A systematic review of methodology developed for meta-analyses with time-to-event outcomes

Sarah Nolan
Department of Biostatistics
Sarah.Nolan@liv.ac.uk
Systematic reviews and Meta-analysis

Summary of results of a number of independent RCTs which address the same or similar questions.

**Meta-analysis** is a statistical technique used to synthesise the results and obtain a single pooled estimate of overall relative treatment effect

- **Fixed effects or Random effects**

- Increases sample size, precision and power
- Minimises the likelihood of a chance result
- Provides more information on effect size which single studies would not have the power to detect
Aggregate Data (AD) Meta-Analysis

Summary statistics extracted from published literature
• Quick and inexpensive to perform
• Rely on authors publishing sufficient information to perform a meta-analysis

Individual Patient Data (IPD) Meta-Analysis

Original data collected from trial authors and re-analysis performed
• Consistency checks and further analysis / updates
• Resource intensive and time consuming
• ‘Gold standard’ for time to event outcomes
AD or IPD Meta-Analysis?

Inconsistency in reporting survival outcomes in trials Log (HR) and Var (Log (HR)) required for Meta-Analysis

- 11 papers didn’t report total number of patients
- 62% didn’t clearly define an endpoint
- 50% of papers gave no information about extent of follow up
- Less than a sixth reported any estimate of treatment effect
- Over 90% of papers presented survival plots and p-values

- Around 95% of papers presented survival plots and p-values
- 74% of papers reported number of events
- 52% of papers reported a hazard ratio
Objective
Perform a methodology review of existing methods developed for the meta-analysis of both aggregate data and individual patient data with time-to-event outcomes

- Published and unpublished studies (inc. abstracts)
- All study designs (not restricted to RCTs)
- References of included studies

Databases
MEDLINE (PubMed) and Cochrane Methodology Register

Search strategy
Key words: meta-analysis, survival, time-to-event, failure time, review methodology, statistical models
Inclusion Criteria

- Meta-analytic methods only
- Methods for analysis of a single trial are excluded
- Methods related to network meta-analysis (NMA) or mixed treatment comparison (MTC) are excluded
- Time-to-event (censored) outcomes only
- Binary or count data outcomes related to survival or mortality without censoring are excluded
- Methodology papers only (development or discussion of statistical methods)
- Applied papers with ‘standard’ methods are excluded
Database Searches
183 Studies

Abstracts
152 studies

Full papers
105 studies

Eligible Studies
70 studies

Duplicates Removed
31 studies

Excluded
47 studies

Excluded
34 studies

Exclusion Reasons
82 studies
29 not methodology
19 not censored data
32 not meta-analysis
2 other

References of included studies
18 studies

Total: 88 studies
### Included Studies

<table>
<thead>
<tr>
<th>Methodology</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>General methodology</td>
<td>11</td>
</tr>
<tr>
<td>Aggregate data models</td>
<td>6</td>
</tr>
<tr>
<td>Individual patient data models</td>
<td>18</td>
</tr>
<tr>
<td>Combining whole survival curves</td>
<td>12</td>
</tr>
<tr>
<td>Multivariate data</td>
<td>12</td>
</tr>
<tr>
<td>Indirect methods of estimation</td>
<td>14</td>
</tr>
<tr>
<td>Alternatives to Hazard Ratios</td>
<td>14</td>
</tr>
<tr>
<td>Non-Randomised studies</td>
<td>4</td>
</tr>
<tr>
<td>Extension methods (QoL and CEA)</td>
<td>7</td>
</tr>
</tbody>
</table>
Included Studies

General methodology (12 studies)
• Older papers (1970-1998)
• Use and efficiency of the log rank test

Aggregate Data Models (6 studies)
• Comparisons of parametric distributions
• Use of non-parametric modelling
• Models for non-proportional hazards

Non randomised studies (4 studies)
• Use of historical (non-randomised) controls
• Meta-analysis of epidemiological survival studies
Combining whole survival curves (11 studies)

Pooling data at a single fixed time ignores the shape of the whole survival curve

Problems combining trials with variable follow up times
  o Cannot be assumed baseline risk and treatment effects are constant over the whole time frame

Earle et al (2000) compare five methods

Choice of method depends on scope of problem
  o Interest in long or short term survival?
  o Interest in comparing groups or overall survival?
Indirect methods (14 studies)

Methods of estimating indirectly log(Hazard Ratio) and variance


**Numerical estimations**
- Observed / Expected numbers of events in each group
- Total number of patients in each group
- Log Rank test p-value

**Graphical estimations (survival curves)**
- Numbers of events and effective numbers at risk read directly from the published survival curve
- Dependent on quality of the printed curve
- Dependent on assumptions of censoring and event rates
## Included Studies

<table>
<thead>
<tr>
<th>Alternatives to Hazard Ratios</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log (Odds Ratio) or Log (Relative Risk)</td>
<td>8</td>
</tr>
<tr>
<td>Pooled Median Survival times</td>
<td>3</td>
</tr>
<tr>
<td>P Values</td>
<td>1</td>
</tr>
<tr>
<td>Other (Percentile Ratio, EQuRR*)</td>
<td>2</td>
</tr>
</tbody>
</table>

*EQuRR : Extreme Quartile Risk Ratio (Ioannidis and Lau 1998)*

<table>
<thead>
<tr>
<th>Individual patient data models</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modelling Covariates / Interactions</td>
<td>4</td>
</tr>
<tr>
<td>Heterogeneity</td>
<td>8</td>
</tr>
<tr>
<td>Random effects</td>
<td>5</td>
</tr>
<tr>
<td>Comparisons with AD models</td>
<td>5</td>
</tr>
</tbody>
</table>
Multivariate survival data (12 studies)

Clinical practice is rarely decided based on a single outcome at a single time point

Methods are required to analyse multiple outcomes at multiple time points to estimate an overall treatment effect over time

• Short and long term survival
• Disease free and overall survival
• Surrogate outcomes
• Account for correlation between related outcomes
• Fixed or random effects
• Models for IPD or AD
Findings so far

• Log(Hazard Ratio) and variance required for meta-analysis are often not reported

• Indirect methods of estimating Log(HR) and Var(Log(HR)) from published summary statistics have been developed

• Older papers often approximate Log(HR) with Log(OR) or Log(RR)

• Recent methodological developments focus on ‘gold standard’ IPD meta-analysis
What’s next?
Continue reviewing included studies

• Which methods are used in practice?
• What are the current problems?
• Where are the ‘gaps’ in methodology / areas requiring future research?
Acknowledgements

PhD supervisors: Dr Catrin Tudur Smith
Prof Tony Marson

Second Reviewer: Jennifer Pulman

References

3) Ioannidis JP and Lau J. Heterogeneity of the baseline risk with patient populations of clinical trials: a proposed evaluation algorithm