Quality of meta-analyses and why they sometimes lead to different conclusions

Jean-Pierre Pignon, MD, PhD

Service de Biostatistique et d’Epidémiologie, Ligue Nationale Contre le Cancer meta-analysis platform, Gustave Roussy Cancer Campus, CESP, Inserm U1018, Paris Sud & Paris-Saclay University, Villejuif, France
Conflicts of Interest: none
Meta-analyses of randomized trials in oncology*

*meta-analysis[Publication Type] AND (cancer OR oncology) AND (randomised clinical trial OR RCT) AND China

376

144
Meta-analyses of randomized trials in oncology

- Increase in number (376 publications in 2014)
- More often overlapping i.e. evaluating same intervention
- Quality varies
- There is no good meta-analysis without systematic review
- Most use time dependent endpoint
Overlapping meta-analyses (MAs)
(from Siontis et al. BMJ 2013;347)

- Survey of a sample of 73 MAs published in 2010 on effectiveness of a medical intervention, identified 138 other MAs of the same intervention until Feb 2013
- 67% (49/73) had at least one overlapping MA, 50% had 2 and 5% had 8 or more
- Among the 49 interventions with more than 2 MAs, 17 had at least one author in common for at least 2 MAs: corresponding either to an update (7), a partial overlap (9), or an exact duplicate (1)
Example: Chemotherapy in locally advanced head and neck squamous cell carcinoma

- **35 meta-analyses** (17 papers) by 13 groups between 1980 and 2015 with 1 to 6 meta-analyses by paper
- Three groups published 8 papers, including 2 updates and 2 partial overlap
- **14 questions** according to timing, tumor site, and type of chemo
- **11 questions** related to adding chemo to locoregional treatment (LRT)
- From **broad questions** “adding chemo to LRT” (87 RCTs) to very **specific questions** “adding induction chemo before surgery in oral cancer” (2 RCTs)
Algorithm to interpret discordant systematic reviews (Jadad et al. CMAJ 1997)

Pragmatic guide to interpret discordant systematic reviews

Two steps:

1. Are the reviews valid? Major flaws $\Rightarrow$ unsuitable in guiding decision

2. Are the differences among the discordant reviews important?
   "A decision maker may consider differences between 2 reviews to be unimportant if the estimated treatment effects are of different magnitude but in the same direction, and are statistically significant and clinically important"
Quality of meta-analyses publications

- Several scales:
  - Oxman & Guyatt index (10 items),
  - AMSTAR (11 items)
  - ROBIS (>21 items)
- To be interpreted with caution: failure in few items may be sufficient to disqualify a meta-analysis (MA)
- Different from recommendations for publication (PRISMA)
- Protocol and trial quality evaluation are systematic in Cochrane review hence Cochrane reviews are better than other meta-analyses
- Individual patient data MA is better than aggregate data MA
Factors related to quality of meta-analyses

- **Protocol**, statistician & clinician **authorship**
- Journal with **high impact factor**, with peer review
- Quality of **trials search**: no language restriction, search of grey literature (abstract, unpublished RCT) through meeting proceedings and trial registries
- Detailed evaluation of **trial quality**
- PRISMA **flow chart**
- Use of **hazard ratio** and not survival rate or median survival ratio to estimate effect on survival endpoint
Do the overlapping MAs ask the same question?

No

- Select “the meta-analysis closest to the question to be solved “ (Jadad et al); e.g., induction chemo in head & neck cancer (HNC): any chemo, platin-based or 5FU-platin?

- But meta-analysis that “considers multiple, ..., available treatment options for the same condition ....can offer more complete pictures of the evidence” (Siontis et al)

  e.g., addition of chemo to loco-regional treatment in HNC, study of best timing by direct and indirect comparisons; network MA of loco-regional and systemic treatment in HNC
Do the meta-analyses ask the same question?

Yes

similar conclusions

- **Duplicate** publications of the same MA (compare authors’ name)
- **Meta-analysis update** ➔ When to update MA?
  » Quantitative methods proposed by Moher et al (Cochrane DSR 2008)
  » Approximate rule: <10% new data, update not useful, except if contradictory results

discordant conclusions

➔ include the same trials?
Do meta-analyses include the same trials?

