Introduction to Systematic Reviews & meta-analysis

Madhukar Pai, MD, PhD
Associate Professor
Department of Epidemiology & Biostatistics
McGill University, Montreal, Canada

Email: madhukar.pai@mcgill.ca
Knowledge Synthesis for Knowledge Translation

CIHR defines knowledge translation as 'a dynamic and iterative process that includes the synthesis, dissemination, exchange and ethically-sound application of knowledge to improve the health of Canadians, provide more effective health services and products and strengthen the healthcare system'. This definition highlights the importance of knowledge synthesis in knowledge translation activities.

CIHR defines synthesis as 'the contextualization and integration of research findings of individual research studies within the larger body of knowledge on the topic. A synthesis must be reproducible and transparent in its methods, using quantitative and/or qualitative methods. It could take the form of a systematic review; follow the methods developed by The Cochrane Collaboration; result from a consensus conference or expert panel and may synthesize qualitative or quantitative results. Realist syntheses, narrative syntheses, meta-analyses, meta-syntheses and practice guidelines are all forms of synthesis.'
<table>
<thead>
<tr>
<th>Type of review</th>
<th>Sources of evidence used</th>
<th>Type of question</th>
<th>Example of question</th>
<th>Specific methods used</th>
</tr>
</thead>
</table>
| Intervention review [12]      | Qualitative or quantitative | Does the intervention of interest work for a particular outcome?                  | What is the effectiveness of vitamin E for the treatment of cardiovascular disease? [13]                   | - Studies: often limited to data from experimental studies  
- Risk of bias assessment: often focused on experimental studies  
- Analysis: may be qualitative or quantitative (meta-analysis)  
- Studies: often limited to data from experimental studies  
- Risk of bias assessment: often focused on experimental studies  
- Analysis: often allows both direct and indirect comparisons (ie, when the strategies have not been directly compared) |
| Network meta-analysis [14]    | Quantitative             | Does the intervention of interest work for a particular outcome?                  | What are the odds of developing diabetes during long-term treatment with an initial class of antihypertensive drug? [15] | - Studies: focused on studies of test accuracy for example  
- Risk of bias: focused on issues pertinent to diagnostic studies  
- Analysis: combined results will be based on the summary statistic and sources of heterogeneity, especially variation in diagnostic thresholds |
| Diagnostic test review [16]   | Qualitative or quantitative | How well does the diagnostic test work for a particular group of patients?       | What is the diagnostic accuracy of sentinel node biopsy, positron-emission tomography, magnetic resonance imaging, and computed tomography in determining lymph node status in patients with cervical cancer? [17] | - Studies: often limited to data from observational studies  
- Risk of bias: focuses on issues pertinent to genome association studies  
- Analysis:  
  - Quantitative test of bias and confounding usually performed (i.e., Hardy–Weinberg Equilibrium)  
  - Statistical methods used to combine results will be strongly based on heterogeneity |
| Human genome epidemiology reviews [18] | Quantitative | Which genes are associated with particular outcomes?                            | What is the susceptibility of 160A allele carriers to seven types of cancers? [19]                          | - Studies: often limited to data from observational studies  
- Risk of bias: focuses on issues pertinent to genome association studies  
- Analysis:  
  - compares the outcome for groups with different prognostic variables  
  - Must adjust for many confounding variables  
  - May involve survival analysis if reported outcome is time to event  
- Questions: framed in broad open-ended format  
- Literature search: involves browsing relevant perspectives and approaches, finding seminal conceptual articles by tracking references of references  
- Analysis: involves mapping and comparing storylines  
- Question: scope of the review involves additional steps, such as identifying key theories to be explored  
- Literature search: involves searching for theories explaining why the program works  
- Data abstraction: includes contextual data, such as theory, process detail, and historical aspects of the study  
- Analysis: often involves an iterative process  
- Question: question must be answerable with qualitative research  
- Literature search: involves searching electronic databases beyond the medical domain as well as searching books and theses  
- Data abstraction: involves abstracting metaphors, quotations, and common themes  
- Analysis: focuses on narrative summary of the evidence, data may be arranged in a matrix to show the commonalities of themes across studies |
| Prognostic review [20]        | Quantitative             | How can you predict a disease outcome more accurately or efficiently?            | Does B-type natriuretic peptide (BNP) predict mortality or other cardiac endpoints in persons diagnosed with coronary artery disease? [21] | - Studies: often limited to data from observational studies  
- Risk of bias: often focused on issues pertinent to prognostic studies  
- Analysis:  
  - compares the outcome for groups with different prognostic variables  
  - Must adjust for many confounding variables  
  - May involve survival analysis if reported outcome is time to event  
- Questions: framed in broad open-ended format  
- Literature search: involves browsing relevant perspectives and approaches, finding seminal conceptual articles by tracking references of references  
- Analysis: involves mapping and comparing storylines  
- Question: scope of the review involves additional steps, such as identifying key theories to be explored  
- Literature search: involves searching for theories explaining why the program works  
- Data abstraction: includes contextual data, such as theory, process detail, and historical aspects of the study  
- Analysis: often involves an iterative process  
- Question: question must be answerable with qualitative research  
- Literature search: involves searching electronic databases beyond the medical domain as well as searching books and theses  
- Data abstraction: involves abstracting metaphors, quotations, and common themes  
- Analysis: focuses on narrative summary of the evidence, data may be arranged in a matrix to show the commonalities of themes across studies |
| Meta-narrative review [22]    | Qualitative              | How best can one explain complex bodies of evidence?                            | How best to explain the diffusion of innovation in evidence-based medicine? [22]                           | - Studies: often limited to data from observational studies  
- Risk of bias: focuses on issues pertinent to genome association studies  
- Analysis:  
  - compares the outcome for groups with different prognostic variables  
  - Must adjust for many confounding variables  
  - May involve survival analysis if reported outcome is time to event  
- Questions: framed in broad open-ended format  
- Literature search: involves browsing relevant perspectives and approaches, finding seminal conceptual articles by tracking references of references  
- Analysis: involves mapping and comparing storylines  
- Question: scope of the review involves additional steps, such as identifying key theories to be explored  
- Literature search: involves searching for theories explaining why the program works  
- Data abstraction: includes contextual data, such as theory, process detail, and historical aspects of the study  
- Analysis: often involves an iterative process  
- Question: question must be answerable with qualitative research  
- Literature search: involves searching electronic databases beyond the medical domain as well as searching books and theses  
- Data abstraction: involves abstracting metaphors, quotations, and common themes  
- Analysis: focuses on narrative summary of the evidence, data may be arranged in a matrix to show the commonalities of themes across studies |
| Realist review [23]           | Qualitative              | How do complex programs work (or why do they fail) in certain contexts and settings? | Which aspects of school feeding programs in disadvantaged children determine success and failure in various situations? [24] | - Studies: often limited to data from observational studies  
- Risk of bias: focuses on issues pertinent to genome association studies  
- Analysis:  
  - compares the outcome for groups with different prognostic variables  
  - Must adjust for many confounding variables  
  - May involve survival analysis if reported outcome is time to event  
- Questions: framed in broad open-ended format  
- Literature search: involves browsing relevant perspectives and approaches, finding seminal conceptual articles by tracking references of references  
- Analysis: involves mapping and comparing storylines  
- Question: scope of the review involves additional steps, such as identifying key theories to be explored  
- Literature search: involves searching for theories explaining why the program works  
- Data abstraction: includes contextual data, such as theory, process detail, and historical aspects of the study  
- Analysis: often involves an iterative process  
- Question: question must be answerable with qualitative research  
- Literature search: involves searching electronic databases beyond the medical domain as well as searching books and theses  
- Data abstraction: involves abstracting metaphors, quotations, and common themes  
- Analysis: focuses on narrative summary of the evidence, data may be arranged in a matrix to show the commonalities of themes across studies |
| Meta-ethnography review [25]  | Qualitative              | How can qualitative evidence explain why certain interventions work and others do not? | What are the types of factors that could influence adherence to tuberculosis treatment from the patient’s experience? [25] | - Studies: often limited to data from observational studies  
- Risk of bias: focuses on issues pertinent to genome association studies  
- Analysis:  
  - compares the outcome for groups with different prognostic variables  
  - Must adjust for many confounding variables  
  - May involve survival analysis if reported outcome is time to event  
- Questions: framed in broad open-ended format  
- Literature search: involves browsing relevant perspectives and approaches, finding seminal conceptual articles by tracking references of references  
- Analysis: involves mapping and comparing storylines  
- Question: scope of the review involves additional steps, such as identifying key theories to be explored  
- Literature search: involves searching for theories explaining why the program works  
- Data abstraction: includes contextual data, such as theory, process detail, and historical aspects of the study  
- Analysis: often involves an iterative process  
- Question: question must be answerable with qualitative research  
- Literature search: involves searching electronic databases beyond the medical domain as well as searching books and theses  
- Data abstraction: involves abstracting metaphors, quotations, and common themes  
- Analysis: focuses on narrative summary of the evidence, data may be arranged in a matrix to show the commonalities of themes across studies |
Converting knowledge into health

Bridging the Gap Between Knowledge and Health
The Epidemiologist as Accountable Health Advocate ("AHA!")

David W. Dowdy and Madhukar Pai
Why we need systematic reviews
Case Study 1: “Egg on their faces: the story of human albumin solution”*

- Human albumin solution, a blood product, has been used in the treatment of blood loss and burns since the attack on Pearl Harbour over half a century ago.
- In the UK alone, an estimated 100,000 patients are treated with human albumin solution each year, at a cost to the NHS of close to 12 million.
- In 1996, the global albumin market was worth £900,000.
- But is human albumin administration beneficial?

“Egg on their faces: the story of human albumin solution”

- To answer this question a systematic review of controlled trials comparing albumin with crystalloid was conducted by the Cochrane Injuries Group.
- 30 RCTs including 1419 randomised patients identified.
- A meta-analysis showed that the risk of death among those treated with albumin was higher than in the comparison groups.
- The pooled risk ratio was 1.68 (95% CI 1.26, 2.23)
- The data suggested that for every seventeen critically ill patients treated with albumin there is one extra death.

“Egg on their faces: the story of human albumin solution”

- The results were widely reported in the media and stimulated an immediate response from the regulatory agencies, the industry and the medical profession.
- The industry launched a “Albumin Support Programme” to resuscitate the ailing US $ 1.5 billion global albumin market.
  - The objective was to disseminate evidence supporting albumin:
    - (1) the preparation of literature reviews supporting the use of albumin to be sent to leading regulatory authorities
    - (2) preparation and dissemination of a Cochrane critique dossier
    - (3) the establishment of a medical advisory panel to write articles supporting the use of albumin.
- The industry set aside US $2.2 million for the program.

“Egg on their faces: the story of human albumin solution”

• “Despite vigorous attempts by the plasma products industry to limit the impact of the systematic review on albumin sales, the use of albumin declined steeply.
• Throughout the UK albumin sales fell by 40%.
• The decline in albumin use occurred despite vigorous criticism of the review in the letters pages of the BMJ.
• The decline in albumin sales is a clear indication that doctors took into account the evidence presented in the systematic review and that many doctors changed their practice in response.”

Patient Survival after Human Albumin Administration

A Meta-Analysis of Randomized, Controlled Trials

Mahlon M. Wilkes, PhD, and Roberta J. Navickis, PhD

Purpose: To test the hypothesis that albumin administration is not associated with excess mortality.

Data Sources: Computer searches of the MEDLINE and EMBASE databases, the Cochrane Library, and Internet documents; hand searching of medical journals; inquiries to investigators and medical directors; and review of reference lists.

Study Selection: Randomized, controlled trials comparing albumin therapy with crystalloid therapy, no albumin, or lower doses of albumin.

Data Extraction: Two investigators independently extracted data. The primary end point was relative risk for death. Criteria used to assess methodologic quality were blinding, method of allocation concealment, presence of mortality as a study end point, and crossover. Small-trial bias was also investigated.

