Joint models for a longitudinal marker and multivariate survival data

Loïc Ferrer

Institut Curie, INSERM U900, Saint-Cloud, France

email: loic.ferrer@curie.fr

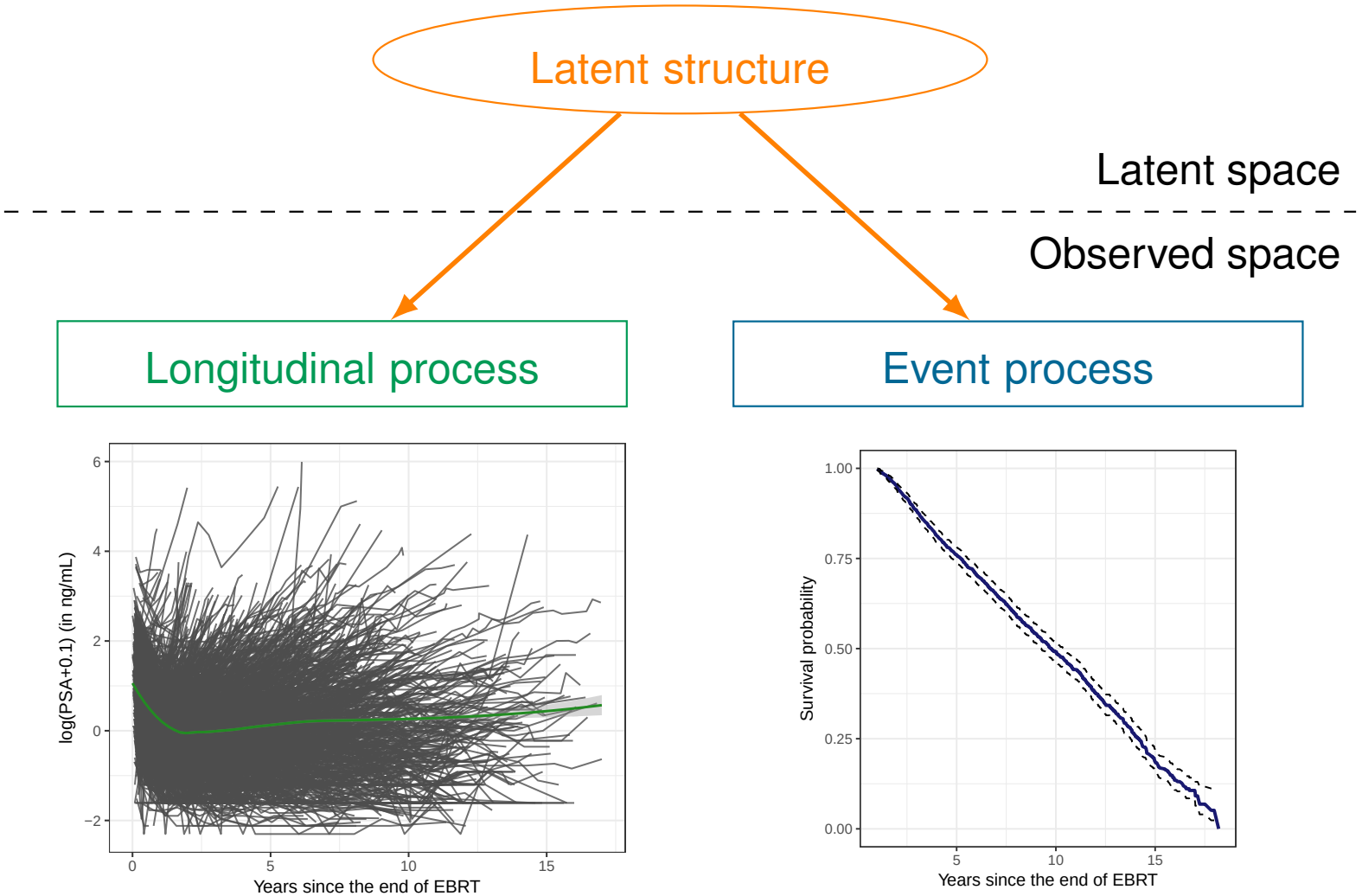
8th scientific day of the SMAC club

Bordeaux

January 24, 2018

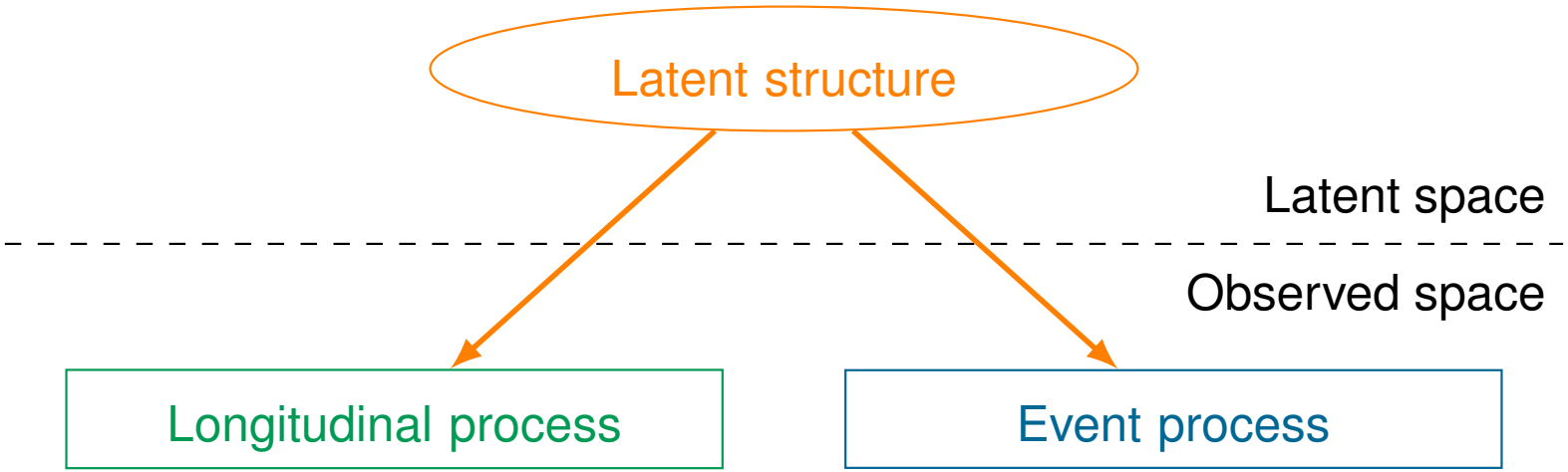
Joint modelling principle

Simultaneous modelling of correlated longitudinal and event processes



Joint modelling principle

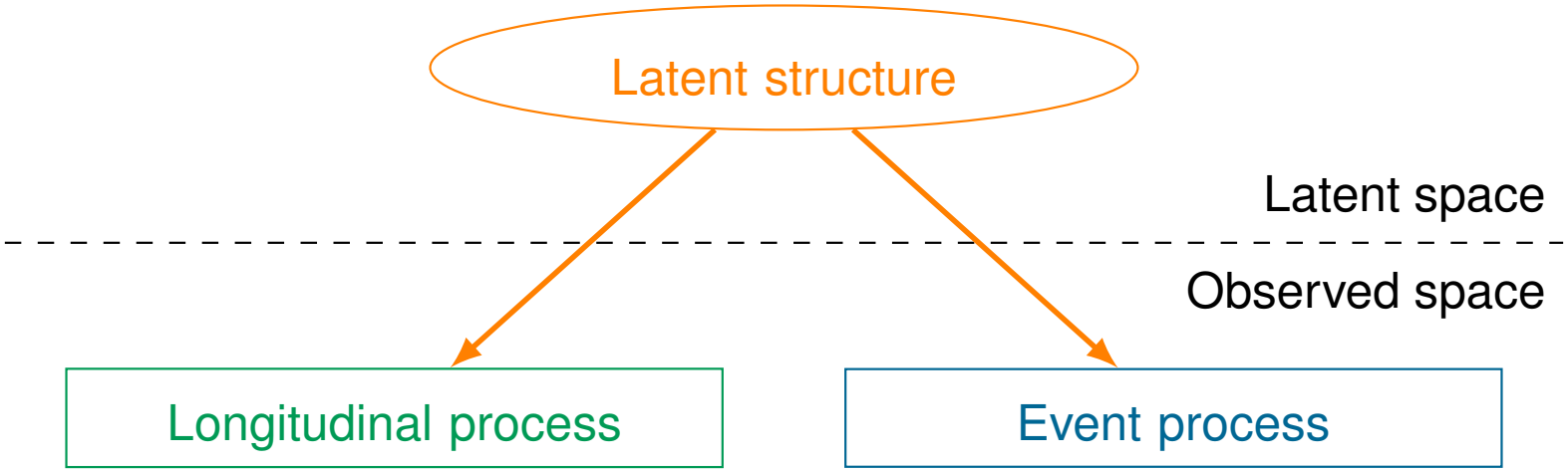
Simultaneous modelling of correlated longitudinal and event processes



- Objectives:
  - Describe the longitudinal process stopped by the event
  - Explore the association between the two processes
  - Predict the risk of event adjusted for the longitudinal process

Joint modelling principle

Simultaneous modelling of correlated longitudinal and event processes

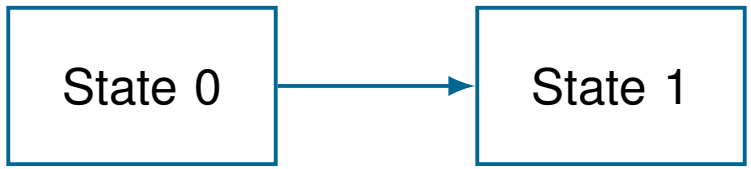


- Latent structure:
  - Function of shared random effects (shared random effect models) [1]
    - Homogeneous population
    - Specification and quantification of the association between the two processes
  - Latent classes (joint latent class models) [2]
    - Heterogeneous population
    - No assumption on the association

1. [Rizopoulos, 2012]
2. [Proust-Lima et al., 2014]

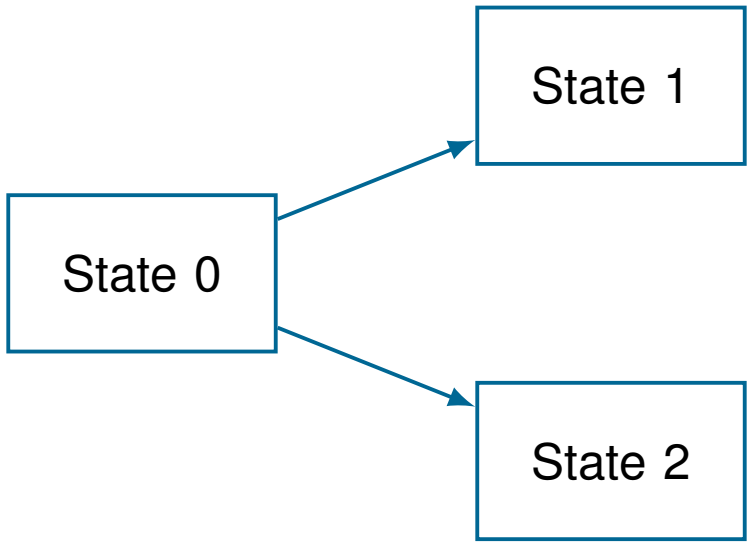
Classical joint modelling

- Classical joint models are developed for
  - A Gaussian longitudinal marker
  - Survival data with one single type of event



