Examples of joint models for multivariate longitudinal and multistate processes in chronic diseases

Cécile Proust-Lima

works with Loïc Ferrer, Anaïs Rouanet and Hélène Jacqmin-Gadda

INSERM U897, Epidemiology and Biostatistics, Bordeaux, France
Univ. Bordeaux, ISPED, Bordeaux, France
cecel.proust-lima@inserm.fr

Workshop on Flexible Models for Longitudinal and Survival Data with Applications in Biostatistics
Joint modelling principle

Simultaneous modelling of correlated longitudinal and survival data

- Longitudinal marker
- Latent structure
- Time to event

Objectives:

▶ Describe the longitudinal process stopped by the event
▶ Predict the risk of event adjusted for the longitudinal process
▶ Explore the association between the two processes
Joint modelling principle

Simultaneous modelling of correlated longitudinal and survival data

Objectives:
- describe the longitudinal process stopped by the event
- predict the risk of event adjusted for the longitudinal process
- explore the association between the two processes
2 main families of joint models

- **Longitudinal marker**
- **Time to event**

**Mixed model** (usually linear)

**Survival model** (usually proportional hazards)

**Link with the latent structure:**
- Random effects from the mixed model (shared random effect models)
- Latent class structure (joint latent class models)
Shared random-effect model (SREM) (Rizopoulos, 2012)

- **Shared random-effects distribution**: \( b_i \sim N(\mu, B) \)

- **Linear mixed model** for the biomarker trajectory:
  \[
  Y_i(t_{ij}) = Y_i(t_{ij})^* + \epsilon_{ij} = Z_i(t_{ij})^T b_i + X_{Li}(t_{ij})^T \beta + \epsilon_{ij} \text{ with } \epsilon_{ij} \sim N(0, \sigma^2_{\epsilon})
  \]

- **Proportional hazard model** including marker trajectory characteristics:
  \[
  \lambda(t \mid b_i) = \lambda_0(t)e^{X_{Si}(t)^T \delta + W_i(b_i, \beta, t)^T \eta}
  \]

→ JM, JMBayes in R, stjm in Stata, JMFit in SAS
Joint latent class model (JLCM) (Proust-Lima et al., 2014)

- **Shared latent class** \((c_i)\) membership:
  \[
  \pi_{ig} = P(c_i = g | X_{pi}) = \frac{e^{\xi_{0g} + X_{Ci}^T \xi_g}}{\sum_{l=1}^{G} e^{\xi_{0l} + X_{Ci}^T \xi_l}} \quad \text{with} \ \xi_{0G} = 0 \ & \xi_{1G} = 0
  \]

- **Class-specific linear mixed model** for the biomarker trajectory:
  \[
  Y_i(t_{ij}) | c_i = g = Z_i(t_{ij})^T b_{ig} + X_{Li}(t_{ij})^T \beta_g + \epsilon_{ij} \quad \text{with} \ b_{ig} \sim \mathcal{N}(\mu_g, B_g), \ \epsilon_{ij} \sim \mathcal{N}(0, \sigma^2_\epsilon)
  \]

- **Class-specific proportional hazard model**:
  \[
  \lambda(t | c_i = g) = \lambda_{0g}(t)e^{X_{Ti}(t)\delta_g}
  \]

→ lamm in R
Remarks (Proust-Lima et al., 2014)

Shared random effect models:
- extension of the standard time-to-event models
- assessment of specific associations (surrogacy)
- quantification of the association

Joint latent class models:
- heterogeneous population
- no assumption on the association
- useful for predictive tools

In any case, most developments for:
- a Gaussian longitudinal marker
- a right-censored time to event

→ but more complex data in most cohort studies on chronic diseases
Remarks (Proust-Lima et al., 2014)

Shared random effect models:
- extension of the standard time-to-event models
- assessment of specific associations (surrogacy)
- quantification of the association

Joint latent class models:
- heterogeneous population
- no assumption on the association
- useful for predictive tools

In any case, most developments for:
- a Gaussian longitudinal marker
- a right-censored time to event

→ but more complex data in most cohort studies on chronic diseases
In chronic diseases

- **Longitudinal part:**
  - multiple markers of progression
  - markers of different nature
  - Gaussian, binary, poisson
  - ordinal
  - continuous but non Gaussian

- **Survival part:**
  - competing risks
  - recurrent events
  - multiple events
  - succession of different events

Examples of developments through the study of:
- Progression of localized Prostate cancer after treatment
- Natural history of Alzheimer's disease
In chronic diseases

- **Longitudinal part:**
  - multiple markers of progression
  - markers of different nature
  - Gaussian, binary, poisson
  - ordinal
  - continuous but non Gaussian

- **Survival part:**
  - competing risks
  - recurrent events
  - multiple events
  - succession of different events

3 examples of developments through the study of
- progression of localized Prostate cancer after treatment
- natural history of Alzheimer’s disease
Progression of localized Prostate Cancer after a treatment by radiation therapy
Localized prostate cancer

- Monitoring of patients after radiation therapy for a localized Prostate cancer:
  - prognostic factors at diagnosis (T-stage, Gleason, dose of RT, ...)
  - repeated measures of PSA (prostate specific antigen) collected in routine

- Interest in predicting the risk of progression
  - multiple types: local recurrence, distant recurrence, death
  - problem of initiation of new treatment: hormonal treatment
**Dynamic prediction of clinical recurrence of any type**

(Sène et al., SMMR 2014):

- **Individualized probability of clinical recurrence:**
  - in the next three years
  - for a man naive of HT
  - according to hypothetical times of initiation of HT (time-dependent covariate)

---

![Graph showing PSA measures and probability of recurrence over time.

