Advanced survival models and prediction for correlated data

Virginie Rondeau[1] ; Agnieszka Krol[1]

[1]INSERM U 1219, University of Bordeaux,

Virginie.Rondeau@isped.u-bordeaux2.fr

Short Course - Karolinska Institutet
Stockholm, October 24-25, 2016

Outline of the short course

- **Part 1** : **Standard frailty models**
  - Heterogeneity in survival data
  - Standard frailty models
- **Part 2** : **Extension of standard frailty models**
  - Nested frailty models
  - Additive frailty models
- **Part 3** : **Joint frailty models for recurrent events and terminal event**
  - Joint frailty models
  - Application
  - Prediction in joint frailty models
- **Part 4** : **Joint models for a longitudinal biomarker and terminal event**
  - Joint frailty models
  - Application
  - Prediction in joint frailty models

Inserm
Institut national
de la santé et de la recherche médicale

PART 1 :
Standard frailty models

Heterogeneity

- **Clustered data**
  ex : families, hospitals, trials,
- **Recurrent events**
  ex : recurrences of breast cancer

$\rightarrowtail$Heterogeneity in the population

Heterogeneity for clustered data (1)

- **Heterogeneity between groups**
  (= correlation intra-group)
  ie, similar survival times for the patients of the same group
  Correlation linked to a set of characteristics specific to each group :
  - cluster-level factors **measured : explain a part of the heterogeneity**
  - cluster-level factors **unmeasured : residual correlation**

  $\rightarrowtail$ use of **standard shared frailty models**

Heterogeneity for clustered data (2)

- for cluster $i = 1, ..., G$ and subject $j = 1, ..., n_i$ from cluster i
  Assumptions :
  $\rightarrowtail$ **independence** of survival times for **two different groups** (T_{ij} and $T_{i'j'}$)
  $\rightarrowtail$ **non independence** for two survival times of the **same cluster** (T_{ij} and $T_{ij'}$)
- *Example :* A multicentric cohort : subjects share the same occupational exposure in the same industry

Heterogeneity for recurrent outcomes (1)

Recurrent events :

- Patients may experience the outcome of interest more than once over a period of observation
- Naturally ordered failure times

$\rightarrowtail$ different events "within" an individual are correlated
$\rightarrowtail$ use of **standard shared frailty models**

Heterogeneity for recurrent outcomes (2)

Examples :

- HIV Patients : different opportunistic infections
- Patients with recurrent cardiovascular events : different heart attacks
- Patients with different hospitalisations
- Children with recurrent asthma attacks
- Patients with a first cancer (breast, lymphoma, bladder) : relapses of their first cancer

Heterogeneity for recurrent outcomes (3)

- **repeated outcomes (T_{ij})**
  j^{th} observation ($j = 1, ..., n_i$) of patient i ($i = 1, ..., G$)
  $\rightarrowtail$ intra-subject correlation
- **Choice of the time-scale** ?
  gap-time = time between two recurrent events
  or
  calendar time = time since inclusion

Recurrent outcomes (4)

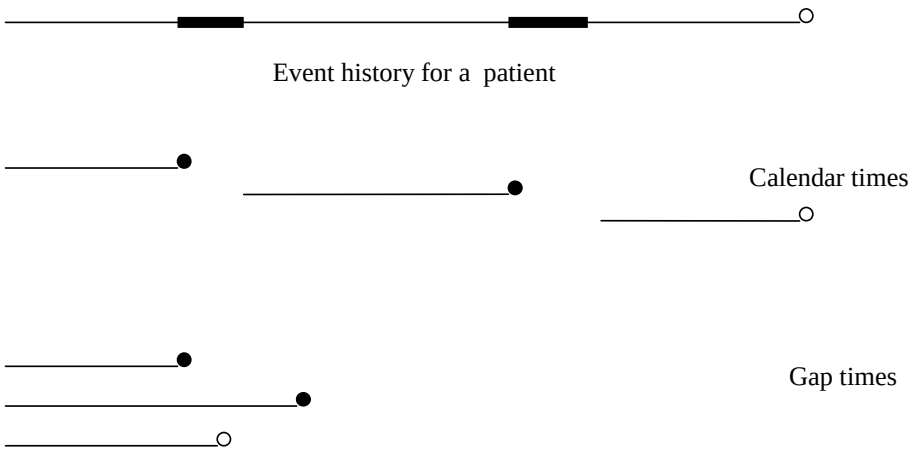


Figure 1: Event history for a patient with recurrent events together with the calendar times and the gap times; ● represents an event ○ a censoring time and ▬ the non at risk periods.

Choice of the timescale ?

Using clinical aspects

- time-between-events (*gap times* between 2 observations)
  after each event we reset the counter 0 (T0=0).
  ex : if after a first event, the risk for a second event increases, otherwise
- time-to-event (*calendar times*)
  (T0 $\neq$0 for $j > 0$, the beginning of the at risk period is not reset to zero) here, the risk of a new recurrent event is not altered by a previous event
- time since inclusion (*total time*)
  a subject is at risk for the k^{th} event since its entry into the study (T0= 0) $\rightarrow$ NO !

Illustration : rehospitalizations of patients with colorectal cancers

id	t.start	t.stop	time	event	chemo	sex	dukes	charlson
1	0	24	24	1	Treated	F	stage D	3
1	24	457	433	1	Treated	F	stage D	0 (ref)
1	457	1037	580	0	Treated	F	stage D	0 (ref)
2	0	489	489	1	Untreated	M	stage C	0 (ref)
2	489	1182	693	0	Untreated	M	stage C	0 (ref)
3	0	15	15	1	Untreated	M	stage C	3
3	15	783	768	0	Untreated	M	stage C	3
4	0	163	163	1	Treated	F	stage A-B	0 (ref)
4	163	288	125	1	Treated	F	stage A-B	0 (ref)
4	288	638	350	1	Treated	F	stage A-B	0 (ref)
4	638	686	48	1	Treated	F	stage A-B	0 (ref)
4	686	2048	1362	0	Treated	F	stage A-B	0 (ref)
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Reminder : survival analysis (1)

- Classical survival analysis : Cox proportional hazard models

  assumption = independence of the survival times
  (at least given the observed covariates)

  $\rightarrowtail$ assumption necessary for the estimation of the parameters
- Example : likelihood for right-censored data

$$V(\beta) = \prod_{j=1}^{n} \lambda(Y_j|X_j)^{\delta_j} S(Y_j|X_j)$$

(independence of the n observations)

Reminder : survival analysis (2)

- **assumption not valid :**
  - when studying patients from different clusters
  ex : families, hospitals, geographical areas
  $\rightarrowtail$ share the same environment (ex : diet, life-style, clinical practices, air pollution)
  - when studying recurrent events per patient
- **using a standard survival analysis in case of correlated data :** under-estimation of the standards errors of the regression parameters, especially for cluster specific covariates

Shared frailty model
= survival model with random effects

With right censored data

- **Notations :**
  j^{th} subject $(j = 1, ..., n_i)$ from i^{th} cluster $(i = 1, ..., G)$
  T_{ij} survival time and C_{ij} censoring time
  $Y_{ij} = \min(T_{ij}, C_{ij})$ observation time
  $\delta_{ij} = I_{\{T_{ij} \leq C_{ij}\}}$ censoring inicator

Shared frailty model

- **The model**

$$\lambda_{ij}(t|Z_i) = Z_i \lambda_0(t) \exp(\beta' X_{ij})$$

- $\lambda_0(t)$ the baseline hazard function
- $X_{ij} = (X_{1ij}, ..., X_{pij})'$ vector of the explanatory variables
- β vector of the regression coefficients
- Z_i random effects (or frailty variables) shared by all individuals from the same cluster
  $=$ all unobserved risk factors
  if $Z_i > 1$ hazard increased (frail subjects die earlier)

In the specific case where $n_i = 1$, the model is more an **overdispersion** model

Assumptions of the frailty model

$\rightarrowtail$ **independence** of the survival times **in each** cluster given the random effects

$\rightarrowtail$ **independence** of the survival times **between** clusters

$\rightarrowtail$ **Proportionality** of the hazards conditionally on the frailties, but not marginally

$\rightarrowtail$ choose a **distribution** for the random effects :

- gamma distribution (good mathematical properties) :
  Z_i iid et $\sim$ gamma $f_Z(z) = \frac{z^{(1/\theta-1)} \exp\{-z/\theta\}}{\Gamma(1/\theta)\theta^{1/\theta}}$
  $\mathrm{E}(\mathrm{Z_i}) = 1$ et $\mathrm{var}(\mathrm{Z_i}) = \theta$
- log-normal frailty
- positive stable frailty ...

