Multivariate Bayesian models for longitudinal patient-reported outcomes and survival data in cancer clinical trials

Laura A. Hatfield
Division of Biostatistics
School of Public Health
University of Minnesota

November 8, 2010

Joint work with:

Bradley P. Carlin, Mark E. Boye, Michelle D. Hackshaw
Motivation

Malignant Pleural Mesothelioma (MPM)

- 2nd most important occupational cancer among industrial workers
- Treatment options: chemotherapy, supportive care, surgery, radiation
- Pemetrexed (Alimta) with cisplatin is only FDA approved first-line therapy (Vogelzang et al. 2003)
Motivation

Malignant Pleural Mesothelioma (MPM)

- 2nd most important occupational cancer among industrial workers
- Treatment options: chemotherapy, supportive care, surgery, radiation
- Pemetrexed (Alimta) with cisplatin is only FDA approved first-line therapy (Vogelzang et al. 2003)

How does treatment affect symptoms and survival?

- FDA wants PROs to supplement survival (Lipscomb et al. 2007)
- May consider PROs in cancer drug approval (Rock et al. 2007)
Clinical Trial Data: Observation Schedule

### Control

<table>
<thead>
<tr>
<th>Days</th>
<th>Patient</th>
<th>PRO</th>
<th>Progression</th>
<th>Death</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>200</td>
<td>400</td>
<td>600</td>
<td>800</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>15</td>
<td>20</td>
<td>25</td>
</tr>
</tbody>
</table>

### Treatment

<table>
<thead>
<tr>
<th>Days</th>
<th>Patient</th>
<th>PRO</th>
<th>Progression</th>
<th>Death</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>200</td>
<td>400</td>
<td>600</td>
<td>800</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>15</td>
<td>20</td>
<td>25</td>
</tr>
</tbody>
</table>
Symptoms

Patient-Reported Outcomes (PROs) on a visual analog scale

How much fatigue do you have?

None                 As much as it could be

► **6 symptoms:** anorexia, fatigue, cough, dyspnea, hemoptysis, and pain

► **3 global measures:** symptom distress, interference with carrying out normal activities, and quality of life
Zero Inflated PROs

Fatigue Score at Baseline

Density

0 0.5 1.0 1.5 2.0 2.5 3.0 3.5

0 20 40 60 80 100

Fatigue Score at Baseline
Zero Inflated PROs

Fatigue Score at Baseline

Density

Density vs Fatigue Score at Baseline
Zero Inflated PROs

Fatigue Score at Baseline

Density

$\text{Beta}(1,1.7)$
PRO Items at Baseline
PRO Items at Baseline

Interference and anorexia
weakly correlated

Interference and QoL
strongly correlated
Single PRO Item (Fatigue) Across Time
Single PRO Item (Fatigue) Across Time

Time 4 and Baseline strongly correlated

Time 4 and Time 3 strongly correlated
Survival Outcome: Progression Free Survival (PFS)

Proportion Alive and Unprogressed

<table>
<thead>
<tr>
<th></th>
<th>K–M Ctrl</th>
<th>K–M Trt</th>
<th>Weibull Ctrl</th>
<th>Weibull Trt</th>
</tr>
</thead>
</table>

Months

PFS

Proportion Alive and Unprogressed vs Months

K–M Ctrl, K–M Trt, Weibull Ctrl, Weibull Trt
Covariates

- Race (= 0 if white/Caucasian)
- Gender (= 0 if male)
- Dichotomized Karnofsky performance score: (= 0 if 90 – 100)
- Dichotomized stage of disease (= 0 if Stage I/II)
- Vitamin supplementation (= 0 if fully supplemented)
- Treatment group (= 0 if cisplatin only)
- Time (days): std’ized by mean (75) and std dev (69)
Notation

\( Y_i(s) \)  observed longitudinal outcomes,  \( i^{th} \) subject at time  \( s \)

\( X_i(s) \)  observed covariates

\( U_i(s) \)  unobserved (latent) process

\( T_i \)  event or censoring time, where  \( \delta_i = 1 \)  if censored
Longitudinal Model: Zero-inflation