- PSA measures represented by crosses.
- Time of prediction indicated by dashed line.
- Probability of recurrence from M4b and M2c.
- Initiation of HT at different time points: now, in 1 year, in 2 years, if no HT.
- Time (years) since end of EBRT on x-axis.
- Log(PSA + 0.1) on y-axis.
- Probability of recurrence on y-axis.

---

Cécile Proust-Lima (INSERM) | Joint models for multiple outcomes
Dynamic prediction of clinical recurrence of any type
(Sène et al., SMMR 2014):

- Individualized probability of clinical recurrence:
  - in the next three years
  - for a man naive of HT
  - according to hypothetical times of initiation of HT (time-dependent covariate)
Dynamic prediction of clinical recurrence of any type
(Sène et al., SMMR 2014):

- Individualized probability of clinical recurrence:
  - in the next three years
  - for a man naive of HT
  - according to hypothetical times of initiation of HT (time-dependent covariate)

![Graph showing PSA measures and probability of recurrence over time since end of EBRT.](image)
Multiple types of progression

- Clinical progression is a multistate process:

  - Importance of distinguishing the different types to:
    - clarify the impact of PSA dynamics and other prognostic factors on each transition
    - predict type-specific progressions
The joint longitudinal and multistate model \cite{Ferrer2015} \noindent

Notations:

- **multi-state process**
  - \( E_i = \{E_i(t), T_{i0} \leq t \leq C_i \} \) is a non-homogeneous Markov process
    - \( E_i(t) \) takes values in the finite state space \( S = \{0, 1, \ldots, M\} \)
    - \( T_{i0} \) the left truncation time, \( C_i \) the right censoring time
  - \( T_i = (T_{i1}, \ldots, T_{imi})^\top \) the \( m_i \) observed times with \( T_{ir} < T_{i(r+1)}, \forall r \in S \)
  - \( \delta_i = (\delta_{i1}, \ldots, \delta_{imi})^\top \) the vector of observed transition indicators

- **longitudinal process**
  - \( Y_i = (Y_{i1}, \ldots, Y_{ini})^\top \) the \( n_i \) measures of the marker collected at times \( t_{i1}, \ldots, t_{ini} \), with \( t_{ini} \leq T_{i1} \)
The joint longitudinal and multistate model \cite{Ferrer2015} (cont'd)

- **Longitudinal part : mixed model**

  \[
  Y_{ij} = Y_i^*(t_{ij}) + \epsilon_{ij} = X_{Li}(t_{ij})^\top \beta + Z_i(t_{ij})^\top b_i + \epsilon_{ij}
  \]

  \[b_i \sim \mathcal{N}(0, D), \quad \epsilon_i = (\epsilon_{i1}, \ldots, \epsilon_{in_i})^\top \sim \mathcal{N}(0, \sigma^2 I_{ni}), \quad b_i \perp \epsilon_i\]

- **Survival part : multistate model**

  \[
  \lambda_{hk}^i(t|b_i) = \lim_{dt \to 0} \frac{\Pr(E_i(t + dt) = k|E_i(t) = h; b_i)}{dt} = \lambda_{hk,0}(t) \exp(X_{Thk,i}^\top \gamma_{hk} + W_{hk,i}(b_i, t)^\top \eta_{hk}), \quad \text{for } h, k \in S,
  \]

  \[\lambda_{hk,0}(t) \quad \text{parametric baseline intensity, } X_{Thk,i} \quad \text{prognostic factors}\]

  \[W_{hk,i}(b_i, t) \quad \text{the dependence structure} \quad \text{(Sène et al., J sfds 2014)}\]

  \[
  \begin{align*}
  \star \quad W_{hk,i}(b_i, t) &= Y_i^*(t) \quad \longrightarrow \quad \text{(true current level)} \\
  \star \quad W_{hk,i}(b_i, t) &= \partial Y_i^*(t) / \partial t \quad \longrightarrow \quad \text{(true current slope)} \\
  \star \quad W_{hk,i}(b_i, t) &= (Y_i^*(t), \partial Y_i^*(t) / \partial t)^\top \quad \longrightarrow \quad \text{(both)} \\
  \star \quad \ldots
  \end{align*}
  \]
Maximum Likelihood Estimation

- Likelihood function using $Y_i \indep_{b_i} T_i$,

$$L(\theta) = \prod_{i=1}^{N} \int_{\mathbb{R}^q} f_Y(Y_i|b_i; \theta) \, f_T(T_i, \delta_i|b_i; \theta) \, f_b(b_i; \theta) \, db_i$$

with:

- Random effects part: $b_i \sim \mathcal{N}_q(0, D)$
- Longitudinal part: $Y_i|b_i \sim \mathcal{N}_{n_i}(X_i^\top \beta + Z_i^\top b_i, \sigma^2 I_{n_i})$
- Multi-state part:

$$f_T(T_i, \delta_i|b_i; \theta) = \prod_{r=0}^{m_i-1} \left[ P_{E_i(T_{ir}), E_i(T_{ir})}^i(T_{ir}, T_{i(r+1)}|b_i) \times \lambda_{E_i(T_{ir}), E_i(T_{i(r+1)})}^i(T_{i(r+1)}|b_i) \delta_{i(r+1)} \right]$$

with $P_{hh}(s, t) = \exp \left( \int_s^t \lambda_{hh}(u) \, du \right) = \exp \left( - \sum_{k \neq h} \int_s^t \lambda_{hk}(u) \, du \right)$
Implementation under R

- Relies on **JM** package

- Implementation procedure decomposed into four steps:
  1. `lme()` function (**nlme** package) to initialise the parameters in the longitudinal sub-model;
  2. `mstate` package to adapt the data to the multi-state framework;
  3. `coxph()` function (**survival** package) to initialise the parameters in the multi-state sub-model;
  4. `JMstateModel()` function (extension of `jointModel()`) to estimate all the parameters of the joint multi-state model.

