Statistical challenges of meta-analyses of randomised clinical trials in a regulatory setting

Frank Pétavy

ISCTM 14th Annual Scientific Meeting, Washington D.C.

Presented by Frank Pétavy on 21 February 2018
Biostatistics and Methodology Support, European Medicines Agency
Disclaimer

The views expressed in this presentation are the personal views of the speakers and may not be understood or quoted as being made on behalf of or reflecting the position of the EMA or one of its committees or working parties.

Copyright: European Medicines Agency.
Acknowledgements

Andrew Thomson (EMA)

Steven Teerenstra (MEB/CBG, Radboud University Nijmegen)
Outline

- Motivating example: cardiovascular risk with varenicline
- General reflections on challenges in meta-analyses, on opportunities with individual patient meta-analyses and on estimands in clinical trials for safety
Motivating Example

• Several **meta-analyses** of **cardiovascular adverse events** in trials for varenicline (Champix, approved in the EU and US in 2006). **Two in particular:**
  - Meta-analysis from an academic group (Singh et al) gave a signal of increased cardiovascular (CV) risk.
  - In response to this and upon request from the FDA, the MAH (Pfizer) performed a meta-analysis of their Phase II-IV trials.

• **Varenicline:**
  - Highly selective neuronal α4β2 nicotinic acetylcholine receptor partial agonist.
  - Reduces craving / withdrawal of smoking cessation.
Singh et al. meta-analysis: studies ¹
(CMAJ, epub July 2011)

14 randomised clinical trials (4908 varenicline, 3308 placebo)

13 in smokers, 1 in smokeless tobacco

13 with no clinically significant CV disease within 6 months prior to entry, 1 stable CV disease at least 2 months prior entry

7-52 weeks of treatment duration

24-52 weeks of study duration
Singh et al. meta-analysis: endpoints

• **Primary endpoint:** any ischemic or arrhythmic adverse event:
  - Myocardial infarction, stroke, sudden or CV death, unstable angina, coronary revascularization, coronary artery disease, arrhythmias, transient ischemic attacks, congestive heart failure

• **Of interest:** definition of serious AE (requiring intervention, hospitalisation) not standardised over studies

• **Secondary endpoint:** all-cause mortality
Singh et al. meta-analysis: results

<table>
<thead>
<tr>
<th>Study</th>
<th>Cardiovascular events, n/N</th>
<th>Weight, %</th>
<th>Peto OR (95% CI)</th>
<th>Decreased risk with varenicline</th>
<th>Increased risk with varenicline</th>
</tr>
</thead>
</table>
| Protocol A3051080 
16 | 1/394 | 0/199 | 1.2 | 4.50 (0.07–285.96) |                                |
| Protocol A3051095 
17 | 1/493 | 0/166 | 1.0 | 3.81 (0.04–347.82) |                                |
| Fagerstrom et al. 
18 | 0/214 | 1/218 | 1.4 | 0.14 (0.00–6.95) |                                |
| Gonzales et al. 
19 | 2/352 | 2/344 | 5.4 | 0.98 (0.14–6.97) |                                |
| Jorenby et al. 
20 | 1/344 | 1/341 | 2.7 | 0.99 (0.06–15.88) |                                |
| Nakamura et al. 
21 | 1/465 | 0/154 | 1.0 | 3.79 (0.04–352.44) |                                |
| Niaura et al. 
22 | 2/160 | 0/160 | 2.7 | 7.44 (0.46–119.40) |                                |
| Nides et al. 
23 | 1/383 | 0/127 | 1.0 | 3.79 (0.04–352.09) |                                |
| Oncken et al. 
24 | 2/518 | 0/129 | 1.7 | 3.49 (0.11–112.44) |                                |
| Rigotti et al. 
9 | 25/355 | 20/359 | 57.3 | 1.28 (0.70–2.34) |                                |
| Tashkin et al. 
25 | 5/250 | 2/254 | 9.4 | 2.42 (0.55–10.74) |                                |
| Tonstad et al. 
26 | 4/603 | 0/607 | 5.4 | 7.48 (1.05–53.20) |                                |
| Tsai et al. 
27 | 1/126 | 0/124 | 1.4 | 7.27 (0.14–366.57) |                                |
| Williams et al. 
28 | 6/251 | 1/126 | 8.3 | 2.40 (0.49–11.67) |                                |
| Overall        | 52/4908 | 27/3308 | 100.0 | 1.72 (1.09–2.71) |                                |