Estimation of the parameters

$\xi = (\hat{\beta}, \hat{\theta}, \hat{\lambda}_0(t))$

- **Marginal log-likelihood for right censored-data**

$$V(Y_{ij}) = \prod_{i=1}^{G} \int_0^{+\infty} \prod_{j=1}^{n_i} \lambda_{ij}(Y_{ij}|Z_i)^{\delta_{ij}} \times S_{ij}(Y_{ij}|Z_i) g(Z_i) dZ_i$$

$\rightarrowtail$ marginal log-likelihood :

$l(Y_{ij}) = \log(V(Y_{ij})) = \log(\prod_{i=1}^{G} V_i(Y_{ij})) = \sum_{i=1}^{G}(\log(V_i(Y_{ij})))$

Estimation of the parameters

- Marginal log-likelihood for **right-censored data** and **gamma frailties**

$$
\begin{aligned}
l(Y_{ij}) = \sum_{i=1}^{G} \Bigg\{ & \sum_{j=1}^{n_i} \delta_{ij} \{ \beta' X_{ij} + \ln(\lambda_0(Y_{ij})) \} \\
& - (1/\theta + m_i) \ln \left[1 + \theta \sum_{j=1}^{n_i} \Lambda_0(Y_{ij}) \exp(\beta' X_{ij}) \right] \\
& + I_{\{m_i \neq 0\}} \sum_{k=1}^{m_i} [\ln(1 + \theta(m_i - k))] \Bigg\}
\end{aligned}
$$

with m_i the number of events in cluster i

$\longrightarrow$ **analytical solution** for the integration

Estimation of the parameters

- for **right-censored** and **left-truncated** data

- $\mathbf{Y_i} = (Y_{i1}, ..., Y_{in_i})$ the n_i observation times from cluster i,
  $\mathcal{L}_i = (\mathcal{L}_{i1}, ..., \mathcal{L}_{in_i})$ the left truncated times
  ex : age as the basic timescale in a cohort of patients included at 65 years (and over)
- With left truncation times, the frailty distribution among survivors in a cluster is used

$$V_i(.) = \int_{Z_i} \frac{V_i(\mathbf{Y_i}|Z_i)}{S_i(\mathcal{L}_i|Z_i)} f(Z_i|T_{ij} > \mathcal{L}_{ij}, j = 1, ..., n_i) dZ_i$$

with

$$f(Z_i|T_{ij} > \mathcal{L}_{ij}, \forall j) = \frac{(1/\theta + \Lambda_i(\mathcal{L}_i))^{1/\theta} Z_i^{(1/\theta - 1)} \exp(-Z_i(1/\theta + \Lambda_i(\mathcal{L}_i)))}{\Gamma(1/\theta)}$$

(Rondeau, LIDA 2003 ; Lawless, Stat Med 1999)

Estimation of the parameters

- for **right-censored** and **left-truncated** data

$$
\begin{aligned}
l(Y_{ij}) = \sum_{i=1}^{G} \Bigg\{ & \sum_{j=1}^{n_i} \delta_{ij} \{\beta' X_{ij} + \ln(\lambda_0(Y_{ij}))\} \\
& -(1/\theta + m_i) \ln \left[1 + \theta \sum_{j=1}^{n_i} \Lambda_0(Y_{ij}) \exp(\beta' X_{ij}) \right] \\
& + I_{\{m_i \neq 0\}} \sum_{k=1}^{m_i} [\ln(1 + \theta(m_i - k))] \\
& + 1/\theta \ln(1 + \theta \sum_{j=1}^{n_i} \Lambda_0(\mathcal{L}_{ij}) \exp(\beta' X_{ij})) \Bigg\}
\end{aligned}
$$

Estimation of the parameters

Estimation in a maximum likelihood framework (frequentist) :

- with penalized partial likelihood *(Therneau, Springer 2000)*

$$ppl = \underbrace{pl(\beta, Z^*; \text{data})}_{partial.log-lik} - \underbrace{g(Z^*; \theta)}_{penalisation}$$

- with the EM algorithm (and Breslow estimator for $\Lambda_0(.)$) *(Nielsen, Scand J Stat 1992 ; Klein, Biom 1992 ; Parner, thesis 1997)*
- with penalized likelihood *(Rondeau, LIDA 2003)*

$$pl(\lambda_0(.), \beta, \theta) = l(.) - \underbrace{\kappa \int_0^\infty \lambda_0''^2(t)dt}_{penalisation}$$

The baseline hazard function in the shared frailty models

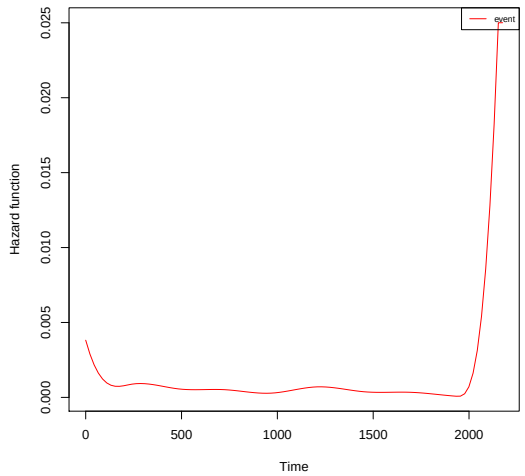
$$\lambda_{ij}(t|Z_i) = Z_i \lambda_0(t) \exp(\beta' X_{ij})$$

- Semi-parametric approach (approximation with splines)
- Parametrical approach (Weibull, Piecewise constant)

The hazard function : approximation with splines

$$\lambda_{ij}(t|Z_i) = Z_i \lambda_0(t) \exp(\beta' X_{ij})$$

- $\lambda_0(t) = \sum_{i=1}^{m} \eta_i M_i(.)$
  with $M_i(.)$ Cubic M-splines of order 3,
  an $\eta = (\eta_1, ..., \eta_m)$ the vector of splines coefficients
- We can use equidistant or percentiles knots



The baseline hazard function : piecewise constant

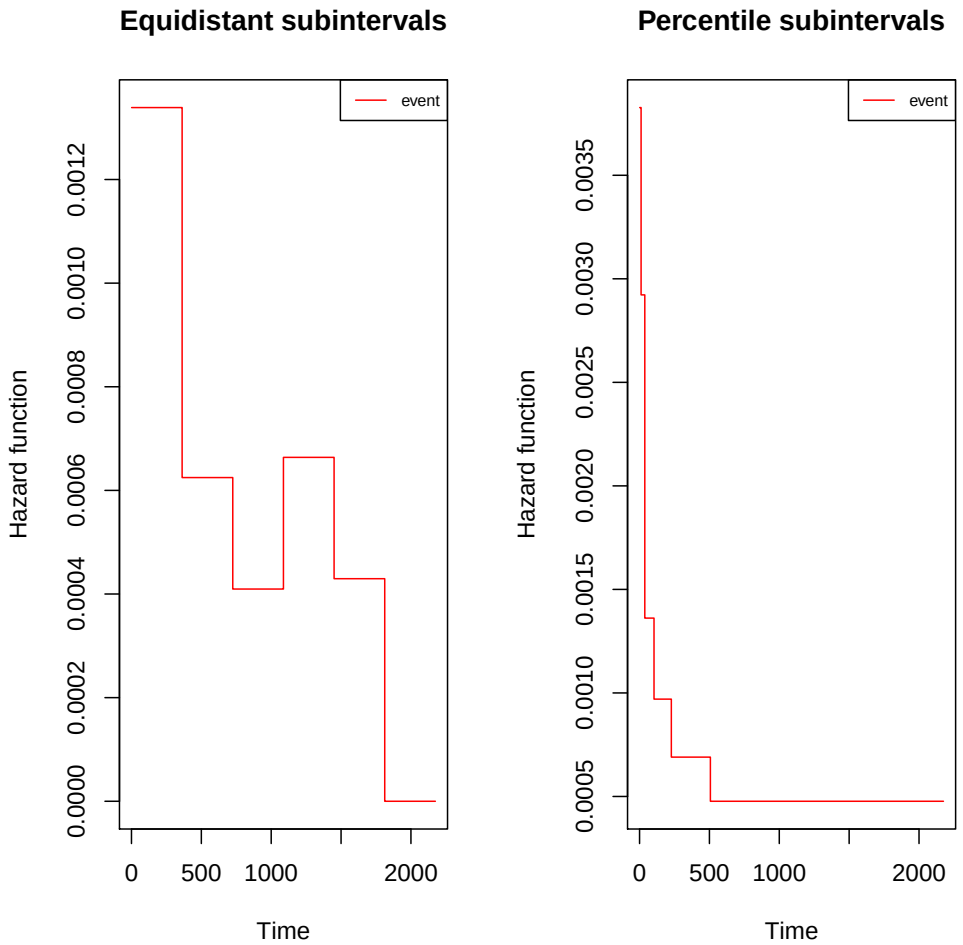
$$\lambda_{ij}(t|Z_i) = Z_i \lambda_0(t) \exp(\beta' X_{ij})$$

$$\lambda_0(t) = \sum_{i=1}^{n_{int}} I_{\{t \in (t_{i-1}, t_i)\}} c_i$$

in the interval $[0, \tau]$ and τ the last observed time among N individuals and n_{int} the number of subintervals

