Challenges in Assessing Adverse Effects

Yoon K Loke
Cochrane Adverse Effects Methods Group
(supported by Cochrane Methods Innovation Fund)
Today’s Content

- Adverse Effects (AE) are Unintended Outcomes
- How and why synthesizing AE data is more challenging
- Intended Outcomes
- Formulating relevant and important questions
- Constructing a PICO
- Integrating AE review with intended outcomes
- Relevant study designs
- Interpreting zero events and Risk of bias
- Outcomes tables and tackling selective non-reporting
- Danger of post-hoc decisions in AE reviews
Typical Review focuses on Intended Outcomes (Benefit)

- Pre-specified / defined primary outcome (usually beneficial effect of intervention)
- Outcome is main focus of research study, thus rigorous monitoring
- Power calculation to plan sample size
- Transparent reporting of data for primary outcome
Adverse Effects: Tiger Country

Photo by digitalART2, available under Creative Commons License
How AEs differ from Intended Outcomes

• Seldom considered as primary interest
  • Not prespecified or defined, thus inconsistent measurement/coding, or missed altogether
  • Study not powered to detect significant differences in secondary outcomes
  • Most AE are less frequent than beneficial outcomes, so effect estimates are imprecise
  • AE poorly reported due to focus on reporting main outcomes
Almost limitless, diverse range of AEs

- Impossible for single study to capture all types of AE (common or rare, occurs shortly after intervention or long term)
- Some can be predicted (e.g. wound infection from surgery)
- Some new or unexpected; may not be correctly diagnosed (only become apparent on post-hoc analysis of emergent data)
- Only certain AE are reported; others selectively non-reported
- Multiple statistical testing – false alarms
### Example: AE Reporting (GSK trial)

<table>
<thead>
<tr>
<th>Serious Adverse Events - On-Therapy, n (%)</th>
<th>RSG N=1456</th>
<th>MET N=1454</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myocardial infarction</td>
<td>20 (1.4)</td>
<td>15 (1.0)</td>
</tr>
<tr>
<td>Angina pectoris</td>
<td>8 (0.5)</td>
<td>19 (1.3)</td>
</tr>
<tr>
<td>Coronary artery disease</td>
<td>12 (0.8)</td>
<td>16 (1.1)</td>
</tr>
<tr>
<td>Angina unstable</td>
<td>8 (0.6)</td>
<td>7 (0.5)</td>
</tr>
<tr>
<td>Acute myocardial infarction</td>
<td>3 (0.2)</td>
<td>3 (0.2)</td>
</tr>
<tr>
<td>Myocardial ischemia</td>
<td>2 (0.1)</td>
<td>2 (0.1)</td>
</tr>
<tr>
<td>Coronary artery stenosis</td>
<td>3 (0.2)</td>
<td>2 (0.1)</td>
</tr>
<tr>
<td>Acute coronary syndrome</td>
<td>0</td>
<td>3 (0.2)</td>
</tr>
</tbody>
</table>

- Very similar Cardiac events split under multiple different categories
- Impossible to judge extent of duplication (same event coded more than once?)
- Or if number of events = number of patients
Common endpoint: Composite AE

• Widely used and reported for comparing AE rates intervention vs. control:
  • Total Serious AE
  • Total Withdrawals due to AE

• But suffers same flaws as any other composite:
  • Huge mish-mash of diverse events
  • Elevated risk of a rare AE obscured by common AE
  • Some AE are due to treatment failure/worsening disease
  • Who decides what the reason for withdrawal is? Often complex or multifactorial.
Decision Point: Review Question

• What do you want to achieve with your AE review?
• Different points of view
  • I have read that Treatment A is associated with brain haemorrhage
  • Patients want to know if New Treatment C genuinely has fewer stomach and skin AE than existing Treatment D
  • I don’t really have any specific AE in mind, but I just want to have a general look around the Included trials to see if anything suspicious pops up
Formulating Review Question (2)

• Impossible to pre-specify all conceivable AE

• Three pragmatic approaches:
  • Focused – specific evaluation of a few important AE
  • Broader exploratory – ad hoc evaluation of any or all AE that happened to be reported in the Included studies.
  • A bit of both
Targeted /Hypothesis-Testing

• Pre-specify a few important events of interest (in the same way as intended outcomes)
• Can do scoping search for relevant events e.g.
  • Mechanistic plausibility (wound infection with surgery; bleeding with drugs that block blood clotting)
  • Signals identified in early studies (phase I/II trials, regulatory documents)
• This evaluates presence or absence of association between intervention and important AE
• Fails to pick up new or unexpected AE
Exploratory/ Hypothesis-Generating

- Does not name any particular AE outcomes
- Reviewers check all Included/Relevant studies, to fish out all or any AE
- Compile potentially huge list of disparate items
- Detects new or unexpected issues but:
  - Difficult to synthesize large chunks of varied data
  - Affected by multiple testing and post-hoc decisions
- Generates potential new signals, rather than confirmatory – further focused/hypothesis testing evaluation is needed as follow-up
A Bit of Both

• Review can conceivably have hybrid approach:
  • Main focus on a few important AE
  • Subsidiary exploratory section on any new AE
Constructing a PICO

• AE data potentially available from any study that fulfils Participant-Intervention-Comparator criteria

• Difficulties in directly comparing benefit vs. harm if:
  • Studies included in the benefit meta-analysis are different from those in the harms meta-analysis
  • AE data available in diverse participants that are not necessarily covered in single review e.g. aspirin used in headache, stroke, heart attacks. This may require separate review covering aspirin/AE in all populations.
Decision Point: Study Designs

• Can I just use RCTs or should I broaden selection to non-randomized designs?

