Why meta-analysis should be based on individual patient data

Marc Buyse

IDDI, Louvain-la-Neuve, and Hasselt University, Diepenbeek, Belgium

ROeS Seminar 2007, 10 September 2007
Bern, Switzerland
Types of meta-analyses

- Based on the literature
- Based on summary statistics
- Based on individual patient data (IPD)
Disadvantages of IPD meta-analyses

- Time
- Cost
- Feasibility
Advantages of IPD meta-analyses

- Reduction of bias
- Maximization of power
- Quality checks
- Investigations of heterogeneity
- Further analyses
Advantages of IPD meta-analyses

- Reduction of bias
  - Publication bias
  - Published statistics are often biased
- Maximization of power
- Quality checks
- Investigations of heterogeneity
- Further analyses
Advantages of IPD meta-analyses

- Reduction of bias
- Maximization of power
  - Power is lost in time to event analyses if event is frequent and time is ignored
  - Methods to extract information from Kaplan-Meier curves require assumptions
- Quality checks
- Investigations of heterogeneity
- Further analyses
Advantages of IPD meta-analyses

- Reduction of bias
- Maximization of power
- Quality checks
  - Validity of randomization can partially be checked
  - Patient characteristics can be explored
- Investigations of heterogeneity
- Further analyses
## Advantages of IPD meta-analyses

<table>
<thead>
<tr>
<th>Dates of entry</th>
<th>Control</th>
<th>Experimental</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>November 1981 to March 1982</td>
<td>0</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>April 1982 to July 1985</td>
<td>59</td>
<td>56</td>
<td>115</td>
</tr>
<tr>
<td>August 1985 to October 1985</td>
<td>8</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>67</td>
<td>68</td>
<td>135</td>
</tr>
</tbody>
</table>
Advantages of IPD meta-analyses

- Reduction of bias
- Maximization of power
- Quality checks
- Investigations of heterogeneity
  - Heterogeneity can be due to design and/or patient mix, or to random variation
  - Meta-regression potentially subject to ecologic and/or baseline risk bias
- Further analyses
Ecologic bias

Meta-regression (in which the units of analysis are the trials) may be subject to ecologic bias and imply incorrect relationship between treatment effect and prognostic factors at the individual level.

Ecologic bias

Example:

Absolute risk reductions (ARR) with adjuvant treatment for early colorectal cancer in three trials including patients with stage II and stage III disease

Ref: Buyse and Piedbois, Seminars Oncol 2001;S1:20.
Ecologic bias

Overall ARR Trial 1

Trial 2

Trial 3

Proportion of Stage III
### Ecologic bias

<table>
<thead>
<tr>
<th></th>
<th>(\text{ARR}_{\text{III}})</th>
<th>(\text{ARR}_{\text{II}})</th>
<th>(\text{Prop}_{\text{III}})</th>
<th>(\text{Prop}_{\text{II}})</th>
<th>(\text{ARR})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial 1</td>
<td>0.4</td>
<td>0.2</td>
<td>0.1</td>
<td>0.9</td>
<td>0.22</td>
</tr>
<tr>
<td>Trial 2</td>
<td>0.2</td>
<td>0.1</td>
<td>0.5</td>
<td>0.5</td>
<td>0.15</td>
</tr>
<tr>
<td>Trial 3</td>
<td>0.08</td>
<td>0.04</td>
<td>0.9</td>
<td>0.1</td>
<td>0.08</td>
</tr>
</tbody>
</table>

\(\text{ARR}_{\text{III}} = 2 \text{ ARR}_{\text{II}}\)
Ecologic bias

- **at trial level:**
  Overall absolute risk reduction decreases as the proportion of stage III patients increases, suggesting that treatment effect is larger in lower risk patients:
  \[ \text{ARR}_{\text{III}} < \text{ARR}_{\text{II}} \]

- **at individual level:**
  Absolute risk reduction is in fact **twice** larger in stage III patients than in stage II!
  \[ \text{ARR}_{\text{III}} = 2 \text{ARR}_{\text{II}} \]
Baseline risk bias