Heterogeneity: $I^2 = 0\%$

Figure 2: Meta-analysis of double-blind placebo-controlled randomized trials of the risk of serious adverse cardiovascular events associated with the use of varenicline. An odds ratio (OR) greater than 1.0 indicates an increased risk of a serious adverse cardiovascular event. CI = confidence interval.
CHMP assessment, November 2011

Champix marketing-authorisation holder should:

- Introduce **changes in SmPC**
  - **Include more information on cardiovascular events** → appropriate warnings on risk of cardiovascular events (section 4.4 - Special warnings and precautions for use)
  - The text has been made **more factual** by including information on the **frequency of cardiovascular events** in general and specific events such as myocardial infarction in each treatment group.
Ware et al meta-analysis: studies

- Done by the MAH (marketing authorisation holder)
- Included all Pfizer Phase II-IV randomised clinical trials
  - Blinded, ≥12 weeks, with report before 30th June 2011
  - And two open-label, active-controlled, RCT
- No non-Pfizer trials (allowed access to individual patient data)
Ware et al meta-analysis: endpoints

• **Endpoints**
  
  – MACE (Major Adverse Cardiovascular Events): CV death, non-fatal myocardial infarction, non-fatal stroke
  
  – MACE+: MACE plus new onset or worsening peripheral vascular disease (PVD) requiring intervention; hospitalization for unstable angina; coronary revascularization

• **Summary measures:** time-to-event + incidence rates
Ware et al meta-analysis: results

![Data Table]

**Note:** many studies with 0 events
Methodological questions on meta-analysis (1)

• Are the ‘appropriate’ studies included?
  – Is the search strategy adequate?
  – (publication) bias?
  – Quality of study (allocation concealment, blinding of personnel/patients, reporting withdrawals, ability to capture/report adverse events)

• Are the endpoints adequate (for the safety question)?
Methodological questions on meta-analysis (2)

• Are the studies similar enough to be combined (heterogeneity)?
  – Assessment bias
  – Different populations
  – Treatment duration, follow-up

• Is the **statistical method of combining** appropriate?

• Is the **interpretation** of the meta-analysis correct?
Singh et al meta-analysis: search strategy

<table>
<thead>
<tr>
<th>Database</th>
<th>Search Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEDLINE (OvidSP) using Haynes</td>
<td>Varenicline.mp AND (randomized controlled trial.pt. OR randomized.mp. OR placebo.mp.)</td>
</tr>
<tr>
<td>optimized filter: search March</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td></td>
</tr>
<tr>
<td>EMBASE using OvidSP with</td>
<td>Varenicline.mp</td>
</tr>
<tr>
<td>randomized controlled trial filter</td>
<td></td>
</tr>
<tr>
<td>applied: search March 2011</td>
<td></td>
</tr>
<tr>
<td>Cochrane Controlled Trials</td>
<td>(varenicline)</td>
</tr>
<tr>
<td>Register: search March 2011</td>
<td></td>
</tr>
<tr>
<td>CLINICALTRIALS.GOV: search</td>
<td>(varenicline [Limits: Completed Studies])</td>
</tr>
<tr>
<td>September 2010</td>
<td></td>
</tr>
<tr>
<td>CLINICALSTUDYRESULTS.ORG: search</td>
<td>(varenicline)</td>
</tr>
<tr>
<td>September 2010</td>
<td></td>
</tr>
</tbody>
</table>
Singh et al meta-analysis: endpoints

Different studies had **different definitions of serious adverse events**
Singh et al meta-analysis: endpoints

### Appendix 3: Risk-of-bias assessment and methodologic details of randomized controlled trials of varenicline included in the analysis of serious adverse cardiovascular events