Data Synthesis: Fifty-five trials involving surgery or trauma, burns, hypoalbuminemia, high-risk neonates, ascites, and other indications were included. Albumin administration did not significantly affect mortality in any category of indications. For all trials, the relative risk for death was 1.11 (95% CI, 0.95 to 1.28). Relative risk was lower among trials with blinding (0.73 [CI, 0.48 to 1.12]; n = 7), mortality as an end point (1.00 [CI, 0.84 to 1.18]; n = 17), no crossover (1.04 [CI, 0.89 to 1.22]; n = 35), and 100 or more patients (0.94 [CI, 0.77 to 1.14]; n = 10). In trials with two or more such attributes, relative risk was further reduced.

Conclusions: Overall, no effect of albumin on mortality was detected; any such effect may therefore be small. This finding supports the safety of albumin. The influence of methodologic quality on relative risk for death suggests the need for further well-designed clinical trials.


For author affiliations, current addresses, and contributions, see end of text. See editorial comment on pp 205-208.
Case study 2: The Vioxx story

- On Sept 30, 2004, Merck announced the withdrawal of rofecoxib (Vioxx) because of an increased cardiovascular risk in patients taking the drug for >18 months.
- Decision was based on the 3-year results of the unpublished APPROVe study, a RTC of rofecoxib for the prevention of colorectal polyps in patients with a history of colorectal adenomas.
- By 2004, rofecoxib had been taken by ~80 million people (sales US$2.5 billion).
- Juni et al. did a meta-analysis of 18 RCTs and 11 observational studies.
- By the end of 2000 (52 events, 20742 patients) the relative risk from RTCs was 2.30 (95% CI 1.22–4.33, p=0.010), and 1 year later (64 events, 21432 patients) it was 2.24 (1.24–4.02, p=0.007).
- Juni et al. concluded that “rofecoxib should have been withdrawn several years earlier; the reasons why manufacturer and drug licensing authorities did not continuously monitor and summarise the accumulating evidence need to be clarified.”

Juni et al. Lancet 2004 Dec 4-10;364(9450):2021-9
Case study 2: The Vioxx story

Relative risk (95% CI) of myocardial infarction

- Ehrich et al (2001)\textsuperscript{19}
- Extension of Ehrich et al (2001)\textsuperscript{19}
- Cannon et al (2000)\textsuperscript{14}
- Day et al (2000)\textsuperscript{17}
- Hawkey et al (2000)\textsuperscript{15}
- Truitt et al (2001)\textsuperscript{21}
- Saag et al (2000 A)\textsuperscript{18}
- Kivitz et al (2004)\textsuperscript{22}
- Extension of Schnitzer et al (1999)\textsuperscript{24}
- Bombardier et al (2000)\textsuperscript{4}
- Geba et al (2001)\textsuperscript{20}
- Truitt et al (2001 A)\textsuperscript{25}
- Lisse et al (2003)\textsuperscript{23}
- Extension of Truitt et al (2001 A)\textsuperscript{25}
- Extension of Geusens et al (2002)\textsuperscript{26}
- Katz et al (2003)\textsuperscript{28}

Combined 2.24 (95% CI 1.24–4.02)

Juni et al. Lancet 2004 Dec 4-10;364(9450):2021-9
Case study 2: The Vioxx story

<table>
<thead>
<tr>
<th>Year</th>
<th>Patients</th>
<th>Events</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>523</td>
<td>1</td>
<td>0.916</td>
</tr>
<tr>
<td>1998</td>
<td>615</td>
<td>2</td>
<td>0.736</td>
</tr>
<tr>
<td></td>
<td>1399</td>
<td>5</td>
<td>0.828</td>
</tr>
<tr>
<td></td>
<td>2208</td>
<td>6</td>
<td>0.996</td>
</tr>
<tr>
<td></td>
<td>2983</td>
<td>8</td>
<td>0.649</td>
</tr>
<tr>
<td></td>
<td>3324</td>
<td>9</td>
<td>0.866</td>
</tr>
<tr>
<td>1999</td>
<td>4017</td>
<td>12</td>
<td>0.879</td>
</tr>
<tr>
<td></td>
<td>5059</td>
<td>13</td>
<td>0.881</td>
</tr>
<tr>
<td>2000</td>
<td>5393</td>
<td>16</td>
<td>0.855</td>
</tr>
<tr>
<td></td>
<td>13269</td>
<td>40</td>
<td>0.070</td>
</tr>
<tr>
<td></td>
<td>14247</td>
<td>44</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>15156</td>
<td>46</td>
<td>0.025</td>
</tr>
<tr>
<td></td>
<td>20742</td>
<td>52</td>
<td>0.010</td>
</tr>
<tr>
<td>2001</td>
<td>20742</td>
<td>58</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>20742</td>
<td>63</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>21432</td>
<td>64</td>
<td>0.007</td>
</tr>
</tbody>
</table>

Relative risk (95% CI) of myocardial infarction

Combined: 2.24
(95% CI 1.24-4.02)

Juni et al. Lancet 2004 Dec 4-10;364(9450):2021-9
Case study 3. Does male circumcision reduce risk of HIV?

International Journal of STD & AIDS 1999; 10: 8–16

REVIEW ARTICLE

Circumcision and HIV infection: review of the literature and meta-analysis

R S Van Howe MD FAAP

Department of Pediatrics, Marshfield Clinic, Lakeland Center, USA

Summary: Thirty-five articles and a number of abstracts have been published in the medical literature looking at the relationship between male circumcision and HIV infection. Study designs have included geographical analysis, studies of high-risk patients, partner studies and random population surveys. Most of the studies have been conducted in Africa. A meta-analysis was performed on the 29 published articles where data were available. When the raw data are combined, a man with a circumcised penis is at greater risk of acquiring and transmitting HIV than a man with a non-circumcised penis (odds ratio (OR)=1.06, 95% confidence interval (CI)=1.01–1.12). Based on the studies published to date, recommending routine circumcision as a prophylactic measure to prevent HIV infection in Africa, or elsewhere, is scientifically unfounded.
OR = 1.06

<table>
<thead>
<tr>
<th>Author</th>
<th>Intact+</th>
<th>Intact−</th>
<th>Circ+</th>
<th>Circ−</th>
<th>Odds ratio</th>
<th>95% CI</th>
<th>Relative risk</th>
<th>Attributable risk (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truck drivers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bwayo⁶</td>
<td>92</td>
<td>86</td>
<td>160</td>
<td>612</td>
<td>4.09</td>
<td>2.91-5.76</td>
<td>2.49</td>
<td>59.9</td>
</tr>
<tr>
<td>Bwayo⁷</td>
<td>22</td>
<td>46</td>
<td>37</td>
<td>200</td>
<td>2.59</td>
<td>1.39-4.79</td>
<td>2.07</td>
<td>51.7</td>
</tr>
<tr>
<td>STD clinics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kreiss²³</td>
<td>59</td>
<td>18</td>
<td>254</td>
<td>168</td>
<td>2.17</td>
<td>1.24-3.81</td>
<td>1.27</td>
<td>21.4</td>
</tr>
<tr>
<td>Hira⁸</td>
<td>418</td>
<td>172</td>
<td>10</td>
<td>10</td>
<td>2.43</td>
<td>0.99-5.94</td>
<td>1.42</td>
<td>29.4</td>
</tr>
<tr>
<td>Cameron¹¹</td>
<td>18</td>
<td>61</td>
<td>6</td>
<td>208</td>
<td>10.23</td>
<td>3.89-26.90</td>
<td>8.13</td>
<td>87.7</td>
</tr>
<tr>
<td>Pépin⁹¹¹</td>
<td>5</td>
<td>42</td>
<td>13</td>
<td>243</td>
<td>2.22</td>
<td>0.75-6.57</td>
<td>2.09</td>
<td>52.3</td>
</tr>
<tr>
<td>Greenblatt¹²</td>
<td>11</td>
<td>28</td>
<td>8</td>
<td>68</td>
<td>3.34</td>
<td>1.21-9.18</td>
<td>2.68</td>
<td>62.7</td>
</tr>
<tr>
<td>Diallo¹⁰</td>
<td>38</td>
<td>46</td>
<td>212</td>
<td>873</td>
<td>3.40</td>
<td>2.16-5.36</td>
<td>2.32</td>
<td>56.8</td>
</tr>
<tr>
<td>Simonsen¹³</td>
<td>17</td>
<td>70</td>
<td>21</td>
<td>232</td>
<td>2.68</td>
<td>1.34-5.37</td>
<td>2.35</td>
<td>57.5</td>
</tr>
<tr>
<td>Tyndall²</td>
<td>55</td>
<td>93</td>
<td>105</td>
<td>527</td>
<td>4.59</td>
<td>3.20-6.58</td>
<td>2.87</td>
<td>65.2</td>
</tr>
<tr>
<td>Nasio³⁸</td>
<td>86</td>
<td>78</td>
<td>137</td>
<td>580</td>
<td>4.67</td>
<td>3.26-6.68</td>
<td>2.74</td>
<td>63.6</td>
</tr>
<tr>
<td>Mehenda²²</td>
<td>837</td>
<td>3411</td>
<td>38</td>
<td>253</td>
<td>1.63</td>
<td>1.15-2.32</td>
<td>1.51</td>
<td>33.7</td>
</tr>
<tr>
<td>Bollinger²⁵</td>
<td>50</td>
<td>241</td>
<td>1</td>
<td>34</td>
<td>2.90</td>
<td>0.37-22.60</td>
<td>2.58</td>
<td>61.2</td>
</tr>
<tr>
<td>Chiasson²²</td>
<td>36</td>
<td>797</td>
<td>14</td>
<td>542</td>
<td>1.75</td>
<td>0.93-3.27</td>
<td>1.72</td>
<td>41.7</td>
</tr>
<tr>
<td>TB patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sassan²⁰</td>
<td>75</td>
<td>18</td>
<td>415</td>
<td>221</td>
<td>2.22</td>
<td>1.29-3.81</td>
<td>1.24</td>
<td>19.1</td>
</tr>
<tr>
<td>High-risk total</td>
<td>1849</td>
<td>5207</td>
<td>1431</td>
<td>4751</td>
<td>1.18</td>
<td>1.09-1.28</td>
<td>1.13</td>
<td>11.7</td>
</tr>
<tr>
<td>Partner studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hunter³¹</td>
<td>43</td>
<td>330</td>
<td>165</td>
<td>3765</td>
<td>2.97</td>
<td>2.09-4.24</td>
<td>2.75</td>
<td>63.6</td>
</tr>
<tr>
<td>Carael³³</td>
<td>90</td>
<td>105</td>
<td>34</td>
<td>45</td>
<td>1.13</td>
<td>0.67-1.92</td>
<td>1.07</td>
<td>6.8</td>
</tr>
<tr>
<td>Chao³⁴</td>
<td>442</td>
<td>4844</td>
<td>75</td>
<td>232</td>
<td>0.28</td>
<td>0.21-0.37</td>
<td>0.34</td>
<td>65.8†</td>
</tr>
<tr>
<td>Mose³²</td>
<td>24</td>
<td>16</td>
<td>12</td>
<td>17</td>
<td>2.13</td>
<td>0.80-5.62</td>
<td>1.45</td>
<td>31.0</td>
</tr>
<tr>
<td>Allen³⁸</td>
<td>275</td>
<td>612</td>
<td>132</td>
<td>324</td>
<td>1.10</td>
<td>0.86-1.41</td>
<td>1.07</td>
<td>6.8</td>
</tr>
<tr>
<td>Sedlin³⁸</td>
<td>32</td>
<td>26</td>
<td>33</td>
<td>40</td>
<td>1.49</td>
<td>0.75-2.98</td>
<td>1.22</td>
<td>18.1</td>
</tr>
<tr>
<td>Konde-Lucy⁴⁵</td>
<td>153</td>
<td>1516</td>
<td>6</td>
<td>127</td>
<td>2.13</td>
<td>0.93-4.93</td>
<td>2.03</td>
<td>50.8</td>
</tr>
<tr>
<td>Partner total</td>
<td>1059</td>
<td>7449</td>
<td>457</td>
<td>4550</td>
<td>1.42</td>
<td>1.26-1.59</td>
<td>1.36</td>
<td>26.7</td>
</tr>
<tr>
<td>Random population surveys</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barongo³⁷</td>
<td>55</td>
<td>1356</td>
<td>42</td>
<td>642</td>
<td>0.62</td>
<td>0.41-0.94</td>
<td>0.63</td>
<td>36.5†</td>
</tr>
<tr>
<td>Grosskurth⁸</td>
<td>158</td>
<td>4664</td>
<td>61</td>
<td>1026</td>
<td>0.58</td>
<td>0.43-0.78</td>
<td>0.59</td>
<td>40.9†</td>
</tr>
<tr>
<td>Van de Perre⁴⁰</td>
<td>46</td>
<td>224</td>
<td>6</td>
<td>26</td>
<td>0.89</td>
<td>0.35-2.28</td>
<td>0.91</td>
<td>9.1†</td>
</tr>
<tr>
<td>Seed⁹</td>
<td>171</td>
<td>422</td>
<td>51</td>
<td>192</td>
<td>1.52</td>
<td>1.07-2.17</td>
<td>1.37</td>
<td>27.2</td>
</tr>
<tr>
<td>Malamba⁴⁴</td>
<td>111</td>
<td>114</td>
<td>21</td>
<td>47</td>
<td>2.17</td>
<td>1.22-3.88</td>
<td>1.60</td>
<td>56.8</td>
</tr>
<tr>
<td>Quigley⁶⁶</td>
<td>101</td>
<td>272</td>
<td>48</td>
<td>121</td>
<td>0.94</td>
<td>0.62-1.40</td>
<td>0.98</td>
<td>4.7†</td>
</tr>
<tr>
<td>Urassa 1³⁸</td>
<td>56</td>
<td>1301</td>
<td>42</td>
<td>600</td>
<td>0.61</td>
<td>0.41-0.90</td>
<td>0.63</td>
<td>36.9†</td>
</tr>
<tr>
<td>Urassa 2</td>
<td>105</td>
<td>2040</td>
<td>32</td>
<td>426</td>
<td>0.69</td>
<td>0.46-1.03</td>
<td>0.70</td>
<td>29.9†</td>
</tr>
<tr>
<td>Urassa 3</td>
<td>38</td>
<td>309</td>
<td>19</td>
<td>158</td>
<td>1.02</td>
<td>0.57-1.83</td>
<td>1.02</td>
<td>2.0</td>
</tr>
<tr>
<td>Urassa 4</td>
<td>112</td>
<td>716</td>
<td>54</td>
<td>692</td>
<td>2.00</td>
<td>1.43-2.82</td>
<td>1.87</td>
<td>46.5</td>
</tr>
<tr>
<td>Urassa 5</td>
<td>101</td>
<td>365</td>
<td>48</td>
<td>136</td>
<td>0.78</td>
<td>0.53-1.16</td>
<td>0.83</td>
<td>16.9†</td>
</tr>
<tr>
<td>Random total</td>
<td>1054</td>
<td>11723</td>
<td>424</td>
<td>4066</td>
<td>0.86</td>
<td>0.77-0.97</td>
<td>0.87</td>
<td>12.6†</td>
</tr>
<tr>
<td>Total</td>
<td>3962</td>
<td>24379</td>
<td>2312</td>
<td>13367</td>
<td>0.94</td>
<td>0.89-0.99</td>
<td>0.95</td>
<td>5.2†</td>
</tr>
</tbody>
</table>