- using **equidistant intervals** between two knots (all the subintervals are of the same length)
- or using **percentiles** (in each subinterval the same number of events is observed)

The baseline hazard function : piecewise constant
Examples "nb.int=6"



The baseline hazard function : Weibull baseline hazard function

$$\lambda_{ij}(t|Z_i) = Z_i \lambda_0(t) \exp(\beta' X_{ij})$$

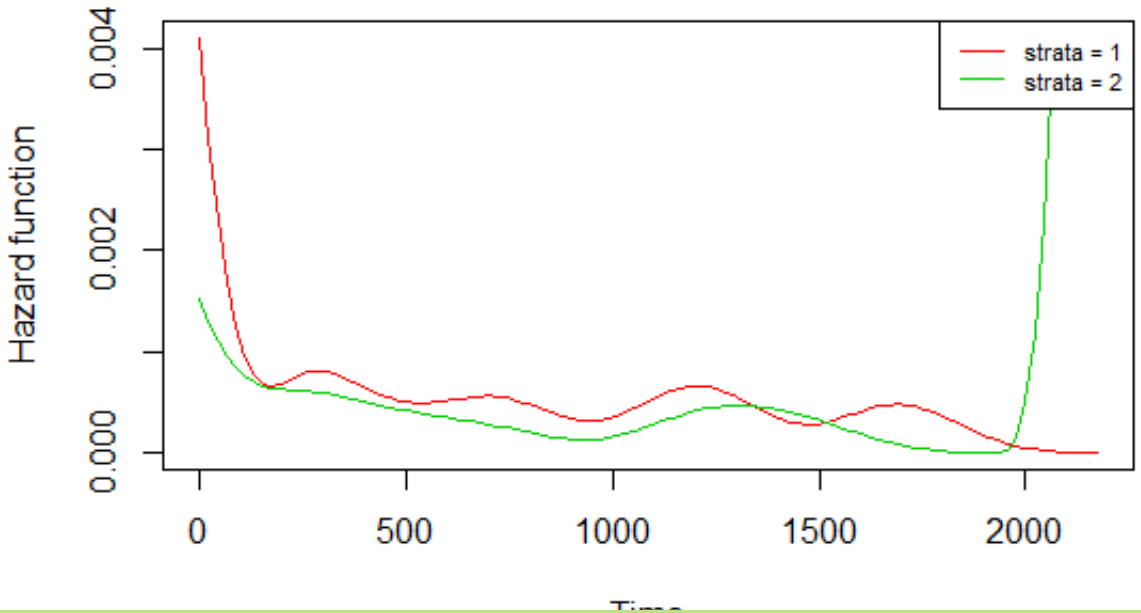
- $\lambda_0(t) = (at^{a-1})/b^a$
- with $a > 0$ the shape parameter and $b > 0$ the scale parameter

Stratification in the shared frailty models

- The **stratified** model

$$\lambda_{ij}(t|Z_i) = Z_i \lambda_{0h}(t) \exp(\beta' X_{ij})$$

- $\lambda_{0h}(t)$ the baseline hazard function different for each category of the variable of stratification
- *Example* : stratification on gender



Programs available (to my knowledge)

- SAS : PHREG /only log-normal frailty model

```
proc phreg data=mydata;
    class ID Treat Type;
    model Time*Status(0)=Treat I Type; #
        interaction
    random ID;
    hazardratio  Treat;
    run;
```

- SAS :NLMIXED /parametrical /interval censoring

  (Bellamy, Stat Med 2004)

```
proc nlmixed data=asthma qpoints=10;
parms beta0=3.9 eta1=-0.2 p=0.9 theta=0.01;
bounds p>0, theta >0;
ebetaxb=exp(-beta0 + beta1*1ri1 + b));
lambda=exp(-beta0);
s1=exp(-(t1*ebetaxb)**(1/p));
su=exp(-(t2*ebetaxb)**(1/p));
ft=((lambda*p)*(lambda*t1)**(p-1))*ebetaxb**1/p;
#ctype(censoring type): 1=exact 2=left 3=right 4=
    int
if ctype=1 then lik=ft;
 else if ctype=2 then lik=1-su;
  else if ctype=3 then lik=s1;else lik=s1-su;
llik=log(lik); model y ~ general(llik);
random b~normal(0,theta) subject=clusidz;run;
```

Programs available

- STATA : streg /parametrical

```
streg age smoking, dist(weib) frailty(gamma) nlog
```

```
streg age smoking, dist(weib) frailty(invgauss)
    nlog
```

Programs available

- MIXGSUR
  http ://tigger.uic.edu/∼ hedeker/
- Survival Kit / Bayesian approach
  http ://www.nas.boku.ac.at/1897.html

R Packages on CRAN

- R : parfm / parametrical / different frailty distribution
- R : frailtyHL / via H-likelihood (Laplace approximation)
- R : survival / via penalized partial likelihood

```
coxph(Surv(time,status)~age+sexe+frailty(gpe
    ),data=dataR)
```

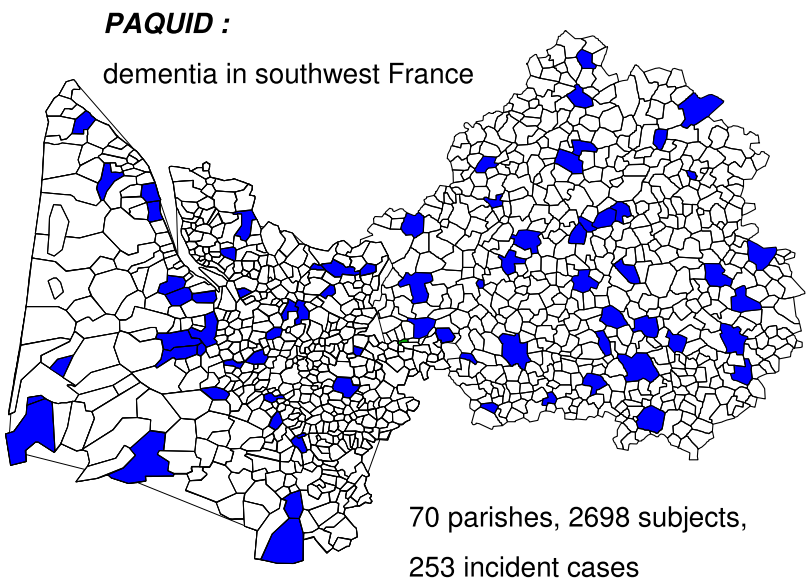
- R : frailtypack / via penalized likelihood or parametrical

```
frailtyPenal(Surv(time,status)~age+sexe+
    cluster(gpe),
n.knots=12,kappa=150000,data=dataR)
```

Example using left truncated data : the Paquid cohort

Aim : study the relationship between aluminum in tap water and Alzheimer disease

- Paquid cohort with 3777 persons from South west France
- Subjects 65 years and older, followed every 2 years
- Randomly selected from 75 geographical areas



Example using left truncated data : the Paquid cohort

- Right censored and left truncated data (age chosen as the basic timescale)

```
    frailtyPenal(Surv(age0,age_dem,status)~sexe+
       educational+silica+alu+cluster(area),
n.knots=12,kappa=150000,data=Paquid)
```

RESULTS of the Paquid cohort

Variable	Cox model		Frailty model	
	Partial likelihood		Penalized likelihood	
	RR	β(SE)	RR	β(SE)
No adjustment	-	-	-	-
			$\theta(SE) = 0.071(0.053)$	
Aluminium*	2.12	0.751(0.254)	2.19	0.783(0.267)
			$\theta(SE) = 0.036(0.038)$	

* adjustment for age (non parametrically), sexe, silica, educational level

Prediction using a gamma shared frailty model

Example : For clustered patients, assess the individual probability of dying for a patient j on the prediction window $[s, s+h]$ given the prognostic factors before s (i.e., given the history of the patient) :

$$P_{ij}(s, s+h|\xi, X_{ij}) = Pr(T_{ij} \leq s+h | T_{ij} \geq s, X_{ij}, \xi)$$

Prediction using a gamma shared frailty model

Two possible approaches :

- First : Conditional prediction (prediction for a given cluster)

$$P^{cond}(T_{ij} \leq s+h | T_{ij} \geq s, X_{ij}, z_i, \xi) = \frac{S_{ij}(s|X_{ij}, z_i, \xi) - S_{ij}(s+h|X_{ij}, z_i, \xi)}{S_{ij}(s|X_{ij}, z_i, \xi)}$$

estimation of the random effects z_i for each subject (using the posterior distribution)

$E_{post}(z_i) = \frac{m_i + 1/\hat{\theta}}{\sum_{j=1}^{n_i} \hat{\Lambda}_{ij}(t) + 1/\hat{\theta}}$ and $var_{post}(z_i) = \frac{m_i + 1/\hat{\theta}}{(\sum_{j=1}^{n_i} \hat{\Lambda}_{ij}(t) + 1/\hat{\theta})^2}$

with m_i the number of events in the cluster i

$\hookrightarrow$ less easy to use in practice (for a specific cluster belonging to the development sample).