Symptom probability model: $Pr[Y_i(s) > 0]$
Longitudinal Model: Zero-inflation

Symptom probability model: $Pr[Y_i(s) > 0]$
Symptom severity model: $E[Y_i(s)|Y_i(s) > 0]$
Longitudinal Model: Zero-inflation

Symptom probability model: \( Pr[Y_i(s) > 0] \)
Symptom severity model: \( E[Y_i(s) | Y_i(s) > 0] \)

Zero-Inflation Literature

**Count data**
- Binomial
- Poisson
Longitudinal Model: Zero-inflation

Symptom probability model: \( Pr[Y_i(s) > 0] \)
Symptom severity model: \( E[Y_i(s)|Y_i(s) > 0] \)

Zero-Inflation Literature

**Count data**
- Binomial
- Poisson

**Continuous data**
- Probit/logistic + lognormal
  (Zhang et al. 2006; Robinson et al. 2006)
- Logistic + beta
  (our work)
Zero Inflated Beta (ZIB) model:

\[ X \sim ZIB(\omega, \mu, \phi), \ X \in \{0\} \cup (0, 1) \]

- \( \omega \): probability of \( X \in (0, 1) \)
- \( \mu \): mean of \( X \in (0, 1) \)
- \( \phi \): dispersion of \( X \in (0, 1) \)
Longitudinal Model

Zero Inflated Beta (ZIB) model:

\[ X \sim ZIB(\omega, \mu, \phi), \; X \in \{0\} \cup (0, 1) \]
\[ \omega = \text{probability of } X \in (0, 1) \]
\[ \mu = \text{mean of } X \in (0, 1) \]
\[ \phi = \text{dispersion of } X \in (0, 1) \]

- \( X \sim ZIB(\omega, \mu, \phi) \) corresponds to \( X = ZB \)
- \( Z \sim Bernoulli(\omega) \perp B \sim Beta(\mu\phi, (1 - \mu)\phi) \)
Zero Inflated Beta (ZIB) model:

\[ X \sim ZIB(\omega, \mu, \phi), \; X \in \{0\} \cup (0, 1) \]
\[ \omega = \text{probability of } X \in (0, 1) \]
\[ \mu = \text{mean of } X \in (0, 1) \]
\[ \phi = \text{dispersion of } X \in (0, 1) \]

- \( X \sim ZIB(\omega, \mu, \phi) \) corresponds to \( X = ZB \)
- \( Z \sim Bernoulli(\omega) \perp B \sim Beta(\mu \phi, (1 - \mu) \phi) \)
- \( E(B) = \mu, \; Var(B) = \frac{\mu(1-\mu)}{\phi+1} \)
- \( E(X) = \omega \mu, \; Var(X) = \omega \mu \left[ \frac{\phi \mu + 1}{\phi + 1} - \omega \mu \right] \)
Joint Models: Conceptual Framework

[ Longitudinal ]
↑
[ Latent Processes ]
↓
[ Survival ]
Joint Models: Conceptual Framework

[ Longitudinal ]

↑

[ Latent Processes ]

↓

[ Survival ]

- Xu and Zeger (2001) for an example of a popular approach
- Tsiatis and Davidian (2004) for a comprehensive (non-Bayesian) review
Linking via value of Latent Process

Longitudinal models:
- $X_{0i}(s)\beta_0 + U_{0i}(s)$
- $X_{1i}(s)\beta_1 + U_{1i}(s)$

Latent processes:
- $U_{0i}(s) = f_0(u_{0i})$
- $U_{1i}(s) = f_1(u_{1i})$

Survival model:
- $X_{2i}\beta_2 + A(\alpha, U_{0i}(s), U_{1i}(s))$
Linking via value of Latent Process

Longitudinal models:
\[ X_{0i}(s)\beta_0 + U_{0i}(s) \]
\[ X_{1i}(s)\beta_1 + U_{1i}(s) \]