- Likelihood computed and optimised using:
  - numerical integration algorithms (Gaussian quadratures)
  - optimisation algorithms (EM + quasi-Newton)
2 cohorts of men with localized prostate cancer treated by radiotherapy (N=1474)

Repeated measures of PSA

Multi-state representation of the clinical progressions

10 (3, 21) measurements per patient (50th, 5th, 95th) %iles

\[ \Upsilon = \begin{pmatrix} 533 & 144 & 227 & 47 & 523 \\ 0 & 20 & 90 & 10 & 24 \\ 0 & 0 & 106 & 33 & 178 \\ 0 & 0 & 0 & 13 & 77 \\ 0 & 0 & 0 & 0 & 802 \end{pmatrix} \]
Longitudinal sub-model specification

\[ Y_{ij} = Y_i^*(t_{ij}) + \epsilon_{ij} \]
\[ = (\beta_0 + X_{L0i}^T \beta_{0,cov} + b_{i0}) + \]
\[ (\beta_1 + X_{L1i}^T \beta_{1,cov} + b_{i1}) \times f_1(t_{ij}) + \]
\[ (\beta_2 + X_{L2i}^T \beta_{2,cov} + b_{i2}) \times f_2(t_{ij}) + \epsilon_{ij} \]

- \( f_1(t) = (1 + t)^{-1.2} - 1 \) and \( f_2(t) = t \)
- \( b_i = (b_{i0}, b_{i1}, b_{i2})^T \sim \mathcal{N}(0, D) \), \( D \) unstructured, \( \epsilon_i \sim \mathcal{N}(0, \sigma^2 I_{n_i}) \)
- \( X_{L0i}, X_{L1i} \) and \( X_{L2i} \) were obtained using a backward stepwise procedure.
Multi-state sub-model specification

\[ \lambda_{hk}^i(t|b_i) = \lambda_{hk,0}(t) \exp \left( X_{T,hk,i}^\top \gamma_{hk} + \left( \frac{Y_i^*(t)}{\partial Y_i^*(t)/\partial t} \right)^\top \left( \eta_{hk,\text{level}} \eta_{hk,\text{slope}} \right) \right) \]

- Log-baseline intensities approximated by B-splines
- Proportionality assumptions between several baseline intensities
- Backward stepwise procedure to select the prognostic factors
- Dependence function chosen by Wald tests
## Results

Estimates of the association parameters between the longitudinal and multi-state processes

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>StdErr</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level : 01</td>
<td>0.37</td>
<td>0.09</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Level : 02</td>
<td>0.51</td>
<td>0.07</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Level : 03</td>
<td>0.45</td>
<td>0.11</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Level : 04</td>
<td>−0.17</td>
<td>0.05</td>
<td>0.001</td>
</tr>
<tr>
<td>Level : 12</td>
<td>−0.16</td>
<td>0.10</td>
<td>0.110</td>
</tr>
<tr>
<td>Level : 13</td>
<td>−0.41</td>
<td>0.20</td>
<td>0.042</td>
</tr>
<tr>
<td>Level : 14</td>
<td>0.10</td>
<td>0.14</td>
<td>0.487</td>
</tr>
<tr>
<td>Level : 23</td>
<td>−0.15</td>
<td>0.09</td>
<td>0.120</td>
</tr>
<tr>
<td>Level : 24</td>
<td>0.00</td>
<td>0.05</td>
<td>0.412</td>
</tr>
<tr>
<td>Level : 34</td>
<td>0.04</td>
<td>0.08</td>
<td>0.609</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>StdErr</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope : 01</td>
<td>2.54</td>
<td>0.31</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Slope : 02</td>
<td>3.04</td>
<td>0.25</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Slope : 03</td>
<td>2.43</td>
<td>0.49</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Slope : 04</td>
<td>1.03</td>
<td>0.32</td>
<td>0.001</td>
</tr>
<tr>
<td>Slope : 12</td>
<td>2.01</td>
<td>0.61</td>
<td>0.001</td>
</tr>
<tr>
<td>Slope : 13</td>
<td>3.18</td>
<td>0.80</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Slope : 14</td>
<td>−0.20</td>
<td>1.27</td>
<td>0.873</td>
</tr>
<tr>
<td>Slope : 23</td>
<td>0.97</td>
<td>0.67</td>
<td>0.150</td>
</tr>
<tr>
<td>Slope : 24</td>
<td>0.29</td>
<td>0.52</td>
<td>0.583</td>
</tr>
<tr>
<td>Slope : 34</td>
<td>−0.79</td>
<td>0.78</td>
<td>0.313</td>
</tr>
</tbody>
</table>

### Multi-state process

![Multi-state process diagram](image)

Prognostic factors: advanced initial stage not always associated with intensities of transitions between health states after adjustment on PSA

---

Cécile Proust-Lima (INSERM)  
Joint models for multiple outcomes  
July 2015
## Results