Case study 3. Does male circumcision reduce risk of HIV?

Circumcision in men and the prevention of HIV infection: a ‘meta-analysis’ revisited

Nigel O’Farrell MD MRCP¹ and Matthias Egger MD MFPHM²

¹Department of Genitourinary Medicine, Mile End Clinic, Bristol Royal Infirmary, Bristol, and ²MRC Health Services Research Collaboration, Department of Social Medicine, University of Bristol, Bristol, UK

Summary: There is debate on the role of male circumcision in HIV transmission. Most case-control and cohort studies from Africa have shown an association between a lack of circumcision and an increased risk of HIV infection in men. The evidence is conflicting, however, with cross-sectional surveys from Tanzania and Rwanda either showing no relationship or an association in the opposite direction. A recent review and meta-analysis of the literature concluded that the risk of HIV infection was lower in uncircumcised men (combined odds ratio 0.94, 95% confidence interval 0.89 to 0.99). However, the analysis was performed by simply pooling the data from 33 diverse studies, which is an inappropriate method for combining studies.

We re-analysed the data, stratifying by study, and found that an intact foreskin was associated with an increased risk of HIV infection: combined odds ratio 1.43 (1.32 to 1.54) with a fixed effect model and 1.67 (1.25 to 2.24) with a random effect model. There was significant between-study heterogeneity (P<0.0001) which was partly explained by stronger associations in studies in high-risk groups. The results from this re-analysis thus support the contention that male circumcision may offer protection against HIV infection, particularly in high-risk groups where genital ulcers and other STDs ‘drive’ the HIV epidemic. A systematic review is required to clarify this issue. Such a review should be based on an extensive search for relevant studies, published and unpublished, and should include a careful assessment of the design and methodological quality of studies. Much emphasis should be given to the exploration of possible sources of heterogeneity. In view of the continued high prevalence and incidence of HIV in many countries in sub-Saharan Africa, the question of whether circumcision could contribute to prevent infections is of great importance, and a sound systematic review of the available evidence should be performed without delay.
Case study 3. Does male circumcision reduce risk of HIV?

HIV and male circumcision—a systematic review with assessment of the quality of studies

N Siegfried, M Muller, J Deeks, J Volmink, M Egger, N Low, S Walker, and P Williamson

This Cochrane systematic review assesses the evidence for an interventional effect of male circumcision in preventing acquisition of HIV-1 and HIV-2 by men through heterosexual intercourse. The review includes a comprehensive assessment of the quality of all 37 included observational studies. Studies in high-risk populations consisted of four cohort studies, 12 cross-sectional studies, and three case-control studies; general population studies consisted of one cohort study, 16 cross-sectional studies, and one case-control study. There is evidence of methodological heterogeneity between studies, and statistical heterogeneity was highly significant for both general population cross-sectional studies ($\chi^2=132.34; \text{df}=15; p<0.00001$) and high-risk cross-sectional studies ($\chi^2=29.70; \text{df}=10; p=0.001$). Study quality was very variable and no studies measured the same set of potential confounding variables. Therefore, conducting a meta-analysis was inappropriate. Detailed quality assessment of observational studies can provide a useful visual aid to interpreting findings. Although most studies show an association between male circumcision and prevention of HIV, these results may be limited by confounding, which is unlikely to be adjusted for.

Lancet Infect Dis 2005; 5:165-73

NS and JV are at the South African Cochrane Centre, Medical Research Council, South Africa; NS is currently a Nuffield Medical Fellow at The University of Oxford, Oxford, UK; JV is also at the Primary Health Care Directorate, University of Cape Town, Cape Town, South Africa; MM is at the Institute for Maritime Technology, Simon’s Town, South Africa; JD is at the Centre for Statistics in Medicine, Institute of Health Sciences,
Randomized, Controlled Intervention Trial of Male Circumcision for Reduction of HIV Infection Risk: The ANRS 1265 Trial

Betten Buvuma, Willy, Dirck Tijou, Zimkakera Langa, Foluke Selvagrip-Tambusaro, Rimi Sita, Adrian Pare

Abstract

Background

Observational studies suggest that male circumcision may provide protection against HIV infection. A randomized, controlled intervention trial was conducted in a general population of South Africa to test this hypothesis.

Methods and Findings

A total of 3,234 normcircumcised men, aged 18-24 years, were randomized to a control or intervention group with follow-up visits at months 1, 3, 12, and 24. Male circumcision was offered to the intervention group immediately after randomization and the control group at the end of the follow-up. The group-censored data were analyzed in intention-to-treat, univariate, and multivariate analysis using poisson exponential proportional hazards models. Rate ratios (RRs) of HIV incidence were determined with 95% CI. Protection against HIV infection was calculated in 1-24 months. The rates were analyzed. There were 21 HIV infections (prevalence rate: 0.68 per 100 person-years) in the intervention group and 28 (1.01 per 100 person-years) in the control group, corresponding to an RR of 0.65 (95% CI: 0.24-2.08; p = 0.001). The Kaplan-Meier analysis showed a 63% (95% CI: 0.53-0.73) protection in the intervention group compared to the control group, and to the control group, respectively in the intervention group compared to the control group. Among men with behavioral factors, including sexual behavior that increased sharply in the intervention group, condom use, and health-seeking behavior, the protection was of 63% (95% CI: 0.53 - 0.73).

Conclusion

Male circumcision provides a degree of protection against acquiring HIV infection, equivalent to what a vaccine of high efficacy would have achieved. Male circumcision may provide an important tool in reducing the spread of HIV infection in sub-Saharan Africa. Preliminary and partial results were presented at the International AIDS Society 2003 Conference, on 26 July 2003, in Rio de Janeiro, Brazil.

Male circumcision for HIV prevention in men in Rakai, Uganda: a randomized controlled trial

Amadi W. Ged, Godfrey Gakula, David Szerk, Frederick Mugamba, Stephen Pfeffer, Fred Nakubere, Nadeen Gumbo, Lawrence M. Mbiti, Muhamed A. Ouoba, Michael J. Crow, Nicholas N. Onyekwere, Fred Wahhbe-Mahng, Mariana Calho, Carol M. Williams, Fous Qureshi, Steven A. Elion, Lunc Lengendë, Thomas Q. Culican, Robert W. Mbiti

Summary

Background Ecological and observational studies suggest that male circumcision reduces the risk of HIV acquisition in men. Our aim was to investigate the effect of male circumcision on HIV incidence in men.

Methods 496 uncircumcised, HIV-negative men aged 15-49 years who agreed to HIV testing and counselling were enrolled in this randomized trial in rural Rakai district, Uganda. Men were randomly assigned to receive immediate circumcision (n=247) or circumcision delayed for 24 months (n=249). HIV testing, physical examination, and interviews were repeated at 6, 12, and 24 month follow-up visits. The primary outcome was HIV incidence. Analyses were done on a modified intention-to-treat basis. This trial is registered with ClinicalTrials.gov, with the number NCT00415984.

Findings Baseline characteristics of the men in the intervention and control groups were very similar at enrolment. Retention rates were much the same in the two groups, with 99-92% of participants retained at all time points. In the modified intention-to-treat analysis, HIV incidence over 24 months was 0.66 cases per 100 person-years in the intervention group and 1.33 cases per 100 person-years in the control group (estimated efficacy of intervention 15%; 95% CI: 22.72-72; p=0.001). The averted efficacy was 55% (95% CI: 22.72-72; p=0.001) efficacy from the Kaplan-Meier estimates. Male circumcision reduced the risk of HIV infection by 63% (95% CI: 0.24-2.08; p=0.001). Incidence was lower in the intervention group than in the control group in all sociodemographic, behavioral, and sexually transmitted disease subgroups. Moderate or severe adverse events occurred in 34 (1.15%) circumcision all resolved with treatment. Behaviors were much the same in both groups during follow-up.

Interpretation Male circumcision reduced HIV incidence in men without behavioural disinhibition. Circumcision can be recommended for prevention in men.

Male circumcision for HIV prevention in young men in Kisu, Kenya: a randomised controlled trial

Robert C. Bentley, Stephen Hoare, Erathe Pake, Kusuma-Age, John Mackline, John N. King, Carolyn A. Williams, Richard T. Campbell, Jakes A. Mbugua-Akinbo

Summary

Background Male circumcision could provide substantial protection against acquisition of HIV-1 infection. Our aim was to determine whether male circumcision had a protective effect against HIV infection, and to assess safety and changes in sexual behavior related to this intervention.