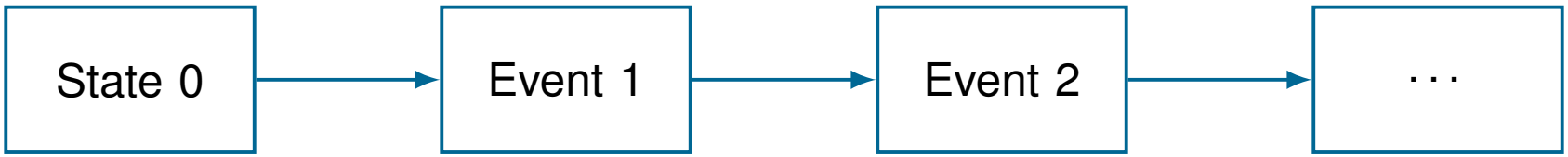
Extension of the classical joint models

- Many applications require joint models with
  - More complex longitudinal processes
  - More complex event processes
    - Survival data with competing events
    - Recurrent events
    - Multi-state process with possible multiple transitions



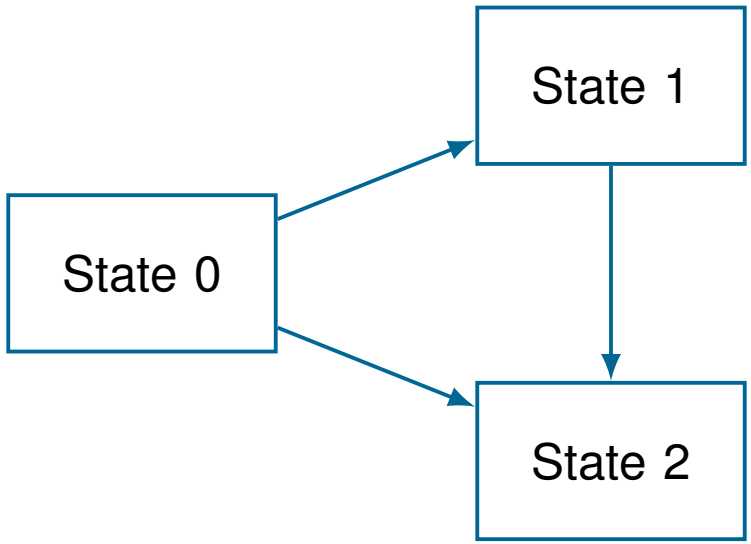
Extension of the classical joint models

- Many applications require joint models with
  - More complex longitudinal processes
  - More complex event processes
    - Survival data with competing events
    - Recurrent events
    - Multi-state process with possible multiple transitions



Extension of the classical joint models

- Many applications require joint models with
  - More complex longitudinal processes
  - More complex event processes
    - Survival data with competing events
    - Recurrent events
    - Multi-state process with possible multiple transitions



Joint multi-state models

- In the literature, very few papers have focused on the succession of events in the joint model setting [3]
- Focus on
  - A joint multi-state model with shared random effects [4]
    - application: link between PSA & multiple clinical progressions in prostate cancer
  - A joint multi-state model with latent classes [5]
    - application: distinction of profiles of cognitive decline associated with risks of dementia and death in elderly people

3. [Hickey et al., 2018]
4. [Ferrer et al., 2016]
5. [Rouanet et al., 2016]

Research Article

Statistics in Medicine

Received 3 June 2015, Accepted 24 March 2016 Published online in Wiley Online Library

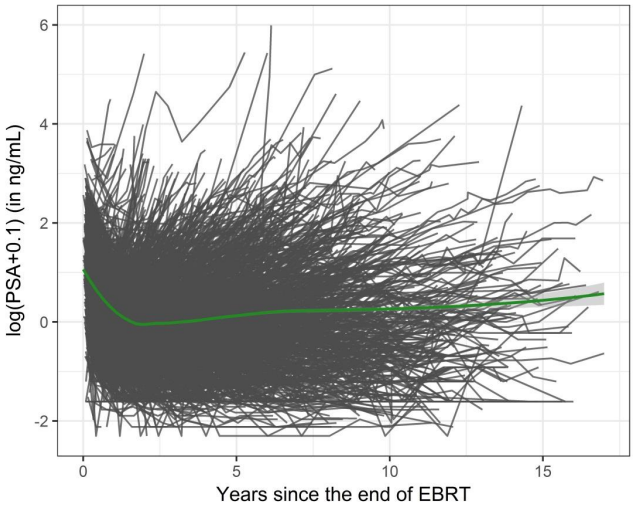
(wileyonlinelibrary.com) DOI: 10.1002/sim.6972

Joint modelling of longitudinal and multi-state processes: application to clinical progressions in prostate cancer

Loïc Ferrer,[a*†] Virginie Rondeau,[a] James Dignam,[b] Tom Pickles,[c] Hélène Jacqmin-Gadda[a] and Cécile Proust-Lima[a]

Classical modelling in prostate cancer

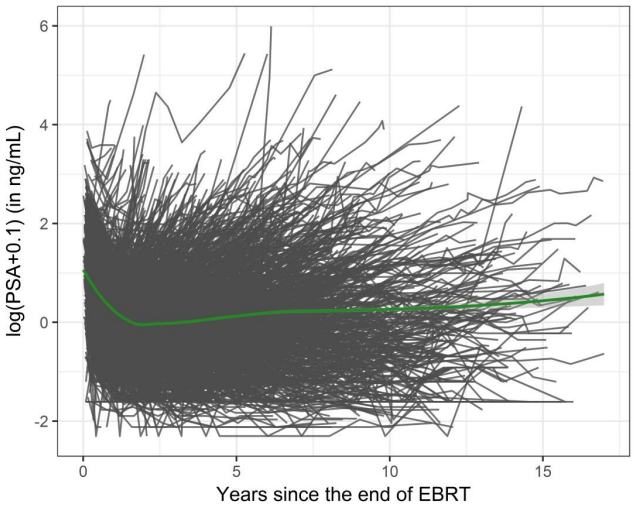
- Longitudinal PSA & clinical progression

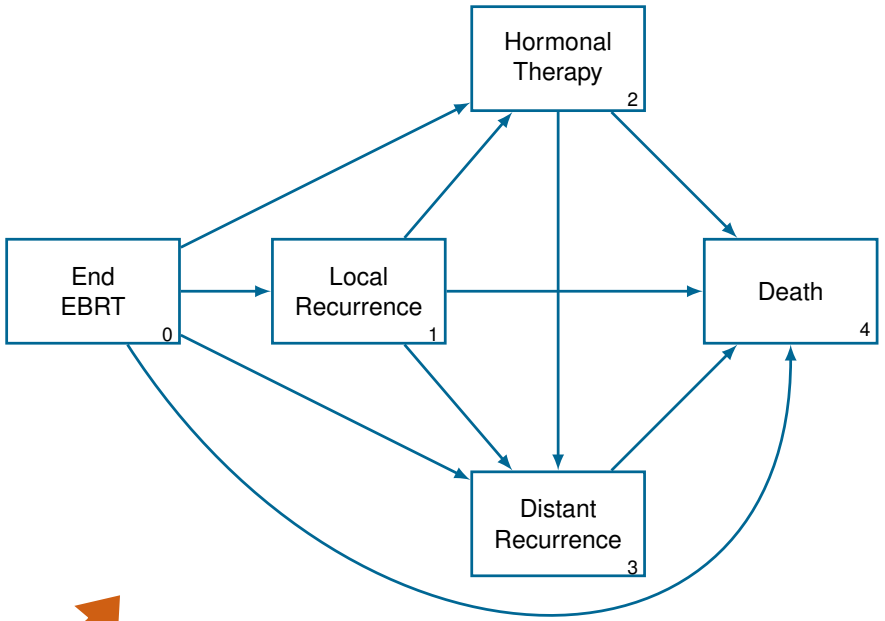


- No distinction between the clinical progressions
- No modelling of the full disease progression

Multiple clinical progressions in cancer

- Longitudinal PSA & multiple clinical progressions





- Objectives
  - → Understanding the PSA evolution during the patient's follow-up
  - → Distinguishing & evaluating the impact of PSA dynamics and other prognostic factors on disease evolution

Notations

- Let us consider two observed processes: one longitudinal and one multi-state
- Multi-state process
  - $E_i = \{E_i(t), T_{i0} \leq t \leq C_i\}$ non-homogeneous Markovian process
    - $E_i(t)$ with values in the finite state space $S = \{0, 1, \ldots, M\}$
    - T_{i0} left truncature time, C_i right censoring time
  - $T_i = (T_{i1}, \ldots, T_{im_i})^\top$ the m_i observed time(s); $T_{ir} < T_{i(r+1)}, \forall r \in S$
  - $\delta_i = (\delta_{i1}, \ldots, \delta_{im_i})^\top$ the vector of indicators of observed transition(s)
- Longitudinal process
  - $Y_i = (Y_{i1}, \ldots, Y_{in_i})^\top$ the n_i measure(s) of marker collected at time(s) $t_{i1}, \ldots, t_{in_i}$, with $t_{in_i} \leq T_{im_i}$

Joint multi-state model

$$\begin{cases} Y_{ij} &= Y_i^*(t_{ij}) + \epsilon_{ij} \\ &= X_i^L(t_{ij})^\top \beta + Z_i(t_{ij})^\top b_i + \epsilon_{ij} \\ \\ \lambda_{hk}^i(t) &= \lim_{\mathrm{d}t \to 0} \dfrac{\Pr(E_i(t+\mathrm{d}t) = k | E_i(t) = h)}{\mathrm{d}t} \\ &= \lambda_{hk,0}(t) \exp(X_{hk,i}^{E\top} \gamma_{hk} + W_{hk,i}(b_i, t)^\top \eta_{hk}), \quad \text{for } (h,k) \in S^2, \end{cases}$$

- $b_i \sim \mathcal{N}_q(0, B)$, $\epsilon_i = (\epsilon_{i1}, \ldots, \epsilon_{in_i})^\top \sim \mathcal{N}_{n_i}(0, \sigma^2 I)$, $b_i \perp\!\!\!\perp \epsilon_i$
- $\lambda_{hk,0}(t)$ parametric baseline intensity, $X_{hk,i}^E$ prognostic factors
- $W_{hk,i}(b_i, t)$ structure of dependence, e.g.
  - $W_{hk,i}(b_i, t) = Y_i^*(t)$ $\longrightarrow$ (true current level)
  - $W_{hk,i}(b_i, t) = \partial Y_i^*(t)/\partial t$ $\longrightarrow$ (true current slope)
  - $W_{hk,i}(b_i, t) = \left(Y_i^*(t), \partial Y_i^*(t)/\partial t\right)^\top$ $\longrightarrow$ (both)
  - ...