### Estimates of the association parameters between the longitudinal and multi-state processes

<table>
<thead>
<tr>
<th>Value</th>
<th>StdErr</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level : 01</td>
<td>0.37</td>
<td>0.09</td>
</tr>
<tr>
<td>Level : 02</td>
<td>0.51</td>
<td>0.07</td>
</tr>
<tr>
<td>Level : 03</td>
<td>0.45</td>
<td>0.11</td>
</tr>
<tr>
<td>Level : 04</td>
<td>−0.17</td>
<td>0.05</td>
</tr>
<tr>
<td>Level : 12</td>
<td>−0.16</td>
<td>0.10</td>
</tr>
<tr>
<td>Level : 13</td>
<td>−0.41</td>
<td>0.20</td>
</tr>
<tr>
<td>Level : 14</td>
<td>0.10</td>
<td>0.14</td>
</tr>
<tr>
<td>Level : 23</td>
<td>−0.15</td>
<td>0.09</td>
</tr>
<tr>
<td>Level : 24</td>
<td>0.00</td>
<td>0.05</td>
</tr>
<tr>
<td>Level : 34</td>
<td>0.04</td>
<td>0.08</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Value</th>
<th>StdErr</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope : 01</td>
<td>2.54</td>
<td>0.31</td>
</tr>
<tr>
<td>Slope : 02</td>
<td>3.04</td>
<td>0.25</td>
</tr>
<tr>
<td>Slope : 03</td>
<td>2.43</td>
<td>0.49</td>
</tr>
<tr>
<td>Slope : 04</td>
<td>1.03</td>
<td>0.32</td>
</tr>
<tr>
<td>Slope : 12</td>
<td>2.01</td>
<td>0.61</td>
</tr>
<tr>
<td>Slope : 13</td>
<td>3.18</td>
<td>0.80</td>
</tr>
<tr>
<td>Slope : 14</td>
<td>−0.20</td>
<td>1.27</td>
</tr>
<tr>
<td>Slope : 23</td>
<td>0.97</td>
<td>0.67</td>
</tr>
<tr>
<td>Slope : 24</td>
<td>0.29</td>
<td>0.52</td>
</tr>
<tr>
<td>Slope : 34</td>
<td>−0.79</td>
<td>0.78</td>
</tr>
</tbody>
</table>

### Multi-state process

![Multi-state process diagram](image)

**Prognostic factors**:
advanced initial stage
not always associated with intensities of transitions between health states after adjustment on PSA
## Results

Estimates of the association parameters between the longitudinal and multi-state processes

<table>
<thead>
<tr>
<th>Level</th>
<th>Value</th>
<th>StdErr</th>
<th>( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level : 01</td>
<td>0.37</td>
<td>0.09</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Level : 02</td>
<td>0.51</td>
<td>0.07</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Level : 03</td>
<td>0.45</td>
<td>0.11</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Level : 04</td>
<td>−0.17</td>
<td>0.05</td>
<td>0.001</td>
</tr>
<tr>
<td>Level : 12</td>
<td>−0.16</td>
<td>0.10</td>
<td>0.110</td>
</tr>
<tr>
<td>Level : 13</td>
<td>−0.41</td>
<td>0.20</td>
<td>0.042</td>
</tr>
<tr>
<td>Level : 14</td>
<td>0.10</td>
<td>0.14</td>
<td>0.487</td>
</tr>
<tr>
<td>Level : 23</td>
<td>−0.15</td>
<td>0.09</td>
<td>0.120</td>
</tr>
<tr>
<td>Level : 24</td>
<td>0.00</td>
<td>0.05</td>
<td>0.412</td>
</tr>
<tr>
<td>Level : 34</td>
<td>0.04</td>
<td>0.08</td>
<td>0.609</td>
</tr>
</tbody>
</table>

| Slope : 01 | 2.54  | 0.31   | < 0.001       |
| Slope : 02 | 3.04  | 0.25   | < 0.001       |
| Slope : 03 | 2.43  | 0.49   | < 0.001       |
| Slope : 04 | 1.03  | 0.32   | 0.001         |
| Slope : 12 | 2.01  | 0.61   | 0.001         |
| Slope : 13 | 3.18  | 0.80   | < 0.001       |
| Slope : 14 | −0.20 | 1.27   | 0.873         |
| Slope : 23 | 0.97  | 0.67   | 0.150         |
| Slope : 24 | 0.29  | 0.52   | 0.583         |
| Slope : 34 | −0.79 | 0.78   | 0.313         |

### Multi-state process

Prognostic factors: advanced initial stage not always associated with intensities of transitions between health states after adjustment on PSA.


## Results

### Estimates of the association parameters between the longitudinal and multi-state processes

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>StdErr</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level : 01</td>
<td>0.37</td>
<td>0.09</td>
<td>$&lt; 0.001$</td>
</tr>
<tr>
<td>Level : 02</td>
<td>0.51</td>
<td>0.07</td>
<td>$&lt; 0.001$</td>
</tr>
<tr>
<td>Level : 03</td>
<td>0.45</td>
<td>0.11</td>
<td>$&lt; 0.001$</td>
</tr>
<tr>
<td>Level : 04</td>
<td>-0.17</td>
<td>0.05</td>
<td>0.001</td>
</tr>
<tr>
<td>Level : 12</td>
<td>-0.16</td>
<td>0.10</td>
<td>0.110</td>
</tr>
<tr>
<td>Level : 13</td>
<td>-0.41</td>
<td>0.20</td>
<td>0.042</td>
</tr>
<tr>
<td>Level : 14</td>
<td>0.10</td>
<td>0.14</td>
<td>0.487</td>
</tr>
<tr>
<td>Level : 23</td>
<td>-0.15</td>
<td>0.09</td>
<td>0.120</td>
</tr>
<tr>
<td>Level : 24</td>
<td>0.00</td>
<td>0.05</td>
<td>0.412</td>
</tr>
<tr>
<td>Level : 34</td>
<td>0.04</td>
<td>0.08</td>
<td>0.609</td>
</tr>
<tr>
<td>Slope : 01</td>
<td>2.54</td>
<td>0.31</td>
<td>$&lt; 0.001$</td>
</tr>
<tr>
<td>Slope : 02</td>
<td>3.04</td>
<td>0.25</td>
<td>$&lt; 0.001$</td>
</tr>
<tr>
<td>Slope : 03</td>
<td>2.43</td>
<td>0.49</td>
<td>$&lt; 0.001$</td>
</tr>
<tr>
<td>Slope : 04</td>
<td>1.03</td>
<td>0.32</td>
<td>0.001</td>
</tr>
<tr>
<td>Slope : 12</td>
<td>2.01</td>
<td>0.61</td>
<td>0.001</td>
</tr>
<tr>
<td>Slope : 13</td>
<td>3.18</td>
<td>0.80</td>
<td>$&lt; 0.001$</td>
</tr>
<tr>
<td>Slope : 14</td>
<td>-0.20</td>
<td>1.27</td>
<td>0.873</td>
</tr>
<tr>
<td>Slope : 23</td>
<td>0.97</td>
<td>0.67</td>
<td>0.150</td>
</tr>
<tr>
<td>Slope : 24</td>
<td>0.29</td>
<td>0.52</td>
<td>0.583</td>
</tr>
<tr>
<td>Slope : 34</td>
<td>-0.79</td>
<td>0.78</td>
<td>0.313</td>
</tr>
</tbody>
</table>