Methods We did a randomised controlled trial of 2,784 men aged 18–36 years in Kisu, Kenya. Men were randomly assigned to an intervention group (circumcision: n=1391) or a control group (delayed circumcision, 1393), and assessed by HIV testing, medical examinations, and behavioral interviews during follow-ups at 1, 3, 6, 12, 18, and 24 months. HIV seroconversion was estimated in an intention-to-treat analysis. This trial is registered with ClinicalTrials.gov, with the number NCT00169971.

Findings The trial was stopped early on December 12, 2006, after a third interim analysis reviewed by the data and safety monitoring board. The median length of follow-up was 24.0 (IQR 0.4%) participants. 22 men in the intervention group and 47 in the control group had tested positive for HIV when the study was stopped. The 2-year HIV incidence was 2.1% (95% CI 1.2–3.1) in the circumcision group and 4.2% (3.8–5.5) in the control group (p=0.003). The crude protective effect of circumcision in men was 50% (95% CI 22–72). Adverse events related to the intervention (21 events in 5-9% of those circumcised) resolved quickly. No behavioral risk compensation after circumcision was observed.

Interpretation Male circumcision significantly reduces the risk of HIV-acquisition in young men in Africa. Where appropriate, voluntary, safe, and affordable circumcision services should be integrated with other HIV prevention interventions and provided as expeditiously as possible.
Meta-analysis of 3 RCTs shows strong, consistent effect

Table 1  Study characteristics and outcomes

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Setting</th>
<th>Population</th>
<th>n</th>
<th>Intervention</th>
<th>Control</th>
<th>Relative risk (95% confidence interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auvert et al. (2005)*</td>
<td>Randomized trial</td>
<td>Orange farm, South Africa</td>
<td>Males aged 18–24 years</td>
<td>3128</td>
<td>20/1546</td>
<td>49/1582</td>
<td>0.42 (0.25–0.70)</td>
</tr>
<tr>
<td>Bailey et al. (2007)**</td>
<td>Randomized trial</td>
<td>Kisumu, Kenya</td>
<td>Males aged 18–24 years</td>
<td>2780</td>
<td>19/1368</td>
<td>46/1392</td>
<td>0.41 (0.24–0.70)</td>
</tr>
<tr>
<td>Gray et al. (2007)</td>
<td>Randomized trial</td>
<td>Rakai district, Uganda</td>
<td>Males aged 15–49 years</td>
<td>4996</td>
<td>22/2474</td>
<td>45/2622</td>
<td>0.50 (0.30–0.83)</td>
</tr>
</tbody>
</table>
*3274 randomized, 3128 included in analysis.
**2784 randomized, 2780 included in analysis.

Study name  | Statistics for each study  | Risk ratio and 95% CI
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Auvert, RSA</td>
<td>0.42 0.25 0.70 0.001</td>
<td></td>
</tr>
<tr>
<td>Bailey, Kenya</td>
<td>0.41 0.24 0.70 0.001</td>
<td></td>
</tr>
<tr>
<td>Gray, Uganda</td>
<td>0.50 0.30 0.83 0.007</td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td>0.44 0.33 0.60 &lt;0.0001</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1 Random effects meta-analysis.

There is strong evidence that medical male circumcision reduces the acquisition of HIV by heterosexual men by between 38% and 66% over 24 months. Incidence of adverse events is very low, indicating that male circumcision, when conducted under these conditions, is a safe procedure. Inclusion of male circumcision into current HIV prevention measures guidelines is warranted, with further research required to assess the feasibility, desirability, and cost-effectiveness of implementing the procedure within local contexts.
Case study 3. Does male circumcision reduce risk of HIV?

WHO AND UNAIDS ANNOUNCE RECOMMENDATIONS FROM EXPERT MEETING ON MALE CIRCUMCISION FOR HIV PREVENTION

Paris, 28 March 2007 -- In response to the urgent need to reduce the number of new HIV infections globally, the World Health Organization (WHO) and the UNAIDS Secretariat convened an international expert consultation to determine whether male circumcision should be recommended for the prevention of HIV infection. Based on the evidence presented, which was considered to be compelling, experts attending the consultation recommended that male circumcision now be recognized as an additional important intervention to reduce the risk of heterosexually acquired HIV infection in men. The international consultation, which was held from 6-8 March 2007 in Montreux, Switzerland, was attended by participants representing a wide range of stakeholders, including governments, civil society, researchers, human rights and women’s health advocates, young people, funding agencies and implementing partners.

"The recommendations represent a significant step forward in HIV prevention," said Dr Kevin De Cock, Director, HIV/AIDS Department, World Health Organization. "Countries with high rates of heterosexual HIV infection and low rates of male circumcision now have an additional intervention which can reduce the risk of HIV infection in heterosexual men. Scaling up male circumcision in such countries will result in immediate benefit to individuals. However, it will be a number of years before we can expect to see an impact on the epidemic from such investment."

There is now strong evidence from three randomized controlled trials undertaken in Kisumu, Kenya, Rakai District, Uganda and Orange Farm, South Africa that male circumcision reduces the risk of heterosexually acquired HIV infection in men by approximately 60%. This evidence supports the findings of numerous observational studies that have also suggested that the geographical correlation long described between lower HIV prevalence and high rates of male circumcision in some countries in Africa, and more recently elsewhere, is, at least in part, a causal association. Currently, an estimated 665 million men, or 30% of men worldwide, are estimated to be uncircumcised.
These were just a few case studies

- But to practice medicine and public health, we need evidence on a large number of issues...
As a teacher and a mother, Victoria found herself catching colds all the time. In her spare time, Victoria took to her kitchen to wage war on the common cold. Within six months she had created the prototype for Airborne, her all-natural cold fighter. Her friends and family started using it and Victoria says no one was getting sick. So she and her husband set up shop in their home and began to market Airborne. The accounting office was in the dining room, one of the bedrooms was the marketing office and the bathroom was shipping and receiving. The orders started pouring in and in the first year, Victoria made $25,000— the same as her teaching salary.
Snake Oil?
Scientific evidence for popular health supplements
showing tangible human health benefits when taken orally by an adult with a healthy diet.

- Folic acid
- Fish oil / omega 3
- Garlic
- Dark chocolate
- Calcium + vit. D
- Olive leaf extract
- Omega 6
- Omega 3
- Zinc
- Vitamin D
- Green tea
- St John's wort
- Vitamin K2
- Ginger
- Magnesium + vitamin B6
- Honey
- Probiotics
- Iron
- Coconut oil
- Whole grains
- Antioxidants
- Lavender
- L-arginine
- L-carnitine
- Glucosamine
- Vitamin C
- Vitamin D
- Beta-carotene
- Retinol
- Copper
- Vitamin E
- Resveratrol

EVIDENCE
GOOD
POPULARITY
ONE TO WATCH
(see text box for promising results)

Source: PubMed, Cochrane.org
Large human placebo-controlled trials only.
Research: Miriam Quick
Information@beautiful.net / andyp@kins.org
See Data
How can we separate the good from all the bad science?

http://www.badscience.net/
What is evidence-based medicine?

The practice of EBM is the integration of
• individual clinical expertise
  with the
• best available external clinical evidence from systematic research and
• patient’s values and expectations

http://www.cebm.net/index.asp
What is evidence-based public health?

Evidence-based public health is “the development, implementation, and evaluation of effective programs and policies in public health through application of principles of scientific reasoning including systematic uses of data and program planning models.”

The importance of research synthesis

- We need evidence for both clinical practice and for public health decision making
- Where does evidence come from?
  - An good review is a state-of-the-art synthesis of current evidence on a given research question
  - Given the explosion of medical literature, and the fact that time is always scarce, review articles play a big role in decision-making
The importance of research synthesis

Managing evidence-based knowledge: the need for reliable, relevant and readable resources

Sharon Straus MD MSc, R. Bryan Haynes MD PhD

The sheer volume of research-based evidence is one of the main barriers to better use of knowledge. About 10 years ago, if general internists wanted to keep abreast of the primary clinical literature, they would have needed to read 17 articles daily. Today, with more than 1000 articles indexed daily by MEDLINE, that figure is likely double. The problem is compounded by the inability of clinicians to afford more than a few seconds at a time in their practices for finding and assimilating evidence. These challenges highlight the need for better infrastructure in the management of evidence-based knowledge.
Seventy-Five Trials and Eleven Systematic Reviews a Day: How Will We Ever Keep Up?

Hilda Bastian¹*, Paul Glasziou², Iain Chalmers³

¹ German Institute for Quality and Efficiency in Health Care (IQWIG), Cologne, Germany, ² Centre for Research in Evidence-Based Practice, Faculty of Health Sciences, Bond University, Gold Coast, Australia, ³ James Lind Library, James Lind Initiative, Oxford, United Kingdom

Thirty years ago, and a quarter of a century after randomised trials had become widely accepted, Archie Cochrane reproached the medical profession for not having managed to organise a “critical summary, by speciality or subspeciality, adapted periodically, of all relevant randomised controlled trials” [1]. Thirty years after Cochrane’s reproach we feel it is timely to consider the extent to which health professionals, the public and policymakers could now use “critical summaries” of trials for their decision-making.

Summary Points

- When Archie Cochrane reproached the medical profession for not having critical summaries of all randomised controlled trials, about 14 reports of trials were being published per day. There are now 75 trials, and 11 systematic reviews of trials, per day and a plateau in growth has not yet been reached.
- Although trials, reviews, and health technology assessments have undoubtedly had major impacts, the staple of medical literature synthesis remains the non-systematic narrative review. Only a small minority of trial reports are being analysed in up-to-date systematic reviews. Given the constraints, Archie Cochrane’s vision will not be achieved without some serious changes in course.
- To meet the needs of patients, clinicians, and policymakers, unnecessary trials need to be reduced, and systematic reviews need to be prioritised. Streamlining and innovation in methods of systematic reviewing are necessary to enable valid answers to be found for most patient questions. Finally, clinicians and patients require open access to these important resources.

The Landscape

Keeping up with information in health care has never been easy. Even in 1753,
The Ascent of Evidence (and the exhaustion of Man)
The importance of research synthesis

• Given that most clinicians and public health professionals do not have the time to track down all the original articles, critically read them, and obtain the evidence they need for their questions,
  ▫ Systematic reviews and clinical practice guidelines may be their best source of evidence
    • Several “pre-digested” sources of evidence are currently available
    • The EBM movement is heavily dependent on these pre-appraised evidence sources
# Hierarchy of evidence