Meta-regression on baseline risk (in which the units of analysis are trials) is *always* subject to bias and may imply incorrect relationship between treatment effect and prognostic factors

Baseline risk bias
Baseline risk bias

ARR = $R_T - R_C$, absolute risk reduction, where

$R_C$, risk in control group (= baseline risk)

$R_T$, risk in treated group

ARR = $\alpha + \beta R_C$

Even when $\beta = 0$, $E(\beta) = -\sigma_w^2 / (\sigma_b^2 + \sigma_w^2)$

where $\sigma_w^2$ is (common) variance within trials

and $\sigma_b^2$ is variance between trials
Baseline risk bias

- **at trial level:**
  Treatment benefit increases as baseline risk increases, suggesting that treatment effect is larger in higher risk patients.

- **at individual level:**
  Treatment benefit is in fact exactly the same in high risk as in low risk patients.
Multi-level information

**Trial level** ($i = 1 \ldots I$):
Summary statistics on endpoints and treatment effects

**Individual level** ($j = 1 \ldots Ji$)
Observations of individual endpoints
Validation of surrogate endpoints

Randomized treatment

Intermediate endpoint, potential surrogate

True endpoint

Ref: Buyse and Molenberghs, Biometrics 1998, 54:1014
Validation of surrogate endpoints

Effect of surrogate on true endpoint
Validation of surrogate endpoints

Effect of treatment on surrogate

T

S

Trt
Validation of surrogate endpoints

Effect of treatment on true endpoint
Effect of treatment on surrogate

Effect of surrogate on true endpoint

Effect of treatment on true endpoint must be fully captured by surrogate

Proportion of effect explained

Effect of treatment on surrogate

Effect of surrogate on true endpoint

Effect of treatment on true endpoint must be fully captured by surrogate

“Full capture of effect” in fact implies perfect individual-level correlation
Multilevel approach

Effects of treatment on surrogate and on true endpoint must be correlated

Ref: Buyse et al, Biostatistics 2000;1:49;
Gail, Pfeiffer and van Houwelingen, Biostatistics 2000;1:231.
Multilevel approach

This individual-level correlation can be shown in a single trial.

Effects of treatment on surrogate and on true endpoint must be correlated.
Multilevel approach

Effects of treatment on surrogate and on true endpoint must be correlated.

This trial-level correlation requires several trials.

Surrogate and true endpoint must be correlated.
General model

Model (assuming S and T normally distributed):

\[
S_{ij} | Z_{ij} = \mu_{Si} + \alpha_i Z_{ij} + \varepsilon_{Sij}.
\]

\[
T_{ij} | Z_{ij} = \mu_{Ti} + \beta_i Z_{ij} + \varepsilon_{Tij}.
\]

\[
\Sigma = \begin{pmatrix}
\sigma_{SS} & \sigma_{ST} \\
\sigma_{TS} & \sigma_{TT}
\end{pmatrix}
\]
Correlation between endpoints

Individual-level surrogacy:

\[ R^2_{\text{indiv}} = \frac{\sigma_{ST}^2}{\sigma_{SS} \sigma_{TT}} \]
Random effects model

\[
\begin{pmatrix}
\mu_{si} \\
\mu_{ti}
\end{pmatrix} = \begin{pmatrix}
\alpha_i \\
\beta_i
\end{pmatrix} + \begin{pmatrix}
\mu_{si} \\
\mu_{ti}
\end{pmatrix}
\]

\[
P = \begin{pmatrix}
dss & dsa & dtt & dab & dbb \\
dsr & dsb & dtt & dab & dbb \\
dta & dtt & dtt & dab & dbb \\
dab & dtt & dtt & dtt & dbb
\end{pmatrix}
\]
Correlation between treatment effects

Trial-level surrogacy:

\[ R_{\text{trial(f)}}^2 = \left( \begin{array}{c} d_{sb} \\ d_{ab} \end{array} \right)^T \left( \begin{array}{cc} d_{ss} & d_{sa} \\ d_{sa} & d_{aa} \end{array} \right)^{-1} \left( \begin{array}{c} d_{sb} \\ d_{ab} \end{array} \right) \]
Early colorectal cancer: DFS as a surrogate for survival