Estimation and implementation

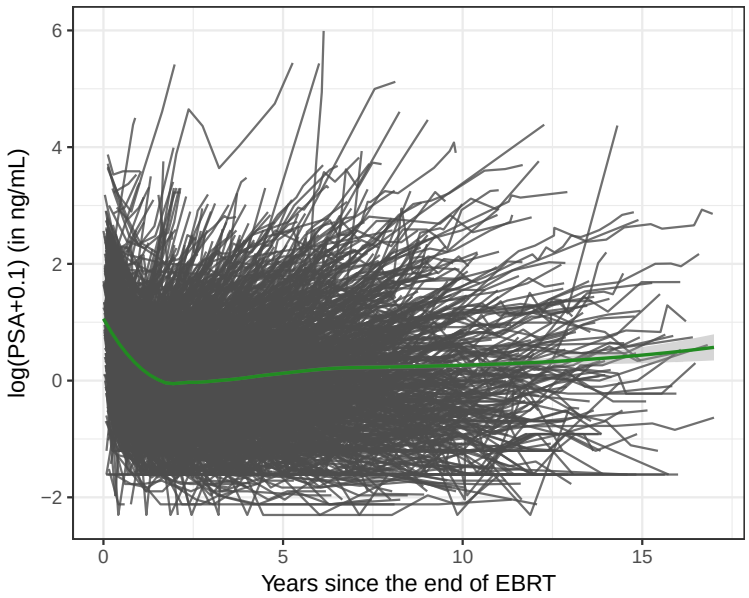
- Maximum likelihood approach
  - Using $Y_i \perp\!\!\!\perp_{b_i} E_i$
  - Likelihood function

$$L(\theta) = \prod_{i=1}^{N} \int_{\mathbb{R}^q} f_Y(Y_i|b_i;\theta)\ f_E(E_i|b_i;\theta)\ f_b(b_i;\theta)\,\mathrm{d}b_i$$

- Implementation in R
  - Combination and extension of the existing R packages `JM` and `mstate`
  - Codes with detailed examples available at `https://github.com/LoicFerrer/JMstateModel`
  - Likelihood computed and optimised using
    - numerical integration algorithms (Gaussian quadratures: multi-step pseudo-adaptive Gauss-Hermite quadratures for the integral over random effects)
    - optimisation algorithms (EM + quasi-Newton)

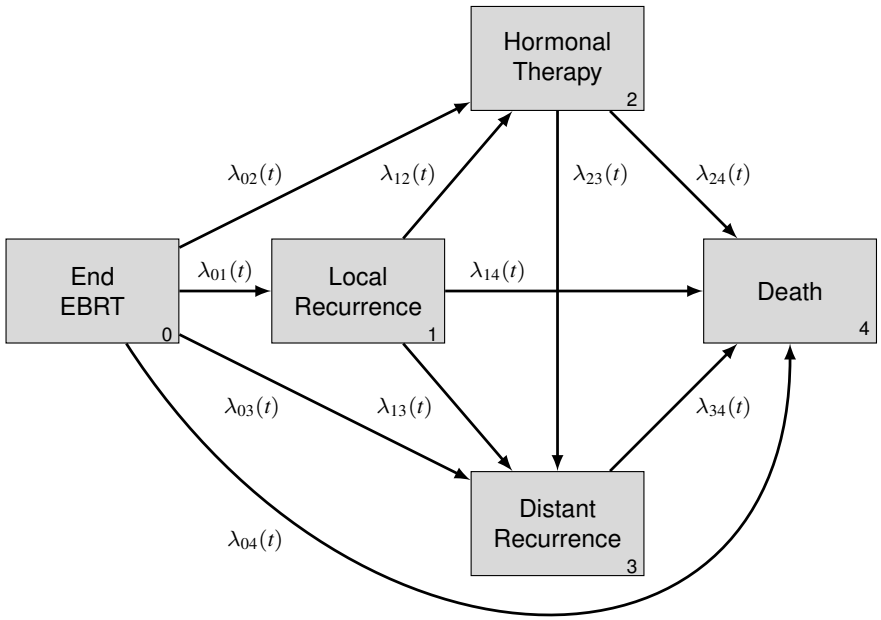
Application

- 2 cohorts of men with localised prostate cancer treated by radiotherapy (N=1474)
- Longitudinal biomarker: PSA
- Multi-state representation of the clinical progressions



10, [3–21] measures per patient

(50th, [5th – 95th] %iles)



$$\Upsilon = \begin{pmatrix} 533 & 144 & 227 & 47 & 523 \\ 0 & 20 & 90 & 10 & 24 \\ 0 & 0 & 106 & 33 & 178 \\ 0 & 0 & 0 & 13 & 77 \\ 0 & 0 & 0 & 0 & 802 \end{pmatrix}$$

matrix of direct transitions

Specification of the joint multi-state model

- Model inspired from the literature with one unique type of event

$$
\begin{cases}
\begin{aligned}
Y_{ij} &= Y_i^*(t_{ij}) + \epsilon_{ij} \\
&= \left(\beta_0 + X_i^{L0\top}\beta_{0,\texttt{cov}} + b_{i0}\right) + \\
&\quad \left(\beta_1 + X_i^{L1\top}\beta_{1,\texttt{cov}} + b_{i1}\right) \times \left((1+t_{ij})^{-1.2} - 1\right) + \\
&\quad \left(\beta_2 + X_i^{L2\top}\beta_{2,\texttt{cov}} + b_{i2}\right) \times t_{ij} + \epsilon_{ij}
\end{aligned} \\
\lambda_{hk}^i(t) = \lambda_{hk,0}(t)\exp\left(X_{hk,i}^{E\top}\gamma_{hk} + \begin{pmatrix} g(Y_i^*(t)) \\ \partial Y_i^*(t)/\partial t \end{pmatrix}^\top \begin{pmatrix} \eta_{hk,\texttt{level}} \\ \eta_{hk,\texttt{slope}} \end{pmatrix}\right)
\end{cases}
$$

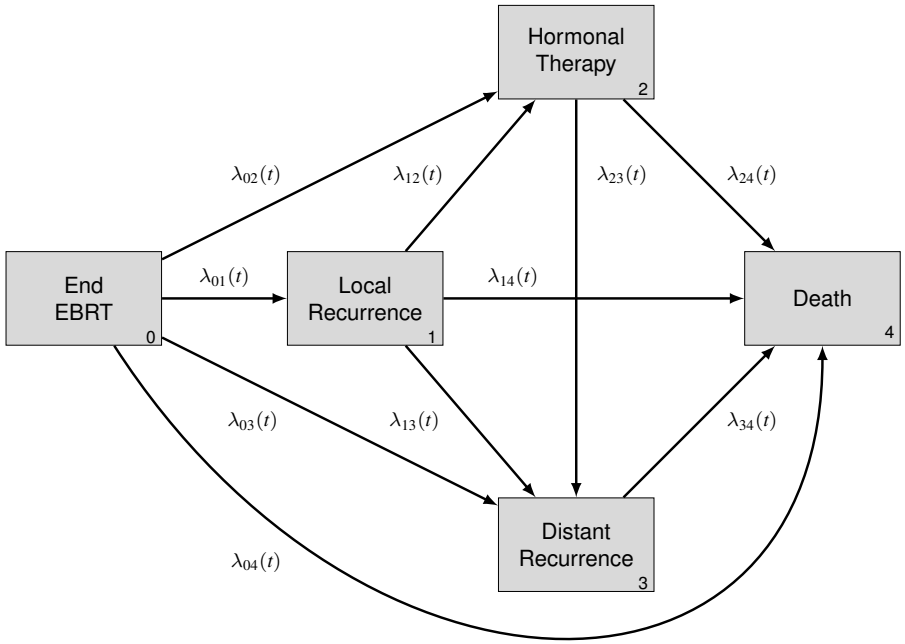
- $g(Y_i^*(t)) = \text{logit}^{-1}((Y_i^*(t) - 0.71)/0.44)$
- $\epsilon_i \sim \mathcal{N}(0, \sigma^2 I_{n_i})$
- $b_i = (b_{i0}, b_{i1}, b_{i2})^\top \sim \mathcal{N}(0, D)$, D unstructured

Results

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

	Value	StdErr	p-value
Level : 01	3.32	0.41	< 0.001
Level : 02	4.89	0.39	< 0.001
Level : 03	2.94	0.68	< 0.001
Level : 04	-0.41	0.23	0.071
Level : 12	1.90	0.83	0.023
Level : 13	-2.30	1.32	0.081
Level : 14	-0.07	0.88	0.939
Level : 23	-0.29	1.04	0.778
Level : 24	-0.48	0.62	0.440
Level : 34	-0.02	0.57	0.974
.....			
Slope : 01	1.33	0.33	< 0.001
Slope : 02	1.60	0.24	< 0.001
Slope : 03	1.74	0.54	0.001
Slope : 04	0.59	0.35	0.088
Slope : 12	0.46	0.58	0.336
Slope : 13	3.82	1.07	< 0.001
Slope : 14	0.70	1.02	0.495
Slope : 23	0.23	0.51	0.651
Slope : 24	0.64	0.23	0.005
Slope : 34	-0.56	0.42	0.186

Multi-state representation of the clinical progressions



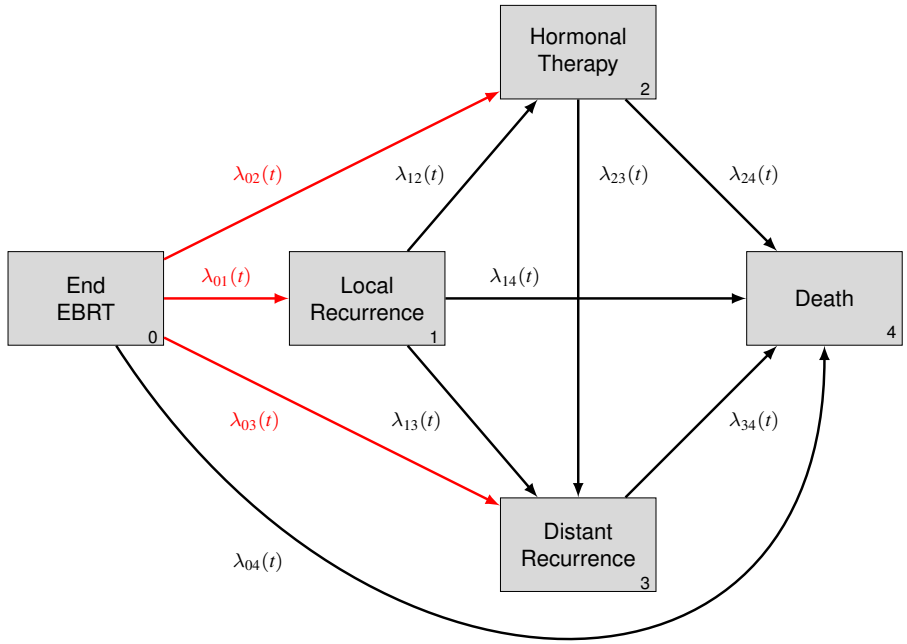
- Prognostic factors: advanced initial stage not always associated with intensities of transitions between health states after adjustment on PSA

