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

Cécile Proust-Lima

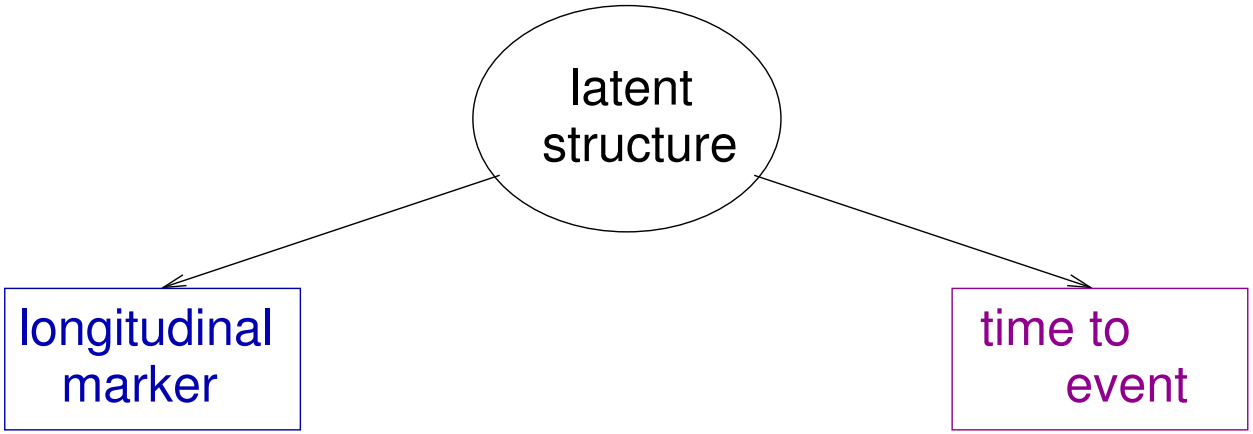
works with Loïc Ferrer, Anaïs Rouanet
and Hélène Jacqmin-Gadda

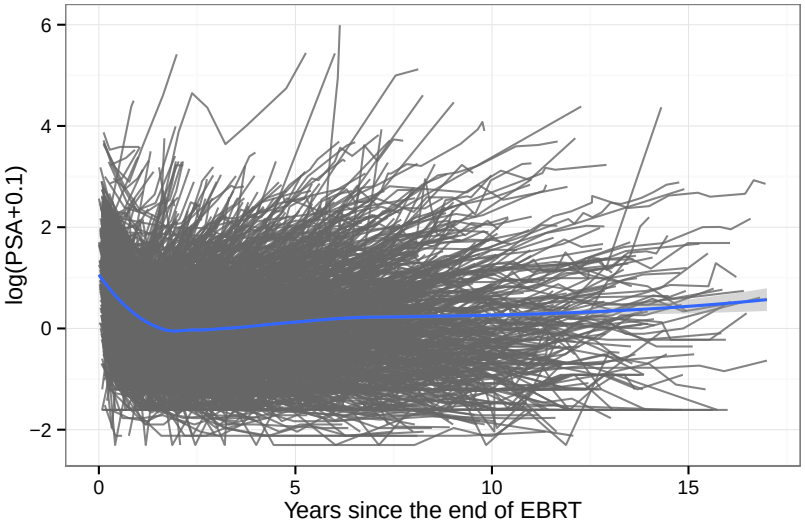
INSERM U897, Epidemiology and Biostatistics, Bordeaux, France
Univ. Bordeaux, ISPED, Bordeaux, France
cecile.proust-lima@inserm.fr

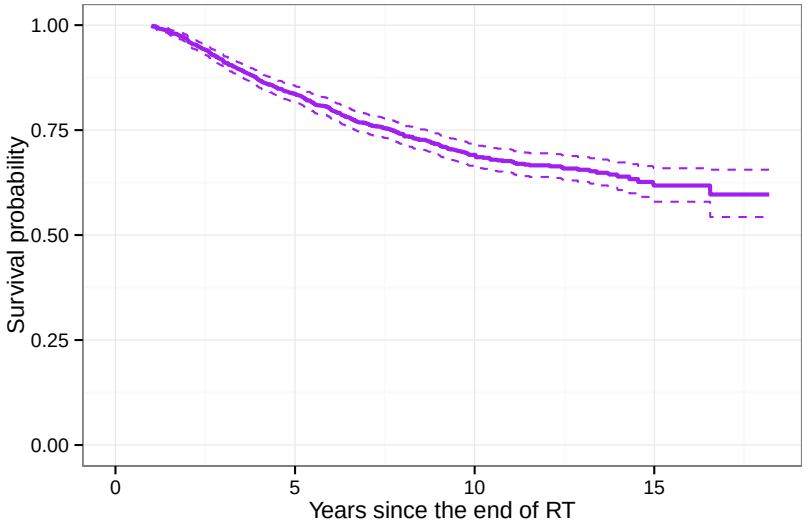
Workshop on Flexible Models for Longitudinal and Survival Data with Applications in Biostatistics

Joint modelling principle

Simultaneous modelling of correlated longitudinal and survival data

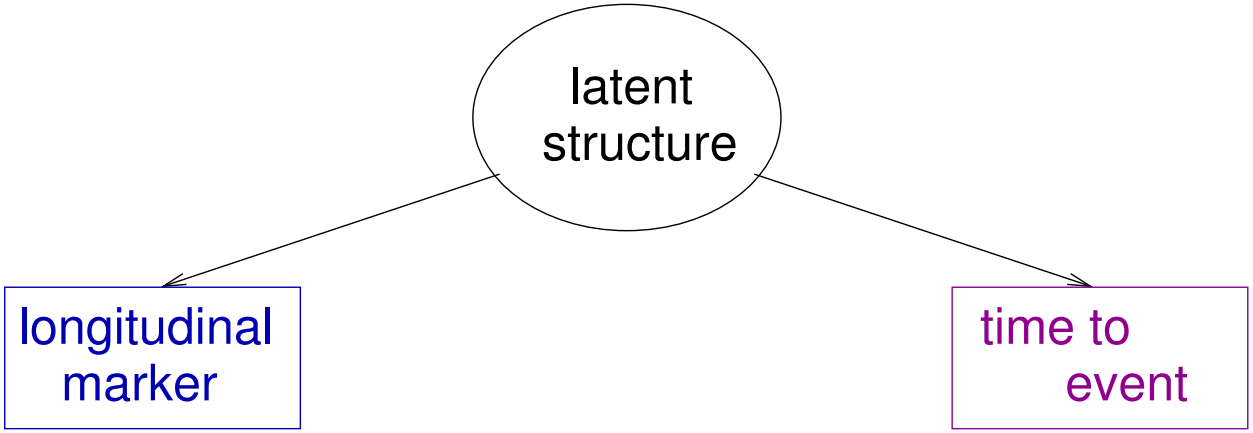






Joint modelling principle

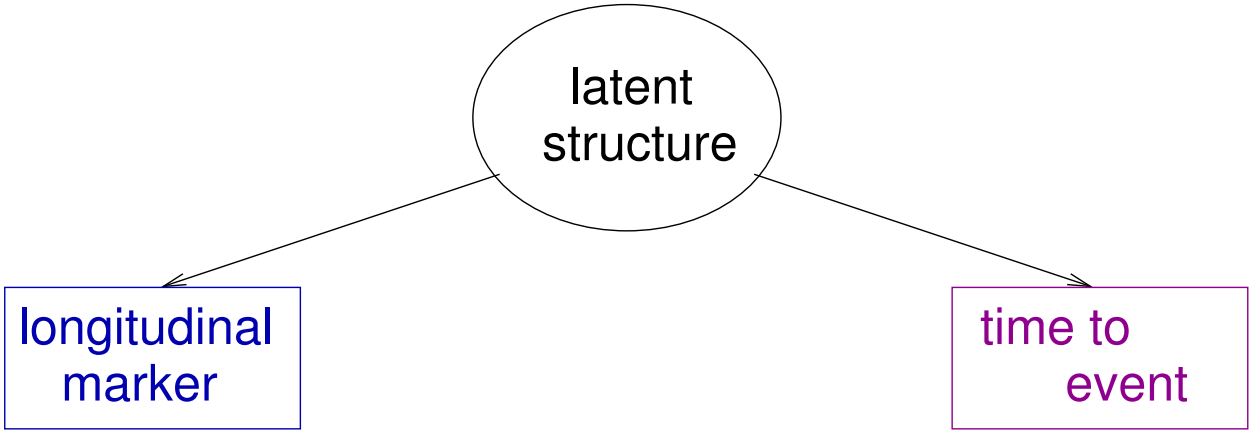
Simultaneous modelling of correlated longitudinal and survival data



Objectives :

- describe the longitudinal process stopped by the event
- predict the risk of event ajusted for the longitudinal process
- explore the association between the two processes

2 main families of joint models



Mixed model
(usually linear)

Survival model
(usually proportional hazards)

Link with the latent structure :

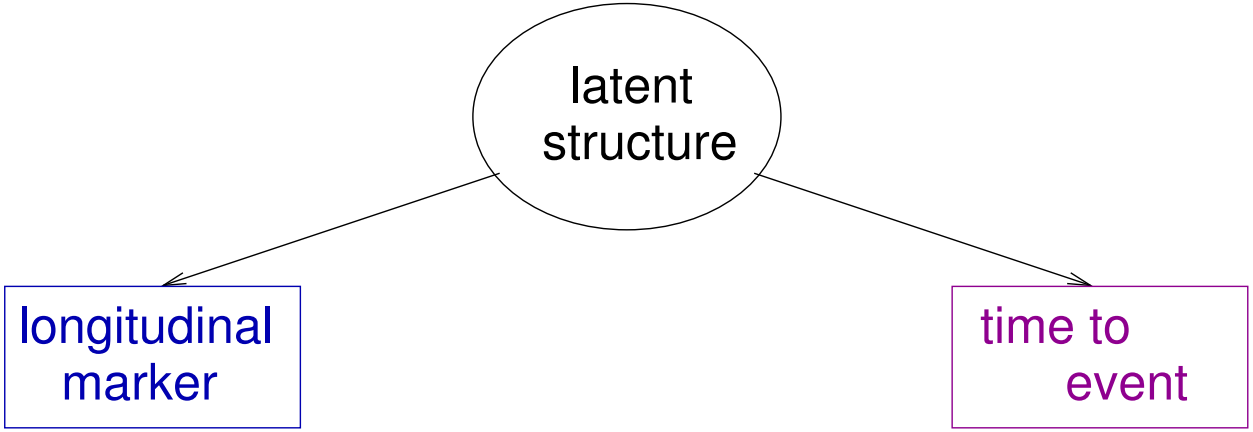
- random effects from the mixed model

  (shared random effect models)

- latent class structure

  (joint latent class models)

Shared random-effect model (SREM) (Rizopoulos, 2012)



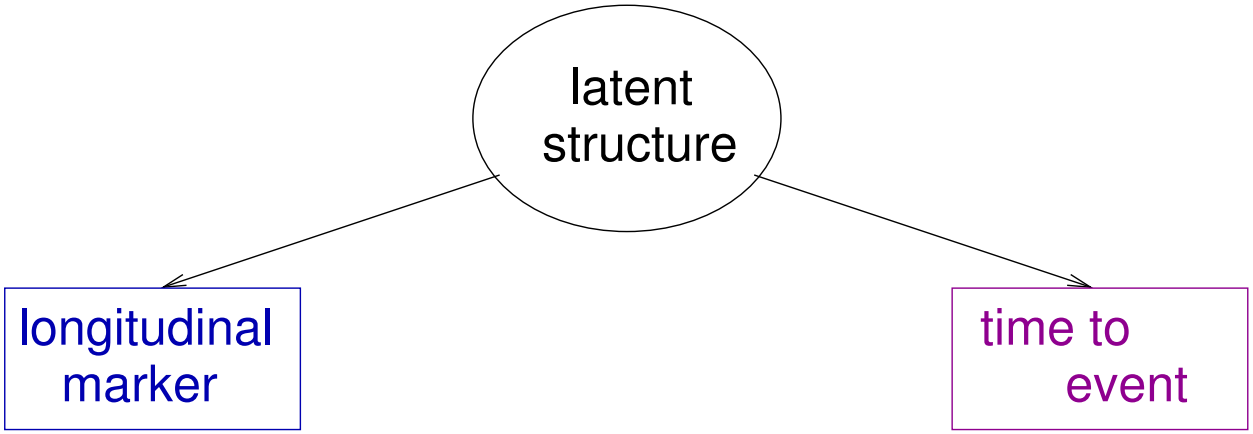
- Shared random-effects distribution : $b_i \sim \mathcal{N}(\mu, B)$
- Linear mixed model for the biomarker trajectory :

  $Y_i(t_{ij}) = Y_i(t_{ij})^* + \epsilon_{ij} = Z_i(t_{ij})^T b_i + X_{Li}(t_{ij})^T \beta + \epsilon_{ij}$ with $\epsilon_{ij} \sim \mathcal{N}(0, \sigma_\epsilon^2)$
- Proportional hazard model including marker trajectory characteristics :

  $\lambda(t \mid b_i) = \lambda_0(t) e^{X_{Si}(t)^T \delta + W_i(b_i, \beta, t)^T \eta}$

$\rightarrow$ 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 :

$$\pi_{ig} = P(c_i = g|X_{pi}) = \frac{e^{\xi_{0g}+X_{Ci}{}^{\top}\xi_{1g}}}{\sum_{l=1}^{G} e^{\xi_{0l}+X_{Ci}/top\xi_{1l}}} \text{ with } \xi_{0G} = 0 \text{ \& } \xi_{\mathbf{1G}} = \mathbf{0}$$

- Class-specific linear mixed model for the biomarker trajectory :

$$Y_i(t_{ij}) \mid_{c_i=g} = Z_i(t_{ij})^T b_{ig} + X_{Li}(t_{ij})^{\top}\beta_g + \epsilon_{ij} \text{ with } b_{ig} \sim \mathcal{N}(\mu_g, B_g),\ \epsilon_{ij} \sim \mathcal{N}(0, \sigma_\epsilon^2)$$

- Class-specific proportional hazard model :

$$\lambda(t \mid c_i = g) = \lambda_{0g}(t)e^{X_{Ti}(t)\delta_g}$$

$\rightarrow$ lcmm 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

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

In chronic diseases

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

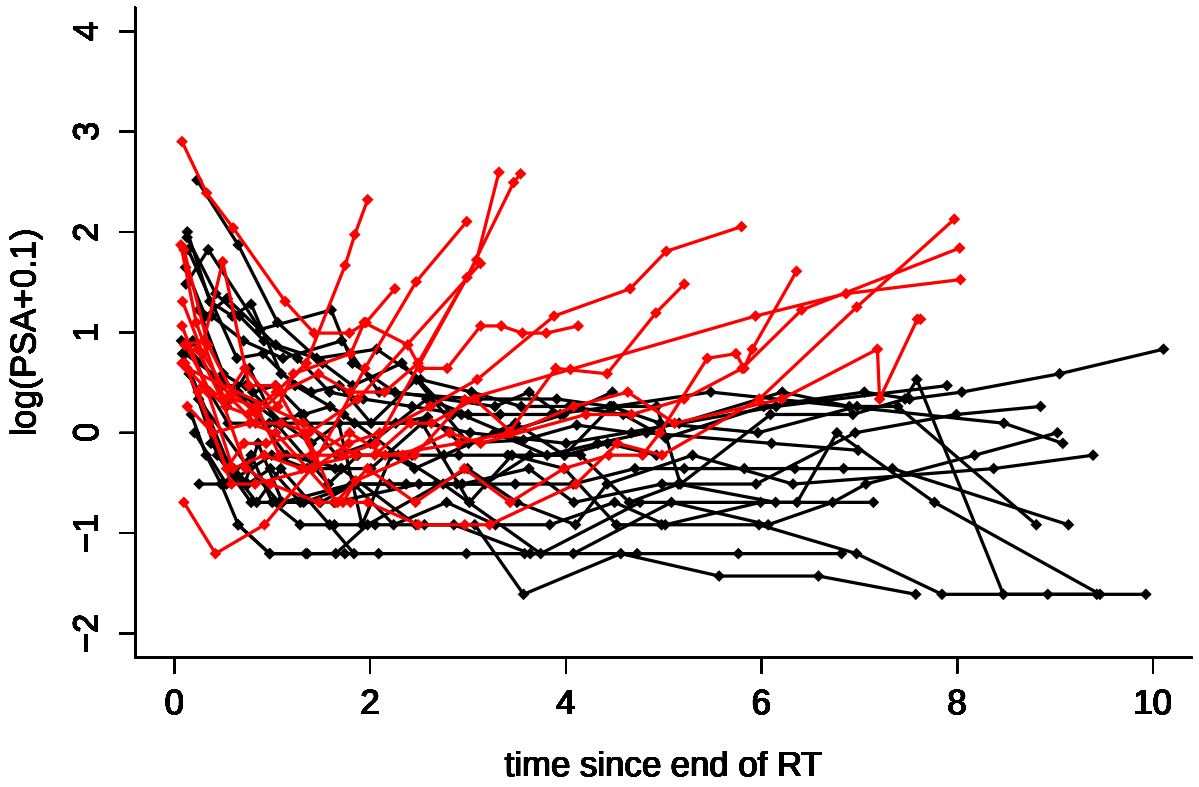
- Survival part :
  - **competing risks**
  - recurrent events
  - multiple events
  - **succession of different events**

