

# Modélisation conjointe de données longitudinales et de temps d'événements dans un cadre multi-états

Loïc Ferrer<sup>1</sup>, James Dignam<sup>2</sup>, Virginie Rondeau<sup>1</sup>,  
Tom Pickles<sup>3</sup> et Cécile Proust-Lima<sup>1</sup>

<sup>(1)</sup>INSERM U897 - Équipe Biostatistique, ISPED, Université Bordeaux 2

<sup>(2)</sup>Department of Health Studies, University of Chicago

<sup>(3)</sup>Department of Radiation Oncology, University of British Columbia

3<sup>èmes</sup> Rencontres R

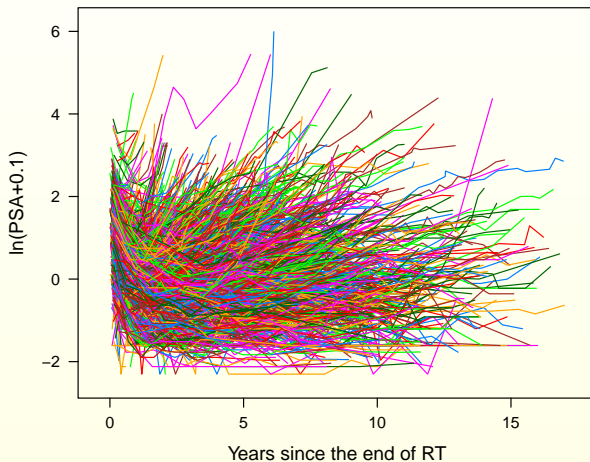
27 Juin 2014

# Contexte clinique

- ▶ **Cancer de la prostate** localisé et traité par radiothérapie
- ▶ **Marqueur longitudinal** : Prostate-Specific-Antigen (PSA)
- ▶ **Événements** : Rechutes cliniques + hormonothérapie
  - ▶ Locale
  - ▶ Distante
  - ▶ Décès dû au cancer de la prostate
  - ▶ Décès d'une autre cause
  - ▶ Hormonothérapie
- ▶ **Objectifs** :
  1. Étudier l'évolution des PSA au cours du suivi du patient
  2. Évaluer l'impact des facteurs pronostiques classiques sur chaque phase du processus multi-états
  3. Comprendre le lien entre PSA et rechutes cliniques

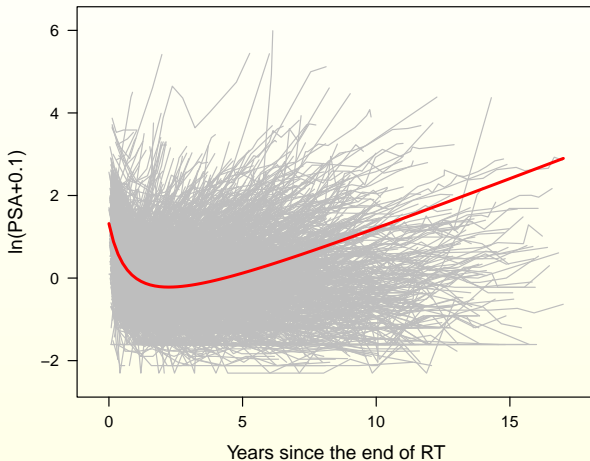
# Marqueur longitudinal

- ▶ 1475 patients, 10 (3, 21) mesures par patient (Médiane, (5<sup>ème</sup>, 95<sup>ème</sup> %iles))
- ▶ Mesures répétées jusqu'à l'apparition du premier événement



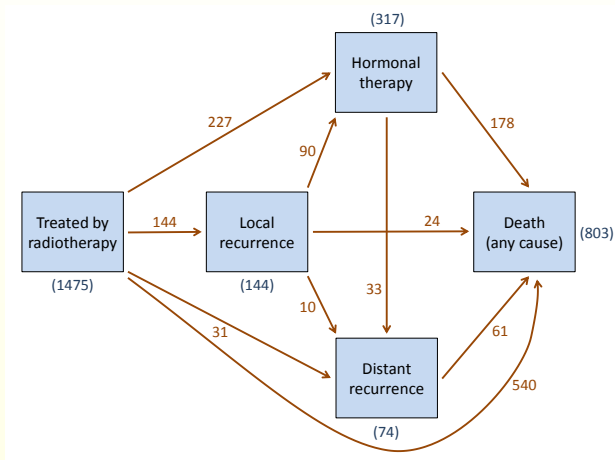
# Marqueur longitudinal

- ▶ 1475 patients, 10 (3, 21) mesures par patient (Médiane, (5<sup>ème</sup>, 95<sup>ème</sup> %iles))
- ▶ Mesures répétées jusqu'à l'apparition du premier événement