Results

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

	Value	StdErr	p-value
Level : 01	3.32	0.41	< 0.001
Level : 02	4.89	0.39	< 0.001
Level : 03	2.94	0.68	< 0.001
Level : 04	-0.41	0.23	0.071
Level : 12	1.90	0.83	0.023
Level : 13	-2.30	1.32	0.081
Level : 14	-0.07	0.88	0.939
Level : 23	-0.29	1.04	0.778
Level : 24	-0.48	0.62	0.440
Level : 34	-0.02	0.57	0.974
.....			
Slope : 01	1.33	0.33	< 0.001
Slope : 02	1.60	0.24	< 0.001
Slope : 03	1.74	0.54	0.001
Slope : 04	0.59	0.35	0.088
Slope : 12	0.46	0.58	0.336
Slope : 13	3.82	1.07	< 0.001
Slope : 14	0.70	1.02	0.495
Slope : 23	0.23	0.51	0.651
Slope : 24	0.64	0.23	0.005
Slope : 34	-0.56	0.42	0.186

Multi-state representation of the clinical progressions



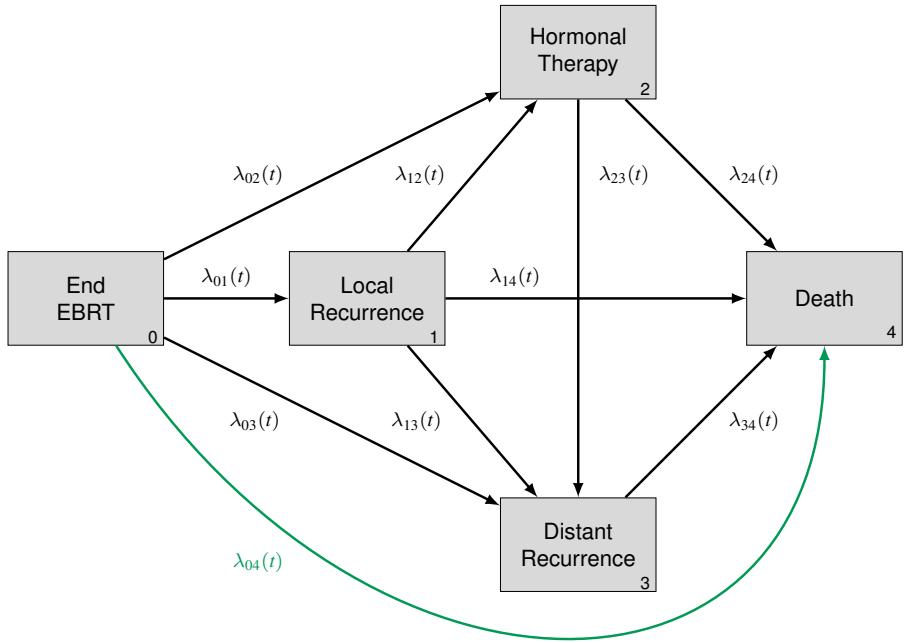
- Prognostic factors: advanced initial stage not always associated with intensities of transitions between health states after adjustment on PSA

Results

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

	Value	StdErr	p-value
Level : 01	3.32	0.41	< 0.001
Level : 02	4.89	0.39	< 0.001
Level : 03	2.94	0.68	< 0.001
Level : 04	-0.41	0.23	0.071
Level : 12	1.90	0.83	0.023
Level : 13	-2.30	1.32	0.081
Level : 14	-0.07	0.88	0.939
Level : 23	-0.29	1.04	0.778
Level : 24	-0.48	0.62	0.440
Level : 34	-0.02	0.57	0.974
.....			
Slope : 01	1.33	0.33	< 0.001
Slope : 02	1.60	0.24	< 0.001
Slope : 03	1.74	0.54	0.001
Slope : 04	0.59	0.35	0.088
Slope : 12	0.46	0.58	0.336
Slope : 13	3.82	1.07	< 0.001
Slope : 14	0.70	1.02	0.495
Slope : 23	0.23	0.51	0.651
Slope : 24	0.64	0.23	0.005
Slope : 34	-0.56	0.42	0.186

Multi-state representation of the clinical progressions



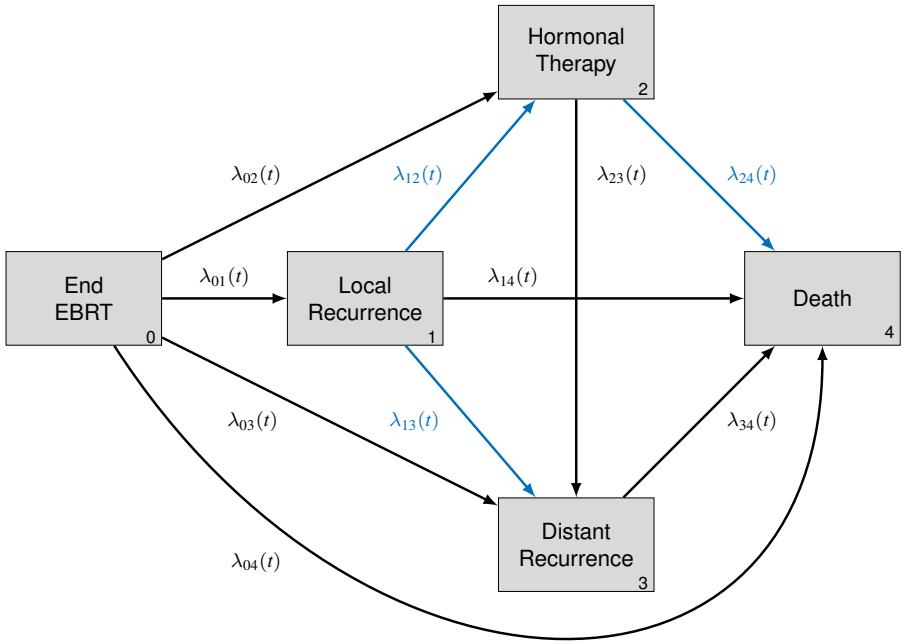
- Prognostic factors: advanced initial stage not always associated with intensities of transitions between health states after adjustment on PSA

Results

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

	Value	StdErr	p-value
Level : 01	3.32	0.41	< 0.001
Level : 02	4.89	0.39	< 0.001
Level : 03	2.94	0.68	< 0.001
Level : 04	-0.41	0.23	0.071
Level : 12	1.90	0.83	0.023
Level : 13	-2.30	1.32	0.081
Level : 14	-0.07	0.88	0.939
Level : 23	-0.29	1.04	0.778
Level : 24	-0.48	0.62	0.440
Level : 34	-0.02	0.57	0.974
.....			
Slope : 01	1.33	0.33	< 0.001
Slope : 02	1.60	0.24	< 0.001
Slope : 03	1.74	0.54	0.001
Slope : 04	0.59	0.35	0.088
Slope : 12	0.46	0.58	0.336
Slope : 13	3.82	1.07	< 0.001
Slope : 14	0.70	1.02	0.495
Slope : 23	0.23	0.51	0.651
Slope : 24	0.64	0.23	0.005
Slope : 34	-0.56	0.42	0.186

Multi-state representation of the clinical progressions

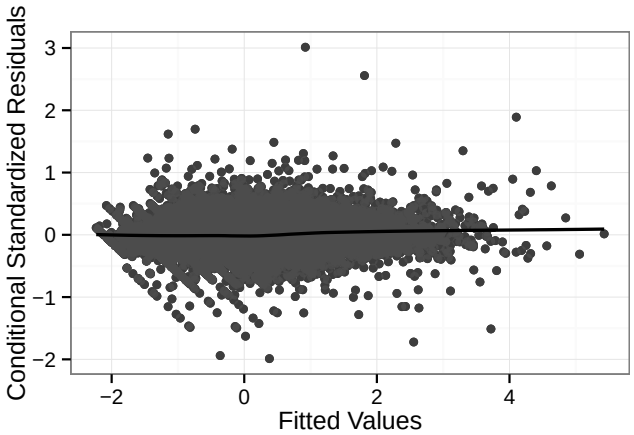


- Prognostic factors: advanced initial stage not always associated with intensities of transitions between health states after adjustment on PSA