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

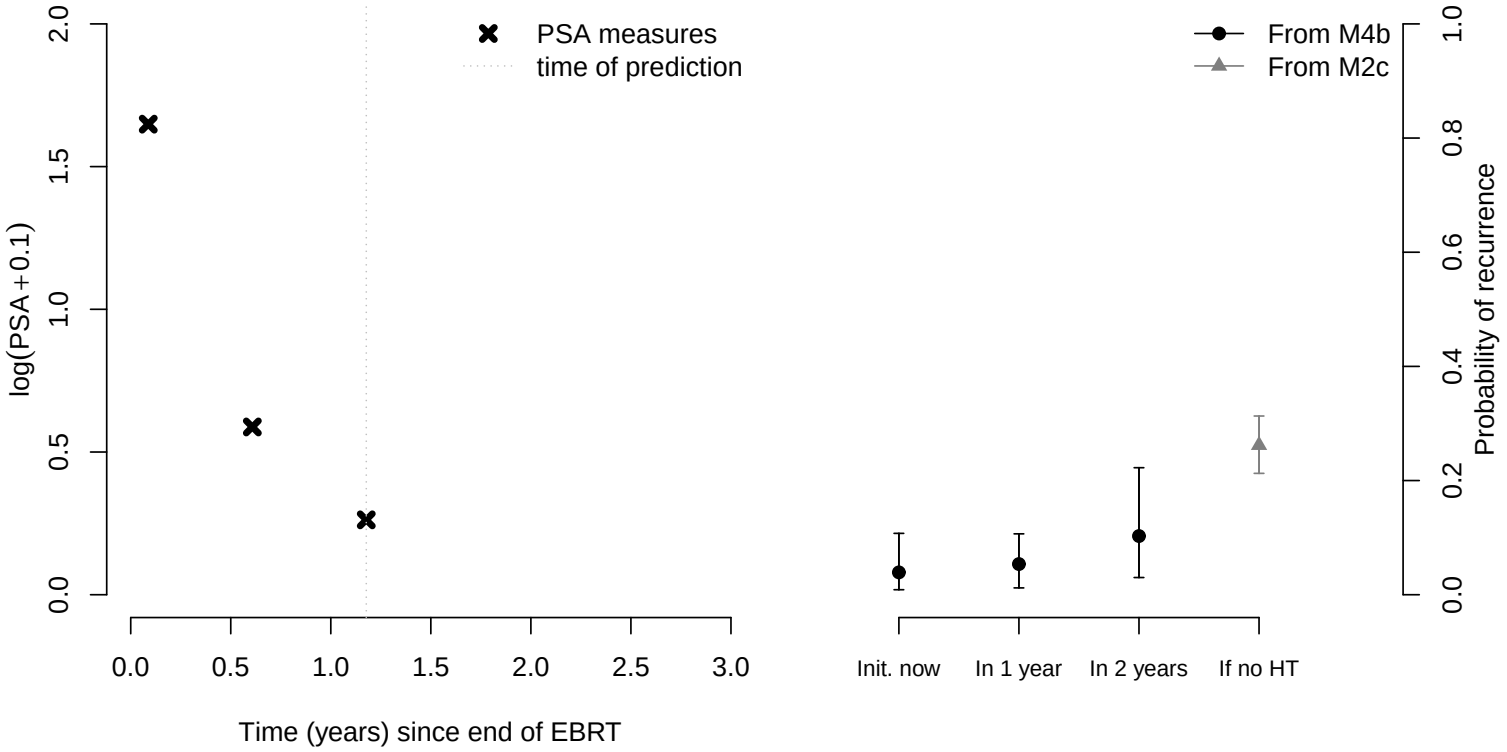


- Interest in predicting the risk of progression
  - multiple types : local recurrence, distant recurrence, death
  - problem of initiation of new treatment : hormonal treatment

Dynamic prediction of clinical recurrence of any type

(Sène et al., SMMR 2014) :

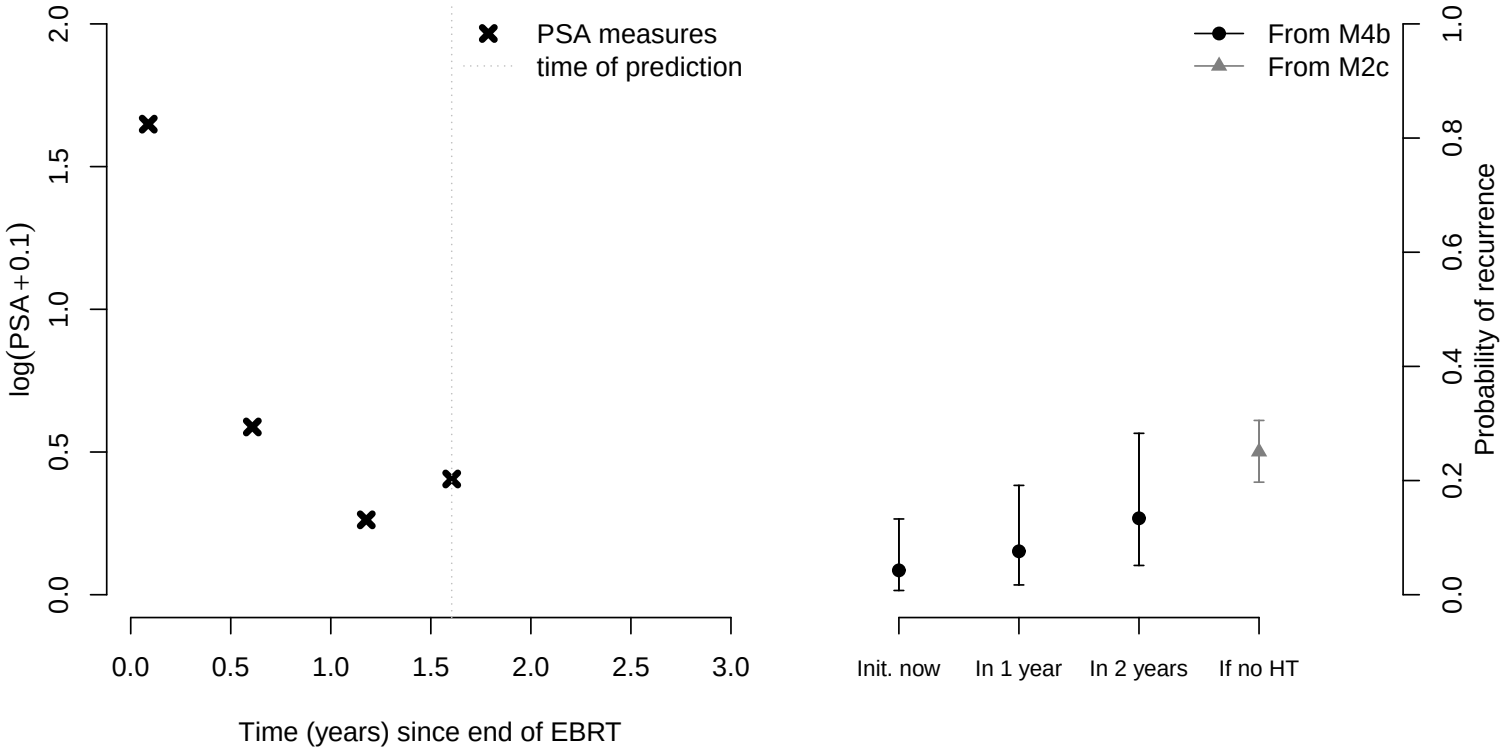
- Individualized probability of clinical recurrence :
  - in the next three years
  - for a man naive of HT
  - according to hypothetical times of initiation of HT (time-dependent covariate)



Dynamic prediction of clinical recurrence of any type

(Sène et al., SMMR 2014) :

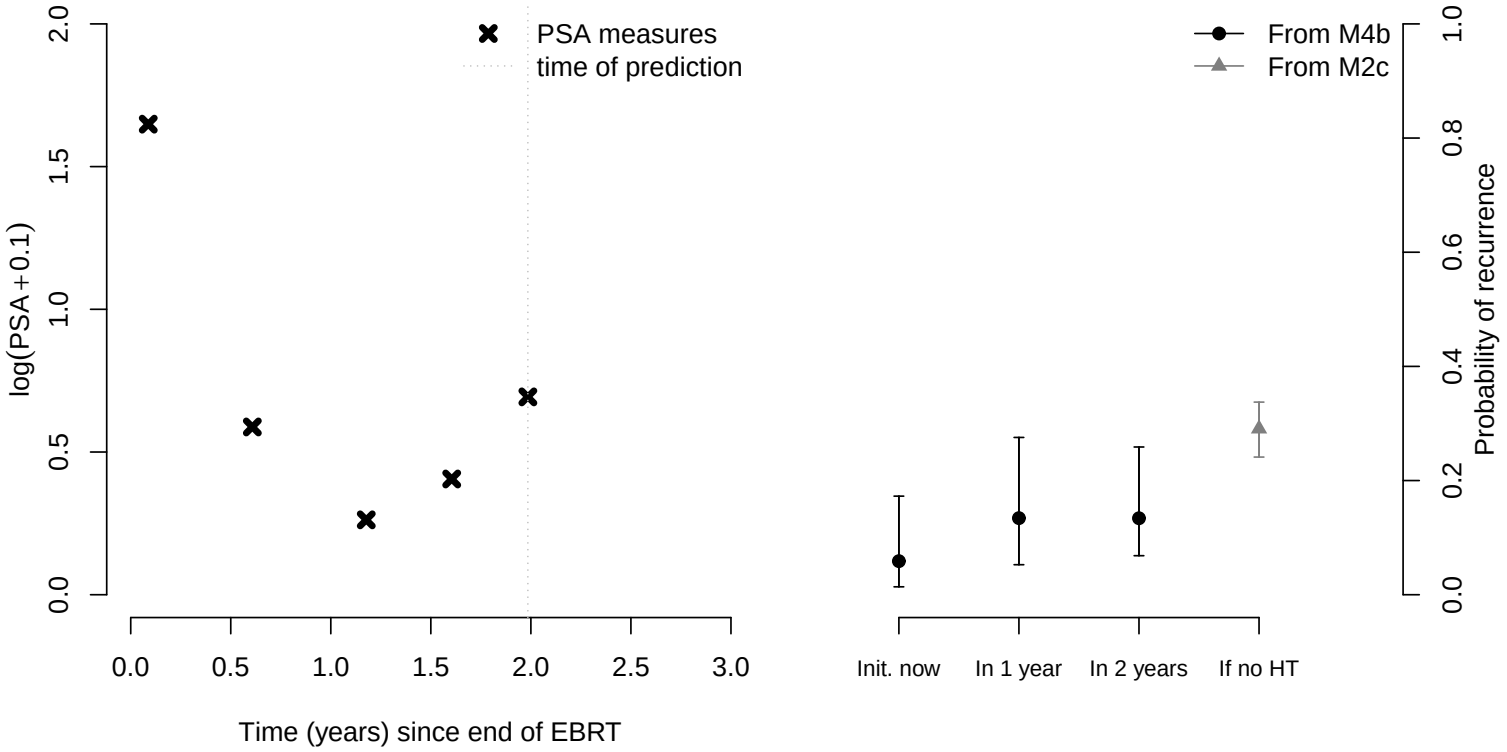
- Individualized probability of clinical recurrence :
  - in the next three years
  - for a man naive of HT
  - according to hypothetical times of initiation of HT (time-dependent covariate)



Dynamic prediction of clinical recurrence of any type

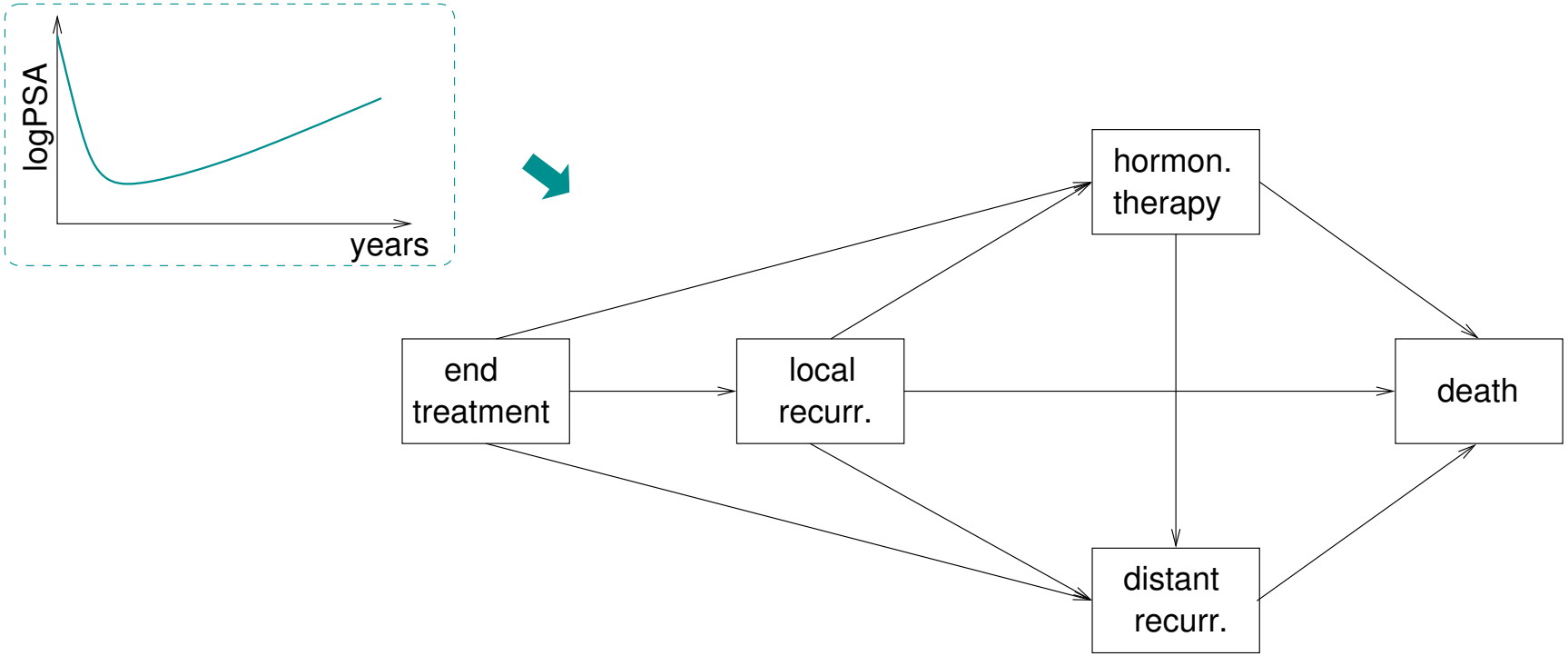
(Sène et al., SMMR 2014) :

- 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)



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 (Ferrer et al., arXiv 2015)

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
  - $T_i = (T_{i1}, \ldots, T_{im_i})^\top$ the m_i observed times with $T_{ir} < T_{i(r+1)}, \forall r \in S$
  - $\delta_i = (\delta_{i1}, \ldots, \delta_{im_i})^\top$ the vector of observed transition indicators
- longitudinal process
  - $Y_i = (Y_{i1}, \ldots, Y_{in_i})^\top$ the n_i measures of the marker collected at times $t_{i1}, \ldots, t_{in_i}$, with $t_{in_i} \leq T_{i1}$

The joint longitudinal and multistate model (Ferrer et al., arXiv 2015) (cont'd)

- Longitudinal part : mixed model

$$\begin{aligned} Y_{ij} &= Y_i^*(t_{ij}) + \epsilon_{ij} \\ &= X_{Li}(t_{ij})^\top \beta + Z_i(t_{ij})^\top b_i + \epsilon_{ij} \end{aligned}$$

  - $b_i \sim \mathcal{N}(0, D)$, $\epsilon_i = (\epsilon_{i1}, \ldots, \epsilon_{in_i})^\top \sim \mathcal{N}(0, \sigma^2 \mathrm{I}_{n_i})$, $b_i \perp\!\!\!\perp \epsilon_i$

- Survival part : multistate model

$$\begin{aligned} \lambda_{hk}^i(t|b_i) &= \lim_{\mathrm{d}t \to 0} \frac{\Pr(E_i(t+\mathrm{d}t) = k | E_i(t) = h; b_i)}{\mathrm{d}t} \\ &= \lambda_{hk,0}(t) \exp(X_{Thk,i}^\top \gamma_{hk} + W_{hk,i}(b_i, t)^\top \eta_{hk}), \quad \text{for } h, k \in S, \end{aligned}$$

  - $\lambda_{hk,0}(t)$ parametric baseline intensity, $X_{Thk,i}$ prognostic factors
  - $W_{hk,i}(b_i, t)$ the dependence structure (Sène et al., J sfds 2014)
    - ★ $W_{hk,i}(b_i, t) = Y_i^*(t)$ $\longrightarrow$ (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)
    - ★ ...