<table>
<thead>
<tr>
<th>Study</th>
<th>Sequence generation</th>
<th>Allocation concealment</th>
<th>Adequate monitoring of adverse events</th>
<th>Drug and dose</th>
<th>Withdrawal rates [%]</th>
<th>Losses [n]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protocol A205.100, 2010</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Adequate</td>
<td>Varenicline 1 mg bid</td>
<td>19.3 ± 2.3</td>
<td></td>
</tr>
<tr>
<td>Protocol A206.109, 2010</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Adequate</td>
<td>Placebo</td>
<td>22.2 ± 2.5</td>
<td></td>
</tr>
</tbody>
</table>

- **Gonzales et al., 2006**
  - Adequate monitoring of adverse events
  - No safety assessments in follow-up period
  - Adverse events resulting in hospital admission, disability or death within 7 d of last drug dose were classified as serious

- **Niura et al., 2008**
  - Adequate monitoring of adverse events
  - No adverse events recorded through 10 d after last study dose

- **Nides et al., 2006**
  - Adequate monitoring of adverse events
  - No safety assessments in follow-up period
  - Adverse events resulting in hospital admission, disability or death within 7 d of last drug dose were classified as serious

- **Rigetti et al., 2006**
  - Adequate monitoring of adverse events
  - No safety assessments in follow-up period
  - Adverse events resulting in hospital admission, disability or death within 7 d of last drug dose were classified as serious

- **Tashjian et al., 2010**
  - Adequate monitoring of adverse events
  - No safety assessments in follow-up period
  - Adverse events resulting in hospital admission, disability or death within 7 d of last drug dose were classified as serious

- **Tomasi et al., 2010**
  - Adequate monitoring of adverse events
  - No safety assessments in follow-up period
  - Adverse events resulting in hospital admission, disability or death within 7 d of last drug dose were classified as serious

- **Liu et al., 2007**
  - Adequate monitoring of adverse events
  - No safety assessments in follow-up period
  - Adverse events resulting in hospital admission, disability or death within 7 d of last drug dose were classified as serious

- **Williams et al., 2008**
  - Adequate monitoring of adverse events
  - No safety assessments in follow-up period
  - Adverse events resulting in hospital admission, disability or death within 7 d of last drug dose were classified as serious

- **Aubin et al., 2008**
  - Adequate monitoring of adverse events
  - No safety assessments in follow-up period
  - Adverse events resulting in hospital admission, disability or death within 7 d of last drug dose were classified as serious

---

No safety assessments in follow-up period; adverse events resulting in hospital admission, disability or death within 7 d of last drug dose were classified as serious.
Singh et al meta-analysis: endpoints