↑

Latent processes:
\[ U_{0i}(s) = f_0(u_{0i}) \]
\[ U_{1i}(s) = f_1(u_{1i}) \]

↓

Survival model:
\[ X_{2i}\beta_2 + A(\alpha, U_{0i}(s), U_{1i}(s)) \]

For example:
\[ U_{0i}(s) = u_{01i} + u_{01i}s \]
\[ U_{1i}(s) = u_{11i} + u_{12i}s \]
\[ A(\alpha, U_{0i}(s), U_{1i}(s)) = \alpha_0 U_{0i}(s) + \alpha_1 U_{1i}(s) \]

i.e., linear trajectories underlie both longitudinal and survival
Linking via Random Effects

Longitudinal models:
\[ X_{0i}(s)\beta_0 + U_{0i}(s) \]
\[ X_{1i}(s)\beta_1 + U_{1i}(s) \]

↑

Latent processes:
\[ U_{0i}(s) = f_0(u_{0i}) \]
\[ U_{1i}(s) = f_1(u_{1i}) \]

↓

Survival model:
\[ X_{2i}\beta_2 + A(\alpha, u_{0i}, u_{1i}) \]
Linking via Random Effects

Longitudinal models: \[ X_{0i}(s)\beta_0 + U_{0i}(s) \]
\[ X_{1i}(s)\beta_1 + U_{1i}(s) \]

↑

Latent processes: \[ U_{0i}(s) = f_0(u_{0i}) \]
\[ U_{1i}(s) = f_1(u_{1i}) \]

↓

Survival model: \[ X_{2i}\beta_2 + A(\alpha, u_{0i}, u_{1i}) \]

For example:

\[ U_{0i}(s) = u_{01i} + u_{02i}s \]
\[ U_{1i}(s) = u_{11i} + u_{12i}s \]
\[ A(\alpha, u_{0i}, u_{1i}) = \alpha_0 u_{02i} + \alpha_1 u_{12i} \]

i.e., only the slopes of the linear trajectories contribute to survival
Shared versus Correlated Random Effects

Shared Random Effects

Long’l: $U_{0i}(s) = f_0(\alpha_0, u_i), U_{1i}(s) = f_1(\alpha_1, u_i)$

Survival: $A(\alpha_2, u_i)$
Shared versus Correlated Random Effects

**Shared Random Effects**

Long’l: $U_{0i}(s) = f_0(\alpha_0, u_i), U_{1i}(s) = f_1(\alpha_1, u_i)$

Survival: $A(\alpha_2, u_i)$

**Correlated Random Effects**

Long’l: $U_{0i}(s) = f_0(u_{0i}), U_{1i}(s) = f_1(u_{1i})$

Survival: $A(\alpha, u_{2i})$

RE Dist’n: $(u_{0i}, u_{1i}, u_{2i})^T \sim G$
Shared versus Correlated Random Effects

Shared Random Effects

Long’l: \( U_{0i}(s) = f_0(\alpha_0, u_i), U_{1i}(s) = f_1(\alpha_1, u_i) \)

Survival: \( A(\alpha_2, u_i) \)

Correlated Random Effects

Long’l: \( U_{0i}(s) = f_0(u_{0i}), U_{1i}(s) = f_1(u_{1i}) \)

Survival: \( A(\alpha, u_{2i}) \)

RE Dist’n: \((u_{0i}, u_{1i}, u_{2i})^T \sim G\)

Hybrid Approach

Long’l: \( U_{0i}(s) = f_0(u_{0i}), U_{1i}(s) = f_1(u_{1i}) \)

Survival: \( A(\alpha, u_{0i}, u_{1i}) \)

RE Dist’n: \((u_{0i}, u_{1i})^T \sim G\)
DIC deviance information criterion

Deviance: \( D = -2 \log(p(y|\theta)) \)

Effective complexity: \( p_D = \bar{D}(\theta) - D(\bar{\theta}) \)