# Différents types d'événements

## ► Transitions du sous-modèle multi-états



# Contexte statistique

- ▶ Idée de la modélisation conjointe
  - ▶ **1 modèle mixte** pour les mesures répétées du marqueur longitudinal
  - ▶ **1 modèle de survie** pour les temps d'événements
  - ▶ **1 structure latente commune** pour lier les deux processus (classes latentes ou effets aléatoires partagés)
- ▶ Avantages de la modélisation conjointe
  - ▶ Modélisation simultanée des processus longitudinal et de survie
  - ▶ Réduction des biais
  - ▶ Quantification de l'association entre les deux processus
- ▶ Modèles conjoints à effets aléatoires partagés sous R :
  - ▶ Package JM de Rizopoulos
  - ▶ Limites : Modélisation conjointe d'un marqueur longitudinal et de temps d'événement unique, ou d'événements compétitifs
  - ▶ JM étendu au cadre multi-états

# Sous-modèle linéaire mixte

## ► Trajectoire du marqueur longitudinal

$$\begin{aligned}
 Y_i(t_{ij}) &= \ln(\text{PSA}_i(t_{ij}) + 0.1) = Y_i^*(t_{ij}) + \epsilon_i(t_{ij}) \\
 &= X_{Li}(t_{ij})^T \beta + Z_i(t_{ij})^T b_i + \epsilon_i(t_{ij}) \\
 &= (X_{0i}^T \beta_0 + b_{0i}) + (X_{1i}^T \beta_1 + b_{1i}) f_1(t_{ij}) \\
 &\quad + (X_{2i}^T \beta_2 + b_{2i}) f_2(t_{ij}) + \epsilon_i(t_{ij})
 \end{aligned}$$

- $Y_i = (Y_i(t_{i1}), \dots, Y_i(t_{in_i}))$  les  $n_i$  mesures observées du biomarqueur pour l'individu  $i$
- $Y_i^*$  les vraies (non-observées) mesures répétées du biomarqueur
- $\beta$  vecteur des effets fixes
- $b_i \sim N(\mathbf{0}, D)$  vecteur des effets aléatoires
- $\epsilon_i \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{n_i})$

# Sous-modèle linéaire mixte

## ► Trajectoire du marqueur longitudinal

$$\begin{aligned}
 Y_i(t_{ij}) &= \ln(\text{PSA}_i(t_{ij}) + 0.1) = Y_i^*(t_{ij}) + \epsilon_i(t_{ij}) \\
 &= X_{Li}(t_{ij})^T \beta + Z_i(t_{ij})^T b_i + \epsilon_i(t_{ij}) \\
 &= (X_{0i}^T \beta_0 + b_{0i}) + (X_{1i}^T \beta_1 + b_{1i}) f_1(t_{ij}) \\
 &\quad + (X_{2i}^T \beta_2 + b_{2i}) f_2(t_{ij}) + \epsilon_i(t_{ij})
 \end{aligned}$$

- $Y_i = (Y_i(t_{i1}), \dots, Y_i(t_{in_i}))$  les  $n_i$  mesures observées du biomarqueur pour l'individu  $i$
- $Y_i^*$  les vraies (non-observées) mesures répétées du biomarqueur
- $\beta$  vecteur des effets fixes
- $b_i \sim N(\mathbf{0}, D)$  vecteur des effets aléatoires
- $\epsilon_i \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{n_i})$



# Sous-modèle linéaire mixte

## ► Trajectoire du marqueur longitudinal

$$\begin{aligned}
 Y_i(t_{ij}) &= \ln(\text{PSA}_i(t_{ij}) + 0.1) = Y_i^*(t_{ij}) + \epsilon_i(t_{ij}) \\
 &= X_{Li}(t_{ij})^T \beta + Z_i(t_{ij})^T b_i + \epsilon_i(t_{ij}) \\
 &= (X_{0i}^T \beta_0 + b_{0i}) + (X_{1i}^T \beta_1 + b_{1i}) f_1(t_{ij}) \\
 &\quad + (X_{2i}^T \beta_2 + b_{2i}) f_2(t_{ij}) + \epsilon_i(t_{ij})
 \end{aligned}$$