<table>
<thead>
<tr>
<th>Grade of Recommendation</th>
<th>Level of Evidence</th>
<th>Therapy/Prevention, Aetiology/Harm</th>
<th>Prognosis</th>
<th>Diagnosis</th>
<th>Economic analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1a</td>
<td>SR (with homogeneity*) of RCTs</td>
<td>SR (with homogeneity*) of inception cohort studies; or a CPG validated on a test set.</td>
<td>SR (with homogeneity*) of Level 1 diagnostic studies; or a CPG validated on a test set.</td>
<td>SR (with homogeneity*) of Level 1 economic studies</td>
</tr>
<tr>
<td></td>
<td>1b</td>
<td>Individual RCT (with narrow Confidence Interval)</td>
<td>Individual inception cohort study with ≥ 80% follow-up</td>
<td>Independent blind comparison of an appropriate spectrum of consecutive patients, all of whom have undergone both the diagnostic test and the reference standard.</td>
<td>Analysis comparing all (critically-validated) alternative outcomes against appropriate cost measurement, and including a sensitivity analysis incorporating clinically sensible variations in important variables.</td>
</tr>
<tr>
<td></td>
<td>1c</td>
<td>All or none⁵</td>
<td>All or none case-series</td>
<td>Absolute SpPins and SnNouts¹¹</td>
<td>Clearly as good or better, but cheaper. Clearly as bad or worse but more expensive. Clearly better or worse at the same cost.</td>
</tr>
<tr>
<td>B</td>
<td>2a</td>
<td>SR (with homogeneity*) of cohort studies</td>
<td>SR (with homogeneity*) of either retrospective cohort studies or untreated control groups in RCTs.</td>
<td>SR (with homogeneity*) of Level ≥2 diagnostic studies</td>
<td>SR (with homogeneity*) of Level ≥2 economic studies</td>
</tr>
<tr>
<td></td>
<td>2b</td>
<td>Individual cohort study (including low quality RCT; e.g., &lt;80% follow-up)</td>
<td>Retrospective cohort study or follow-up of untreated control patients in an RCT; or CPG not validated in a test set.</td>
<td>Independent blind comparison but either in non-consecutive patients, or confined to a narrow spectrum of study individuals (or both), all of whom have undergone both the diagnostic test and the reference standard; or a diagnostic CPG not validated in a test set.</td>
<td>Analysis comparing a limited number of alternative outcomes against appropriate cost measurement, and including a sensitivity analysis incorporating clinically sensible variations in important variables.</td>
</tr>
<tr>
<td></td>
<td>2c</td>
<td>&quot;Outcomes&quot; Research</td>
<td>&quot;Outcomes&quot; Research</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>3a</td>
<td>SR (with homogeneity*) of case-control studies</td>
<td></td>
<td>Independent blind comparison of an appropriate spectrum, but the reference standard was not applied to all study patients</td>
<td>Analysis without accurate cost measurement, but including a sensitivity analysis incorporating clinically sensible variations in important variables.</td>
</tr>
<tr>
<td></td>
<td>3b</td>
<td>Individual Case-Control Study</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>4</td>
<td>Case-series (and poor quality cohort and case-control studies)⁶⁶</td>
<td>Case-series (and poor quality prognostic cohort studies)⁷⁷</td>
<td>Reference standard was not applied independently or not applied blindly</td>
<td>Analysis with no sensitivity analysis</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Expert opinion without explicit critical appraisal, or based on physiology, bench research or &quot;first principles&quot;</td>
<td>Expert opinion without explicit critical appraisal, or based on physiology, bench research or &quot;first principles&quot;</td>
<td>Expert opinion without explicit critical appraisal, or based on economic theory</td>
<td>Expert opinion without explicit critical appraisal, or based on economic theory</td>
</tr>
</tbody>
</table>

---

⁵⁶ SR with homogeneity* of the same cohort studies; or a CPG validated on a test set.⁷⁷ SR with homogeneity* of the same case-control studies, or a CPG validated on a test set.
Guidelines and recommendations: GRADE

GRADE: an emerging consensus on rating quality of evidence and strength of recommendations

Guidelines are inconsistent in how they rate the quality of evidence and the strength of recommendations. This article explores the advantages of the GRADE system, which is increasingly being adopted by organisations worldwide.
Systematic reviews are used to judge quality of evidence

Box 2 | Quality of evidence and definitions

**High quality**— Further research is very unlikely to change our confidence in the estimate of effect

**Moderate quality**— Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate

**Low quality**— Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate

**Very low quality**— Any estimate of effect is very uncertain

Factors in deciding on quality of evidence

Factors that might decrease quality of evidence
- Study limitations
- Inconsistency of results
- Indirectness of evidence
- Imprecision
- Publication bias
- Factors that might increase quality of evidence
- Large magnitude of effect
- Plausible confounding, which would reduce a demonstrated effect
- Dose-response gradient

http://www.gradeworkinggroup.org/
But evidence is just one component that determines strength of recommendations

<table>
<thead>
<tr>
<th>Factor</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balance between desirable and undesirable effects</td>
<td>The larger the difference between the desirable and undesirable effects, the higher the likelihood that a strong recommendation is warranted. The narrower the gradient, the higher the likelihood that a weak recommendation is warranted</td>
</tr>
<tr>
<td>Quality of evidence</td>
<td>The higher the quality of evidence, the higher the likelihood that a strong recommendation is warranted</td>
</tr>
<tr>
<td>Values and preferences</td>
<td>The more values and preferences vary, or the greater the uncertainty in values and preferences, the higher the likelihood that a weak recommendation is warranted</td>
</tr>
<tr>
<td>Costs (resource allocation)</td>
<td>The higher the costs of an intervention—that is, the greater the resources consumed—the lower the likelihood that a strong recommendation is warranted</td>
</tr>
</tbody>
</table>

http://www.gradeworkinggroup.org/
Doing New Research? Don’t Forget the Old

Nobody should do a trial without reviewing what is known

Mike Clarke

On May 2, 1898, George Gould used his address to the founding meeting of the Association of Medical Librarians in Philadelphia to present a vision of the future of health information. ‘I look forward,’ he said, ‘to such an organisation of the literary records of medicine that a puzzled worker in any part of the civilised world shall in an hour be able to gain a knowledge pertaining to a subject of the experience of every other man in the world’ [1]. Has his vision been realised?

Good quality, but some of it is not. Thus, anyone wishing to use the health literature to make well-informed decisions must both identify the relevant research from amidst this vast amount of information and then appraise it. This is an impossible task for many. Even though making access to the literature easier and cheaper will increase the ability of people to find research, it will also reveal just how much information there is out there and how daunting is the task of making sense of it.

with one or more search engines?
Almost certainly, as the speed of the search increased through these four


Copyright: © 2004

Box 1. Practical Suggestions for Researchers

- Conduct a systematic review of your research question before embarking on a new study, or identify a relevant review done by someone else.
- Design your study to take account of the relevant successes and failures of the prior studies, and of the evidence within them.
- Discuss the findings of your study in the context of an updated systematic review of relevant research.
- Publish the systematic review within, alongside, or shortly after the report of your study.
- Provide information from your study to others doing systematic reviews of similar topics.
Prof Archibald Cochrane, CBE (1909 - 1988)

- The Cochrane Collaboration is named in honour of Archie Cochrane, a British researcher.
- In 1979 he wrote, "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 randomized controlled trials”

Source: http://www.cochrane.org/cochrane/archieco.htm
The Cochrane Collaboration

- Archie Cochrane’s challenge led to the establishment during the 1980s of an international collaboration to develop the Oxford Database of Perinatal Trials.

Source: http://www.cochrane.org/cochrane/archieco.htm
COCHRANE DATABASE OF SYSTEMATIC REVIEWS

Issue 4 of 12, Apr 2013 | Contents

BROWSE BY TOPICS

- Anaesthesia & pain control (834)
- Blood disorders (417)
- Cancer (1448)
- Child health (3746)
- Complementary & alternative medicine (558)
- Consumer & communication strategies (182)
- Dentistry & oral health (282)
- Developmental, psychosocial, & learning problems (232)
- Ear, nose, & throat (298)

Other Browse Options

EXPAND

SPECIAL COLLECTIONS

- World day for Safety and Health at Work 2013
- Preventing falls and fall-related injuries in older people
- Tuberculosis
- Cochrane Evidence Aid: resources for earthquakes

EDITORIALS

- Assessing risk of bias in randomised clinical trials included in Cochrane Reviews: the why is easy, the how is a challenge
  - Åberg Hjortdorffson, Isabelle
  - Broun, Lucy Turner, Douglas G
  - Altman & David Moher

- It's time for AllTrials registered and reported
  - Tracey Brown

- Prevention of occupational diseases: implementing the evidence
  - Jos Verbeek, Thais Morata, Jani Ruotsalainen & Harri Vainio

HIGHLIGHTED NEW AND UPDATED COCHRANE REVIEWS

- School-based programmes for preventing smoking
- Interventions to improve antibiotic prescribing practices for hospital inpatients
- Vaccines for measles, mumps and rubella in children
- Swimming training for asthma in children and adolescents aged 18 years and under
- Multimedia educational interventions for consumers about prescribed and over-the-counter medications
- Training to recognise the early signs of recurrence in schizophrenia
- Computer-based diabetes self-management interventions for adults with type 2 diabetes mellitus
2500 SRs per year, of which about 20% are Cochrane reviews (estimate by Moher et al. PLoS Med 2007)
Secondary journals
Are textbooks good sources of current evidence?

• Not always!
  ▫ They are better for background questions than foreground questions
  ▫ They are not updated frequently and often lag behind current evidence by many years

• Exceptions:
  ▫ ACP Medicine [Scientific American Medicine]
  ▫ UpToDate
  ▫ Clinical Evidence
  ▫ Harrison’s Online
  ▫ Emedicine (totally online text)
Evidence vs. textbook recommendations

A. Thrombolytic Therapy

<table>
<thead>
<tr>
<th>Year</th>
<th>Cumulative RCTs</th>
<th>Odds Ratio (Log Scale)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>1965</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>1970</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>1975</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>1980</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>1985</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td></td>
<td>43</td>
<td></td>
</tr>
<tr>
<td></td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>70</td>
<td></td>
</tr>
</tbody>
</table>

P<.01

Favors Treatment  Favors Control

Textbook/Review Recommendations

<table>
<thead>
<tr>
<th>Routine</th>
<th>Specific</th>
<th>Rare/Never</th>
<th>Experimental</th>
<th>Not Mentioned</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6</td>
</tr>
</tbody>
</table>

Antman et al. JAMA 1992
Are textbooks good sources of current evidence?
Systematic reviews are done in different domains

Tuberculosis among Health-Care Workers in Low- and Middle-Income Countries: A Systematic Review

Rajnish Joshi¹,², Arthur L. Reingold¹, Dick Menzies³, Madhukar Pai³*

¹ Division of Epidemiology, School of Public Health, University of California Berkeley, Berkeley, California, United States of America, ² Department of Medicine, Mahatma Gandhi Institute of Medical Sciences, Sevagram, Maharashtra, India, ³ Montreal Chest Institute, McGill University, Montreal, Canada

Meta-analysis of “rates”
Systematic reviews are done in different domains

**Review**

**Annals of Internal Medicine**

**Accuracy of Rapid Influenza Diagnostic Tests**

A Meta-analysis

Caroline Chartrand, MD, MSc; Mariska M.G. Leeﬂang, DVM, PhD; Jessica Minion, MD, MSc; Timothy Brewer, MD, MPH; and Madhukar Pai, MD, PhD

**Microscopic-observation drug susceptibility and thin layer agar assays for the detection of drug resistant tuberculosis: a systematic review and meta-analysis**

Jessica Minion, Erika Leung, Dick Menzies, Madhukar Pai

*Lancet Infect Dis 2010*

Meta-analysis of “diagnostic accuracy [diagnosis]”
Systematic reviews are done in different domains.
Systematic reviews are done in different domains

**REVIEW ARTICLE**

**Risk of Tuberculosis From Exposure to Tobacco Smoke**

*A Systematic Review and Meta-analysis*

Michael N. Bates, PhD; Asheena Khalakdina, PhD; Madhukar Pai, MD, PhD; Lisa Chang, MPH; Fernanda Lessa, MD, MPH; Kirk R. Smith, PhD

*Arch Intern Med.* 2007;167:335-342

Meta-analysis of “observational studies [etiology]”
Are these the same or different?

- Traditional, narrative review
- Systematic review
- Overview
- Meta-analysis
- Pooled analysis
All reviews
(also called overviews)

Types of review articles

Meta-analyses

Systematic reviews

Individual patient data (IPD) meta-analyses

Reviews that are not systematic (traditional, narrative reviews)

In practice, not all meta-analyses are conducted as part of systematic reviews.
Some definitions

• Traditional, narrative reviews, usually written by experts in the field, are qualitative, narrative summaries of evidence on a given topic. Typically, they involve informal and subjective methods to collect and interpret information.