\( DIC = \bar{D}(\theta) + p_D \)
**Tools for Model Choice**

**DIC** deviance information criterion
- Deviance: $D = -2 \log(p(y|\theta))$
- Effective complexity: $p_D = \bar{D}(\theta) - D(\bar{\theta})$
- $DIC = \bar{D}(\theta) + p_D$

**RMSPE** root mean squared prediction error
- Validation data: $y^{(2)}$, Fitting data: $y^{(1)}$
- Posterior predictions: $y_i^*|y^{(1)}, i \in y^{(2)}$
- $RMSE = \sqrt{\frac{1}{n^{(2)}} \sum_i (y_i^* - y_i^{true})^2}$
**Tools for Model Choice**

**DIC** deviance information criterion

- Deviance: $D = -2 \log(p(y|\theta))$
- Effective complexity: $p_D = \bar{D}(\theta) - D(\bar{\theta})$
- $\text{DIC} = \bar{D}(\theta) + p_D$

**RMSPE** root mean squared prediction error

- Validation data: $y^{(2)}$
- Fitting data: $y^{(1)}$
- Posterior predictions: $y_i^*|y^{(1)}$, $i \in y^{(2)}$
- $\text{RMSE} = \sqrt{\frac{1}{n^{(2)}} \sum_i (y_i^* - y_i^{\text{true}})^2}$

**LPML** log pseudo maximum likelihood

- Conditional predictive ordinate: $CPO_i = \int f(y_i|\theta) p(\theta|y^{(1)}) \, d\theta$, $i \in y^{(2)}$
- Sum of the logs is like a log likelihood: $\text{LMPL} = \sum_i \log(CPO_i)$
### Choosing Random Effect Structure

<table>
<thead>
<tr>
<th>Model</th>
<th>DIC</th>
<th>pD</th>
<th>LPML</th>
<th>RMSPE</th>
<th>LPML</th>
<th>RMSPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>shared effects</td>
<td>2683</td>
<td>368</td>
<td>221</td>
<td>0.368</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>correl r.e.’s</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>long’l r.e.’s</td>
<td>2238</td>
<td>507</td>
<td>242</td>
<td>0.366</td>
<td>-413</td>
<td>202</td>
</tr>
<tr>
<td>hybrid</td>
<td>2212</td>
<td>512</td>
<td>244</td>
<td>0.368</td>
<td>-410</td>
<td>211</td>
</tr>
</tbody>
</table>
Choosing Random Effect Structure

<table>
<thead>
<tr>
<th>Model</th>
<th>DIC</th>
<th>pD</th>
<th>LPML</th>
<th>RMSPE</th>
<th>LPML</th>
<th>RMSPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>shared effects</td>
<td>2683</td>
<td>368</td>
<td>221</td>
<td>0.368</td>
<td>-410</td>
<td>210</td>
</tr>
<tr>
<td>correl r.e.’s</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>long’l r.e.’s</td>
<td>2238</td>
<td>507</td>
<td>242</td>
<td>0.366</td>
<td>-413</td>
<td>202</td>
</tr>
<tr>
<td>hybrid</td>
<td>2212</td>
<td>512</td>
<td>244</td>
<td>0.368</td>
<td>-410</td>
<td>211</td>
</tr>
</tbody>
</table>
Consider in four places:

\( F_0 \) Fixed in probability model: \( X_{0i}(s) \)
\( R_0 \) Random in probability model: \( U_{0i}(s) \)
\( F_1 \) Fixed in severity model: \( X_{1i}(s) \)
\( R_1 \) Random in severity model: \( U_{1i}(s) \)
Time Parameterization

Consider in four places:

- $F_0$: Fixed in probability model: $X_{0i}(s)$
- $R_0$: Random in probability model: $U_{0i}(s)$
- $F_1$: Fixed in severity model: $X_{1i}(s)$
- $R_1$: Random in severity model: $U_{1i}(s)$

Yields a compact notation:

- $Q$: quadratic
- $L$: linear
- $I$: intercept-only

$F_0 R_0 F_1 R_1$

e.g., $Q I L L$
Indicate Q trends may be needed in $F_0$ (top)
Indicate Q trends may be needed in $F_0$ (top) and $F_1$ (bottom)
Choosing a Time Parameterization

