<td>214</td>
<td>6 months–10 years</td>
<td>GE with mild to moderate dehydration and vomiting in the preceding 4 hours</td>
<td>Ondansetron</td>
<td>PO</td>
<td>1–2 weeks</td>
</tr>
<tr>
<td>Reeves et al., 2002</td>
<td>ED</td>
<td>107</td>
<td>1 month–22 years</td>
<td>GE and vomiting requiring IV rehydration</td>
<td>Ondansetron</td>
<td>IV</td>
<td>5–7 days</td>
</tr>
<tr>
<td>Roslund et al., 2007</td>
<td>ED</td>
<td>106</td>
<td>1–10 years</td>
<td>GE with failed oral rehydration attempt in ED</td>
<td>Ondansetron</td>
<td>PO</td>
<td>1 week</td>
</tr>
<tr>
<td>Stork et al., 2006</td>
<td>ED</td>
<td>137</td>
<td>6 months–12 years</td>
<td>GE, recurrent emesis, mild to moderate dehydration, and failed oral hydration</td>
<td>Ondansetron and dexamethasone</td>
<td>IV</td>
<td>1 and 2 days</td>
</tr>
</tbody>
</table>

ED, emergency department; GE, gastroenteritis; IV, intravenous; PO, by mouth. Adapted from [135].

doi:10.1371/journal.pmed.1000100.t002
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
Optional part of a systematic review

- Systematic reviews
- Meta-analyses
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: 01 Immediate hypothermia versus normothermia
Outcome: 01 Death at final follow-up

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Hypothermia Events</th>
<th>Hypothermia Total</th>
<th>Control Events</th>
<th>Control 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>Aibiki 2000</td>
<td>1</td>
<td>15</td>
<td>3</td>
<td>11</td>
<td>2.7%</td>
<td>0.19 [0.02, 2.15]</td>
<td></td>
</tr>
<tr>
<td>Clifton 1992</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>0.7%</td>
<td>1.00 [0.05, 22.18]</td>
<td></td>
</tr>
<tr>
<td>Clifton 1993</td>
<td>8</td>
<td>23</td>
<td>8</td>
<td>22</td>
<td>4.5%</td>
<td>0.93 [0.28, 3.16]</td>
<td></td>
</tr>
<tr>
<td>Clifton 2001</td>
<td>53</td>
<td>190</td>
<td>48</td>
<td>178</td>
<td>30.0%</td>
<td>1.05 [0.66, 1.66]</td>
<td></td>
</tr>
<tr>
<td>Hirayama 1994</td>
<td>4</td>
<td>12</td>
<td>5</td>
<td>10</td>
<td>3.0%</td>
<td>0.50 [0.09, 2.81]</td>
<td></td>
</tr>
<tr>
<td>Jiang 2000</td>
<td>11</td>
<td>43</td>
<td>20</td>
<td>44</td>
<td>12.3%</td>
<td>0.41 [0.17, 1.02]</td>
<td></td>
</tr>
<tr>
<td>Marion 1997</td>
<td>9</td>
<td>39</td>
<td>10</td>
<td>42</td>
<td>6.2%</td>
<td>0.96 [0.34, 2.69]</td>
<td></td>
</tr>
<tr>
<td>Meissner 1998</td>
<td>3</td>
<td>12</td>
<td>3</td>
<td>13</td>
<td>1.8%</td>
<td>1.11 [0.18, 6.97]</td>
<td></td>
</tr>
<tr>
<td>Shiozaki 1999</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>8</td>
<td>Not estimable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shiozaki 2001</td>
<td>8</td>
<td>45</td>
<td>6</td>
<td>46</td>
<td>4.1%</td>
<td>1.44 [0.46, 4.55]</td>
<td></td>
</tr>
<tr>
<td>Yan 2001</td>
<td>13</td>
<td>24</td>
<td>16</td>
<td>20</td>
<td>6.7%</td>
<td>0.30 [0.08, 1.15]</td>
<td></td>
</tr>
<tr>
<td>Zhang 2000</td>
<td>41</td>
<td>123</td>
<td>50</td>
<td>123</td>
<td>28.0%</td>
<td>0.73 [0.43, 1.23]</td>
<td></td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>539</strong></td>
<td><strong>522</strong></td>
<td>100.0%</td>
<td></td>
<td>0.80</td>
<td>[0.61, 1.04]</td>
<td></td>
</tr>
</tbody>
</table>

