Practical methods to pool multi-study joint longitudinal and time to event data

Maria Sudell, Dr Ruwanthi Kolamunnage-Dona, Dr Catrin Tudur Smith
University of Liverpool
mesudell@liverpool.ac.uk
Joint longitudinal and time-to-event data (single study)

- Methods to simultaneously model potentially related longitudinal and time-to-event data
- Can produce less biased more efficient results than standalone cases where linked longitudinal and time-to-event data exists

Longitudinal Sub-model
\[ Y_{ij} = g(W_1, \theta_1) + \varepsilon_{ij} \]
\[ Y_{ij} = X_1\beta_1 + Z_i^{(2)}b_i^{(2)} + \varepsilon_{ij} \]

Association Structure
\[ W_1 \propto W_2 \]

Time-to-event Sub-model
\[ \lambda_i(t) = h(W_2, \theta_2) \]
\[ \lambda_i(t) = \lambda_0(t) \exp(X_2\beta_2 + W_2) \]
Meta-Analysis (MA)

- Systematic pooling of results from multiple studies
- Allows increased precision, identification of effect sizes too small to be identified in single studies, and allows questions additional to those originally posed in the data to be answered
- Gold standard – Individual Participant/Patient Data (IPD) meta-analyses, where data for each individual recorded in studies identified in the meta-analysis is available.
Joint longitudinal and time-to-event data (multi-study)

- Data available from multiple studies
- Clustering of data within studies must be accounted for (e.g. through random effects, interaction terms, stratified baseline hazard)

Longitudinal Sub-model

\[ Y_{kj} = g(W_1, \theta) + \varepsilon_{kj} \]

\[ Y_{kij} = X_1\beta_1 + Z_{k}b^{(2)}_{k} + Z_{k}^{(3)}b^{(3)}_{k} + \varepsilon_{kij} \]

Time-to-event Sub-model

\[ \lambda_{ki}(t) = h(W_2, \alpha) \]

\[ \lambda_{ki}(t) = \lambda_0(t) \exp(X_2\beta + W_2) \]

Association Structure

\[ W_1 \propto W_2 \]
Approaches to modelling multi-study IPD joint data

• Two main approaches – one stage or two stage

• Two stage approaches
  • Separate joint models fitted to data from each study
  • Results from each study pooled using standard meta-analytic techniques

• One stage approaches
  • Joint model fitted to meta-dataset (containing data from all studies)
  • Clustering of data must be accounted for
Real Data – subset of the INDANA dataset

- IPD from multiple studies investigating the effect of no treatment versus any treatment for hypertensive patients

- Longitudinal data measured at baseline, 6 months, then annually thereafter to maximum of 7 years. Measurement patterns varied between studies

- Examining longitudinal outcome systolic blood pressure and time-to-event outcome time to death

- Evidence of a changepoint in the data at 6 month, so \( \exp(-3 * time) \) term included in the model
Two stage methods
Two stage methods - overview

Stage 1: Joint model fitted to data from each study

Longitudinal Sub-model

\[ Y_{ki} = \beta_{10} + \beta_{11} time_{kij} + \beta_{12} \exp(-3 \times time_{kij}) + \beta_{13} treat_{ki} + b_{0ki}^{(2)} + b_{1ki}^{(2)} time_{kij} + \varepsilon_{kij} \]

Association Structure

\[ W_{ki}^{(2)} = \alpha^{(2)} (b_{0ki}^{(2)} + b_{1ki}^{(2)} time_{kij}) \]

Time-to-event Sub-model

\[ \lambda_{ki}(t) = \lambda_0(t) \exp(\beta_{21} treat_{ki} + W_{ki}^{(2)}) \]

Stage 2: Study specific parameters pooled using standard meta-analytic techniques

- Inverse variance method used (DerSimonian method used for random meta-analyses)
- Both fixed and random effects meta-analyses fitted and compared
- Separate meta-analyses for each parameter of interest
Two stage methods – real data