Prediction using a gamma shared frailty model

- Second : marginal predictive probability
  averaged prediction over the population

$$P^{marg}(s, s+h; \xi) = \frac{\int_0^\infty (S_{ij}(s|X_{ij}, z) - S_{ij}(s+h|X_{ij}, z)) \times g(z)dz}{\int_0^\infty S_{ij}(s|X_{ij}, z) \times g(z)dz}$$

Inserm
Institut national
de la santé et de la recherche médicale

Example 1 :
Rehospitalisations in colorectal cancers

Example 1 : rehospitalisations in colorectal cancers

★ "hospital readmission" *(Gonzalez, J Epi Com Health 2005)*

★ prospective cohort study (Barcelona, Spain)

★ hospital readmission among patients diagnosed with colorectal cancer and after surgical procedure

★ 403 patients diagnosed 1996-1998, followed until 2002

★ 861 observations : 458 rehospitalisations = recurrent events

★ adjustment for : sexe, Dukes's stage, comorbidity (Charlson score), chemiotherapy (yes/no)

Example 1 : rehospitalisations in colorectal cancers

Questions :

- intra subject correlation of the rehospitalisation times ?
- prognostic factors linked to the risk of rehospitalisation ?

Example 1 : rehospitalisations in colorectal cancers, DATA

id	t.start	t.stop	time	event	chemo	sex	dukes	charlson
1	0	24	24	1	Treat	F	D	3
1	24	457	433	1	Treat	F	D	0
1	457	1037	580	0	Treat	F	D	0
2	0	489	489	1	NonT	M	C	0
2	489	1182	693	0	NonT	M	C	0
3	0	15	15	1	NonT	M	C	3
3	15	783	768	0	NonT	M	C	3
4	0	163	163	1	Treat	F	A-B	0
4	163	288	125	1	Treat	F	A-B	0
4	288	638	350	1	Treat	F	A-B	0

Example 1 : rehospitalisations

```
#  R package FRAILTYPACK
  library(frailtypack)
# dataset
  data(readmission)

# Standard shared frailty model: GAP time
 fit.gap<-frailtyPenal(Surv(time,event)~
  as.factor(dukes)+cluster(id)+strata(sex),
  n.knots=10,kappa=c(10000,10000),data=readmission
      )
```

Output 1 : standard gamma frailty model

```
 > fit.gap
Call:
frailtyPenal(formula = Surv(time, event) ~ as.factor(dukes) +
    cluster(id) + strata(sex), data = readmission, n.knots = 10,
    kappa = c(10000, 10000))


  Shared Gamma Frailty model parameter estimates
  using a Penalized Likelihood on the hazard function
  (Stratification structure used) : 2 strata

           coef exp(coef) SE coef (H) SE coef (HIH)       z          p
dukesC 0.445073   1.56060    0.149604      0.149604 2.97500 2.9299e-03
dukesD 1.288954   3.62899    0.182517      0.182517 7.06209 1.6401e-12

        chisq df global p
dukes 49.8775  2 1.48e-11

    Frailty parameter, Theta: 0.69748 (SE (H): 0.14535 ) p = 7.9879e-07

      penalized marginal log-likelihood = -3242.99
      Convergence criteria:
      parameters = 0.000658 likelihood = 5.87e-05 gradient = 2.99e-07

      LCV = the approximate likelihood cross-validation criterion
            in the semi parametrical case     = 3.7979

      n= 861
      n events= 458  n groups= 403
      number of iterations:  20
      Exact number of knots used:  10
      Value of the smoothing parameter:  10000 10000, DoF:  11.24
```

Example 1 : rehospitalisations

```
# GAP time, gaussian frailty
 fit.gap.g<-frailtyPenal(Surv(time,event)~
  as.factor(dukes)+cluster(id)+strata(sex),
  n.knots=10,kappa=c(10000,10000),data=readmission
      , RandDist='LogN')
```

Output 2 : standard Log-normal frailty model

```
> fit.gap.g
Call:
frailtyPenal(formula = Surv(time, event) ~ as.factor(dukes) +
    cluster(id) + strata(sex), data = readmission, n.knots = 10,
    kappa = c(10000, 10000), RandDist = 'LogN')


  Shared Log-Normal Frailty model parameter estimates
  using a Penalized Likelihood on the hazard function
  (Stratification structure used) : 2 strata

           coef exp(coef) SE coef (H) SE coef (HIH)       z          p
dukesC 0.425704   1.53067    0.152739      0.152739 2.78713 5.3177e-03
dukesD 1.308382   3.70018    0.183047      0.183047 7.14779 8.8185e-13

        chisq df global p
dukes 51.2757  2 7.34e-12

    Frailty parameter, Sigma Square: 0.612921 (SE (H): 0.122791 ) p = 2.9946e-07

      penalized marginal log-likelihood = -3238.49
      Convergence criteria:
      parameters = 8.02e-05 likelihood = 0.000601 gradient = 7.18e-07

      LCV = the approximate likelihood cross-validation criterion
            in the semi parametrical case     = 3.79268

      n= 861
      n events= 458  n groups= 403
      number of iterations:  20
      Exact number of knots used:  10
      Value of the smoothing parameter:  10000 10000, DoF:  11.24
```

Example 1 : rehospitalisations

```
# Calendar time, parametrical
fit.cal <- frailtyPenal ( Surv (t.start ,t.stop ,
    event )~
   as.factor(dukes)+cluster(id)+strata(sex),
      recurrentAG=T,
   n.knots=10,kappa=c(10000,10000),data=readmission
      )
```

Output 3 : with calendar time scale

```
> fit.cal
Call:
frailtyPenal(formula = Surv(t.start, t.stop, event) ~ as.factor(dukes) +
    cluster(id) + strata(sex), data = readmission, recurrentAG = T,
    n.knots = 10, kappa = c(10000, 10000))

      Calendar timescale

  Shared Gamma Frailty model parameter estimates
  using a Penalized Likelihood on the hazard function
  (Stratification structure used) : 2 strata

           coef exp(coef) SE coef (H) SE coef (HIH)       z          p
dukesC 0.490344   1.63288    0.176245      0.176245 2.78217 5.3996e-03
dukesD 1.612233   5.01400    0.217164      0.217164 7.42404 1.1358e-13

        chisq df global p
dukes 55.2113  2 1.03e-12

    Frailty parameter, Theta: 1.31463 (SE (H): 0.19494 ) p = 7.7169e-12

      penalized marginal log-likelihood = -3311.8
      Convergence criteria:
      parameters = 0.000645 likelihood = 0.000279 gradient = 5.64e-06

      LCV = the approximate likelihood cross-validation criterion
            in the semi parametrical case     = 3.87782

      n= 861
      n events= 458  n groups= 403
      number of iterations:  19
      Exact number of knots used:  10
      Value of the smoothing parameter:  10000 10000, DoF:  12.00
```

The approximate likelihood cross-validation criterion (LCV)

- measures the relative goodness of fit among a collection of models
- lower values indicate a better fitting
- $$LCV_a = \frac{1}{n}(trace(H_{pl}^{-1} H_l) - l(.))$$
  - H_{pl} minus the converged hessian of the penalized log-likelihood,
  - H_l minus the converged hessian of the log-likelihood
  - $l(.)$ is the full log-likelihood
- If parametrical approach $trace(H_{pl}^{-1} H_l) =$ number of parameters
  $LCV = \frac{1}{n}(np - l(.)) \sim$ AIC criterion

(O'Sullivan, J Sci Stat Comp 1988 ; Commenges, Scnad J Stat 2007)

Prediction of the frailties

AIM :

- Individual predictions of frailties
- Identify graphically outliers

Prediction of the frailties

Bayesian approach : *posterior* dist of the random effects

$$f_{z_i}(z|\tilde{T}_i,\beta,\alpha_0(.),\theta) = \frac{\overbrace{f(\tilde{T}_i|z_i,\beta,\alpha_0(.),\theta)}^{\text{conditional like.}} \overbrace{f_{z_i}(z|\theta)}^{\text{a priori distr.}}}{\underbrace{f(\tilde{T}_i,\beta,\alpha_0(.),\theta)}_{\text{marginal like.}}}$$