- $Y_i = (Y_i(t_{i1}), \dots, Y_i(t_{in_i}))$  les  $n_i$  mesures observées du biomarqueur pour l'individu  $i$
- $Y_i^*$  les vraies (non-observées) mesures répétées du biomarqueur
- $\beta$  vecteur des effets fixes
- $b_i \sim N(\mathbf{0}, D)$  vecteur des effets aléatoires
- $\epsilon_i \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{n_i})$

# Sous-modèle de survie

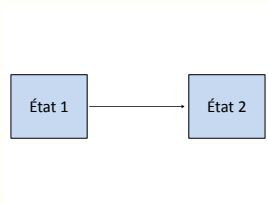
- ▶ **Standard** (type d'événement unique) :

$$\lambda_i(t|X_{Si}, \mathbf{b}_i) = \lambda_0(t) \exp(X_{Si}^T \gamma + W_i(\mathbf{b}_i, t)^T \eta)$$

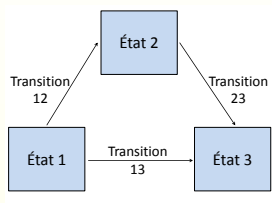
- ▶ On observe  $T_i = \min(T_i^*, C_i)$
- ▶  $\lambda_0(\cdot)$  la fonction de **risque de base**
- ▶  $X_{Si}$  vecteur des **facteurs pronostiques**
- ▶  $W_i(\mathbf{b}_i, t)^T$  définit la structure de dépendance entre les deux processus :
  - $W_i(\mathbf{b}_i, t) = Y_i^*(t)$   $\longrightarrow$  (niveau courant non bruité)
  - $W_i(\mathbf{b}_i, t) = \partial Y_i^*(t)/\partial t$   $\longrightarrow$  (pente courante non bruitée)
  - $W_i(\mathbf{b}_i, t) = (Y_i^*(t), \partial Y_i^*(t)/\partial t)^T$   $\longrightarrow$  (les deux)
- ▶  $\eta$  quantifie le lien entre les deux processus

# Sous-modèle multi-états

Survie



Multi-états



## ► Multi-états : Spécification de la transition

- On observe  $T_i = (T_{i,1}, \dots, T_{i,m_i+1})$  le temps d'entrée dans l'état initial et les  $m_i$  temps exacts de transitions jusqu'à  $C_i$  le temps de censure
- Pour chaque transition  $j \rightarrow k$  :

$$\lambda_{jk}^i(t|X_{Si}, \mathbf{b}_i) = \lambda_{jk,0}(t) \exp(X_{Si}^T \gamma_{jk} + W_i(\mathbf{b}_i, t)^T \eta_{jk})$$

## Approche par maximum de vraisemblance

- ▶ Risque de base paramétrique : Weibull, Gamma, B-splines, ...
- ▶ En temps continu, la contribution à la vraisemblance pour le sujet  $i$  dans le modèle multi-états classique s'écrit :

$$L_{mstate,i}(\theta_t | b_i) = \left[ \prod_{r=1}^{m_i} P_{E_r, E_r}^i(T_{i,r}, T_{i,(r+1)} - | b_i) \lambda_{E_r, E_{(r+1)}}^i(T_{i,r+1} | b_i) \right] \\ \times P_{E_{(m_i+1)}, E_{(m_i+1)}}^i(T_{i,m_i+1}, C_i | b_i)$$

avec  $P_{jk}(s, t) = \Pr(E(t) = k, E(s) = j)$

- ▶ Comme  $Y_i \perp\!\!\!\perp_{b_i} T_i$ , la contribution individuelle à la vraisemblance dans le modèle conjoint s'écrit :

$$L_{joint,i}(\theta) = \int_{b_i} L_{longit,i}(\theta_y | b_i) L_{mstate,i}(\theta_t | b_i) L_{random,i}(\theta_b) db_i$$

avec  $\theta = (\theta_t^T, \theta_y^T, \theta_b^T)^T$  le vecteur de l'ensemble des paramètres.