### Separate Longitudinal model

<table>
<thead>
<tr>
<th>Study</th>
<th>TE</th>
<th>seTE</th>
<th>Mean difference</th>
<th>MD</th>
<th>95% CI</th>
<th>Weight (fixed)</th>
<th>Weight (random)</th>
</tr>
</thead>
<tbody>
<tr>
<td>COOP</td>
<td>-10.90</td>
<td>1.1280</td>
<td>-10.90</td>
<td>-13.11; -8.69</td>
<td>1.6%</td>
<td>13.6%</td>
<td></td>
</tr>
<tr>
<td>EWPFH</td>
<td>-11.38</td>
<td>1.0509</td>
<td>-11.38</td>
<td>-13.44; -9.32</td>
<td>1.0%</td>
<td>14.1%</td>
<td></td>
</tr>
<tr>
<td>MRC1</td>
<td>-8.18</td>
<td>0.2015</td>
<td>-8.18</td>
<td>-8.57; -7.78</td>
<td>51.3%</td>
<td>18.8%</td>
<td></td>
</tr>
<tr>
<td>MRC2</td>
<td>-10.06</td>
<td>0.3334</td>
<td>-10.06</td>
<td>-11.31; -8.01</td>
<td>18.7%</td>
<td>18.4%</td>
<td></td>
</tr>
<tr>
<td>SHEP</td>
<td>-5.99</td>
<td>0.8089</td>
<td>-5.99</td>
<td>-7.59; -4.38</td>
<td>21.8%</td>
<td>18.5%</td>
<td></td>
</tr>
<tr>
<td>STOP</td>
<td>-11.94</td>
<td>0.6760</td>
<td>-11.94</td>
<td>-13.27; -10.62</td>
<td>4.6%</td>
<td>16.6%</td>
<td></td>
</tr>
</tbody>
</table>

Fixed effect model
Random effects model
Heterogeneity: $I^2 = 95\%, \tau^2 = 3.209, p < 0.01$

### Joint model

<table>
<thead>
<tr>
<th>Study</th>
<th>TE</th>
<th>seTE</th>
<th>Mean difference</th>
<th>MD</th>
<th>95% CI</th>
<th>Weight (fixed)</th>
<th>Weight (random)</th>
</tr>
</thead>
<tbody>
<tr>
<td>COOP</td>
<td>-10.90</td>
<td>1.1716</td>
<td>-10.90</td>
<td>-13.19; -8.00</td>
<td>1.5%</td>
<td>13.4%</td>
<td></td>
</tr>
<tr>
<td>EWPFH</td>
<td>-11.38</td>
<td>1.2027</td>
<td>-11.38</td>
<td>-13.74; -8.03</td>
<td>1.4%</td>
<td>13.2%</td>
<td></td>
</tr>
<tr>
<td>MRC1</td>
<td>-8.18</td>
<td>0.1356</td>
<td>-8.18</td>
<td>-8.65; -7.70</td>
<td>54.7%</td>
<td>19.2%</td>
<td></td>
</tr>
<tr>
<td>MRC2</td>
<td>-10.68</td>
<td>0.3356</td>
<td>-10.68</td>
<td>-11.32; -10.00</td>
<td>19.2%</td>
<td>18.6%</td>
<td></td>
</tr>
<tr>
<td>SHEP</td>
<td>-6.99</td>
<td>0.3185</td>
<td>-6.99</td>
<td>-7.61; -6.36</td>
<td>20.3%</td>
<td>18.8%</td>
<td></td>
</tr>
<tr>
<td>STOP</td>
<td>-11.94</td>
<td>0.7403</td>
<td>-11.94</td>
<td>-13.39; -10.49</td>
<td>3.7%</td>
<td>16.5%</td>
<td></td>
</tr>
</tbody>
</table>

Fixed effect model
Random effects model
Heterogeneity: $I^2 = 95\%, \tau^2 = 3.854, p < 0.01$
Two stage methods – real data
Two stage methods – real data