In the case of prior gamma frailties :

$f_z(z|\theta) = \frac{z^{1/\theta-1}\exp(-z/\theta)}{\Gamma(1/\theta)\theta^{1/\theta}} \sim \Gamma(1/\theta;1/\theta)$

we then obtain a gamma dist. a posteriori

$f(\tilde{T}_i|z_i,\beta,\alpha_0(.),\theta) \sim \Gamma(m_i+1/\theta;\sum_{j=1}^{n_i}\Lambda_{ij}(t)+1/\theta))$

a posteriori mean (replacing θ et $\Lambda(.)$ by their estimators) :

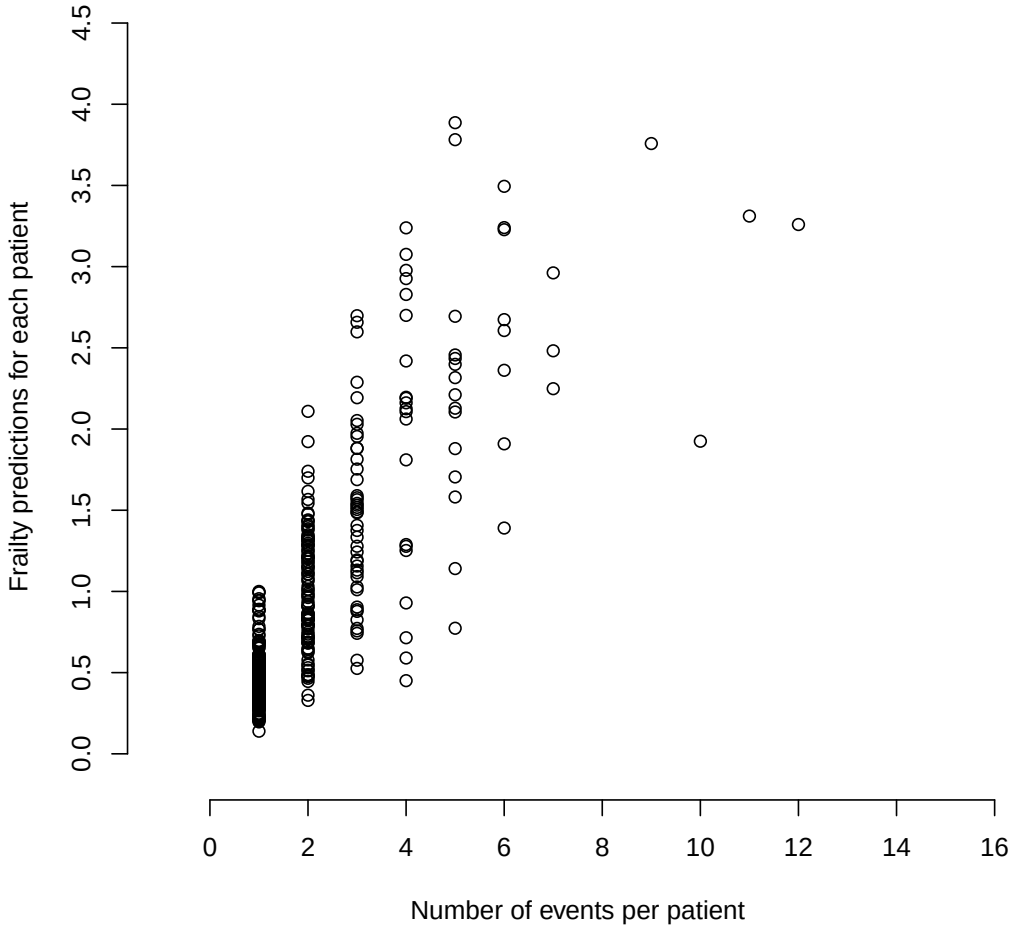
$$\mathrm{E}_{post}(z_i) = \frac{m_i+1/\theta}{\sum_{j=1}^{n_i}\Lambda_{ij}(t)+1/\theta}$$

Example 1 : rehospitalisations

Prediction of the random effects

```
nb.events<-as.vector(table(readmission$id))
frailtypred <- fit.cal$frailty.pred

plot(nb.events,frailtypred,xlab='Number of events
    per patient'
,ylab='Frailty predictions for each patient', type
    ='p',
axes=F,
pch=1,ylim=c(-0.1,4.5),xlim=c(-1,16))
axis(1,round(seq(0,16,length=9),digit=0))
axis(2,round(seq(0,4.5,length=10),digit=1))
```

Individual frailty predictions according to the number of rehospitalisations



The higher the number of infections, the higher the frailty is

Interval-censored data in shared frailty models

Epidemiological motivation :

Prospective cohort : the Three-city study (3C)

- population-based study of 1296 couples followed (2592 subjects)
- 65 years and over
- clinical examinations every 2 years during 10 years
- Age : important factor

Interval-censored data in shared frailty models

Epidemiological motivation :

- Analyze of clustered and interval-censored outcomes
  Greater risk of dementia in couple when spouse has dementia ?
  - Couples are natural clusters, does it exist an intra-couple correlation ?
  - common habits, common diet, common environmental factors
  - a chronic and severe stress, as a specific event may also explain an intra-couple correlation
  - Results from *Norton et al. 2010 : "a subject whose spouse experienced incident dementia onset had a six times greater risk for incident dementia as subjects whose spouses were dementia free (RR=6.0, 95 % CI=2.2-16.2)"*

Incomplete data

- right censored (lost to follow-up, end of follow-up, death)
- Age chosen as the basic timescale : left truncation
- The interval-censored data when subjects are followed periodically for the event of interest
  $\rightarrow$ a continuous-time model for the biological system, but a discrete-time observation scheme
  $\rightarrow$ the event time T is not directly observable but may be detected in some periodic examination interval, denoted as $[L, R]$ where L is the left examination time and R is the right examination time

Sun et al. 2006, Springer

Interval-censored data in shared frailty models

Notation

subject j	$j = 1, ..., N_i$
from group i	$i = 1, ..., G$
$T_{ij} = min(X_{ij}, C_{ij}, D_{ij})$	follow-up times
$R_i = \{j \mid T_{ij} > min(C_{ij}, D_{ij})\}$	index for right censored
$D_i = \{j \mid L_{ij} < T_{ij} \leq U_{ij}\}$	index for interval-censored

Conditional contribution for the marginal log likelihood

$$V_i(\mathbf{T_i}|\omega_i) = \prod_{j \in R_i} S_{ij}(C_{ij}|\omega_i) \prod_{j \in D_i} \{S_{ij}(L_{ij}|\omega_i) - S_{ij}(U_{ij}|\omega_i)\}$$

Interval-censored data in shared frailty models

- R : frailtypack

```
shared.ic <-frailtyPenal(SurvIC(age0,ageL,ageU,
    dem0_5)~
cluster(couple)+sexe+dipniv02+as.factor(centre)
+depres0 + as.factor(diab) + apoE4
+ as.factor(bmi) + atcdavc,
data=data_shared,n.knots=8,kappa=817)
```

Results : risk of dementia

	Cox Proport. hazard model		Gamma Frailty models			
			without interval-censoring**		with interval-censoring	
No adjustment			$\theta = 0.15$ (0.33)		$\theta = 0.19$ (0.31)	
Covariates*	RR	(95% CI)	RR	(95% CI)	RR	(95% CI)
Sex (men vs women)	**1.34**	**(1.01-1.80)**	**1.35**	**(1.01-1.81)**	**1.34**	**(1.00-1.80)**
Educ level (vs primary)						
primary school	1		1		1	
secondary school	0.70	(0.44-1.11)	0.72	(0.44-1.16)	0.73	(0.44-1.20)
high educational	0.57	(0.36-0.93)	0.59	(0.35-0.99)	0.60	(0.36-1.02)
Depressive symptoms (yes vs no)	**1.55**	**(1.02-2.35)**	**1.59**	**(1.02-2.48)**	**1.60**	**(1.02-2.51)**
Diabetes Status						
Without diabetes	1		1		1	
hyperglycemia	0.74	(0.32-1.70)	0.68	(0.28-1.65)	0.67	(0.27-1.64)
diabetes	**1.83**	**(1.24-2.70)**	**1.85**	**(1.23-2.78)**	**1.86**	**(1.23-2.82)**
APOE4	**2.19**	**(1.63-2.96)**	**2.28**	**(1.64-3.17)**	**2.32**	**(1.67-3.24)**
BMI						
< 21	**1.71**	**(1.08-2.69)**	**1.79**	**(1.09-2.93)**	**1.81**	**(1.10-2.99)**
$[21 - 27[$	1		1		1	
$[27 - 30[$	1.31	(0.92-1.86)	1.32	(0.92-1.90)	1.33	(0.92-1.93)
≥ 30	**1.74**	**(1.17-2.60)**	**1.79**	**(1.17-2.73)**	**1.81**	**(1.18-2.78)**
History of stroke	**1.90**	**(1.02-3.53)**	**2.11**	**(1.03-4.34)**	**2.14**	**(1.03-4.44)**
Frailty var (SE)			$\theta = 0.18$ (0.26)		$\theta = 0.23$ (0.27)	
LCV***	0.398		0.400		0.326	