- ▶ Rappel :  $b_i \sim \mathcal{N}_{q_b}(0, D)$  et  $Y_i | b_i \sim \mathcal{N}_{n_i}(X_{L_i}^T \beta + Z_i^T b_i, \sigma^2 I_{n_i})$

## Approche par maximum de vraisemblance

- ▶ Risque de base paramétrique : Weibull, Gamma, B-splines, ...
- ▶ En temps continu, la contribution à la vraisemblance pour le sujet  $i$  dans le modèle multi-états classique s'écrit :

$$L_{mstate,i}(\theta_t | b_i) = \left[ \prod_{r=1}^{m_i} P_{E_r, E_r}^i(T_{i,r}, T_{i,(r+1)} - | b_i) \lambda_{E_r, E_{(r+1)}}^i(T_{i,r+1} | b_i) \right] \\ \times P_{E_{(m_i+1)}, E_{(m_i+1)}}^i(T_{i,m_i+1}, C_i | b_i)$$

avec  $P_{jk}(s, t) = \Pr(E(t) = k, E(s) = j)$

- ▶ Comme  $Y_i \perp\!\!\!\perp_{b_i} T_i$ , la contribution individuelle à la vraisemblance dans le modèle conjoint s'écrit :

$$L_{joint,i}(\theta) = \int_{b_i} L_{longit,i}(\theta_y | b_i) L_{mstate,i}(\theta_t | b_i) L_{random,i}(\theta_b) db_i$$

avec  $\theta = (\theta_t^T, \theta_y^T, \theta_b^T)^T$  le vecteur de l'ensemble des paramètres.

- ▶ Rappel :  $b_i \sim \mathcal{N}_{q_b}(0, D)$  et  $Y_i | b_i \sim \mathcal{N}_{n_i}(X_{L_i}^T \beta + Z_i^T b_i, \sigma^2 I_{n_i})$

# Implémentation sous R

## ► Modèle conjoint multi-états

$$\begin{cases} Y_i(t_{ij}) &= X_{Li}(t_{ij})^T \beta + Z_i(t_{ij})^T \mathbf{b}_i + \epsilon_i(t_{ij}) \\ \lambda_{jk}^i(t|X_{Si}, \mathbf{b}_i) &= \lambda_{jk,0}(t) \exp(X_{Si}^T \gamma_{jk} + W_i(\mathbf{b}_i, t)^T \eta_{jk}) \end{cases}$$

## ► JM

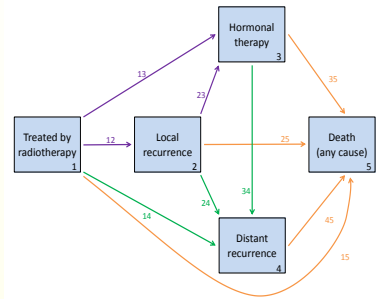
- **lme** {*nlme*} pour initialiser les paramètres dans le sous-modèle linéaire mixte
- **coxph** {*survival*} pour initialiser les paramètres dans le sous-modèle à risques proportionnels
- **jointModel** {*JM*} pour le modèle conjoint (événement unique ou événements compétitifs)

## ► "JMSTATE"

- **lme** {*nlme*}
- **coxph** {*mstate*} pour initialiser les paramètres dans le sous-modèle multi-états
- fonction **jointModel** {*JM*} étendue au cadre multi-états

# Réduction du nombre de paramètres

- ▶ Hypothèses de proportionnalité sur les 10 risques de base



- ▶ 3 risques de base à estimer :  $\lambda_{12,0}(t)$ ,  $\lambda_{14,0}(t)$  et  $\lambda_{15,0}(t)$

- ▶ On considère que certains facteurs pronostiques ont le même effet sur plusieurs transitions
- ▶ Critères de sélection du meilleur modèle

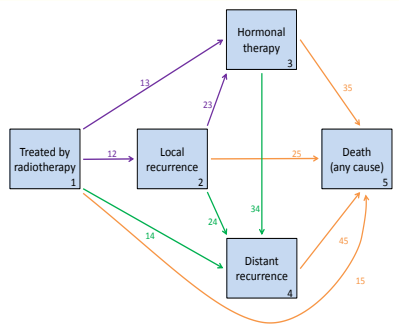
Modèle	L	AIC	BIC	p
Reduced.slope	-12731.4	25630.8	26075.7	84
<b>Reduced.both</b>	<b>-12671.1</b>	<b>25530.2</b>	<b>26028.1</b>	<b>94</b>