Maximum Likelihood Estimation

- Likelihood function using $Y_i \perp\!\!\!\perp_{b_i} T_i$,

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

with :

- Random effects part : $b_i \sim \mathcal{N}_q(0, D)$
- Longitudinal part : $Y_{i\ |b_i} \sim \mathcal{N}_{n_i}(X_{Li}^\top \beta + Z_i^\top b_i, \sigma^2 \mathrm{I}_{n_i})$
- Multi-state part :

$$f_T(T_i,\delta_i|b_i;\theta) = \prod_{r=0}^{m_i-1} \Big[P^i_{E_i(T_{ir}),E_i(T_{ir})}(T_{ir}, T_{i(r+1)}|b_i) \times \lambda^i_{E_i(T_{ir}),E_i(T_{i(r+1)})}(T_{i(r+1)}|b_i)^{\delta_{i(r+1)}} \Big]$$

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

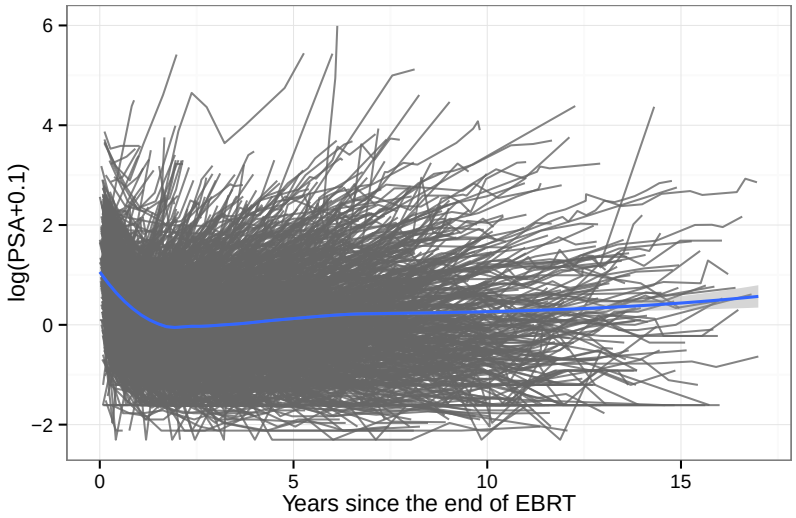
Implementation under R

- 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)

Application

- 2 cohorts of men with localized prostate cancer treated by radiotherapy (N=1474)

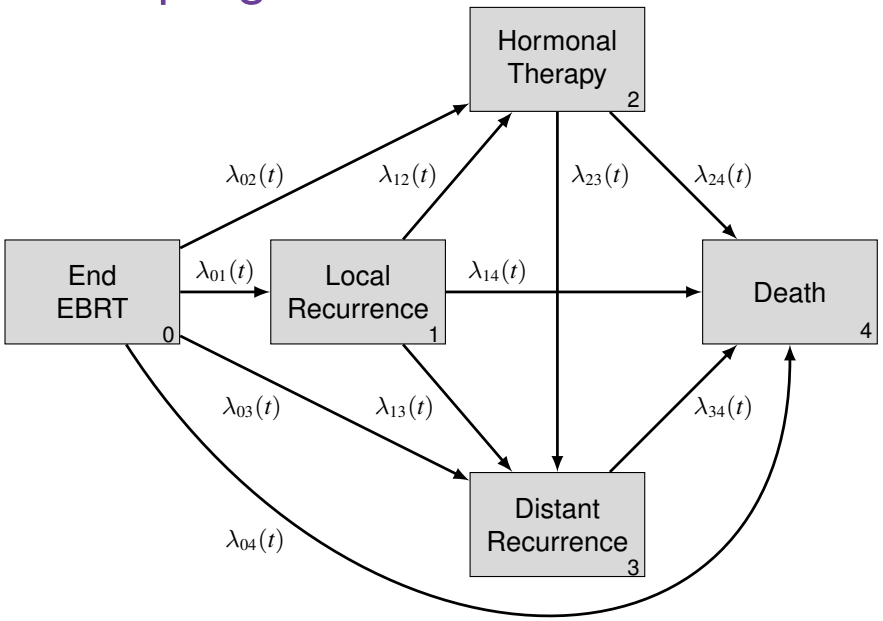
- Repeated measures of PSA



10 (3, 21) measurements per patient

(50th, (5th, 95th) %iles)

- Multi-state representation of the clinical progressions



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

matrix of direct transitions

Specification of the joint model (1/2)

- Longitudinal sub-model specification

$$
\begin{aligned}
Y_{ij} &= Y_i^*(t_{ij}) + \epsilon_{ij} \\
&= \left(\beta_0 + X_{L0i}^\top \beta_{0,\mathrm{cov}} + b_{i0}\right) + \\
&\quad \left(\beta_1 + X_{L1i}^\top \beta_{1,\mathrm{cov}} + b_{i1}\right) \times f_1(t_{ij}) + \\
&\quad \left(\beta_2 + X_{L2i}^\top \beta_{2,\mathrm{cov}} + b_{i2}\right) \times f_2(t_{ij}) + \epsilon_{ij}
\end{aligned}
$$

  - $f_1(t) = (1+t)^{-1.2} - 1$ and $f_2(t) = t$
  - $b_i = (b_{i0}, b_{i1}, b_{i2})^\top \sim \mathcal{N}(0, D)$, 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.

Specification of the joint model (2/2)

- Multi-state sub-model specification

$$\lambda_{hk}^{i}(t|b_i) = \lambda_{hk,0}(t)\exp\left(X_{T,hk,i}^{\top}\gamma_{hk} + \begin{pmatrix} Y_i^*(t) \\ \partial Y_i^*(t)/\partial t \end{pmatrix}^{\top} \begin{pmatrix} \eta_{hk,\texttt{level}} \\ \eta_{hk,\texttt{slope}} \end{pmatrix}\right)$$

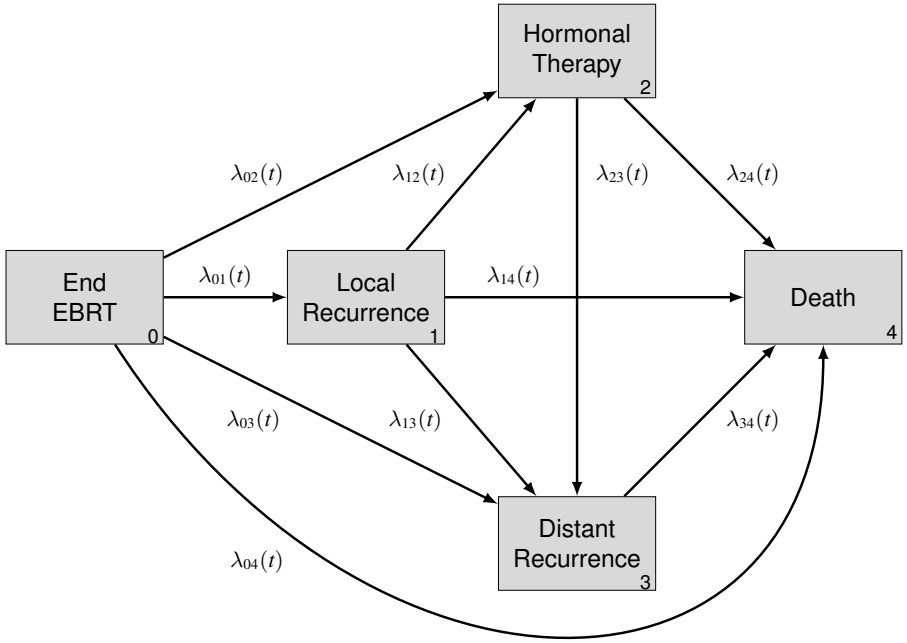
  - Log-baseline intensities approximated by B-splines
  - Proportionality assumptions between several baseline intensities
  - Backward stepwise procedure to select the prognostic factors
  - Dependence function chosen by Wald tests

Results

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

	Value	StdErr	p-value
Level : 01	0.37	0.09	< 0.001
Level : 02	0.51	0.07	< 0.001
Level : 03	0.45	0.11	< 0.001
Level : 04	-0.17	0.05	0.001
Level : 12	-0.16	0.10	0.110
Level : 13	-0.41	0.20	0.042
Level : 14	0.10	0.14	0.487
Level : 23	-0.15	0.09	0.120
Level : 24	0.00	0.05	0.412
Level : 34	0.04	0.08	0.609
	.....		
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



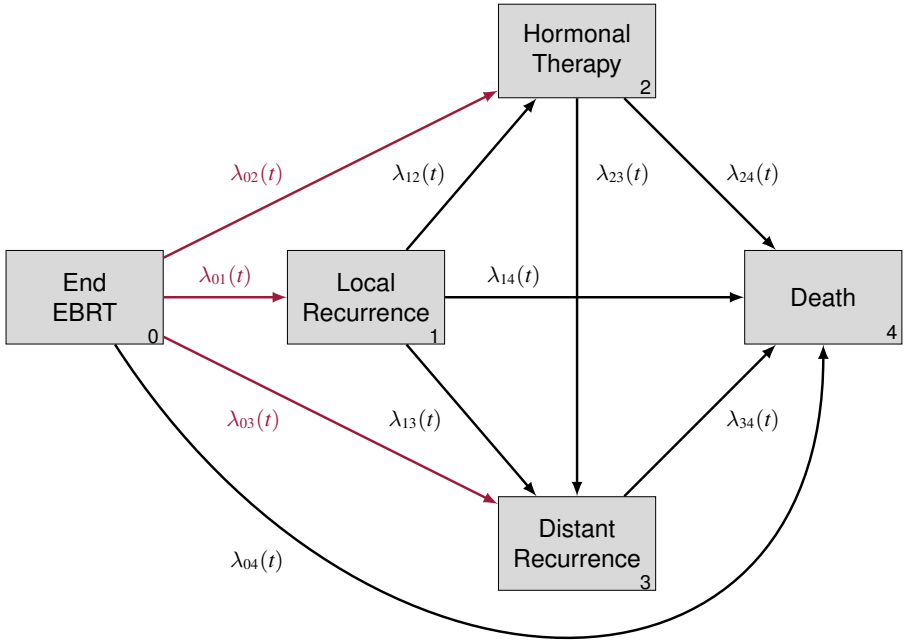
Prognostic factors : advanced initial stage not always associated with intensities of transitions between health states after adjustment on PSA

Results

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

	Value	StdErr	p-value
Level : 01	0.37	0.09	< 0.001
Level : 02	0.51	0.07	< 0.001
Level : 03	0.45	0.11	< 0.001
Level : 04	-0.17	0.05	0.001
Level : 12	-0.16	0.10	0.110
Level : 13	-0.41	0.20	0.042
Level : 14	0.10	0.14	0.487
Level : 23	-0.15	0.09	0.120
Level : 24	0.00	0.05	0.412
Level : 34	0.04	0.08	0.609
	.....		
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



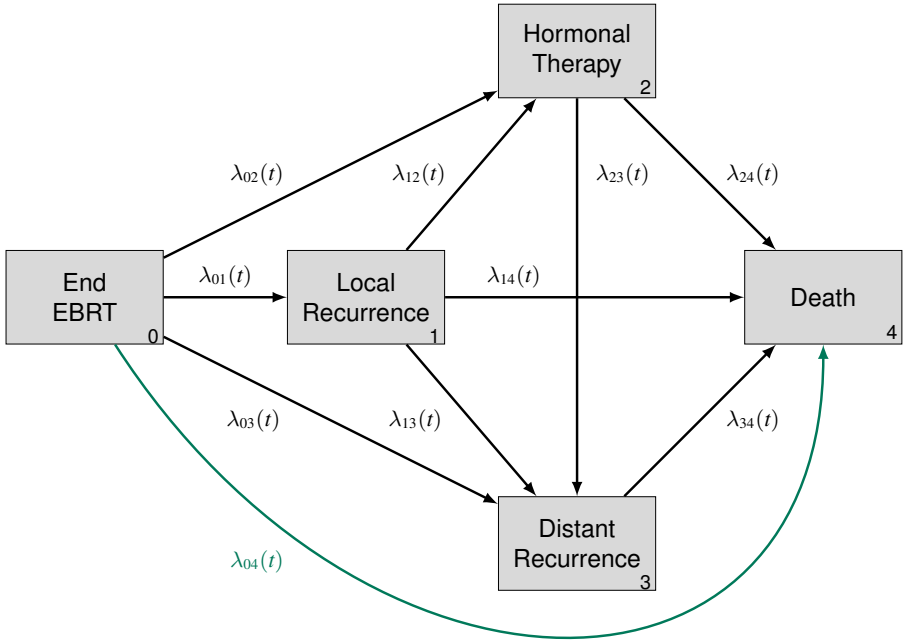
Prognostic factors : advanced initial stage not always associated with intensities of transitions between health states after adjustment on PSA

Results

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

	Value	StdErr	p-value
Level : 01	0.37	0.09	< 0.001
Level : 02	0.51	0.07	< 0.001
Level : 03	0.45	0.11	< 0.001
Level : 04	-0.17	0.05	0.001
Level : 12	-0.16	0.10	0.110
Level : 13	-0.41	0.20	0.042
Level : 14	0.10	0.14	0.487
Level : 23	-0.15	0.09	0.120
Level : 24	0.00	0.05	0.412
Level : 34	0.04	0.08	0.609
	.....		
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



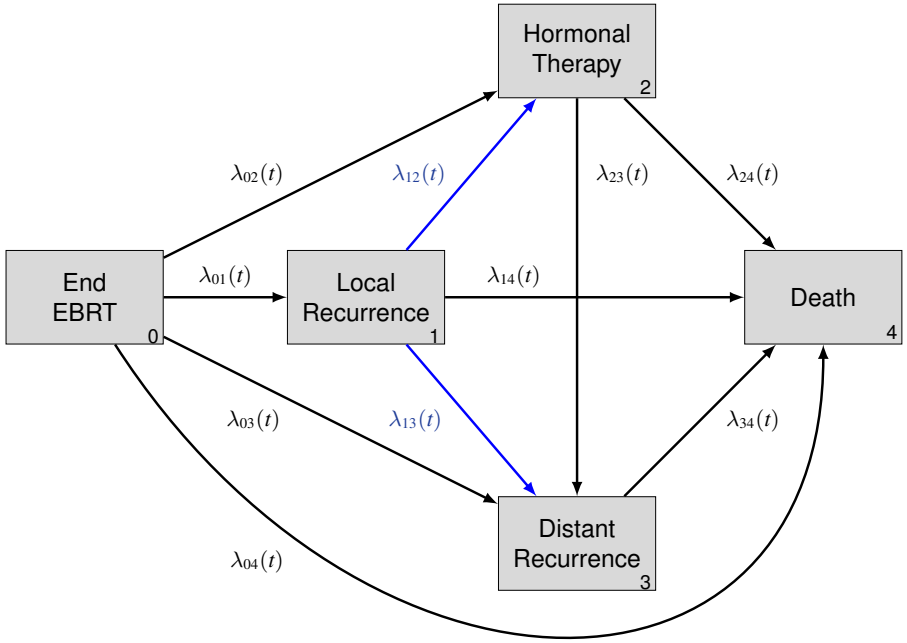
Prognostic factors : advanced initial stage not always associated with intensities of transitions between health states after adjustment on PSA

Results

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

	Value	StdErr	p-value
Level : 01	0.37	0.09	< 0.001
Level : 02	0.51	0.07	< 0.001
Level : 03	0.45	0.11	< 0.001
Level : 04	−0.17	0.05	0.001
Level : 12	−0.16	0.10	0.110
Level : 13	−0.41	0.20	0.042
Level : 14	0.10	0.14	0.487
Level : 23	−0.15	0.09	0.120
Level : 24	0.00	0.05	0.412
Level : 34	0.04	0.08	0.609
	.....		
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

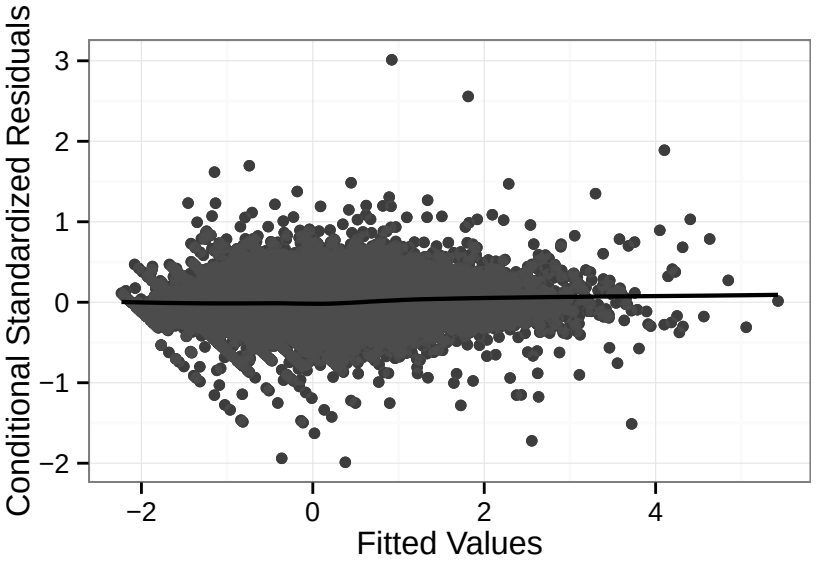
Multi-state process

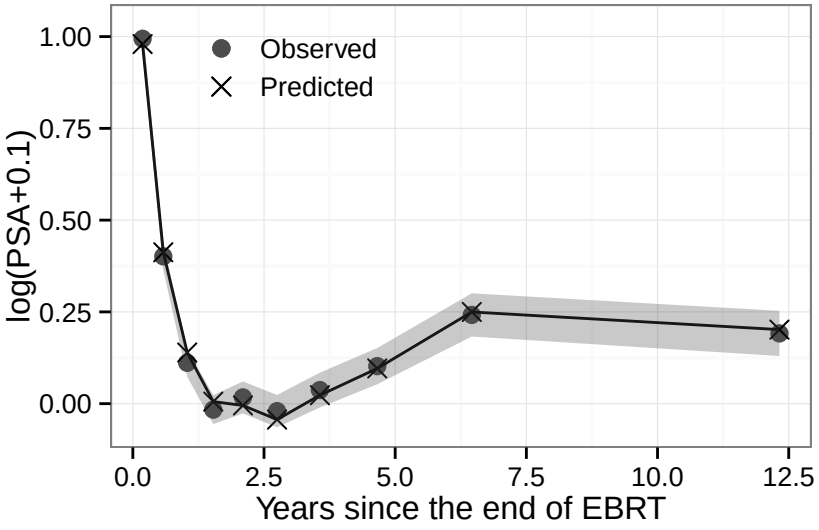


Prognostic factors : advanced initial stage not always associated with intensities of transitions between health states after adjustment on PSA