* adjusted for center, ** at midpoint of the interval
*** approximate Cross-validation criterion

PART 2 :
Extension of Standard frailty models

- Nested frailty models
- Additive frailty models

Nested frailty models

- Example :
  - modelling the effect of air pollution on mortality : two levels of regrouping (city and geographical area)
  - recurrent infection times of patients from different hospitals

(Sastry, JASA 1997 ; Manda, Aus and NZ J of stat 2001 ; Rondeau, stat med 2006)

Nested frailty models

- Model :

$$\lambda_{\mathbf{ijk}}(\mathbf{t}|\mathbf{v_i}, \mathbf{w_{ij}}) = \mathbf{v_i w_{ij}} \lambda_{\mathbf{0}}(\mathbf{t}) \exp(\beta' \mathbf{X_{ijk}})$$

with,
$i = 1, G$ (cluster - ex : city)
$j = 1, J_i$ (sub cluster - ex : family)
$k = 1, K_{ij}$ (subjects - ex : members of each family j)
v_i iid gamma, $\mathrm{E}(v_i) = 1$ et $\mathrm{var}(v_i) = \alpha$
w_{ij} iid gamma, $\mathrm{E}(w_{ij}) = 1$ et $\mathrm{var}(w_{ij}) = \eta$
v_i et w_{ij} iid
$Y_{ijk} = \min(T_{ijk}, C_{ijk})$ observations

Nested frailty models

- Estimation of the parameters : $\xi = (\beta, \alpha, \eta, \lambda_0(t))$

$\Rightarrow$ full log-likelihood for left-truncated and right-censored data :

$$l(\lambda_0(.), \beta, \alpha, \eta) =$$

$$\sum_{i=1}^{G} \left\{ \sum_{j=1}^{J_i} \sum_{k=1}^{K_{ij}} \delta_{ijk} \{ \beta' X_{ijk} + \ln(\lambda_0(t_{ijk})) \} \right.$$

$$+ \sum_{j=1}^{J_i} \left[I_{\{m_i > 1\}} \sum_{k=1}^{m_{ij}} \ln(1 + \eta(m_{ij} - k)) \right]$$

$$+ \ln \int \frac{v_i^{(1/\alpha - 1 + m_i)} \exp(-v_i/\alpha)}{\prod_j (\eta v_i \sum_k \Lambda_{ijk}(t) + 1)^{(1/\eta + m_{ij})}} \partial v_i$$

$$\left. - \ln \int \frac{v_i^{(1/\alpha - 1)} \exp(-v_i/\alpha)}{\prod_i (\eta v_i \sum_k \Lambda_{ijk}(\mathcal{L}) + 1)^{(1/\eta)}} \partial v_i \right\}$$

Inserm
Institut national
de la santé et de la recherche médicale

Example 2 : PAARC study

Air pollution and cardiopulmonary mortality

Example 2 : PAARC study

Aim : Analyze the long-term effect of air pollution on cardiopulmonary mortality taking into account the clustering

★ 11 504 subjets initialy aged 25-59 years, randomly selected on electoral lists between 1974-1976

★ pollution measurements, mean between 1974 and 1976 : total suspended particles (TSP), black smoke, SO2, NO2, NO at a centrally located pollution monitoring station

★ mortality in 2000-2001 using the national registry and the department SC8 of INSERM

★ in 24 areas of seven french cities with different air pollution

★ 105 to 553 subjects in each area

★ Data right-censored and left truncated (age as teh basic time-scale)

Example 2 : PAARC study

Questions :

$\rightarrowtail$ intra-city correlation ?

$\rightarrowtail$ intra-area correlation ?

$\rightarrowtail$ influence of correlation on air pollution (area-specific variables) ?

Example 2 : Air pollution and cardiopulmonary mortality

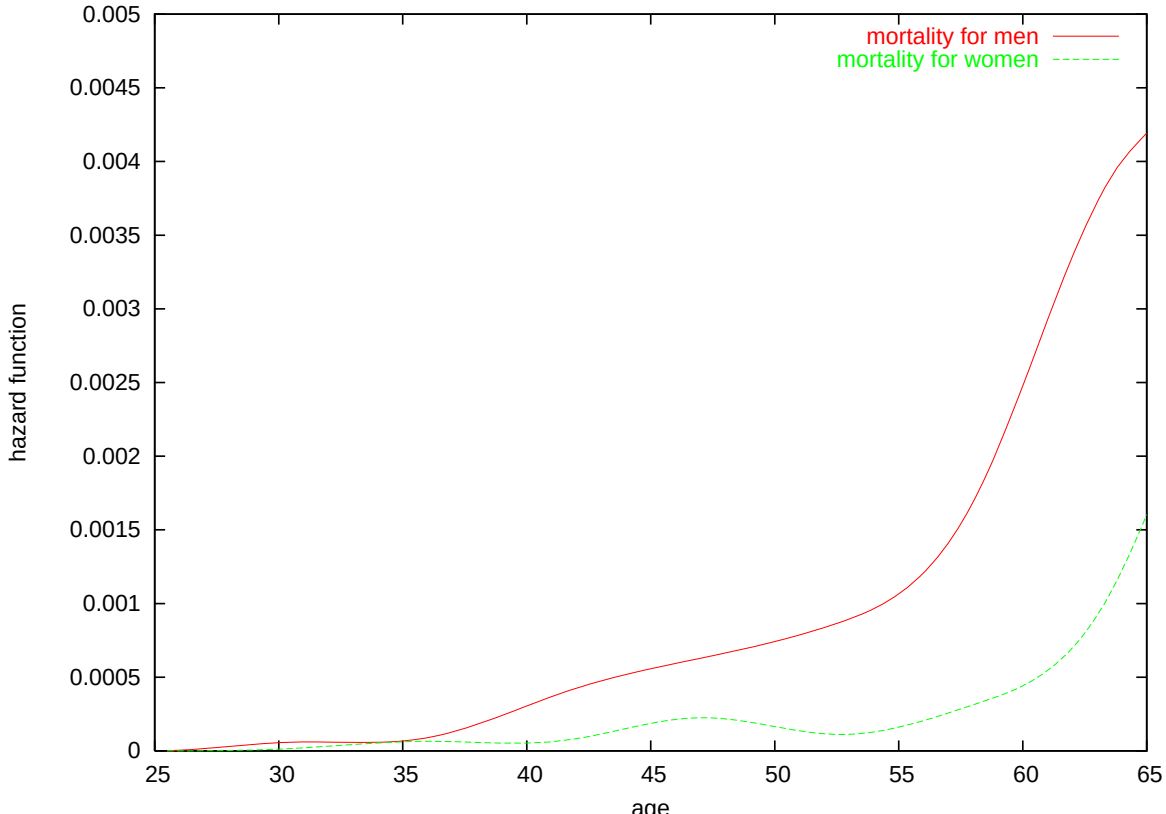
```
fittsp.nest<-frailtyPenal(Surv(age,agedc10,cens401
    ==0)~
tsp+exf+pf+mf+gf+nivetu2+nivetu3+sexe+ subcluster(
    zonrec)+cluster(ville)
,n.knots=8,kappa=1000,data=paarcMCP10)
```