<table>
<thead>
<tr>
<th>Model</th>
<th>DIC</th>
<th>pD</th>
<th>LPML</th>
<th>RMSPE</th>
<th>Longitudinal</th>
<th>Survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>LILI</td>
<td>2212</td>
<td>512</td>
<td>244</td>
<td>0.368</td>
<td>-410</td>
<td>211</td>
</tr>
<tr>
<td>LILL</td>
<td>1768</td>
<td>708</td>
<td>337</td>
<td>0.381</td>
<td>-404</td>
<td>221</td>
</tr>
<tr>
<td>LIQL</td>
<td>1766</td>
<td>712</td>
<td>335</td>
<td>0.380</td>
<td>-403</td>
<td>220</td>
</tr>
<tr>
<td>QIPI</td>
<td>2190</td>
<td>515</td>
<td>251</td>
<td>0.367</td>
<td>-409</td>
<td>211</td>
</tr>
<tr>
<td>QIQL</td>
<td>1759</td>
<td>716</td>
<td>338</td>
<td>0.380</td>
<td>-403</td>
<td>220</td>
</tr>
</tbody>
</table>
Choosing a Time Parameterization

<table>
<thead>
<tr>
<th>Model</th>
<th>DIC</th>
<th>pD</th>
<th>LPML</th>
<th>RMSPE</th>
<th>Survival LPML</th>
<th>RMSPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>LILI</td>
<td>2212</td>
<td>512</td>
<td>244</td>
<td>0.368</td>
<td>-410</td>
<td>211</td>
</tr>
<tr>
<td>LILL</td>
<td>1768</td>
<td>708</td>
<td>337</td>
<td>0.381</td>
<td>-404</td>
<td>221</td>
</tr>
<tr>
<td>LIQL</td>
<td>1766</td>
<td>712</td>
<td>335</td>
<td>0.380</td>
<td>-403</td>
<td>220</td>
</tr>
<tr>
<td>QIQI</td>
<td>2190</td>
<td>515</td>
<td>251</td>
<td>0.367</td>
<td>-409</td>
<td>211</td>
</tr>
<tr>
<td>QIQL</td>
<td><strong>1759</strong></td>
<td>716</td>
<td><strong>338</strong></td>
<td>0.380</td>
<td><strong>-403</strong></td>
<td>220</td>
</tr>
</tbody>
</table>
Hybrid ZIB QIQL Joint Model

\[ Y_i(s) \sim ZIB(\omega_i(s), \mu_i(s), \phi) \]
\[ \text{logit}(\omega_i(s)) = X_0i(s)\beta_0 + U_0i(s) \]
\[ \text{logit}(\mu_i(s)) = X_1i(s)\beta_1 + U_1i(s) \]
\[ U_0i(s) = u_{0i1} \]
\[ U_1i(s) = u_{1i1} + u_{1i2}s \]
\[ T_i \sim \text{Weibull}(\gamma, \lambda_i) \]
\[ \log(\lambda_i) = X_2i\beta_2 + \alpha_1 u_{0i1} + \alpha_2 u_{1i1} + \alpha_3 u_{1i2} \]

- Intercepts only for probability trajectory: \( U_{0i}(s) \)
- Intercepts + slopes for severity trajectory: \( U_{1i}(s) \)
- Simple linear RE combination in survival model: \( A(\alpha, u_{0i}, u_{1i}) \)
Two Phases of Analysis