- 43 treatment arms in 18 randomized trials (20,898 patients)
  - 9 surgery alone control groups
  - 34 5FU-based experimental treatment groups
- Endpoints: disease-free survival and survival

Ref: Sargent et al, JCO 2005;23:8664.
Early colorectal cancer: DFS as a surrogate for survival

Correlation between DFS hazard ratio and OS hazard ratio

Association between 3-yr DFS and 5-yr OS

Ref: Sargent et al, JCO 2005;23:8664.
Correlation between summary statistics on endpoints

\[ R^2 = 0.89 \]
Correlation between treatment effects

\[ R^2 = 0.90 \]
Early colorectal cancer: DFS as a surrogate for survival

- Remarkably high correlations between endpoints \textit{as well as} treatment effects

- However,
  - Individual-level surrogacy based on summary statistics \textit{is} loss of information important, and are results sensitive to choice?
  - Uncertainty in summary statistics is ignored: \textit{is} correlation over-estimated?
  - Only 5FU-based treatments
Advanced colorectal cancer: PFS as a surrogate for survival

- 13 trials on 4,352 patients
- Treatments (5FU/LV common arm):
  - Training set (10 trials, 3,089 patients)
    5FU Bolus or ralitrexed vs. 5FU/L
  - Validation set (3 trials, 1,263 patients)
    oxaliplatin or irinotecan + 5FU/LV vs. 5FU/LV

Ref: Buyse et al, J Clin Oncol (in press, 2007)
Advanced colorectal cancer

Correlation between PFS hazard ratio and OS hazard ratio

Association between PFS and OS

Overall survival
<table>
<thead>
<tr>
<th>Study</th>
<th>Events/ Patients</th>
<th>Events/ Patients</th>
<th>Hazard Ratio</th>
<th>95% CI</th>
<th>Hazard Ratio</th>
</tr>
</thead>
</table>
| (a) Historical trials
| 5FU+LV | 5FU |
| Crema           | 49 / 49         | 99 / 99         | 0.26         | [0.16-0.41] |              |
| Crema           | 49 / 49         | 99 / 99         | 0.42         | [0.28-0.64] |              |
| NCCTG           | 69 / 69         | 137 / 137       | 0.66         | [0.48-0.91] |              |
| NCCTG           | 68 / 69         | 136 / 137       | 0.78         | [0.58-1.08] |              |
| Siena           | 91 / 91         | 94 / 94         | 0.86         | [0.64-1.15] |              |
| Siena           | 91 / 91         | 94 / 94         | 0.68         | [0.5-0.92]  |              |
| EORTC           | 162 / 164       | 315 / 326       | 0.86         | [0.71-1.05] |              |
| EORTC           | 142 / 164       | 283 / 326       | 0.93         | [0.76-1.14] |              |
| SWOG            | 93 / 93         | 176 / 178       | 0.89         | [0.69-1.15] |              |
| SWOG            | 93 / 93         | 176 / 178       | 0.96         | [0.75-1.25] |              |
| SAKK            | 154 / 157       | 150 / 152       | 0.92         | [0.74-1.16] |              |
| SAKK            | 153 / 157       | 150 / 152       | 0.89         | [0.71-1.12] |              |
| HECOG           | 58 / 67         | 61 / 68         | 0.95         | [0.66-1.36] |              |
| Subtotal (PFS)  | 676 / 690       | 1032 / 1054     | 0.81         | [0.73-0.9]  |              |
| Subtotal (OS)   | 643 / 690       | 991 / 1054      | 0.84         | [0.76-0.93] |              |
| (b) Historical trials
| 5FU+LV | Raltitrexed |
| TCCSG-US        | 205 / 216       | 192 / 206       | 0.62         | [0.51-0.76] |              |
| TCCSG-US        | 163 / 216       | 135 / 206       | 0.74         | [0.59-0.92] |              |
| TCCSG-EU1       | 209 / 246       | 195 / 243       | 0.75         | [0.61-0.91] |              |
| TCCSG-EU1       | 123 / 246       | 119 / 243       | 0.87         | [0.67-1.12] |              |
| TCCSG-EU2       | 211 / 222       | 204 / 212       | 0.93         | [0.77-1.13] |              |
| TCCSG-EU2       | 165 / 222       | 152 / 212       | 1            | [0.81-1.25] |              |
| Subtotal (PFS)  | 625 / 684       | 591 / 661       | 0.76         | [0.68-0.86] |              |
| Subtotal (OS)   | 451 / 684       | 407 / 661       | 0.87         | [0.76-0.99] |              |
| (c) Validation trials
| 5FU+LV | 5FU+LV+Irinotecan |
| Irinotecan-EU   | 161 / 188       | 159 / 199       | 1.32         | [1.05-1.64] |              |
| Irinotecan-EU   | 136 / 188       | 126 / 199       | 1.31         | [1.02-1.67] |              |
| Irinotecan-US   | 209 / 226       | 201 / 230       | 1.22         | [1-1.48]    |              |
| Irinotecan-US   | 184 / 226       | 166 / 230       | 1.24         | [1-1.53]    |              |
| Subtotal (PFS)  | 370 / 414       | 360 / 429       | 1.26         | [1.09-1.46] |              |
| Subtotal (OS)   | 320 / 414       | 294 / 429       | 1.27         | [1.08-1.49] |              |
| (d) Validation trial
| 5FU+LV | 5FU+LV+Oxaliplatin |
| Oxaliplatin-EU  | 177 / 210       | 162 / 210       | 1.53         | [1.23-1.9]  |              |
| Oxaliplatin-EU  | 131 / 210       | 120 / 210       | 1.21         | [0.94-1.55] |              |
| Subtotal (PFS)  | 177 / 210       | 162 / 210       | 1.53         | [1.23-1.9]  |              |
| Subtotal (OS)   | 131 / 210       | 120 / 210       | 1.21         | [0.94-1.55] |              |
Individual level - summary statistics