**Association parameter**

**Joint model**

<table>
<thead>
<tr>
<th>Study</th>
<th>TE</th>
<th>seTE</th>
<th>Mean difference</th>
<th>MD</th>
<th>95%-CI</th>
<th>Weight (fixed)</th>
<th>Weight (random)</th>
</tr>
</thead>
<tbody>
<tr>
<td>COOP</td>
<td>-0.00</td>
<td>0.0080</td>
<td></td>
<td>-0.00</td>
<td>[-0.02; 0.01]</td>
<td>8.8%</td>
<td>15.7%</td>
</tr>
<tr>
<td>EWPH</td>
<td>0.01</td>
<td>0.0054</td>
<td></td>
<td>0.01</td>
<td>[0.00; 0.02]</td>
<td>19.4%</td>
<td>18.2%</td>
</tr>
<tr>
<td>MRC1</td>
<td>0.03</td>
<td>0.0050</td>
<td></td>
<td>0.03</td>
<td>[0.02; 0.04]</td>
<td>22.6%</td>
<td>18.6%</td>
</tr>
<tr>
<td>MRC2</td>
<td>-0.00</td>
<td>0.0064</td>
<td></td>
<td>-0.00</td>
<td>[-0.01; 0.01]</td>
<td>13.8%</td>
<td>17.3%</td>
</tr>
<tr>
<td>SHEP</td>
<td>0.01</td>
<td>0.0042</td>
<td></td>
<td>0.01</td>
<td>[0.00; 0.02]</td>
<td>32.1%</td>
<td>19.2%</td>
</tr>
<tr>
<td>STOP</td>
<td>-0.01</td>
<td>0.0131</td>
<td></td>
<td>-0.01</td>
<td>[-0.04; 0.01]</td>
<td>3.3%</td>
<td>11.0%</td>
</tr>
</tbody>
</table>

**Fixed effect model**

0.01 [0.01; 0.02] 100.0%  --

**Random effects model**

0.01 [-0.01; 0.02]  --  100.0%

Heterogeneity: $I^2 = 84\%$, $Q^2 = 0.0002$, $p < 0.01$
Two stage methods – simulations (longitudinal treatment effect coefficient)
Two stage methods – simulations (time-to-event treatment effect coefficient)
Two stage methods - recommendations

Preliminary work

• For each study:
  • Plot longitudinal trajectories separately for those experiencing an event and those censored.
  • Produce Kaplan-Meier plots for e.g. each treatment group
  • Use plots to assess whether an association between longitudinal and time-to-event outcomes is feasible
  • Use plots and clinical background of the data to select:
    • Longitudinal sub-model
    • Time-to-event sub-model
    • Association structure
Two stage methods - recommendations

First Stage

- Group studies such that chosen model structure within each group is identical.
- Within each group, fit identical joint models to data from each study. Model structures can differ between groups.
Two stage methods - recommendations

Second Stage

• For each study extract model parameters, precision estimates and sample size
• Pool estimates within groups using standard MA techniques.
One stage methods
One stage methods - overview

- Same model basic model specification as first stage of two stage work, but now additional terms included to account for between study heterogeneity

<table>
<thead>
<tr>
<th>Group</th>
<th>Method to account for between study heterogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Between study heterogeneity ignored</td>
</tr>
<tr>
<td>1</td>
<td>Fixed interaction term between treatment and study in each sub-model</td>
</tr>
<tr>
<td>2</td>
<td>Fixed study indicator in longitudinal sub-model, study level random treatment effect</td>
</tr>
<tr>
<td>3</td>
<td>Study level random intercept and random treatment effect</td>
</tr>
<tr>
<td>4</td>
<td>Fixed interaction term between treatment and study in longitudinal sub-model, baseline hazard stratified by study</td>
</tr>
<tr>
<td>5</td>
<td>Fixed study indicator in longitudinal sub-model, study level random treatment effect, baseline hazard stratified by study</td>
</tr>
</tbody>
</table>
One stage methods – real data

**Longitudinal Treatment Effect**

### Joint Models

- **Model Group** (Group 0 to Group 5)
- Treatment Effect Estimate

### Separate Models

- **Model Group** (Group 0 to Group 3)
- Treatment Effect Estimate

**Estimate Type**
- Overall
- COOP
- EWPHE
- MRC1
- MRC2
- SHEP
- STOP
One stage methods – real data

Time-to-Event Treatment Effect

Joint Models

Separate Models

Model Group

Model Group

Estimate Type
- Overall
- COOP
- EWPHE
- MRC1
- MRC2
- SHEP
- STOP
One stage methods – real data

![Graph showing association parameter estimates for different groups.](image-url)
Conclusions

• Care must be taken during two stage meta-analyses of joint data to pool only parameters with comparable interpretations

• A variety of methods exist to model multi-study joint data in a one stage analyses, however some may not be appropriate unless the number of studies in the meta-analysis is over a given threshold

• Functions for analysis of multi-study joint data available in R package joineRmeta
Thank you for listening.
Any questions?