### Multi-state process

- **End EBRT**
  - $\lambda_{01}(t)$
- **Local Recurrence**
  - $\lambda_{02}(t)$
  - $\lambda_{12}(t)$
- **Hormonal Therapy**
- **Distant Recurrence**
  - $\lambda_{23}(t)$
  - $\lambda_{24}(t)$
- **Death**
  - $\lambda_{34}(t)$

### Prognostic factors:
- Advanced initial stage not always associated with intensities of transitions between health states after adjustment on PSA.
Diagnostics for the parametric assumptions

- Goodness-of-fit plots for the longitudinal process
  - Conditional standardized residuals versus fitted values
  - Observed and predicted values of the biomarker

![Conditional standardized residuals versus fitted values](image1)

![Observed and predicted values of the biomarker](image2)
Diagnostics for the parametric assumptions

- Goodness-of-fit plots for the longitudinal process
- Goodness-of-fit plots for the multi-state process
  - Predicted transition probabilities from the joint multi-state model and non-parametric probability transitions

![Graph showing transition probabilities](image)
Natural history of Alzheimer’s disease, dementias and cognitive aging
Cognitive aging and dementia

- Dementia (e.g. Alzheimer’s disease) characterized by a progressive decline of cognition

- Most interest in
  - the natural history of dementia
  - risk factors of cognitive decline and dementia
  - dynamic individual prediction of dementia

\[\text{\textcopyright Cécile Proust-Lima (INSERM)}\]

Joint models for multiple outcomes

July 2015 22 / 45
Cognitive aging and dementia

- Dementia (e.g. Alzheimer’s disease) characterized by a progressive decline of cognition

- Most interest in
  - the natural history of dementia
  - risk factors of cognitive decline and dementia
  - dynamic individual prediction of dementia

→ 1st complexity: elderly pathology
  - delayed entry
  - dementia in competition with death
  - diagnosis at pre-established visit times

→ 2nd complexity: cognition is not directly observed
  - cognitive process (trait) defined in continuous time
  - repeated psychometric tests measured in discrete times
    ★ multiple cognitive functions (language, memory, attention,...)
    ★ noisy measures of overall cognition
    ★ limited statistical properties
Latent process mixed model

covariates $X(t)$

\[ \Lambda_i(t) = X_L1i(t) \top \beta + Z_i(t) \top b_i + w_i(t) \]

\[ b_i \sim \text{MVN}(\mu, B) \]

\[ w_i(t) \text{ autocorrelated process} \]

\[ b_{i0} \sim \text{N}(0, 1) \text{ for identifiability} \]

linear link function

\[ \tilde{Y}_{kij} = \Lambda_i(t)_{ijk} + X_L2i(t) \top \gamma_k + \alpha_{ki} + \epsilon_{kij} \]

\[ \alpha_{ki} \sim \text{N}(0, \sigma_\alpha^2) \]

\[ \epsilon_{kij} \sim \text{N}(0, \sigma_\epsilon^2) \]
Latent process mixed model

covariates $X(t)$
time $t$
cognitive process $\Lambda(t)$
test 1 $Y_1$ at $T_{11}$
Latent process mixed model

covariates $X(t)$
time $t$

cognitive process $\Lambda(t)$

test 1 $Y_1$ at $T_{11}$
test $k$ $Y_k$ at $T_{1k}$
test $K$ $Y_K$ at $T_{1K}$
Latent process mixed model

covariates $X(t)$
time $t$

$cognitive$

process

$\Lambda(t)$

$\Lambda_i(t) = X_{L1i}(t)^\top \beta + Z_i(t)^\top b_i + w_i(t)$

$\downarrow b_i \sim MVN(\mu, B)$

$\downarrow w_i(t)$ autocorrelated process

$\downarrow b_{i0} \sim N(0, 1)$ for identifiability
Latent process mixed model

covariates $X(t)$
time $t$

cognitive process $\Lambda(t)$

test 1 $Y_1$ at $T_{11}$
test $k$ $Y_k$ at $T_{1k}$
test $K$ $Y_K$ at $T_{1K}$

$\Lambda_i(t) = X_{L1i}(t)^\top \beta + Z_i(t)^\top b_i + w_i(t)$

$\triangleright b_i \sim MVN(\mu, B)$
$\triangleright w_i(t)$ autocorrelated process
$\triangleright b_{i0} \sim N(0, 1)$ for identifiability

$Y_{kij} = \zeta_{1k} + \zeta_{2k} \tilde{Y}_{kij}$
$\tilde{Y}_{kij} = \Lambda_i(t_{ijk}) + \epsilon_{kij}$

$\epsilon_{kij} \sim N(0, \sigma_{\epsilon_k}^2)$
Latent process mixed model

covariates \(X(t)\) at time \(t\)

cognitive process \(\Lambda(t)\)

test 1 \(Y_1\) at \(T_{11}\)
test \(k\) \(Y_k\) at \(T_{1k}\)
test \(K\) \(Y_K\) at \(T_{1K}\)

\[\Lambda_i(t) = X_{L1i}(t)^\top \beta + Z_i(t)^\top b_i + w_i(t)\]