• “A systematic review is a review in which there is a comprehensive search for relevant studies on a specific topic, and those identified are then appraised and synthesized according to a predetermined and explicit method.”*

## Narrative vs. Systematic Reviews

<table>
<thead>
<tr>
<th>Components of a review</th>
<th>Traditional, narrative reviews</th>
<th>Systematic reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formulation of the question</td>
<td>Usually address broad questions</td>
<td>Usually address focused questions</td>
</tr>
<tr>
<td>Methods section</td>
<td>Usually not present, or not well-described</td>
<td>Clearly described with pre-stated criteria about participants, interventions and outcomes</td>
</tr>
<tr>
<td>Search strategy to identify studies</td>
<td>Usually not described; mostly limited by reviewers’ abilities to retrieve relevant studies; usually not reproducible and prone to selective citation</td>
<td>Clearly described and usually exhaustive; transparent, reproducible and less prone to selective citation</td>
</tr>
<tr>
<td>Quality assessment of identified studies</td>
<td>Usually all identified studies are included without explicit quality assessment</td>
<td>Only high-quality studies are included using pre-stated criteria; if lower-quality studies included, the effects of this are tested in subgroup analyses</td>
</tr>
<tr>
<td>Data extraction</td>
<td>Methods usually not described</td>
<td>Usually undertaken by more than one reviewer onto pre-tested data forms; attempts often made to obtain missing data from authors of primary studies</td>
</tr>
<tr>
<td>Data synthesis</td>
<td>Qualitative description employing the ‘vote counting’ approach, where each included study is given equal weight, irrespective of study size and quality</td>
<td>Meta-analysis assigns higher weights to effect measures from more precise studies; pooled, weighted effect measures with confidence limits provide power and precision to results</td>
</tr>
<tr>
<td>Heterogeneity</td>
<td>Usually dealt with in a narrative fashion</td>
<td>Heterogeneity dealt with by graphical and statistical methods; attempts are often made to identify sources of heterogeneity</td>
</tr>
<tr>
<td>Interpreting results</td>
<td>Prone to cumulative systematic biases and personal opinion</td>
<td>Less prone to systematic biases and personal opinion</td>
</tr>
</tbody>
</table>
Some definitions

• “A meta-analysis is the statistical combination of at least 2 studies to produce a single estimate of the effect of the healthcare intervention under consideration.”*

• Individual patient data meta-analyses (pooled analyses) involves obtaining raw data on all patients from each of the trials directly and then re-analyzing them.

Multidrug Resistant Pulmonary Tuberculosis Treatment Regimens and Patient Outcomes: An Individual Patient Data Meta-analysis of 9,153 Patients


Abstract

Background: Treatment of multidrug resistant tuberculosis (MDR-TB) is lengthy, toxic, expensive, and has generally poor outcomes. We undertook an individual patient data meta-analysis to assess the impact on outcomes of the type, number, and duration of drugs used to treat MDR-TB.

Methods and Findings: Three recent systematic reviews were used to identify studies reporting treatment outcomes of microbiologically confirmed MDR-TB. Study authors were contacted to solicit individual patient data including clinical characteristics, treatment given, and outcomes. Random effects multivariable logistic meta-regression was used to estimate adjusted odds of treatment success. Adequate treatment and outcome data were provided for 9,153 patients with MDR-TB from 32 observational studies. Treatment success, compared to failure/relapse, was associated with use of: later generation quinolones, (adjusted odds ratio [aOR]: 2.5 [95% CI 1.1–6.0]), ofloxacin (aOR: 2.5 [1.6–3.9]), ethionamide or prothionamide (aOR: 1.7 [1.3–2.3]), or use of four or more likely effective drugs in the initial intensive phase (aOR: 2.3 [1.3–3.9]), and three or more likely effective drugs in the continuation phase (aOR: 2.7 [1.7–4.1]). Similar results were seen for the association of treatment success compared to failure/relapse or death: later generation quinolones, (aOR: 2.7 [1.7–4.3]), ofloxacin (aOR: 2.3 [1.3–3.8]), ethionamide or prothionamide (aOR: 1.7 [1.4–2.1]), or use of four or more likely effective drugs in the initial intensive phase (aOR: 2.7 [1.9–3.9]), and three or more likely effective drugs in the continuation phase (aOR: 4.5 [3.4–6.0]).

Conclusions: In this individual patient data meta-analysis of observational data, improved MDR-TB treatment success and survival was associated with use of certain fluoroquinolones, ethionamide, or prothionamide, and greater total number of effective drugs. However, randomized trials are urgently needed to optimize MDR-TB treatment.
Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses

John P.A. Ioannidis MD

Meta-analysis is an important research design for appraising evidence and guiding medical practice and health policy.¹ Meta-analyses draw strength from combining data from many studies. However, even if perfectly done with perfect data, a single meta-analysis that addresses 1 treatment comparison for 1 outcome may offer a short-sighted view of the evidence. This may suffice for decision-making if there is only 1 treatment choice for this condition and only 1 outcome of interest and research results are perfect. However, usually there are many treatments to choose from, many outcomes to consider and research is imperfect. For example, there are 68 antidepressant drugs to choose from,² dozens of scales to measure depression outcomes, and biases abound in research about antidepressants.³,⁴

Key points

- Single meta-analysis of a treatment comparison for a single outcome offers a limited view if there are many treatments or many important outcomes to consider.
- Umbrella reviews assemble together several systematic reviews on the same condition.
- Treatment networks quantitatively analyze data for all treatment comparisons on the same disease.
- Multiple treatments meta-analysis can rank the effectiveness of many treatments in a network.
- Integration of evidence from multiple meta-analyses may be extended across many diseases.
Elements of a Systematic Review

1. Formulate the review question & write a protocol
2. Search for and include primary studies
3. Assess study quality
4. Extract and analyze data
5. Interpret results & write a report


Systematic reviews to support evidence-based medicine, 2nd edition. Khan K et al.
Road map for systematic reviews

A “road map” for systematic reviews of diagnostic test evaluations

1. Define a focused diagnostic review question: (Patient/Disease, Index test, Reference standard, and Outcomes)

2. Search databases: PubMed, EMBASE, BIOSIS, Web of Science, Cochrane CENTRAL, MEDION, and subject-specific databases: Contact authors, experts, companies: citation tracking

3. Identify appropriate databases and sources of diagnostic studies

4. Run searches on all relevant databases and sources

5. Save all citations (titles/abstracts) in a reference manager. Document search strategies that were employed. These citations are ready for first screen (N1)

6. Reviewer 1 screens all titles/abstracts and makes selections for second screen

7. Reviewers meet and resolve disagreements on citations they do not agree on. The final number (N2) selected after this process is ready for second screen (review of full text articles)

8. Get full texts of all articles identified for second screen (N2)

9. Articles considered eligible after full-text review (by 2 reviewers) is the final set of studies for inclusion (n2)

10. Studies included in the final analysis (n2) each article goes a unique ID number

11. Excluded after second screen

12. Keep a log of excluded studies with reasons for exclusion

13. Paper data extraction forms (after pilot test)

14. Reviewer 1 extracts data (including quality assessment) from the final selected articles

15. Reviewers meet and resolve disagreements on data. Compute inter-rater reliability (eg, kappa statistic). The final data after this process are ready for data entry

16. Enter data into database manager software

17. Import data and analyze using software: SROC (Summary Receiver Operating Characteristic) Forest and ROC plots of SE and SP. Look for correlation between TPR and FPR. Search for threshold effect. Perform SROC analyses. Pool measures like LR and DOR only if appropriate. Search for heterogeneity, and reasons for heterogeneity. Consider subgroup and sensitivity analyses

18. Interpret, discuss results, and write the report. Discuss applicability of results, and limitations of the review. Make recommendations for practice or policy, and research

Software suggestions: EndNote, Reference Manager, ProCite

A good SR is a lot of work! especially if you follow the best practice standards...
Finding What Works in Health Care
Standards for Systematic Reviews

These standards are for systematic reviews of comparative effectiveness research of therapeutic medical or surgical interventions

http://books.nap.edu/openbook.php?record_id=13059
All systematic reviews are not meta-analyses!

• “...it is always appropriate and desirable to systematically review a body of data, but it may sometimes be inappropriate, or even misleading, to statistically pool results from separate studies. Indeed, it is our impression that reviewers often find it hard to resist the temptation of combining studies even when such meta-analysis is questionable or clearly inappropriate.”
All reviews are not systematic!

- In 1987, Cynthia Mulrow published an interesting article entitled “The Medical Review Article: State of the Science.”
- She examined 50 review articles published in 4 major general medical journals [Annals of Internal Med; Archives of Internal Med; JAMA; New Engl J Med]
- Findings:
  - 80% addressed a focused review question
  - 2% described the method of locating evidence
  - 2% used explicit criteria for selecting studies for inclusion
  - 2% assessed the quality of the primary studies
  - 6% performed a quantitative analysis

All systematic reviews are not systematic!

Many Reviews Are Systematic but Some Are More Transparent and Completely Reported than Others

The PLoS Medicine Editors

Epidemiology and Reporting Characteristics of Systematic Reviews

David Moher¹,²,³*, Jennifer Tetzlaff¹, Andrea C. Tricco¹,⁴, Margaret Sampson¹, Douglas G. Altman⁵

¹ Chalmers Research Group, Children’s Hospital of Eastern Ontario Research Institute, Ottawa, Canada, ² Department of Paediatrics, Faculty of Medicine, University of Ottawa, Ottawa, Canada, ³ Department of Epidemiology and Community Medicine, Faculty of Medicine, University of Ottawa, Ottawa, Canada, ⁴ Institute of Population Health, University of Ottawa, Ottawa, Canada, ⁵ Centre for Statistics in Medicine, Oxford, United Kingdom
All systematic reviews are not systematic!

- 300 SRs were identified (one month)
- Majority (272 [90.7%]) reported in specialty journals
- Most reviews (213 [71.0%]) were categorized as therapeutic, and included a median of 16 studies
- Reviews typically searched a median of three electronic databases and two other sources
- Most (197/295 [66.8%]) reviews reported information about quality assessment, while few (68/294 [23.1%]) reported assessing for publication bias.
- A little over half (161/300 [53.7%]) reported combining their results statistically, of which most (147/161 [91.3%]) assessed for consistency across studies.
- There were large differences between Cochrane reviews and non-Cochrane reviews in the quality of reporting

When can meta-analyses mislead?

- When a meta-analysis is done outside of a systematic review
- When poor quality studies are included or when quality issues are ignored
- When small and inconclusive studies are included
- When inadequate attention is given to heterogeneity ▫ Indiscriminate data aggregation can lead to inaccurate conclusions
- When reporting biases are a problem ▫ Publication bias ▫ Time lag bias ▫ Duplicate publication bias ▫ Language bias ▫ Outcome reporting bias

Publication of NIH funded trials registered in ClinicalTrials.gov: cross sectional analysis

Joseph S Ross assistant professor of medicine, Tony Tse program analyst at ClinicalTrials.gov, Deborah A Zarin director of ClinicalTrials.gov, Hui Xu postgraduate house staff trainee, Lei Zhou postgraduate house staff trainee, Harlan M Krumholz Harold H Hines Jr professor of medicine and professor of investigative medicine and of public health.

Abstract

Objective To review patterns of publication of clinical trials funded by US National Institutes of Health (NIH) in peer reviewed biomedical journals indexed by Medline.

Design Cross sectional analysis.

Setting Clinical trials funded by NIH and registered within ClinicalTrials.gov (clinicaltrials.gov), a trial registry and results database maintained by the US National Library of Medicine, after 30 September 2005 and updated as having been completed by 31 December 2007, allowing at least 30 months for publication after completion of the trial.

Main outcome measures Publication and time to publication in the biomedical literature, as determined through Medline searches, the last of which was performed in June 2011.

Results Among 635 clinical trials completed by 31 December 2008, 284 (46%) were published in a peer reviewed biomedical journal, indexed by Medline, within 30 months of trial completion. The median period of follow-up after trial completion was 51 months (25th-75th centiles 40-69 months), and 432 (68%) were published overall. Among published trials, the median time to publication was 23 months (14-36 months). Trials completed in either 2007 or 2008 were more likely to be published within 30 months of study completion compared with trials completed before 2007 (64% (196/306) v 56% (96/179); P=0.001).

Conclusions Despite recent improvement in timely publication, fewer than half of trials funded by NIH are published in a peer reviewed biomedical journal indexed by Medline within 30 months of trial completion. Moreover, after a median of 51 months after trial completion, a third of trials remained unpublished.