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Randomization</th>
<th>Sequence generation</th>
<th>Allocation concealment</th>
<th>Adequate monitoring of adverse events</th>
<th>Adequate monitoring of death</th>
<th>Withdrawal rate (%)</th>
<th>Loss to follow-up, %</th>
<th>Drug and dose</th>
<th>Endpoints Monitoring and laboratory tests (including statistical analysis)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singh et al.</td>
<td>Randomized</td>
<td>Centralized</td>
<td>Clear</td>
<td>Clear</td>
<td>Adequate</td>
<td>Adequate</td>
<td>22.2</td>
<td>2.5</td>
<td>Varenicline 1 mg bid</td>
<td>(1) loss of weight, (2) smoking cessation, (3) smoking reduction, (4) smoking relapse</td>
</tr>
<tr>
<td>Ettinger et al.</td>
<td>Randomized</td>
<td>Centralized</td>
<td>Clear</td>
<td>Clear</td>
<td>Adequate</td>
<td>Adequate</td>
<td>14.5</td>
<td>0.1</td>
<td>Varenicline 1 mg bid</td>
<td>(1) loss of weight, (2) smoking cessation, (3) smoking reduction, (4) smoking relapse</td>
</tr>
<tr>
<td>Gonzalez et al.</td>
<td>Randomized</td>
<td>Centralized</td>
<td>Clear</td>
<td>Clear</td>
<td>Adequate</td>
<td>Adequate</td>
<td>14.5</td>
<td>0.1</td>
<td>Varenicline 1 mg bid</td>
<td>(1) loss of weight, (2) smoking cessation, (3) smoking reduction, (4) smoking relapse</td>
</tr>
<tr>
<td>Ameen et al.</td>
<td>Randomized</td>
<td>Centralized</td>
<td>Clear</td>
<td>Clear</td>
<td>Adequate</td>
<td>Adequate</td>
<td>14.5</td>
<td>0.1</td>
<td>Varenicline 1 mg bid</td>
<td>(1) loss of weight, (2) smoking cessation, (3) smoking reduction, (4) smoking relapse</td>
</tr>
<tr>
<td>Nabakuma et al.</td>
<td>Randomized</td>
<td>Centralized</td>
<td>Clear</td>
<td>Clear</td>
<td>Adequate</td>
<td>Adequate</td>
<td>14.5</td>
<td>0.1</td>
<td>Varenicline 1 mg bid</td>
<td>(1) loss of weight, (2) smoking cessation, (3) smoking reduction, (4) smoking relapse</td>
</tr>
<tr>
<td>Bisera et al.</td>
<td>Randomized</td>
<td>Centralized</td>
<td>Clear</td>
<td>Clear</td>
<td>Adequate</td>
<td>Adequate</td>
<td>14.5</td>
<td>0.1</td>
<td>Varenicline 1 mg bid</td>
<td>(1) loss of weight, (2) smoking cessation, (3) smoking reduction, (4) smoking relapse</td>
</tr>
<tr>
<td>Nodis et al.</td>
<td>Randomized</td>
<td>Centralized</td>
<td>Clear</td>
<td>Clear</td>
<td>Adequate</td>
<td>Adequate</td>
<td>14.5</td>
<td>0.1</td>
<td>Varenicline 1 mg bid</td>
<td>(1) loss of weight, (2) smoking cessation, (3) smoking reduction, (4) smoking relapse</td>
</tr>
<tr>
<td>Cho et al.</td>
<td>Randomized</td>
<td>Centralized</td>
<td>Clear</td>
<td>Clear</td>
<td>Adequate</td>
<td>Adequate</td>
<td>14.5</td>
<td>0.1</td>
<td>Varenicline 1 mg bid</td>
<td>(1) loss of weight, (2) smoking cessation, (3) smoking reduction, (4) smoking relapse</td>
</tr>
</tbody>
</table>

Adverse events resulting in hospital admission, disability or death, or congenital anomalies or birth defects were classified as serious. Reported or observed cardiovascular or death events were reviewed and adjudicated under blinded conditions by an independent event committee that used a standard events manual.
Singh et al. meta-analysis: naïve pooling

- **Naïve pooling:** Varenicline 1.06% vs. placebo 0.82% (52/4908 vs 27/3308).
  - 3 trials with multiple arms.

- **Limitations:**
  - Small number of CV events;
  - Differences in length of follow-up;
  - Non-inclusion of trials with no serious AEs;
  - Lumping CV events with different pathophysiological pathways and different degree of clinical importance.
Singh et al meta-analysis: method to combine studies

• **Peto odds ratio:**

  ".. provides best confidence interval coverage and is more powerful and relatively less biased than the random-effects analysis when dealing with low event rates."

• **Comments:**
  - Biased when *treatment groups of different size*
  - Not accounting for differences in *follow-up times* (alternative: time to event analysis)
  - Does not include *0-0 event* studies (alternative: risk difference)
  - Other *weighting* methods available
Ware et al meta-analysis: endpoints

- Homogeneous definition
- Event adjudication:
  - Post hoc, based on trial data
  - Could decrease power to detect a difference (adding non-CV related events or conversely)
  - Changes due to adjudication need to be clear and transparent

Table 2. Distribution of first MACE+ occurring in 38 unique subjects during and 30 days after blinded treatment.