\(b_i \sim \text{MVN}(\mu, B)\)
\(w_i(t)\) autocorrelated process
\(b_{i0} \sim N(0, 1)\) for identifiability

\[Y_{kij} = \zeta_{1k} + \zeta_{2k} \tilde{Y}_{kij}\]
\[
\tilde{Y}_{kij} = \Lambda_i(t_{ijk}) + X_{L2i}(t)^\top \gamma_k + \alpha_{ki} + \epsilon_{kij}
\]

\(\alpha_{ki} \sim N(0, \sigma^2_{\alpha_k})\)
\(\epsilon_{kij} \sim N(0, \sigma^2_{\epsilon_k})\)
Latent process mixed model

covariates $X(t)$
time $t$

cognitive process $\Lambda(t)$

test 1 $Y_1$ at $T_{11}$
test $k$ $Y_k$ at $T_{1k}$
test $K$ $Y_K$ at $T_{1K}$

$\Lambda_i(t) = X_{L1i}(t)^\top \beta + Z_i(t)^\top b_i + w_i(t)$

- $b_i \sim MVN(\mu, B)$
- $w_i(t)$ autocorrelated process
- $b_{i0} \sim N(0, 1)$ for identifiability

linear link function

$Y$ noisy latent process
Latent process mixed model involving nonlinear link functions (Proust-Lima et al., 2015)

\[ \Lambda_i(t) = X_{L1i}(t)^T \beta + Z_i(t)^T b_i + w_i(t) \]

- \( b_i \sim MVN(\mu, B) \)
- \( w_i(t) \) autocorrelated process
- \( b_{i0} \sim N(0, 1) \) for identifiability

- nonlinear link function

Cognitive process \( \Lambda(t) \)

covariates \( X(t) \)
time \( t \)

Y at \( T \)

noisy latent process

test 1 \( Y_1 \) at \( T_{11} \)

... \( \ldots \)

test \( k \) \( Y_k \) at \( T_{1k} \)

... \( \ldots \)

test \( K \) \( Y_K \) at \( T_{1K} \)
Latent process mixed model involving nonlinear link functions \cite{Proust-Lima2015}

\begin{align*}
\Lambda_i(t) &= X_{L1i}(t)^T \beta + Z_i(t)^T b_i + w_i(t) \\
&\quad \text{where } b_i \sim \text{MVN}(\mu, B) \\
&\quad \text{and } w_i(t) \text{ is an autocorrelated process} \\
&\quad \text{with } b_{i0} \sim \text{N}(0, 1) \text{ for identifiability}
\end{align*}

\begin{align*}
Y_{kij} &= H_k(\tilde{Y}_{kij} ; \zeta_k) \\
\tilde{Y}_{kij} &= \Lambda_i(t_{ijk}) + X_{L2i}(t)^T \gamma_k + \alpha_{ki} + \epsilon_{kij}
\end{align*}

\begin{align*}
&\alpha_{ki} \sim \mathcal{N}(0, \sigma_{\alpha_k}^2) \\
&\epsilon_{kij} \sim \mathcal{N}(0, \sigma_{\epsilon_k}^2)
\end{align*}

\(H_k\) = flexible parameterized transformation for outcome \(k\)

\(\rightarrow\) linear, standardised Beta CDF, quadratic I-splines, thresholds, ...
Joint model for multivariate cognitive measures, dementia and death

cognitive measures:
- latent process mixed model
Joint model for multivariate cognitive measures, dementia and death

- Latent process \( \Lambda(t) \)
- Random effects \( u \)
- Cognitive measures:
  - Latent process mixed model
- Times to event \((T, \delta)\)
- Dementia and death?
  - Option 1: First event in competing setting
  - Option 2: Multistate model
Joint model for multivariate cognitive measures, dementia and death

- Latent process
  - Latent classes $c$
- Random effects $u$
  - Shared latent quantity = latent classes
    - Heterogeneous trajectories
    - No assumption on the association

Cognitive measures:
- Latent process mixed model
  - Option 1: First event in competing setting
  - Option 2: Multistate model

Times to event $(T, \delta)$

Dementia and death?

Cécile Proust-Lima (INSERM)
Option 1: competing setting (Proust-Lima et al., ArXiv 2014)

- Times to events = Time to $P$ competing events
  - $T_i = \min(\text{censoring } \tilde{T}_i, \text{ and cause-spec. times } T^*_i, \ldots, T^*_P)$,
  - $\delta_i = 0$ for censored, $\delta_i = p$ otherwise

- class-specific cause-specific proportional hazard models

$$\lambda_p(t) |_{c_i=g} = \lambda_{0p}(t; \nu_{pg}) \exp(X_{Ti}^\top \zeta_{pg})$$

- $\lambda_{0p}$ parametric (splines, Gompertz, Weibull,...)
Maximum Likelihood Estimation

- Likelihood function using $Y_i \perp_{c_i} T_i$,

$$L(\theta) = \prod_{i=1}^{N} \sum_{g=1}^{G} f_Y(Y_i|X_{Li}, Z_i, c_i = g; \theta) \ f_T(T_i, \delta_i|X_{T_i}, c_i = g; \theta) \ P(c_i = g|X_{Ci}; \theta)$$

with:

- $f_Y(Y_i|X_{Li}, Z_i, c_i = g; \theta)$ from the latent process mixed model
  - closed form if only continuous markers: $	imes$ Jacobian of $(H_k)_{k=1,\ldots,K}$
  - by numerical integration otherwise ...