# Résultats

Estimation des paramètres d'association  
entre les processus longitudinal et multi-états

	Value	Std.Err	<i>p</i> -value
Level : 12	0.41	0.09	< 0.001
Level : 13	0.55	0.07	< 0.001
Level : 14	0.55	0.14	< 0.001
Level : 15	-0.05	0.05	0.339
Level : 23	-0.18	0.10	0.071
Level : 24	-0.24	0.20	0.220
Level : 25	0.25	0.16	0.103
Level : 34	-0.06	0.09	0.512
Level : 35	0.03	0.03	0.355
Level : 45	-0.17	0.05	0.002
.....			
Slope : 12	2.43	0.30	< 0.001
Slope : 13	2.88	0.25	< 0.001
Slope : 14	2.67	0.52	< 0.001
Slope : 15	0.40	0.34	0.241
Slope : 23	1.64	0.56	0.004
Slope : 24	3.91	1.09	< 0.001
Slope : 25	-1.61	1.39	0.247
Slope : 34	0.61	0.67	0.366
Slope : 35	0.37	0.33	0.258
Slope : 45	1.45	0.49	0.003

Sous-modèle multi-états



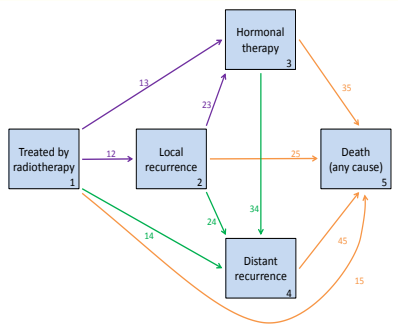


# Résultats

Estimation des paramètres d'association  
entre les processus longitudinal et multi-états

	Value	Std.Err	<i>p</i> -value
Level : 12	0.41	0.09	< 0.001
Level : 13	0.55	0.07	< 0.001
Level : 14	0.55	0.14	< 0.001
Level : 15	-0.05	0.05	0.339
Level : 23	-0.18	0.10	0.071
Level : 24	-0.24	0.20	0.220
Level : 25	0.25	0.16	0.103
Level : 34	-0.06	0.09	0.512
Level : 35	0.03	0.03	0.355
Level : 45	-0.17	0.05	0.002
.....			
Slope : 12	2.43	0.30	< 0.001
Slope : 13	2.88	0.25	< 0.001
Slope : 14	2.67	0.52	< 0.001
Slope : 15	0.40	0.34	0.241
Slope : 23	1.64	0.56	0.004
Slope : 24	3.91	1.09	< 0.001
Slope : 25	-1.61	1.39	0.247
Slope : 34	0.61	0.67	0.366
Slope : 35	0.37	0.33	0.258
Slope : 45	1.45	0.49	0.003

Sous-modèle multi-états

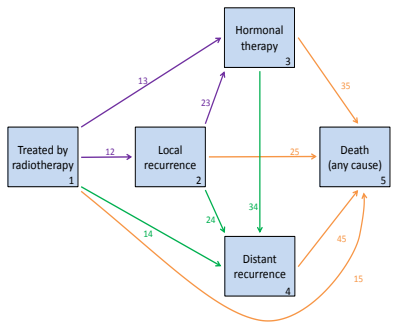


# Résultats

Estimation des paramètres d'association  
entre les processus longitudinal et multi-états

	Value	Std.Err	<i>p</i> -value
Level : 12	0.41	0.09	< 0.001
Level : 13	0.55	0.07	< 0.001
Level : 14	0.55	0.14	< 0.001
Level : 15	-0.05	0.05	0.339
Level : 23	-0.18	0.10	0.071
Level : 24	-0.24	0.20	0.220
Level : 25	0.25	0.16	0.103
Level : 34	-0.06	0.09	0.512
Level : 35	0.03	0.03	0.355
Level : 45	-0.17	0.05	0.002
.....			
Slope : 12	2.43	0.30	< 0.001
Slope : 13	2.88	0.25	< 0.001
Slope : 14	2.67	0.52	< 0.001
Slope : 15	0.40	0.34	0.241
Slope : 23	1.64	0.56	0.004
Slope : 24	3.91	1.09	< 0.001
Slope : 25	-1.61	1.39	0.247
Slope : 34	0.61	0.67	0.366
Slope : 35	0.37	0.33	0.258
Slope : 45	1.45	0.49	0.003