Protocol-Specified Analysis
Consider only the baseline influence of covariates

Subgroup Analysis
Consider two-way interactions (covariate-by-day) and three-way interactions (covariate-by-day-by-treatment)
Protocol Results: Log Odds Ratios for Non-zero LCSS

\[
\log(\text{OR}_{\text{cont vs ref}}) = \beta_{0,\text{cov}}
\]

i.e., (log) odds ratio of non-zero LCSS in contrast (cont) versus reference (ref) group for each covariate (cov)
Protocol Results: Difference in Mean LCSS Severity

\[ \mu_{\text{cont}} - \mu_{\text{ref}} = \left[ \frac{\exp(\beta_{11} + \beta_1, \text{cov})}{1 + \exp(\beta_{11} + \beta_1, \text{cov})} - \frac{\exp(\beta_{11})}{1 + \exp(\beta_{11})} \right] \times 100 \]

i.e., difference between contrast (cont) and reference (ref) groups on the original LCSS visual analog scale for each covariate (cov).
Protocol Results: Difference in Median PFS (months)

Contrast:

- Non-white
- Female
- Karnofsky<90
- Stage III/IV
- Partial/no B12
- Treatment

Symptoms:

- QoL
- Interference
- Symptom Distress
- Pain
- Dyspnea
- Cough
- Fatigue
- Anorexia

\[ m_{\text{cont}} - m_{\text{ref}} = \left[ -\frac{1}{\exp(\beta_{21} + \beta_{2\text{cov}})} \log(.5) \right]^{\frac{1}{\gamma}} - \left[ -\frac{1}{\exp(\beta_{21})} \log(.5) \right]^{\frac{1}{\gamma}} \]

i.e., difference between contrast (cont) and reference (ref) groups in months of median survival for each covariate (cov)
Protocol Results: Treatment Effects on Severity

Fitted LCSS severity for treatment (blue) and control (black). Asterisks indicate significance of treatment-by-day$^2$ (**) or only treatment-by-day (*) interactions.
Protocol Results: Treatment Effects on Severity

Fitted LCSS severity for treatment (blue) and control (black) Asterisks indicate significance of treatment-by-day\(^2\) (***) or only treatment-by-day (*) interactions
Protocol Results: Treatment Effects on Severity

Fitted LCSS severity for treatment (blue) and control (black) Asterisks indicate significance of treatment-by-day² (**) or only treatment-by-day (*) interactions
Subgroup Results: Vitamins, Treatment, and Anorexia

- Fitted trajectories for probability of non-zero anorexia in treatment (blue) and control (black) by vitamin supplementation status.
- Asterisks indicate 3-way interaction (vitamin-by-treatment-by-day) is significant (**)
- Notice narrow range of values; probably not clinically significant
Subgroup Results: Vitamins, Treatment, and Anorexia

- Fitted trajectories for mean of non-zero anorexia in treatment (blue) and control (black) by vitamin supplementation status.
- Asterisk indicates only 2-way interaction (vitamin-by-day) is significant (*).
- Now the range of values is more interesting; probably is clinically significant.
Individual Fitted Probability Trajectories

- Fitted individual-level (blue) and population-level (black)
- Vertical bar marks time of progression/death
- Crosses are binary LCSS observations $I(Y_{ij} > 0)$
Individual Fitted Severity Trajectories

- Fitted individual-level (blue) and population-level (black)
- Vertical bar marks time of progression/death
- Crosses are LCSS severity observations
Multivariate ZIB Joint Model

Longitudinal models:
\[ X_{0ik}(s)\beta_{0k} + U_{0ik}(s) \]
\[ X_{1ik}(s)\beta_{1k} + U_{1ik}(s) \]

Latent processes:
\[ U_{0i1}(s) = f_0(u_{0i1}) \]
\[ U_{1i1}(s) = f_1(u_{1i1}) \]
\[ \ldots \]
\[ U_{0ik}(s) = f_0(u_{0iK}) \]
\[ U_{1iK}(s) = f_1(u_{1iK}) \]

Survival model:
\[ X_{2i}\beta_2 + A(\alpha, u_{0i1}, u_{1i1}, \ldots, u_{0iK}, u_{1iK}) \]

i.e., different pair of latent trajectories underlie each long’l outcome
Multivariate ZIB Joint Model

Longitudinal models:
\[ X_{0ik}(s)\beta_{0k} + \alpha_{0k} U_{0i}(s) \]
\[ X_{1ik}(s)\beta_{1k} + \alpha_{1k} U_{1i}(s) \]