Diagnostics for the parametric assumptions

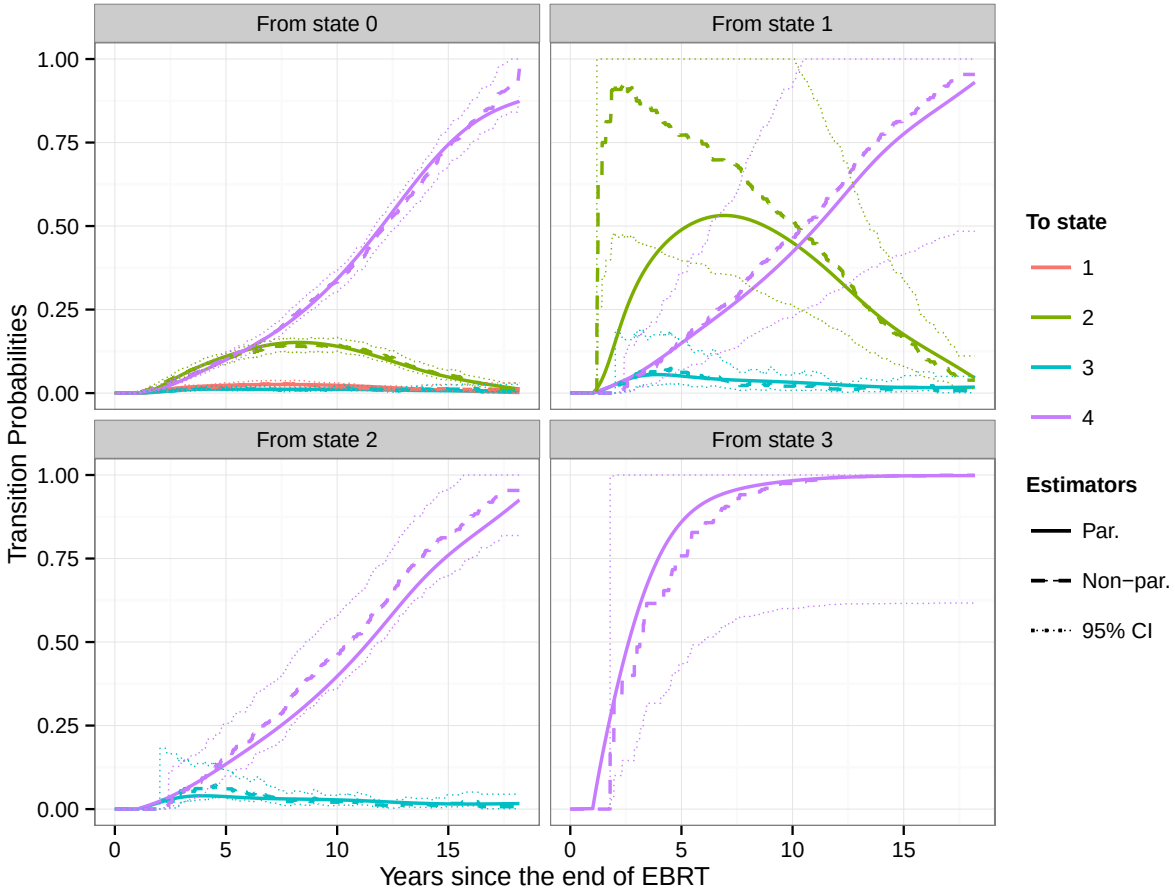
- Goodness-of-fit plots for the longitudinal process
  - Conditional standardized residuals versus fitted values
  - Observed and predicted values of the biomarker





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



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

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

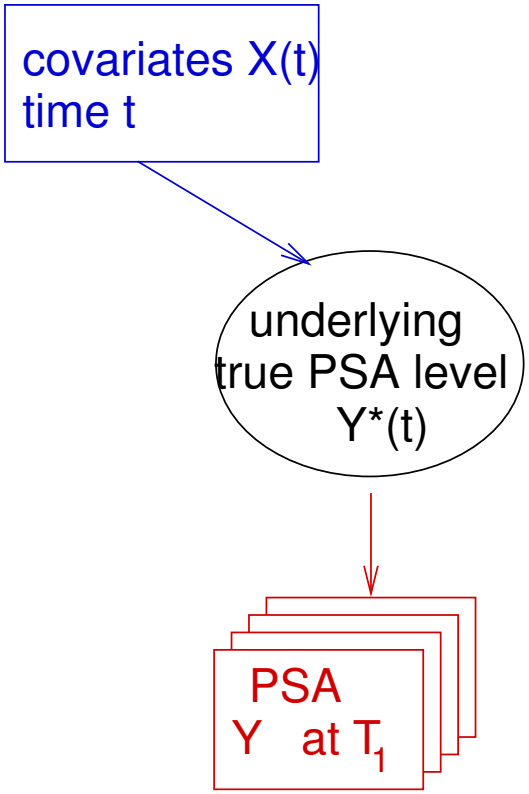
→ 1st complexity : elderly pathology

- delayed entry
- dementia in competition with death
- diagnosis at pre-established visit times

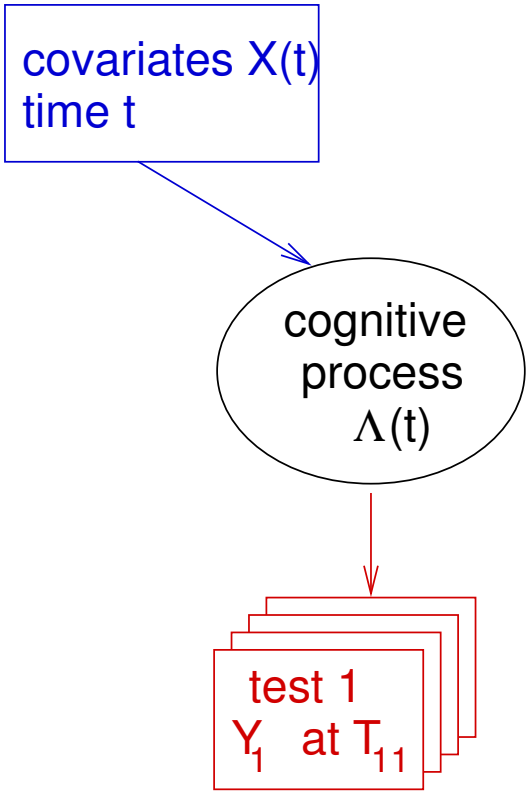
→ 2nd complexity : cognition is not directly observed

- cognitive process (trait) defined in continuous time
- repeated psychometric tests measured in discrete times
  - ★ multiple cognitive functions (langage, memory, attention,...)
  - ★ noisy measures of overall cognition
  - ★ limited statistical properties

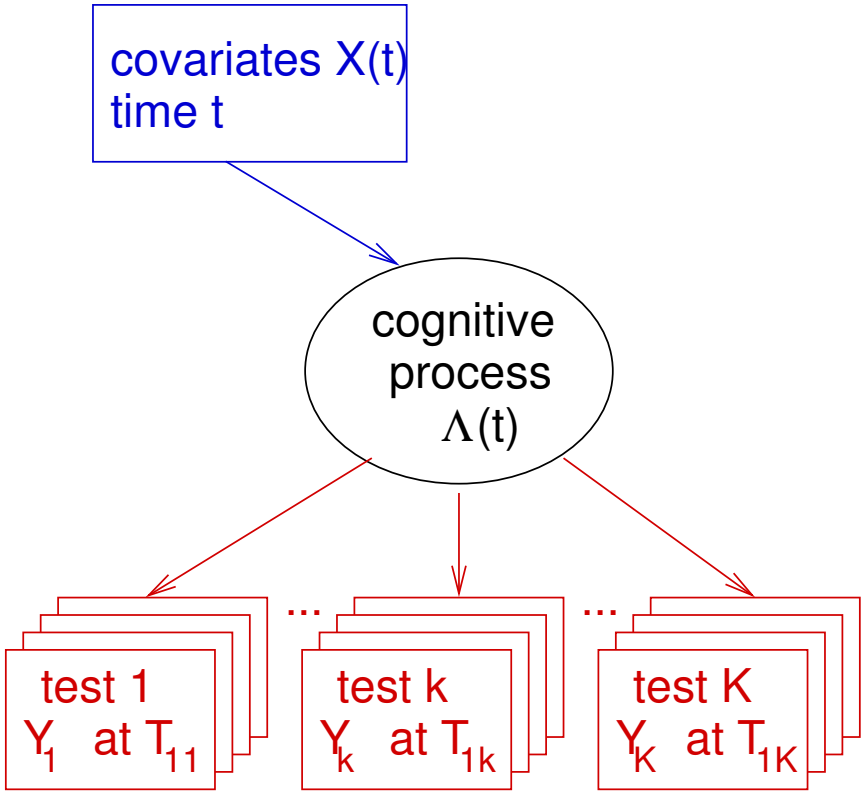
Latent process mixed model



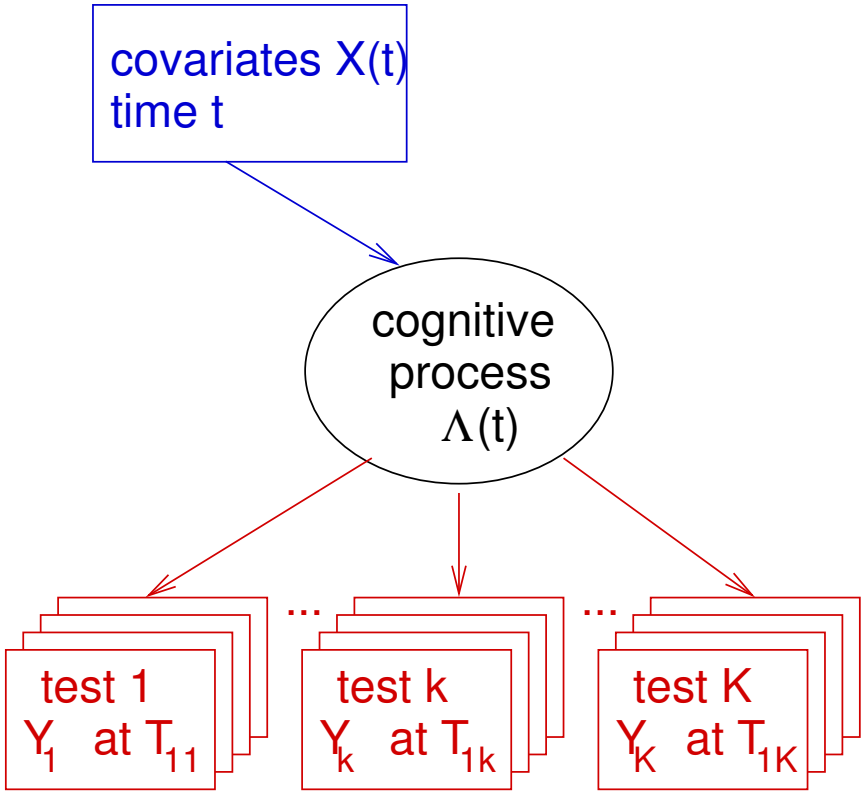
Latent process mixed model



Latent process mixed model



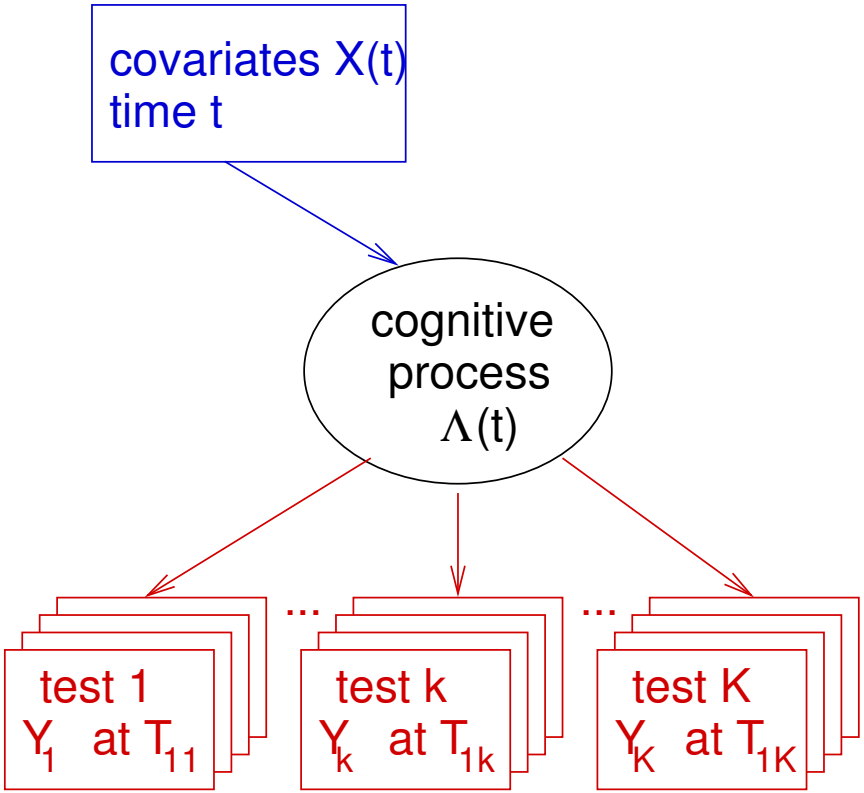
Latent process mixed model



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

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

Latent process mixed model



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

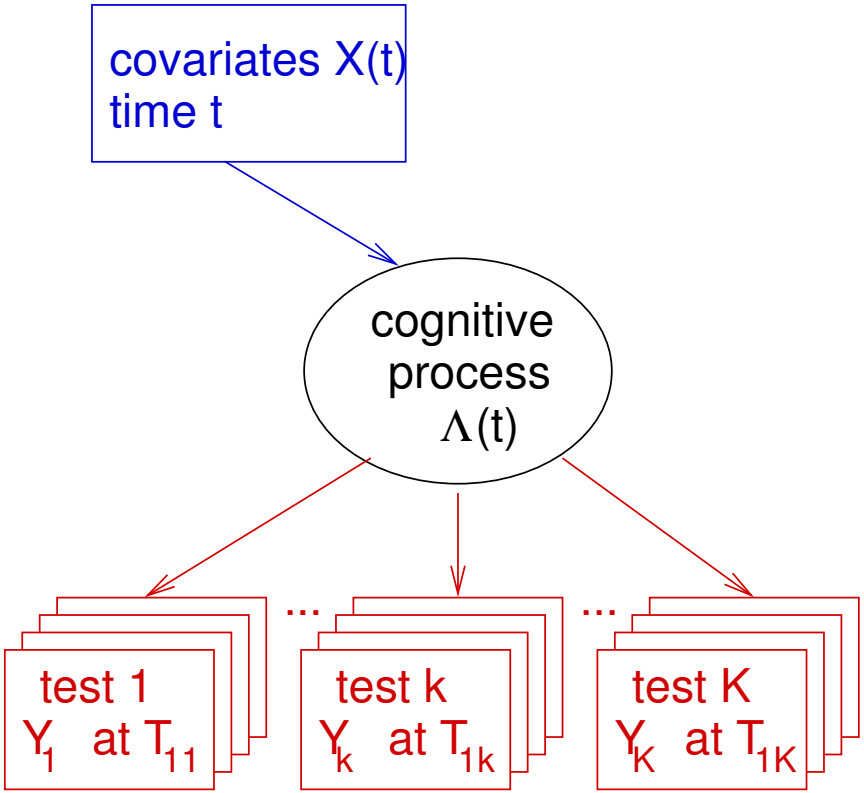
- $b_i \sim MVN(\mu, B)$
- $w_i(t)$ autocorrelated process
- $b_{i0} \sim N(0,1)$ for identifiability

$\leftarrow \quad Y_{kij} = \zeta_{1k} + \zeta_{2k}\tilde{Y}_{kij}$

$\tilde{Y}_{kij} = \Lambda_i(t_{ijk}) + \epsilon_{kij}$

- $\epsilon_{kij} \sim \mathcal{N}(0, \sigma^2_{\epsilon_k})$

Latent process mixed model



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

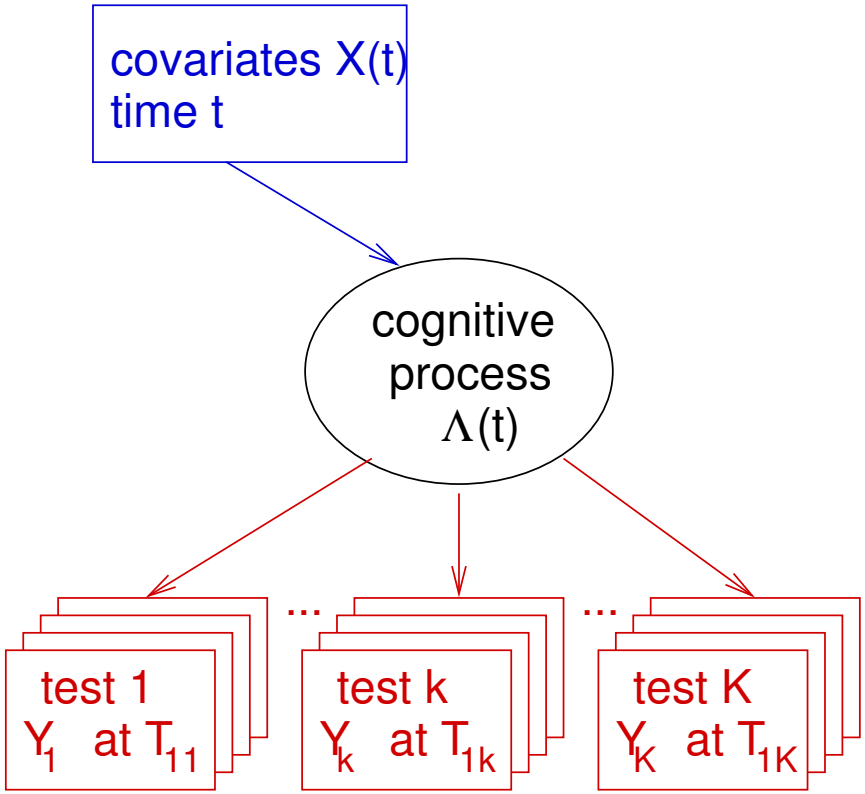
- $b_i \sim MVN(\mu, B)$
- $w_i(t)$ autocorrelated process
- $b_{i0} \sim N(0,1)$ for identifiability