Sous-modèle multi-états

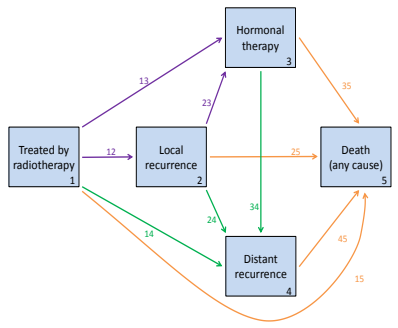


# Résultats

Estimation des paramètres d'association  
entre les processus longitudinal et multi-états

	Value	Std.Err	<i>p</i> -value
Level : 12	0.41	0.09	< 0.001
Level : 13	0.55	0.07	< 0.001
Level : 14	0.55	0.14	< 0.001
Level : 15	-0.05	0.05	0.339
Level : 23	-0.18	0.10	0.071
Level : 24	-0.24	0.20	0.220
Level : 25	0.25	0.16	0.103
<b>Level : 34</b>	-0.06	0.09	0.512
<b>Level : 35</b>	0.03	0.03	0.355
Level : 45	-0.17	0.05	0.002
.....			
Slope : 12	2.43	0.30	< 0.001
Slope : 13	2.88	0.25	< 0.001
Slope : 14	2.67	0.52	< 0.001
Slope : 15	0.40	0.34	0.241
Slope : 23	1.64	0.56	0.004
Slope : 24	3.91	1.09	< 0.001
Slope : 25	-1.61	1.39	0.247
<b>Slope : 34</b>	0.61	0.67	0.366
<b>Slope : 35</b>	0.37	0.33	0.258
Slope : 45	1.45	0.49	0.003

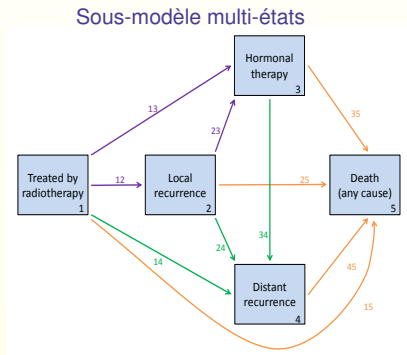
Sous-modèle multi-états



# Résultats

Estimation des paramètres d'association  
entre les processus longitudinal et multi-états

	Value	Std.Err	<i>p</i> -value
Level : 12	0.41	0.09	< 0.001
Level : 13	0.55	0.07	< 0.001
Level : 14	0.55	0.14	< 0.001
Level : 15	-0.05	0.05	0.339
Level : 23	-0.18	0.10	0.071
Level : 24	-0.24	0.20	0.220
Level : 25	0.25	0.16	0.103
Level : 34	-0.06	0.09	0.512
Level : 35	0.03	0.03	0.355
<b>Level : 45</b>	<b>-0.17</b>	<b>0.05</b>	<b>0.002</b>
.....			
Slope : 12	2.43	0.30	< 0.001
Slope : 13	2.88	0.25	< 0.001
Slope : 14	2.67	0.52	< 0.001
Slope : 15	0.40	0.34	0.241
Slope : 23	1.64	0.56	0.004
Slope : 24	3.91	1.09	< 0.001
Slope : 25	-1.61	1.39	0.247
Slope : 34	0.61	0.67	0.366
Slope : 35	0.37	0.33	0.258
<b>Slope : 45</b>	<b>1.45</b>	<b>0.49</b>	<b>0.003</b>



# Discussion

- ▶ **Modèle complet de progression du cancer** : tous les types de progression et le biomarqueur
- ▶ Programme d'estimation **validé par simulations**
- ▶ **Implémentation facile sous R** :
  - ▶ 2 packages : mstate (modèles multi-états) et JM (modèles conjoints standards)
- ▶ Modèles semi-Markoviens
- ▶ **Prédiction précise** des différents types de progression

# Références

## ▶ **Modélisation conjointe en cancer de la prostate :**

- ▶ Proust-Lima, C. et al. (2008). Determinants of change in prostate-specific antigen over time and its association with recurrence after external beam radiation therapy for prostate cancer in five large cohorts. *International Journal of Radiation Oncology Biology Physics*, 72(3), 782-791.
- ▶ Sène, M. (2013). Développement d'outils pronostiques dynamiques dans le cancer de la prostate localisé traité par radiothérapie. *Thèse de Santé Publique – Biostatistiques*, Université Bordeaux 2.