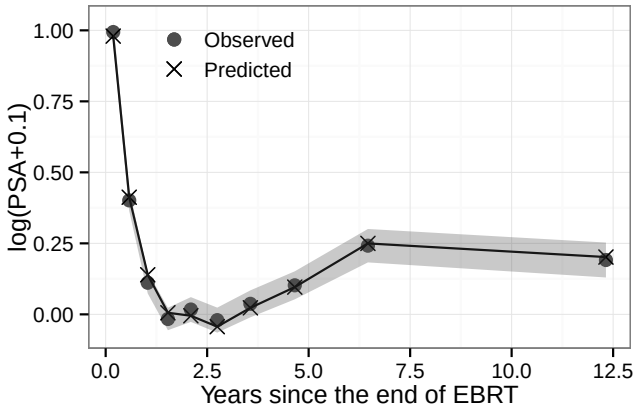
Model diagnostics

- Goodness-of-fit plots for the longitudinal process

  - Conditional standardized residuals versus fitted values

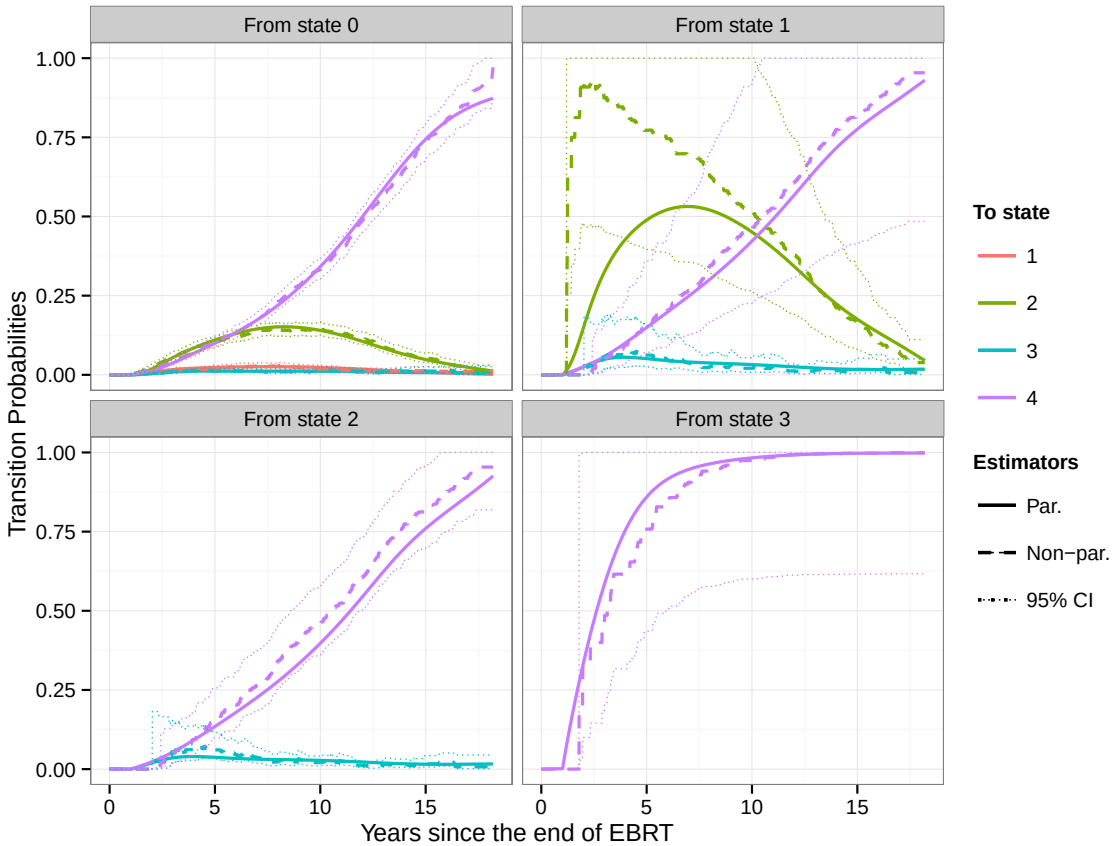


  - Observed and predicted values of the biomarker



Model diagnostics (cont'd)

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



Biometrics DOI: 10.1111/biom.12530

Joint Latent Class Model for Longitudinal Data and Interval-Censored Semi-Competing Events: Application to Dementia

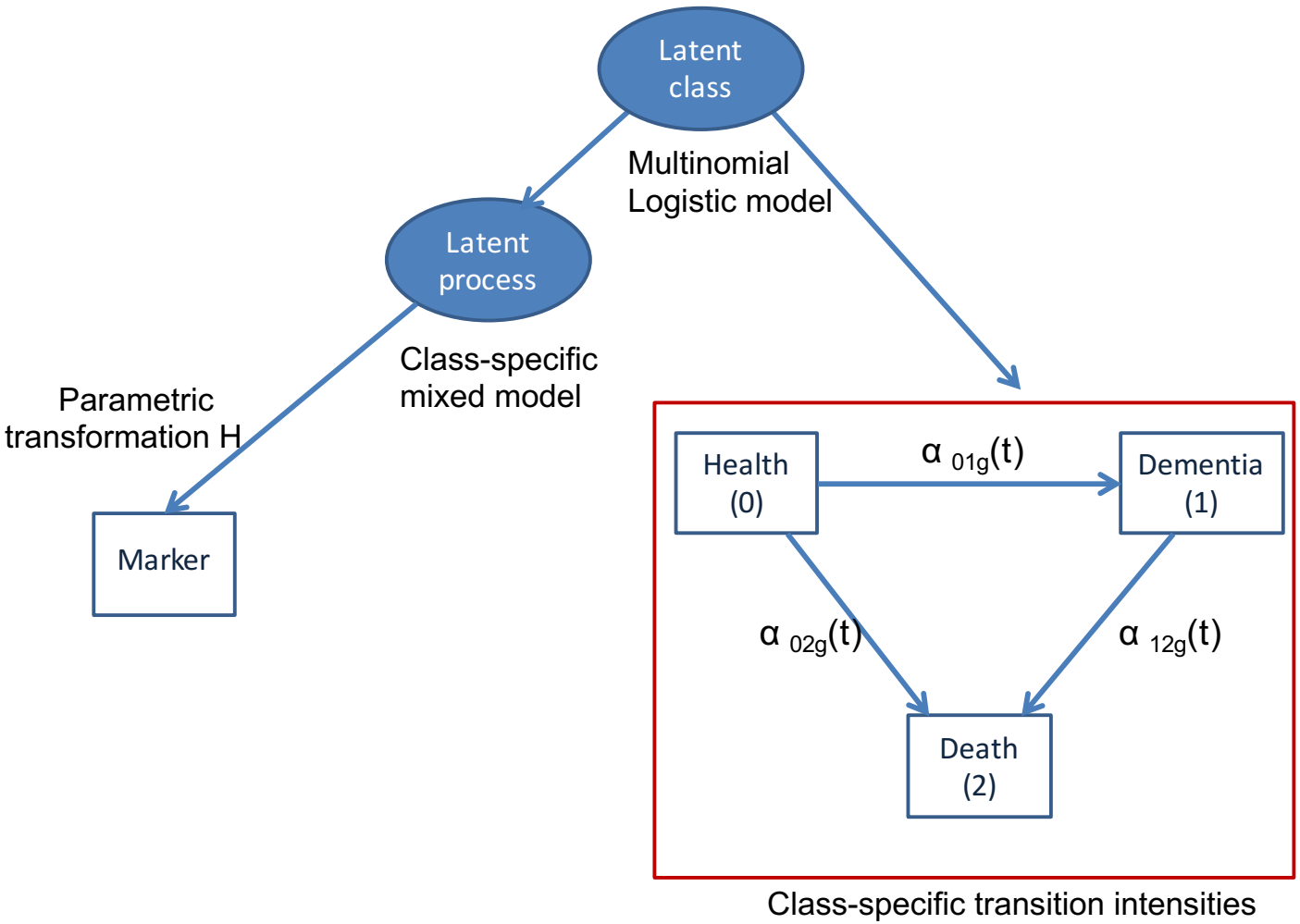
Anaïs Rouanet,[1,2,*] Pierre Joly,[1] Jean-François Dartigues,[1,2] Cécile Proust-Lima,[1,2] and Hélène Jacqmin-Gadda[1,2]

[1]INSERM, Centre INSERM U1912 - Epidemiologie - Biostatistiques, F-33076 Bordeaux, France
[2]Université de Bordeaux, ISPED, 146 rue Léo Saignat, F-33076 Bordeaux, France
[*]*email:* anais.rouanet@isped.u-bordeaux2.fr

Application to dementia

- Aim: To distinguish different profiles of cognitive decline associated with risks of dementia and death
- Data: Paquid Cohort (French prospective cohort: Normal and pathological brain ageing)
  - 3777 subjects from Dordogne and Gironde, aged 65 and over
  - Visits every 2/3 years during 25 years
  - Study of Isaacs Set Test [0-40], verbal fluency
- Methodological challenges
  - Heterogeneity in cognitive decline
  - Correlation between cognitive decline and occurrence of dementia
  - Competing risk of death
  - Time-to-dementia onset interval-censored

Joint latent class illness-death model



Joint model formulation

- Membership probability: $p_{ig} = P(c_i = g|X_{pi})$
- Latent process Λ_i, given the class g:

$$\begin{aligned}\Lambda_i(t_{ij}|c_i = g) &= f_1(X_{ij};\ \beta_g) + f_2(Z_{ij};\ \beta_g)\ u_{ig} \\ &= X_{ij}^T \beta_g + Z_{ij}^T u_{ig}\end{aligned}$$

f_1, f_2: (possibly nonlinear) functions of time, covariates
β_g: class-specific parameters
$u_{ig} \sim \mathcal{N}(0, \sigma_g^2 B)$

- Transformed gaussian marker $\tilde{Y}$:

$$\tilde{Y}_{ij} = H(Y_{ij}; \eta) = \Lambda_i(t_{ij}) + \epsilon_{ij} \text{ with } \epsilon_{ij} \sim \mathcal{N}(0, \sigma_e^2)$$

$H(.; \eta)$: Parametric transformation

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

$$\alpha_{klig}(t) = \alpha_{klg}^0(t)\ e^{X_{ei}^\top \gamma_{klg}}$$

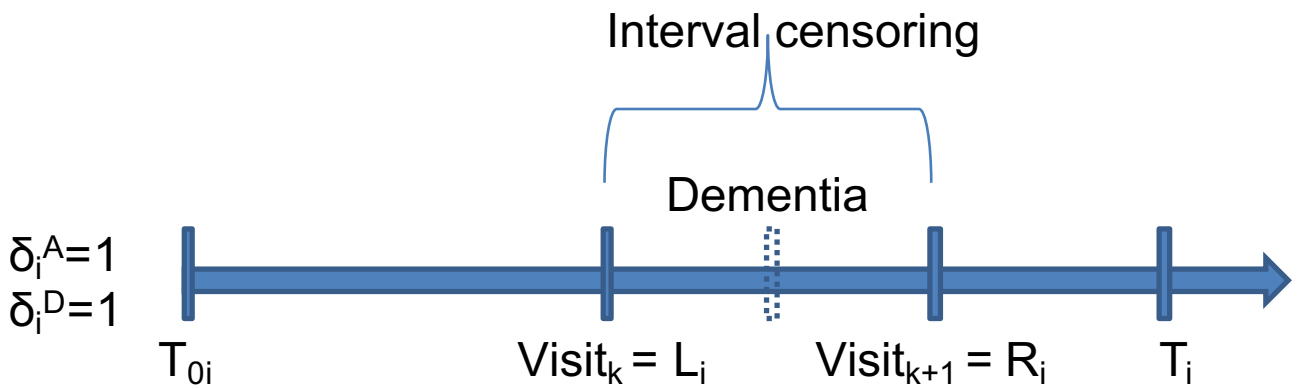
α_{klg}^0: class-specific baseline intensity
γ_{klg}: class-specific regression parameters

Estimation

- Maximum likelihood approach
  - Using $Y_i \perp\!\!\!\perp_g D_i$
  - Log-likelihood function

$$\mathcal{L}(\theta_G) = \sum_{i=1}^{N} \log \left[\sum_{g=1}^{G} p_{ig}\, f(Y_i|c_i = g;\theta_G)\; P(D_i|c_i = g;\theta_G) \right]$$

$$- \sum_{i=1}^{N} \log \left[\sum_{g=1}^{G} p_{ig}\, \mathrm{e}^{-A_{01ig}(T_{0i};\theta_G) - A_{02ig}(T_{0i};\theta_G)} \right]$$



- Marquardt algorithm for a fixed number of latent classes G
- G chosen by Bayesian Information Criterion (BIC) minimisation