$\leftarrow \quad Y_{kij} = \zeta_{1k} + \zeta_{2k}\tilde{Y}_{kij}$

$\tilde{Y}_{kij} = \Lambda_i(t_{ijk}) + \mathbf{X_{L2i}}(t)^\top \gamma_k + \alpha_{ki} + \epsilon_{kij}$

- $\alpha_{ki} \sim \mathcal{N}(0, \sigma^2_{\alpha_k})$
- $\epsilon_{kij} \sim \mathcal{N}(0, \sigma^2_{\epsilon_k})$

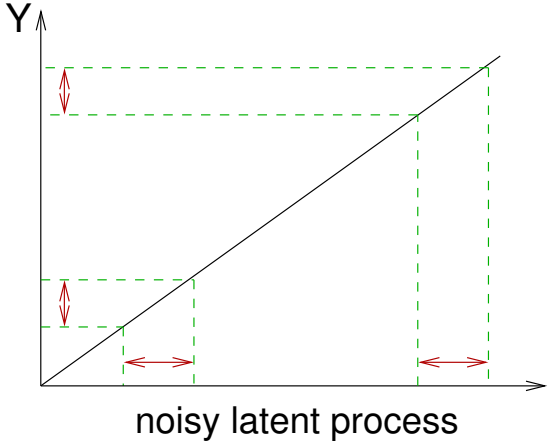
Latent process mixed model



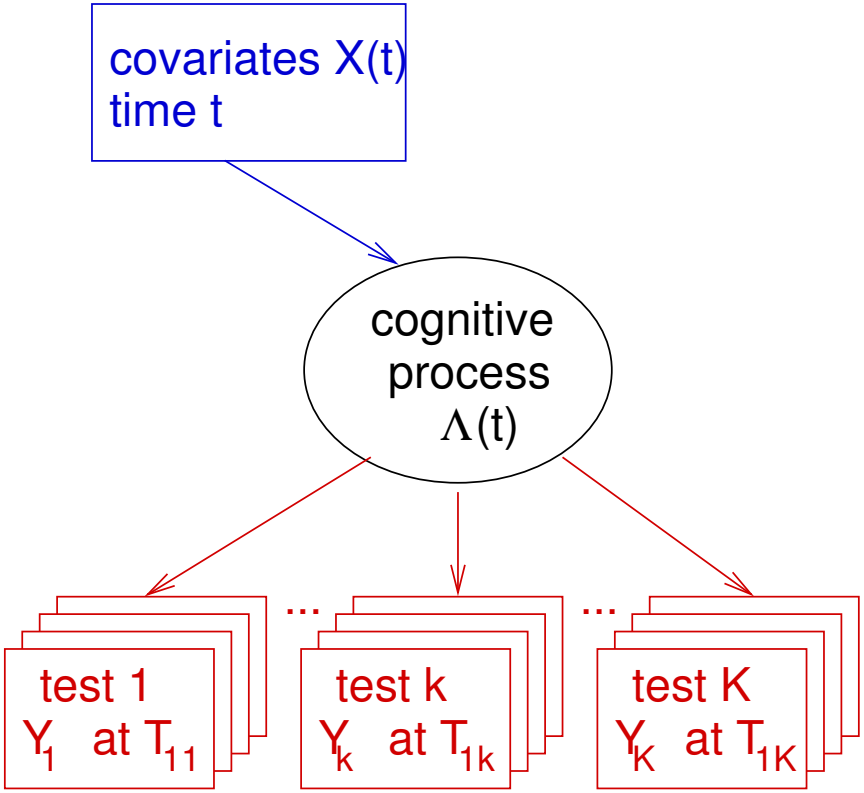
$\leftarrow$ $\Lambda_i(t) = \mathbf{X_{L1i}}(t)^\top \boldsymbol{\beta} + \mathbf{Z_i}(t)^\top \mathbf{b_i} + w_i(t)$

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

$\leftarrow$ linear link function



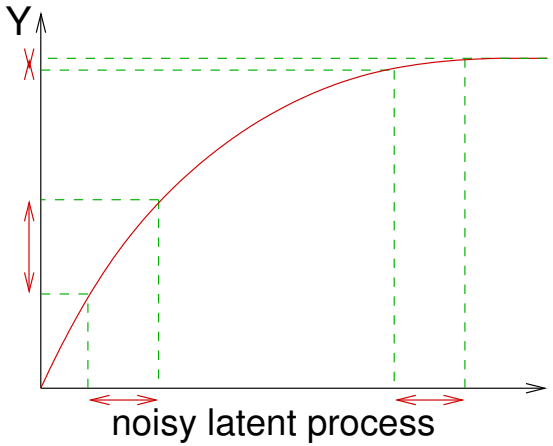
Latent process mixed model involving nonlinear link functions (Proust-Lima et al., 2015)



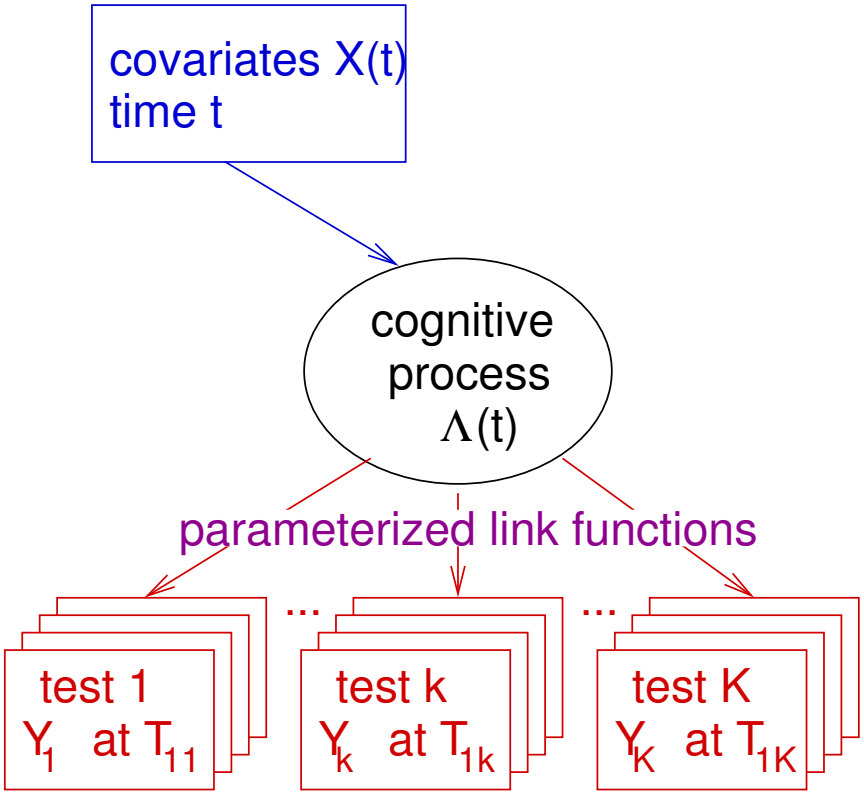
$\leftarrow$ $\Lambda_i(t) = \mathbf{X_{L1i}}(t)^\top \boldsymbol{\beta} + \mathbf{Z_i}(t)^\top \mathbf{b_i} + w_i(t)$

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

$\leftarrow$ nonlinear link function



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



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

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

$\leftarrow \quad Y_{kij} = \boldsymbol{H_k}(\ \tilde{Y}_{kij}\ ; \zeta_k)$

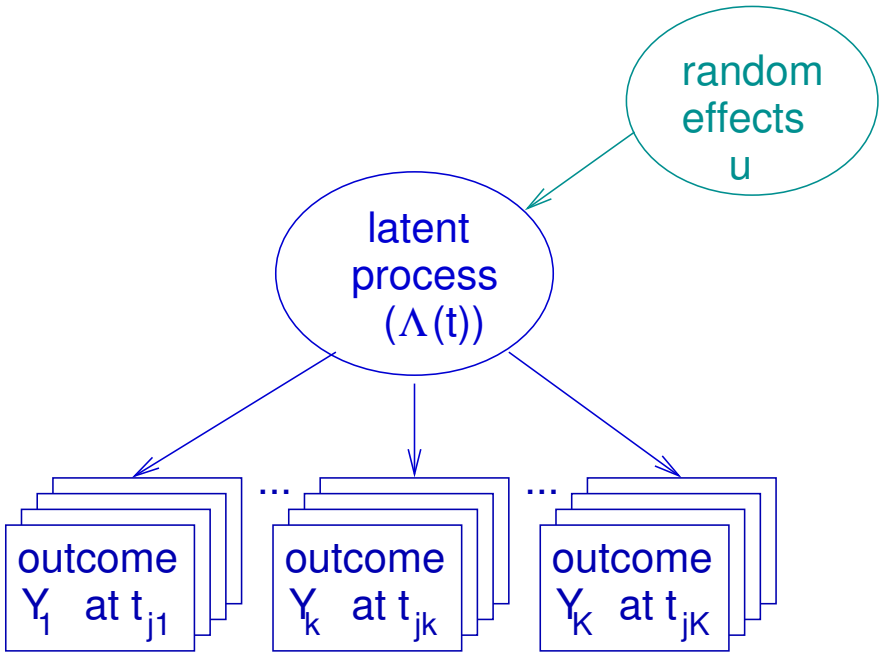
$\tilde{Y}_{kij} = \Lambda_i(t_{ijk}) + \mathbf{X_{L2i}}(t)^\top \gamma_k + \alpha_{ki} + \epsilon_{kij}$

- $\alpha_{ki} \sim \mathcal{N}(0, \sigma^2_{\alpha_k})$
- $\epsilon_{kij} \sim \mathcal{N}(0, \sigma^2_{\epsilon_k})$

H_k = flexible parameterized transformation for outcome k

$\rightarrow$ *linear, standardised Beta CDF, quadratic I-splines, thresholds, ...*

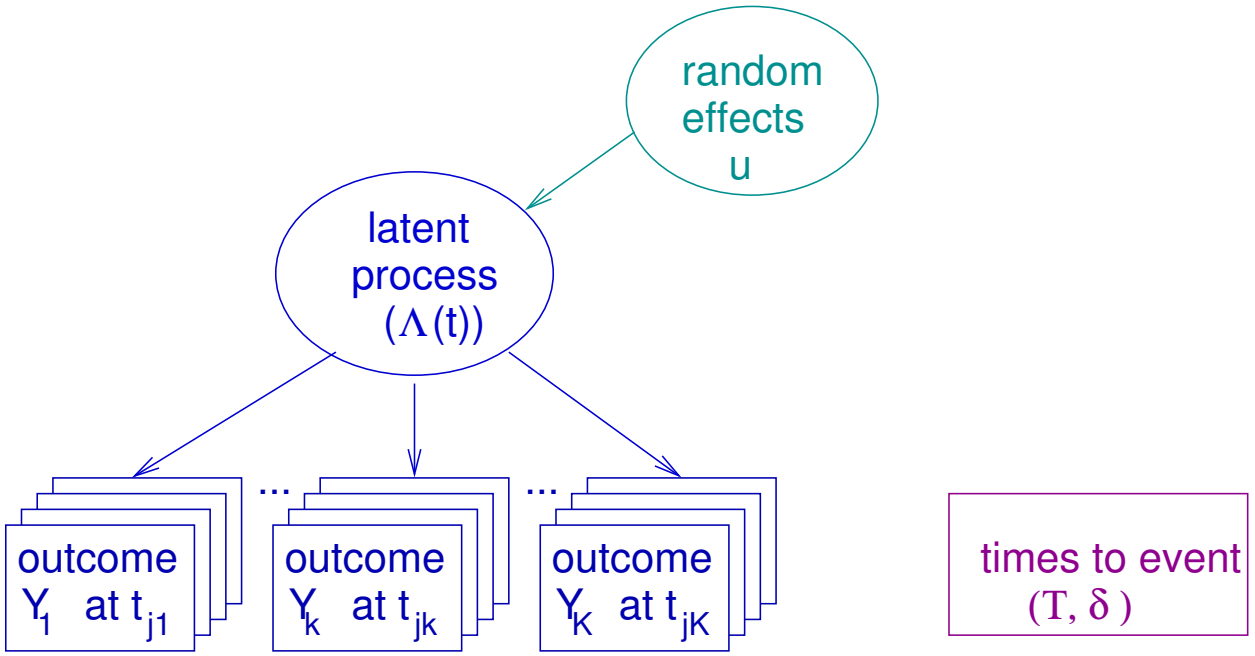
Joint model for multivariate cognitive measures, dementia and death



cognitive measures :

- latent process mixed model

Joint model for multivariate cognitive measures, dementia and death



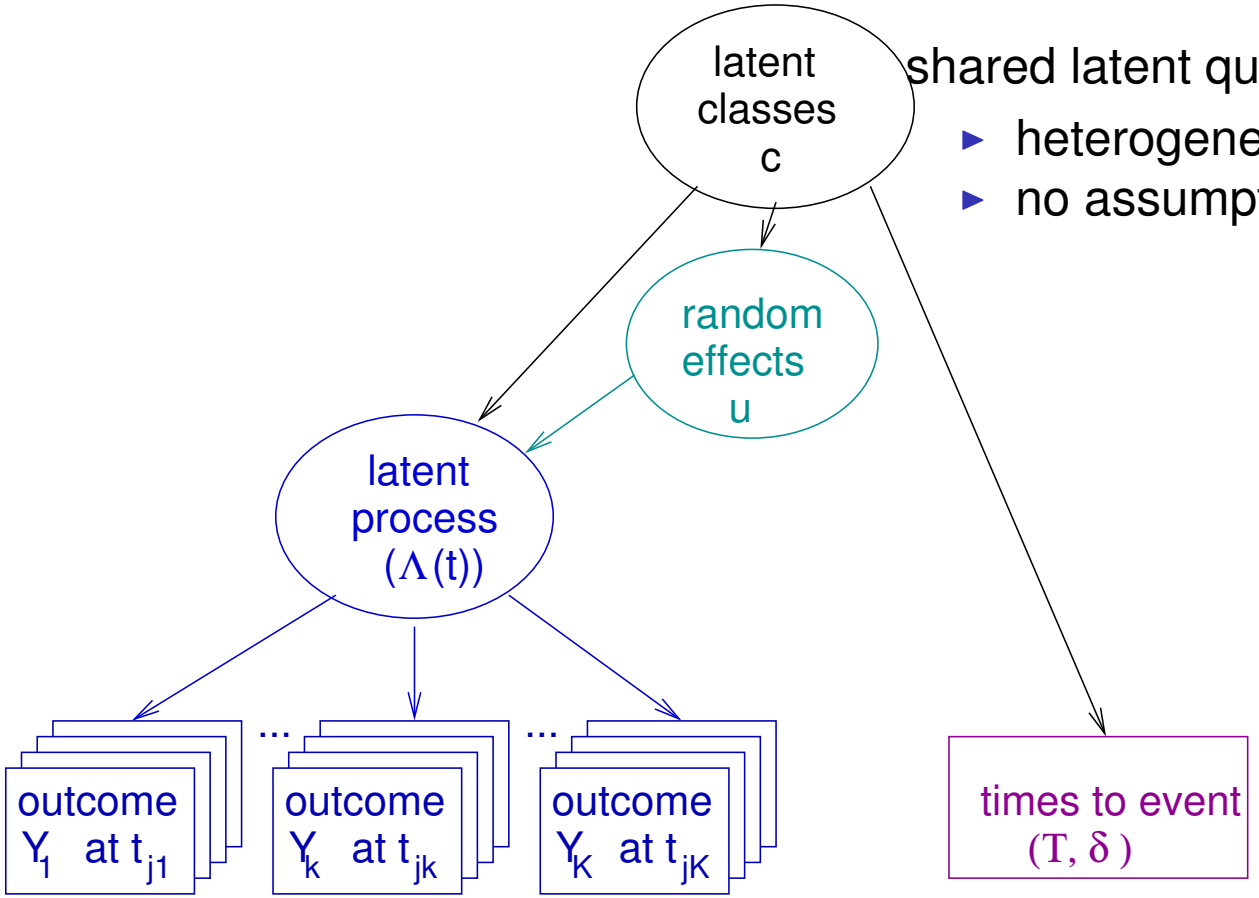
cognitive measures :

- latent process mixed model

dementia and death ?

- option 1 : first event in competing setting
- option 2 : multistate model

Joint model for multivariate cognitive measures, dementia and death



shared latent quantity = latent classes

- heterogeneous trajectories
- no assumption on the association

cognitive measures :

- latent process mixed model

dementia and death ?

- option 1 : first event in competing setting
- option 2 : multistate model

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

- Times to events = Time to P competing events
  - $T_i = \min($censoring $\tilde{T}_i$, and cause-spec. times $T^*_{i1}, ..., T^*_{iP})$,
  - $\delta_i = 0$ for censored, $\delta_i = p$ otherwise

- class-specific cause-specific proportional hazard models

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

  - λ_{0p} parametric (splines, Gompertz, Weibull,...)