## ▶ **Modèles multi-états sous R :**

- ▶ De Wreede, L. C., Fiocco, M., & Putter, H. (2010). The mstate package for estimation and prediction in non- and semi-parametric multi-state and competing risks models. *Computer methods and programs in biomedicine*, 99(3), 261-274.

## ▶ **Modèles conjoints standards sous R :**

- ▶ Rizopoulos, D. (2010). JM : An R package for the joint modelling of longitudinal and time-to-event data. *Journal of Statistical Software*, 35(9), 1-33.

## Sous-modèle multi-états

- ▶ Intensités de transitions

$$\lambda_{jk}^i(t|X_{Si}, b_i) = \lambda_{jk,0}(t) \exp(X_{Si}^T \gamma_{jk} + W_i(b_i, t)^T \eta_{jk})$$

- ▶  $X_{Si}$  : iPSA, T-stage, Gleason, Cohorte

- ▶ Matrice des intensités de transition

$$\mathbf{Q} = \begin{pmatrix} -(\lambda_{12} + \lambda_{13} + \lambda_{14} + \lambda_{15}) & \lambda_{12} & \lambda_{13} & \lambda_{14} & \lambda_{15} \\ 0 & -(\lambda_{23} + \lambda_{24} + \lambda_{25}) & \lambda_{23} & \lambda_{24} & \lambda_{25} \\ 0 & 0 & -(\lambda_{34} + \lambda_{35}) & \lambda_{34} & \lambda_{35} \\ 0 & 0 & 0 & -\lambda_{45} & \lambda_{45} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

- ▶ Matrice des probabilités de transition

$$\mathbf{P} = \begin{pmatrix} P_{11} & P_{12} & P_{13} & P_{14} & P_{15} \\ 0 & P_{22} & P_{23} & P_{24} & P_{25} \\ 0 & 0 & P_{33} & P_{34} & P_{35} \\ 0 & 0 & 0 & P_{44} & P_{45} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

- ▶  $P_{jj}(s, t) = \exp \left[ \int_s^t \lambda_{jj}(u) du \right]$

- ▶ Transitions directes :

$$P_{jk}(s, t) = \int_s^t P_{jj}(s, u^-) \lambda_{jk}(u) P_{kk}(u, t) du$$

- ▶ Transitions indirectes :

$$P_{jl}(s, t) = 1 - \sum_{m \neq k} P_{jm}(s, t)$$

# Implémentation sous R

## ▶ JM

- ▶ 'lme' (package nlme) pour initialiser les paramètres dans le sous-modèle linéaire mixte
- ▶ 'coxph' (survival) pour initialiser les paramètres dans le sous-modèle à risques proportionnels
- ▶ 'jointModel' pour le modèle conjoint (événement unique ou événements compétitifs)

## ▶ Mstate

- ▶ Préparation des données
  - Dans le modèle de Cox classique, chaque ligne représente un patient.
  - Dans Mstate, chaque ligne représente une transition pour laquelle un patient est à risque.
- ▶ 'coxph' (survival) pour le modèle multi-états

## ▶ "JMSTATE"

- ▶ 'lme'
- ▶ Mstate ('coxph')
- ▶ fonction 'jointModel' étendue au cadre multi-états



# Calcul de la vraisemblance sous Mstate

- **Exemple** : Le sujet  $i$  a une rechute distante en  $T_{RD,i}$  puis est censuré en  $C_i$ .