- $f_T(T_i, \delta_i|X_{T_i}, c_i = g; \theta)$ from the cause-specific model
  - overall survival $S_{ig} \times$ instantaneous risk for cause $p$ in $g$ if $\delta_i = p$

- $P(c_i = g|X_{Ci}; \theta)$ from a multinomial logistic model

- Left truncation (entry at $T_{0i}$): $l^{T_0}(\theta) = \log \left( \frac{L(\theta)}{\prod_{i=1}^{N} S_i(T_{0i}; \theta)} \right)$
Implementation, model selection and evaluation

- Iterative (Marquardt) algorithm for a given $G$
  - implemented in HETMIXSURV_V2 parallel Fortran90 program
  - validated in simulation studies (with for instance splines and threshold link functions)
  - implemented in Jointlcmm (R) for 1 marker

- Posterior selection of the optimal number $G$ of latent classes
  - Information measures : AIC, BIC
  - Score Test for conditional independence assumption :
    $\rightarrow$ longitudinal and survival parts are independent conditionally on the latent classes

- Further evaluation of the model using :
  - Posterior classification stemmed from $P(c_i = g | X_i, Y_i, (T_i, \delta_i); \hat{\theta})$
  - Longitudinal/Survival predictions versus observations
    $\rightarrow$ posterior-probability-weighted means over time intervals
Conditional independence assumption:

Class-specific cause-specific proportional hazard model

\[ \lambda_p(t) = \lambda_{0p}(t; \nu_{pg}) \exp(X_{Ti}^T \zeta_{pg}) \]
**Conditional independence assumption**: alternative

Class-specific cause-specific proportional hazard model

\[ \lambda_p(t)^{\mathcal{H}_1} = \lambda_{0p}(t; \nu_{pg}) \exp(X_{Ti}^\top \zeta_{pg} + u_{ig} \kappa_p) \]

\[ \rightarrow \text{Score test for } \mathcal{H}_0 : \kappa_p = 0 \text{ and } \mathcal{H}_0 : \kappa = (\kappa_1^\top, \ldots, \kappa_P^\top)^\top = 0 \]
Clinical background
- semantic memory (verbal fluency, ...) affected long before dementia diagnosis
- could play a role for early prediction of dementia

Objective
- describe profiles of semantic memory decline in association with dementia and death
- predict the risk of dementia from semantic memory history

PAQUID cohort data:
- population-based cohort on cerebral aging
- 65 years and older
- 22 years of follow-up every 2 or 3 years
- subpopulation with genetic information: ApoE4

→ 588 subjects
Dynamics of semantic memory

- 2 longitudinal measures:
  - Isaacs Set Test (\textit{IST15}) (discrete quantitative in \{0-40\})
  - Wechsler similarities test (\textit{WST}) (ordinal in \{0-10\})

- Trajectory according to \textit{age} (natural history)
  - age at entry (\textit{ageT0}), sex (\textit{sex}), education (\textit{EL}), apoE4 (\textit{E4})

- In each latent class \(g\):

\[
\Lambda(\text{age})_{|c_i=g} = b_{0ig} + \beta_1 \text{sex} + \beta_2 \text{EL} + \beta_3 \text{E4} + \beta_4 \text{ageT0} + \\
(b_{1ig} + \beta_5 \text{sex} + \beta_6 \text{EL} + \beta_7 \text{E4}) \times \text{age} \\
(b_{2ig} + \beta_8 \text{sex} + \beta_9 \text{EL} + \beta_{10} \text{E4}) \times \text{age}^2 + w_i(\text{age})
\]

\[
\text{IST15}_{ij} = H_1(\Lambda(\text{age}_{1ij}) + \alpha_{1i} + \epsilon_{1ij} ; \eta_1)
\]

\[
\text{WST}_{ij} = H_2(\Lambda(\text{age}_{2ij}) + \alpha_{2i} + \epsilon_{2ij} ; \eta_2)
\]

with \(b_{ig} \sim \mathcal{N}(\mu_g, B)\), \(w_i \sim \text{Brownian motion}\)
\(\alpha_{2i} \sim \mathcal{N}(0, \sigma_{\alpha}^2)\), \(\epsilon_{ki} \sim \mathcal{N}(0, \sigma_k^2)\), \(k = 1, 2\)
In separated analyses, Splines with 3 internal knots =
good balance between estimation difficulty and goodness-of-fit
Risk of dementia in presence of death

2-cause censored time-to-event with delayed entry:
   age at entry in the cohort
   age at dementia (in between a negative and positive diagnosis)
   age at death in the two years after a negative dementia diagnosis

In each latent class $g$:

$$
\lambda_p(t)|_{c_i=g} = \lambda_{0pg}(t; )e^{\zeta_1 p \text{sex} + \zeta_2 p \text{EL} + \zeta_3 p \text{E4}}, \quad p = 1, 2
$$

$\lambda_{0pg}(t; )$ parametric hazards among Gompertz, Weibull, M-splines (5 knots) and piecewise constant (5 knots)

In separated analyses, Weibull hazards =

good balance between estimation difficulty and goodness-of-fit
Selection of the number of classes in the joint model

![Graphs showing BIC, Global score test, Dementia score test, and Death without dementia score test for different numbers of latent classes.]

1. BIC
2. Global score test
3. Dementia score test
4. Death without dementia score test

Cécile Proust-Lima (INSERM)
Joint models for multiple outcomes
July 2015
4 profiles of semantic decline, dementia & death

Joint models for multiple outcomes
Weighted predictions versus weighted observations

- Isaacs Set Test
  - Mean observation
  - Mean subject-specific prediction

- Wechsler Similarities Test

- Cumulative incidence of dementia
  - Non-parametric estimate
  - 95% bands
  - Prediction

- Cumulative incidence of free-dementia death
  - Class 1 (12.1%)
  - Class 2 (52.2%)
  - Class 3 (11.2%)
  - Class 4 (24.5%)
Individual dynamic prediction: the principle

From a landmark age $s$
- history of the markers $Y_i^{(s)} = \{Y_{kij} \text{ such as } t_{kij} \leq s\}$
- history of the covariates $X_i^{(s)} = \{X_{L1i}(t_{kij}), Z_i(t_{kij}), X_{L2i}(t_{kij}) \text{ such as } t_{kij} \leq s\}$
- other time-independent covariates $X_i$