Model specification

- Mixed model, given latent class g :

$$\Lambda_i(t) = \beta_{0g} + u_{ig}^{(0)} + [\beta_{1g} - \beta_{2g} + u_{ig}^{(1)}]\ (t - \tau_g) + \beta_3\ Educ_i + (\beta_4 - \beta_5)\ Educ_i\ (t - \tau_g) + \beta_6\ Sex_i \quad \text{if } t \leq \tau_g$$

$$\Lambda_i(t) = \beta_{0g} + u_{ig}^{(0)} + [\beta_{1g} + \beta_{2g} + u_{ig}^{(2)}]\ (t - \tau_g) + \beta_3\ Educ_i + (\beta_4 + \beta_5)\ Educ_i\ (t - \tau_g) + \beta_6\ Sex_i \quad \text{if } t \geq \tau_g$$

$u_{ig} = (u_{ig}^{(0)}, u_{ig}^{(1)}, u_{ig}^{(2)})^\top \sim \mathcal{N}(0, \sigma_g^2 B)$, $\epsilon_{ij} \sim \mathcal{N}(0, \sigma_\epsilon^2)$

$\tilde{Y}_{ij} = H(Y_{ij}; \eta) = \Lambda_i(t_{ij}) + \epsilon_{ij}$ with H : Beta cumulative distribution function

- Transition intensities from states $k \to l$ of the multi-state model:

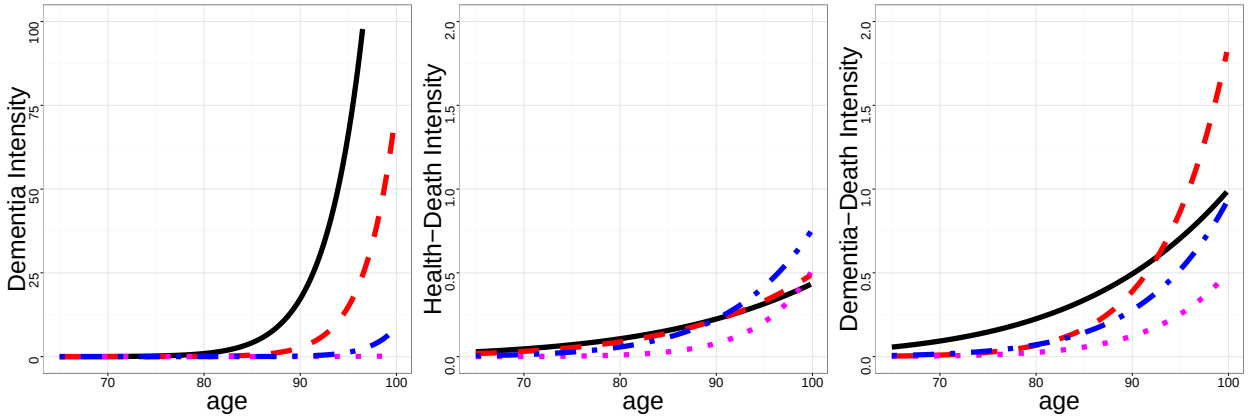
$$\alpha_{klig}(t) = \alpha_{klg}^0(t)\ \mathrm{e}^{\gamma_{kls}\ Sex_i + \gamma_{kle}\ Educ_i}$$

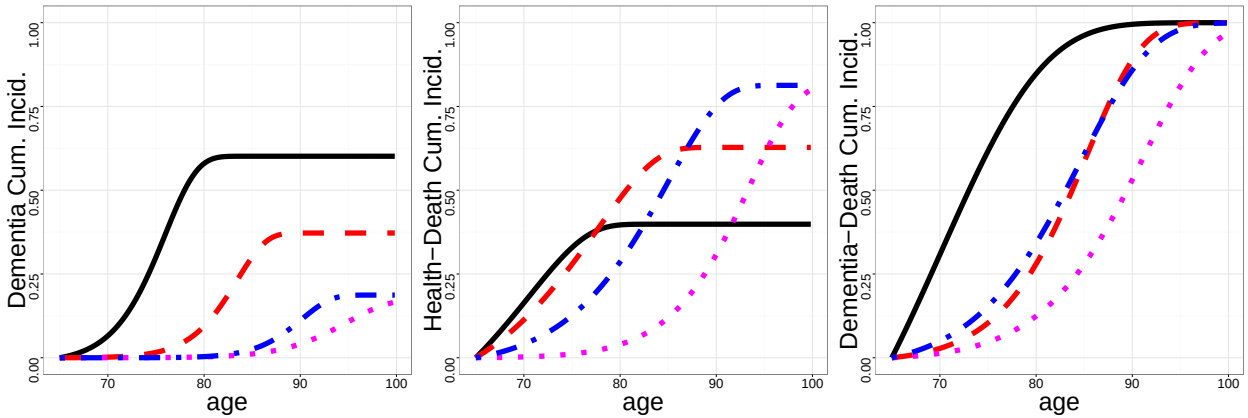
Model choice

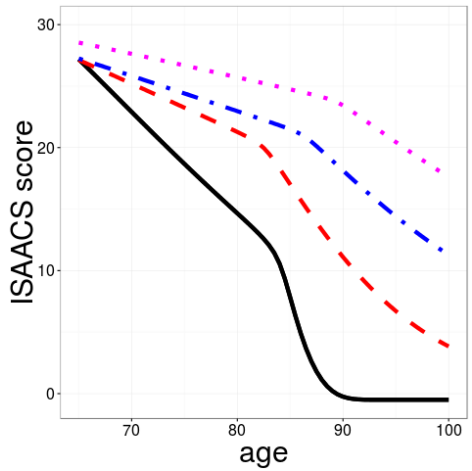
	BIC	
	Markovian	Semi-markovian
G=1	106901	107055
G=2	106270	106356
G=3	106081	106177
G=4	106005	106091
G=5	106027	106107

Mortality among subjects with dementia depends more on age than on the duration of dementia.

The minimum value of BIC with G = 4 classes.







	N	CEP=0	CEP=1	men	women
Class 1	11.12%	23.2	76.8	40.1	59.9
Class 2	32.91%	29.7	70.3	40.7	59.3
Class 3	9.73%	42.9	57.1	42.0	58.0
Class 4	46.24%	38.0	62.0	43.8	56.2

Discrimination of the posterior classification

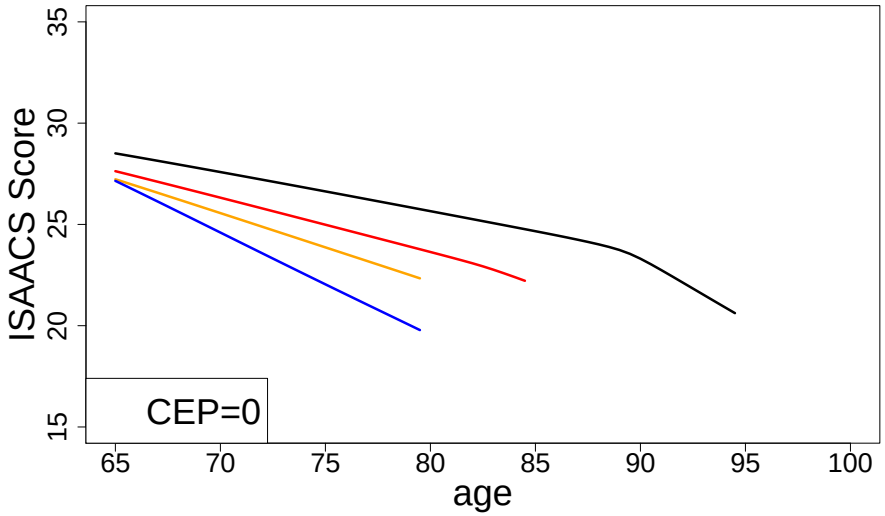
Classification according to:

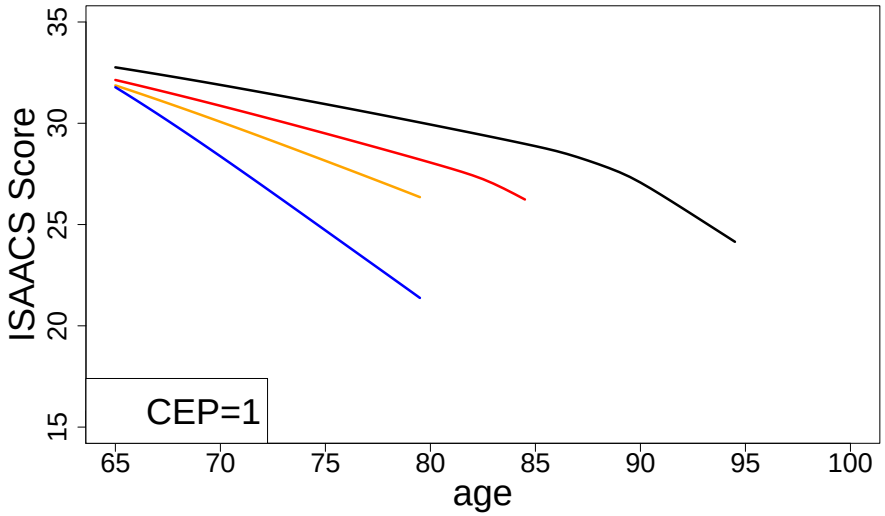
$$\hat{\pi}_{ig}^{Y,D} = P(c_i = g|Y_i, D_i, \hat{\theta}_G) = \frac{f(Y_i|c_i = g; \hat{\theta}_G)P(D_i|c_i = g; \hat{\theta}_G)\ P(c_i = g; \hat{\theta}_G)}{\sum_l^G f(Y_i|c_i = l; \hat{\theta}_G)P(D_i|c_i = l; \hat{\theta}_G)\ P(c_i = l; \hat{\theta}_G)}$$

Class	1	2	3	4
1	71.36	21.92	0.18	6.54
2	12.83	61.54	1.24	24.39
3	0.01	0.47	79.11	20.40
4	0.96	19.11	12.96	66.97

Table 1 – Mean probabilities to belong to each class according to the posterior classification.

Estimated trajectories according to dementia onset & death





Man with a low/high level of education
in 4 different cases:

man alive and dementia-free at 95
$E(Y(t)|T_i^A > 95, T_i^D > 95, \delta_i^A = 0, \delta_i^D = 0; \hat{\theta}_G)$

man alive and dementia-free at 85

man dead dementia-free at 80

man alive with dementia at 80

Discussion

- Joint models extended to several kinds of multivariate survival data
  - See Hickey G.L. et al. (2018) for a full state-of-the-art
- Ferrer et al. (2016) developed a joint multi-state model with shared random effects to
  - Model the disease evolution in its whole
  - Using an easy-to-use function implemented in R
- Rouanet et al. (2016) developed a joint latent class illness-death model accounting for
  - Heterogeneity in the data
  - Competing risk of death
  - Interval censoring
- Model diagnostics
  - Study of the residuals for validating the model assumptions
  - Graphical comparison of the observations and predictions of the model as goodness-of-fit tool

Discussion (cont'd)

- Goodness-of-fit assessment in joint multi-state models using a score test for the inclusion of a Gaussian frailty term
  - for shared random effect models
  - for joint latent class models
- Useful to validate the model assumptions and check its goodness-of-fit
  - Markovian assumption
  - Incomplete adjustment on covariates
  - Presence of non-linear covariate effect
  - ~~Violation of the proportional intensities assumption~~
- Available at `http://github.com/LoicFerrer/JMstateModel/` for shared random effect models
- Submitted for publication in a few days

References

[1.] Rizopoulos D. (2012). *Joint models for longitudinal and time-to-event data: With applications in R*. Chapman and Hall/CRC.