% overall survival at 12 months vs. % progression-free survival at 6 months

\( \rho = 0.32 \) (low!)
Individual level - PFS and OS curves

\[ \rho = 0.82 \text{ (high) with Hougaard copula} \]
Trial level correlation between effects

Treatment effect (HR) on overall survival

Historical trials
Irinotecan-EU
Irinotecan-US
Oxaliplatin-EU

Prediction line
95% Prediction limits

R² = 0.99
Surrogate threshold effect

The “Surrogate Threshold Effect” is the treatment effect on the surrogate that would predict a statistically significant treatment effect on the true endpoint.

Ref: Burzykowski and Buyse, Pharmaceutical Statist 2006;5:173.
Surrogate threshold effect

Treatment effect (HR) on overall survival

0.25 0.5 0.75 1.0 1.25 1.5 1.75

0.25 0.5 0.75 1.0 1.25 1.5 1.75

Surrogate threshold effect

95% Prediction limits

Historical trials

Irinotecan-EU

Oxaliplatin-EU

Irinotecan-US

Prediction line

0.86

1.16
Advanced colorectal cancer

- Remarkably high correlations between endpoints as well as treatment effects, including for new drugs oxaliplatin & irinotecan

- However,
  - Individual-level surrogacy based on summary statistics does not show a strong association ➔ correlation based on individual patient data
  - Influential point for trial-level association ➔ is trial-level association over-estimated?
Conclusion

“If a problem is very important and needs an answer, statisticians should not just make do by calculating a weighted estimate based on published data (…) Presenting poor-quality estimates may mislead the reader and the public health practitioner into believing the estimates are based on good science. That is not the case.”

If a meta-analysis is worth doing at all, it is worth doing well.
Conclusion

If a meta-analysis is worth doing at all, it is worth doing well.

But …

Disadvantages of IPD meta-analyses

- Time
- Cost
- Feasibility

SLIDE # 3
The ONLY satisfactory solution (in the long run…)

*Individual patient data from all randomized clinical trials should fall in the public domain after trial “completion”*

- Re-analyses possible (meta-analyses / detection of low level signals)
- Precedent of clinical trial registration
- Feasible (Gene Expression Omnibus etc.)
- Technically easy (CDISC etc.)
- Ethical