Cumulative incidence for cause $p$ at an horizon of $t$ years

$$P_{pi}(s, t) = P(T_i \leq s + t, \delta_i = p | T_i > s, Y_i^{(s)}, X_i^{(s)}, X_{Ti}, X_{ci}; \theta)$$

- computed using the Bayes theorem and the conditional independence assumption
- with class-specific and cause-specific cumulative incidence approximated by Gauss-Legendre

Monte-Carlo approximation of the posterior distribution of $P_{pi}(s, t)$
Example of individual dynamic prediction

5-year probability of dementia (%):
- 65 years old: 13.0 [7.7, 21.0]
- 80 years old: 16.4 [9.1, 29.4]

5-year probability of death (%):
- 65 years old: 25.1 [18.1, 36.5]
- 80 years old: 36.0 [25.9, 46.3]

Cécile Proust-Lima (INSERM)

Joint models for multiple outcomes

July 2015
Example of individual dynamic prediction

5-year probability of dementia (%) : \(13.0 \, [7.7,21.0]\)
5-year probability of death (%) : \(25.1 \, [18.1,36.5]\)
Example of individual dynamic prediction

5-year probability of dementia (%): [at 80 years old: 13.0 [7.7, 21.0], at 85 years old: 16.4 [9.1, 29.4]]
5-year probability of death (%): [at 80 years old: 25.1 [18.1, 36.5], at 85 years old: 36.0 [25.9, 46.3]]
Option 2: multistate model with interval censoring

(Rouanet et al., ArXiv 2015)

Transition intensity from state $k$ to state $l$ for subject $i$ in class $g$:

$$\alpha_{klg}(t) = \alpha_{klg}^0(t) e^{X_{Ti}' \gamma_{klg}}$$

- $\alpha_{klg}^0$: class-specific baseline intensity
- $\gamma_{klg}$: class-specific regression parameters
Maximum Likelihood Estimation

- Likelihood function using $Y_i \parallel_{c_i} T_i$,

$$L(\theta) = \prod_{i=1}^{N} \sum_{g=1}^{G} f_Y(Y_i|c_i = g; \theta) \ f_D(D_i, \delta_i|c_i = g; \theta) \ P(c_i = g|X_i; \theta)$$

- $P(c_i = g|X_i; \theta)$ from a multinomial logistic model
- $f_Y(Y_i|c_i = g; \theta)$ from the latent process mixed model
- $f_D(D_i, \delta_i|c_i = g; \theta)$ from the multistate model with interval censoring

\[ D_i^T = (T_{0i}, L_i, R_i, \delta_i^A, T_i, \delta_i^D) \]
with $R_i = +\infty$ if $\delta_i^A = 0$.

$$= e^{-A_{01g}(T) - A_{02g}(T)} \alpha_{02g}(T) + \int_{L_i}^{T} e^{-A_{01g}(u) - A_{02g}(u)} \alpha_{01g}(u) e^{-(A_{12g}(T) - A_{12g}(u))} \alpha_{12g}(T) du$$
Application

- From PAQUID study (N=3777)
  - Change over time of Isaacs set test (verbal fluency)
  - in association with dementia and death

- Mixed model:

\[
\Lambda(t)_{c_i=g} = b_{0ig} + \beta_{1g} \text{sex} + \beta_{2g} \text{EL} + (b_{1ig} + \beta_{3g} \text{EL}) \times t + (b_{2ig} + \beta_{4g} \text{EL}) \times t^2
\]

\[
Y_{ij} = \Lambda_i(t_{ij}) + \epsilon_{ij} \quad \text{with} \quad b_{ig} \sim \mathcal{N}(\mu_g, B) \& \epsilon_i \sim \mathcal{N}(0, \sigma^2)
\]

- Proportional transition intensities of the illness-death model:

\[
\alpha_{klig}(t) = \alpha_{klg}^0(t) \ e^{\gamma_{kls} \text{sex}_i + \gamma_{kle} \text{EL}_i}
\]
Four-class Markovian model: poorly educated men

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>EL=0</th>
<th>EL=1</th>
<th>men</th>
<th>women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1</td>
<td>7.3%</td>
<td>42.4</td>
<td>57.6</td>
<td>42.8</td>
<td>57.2</td>
</tr>
<tr>
<td>Class 2</td>
<td>8.3%</td>
<td>25.0</td>
<td>75.0</td>
<td>43.1</td>
<td>56.9</td>
</tr>
<tr>
<td>Class 3</td>
<td>34.2%</td>
<td>29.3</td>
<td>70.7</td>
<td>39.4</td>
<td>60.6</td>
</tr>
<tr>
<td>Class 4</td>
<td>50.5%</td>
<td>37.6</td>
<td>62.4</td>
<td>43.9</td>
<td>56.1</td>
</tr>
</tbody>
</table>
Goodness-of-fit assessment

Class-specific weighted predicted trajectories vs. observed

predicted class-specific cumulative incidences vs. semi-parametric estimator
Concluding remarks

- Joint model methodology
  ▶ extended to multivariate longitudinal markers
  ▶ extended to multistate process for events
  → useful for different purposes in chronic diseases

- Different assumptions for the shared quantity
  ▶ depends on the data
  ▶ depends on the objective

- Parametric assumptions
  ▶ flexible and/or selected distributions according to the data
  ▶ progressive construction of the models, goodness-of-fit (graphs, measures, tests):
  ▶ ... score tests for conditional independence assumptions (between events and marker, between events)
Funding and references

Fundings: INCa/Inserm MULTIPLE, Inserm/Region PhD grant

Further details in:


Sène Mbéry et al. (2014). Statistical Methods in Medical Research. (online)