<table>
<thead>
<tr>
<th>Event</th>
<th>Varenicline, n</th>
<th>Placebo, n</th>
</tr>
</thead>
<tbody>
<tr>
<td>n (subject-years)</td>
<td>4190 (1314)</td>
<td>2812 (838)</td>
</tr>
<tr>
<td>CV death</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Nonfatal MI</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>Nonfatal stroke</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Unstable angina</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Peripheral vascular disease</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Coronary revascularization</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td><strong>26</strong></td>
<td><strong>12</strong></td>
</tr>
<tr>
<td>All-cause mortality</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

MACE+, major CV events plus worsening or any procedure for peripheral vascular disease, hospitalization for angina, or performance of coronary revascularization.
Ware et al meta-analysis: method to combine studies

- Time-to-event
- Risk difference allows inclusion of the 0 event studies
- Modified Peto odds ratio (Peto-type observed minus expected method for incidence rate ratio) – needs to be explained and justified when method not conventional
General considerations on meta-analysis

• One large study but different characteristics vs. other studies;
  – Results of the meta-analysis driven by the one large study
  – Only solution may be one large trial (as opposed to a large meta-analysis)
• Issue of unequal randomisation (when 0 events)
• Issue of 0 events for calculation of summary measures
• Contrast issues with meta-analyses depending on data sources (CT, GP data)
• Early phase studies won’t have SAEs, otherwise development would have stopped, so funnel plot not useful, other data presentations needed
General considerations on meta-analysis

- **Figure 2 from Singh et al:** look at the weight the one study has, and then look at the rate in all the other studies. It’s extremely different, due to the patient characteristics being so different in this study. The question then arises: **is it sensible to pool all studies if/when patient populations are so different?**
General considerations on meta-analysis: removal of trials with 0-0 events

- Particularly problematic when randomisation is unbalanced, e.g. 0/400 vs. 0/200
- In general, when there is unbalanced randomisation it is in favour of the active arm, so excluding these trials (almost) always goes against active and bias the results.
- Using the risk difference, a raised risk may disappear.
Individual patient meta-analysis

- Individual patient data **allows** investigation of:
  - Types of events
  - Observation periods
  - Subgroups
- Analysis - individual patient data **needed for**:
  - Time-to-event (which takes into account different lengths of follow-up between arms)
  - Covariate adjustment
- Quality of meta-analysis increased with IPD vs. aggregate data?
Comments on estimands in CTs for safety

MACE in placebo-controlled studies

<table>
<thead>
<tr>
<th>Observation period</th>
<th>Varenicline (n=4566)</th>
<th>Placebo (n=2812)</th>
<th>Hazard ratio</th>
<th>Risk incidence difference</th>
<th>Risk incidence ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>7</td>
<td>2</td>
<td>2.83 (0.76, 10.55) P-value=0.12</td>
<td>0.007 (-0.002, 0.015) P-value=0.12</td>
<td>2.82 (0.76, 10.52) P-value=0.12</td>
</tr>
<tr>
<td>Treatment + 30 days</td>
<td>13</td>
<td>6</td>
<td>1.95 (0.79, 4.82) P-value=0.15</td>
<td>0.006 (-0.002, 0.015) P-value=0.16</td>
<td>1.92 (0.78, 4.75) P-value=0.16</td>
</tr>
<tr>
<td>Study</td>
<td>18</td>
<td>10</td>
<td>1.66 (0.79, 3.49) P-value=0.18</td>
<td>0.006 (-0.003, 0.015) P-value=0.19</td>
<td>1.65 (0.78, 3.47) P-value=0.19</td>
</tr>
</tbody>
</table>

*patients as denominator, calculated from the data deduced from Tech Report Pfizer. The hazard ratio for MACE during treatment in the PCS (=placebo controlled trial set) is primary analysis.

- Here, the intercurrent event is **treatment discontinuation**
- Choice of **while on treatment** vs. **treatment policy** strategy
References


Thank you for your attention

Acknowledgements: Andrew Thomson, Steven Teerenstra

Further information
Frank.Petavy@ema.europa.eu

European Medicines Agency
30 Churchill Place • Canary Wharf • London E14 5EU • United Kingdom
Telephone +44 (0)20 3660 6000 Facsimile +44 (0)20 3660 5555
Send a question via our website www.ema.europa.eu/contact

Follow us on @EMA_News