Maximum Likelihood Estimation

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

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

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 :
      = multivariate normal $\times$ Jacobian of $(H_k)_{k=1,\ldots,K}$
    - ★ by numerical integration otherwise ...

  - $f_T(T_i, \delta_i|X_{Ti}, c_i = g; \theta)$ 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}) : $l^{T_0}(\theta) = \log\left(\dfrac{L(\theta)}{\prod_{i=1}^{N} S_i(T_{0i}; \theta)}\right)$

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

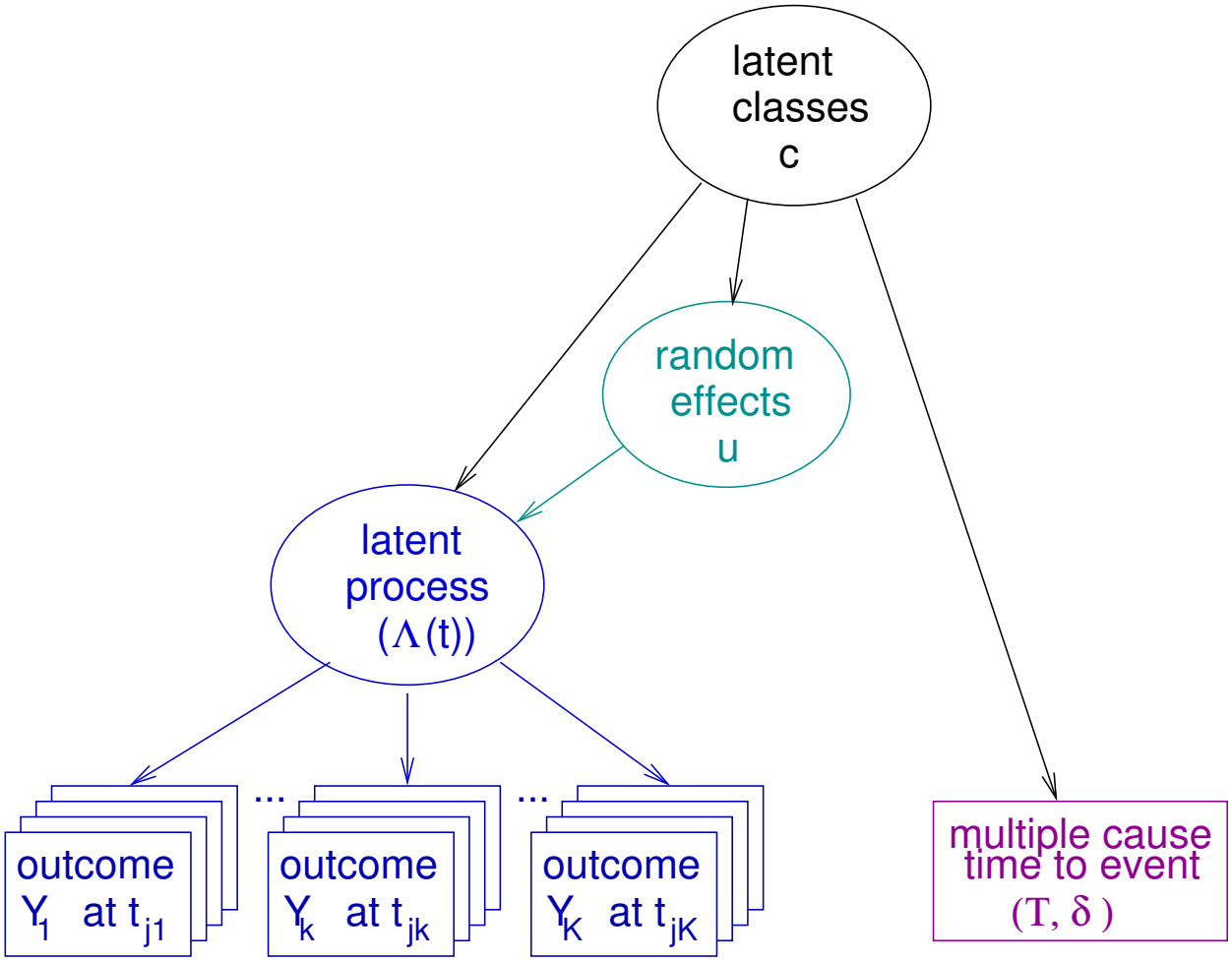
- Posterior selection of the optimal number G of latent classes
  - Information measures : AIC, **BIC**
  - Score Test for conditional independence assumption :
    $\rightarrow$ longitudinal and survival parts are independent conditionally on the latent classes

- 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

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

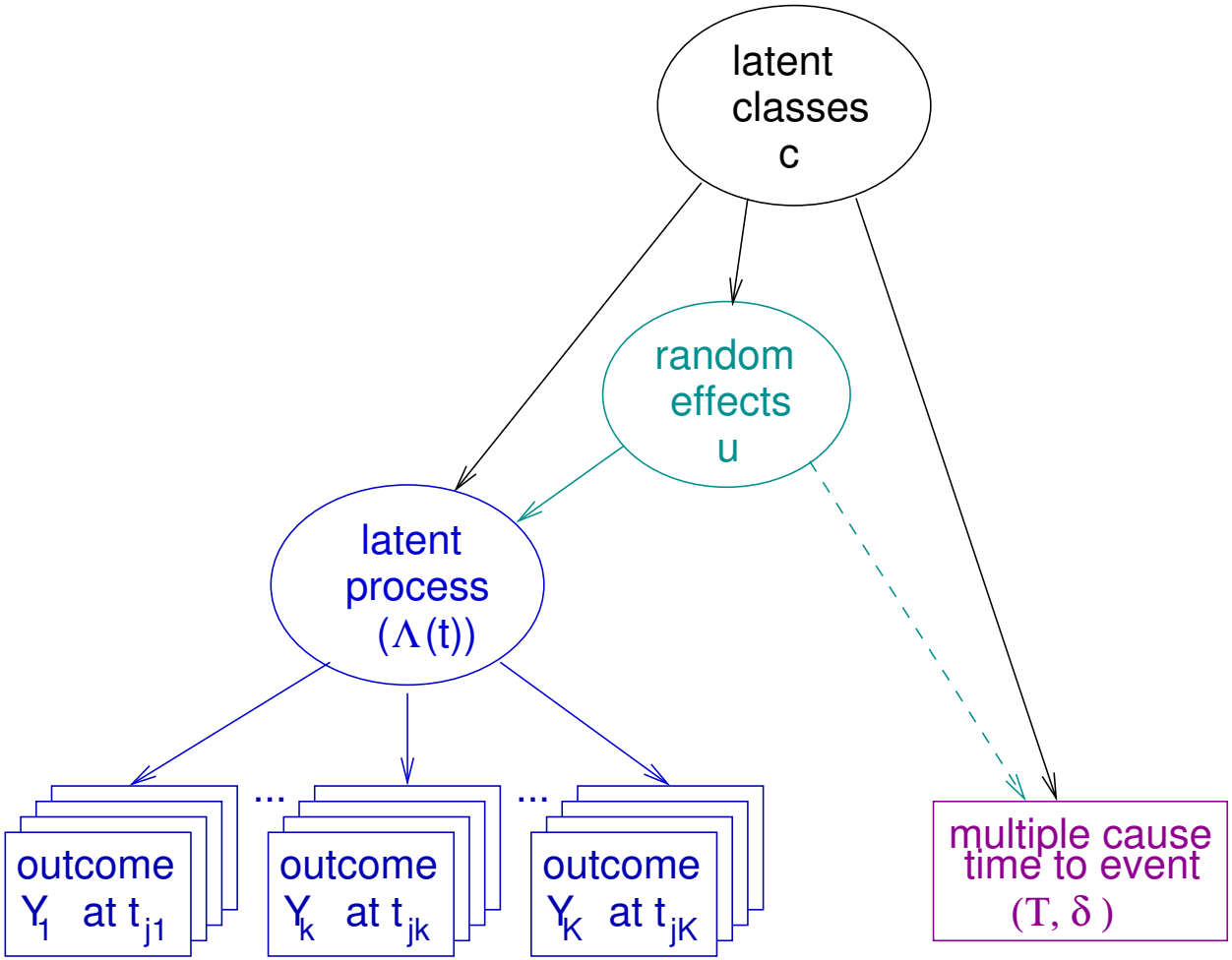


Conditional independence assumption : alternative

Class-specific cause-specific proportional hazard model

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

→ Score test for $\mathcal{H}_0 : \kappa_p = 0$ and $\mathcal{H}_0 : \kappa = (\kappa_1^\top, ..., \kappa_P^\top)^\top = 0$



Application : Semantic memory decline in the elderly

- 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

$\rightarrow$ 588 subjects

Dynamics of semantic memory

- 2 longitudinal measures :
  - Isaacs Set Test (`IST15`) (discrete quantitative in {0-40})
  - Wechsler similarities test (`WST`) (ordinal in {0-10})
- Trajectory according to age (natural history)
  - age at entry (`ageT0`), sex (`sex`), education (`EL`), apoE4 (`E4`)
- In each latent class g :

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

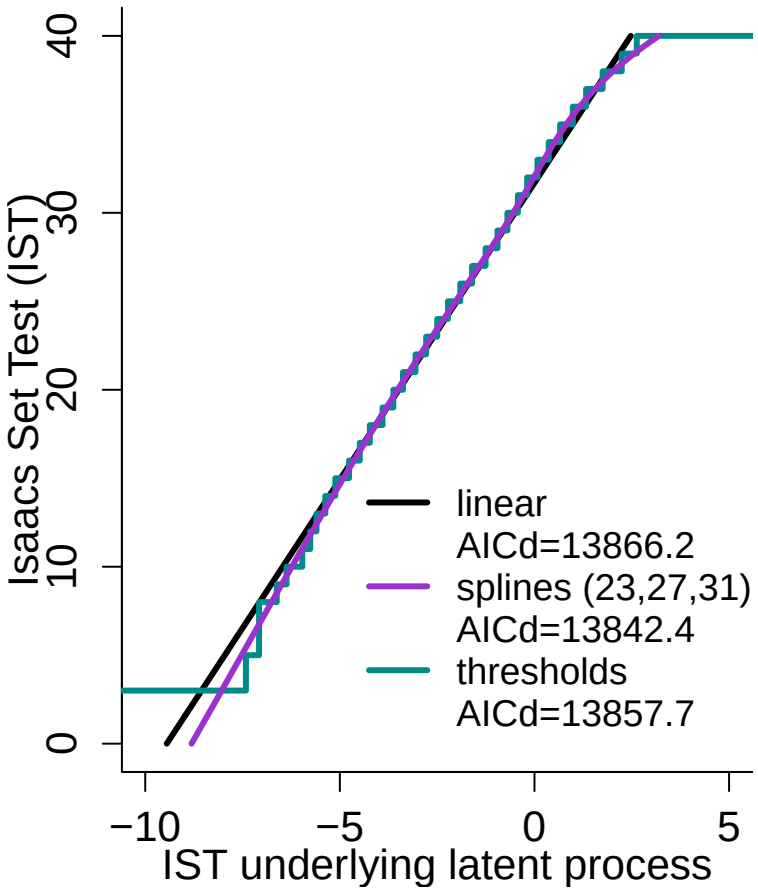
$$\begin{aligned}\texttt{IST15}_{ij} &= \boldsymbol{H_1}(\ \Lambda(\texttt{age}_{1ij}) + \alpha_{1i} + \epsilon_{1ij}\ ;\boldsymbol{\eta_1}) \\ \texttt{WST}_{ij} &= \boldsymbol{H_2}(\ \Lambda(\texttt{age}_{2ij}) + \alpha_{2i} + \epsilon_{2ij}\ ;\boldsymbol{\eta_2})\end{aligned}$$

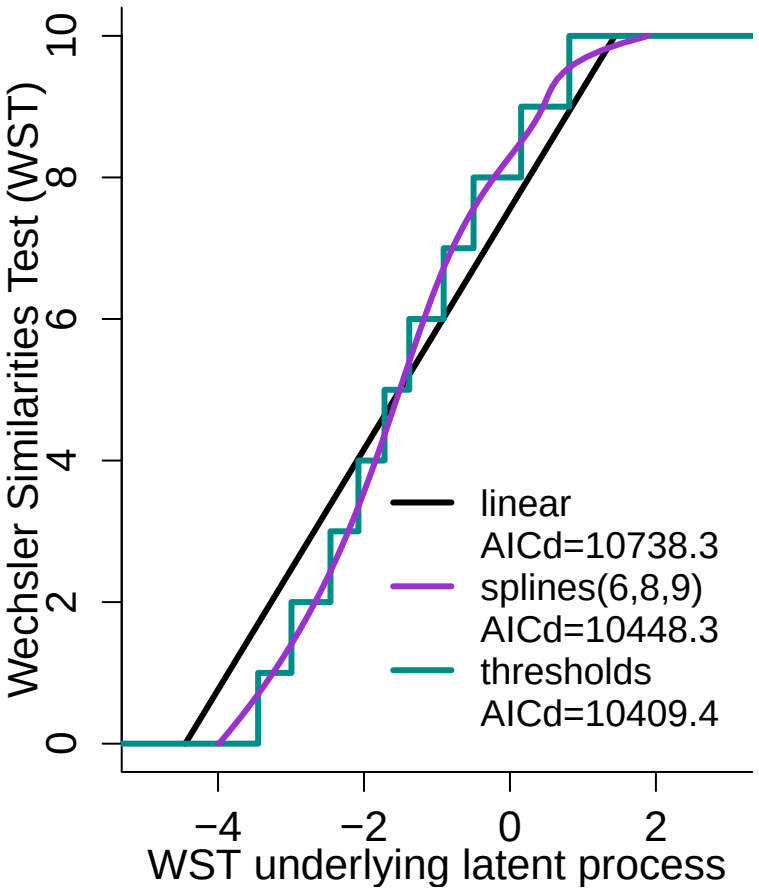
with $b_{ig} \sim \mathcal{N}(\mu_g, B)$, $w_i \sim$ Brownian motion
$\alpha_{2i} \sim \mathcal{N}(0, \sigma_\alpha^2)$, $\epsilon_{ki} \sim \mathcal{N}(0, {\sigma_k}^2)$, $k = 1, 2$

Preliminary analysis : link functions H_1 and H_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 :

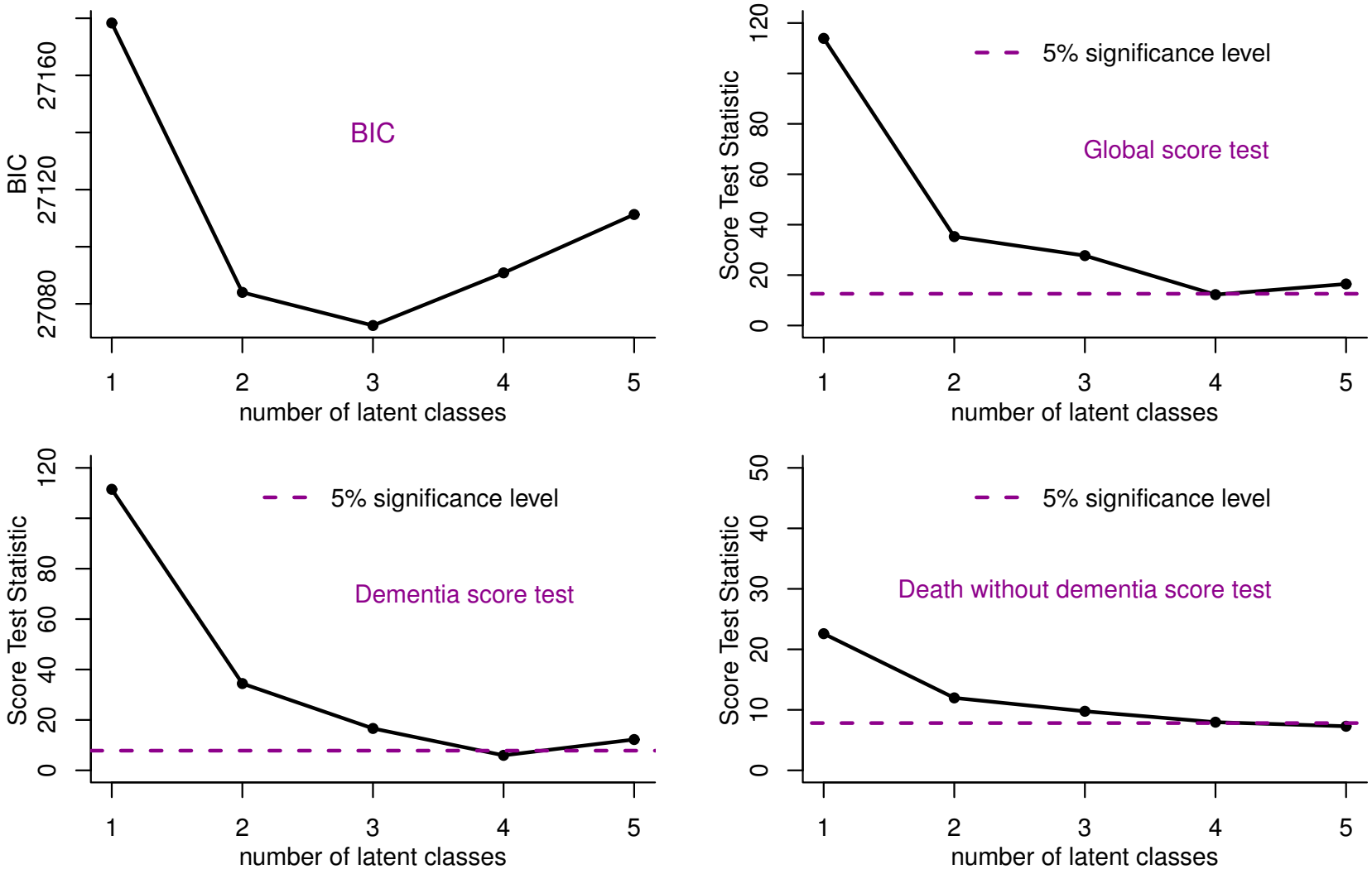
$$\lambda_p(t)|_{c_i=g} = \lambda_{0pg}(t;)e^{\zeta_{1p}\texttt{sex}+\zeta_{2p}\texttt{EL}+\zeta_{3p}\texttt{E4}}, \qquad p = 1, 2$$