Introduction

Today, there is an increasing emphasis on the successful translation of results from research into practice. This requires the timely dissemination of findings. While research results might be submitted directly to regulatory agencies, such as the Food and Drug Administration (FDA), physicians and policy makers generally depend on peer reviewed publications to learn about findings from clinical trials. Extensive research has shown, however, that the results of studies are often not shared publicly in a timely way and that between 25% and 50% of clinical trials remain unpublished even several years after completion, although this work was largely focused on industry funded studies. There are many possible reasons behind the delayed or non-publication of results from clinical trials, including lack of incentive to disseminate negative or unsupportive findings, time constraints, limited resources, changing interests, or even failure to have an article accepted by a journal.

Understanding the patterns of publication of research findings among publicly funded research, as opposed to industry funded research, is important because of the funding and the expectation for public benefit. Within the United States, the National Institutes of Health (NIH) is the leading and largest government agency responsible for biomedical and health related research and invests more than $12bn (about £7600m or €8900m) of public resources in funding research in people or in clinical research. $3.5bn explicitly on clinical trials. These costs do not include the considerable contributions and costs incurred by the participants in the research. Previous work suggests that
While almost all trials with “positive” results on antidepressants had been published, trials with “negative” results submitted to the US Food and Drug Administration, with few exceptions, remained either unpublished or were published with the results presented so that they would appear “positive”
EDITORS CHOICE

Goodbye PubMed, hello raw data

The raw data from trials must be made freely available. Journals clearly have a role to play in making this happen.

This time last year the H1N1 influenza pandemic was burning itself out, having caused, thankfully, far less sickness and death than predicted. Now this year’s seasonal flu epidemic is doing its rounds in the northern hemisphere (p. 134). The UK’s problems with uptake and availability of the flu vaccine seem to have been sorted out, but what interests me is this year’s low key approach to antivirals.

You will remember that neuraminidase inhibitors were promoted by WHO as a key part of influenza prevention and treatment, and that oseltamivir was stockpiled at vast expense by governments around the world. The drug was made widely and easily available, but even so, huge amounts were left unused. You may also remember that serious doubts were raised about its effectiveness.

At the end of 2009 we published an update of the Cochrane review of antivirals as treatment for flu in otherwise healthy adults (BMJ 2009;339:b5106). As reported in a BMJ/Channel 4 investigation, the reviewers had found that, despite repeated requests to the drug company, Roche, they were unable to obtain the trial data necessary to validate their earlier conclusion that oseltamivir reduced complications (BMJ 2009;339:b5374).

This week the Cochrane team explains why their experience with Roche blows a hole in the systematic review enterprise (p. 148). The incomplete information they obtained from Roche merely proved how inadequate the published record on oseltamivir was. The two main published trials don’t mention any adverse events, but the partial study reports from Roche listed 10 serious events, three of which were classified as possibly due to oseltamivir. By laboriously compiling a full list of industry and non-industry trials, they found one large trial by Roche Shanghai that Roche headquarters in Basel hadn’t got on their list. By looking at the regulatory documents, they found that the largest phase III trial of oseltamivir (unpublished) is hardly mentioned in regulatory documents.

From now on, they say, reviewers must have access to all unpublished data, not only from unpublished trials—the usual focus of concern about publication bias—but also from those that have been published in peer reviewed journals. Reviewers must assess entire trial programmes, and so new tools and methods are needed. If the trial reports are incomplete, reviewers should turn to reports from the drug regulators. As Tom Jefferson, the lead author for the Cochrane review, told me, “It’s goodbye PubMed, goodbye Embase.”

The reviewers have posted their new style protocol for this review on the Cochrane site and, recognising the enormity of the task, they are recording how much work is involved. But it must be clear to everyone that such a heroic approach is unsustainable across the whole of healthcare, given the resource constraints on academics and regulators. Which brings us back to what seems to be the only real solution—that the raw data from trials must be made freely available. Journals clearly have a role to play in making this happen, as An-Wen Chan observes in his editorial (p. 117). The International Committee of Medical Journal Editors meets in a few months’ time. This will be on the agenda.

Fiona Godlee, editor, BMJ fgodlee@bmj.com

Cite this as: BMJ 2011;342:d232

To receive Editor’s Choice by email each week, visit bmj.com/cgi/customalert
Optimism bias, non-replicated studies, and selective reporting

Non-Replication and Inconsistency in the Genome-Wide Association Setting

John P.A. Ioannidis

Clinical and Molecular Epidemiology Unit and Evidence-Based Medicine and Clinical Trials Unit, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Biomedical Research Institute-Foundation for Research and Technology-Hellas, Ioannina, Greece; Department of Medicine, Tufts University School of Medicine, Boston, Mass., USA

Almost all articles on cancer prognostic markers report statistically significant results

Panayiotis A. Kyzas, Despina Denaxa-Kyza, John P.A. Ioannidis
Empirical Evidence for Selective Reporting of Outcomes in Randomized Trials
Comparison of Protocols to Published Articles

An-Wen Chan, MD, DPhil
Ashbjörn Hróbjartsson, MD, PhD
Mette T. Haahr, BSc
Peter C. Gotzsche, MD, DrMedSci
Douglas G. Altman, DSc

SELECTIVE PUBLICATION OF STUDIES with statistically significant results has received widespread recognition. In contrast, selective reporting of favorable outcomes within published studies has not undergone comparable empirical investigation. The existence of outcome reporting bias has been widely suspected for years, but direct evidence is limited to case reports that have low generalizability and may themselves be subject to publication bias.

Our study had 3 goals: (1) to determine the prevalence of incomplete outcome reporting in published reports of randomized trials; (2) to assess the association between outcome reporting and statistical significance; and (3) to evaluate the consistency between primary outcomes specified in trial protocols and those defined in the published articles.

METHODS

Context Selective reporting of outcomes within published studies based on the nature or direction of their results has been widely suspected, but direct evidence of such bias is currently limited to case reports.

Objective To study empirically the extent and nature of outcome reporting bias in a cohort of randomized trials.

Design Cohort study using protocols and published reports of randomized trials approved by the Scientific-Ethical Committees for Copenhagen and Frederiksberg, Denmark, in 1994-1995. The number and characteristics of reported and unreported trial outcomes were recorded from protocols, journal articles, and a survey of trialists. An outcome was considered incompletely reported if insufficient data were presented in the published articles for meta-analysis. Odds ratios relating the completeness of outcome reporting to statistical significance were calculated for each trial and then pooled to provide an overall estimate of bias. Protocols and published articles were also compared to identify discrepancies in primary outcomes.

Main Outcome Measures Completeness of reporting of efficacy and harm outcomes and of statistically significant vs nonsignificant outcomes; consistency between primary outcomes defined in the most recent protocols and those defined in published articles.

Results One hundred two trials with 122 published journal articles and 3736 outcomes were identified. Overall, 50% of efficacy and 65% of harm outcomes per trial were incompletely reported. Statistically significant outcomes had a higher odds of being fully reported compared with nonsignificant outcomes for both efficacy (pooled odds ratio, 2.4; 95% confidence interval [CI], 1.4-4.0) and harm (pooled odds ratio, 4.7; 95% CI, 1.8-12.0) data. In comparing published articles with protocols, 62% of trials had at least 1 primary outcome that was changed, introduced, or omitted. Eighty-six percent of survey responders (42/49) denied the existence of unreported outcomes despite clear evidence to the contrary.

Conclusions The reporting of trial outcomes is not only frequently incomplete but also biased and inconsistent with protocols. Published articles, as well as reviews that incorporate them, may therefore be unreliable and overestimate the benefits of an intervention. To ensure transparency, planned trials should be registered and protocols should be made publicly available prior to trial completion.
Selection in Reported Epidemiological Risks: An Empirical Assessment

Fotini K. Kavoura1, George Liberopoulos1, John P. A. Ioannidis1,2*

1 Clinical and Molecular Epidemiology Unit, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece, 2 Department of Medicine, Tufts University School of Medicine, Boston, Massachusetts, United States of America

Funding: The authors received no specific funding for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Academic Editor: Ezequiel L. Franco, McGill University, Canada


Received: August 22, 2006
Accepted: January 8, 2007
Published: March 6, 2007

Background

Epidemiological studies may be subject to selective reporting, but empirical evidence thereof is limited. We empirically evaluated the extent of selection of significant results and large effect sizes in a large sample of recent articles.

Methods and Findings

We evaluated 389 articles of epidemiological studies that reported, in their respective abstracts, at least one relative risk for a continuous risk factor in contrasts based on median, tertile, quartile, or quintile categorizations. We examined the proportion and correlates of reporting statistically significant and nonsignificant results in the abstract and whether the magnitude of the relative risks presented (coined to be consistently \( \geq 1.00 \)) differs depending on the type of contrast used for the risk factor. In 342 articles (87.9%), \( \geq 1 \) statistically significant relative risk was reported in the abstract, while only 169 articles (43.4%) reported \( \geq 1 \) statistically nonsignificant relative risk in the abstract. Reporting of statistically significant results was more common with structured abstracts, and was less common in US-based studies and in cancer outcomes. Among 50 randomly selected articles in which the full text was examined, a median of nine (interquartile range 5–16) statistically significant and six (interquartile range 3–16) statistically nonsignificant relative risks were presented \( (p = 0.25) \). Paradoxically, the smallest presented relative risks were based on the contrasts of extreme quintiles; on average, the relative risk magnitude was 1.41, 1.42, and 1.36-fold larger in contrasts of extreme quartiles, extreme tertiles, and above-versus-below median values, respectively \( (p < 0.001) \).

Conclusions

Published epidemiological investigations almost universally highlight significant associations between risk factors and outcomes. For continuous risk factors, investigators selectively present contrasts between more extreme groups, when relative risks are inherently lower.
Contradicted and Initially Stronger Effects in Highly Cited Clinical Research

John P. A. Ioannidis, MD

Clinical research on important questions about efficacy of medical interventions is sometimes followed by subsequent studies that either reach opposite conclusions or suggest that the original claims were too strong. Such disagreements may upset clinical practice and acquire publicity in both scientific circles and in the lay press. Several empirical investigations have tried to address whether specific types of studies are more likely to be contradicted and to explain observed controversies. For example, evidence exists that small studies may sometimes be refuted by larger ones.1,2

Similarly, there is some evidence on disagreements between epidemiological studies and randomized trials.3-5 Prior investigations have focused on a variety of studies without any particular attention to their relative importance and scientific impact. Yet, most research publications have little impact while a small minority receives

Context Controversy and uncertainty ensue when the results of clinical research on the effectiveness of interventions are subsequently contradicted. Controversies are most prominent when high-impact research is involved.

Objectives To understand how frequently highly cited studies are contradicted or find effects that are stronger than in other similar studies and to discern whether specific characteristics are associated with such refutation over time.

Design All original clinical research studies published in 3 major general clinical journals or high-impact-factor specialty journals in 1990-2003 and cited more than 1000 times in the literature were examined.

Main Outcome Measure The results of highly cited articles were compared against subsequent studies of comparable or larger sample size and similar or better controlled designs. The same analysis was also performed comparatively for matched studies that were not so highly cited.

Results Of 49 highly cited original clinical research studies, 45 claimed that the intervention was effective. Of these, 7 (16%) were contradicted by subsequent studies, 7 others (16%) had found effects that were stronger than those of subsequent studies, 20 (44%) were replicated, and 11 (24%) remained largely unchallenged. Five of 6 highly-cited nonrandomized studies had been contradicted or had found stronger effects vs 9 of 39 randomized controlled trials (P = .008). Among randomized trials, studies with contradicted or stronger effects were smaller (P = .009) than replicated or unchallenged studies although there was no statistically significant difference in their early or overall citation impact. Matched control studies did not have a significantly different share of refuted results than highly cited studies, but they included more studies with “negative” results.

Conclusions Contradiction and initially stronger effects are not unusual in highly cited research of clinical interventions and their outcomes. The extent to which high citations may provoke contradictions and vice versa needs more study. Controversies are most common with highly cited nonrandomized studies, but even the most highly cited randomized trials may be challenged and refuted over time, especially small ones.