Latent processes:
\[ U_{0i}(s) = f_0(u_{0i}) \]
\[ U_{1i}(s) = f_1(u_{1i}) \]

Survival model:
\[ X_{2i}\beta_2 + A(\alpha_2, u_{0i}, u_{1i}) \]
i.e., same two latent trajectories underlie all long’l outcomes
Multi-state Survival Model

1 = Progressed Disease

\[ \lambda_{01} \rightarrow \lambda_{12} \]

0 = Baseline

\[ \lambda_{02} \]

2 = Death

3 survival models:

\[
\log(\lambda_{01}) = X_{2,01}i\beta_{2,01} + A_{01}(\alpha_{01}, u_{0i}, u_{1i})
\]

\[
\log(\lambda_{02}) = X_{2,02}i\beta_{2,02} + A_{02}(\alpha_{02}, u_{0i}, u_{1i})
\]

\[
\log(\lambda_{12}) = X_{2,12}i\beta_{2,12} + A_{12}(\alpha_{12}, u_{0i}, u_{1i})
\]

Linked to longitudinal models via random effects \((u_{0i}, u_{1i})\)

Or could link via value of random trajectories \(U_{0i}(s)\) and \(U_{1i}(s)\)
Learning about random effects

Consider $\mathbf{u}_i = (u_{01i}, u_{11i}, u_{12i})^T$.

Q: How much do we learn about a person’s random effects from long’l versus survival data?

Compare posterior variance using

- all data $\text{Var}(\mathbf{u}_i|\mathbf{y}_i, t_i)$,
- only long’l data $\text{Var}(\mathbf{u}_i|\mathbf{y}_i)$, or
- only survival data $\text{Var}(\mathbf{u}_i|t_i)$.

Since $p(\mathbf{u}_i|\mathbf{y}_i, t_i), p(\mathbf{u}_i|t_i), p(\mathbf{u}_i|\mathbf{y}_i)$ not available in closed form, consider either

- Observed Fisher information matrix
- MCMC sampling from these posteriors
Learning about weakly identified parameters

Identifiability: $\theta_1 \neq \theta_2 \Rightarrow f(X|\theta_1) \neq f(X|\theta_2)$

In a Bayesian context (Dawid 1979):

- **Identified**: $p(\eta|y) \rightarrow I(\eta = \eta_0)$ as $n \rightarrow \infty$
- **Conditionally** unidentified: $p(\theta|\eta, y) = p(\theta|\eta)$
- **Marginally** unidentified: $p(\theta|y) = p(\theta)$
- **Weakly** identified: prior to posterior learning is small
Learning about weakly identified parameters

Identifiability: \( \theta_1 \neq \theta_2 \Rightarrow f(X|\theta_1) \neq f(X|\theta_2) \)

In a Bayesian context (Dawid 1979):
- **Identified**: \( p(\eta|y) \to I(\eta = \eta_0) \) as \( n \to \infty \)
- **Conditionally** unidentified: \( p(\theta|\eta, y) = p(\theta|\eta) \)
- **Marginally** unidentified: \( p(\theta|y) = p(\theta) \)
- **Weakly** identified: prior to posterior learning is small

Quantifying learning

- **Upper bound**: \( p(\theta|y) \to p(\theta|\eta_0) \) as \( n \to \infty \)
- **Potential** learning: \( \text{Var}^{-1}(\theta|\eta) - \text{Var}^{-1}(\theta) \)
  i.e., change in *precision*
- **Remaining** learning: \( \text{Var}^{-1}(\theta|\eta) - \text{Var}^{-1}(\theta|y) \)
Conclusions

- Treatment significantly reduces PRO severity and increases PFS when these outcomes are considered jointly.
- PRO trajectories are associated with PFS times.
- Joint models address informative censoring (thus reducing bias) and can be more efficient.
- Considering multiple PROs simultaneously may improve efficiency further.
- Models we develop are widely applicable to other longitudinal PRO + event time data sets.