RESULTS : Effect of air pollution on cardiopulmonary mortality

	$\hat{\beta}$**(S.E.*)	$\hat{\beta}$**(S.E.*)	$\hat{\beta}$**(S.E.*)	$\hat{\beta}$**(S.E.*)	$\hat{\beta}$**(S.E.*)
	Model I : *No frailty*	Model II : *City-level*	Model III : *Area-level*	Model IV : *Two-level*	Model V : *Area-level*
Smoking vs non					
-Former	0.47 (0.36)	0.50 (0.36)	0.50 (0.36)	0.51 (0.36)	0.51 (0.37)
-Smok($\leq$ 9)	**0.74 (0.35)**	**0.76 (0.35)**	**0.76 (0.35)**	**0.77 (0.35)**	**0.78 (0.35)**
-Smok[10-19]	**0.72 (0.35)**	**0.75 (0.35)**	**0.76 (0.35)**	**0.76 (0.35)**	**0.77 (0.35)**
-Smok($\geq$ 20)	**1.32 (0.30)**	**1.33 (0.30)**	**1.34 (0.30)**	**1.34 (0.29)**	**1.36 (0.30)**
Educ level (vs univers)					
-Secondary	**-0.56 (0.26)**	**-0.55 (0.26)**	**-0.54 (0.26)**	**-0.54 (0.26)**	**-0.57 (0.26)**
-Primary	0.35 (0.24)	0.35 (0.24)	0.34 (0.24)	0.34 (0.24)	0.35 (0.24)
Sex (F vs M)	**-1.40 (0.30)**	**-1.38 (0.30)**	**-1.39 (0.30)**	**-1.38 (0.30)**	**-1.38 (0.31)**
TSP♣	0.021 (0.048)	0.043 (0.059)	0.027 (0.059)	0.043 (0.065)	0.085 (0.079)
City (fixed) (vs city 6)					
-city 1	-	-	-	-	0.27 (0.35)
-city 2	-	-	-	-	-0.19 (0.52)
-city 3	-	-	-	-	-0.55 (0.42)
-city 4	-	-	-	-	-0.08 (0.40)
-city 5	-	-	-	-	-0.51 (0.51)
Variance of frailties					
cities	-	0.019 (0.026)	-	0.018 (0.030)	-
areas	-	-	0.046 (0.065)	0.024 (0.066)	0.002 (0.05)

Age-specific mortality (using stratification)

```
    fit.nested<-frailtyPenal(Surv(age,agedc10,cens401==0)~
tsp+exf+pf+mf+gf+nivetu2+nivetu3+strata(sexe)+
subcluster(zonrec)+cluster(ville)
,n.knots=8,kappa=1000,data=paarcMCP10)

    plot(fit.nested, conf.band=F)
```



PAARC : Conclusion

- increase risk of mortality with a 10 μg /m^3 change in TSP, but non significant
- non significant intra-city nor intra-zone correlation,
- but, better estimation (standard error) using random effects models
- the nested frailty model separates the two levels of correlation
- drawbacks when using fixed effects :
  - at least one event per cluster
  - the sample size increases with the number of parameters
  - numerical issues (with high number of groups)

Inserm
Institut national
de la santé et de la recherche médicale

Example 3 : Chronic Granulomatous Disease, *Fleming, 1991*

recurrent infections in different hospitals

Example 3 : Chronic Granulomatous Disease

Placebo-controlled randomized trial of gamma-interferon (γ-IFN) in CGD

AIM : investigate the effectiveness of γ-IFN on serious infections in CGD patients.

★ 13 hospitals

★ 128 patients (63 in the treated group and65 in the placebo) followed during 1 year

★ 203 observations among them 76 infections (20 in the treated, 56 in the placebo)

★ between 1 and 8 infections per patient

Example 3 : CGD study

Questions :

$\rightarrowtail$ recurrent infection : intra-patient correlation ?

$\rightarrowtail$ clustered data : intra-hospital correlation ?

Example 3 : CGD DATA (calendar timescale)

t0	t1	ic	hospital	patient	ttt
0	293.0000	0	1	1	1
0	255.0000	0	1	2	2
0	213.0000	0	1	3	2
0	203.0000	0	1	4	2
0	219.0000	1	2	5	1
220	373.0000	1	2	5	1
374	414.0000	0	2	5	1
0	8.000000	1	2	6	2
9	26.00000	1	2	6	2
27	152.0000	1	2	6	2
153	241.0000	1	2	6	2
		...			

Example 3 : CGD study

```
fitnest.cgd<-frailtyPenal(Surv(tstart,tstop,status
    )~
treat + subcluster(id)+cluster(center),recurrentAG
    =T
,n.knots=12,kappa=1000,data=cgd)

fitnest.cgd
```

RESULTS :CGD study

	Shared Frailty model *Hospital level only*	Shared Frailty model *Patient level only*	Nested Frailty model
	β (S.E.*)	β (S.E.*)	β (S.E.*)
With gap timescale			
Treatment (γ-IFN)	-1.11 (0.27)	-1.14 (0.35)	-1.08 (0.34)
α (hospital)	0.15 (0.14)	-	0.008 (9.10-5)
η (patient)	-	1.56 (0.68)	1.47 (0.64)
Penalized log-likelihood	-357.62	-350.74	-341.10
With calendar timescale			
Treatment (γ-IFN)	-1.10 (0.26)	-1.04 (0.31)	-1.02 (0.31)
α (hospital)	0.12 (0.13)	-	0.008 (9.10-5)
η (patient)	-	0.83 (0.40)	0.79 (0.39)
Penalized log-likelihood	-352.02	-347.78	-338.09

Additive frailty models

Example : Meta-analysis on an individual patient data
combine results from different randomized trials
$\hookrightarrow$ Two main sources of intertrial heterogeneity in survival data :

- **heterogeneity of the baseline risk**
  due to differences in trial design, in treatment protocols, medical practices or in patient populations
- **heterogeneity of treatment effects across trials**
  reflects differences in patient characteristics and in implementation of the protocol

(Legrand, Stat Med 2005 ; Rondeau, stat med 2008)

Motivation : Meta-Analysis of Chemotherapy in Head and Neck Cancers (MACH-NC)

in a large meta-analysis of randomized trials (n=87) in patients with head and neck cancers

- Study heterogeneity of death between trials
- Study the benefit of adding chemotherapy to locoregional treatment

(Pignon, Lancet 2000 ; Pignon, IJROBP 2007)

Aim

propose a general additive random effects model and an associated estimation method to study :

- a random trial effect
- a random treatment by trial interaction

Correlated additive random effects Cox model

- G independent clusters (ex : trials) $i = 1, ..., G$
- n_i subjects in each cluster $j = 1, ..., n_i$
- T_{ij} = survival times and C_{ij} = censoring times
  $Y_{ij} = \min(T_{ij}, C_{ij})$ observed times
- X_{ij1} = treatment arm

Hazard for the *jth* patient in the *ith* trial :

$$\lambda_{ij}(t|u_i, v_i, \mathbf{X_{ij}}) = \lambda_0(t) \exp(u_i + v_i X_{ij1} + \sum_{k=1}^{p} \beta_k X_{ijk})$$

u_i et v_i random effects for trial i

$$u_i \sim \mathcal{N}(0, \sigma^2), \qquad v_i \sim \mathcal{N}(0, \tau^2), \qquad cov(u_i, v_i) = \rho\sigma\tau$$

Correlated additive random effects Cox model

u_i and v_i = random effects for trial i

$$u_i \sim \mathcal{N}(0, \sigma^2), \qquad v_i \sim \mathcal{N}(0, \tau^2), \qquad cov(u_i, v_i) = \rho\sigma\tau$$

$\rightarrow \sigma^2$ = heterogeneity between trials of the overall underlying baseline risk

$\rightarrow \tau^2$ = heterogeneity between trials of the overall treatment effect

Full marginal log-likelihood :

$$l(\theta) = \ln \prod_{i=1}^{G} \int \int_{\Re} \left[\prod_{j=1}^{n_i} \lambda(T_{ij}|u,v,X_{ij})^{\delta_{ij}} S(T_{ij}|u,v,X_{ij}) \right] f(u,v) du dv$$

$$= \sum_{i=1}^{G} \ln \int \int_{\Re} \exp\{-K_i(u_i,v_i)\} du_i dv_i$$

no analytical solutions of the integrations

⇢ first-order Laplace approximation *(Breslow,1993)*

Penalized log-likelihood :

$$pl(\lambda_0(.),\beta,\sigma,\tau,\rho) = l(\lambda_0(.),\beta,\sigma,\tau,\rho) - \kappa \int_0^{\infty} \lambda_0''^2(t) dt \qquad (1)$$

to expect a smooth baseline hazard $\hat{\lambda}_0(t)$

⇢ approximation on a basis of splines

Inserm
Institut national
de la santé et de la recherche médicale

Example 4 : MACHNC study

Meta-analysis of Chemotherapy for Head and Neck Carcinoma

Example 4 : Meta-analysis of Chemotherapy for Head and Neck Carcinoma

★ Aerodigestive tract (oral cavity, oropharynx, hypopharynx, nasopharynx, larynx) are frequent tumors : 550 000 new cases other the world in 2000

★ standard treatment (without metastasis) = radiotherapy and/or surgery

★ MACH-NC : meta-analysis on individual data, including between 1965 and 2000, 87 randomized trials (101 clusters), and 16360 patients analyzed

Aim : study benefit of adding chemotherapy to locoregional treatment in overall survival of head and neck patients