JAMA. 2005;294:218-228 www.jama.com
Why Most Discovered True Associations Are Inflated

John P. A. Ioannidis

Abstract: Newly discovered true (non-null) associations often have inflated effects compared with the true effect sizes. I discuss here the main reasons for this inflation. First, theoretical considerations prove that when true discovery is claimed based on crossing a threshold of statistical significance and the discovery study is underpowered, the observed effects are expected to be inflated. This has been demonstrated in various fields ranging from early stopped clinical trials to genome-wide associations. Second, flexible analyses coupled with selective reporting may inflate the published discovered effects. The vibration ratio (the ratio of the largest vs. smallest effect on the same association approached with different analytic choices) can be very large. Third, effects may be inflated at the stage of interpretation due to diverse conflicts of interest. Discovered effects are not always inflated, and under some circumstances may be deflated—for example, in the setting of late discovery of associations in sequentially accumulated overpowered evidence, in some types of misclassification from measurement error, and in conflicts causing reverse biases. Finally, I discuss potential approaches to this problem. These include being cautious about newly discovered effect sizes, considering some rational down-adjustment, using analytical methods that correct for the anticipated inflation, ignoring the magnitude of the effect (if not necessary), conducting large studies in the discovery phase, using strict protocols for analyses, pursuing complete and transparent reporting of all results, placing emphasis on replication, and being fair with interpretation of results.

(Epidemiology 2008;19: 640–648)
If exposure and disease are not associated

- False positive study
  - If $\alpha = 0.05$
    - 5 studies show false positive results
      - 5 studies will be published

  - Hot topic Bias
    - Publication Bias
  - Positive results bias

- Likely to be meta-analyzed

THE FALSE POSITIVE RESEARCH CYCLE
(Choi, 1998)

Courtesy: Bernard Choi, PHAC
Why Most Published Research Findings Are False

John P. A. Ioannidis

Summary
There is increasing concern that most current published research findings are false. The probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and, importantly, the ratio of true to no relationships among the relationships probed in each scientific field. In this framework, a research finding is less likely to be true when the studies conducted in a field are smaller; when effect sizes are smaller; when there is a greater number and lesser preselection of tested relationships; where there is greater flexibility in designs, definitions, outcomes, and analytical modes; when there is greater financial and other interest and prejudice; and when more teams are involved in a scientific field in chase of statistical significance. Simulations show that for most study designs and settings, it is more likely for a research claim to be false than true. Moreover, for many current scientific fields, claimed research findings may often be simply accurate measures of the prevailing bias. In this essay, I discuss the implications of these problems for the conduct and interpretation of research.

Factors that influence this problem and some corollaries thereof.

Modeling the Framework for False Positive Findings
Several methodologists have pointed out [9–11] that the high rate of nonreplication (lack of confirmation) of research discoveries is a consequence of the convenient, yet ill-founded strategy of claiming conclusive research findings solely on the basis of a single study assessed by formal statistical significance, typically for a p-value less than 0.05. Research is not most appropriately represented and summarized by p-values, but, unfortunately, there is a widespread notion that medical research articles is characteristic of the field and can vary a lot depending on whether the field targets highly likely relationships or searches for only one or a few true relationships among thousands and millions of hypotheses that may be postulated. Let us also consider, for computational simplicity, circumscribed fields where either there is only one true relationship (among many that can be hypothesized) or the power is similar to find any of the several existing true relationships. The pre-study probability of a relationship being true is \( \frac{R}{(R+1)} \). The probability of a study finding a true relationship reflects the power \( 1 - \beta \) (one minus the Type II error rate). The probability of claiming a relationship when none truly exists reflects the Type I error rate, \( \alpha \). Assuming that \( c \) relationships are being probed in the field, the expected values of the \( 2 \times 2 \) table are given in Table 1. After a research finding has been claimed based on achieving formal statistical significance, the post-study probability that it is true is the positive predictive value, PPV. The PPV is also the complementary probability of what Wacholder et al. have called the false positive report probability [10]. According to the 2

It can be proven that most claimed research findings are false.

should be interpreted based only on p-values. Research findings are defined here as any relationship reaching formal statistical significance, e.g., effective interventions, informative predictors, risk factors, or associations. "Negative" research is also very useful...
MUCH OF WHAT MEDICAL RESEARCHERS CONCLUDE IN THEIR STUDIES IS MISTLETOE, MISLEADING, OR FLAT-OUT WRONG. SO WHY ARE DOCTORS—TO A STRIKING EXTENT—STILL DRAWING UPON MISINFORMATION IN THEIR EVERYDAY PRACTICE?

Dr. John Ioannidis has spent his career challenging his peers by exposing their bad science.

LIES, DAMNED LIES, AND MEDICAL SCIENCE

By DAVID H. FREEDMAN

WRONG

Why Experts Keep Failing Us — And How to Know When Not to Trust Them

SCIENTIFIC AMERICAN

Winner of the 2011 National Magazine Award for General Excellence

An Epidemic of False Claims

Competition and conflicts of interest distort too many medical findings

By John P. A. Ioannidis | May 21, 2011 | 20

False positives and exaggerated results in peer-reviewed scientific studies have reached epidemic proportions in recent years. The problem is rampant in economics, the social sciences and even the natural sciences, but it is particularly egregious in biomedicine. Many studies that claim some drug or treatment is beneficial have turned out not to be true. We need only look to conflicting findings about beta-carotene, vitamin E, hormone treatments, Vioxx and Avandia. Even when effects are genuine, their true magnitude is often smaller than originally claimed.
So, you still want to take this course?

“meta-analysis has made and continues to make major contributions to medical research, clinical decision making, and standards of research reportage. However, it is no panacea. Readers need to examine any meta-analyses critically to see whether researchers have overlooked important sources of clinical heterogeneity among the included trials. They should demand evidence that the authors undertook a comprehensive search, avoiding covert duplicate data and unearthing unpublished trials and data. Lastly, readers and researchers alike need to appreciate that not every systematic review should lead to an actual meta-analysis...”

David Naylor. BMJ 1997;315:617-619
Examples of SRs published by previous course participants

- Insufficient Milk Supply and Breast Cancer Risk: A Systematic Review
  Jacqueline M. Cohen, Jennifer A. Hutcheon, Sofi G. Julien, Michel L. Tremblay, Rebecca Fuhrer

- How Methodologic Differences Affect Results of Economic Analyses: A Systematic Review of Interferon Gamma Release Assays for the Diagnosis of LTBI
  Olivia Oxlade, Marcia Pinto, Anete Trajman, Dick Menzies

- Impact of Pharmacist Care in the Management of Cardiovascular Disease Risk Factors
  Valérie Santschi, PharmD, Arnaud Chiolerio, MD, MSc, Bernard Burnand, MD, MPH; April L. Colosimo, MSc, MLIS; Gilles Paradis, MD, MSc

- The pathways to mental health care of first-episode psychosis patients: a systematic review

- A systematic review and meta-analysis of the impact of tuberculosis on health-related quality of life
  M. Bauer, A. Leavens, K. Schwartzman

- Obesity and C-reactive protein in various populations: a systematic review and meta-analysis
  J. Choi, L. Joseph, L. Pilote
Examples of SRs published by previous course participants

Microscopic-observation drug susceptibility and thin layer agar assays for the detection of drug resistant tuberculosis: a systematic review and meta-analysis

Jessica Minion, Erik Leung, Dick Menzies, Madhukar Pai

Annals of Internal Medicine

Accuracy of Rapid Influenza Diagnostic Tests

A Meta-analysis

Caroline Chartrand, MD, MSc; Mariska M.G. Leeflang, DVM, PhD; Jessica Minion, MD, MSc; Timothy Brewer, MD, MPH; and Madhukar Pai, MD, PhD

Meta-Analysis of Antibiotics and the Risk of Community-Associated Clostridium difficile Infection

Kevin A. Brown, Nagham Khanafer, Nick Daneman, David N. Fisman

Epidemiology Division, Dalhousie School of Public Health, University of Toronto, Toronto, Ontario, Canada; Laboratory of Biometry et de Biologie Évolutive, Université de Lyon, Lyon, France; Division of Infectious Diseases, Department of Medicine, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, Ontario, Canada

GenoType MTBDR assays for the diagnosis of multidrug-resistant tuberculosis: a meta-analysis

D.I. Ling*, A.A. Zwerling* and M. Pai*

Review

Hypertermic Isolated Limb Perfusion for Extremity Soft Tissue Sarcomas: Systematic Review of Clinical Efficacy and Quality Assessment of Reported Trials

N.H. Trabuls, MD, L. Patakalvi, MD, O. NASSIF, MD, R.E. TURCOTTE, MD, FRC, A. NICHOLS, P.S.D., and A.N. MEGUERDITCHIAN, MD, FRC, FACS

Detection of scabies: A systematic review of diagnostic methods

Victor Leung MD†, Mark Miller MD²

I. Nicolau*, D. Ling*, L. Tian, C. Lienhardt, M. Pai*
Examples of SRs published by previous course participants

Scoring systems using chest radiographic features for the diagnosis of pulmonary tuberculosis in adults: a systematic review

Lancelot M. Pinto1,2, Madhukar Pai1,2, Keertan Dheda3, Kevin Schwartzman1,4, Dick Menzies1,3,4, and Karen R. Steingart5

Risk of invasive pneumococcal disease in children and adults with asthma: A systematic review

Constantina Boikos2, Caroline Quach3,4,5,6

Interferon-gamma release assays for tuberculosis screening of healthcare workers: a systematic review

Alice Zwerling1, Susan van den Hof2,3, Jerod Scholten2, Frank Cobelens2,3, Dick Menzies1, Madhukar Pai1

A network meta-analysis of antibiotics for treatment of hospitalised patients with suspected or proven meticillin-resistant *Staphylococcus aureus* infection

Michèle Bally4,5,6, Nandini Dendukuri4,5,6, Alison Sinclair6, Stéphane P. Ahern4,5, Michel Poisson6,5, and James Brophy4,5

Meta-Analysis of Usefulness of D-Dimer to Diagnose Acute Aortic Dissection

Avi Shimony, MD1,2,3, Kristian B. Filion, PhD1, Salvatore Mottillo, BSc1,4,5, Tara Dourian6, and Mark J. Eisenberg, MD, MPH1,4,5

Are *Treponema pallidum* Specific Rapid and Point-of-Care Tests for Syphilis Accurate Enough for Screening in Resource Limited Settings? Evidence from a Meta-Analysis

Yalda Jafari1, Rosanna W. Peeling2, Sushmita Shivkumar1, Christiane Claessens2, Lawrence Joseph1,4, Nitika Pant Pai5


Sushmita Shivkumar, MSc1, Rosanna Peeling, PhD2, Yalda Jafari, MSc1, Lawrence Joseph, PhD2, and Nitika Pant Pai, MD, MPH, PhD1,4
Examples of SRs published by previous course participants

OBSTETRICS

Maternal sleep-disordered breathing and adverse pregnancy outcomes: a systematic review and metaanalysis

Sushmita Pamidi, MD; Lancelot M. Pinto, MD, MSc; Isabelle Marc, MD; Andrea Benedetti, PhD; Kevin Schwartzman, MD, MPH; R. John Kimoff, MD

Delays in diagnosis and treatment of pulmonary tuberculosis in India: a systematic review

C. T. Sreeramareddy, Z. Z. Qin, S. Satyanarayana, R. Subbaraman, M. Pai

*Department of Population Medicine, Faculty of Medicine and Health Science, University Tunku Abdul Rahman, Selangor, Malaysia; †Department of Epidemiology and Biostatistics, McGill International TB Centre, McGill University, Montreal, Quebec, Canada; Division of Infectious Diseases, Brigham and Women’s Hospital, Boston, Massachusetts, USA

Spousal diabetes as a diabetes risk factor: A systematic review and meta-analysis

Aaron Leong, Elham Rahme and Kaberi Dasgupta
If they can do it, so can you!
WHAT HAVE YOU LEARNT FROM THE COCHRANE COLLABORATION?

LIFE IS FULL OF TRIALS