[2.] Proust-Lima C., Séne M., Taylor J.M. & Jacqmin-Gadda H. (2014). Joint latent class models for longitudinal and time-to-event data: A review. *Statistical Methods in Medical Research*, 23(1):74–90.

[3.] Hickey G.L., Philipson P., Jorgensen A. & Kolamunnage-Dona R. (2018). Joint models of longitudinal and time-to-event data with more than one event time outcome: a review. *The International Journal of Biostatistics*, 14(1).

[4.] Ferrer L., Rondeau V., Dignam J., Pickles T., Jacqmin-Gadda, H. & Proust-Lima C. (2016). Joint modelling of longitudinal and multi-state processes: application to clinical progressions in prostate cancer. *Statistics in Medicine*, 35(22):3933–3948.

[5.] Rouanet A., Joly P., Dartigues J-F., Proust-Lima C. & Jacqmin-Gadda H. (2016). Joint Latent Class Model for Longitudinal Data and Interval-Censored Semi-Competing Events: Application to Dementia. *Biometrics*, 72(4):1123–1135.

Numerical approximation of the integral over the random effects

- Likelihood function

$$L(\theta) = \prod_{i=1}^{N} \int_{\mathbb{R}^q} f_Y(Y_i|b_i;\theta)\ f_E(E_i|b_i;\theta)\ f_b(b_i;\theta)\,\mathrm{d}b_i$$

- Adaptive Gauss-Hermite rule $\rightarrow$ Centering and rescaling the integral around its modal value at each step of the optimisation algorithm

Numerical approximation of the integral over the random effects

- Likelihood function

$$L(\theta) = \prod_{i=1}^{N} \int_{\mathbb{R}^q} f_Y(Y_i|b_i;\theta)\ f_E(E_i|b_i;\theta)\ f_b(b_i;\theta)\,\mathrm{d}b_i$$

- Adaptive Gauss-Hermite rule $\rightarrow$ Centering and rescaling the integral around its modal value at each step of the optimisation algorithm
  1. Pseudo-adaptive GH rule: based on the posterior distribution of the random effects from the LMM

$$\widetilde{b}_i = \arg\max_b \{\log f(Y_i, b; \widetilde{\theta}_Y)\}$$

  and their associated covariance matrix

Numerical approximation of the integral over the random effects

- Likelihood function

$$L(\theta) = \prod_{i=1}^{N} \int_{\mathbb{R}^q} f_Y(Y_i|b_i;\theta)\ f_E(E_i|b_i;\theta)\ f_b(b_i;\theta)\,\mathrm{d}b_i$$

- Adaptive Gauss-Hermite rule $\rightarrow$ Centering and rescaling the integral around its modal value at each step of the optimisation algorithm
  1. Pseudo-adaptive GH rule: based on the posterior distribution of the random effects from the LMM

$$\widetilde{b}_i = \arg\max_b \{\log f(Y_i, b; \widetilde{\theta}_Y)\}$$

     and their associated covariance matrix

  2. Multi-step pseudo-adaptive GH rule: based on the posterior distribution of the random effects from the JM

$$\overset{\star}{b}_i = \arg\max_b \{\log f(E_i, Y_i, b; \overset{\star}{\theta})\}$$

     and their associated covariance matrix

Joint multi-state model – Implementation in R (1/4)

- Example of R code

$$\begin{cases} Y_{ij} &= Y_i^*(t_{ij}) + \epsilon_{ij} \\ &= (\beta_0 + \beta_{0,X} X_i + b_{i0}) + (\beta_1 + \beta_{1,X} X_i + b_{i1}) \times t_{ij} + \epsilon_{ij} \\ \lambda_{hk}^i(t|b_i) &= \lambda_{hk,0}(t) \exp\left(\gamma_{hk} X_i + \eta_{hk,\texttt{level}} Y_i^*(t) + \eta_{hk,\texttt{slope}} \partial Y_i^*(t)/\partial t\right) \end{cases}$$

where the multi-state process included three states ($(h,k) \in \{0,1,2\}^2$) and three transitions ($0 \to 1,\ 0 \to 2,\ 1 \to 2$),
the log-baseline intensities are a linear combination of cubic-splines

```
library(mstate)
library(JM)
source("JMstateModel.R")
load("data.RData")

# Initialisation of the longitudinal sub-part
lmeFit <- lme(fixed = Y ~ 1 + X + time + X:time,
              data = data_long,
              random = ~ (1 + time) | id,
              method = "REML",
              control = list(opt = "optim"))
```

Joint multi-state model – Implementation in R (2/4)

```
# Data peparation to the multi-state framework
tmat <- matrix(NA, 3, 3)
tmat[1, 2:3] <- 1:2
tmat[2, 3] <- 3
dimnames(tmat) <- list(from = c("State_0", "State_1", "State_2"),
                       to = c("State_0", "State_1", "State_2"))

covs <- "X"

data_mstate <-
  msprep(time = c(NA, "time_of_State_1", "time_of_State_2"),
         status = c(NA, "State_1", "State_2"),
         data = data_surv,
         trans = tmat,
         keep = covs,
         id = "id")

data_mstate <- expand.covs(data_mstate, covs,
                           append = TRUE, longnames = FALSE)
```

Joint multi-state model – Implementation in R (3/4)

```
# Initialisation of the multi-state sub-part
coxFit <- coxph(Surv(Tstart, Tstop, status) ~
                    X.1 + X.2 + X.3 + strata(trans),
                data = data_mstate,
                method = "breslow", x = TRUE, model = TRUE)
```

Joint multi-state model – Implementation in R (4/4)

```
dForm <- list(fixed = ~ 1 + X,
              indFixed = c(3, 4),
              random = ~ 1,
              indRandom = 2)

# Joint multi-state model
jointFit_both <-
  JMstateModel(lmeObject = lmeFit,
               survObject = coxFit,
               timeVar = "time",
               parameterization = "both",
               method = "spline-PH-aGH",
               interFact = list(value = ~ strata(trans) - 1,
                                slope = ~ strata(trans) - 1,
                                data = data_mstate),
               derivForm = dForm,
               Mstate = TRUE,
               data.Mstate = data_mstate,
               ID.Mstate = "id",
               control = list(GHk = 9, lng.in.kn = 3))
```

Joint multi-state model with an additional frailty

- Model formulation

$$\begin{cases} Y_{ij} &= X_i^L(t_{ij})^\top \beta + Z_i(t_{ij})^\top b_i + \epsilon_{ij} \\ \lambda_{hk}^i(t) &= \lambda_{hk,0}(t) \exp\left(X_{hk,i}^{E\top} \gamma_{hk} + W_{hk,i}(b_i, t)^\top \eta_{hk} + v_i\right), \quad \text{for } (h,k) \in S^2 \end{cases}$$

  - $v_i \sim \mathcal{N}(0, \sigma_v^2)$ the frailty term
  - $v_i \perp\!\!\!\perp b_i$, $v_i \perp\!\!\!\perp \epsilon_i$

- Likelihood function $L(\sigma_v^2, \theta)$

$$L(\sigma_v^2, \theta) = \prod_{i=1}^{N} \int_{\mathbb{R}^{q_b}} f_Y(Y_i|b_i;\theta) f_b(b_i;\theta) \int_{\mathbb{R}} f_E(E_i|b_i, v_i;\theta) f_v(v_i;\theta)\, \mathrm{d}v_i\, \mathrm{d}b_i$$

Score test: score statistic and its variance

$$H_0 : \sigma_v^2 = 0 \qquad \text{vs} \qquad H_1 : \sigma_v^2 > 0$$

- Score statistic $U(0,\theta) = U(\sigma_v^2 = 0, \theta)$

$$U(0,\theta) = \sum_{i=1}^{N} U_i(0,\theta) = \sum_{i=1}^{N} \frac{\partial \log L_i(\sigma_v^2, \theta)}{\partial \sigma_v^2} \Bigg|_{\sigma_v^2=0}$$

  - Analytic expression

- Asymptotic variance corrected for the estimation of the nuisance parameters

$$\text{Var}\left\{U(0,\theta)\right\} = \left(I_{\sigma_v^2\sigma_v^2} - I_{\sigma_v^2\theta} I_{\theta\theta}^{-1} I_{\theta\sigma_v^2}\right) \Big|_{\sigma_v^2=0}$$

  - Forward finite difference method

Score test: test statistic

- Test statistic

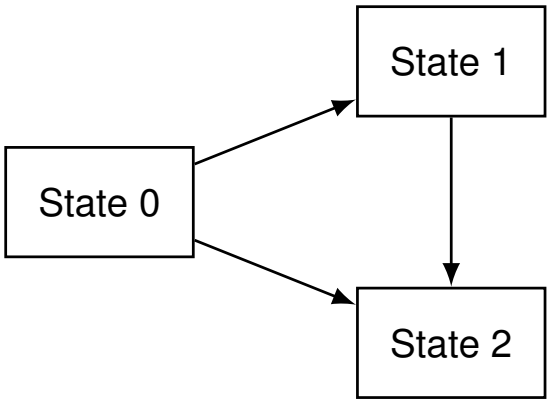
$$T = \begin{cases} 0 & \text{if} \quad U(0, \widehat{\theta}_0) \leq 0 \\ \dfrac{\left\{U(0, \widehat{\theta}_0)\right\}^2}{\text{Var}\left\{U(0, \widehat{\theta}_0)\right\}} & \text{if} \quad U(0, \widehat{\theta}_0) > 0 \end{cases}$$

  - $\widehat{\theta}_0$ the model parameters estimated under the null hypothesis
  - T follows asymptotically a mixture of chi-square distributions