• Depends on types of AE outcomes e.g. I’m worried about osteoporosis tablets and AE:
  • Nausea and stomach pain after taking the tablet
  • A serious rare complication known as osteonecrosis of the jaw (say, 1 in 1000)
  • Atypical bone fractures after 5-10 years of treatment

• Broad-sweep, exploratory reviews of diverse AE – difficult to determine what designs are most suitable
Relevant Study Designs: RCTs

• RCTs – more suited for AE that:
  • Are predictable, defined or well-recognized,
  • Have common background incidence
  • Or develop soon after starting treatment
Relevant Study Designs: Observational

• Non-randomized / database designs – more suited for AE that:
  • Unexpected or not predicted in trial
  • Relatively low background incidence
  • Requires longer term follow-up
Searching and Data Sources

- How far should I search beyond typical databases?
- Methods research - substantial missing AE data can be retrieved from unpublished sources
- Use of unpublished data best suited to reviews that have pre-specified AE of interest (otherwise risk of being swamped with too much data).
- Su Golder will cover this in detail on May 25th
Decision Point: Interpreting Zero Events

- How do we deal with statements such as ‘No significant harm was found’ or ‘Safe and well-tolerated’?

- Multiple potential interpretations:
  - We didn’t measure it/ we didn’t ask participants
  - We measured it but didn’t find anything (true zero)
  - We measured and analysed it but the findings were not statistically significant, so we didn’t report the data
Interpreting Zero Events

- High risk of type II error (false reassurance that intervention is safe) because trials not designed for uncommon/unexpected AE
- Interpretation on absence of significant harm and zero events should be judged in context:
  - Sample size
  - Length of follow-up
  - Adequate definition, monitoring & risk of misclassification
- Conclusions or GRADE should be tempered according to context (e.g. imprecision, likelihood of estimates changing with further study)
Risk of Bias (ROB)

• Update - ROB tools assess each outcome separately e.g. blinding of outcome assessor
  • Not relevant to Mortality AE
  • But relevant to judgement of “cardiovascular cause of death”
  • Participant blinding is relevant to symptom ‘nausea’
• ROB tools not feasible with broad sweep, exploratory AE review that consider lots of outcomes
Selective Non-Reporting

• Inevitable when trials measure hundreds of AE, but can only report a few in published manuscripts
• What is the direction of reporting bias?
  • Choose to report only those with significant harm?
  • Or emphasize safety by focusing on areas where no harm was found?
• Direction of bias depends on standpoint of researcher
• Interpretation of asymmetry testing or funnel plot is challenging
# Example of Outcomes Table

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Review harm outcomes (including odds ratio with 95% confidence interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>1: AE 1</td>
<td></td>
</tr>
<tr>
<td>2: AE 2</td>
<td></td>
</tr>
<tr>
<td><strong>Study 1</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Full:</td>
</tr>
<tr>
<td></td>
<td>OR 1.11 (95% CI, 0.89 to 1.34)</td>
</tr>
<tr>
<td></td>
<td>Not reported</td>
</tr>
<tr>
<td><strong>Study 2</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Not reported</td>
</tr>
<tr>
<td></td>
<td>Partial:</td>
</tr>
<tr>
<td></td>
<td>OR 1.11 (95% CI not reported)</td>
</tr>
<tr>
<td><strong>Study 3</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Not reported</td>
</tr>
<tr>
<td></td>
<td>Partial:</td>
</tr>
<tr>
<td></td>
<td>Authors stated: “No significant difference observed”; Effect estimates not reported</td>
</tr>
</tbody>
</table>
Bias in Review Process – Post-hoc Decisions

• AE reviews particularly susceptible to bias because numerous points where post-hoc decisions are made:
  • Inconsistent outcome definitions – what to extract, which ones to pool (or not)
  • Poor reporting in primary studies – ambiguity in interpretation
  • Exploratory nature of AE reviews, with multiple testing
• Decisions in AE review should be transparently reported
Conclusions

• Important Differences between Reviews of Intended Outcomes and Adverse Effects
• Review Methods mainly determined by initial decision on what AE outcomes are of most interest
• Formulating study question is most important step – the subsequent path flows on from there.
• Diversity of AE and poor reporting are the main challenges that need to be overcome