$$\begin{aligned}
 L_{mstate,i} &= \lambda_{12}^i(T_{RD,i})^{\delta_{12,i}=0} \exp \left[ - \int_0^{T_{RD,i}} \lambda_{12}^i(u) du \right] && 1 \rightarrow 2 : \delta_{12,i} = 0, T_{12,i} = T_{RD,i} \\
 &\times \lambda_{13}^i(T_{RD,i})^{\delta_{13,i}=0} \exp \left[ - \int_0^{T_{RD,i}} \lambda_{13}^i(u) du \right] && 1 \rightarrow 3 : \delta_{13,i} = 0, T_{13,i} = T_{RD,i} \\
 &\times \lambda_{14}^i(T_{RD,i})^{\delta_{14,i}=1} \exp \left[ - \int_0^{T_{RD,i}} \lambda_{14}^i(u) du \right] && 1 \rightarrow 4 : \delta_{14,i} = 1, T_{14,i} = T_{RD,i} \\
 &\times \lambda_{15}^i(T_{RD,i})^{\delta_{15,i}=0} \exp \left[ - \int_0^{T_{RD,i}} \lambda_{15}^i(u) du \right] && 1 \rightarrow 5 : \delta_{15,i} = 0, T_{15,i} = T_{RD,i} \\
 &\times \lambda_{45}^i(C_i)^{\delta_{45,i}=0} \exp \left[ - \int_{T_{RD,i}}^{C_i} \lambda_{45}^i(u) du \right] && 4 \rightarrow 5 : \delta_{45,i} = 0, T_{45,i} = C_i \\
 &= \exp \left[ - \int_0^{T_{RD,i}} (\lambda_{12}^i(u) + \lambda_{13}^i(u) + \lambda_{14}^i(u) + \lambda_{15}^i(u)) du \right] \\
 &\times \lambda_{14}^i(T_{RD,i}) \times \exp \left[ - \int_{T_{RD,i}}^{C_i} (\lambda_{45}^i(u)) du \right] \\
 &= P_{11}^i(0, T_{RD,i}) \lambda_{14}^i(T_{RD,i}) P_{44}^i(T_{RD,i}, C_i)
 \end{aligned}$$

## Écriture de la vraisemblance (1/2)

Comme  $Y_i \perp\!\!\!\perp_{b_i} T_i$ , la contribution individuelle à la vraisemblance dans le modèle conjoint s'écrit :

$$L_{joint,i}(\theta) = \int_{b_i} L_{longit,i}(\theta_y|b_i)L_{mstate,i}(\theta_t|b_i)L_{random,i}(\theta_b) db_i$$

avec  $\theta = (\theta_y^T, \theta_t^T, \theta_b^T)^T$  le vecteur de l'ensemble des paramètres.

- ▶ **Partie longitudinale** ( $Y_i|b_i \sim \mathcal{N}_{n_i}(X_{L_i}^T\beta + Z_i^T b_i, \sigma^2 I_{n_i})$ ) :

$$L_{longit,i}(\theta_y|b_i) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^{n_i} \exp\left[-\frac{\|y_i - X_{L_i}^T\beta - Z_i^T b_i\|^2}{2\sigma^2}\right]$$

- ▶ **Partie effets aléatoires** ( $b_i \sim \mathcal{N}_{q_b}(0, D)$ ) :

$$L_{random,i}(\theta_b) = \left(\frac{1}{\sqrt{2\pi}}\right)^{q_b} \times \left(\frac{1}{\sqrt{\det(D)}}\right) \times \exp\left[-\frac{b_i^T D^{-1} b_i}{2}\right]$$

## Écriture de la vraisemblance (2/2)

- Partie multi-états (qui inclue les effets aléatoires) :

$$\begin{aligned} L_{mstate,i}(\theta_i | b_i) &= \left[ \prod_{r=1}^{m_i} P_{E_r, E_r}^i(T_{i,r}, T_{i,(r+1)} - | b_i) \lambda_{E_r, E_{(r+1)}}^i(T_{i,r+1} | b_i) \right] \\ &\quad \times P_{E_{(m_i+1)}, E_{(m_i+1)}}^i(T_{i,m_i+1}, C_i | b_i) \\ &= \left[ \prod_{r=1}^{m_i} \exp \left( \int_{T_{i,r}}^{T_{i,(r+1)} -} \lambda_{E_r, E_r}^i(u | b_i) du \right) \times \lambda_{E_r, E_{(r+1)}}^i(T_{i,r+1} | b_i) \right] \\ &\quad \times \exp \left( \int_{T_{i,m_i+1}}^{C_i} \lambda_{E_{(m_i+1)}, E_{(m_i+1)}}^i(u | b_i) du \right) \end{aligned}$$

avec  $T_{i,1}$  le temps d'entrée dans l'état initial, les temps exacts de transitions  $T_{i,2}, \dots, T_{i,m_i+1}$ , et le temps de censure  $C_i$ .