**Yes**: same quality? Select the meta-analysis with the highest quality

**No**: same selection criteria? Select the meta-analysis with the largest number of trials, taking into account the quality of trials and meta-analyses

Platin-5FU vs. Taxane-Platin-5FU induction chemo: *same overall conclusion*

<table>
<thead>
<tr>
<th>Reference</th>
<th>Number of Trials</th>
<th>Patients in survival analysis</th>
<th>Type of data</th>
<th>Parameters for survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qin 2012</td>
<td>3</td>
<td>1,241</td>
<td>Aggregated</td>
<td>Risk ratio 3-year surv. rate</td>
</tr>
<tr>
<td>Blanchard 2013</td>
<td>5</td>
<td>1,772</td>
<td>Individual patient</td>
<td>Hazard ratio</td>
</tr>
<tr>
<td>Perl 2013</td>
<td>4</td>
<td>1,454</td>
<td>Aggregated</td>
<td>Hazard ratio</td>
</tr>
<tr>
<td>Qian 2015</td>
<td>5</td>
<td>1,765</td>
<td>Aggregated</td>
<td>Hazard ratio</td>
</tr>
</tbody>
</table>
Why MA sometimes lead to different conclusions?

- Question studied, number and quality of RCTs are the main factors to explain different results in overlapping MAs (systematic review).
- Data collection, in particular for time dependent outcome (OR vs. HR) may explain some differences.
- Statistical methods may explain some differences, for instance when the study lacks power, some significant results with fixed effect model may become non significant with random effect model.
- Importance of sensitivity analyses for the robustness of the conclusion.
- Same data + same methods $\Rightarrow$ same results, but their interpretation may differ.
Conclusions and recommendations (1)

- **Duplicate meta-analyses** are useful, in particular if concerns about existing meta-analysis (MA), but too much duplication is research waste.

- **Meta-analyses** should be updated only when new evidence may affect the previous conclusion.

- **Investigators** should register meta-analyses (PROSPERO), and search MA registries before starting a new MA and at the time of its publication.
Conclusions and recommendations (2)

- Investigators should provide their reasons for doing a new meta-analysis (MA) and discuss the results of similar meta-analyses in their publication
- MA publications should follow PRISMA recommendations
- Editors and reviewers should promote registration, use PRISMA and request investigator to take overlapping meta-analyses into account
References

Back-up slides
Overlapping MA: Jadad algorithm to interpret discordant systematic review (CMAJ 1997)

<table>
<thead>
<tr>
<th>Table 1: Sources of discordance among meta-analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinical question</strong></td>
</tr>
<tr>
<td>Populations of patients</td>
</tr>
<tr>
<td>Interventions</td>
</tr>
<tr>
<td>Outcome measures</td>
</tr>
<tr>
<td>Settings</td>
</tr>
<tr>
<td><strong>Study selection and inclusion</strong></td>
</tr>
<tr>
<td>Selection criteria</td>
</tr>
<tr>
<td>Application of the selection criteria</td>
</tr>
<tr>
<td>Strategies to search the literature</td>
</tr>
<tr>
<td><strong>Data extraction</strong></td>
</tr>
<tr>
<td>Methods to measure outcomes</td>
</tr>
<tr>
<td>End points</td>
</tr>
<tr>
<td>Human error (random or systematic)</td>
</tr>
<tr>
<td><strong>Assessment of study quality</strong></td>
</tr>
<tr>
<td>Methods to assess quality</td>
</tr>
<tr>
<td>Interpretations of quality assessments</td>
</tr>
<tr>
<td>Methods to incorporate quality assessments in review</td>
</tr>
<tr>
<td><strong>Assessment of the ability to combine studies</strong></td>
</tr>
<tr>
<td>Statistical methods</td>
</tr>
<tr>
<td>Clinical criteria to judge the ability to combine studies</td>
</tr>
<tr>
<td><strong>Statistical methods for data synthesis</strong></td>
</tr>
</tbody>
</table>
Overlapping MA: Jadad algorithm to interpret discordant systematic review (CMAJ 1997)

<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results</td>
<td></td>
</tr>
<tr>
<td>Direction of effect</td>
<td>One review favours the experimental treatment and another favours the control treatment</td>
</tr>
<tr>
<td>Magnitude of effect</td>
<td>One review suggests that the intervention results in a 30% reduction in mortality and another suggests that it results in a 5% reduction in mortality</td>
</tr>
<tr>
<td>Statistical significance</td>
<td>One review shows a statistically significant difference between the experimental and the control treatments and another review shows a nonsignificant difference between them</td>
</tr>
</tbody>
</table>

Interpretation of the results
Overlapping MA: Jadad algorithm to interpret discordant systematic review (CMAJ 1997)

A. Same question?
   - Yes
   - No (B)

B. Select the question closest to the problem to be solved

C. Same trials?
   - Yes
   - No (D)

D. Same quality?
   - Yes
   - No (G)

E. Assess and compare
   - data extraction
   - heterogeneity testing
   - data synthesis

F. Select the review with the highest quality

G. Same selection criteria?
   - Yes
   - No (I)

H. Assess and compare
   - search strategies
   - application of selection criteria

I. Assess and compare
   - publication status of primary trials
   - methodologic quality of primary trials
   - language restrictions
   - analysis of data on individual patients
Acknowledgement