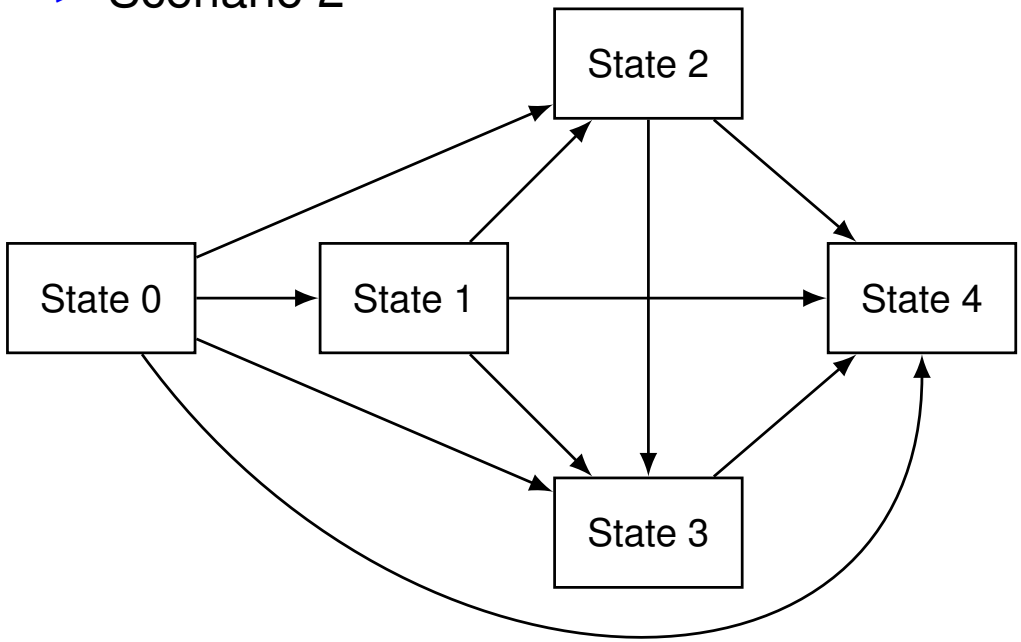
$$T \sim \frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2$$

Assessment by simulation study

- Scenario 1



- Scenario 2



- For each subject $i = 1, \ldots, N$ of the 500 replicates,

$$\begin{cases} Y_i(t) &= Y_i^*(t) + \epsilon_i(t) \\ &= (\beta_0 + \beta_{0,X} X_i + b_{i0}) + (\beta_1 + \beta_{1,X} X_i + b_{i1}) \times t + \epsilon_i(t) \\ \lambda_{hk}^i(t) &= \lambda_{hk,0}(t) \exp\left(\gamma_{hk} X_i + \eta_{hk,0} Y_i^*(t) + \eta_{hk,1} \partial Y_i^*(t)/\partial t + v_i\right) \end{cases}$$

Simulation study – results

- Empirical type-I error rate (nominal level of 5%)

	Scenario 1 (3 states) $\sigma_v^2 = 0$ $(\overline{M} = 0.70)$	Scenario 2 (5 states) $\sigma_v^2 = 0$ $(\overline{M} = 2.84)$
$N = 500$	0.008	0.028
$N = 1000$	0.010	0.054
$N = 1500$	0.020	0.060

- Empirical statistical power (nominal level of 5%)

	Scenario 1 (3 states)		Scenario 2 (5 states)	
	$\sigma_v^2 = 0.5$ $(\overline{M} = 0.74)$	$\sigma_v^2 = 1$ $(\overline{M} = 0.75)$	$\sigma_v^2 = 0.5$ $(\overline{M} = 2.73)$	$\sigma_v^2 = 1$ $(\overline{M} = 2.65)$
$N = 500$	0.278	0.438	0.884	0.990
$N = 1000$	0.568	0.850	0.998	1.000
$N = 1500$	0.846	0.970	1.000	1.000

$\overline{M}$: average number of observed direct transitions per subject

Score statistic

- Score statistic $U(0,\theta)$

$$U(0,\theta) = \sum_{i=1}^{N} \frac{1}{2L_i(0,\theta)} \times$$

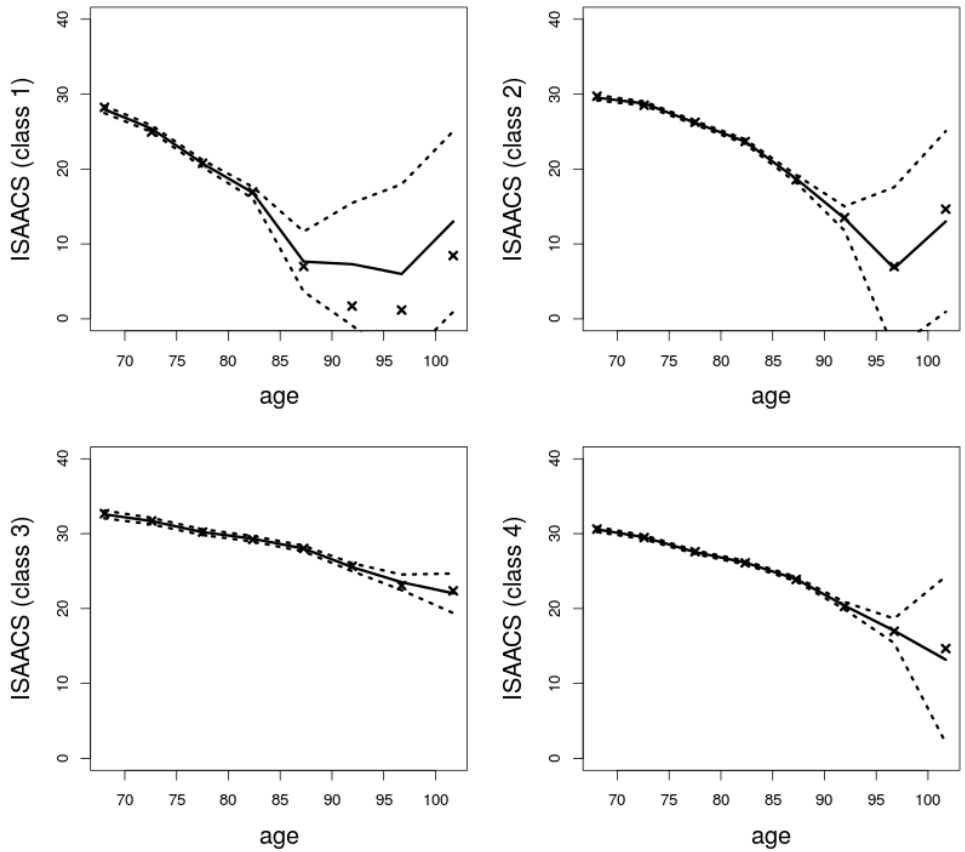
$$\int_{\mathbb{R}^{q_b}} f_Y(Y_i|b_i;\theta) f_E(E_i|b_i;0,\theta) f_b(b_i;\theta) \times$$

$$\left\{ \left[\sum_{r=0}^{m_i-1} \left(\delta_{i(r+1)} + \Lambda^i_{E_i(T_{ir}),E_i(T_{ir})} \left(T_{ir}, T_{i(r+1)}|b_i;0,\theta \right) \right) \right]^2 + \right.$$

$$\left. \sum_{r=0}^{m_i-1} \left(\Lambda^i_{E_i(T_{ir}),E_i(T_{ir})} \left(T_{ir}, T_{i(r+1)}|b_i;0,\theta \right) \right) \right\} \mathrm{d}b_i$$

$$= \sum_{i=1}^{N} \frac{1}{2} \int_{\mathbb{R}^{q_b}} f_b(b_i|Y_i,E_i;0,\theta) \{ \ldots \} dbi$$

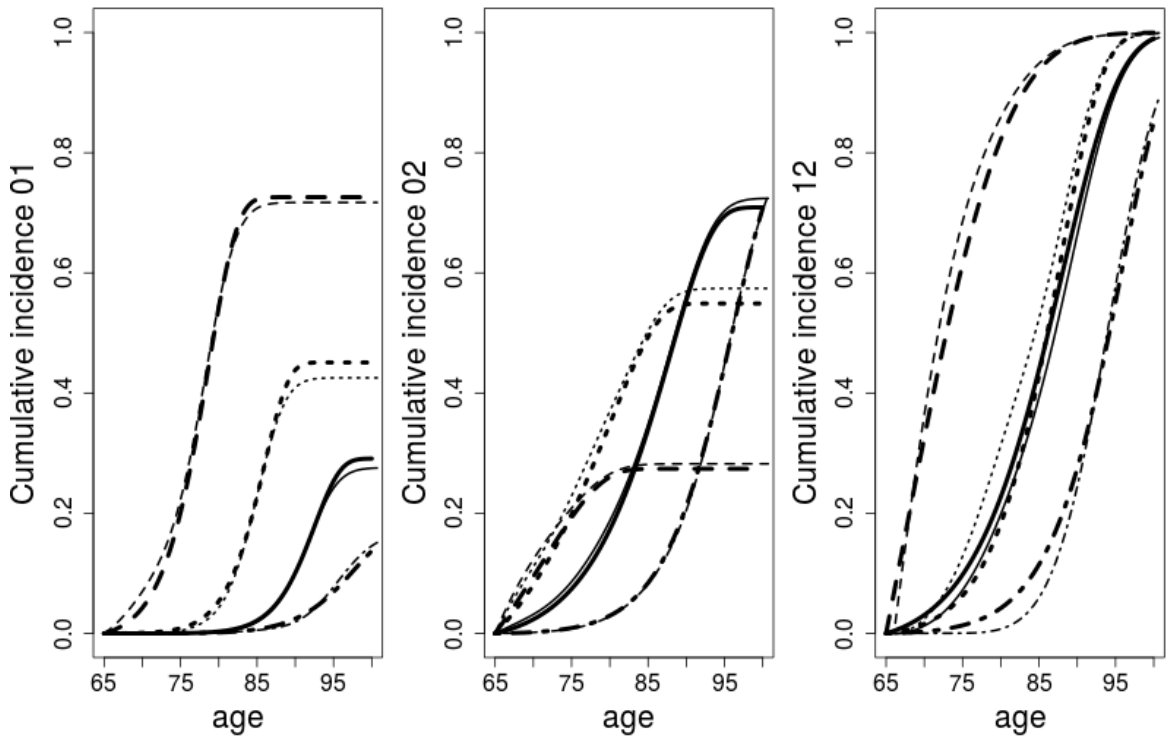
Goodness-of-fit of the longitudinal predictions



$$\hat{\mu}^{o}_{gl} = \frac{\sum_{(i,j)|\tau_l<t_{ij}<\tau_{l+1}} Y_{ij}\, P(c_i = g|Y_i, D_i; \hat{\theta}_G)}{\sum_{(i,j)|\tau_l<t_{ij}<\tau_{l+1}} P(c_i = g|Y_i, D_i; \hat{\theta}_G)}$$

$$\hat{\mu}^{u}_{gl} = \frac{\sum_{(i,j)|\tau_l<t_{ij}<\tau_{l+1}} E(Y_{ij}|c_i = g, \hat{u}_{ig}; \hat{\theta}_G)\, P(c_i = g|Y_i, D_i; \hat{\theta}_G)}{\sum_{(i,j)|\tau_l<t_{ij}<\tau_{l+1}} P(c_i = g|Y_i, D_i; \hat{\theta}_G)}, \text{ with } \hat{u}_{ig} = E(u_{ig}|Y_i, \hat{\theta}_G)$$

Goodness-of-fit of the illness-death predictions



Class-specific cumulative incidences, marginal on covariates:
Joint latent class illness-death model vs. Semi-parametric illness-death model with baseline transition intensities modeled by M-splines.