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

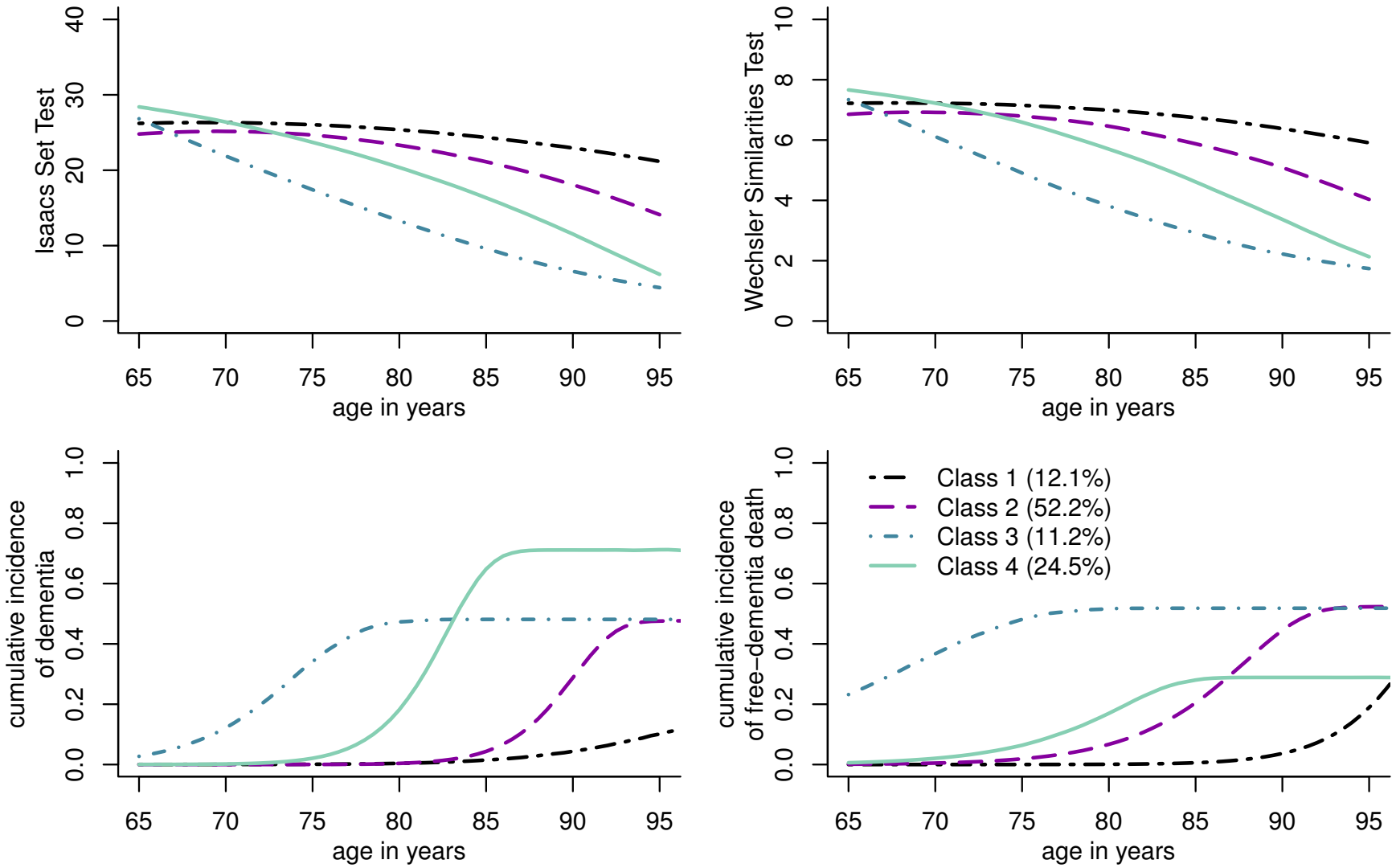
In separated analyses, Weibull hazards =

good balance between estimation difficulty and goodness-of-fit

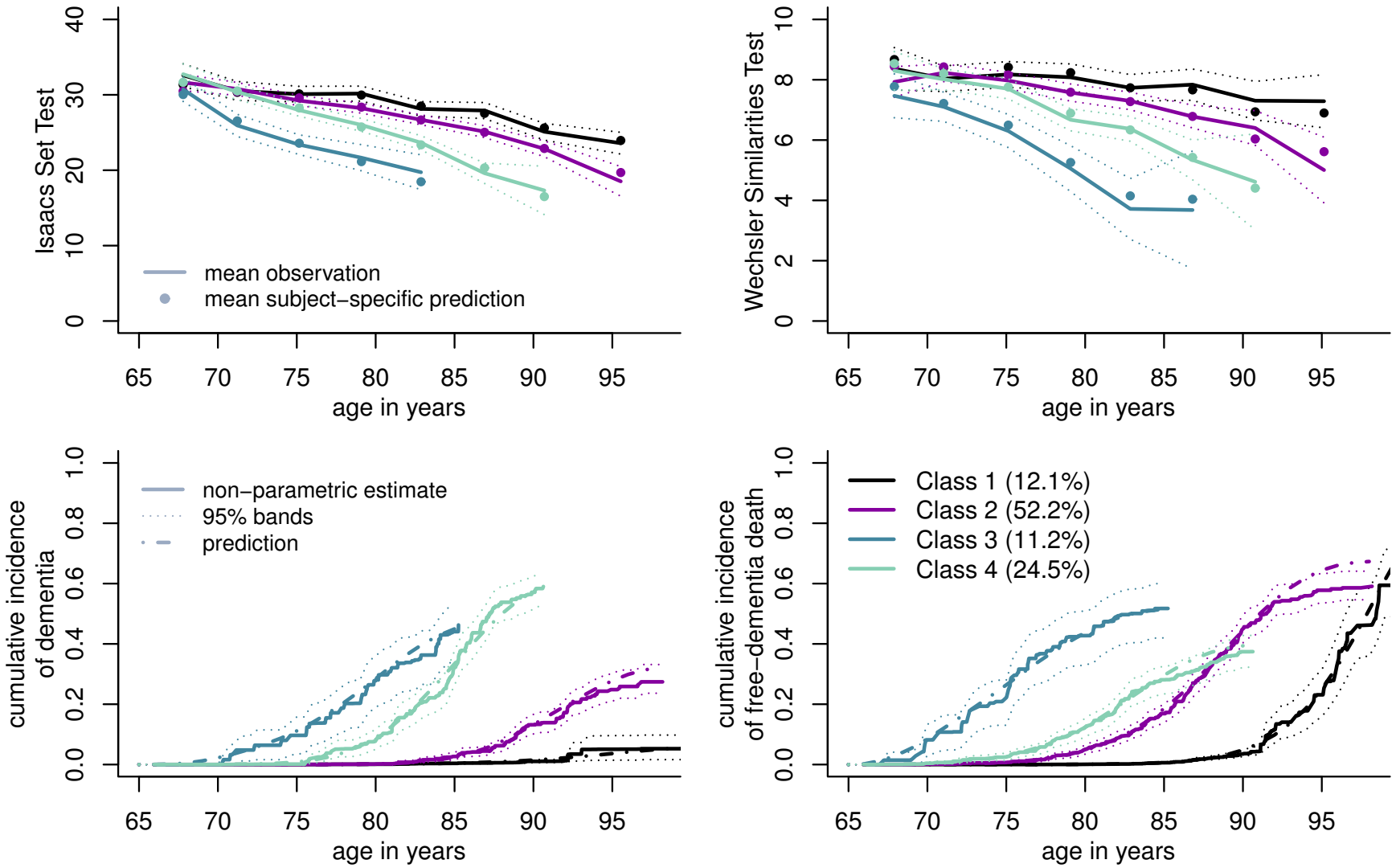
Selection of the number of classes in the joint model



4 profiles of semantic decline, dementia & death



Weighted predictions *versus* weighted observations



Individual dynamic prediction : the principle

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

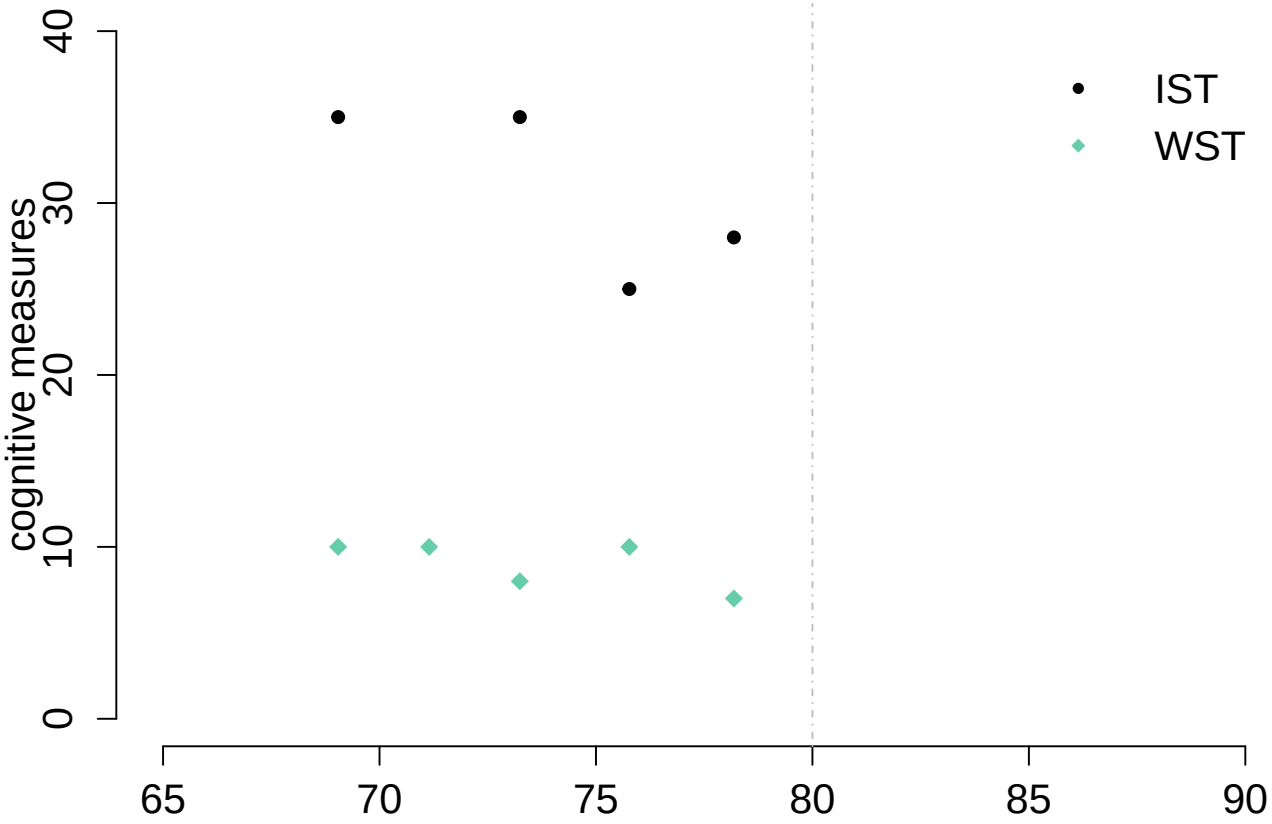
- Cumulative incidence for cause p at an horizon of t years

$$P_{pi}(s,t) = P(T_i \leq s+t, \delta_i = p | T_i > s, Y_i^{(s)}, X_i^{(s)}, X_{Ti}, X_{ci}; \theta) \quad (1)$$

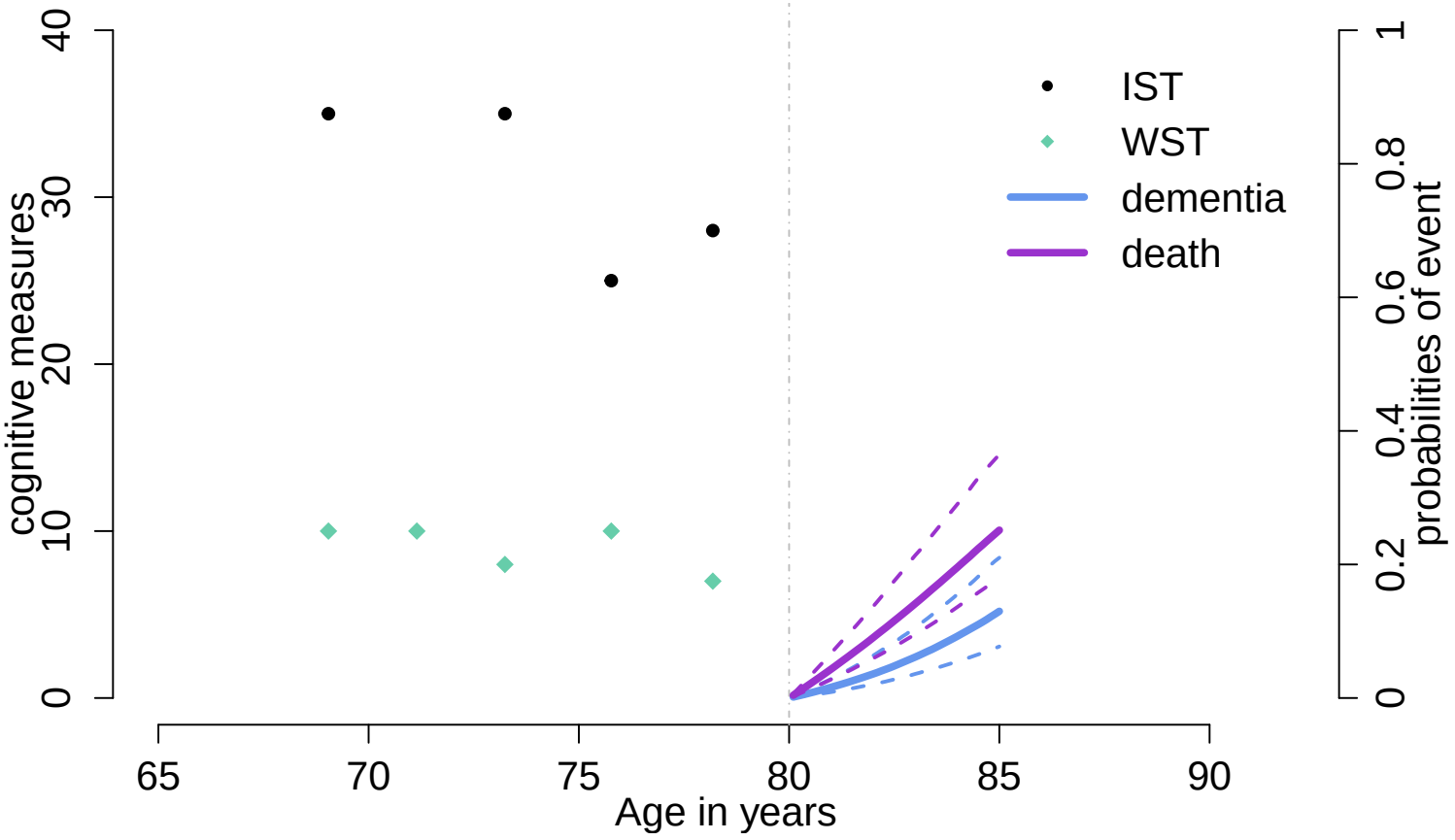
  - computed using the Bayes theorem
    and the conditional independence assumption
  - with class-specific and cause-specific cumulative incidence
    approximated by Gauss-Legendre