Total events 152 170

Heterogeneity: $\chi^2 = 8.53$, df = 10 ($P = 0.58$); $I^2 = 0$

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

Review: Antibiotics for acute otitis media in children
Comparison: 01 Antibiotic versus placebo
Outcome: 01 Pain

### 4.1.1 Pain at 24 hours

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Treatment</th>
<th>Control</th>
<th>Peto Odds Ratio (Peto, Fixed, 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burke 1991</td>
<td>53/112</td>
<td>56/117</td>
<td>0.98 [0.58, 1.64]</td>
</tr>
<tr>
<td>Thalin 1985</td>
<td>58/159</td>
<td>58/158</td>
<td>0.99 [0.63, 1.56]</td>
</tr>
<tr>
<td>vanBuchem 1981a</td>
<td>13/47</td>
<td>11/40</td>
<td>1.01 [0.39, 2.57]</td>
</tr>
<tr>
<td>vanBuchem 1981b</td>
<td>17/48</td>
<td>10/36</td>
<td>1.41 [0.56, 3.55]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>366/751</td>
<td>351/1000</td>
<td>1.03 [0.76, 1.39]</td>
</tr>
</tbody>
</table>

Total events: 141/135
Heterogeneity: $\chi^2 = 0.52, df = 3 (P = 0.91)$; $I^2 = 0$
Test for overall effect: $Z = 0.17 (P = 0.86)$

### 4.1.2 Pain at 2-7 days

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Treatment</th>
<th>Control</th>
<th>Peto Odds Ratio (Peto, Fixed, 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appelman 1991</td>
<td>11/66</td>
<td>10/53</td>
<td>0.86 [0.33, 2.21]</td>
</tr>
<tr>
<td>Burke 1991</td>
<td>20/111</td>
<td>29/114</td>
<td>0.65 [0.34, 1.22]</td>
</tr>
<tr>
<td>Damoiseaux 2000</td>
<td>69/117</td>
<td>89/123</td>
<td>0.55 [0.32, 0.94]</td>
</tr>
<tr>
<td>Halsted 1968</td>
<td>17/62</td>
<td>7/27</td>
<td>1.08 [0.39, 2.97]</td>
</tr>
<tr>
<td>Kaleida 1991</td>
<td>19/488</td>
<td>38/492</td>
<td>0.50 [0.29, 0.85]</td>
</tr>
<tr>
<td>Mygind 1981</td>
<td>15/72</td>
<td>29/77</td>
<td>0.45 [0.22, 0.90]</td>
</tr>
<tr>
<td>Thalin 1985</td>
<td>15/158</td>
<td>25/158</td>
<td>0.57 [0.29, 1.10]</td>
</tr>
<tr>
<td>vanBuchem 1981a</td>
<td>4/38</td>
<td>3/46</td>
<td>1.68 [0.36, 7.87]</td>
</tr>
<tr>
<td>vanBuchem 1981b</td>
<td>5/48</td>
<td>4/38</td>
<td>0.99 [0.25, 3.94]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>1160/128</td>
<td>1000/1000</td>
<td>0.61 [0.48, 0.77]</td>
</tr>
</tbody>
</table>

Total events: 175/234
Heterogeneity: $\chi^2 = 5.36, df = 8 (P = 0.72)$; $I^2 = 0$
Test for overall effect: $Z = 4.02 (P < 0.0001)$
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/crd/index_guidance.htm](http://www.york.ac.uk/inst/crd/index_guidance.htm)

  Download Handbook for free !!