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

Example of individual dynamic prediction

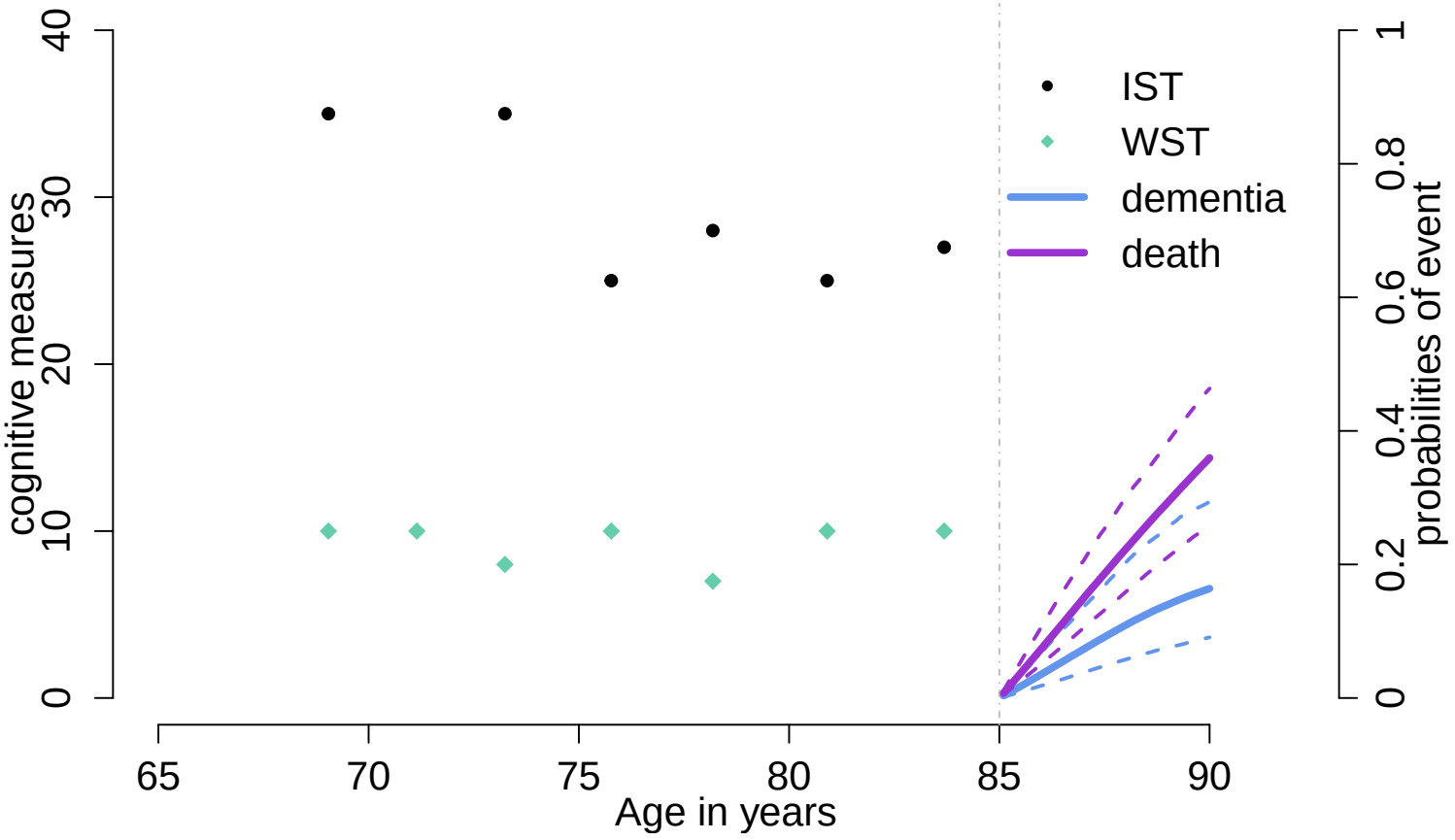


Example of individual dynamic prediction



	at 80 years old
5-year probability of dementia (%) :	13.0 [7.7,21.0]
5-year probability of death (%) :	25.1 [18.1,36.5]

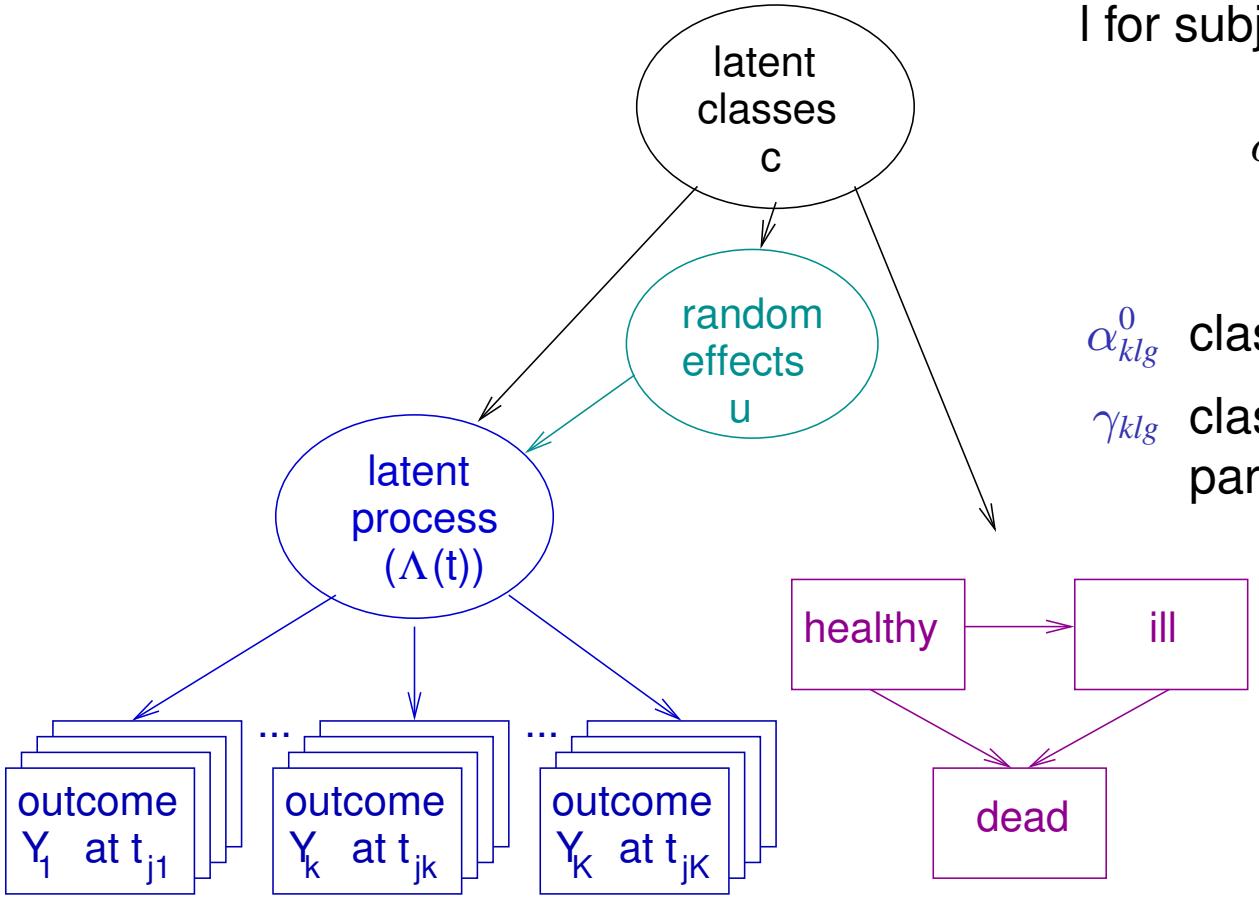
Example of individual dynamic prediction



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

Option2 : multistate model with interval censoring

(Rouanet et al., ArXiv 2015)



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

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

α^0_{klg} class-specific baseline intensity

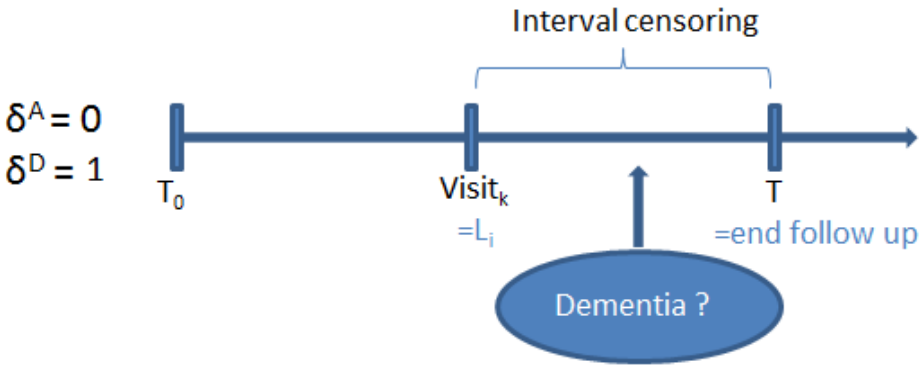
γ_{klg} class-specific regression parameters

Maximum Likelihood Estimation

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

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

  - $P(c_i = g|X_i;\theta)$ 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



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

$$= \mathrm{e}^{-A_{01_g}(T)-A_{02_g}(T)}\ \alpha_{02_g}(T) + \int_{L_i}^{\top} \mathrm{e}^{-A_{01_g}(u)-A_{02_g}(u)}\ \alpha_{01_g}(u)\ \mathrm{e}^{-(A_{12_g}(T)-A_{12_g}(u))}\ \alpha_{12_g}(T)\mathrm{d}u$$

Application

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

- Mixed model :

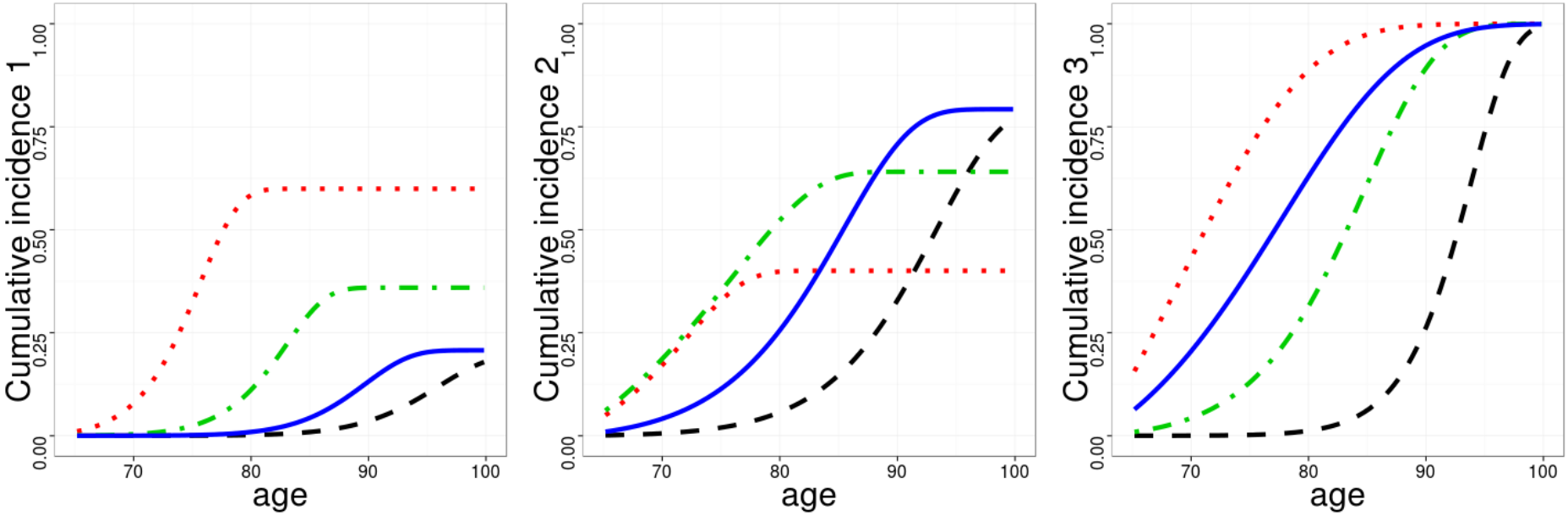
$$\Lambda(\mathbf{t})|_{c_i=g} = b_{0ig} + \beta_{1g}\texttt{sex} + \beta_{2g}\texttt{EL} + (b_{1ig} + \beta_{3g}\texttt{EL}) \times \mathbf{t} + (b_{2ig} + \beta_{4g}\texttt{EL}) \times \mathbf{t}^2$$

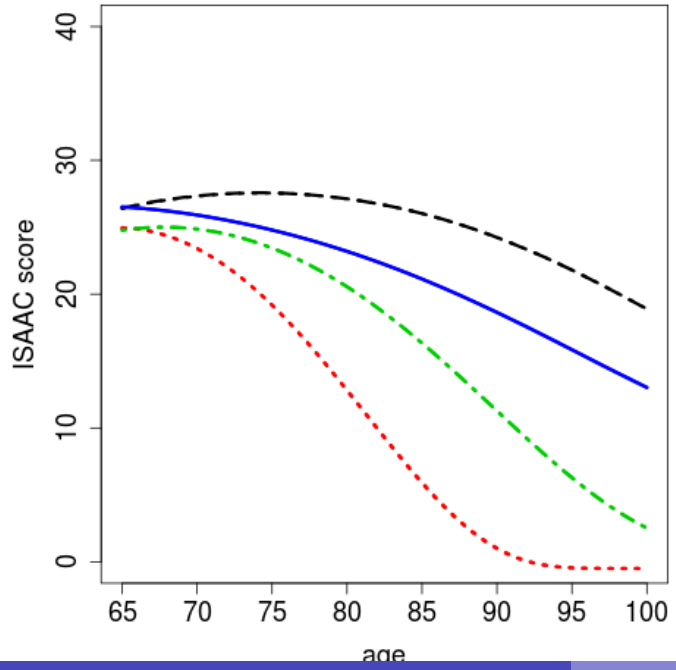
$$Y_{ij} = \Lambda_i(t_{ij}) + \epsilon_{ij} \quad \text{with} \quad b_{ig} \sim \mathcal{N}(\mu_g, B) \& \epsilon_i \sim \mathcal{N}(0, \sigma^2)$$

- Proportional transition intensities of the illness-death model :

$$\alpha_{klig}(t) = \alpha^0_{klg}(t)\ \mathrm{e}^{\gamma_{kls}\ \texttt{sex}_i + \gamma_{kle}\ \texttt{EL}_i}$$

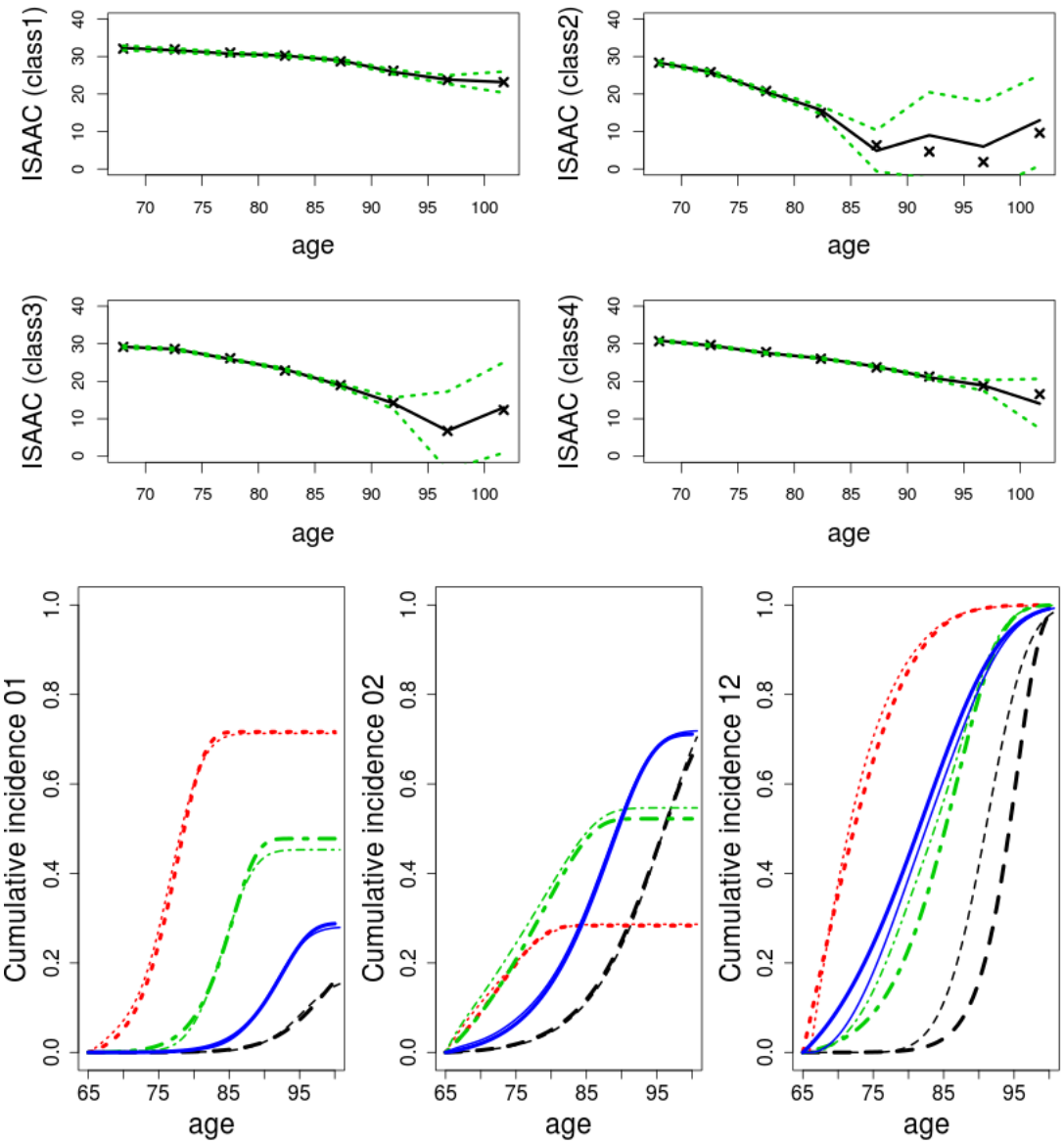
Four-class Markovian model : poorly educated men





	N	EL=0	EL=1	men	women
Class 1	7.3%	42.4	57.6	42.8	57.2
Class 2	8.3%	25.0	75.0	43.1	56.9
Class 3	34.2%	29.3	70.7	39.4	60.6
Class 4	50.5%	37.6	62.4	43.9	56.1

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 :

Ferrer Loïc et al. (2015). arXiv :1506.07496 [stat].

Rouanet Anaïs et al. (2015). arXiv :1506.07415 [stat].

Proust-Lima Cécile et al. (2014). arXiv :1409.7598 [stat].

Proust-Lima Cécile et al. (2014). Statistical Methods in Medical Research, 23(1), 74-90.

Sène Mbéry et al. (2014). Journal de la Société Française de Statistique, 155(1), 134-155. (in english)

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