Example 4 : meta-analysis MACH-NC, DATA

time	status	trial	patkey	sex	chemo	age5160	age60	stage3	stage4	l
11.885	0	1	1	0	1	0	0	0	1	
4.591	1	1	2	1	1	0	1	1	0	
3.236	1	1	3	1	1	0	1	0	1	
6.779	1	1	4	0	0	0	0	0	1	
0.281	1	1	5	0	0	0	1	0	1	
3.260	1	1	6	0	1	0	1	1	0	
4.164	1	1	7	0	0	0	1	1	0	
1.193	1	1	8	0	0	0	0	0	1	
11.143	0	1	9	0	1	0	0	0	1	
11.479	0	1	10	1	0	0	0	0	1	
7.649	1	1	11	0	1	1	0	0	1	
0.369	1	1	12	1	0	0	1	0	1	
7.479	1	1	13	0	0	0	1	0	1	

Analysis of MACH-NC

- investigate the proposed additive random effects
  - treatment (chemotherapy or not) as a fixed effect
  - simultaneously with random treatment-by-trial interactions and random trial effects

$$\lambda_{ij}(t|.) = \lambda_0(t)\exp(u_i + v_i CHEMO + \beta_1 CHEMO + \beta' X)$$

- number of patients per trial varied between 24 and 676 (mean 162)
- a total of 10980 patients (67.1%) died and the number of deaths over trials ranged from 11 to 506 (mean 109)

Analysis of MACH-NC

```
fit.additive<-additivePenal(Surv(time,event)~
    cluster(trial)+
chemo+sex+age5160+age60+stage3+stage4+larynx+slope
    (chemo)
,correlation=TRUE,data=MACHNC,n.knots=8,kappa=200)
```

Results : MACH-NC, 1965-2000 (n=16360, G=101)

	Zero covariance ($\text{cov}(u_i, v_i) = 0$)	Non-zero-covariance ($\text{cov}(u_i, v_i) \neq 0$)
	RR (CI)	RR (CI)
Chemotherapy treated (1) vs control (0)	0.88 (0.83-0.92)	0.88 (0.83-0.93)
σ^2	0.152 (0.026)	0.167 (0.031)
τ^2	0.023 (0.009)	0.029 (0.012)
$cov(u_i, v_i)$	–	-0.018 (0.016)
ρ	–	(ρ =-0.26)
Marginal penalized Log-Likelihood	-24607.06	-24612.28

* Adjusted for Sex, Age, Stage, Site of the tumor

Results : MACH-NC, 1965-2000 (n=16360, G=101)

- adjustment for the period of randomization using three periods (1965-1980 ; 1981-1994 and after 1994)
  $\rightarrow$ no significant higher risk of death for any period of randomization
- 3 separate analyses according to the timing of chemotherapy : adjuvant, neoadjuvant, or concomitant
  $\rightarrow$ significant efficacy of only the concomitant chemotherapy (given concomitantly or alternating with radiotherapy)

Conclusion : correlated frailty models

- Additive random effects model are useful to study :
- heterogeneity across trials of the baseline hazard
- heterogeneity across trials of the treatment effect

- Falsely coercing the covariance parameter between the two random effects to 0 could lead to inadequate results

- Advantages of the maximum penalized likelihood estimation associated with Laplace approximation for estimation

- Easy implementation with R Frailtypack

- Useful in a meta-analysis of clinical trials but also in multi-center clinical trials (with sufficient sample sizes)

Inserm
Institut national
de la santé et de la recherche médicale

PART 3 :
Joint frailty models for recurrent events and terminal event

- Joint frailty models
- Prediction using joint frailty models

Joint Models

- Recurrent events and death **processes** are potentially **correlated**
- Example : Breast cancer relapses and death
- Standard (naive) approach of Cox with time-dependent covariate only for **external covariates** !
- Interest :
  - investigating the **strength of association** between recurrent events and death
  - allows to study impact of **covariates both** on recurrent events and death
  - treat **informative censoring** by death

Joint models : some notations

- D_i time of death for subject i, $i = 1, ..., n$
- X_{ij} time of the *jth* recurrence for subject i
- Z_{ij}^R and Z_i^D covariates vectors for recurrence and death
- λ_{ij}^R and λ_i^D baseline hazards for risk of recurrence or death

Joint models

Joint modeling for the risk of recurrent event (disease relapses) and terminal event (death)

$$\begin{cases} \lambda_{ij}^R(t|u_i) = u_i \lambda_0^R(t) \exp(\beta_1' Z_{ij}^R) \\ \lambda_i^D(t|u_i) = u_i^\alpha \lambda_0^D(t) \exp(\beta_2' Z_i^D) \end{cases}$$

- calendar timescale (time from origin) or gap timescale
- $u_i \sim \Gamma(1/\theta; 1/\theta)$, i.e. $E(u_i) = 1$ and $var(u_i) = \theta$
- θ dependency between recurrent events and death
- α sense and strength of the association (more flexibility)

(Liu et al. Biometrics 2004 ; Rondeau et al. Biostatistics 2007)

Inference in the joint model

Marginal log-likelihood

$$l(\phi) = \sum_i \left\{ \sum_j \delta_{ij} \log r_i(T_{ij}) + \delta_i^* \log \lambda_i(T_i^*) - \log \Gamma(1/\theta) - \frac{1}{\theta} \log \theta \right.$$
$$+ \log \int_0^\infty \omega^{(N_i^R(T_i^*) + \alpha\delta_i^* + 1/\theta - 1)}$$
$$\left. \exp\left(-\omega \int_0^{T_i^*} dR_i(t) - \omega^\alpha \int_0^{T_i^*} d\Lambda_i(t) - \frac{\omega}{\theta} \right) d\omega \right\}$$

Estimation of the parameters :

- Using penalized likelihood *(Rondeau, Biostat 2007)*
- using the EM algorithm *(Liu, Biometrics 2004)*
- ...

Inserm
Institut national
de la santé et de la recherche médicale

Example 5 of recurrent events :

Morbidity and Health care Resources utilization in HIV-infected Children

Example of recurrent events (Desmonde, JAIDS, 2014)

IMPLEMENTATION AND OPERATIONAL RESEARCH: EPIDEMIOLOGY AND PREVENTION

Morbidity and Health care Resource Utilization in HIV-Infected Children After Antiretroviral Therapy Initiation in Côte d'Ivoire, 2004–2009

Sophie Desmonde, MSc,† Jean-Bosco Essanin, MD,‡ Addi E. Aka, MD,§ Eugène Messou, MD, PhD,§ Madeleine Amorissani-Folquet, MD, PhD,‡‖ Virginie Rondeau, PhD,*† Andrea Ciaranello, MD, MPH,¶ and Valériane Leroy, MD, PhD*†*

Background: We describe severe morbidity and health care resource utilization (HCRU) among HIV-infected children on antiretroviral therapy (ART) in Abidjan, Côte d'Ivoire.

Methods: All HIV-infected children enrolled in an HIV-care program (2004–2009) were eligible for ART initiation until database closeout, death, ART interruption, or loss to follow-up. We calculated incidence rates (IRs) of density per 100 child-years (CYs) for severe morbidity, HCRU (outpatient care and inpatient care), and associated factors using frailty models with a Weibull distribution.

Results: Of 332 children with a median age of 5.7 years and median follow-up of 2.5 years, 65.4% were severely immunodeficient by World Health Organ

ratio (aHR): 1.83; 95% CI: 1.35 to 2.47] and to those moderately/severely immunodeficient compared to those not (aHR: 1.57; 95% CI: 1.13 to 2.18 and aHR: 2.53; 95% CI: 1.81 to 3.55, respectively). Of the 464 events, 371 (80%) led to outpatient care (IR: 45.6/100 CYs) and 164 (35%) to inpatient care (IR: 20.2/100 CYs). In adjusted analyses, outpatient care was significantly less frequent in children older than 10 years compared with children younger than 2 years (aHR: 0.49; 95% CI: 0.31 to 0.78) and in those living furthest from clinics compared with those living closest (aHR: 0.65; 95% CI: 0.47 to 0.90). Both inpatient and outpatient HCRU were negatively associated with cotrimoxazole prophylaxis.

Conclusions: Despite ART, HIV-infected children still require

Example of recurrent events (Desmonde, JAIDS, 2014)

★ In Abidjan, Côte d'Ivoire
★ HIV infected children enrolled in an HIV-care program (2004-2009)
★ Children followed from ART initiation until database closeout, death, ART interuption, or loss to follow-up
★ 332 Children, followed-up 2.5 years (median)
★ times to **severe morbidity** (any event classified WHO stage 3 or 4, or any event leading to inpatient day care, hospitalisation or death)
★ times to **health care resource utilization** (HCRU)
★ shared frailty model (Weibull), to study **incidence of recurrent morbidity** and incidence of **recurrent HCRU rates** since ART initiation

Example : recurrent morbidity or HCRU since ART initiation

```
# Joint model / parametrical / morbidity
fit.morb<-frailtyPenal(formula = Surv(time_start,
    time_event, morb) ~
cluster(PAT_ID)+ as.factor(cl_age_art)+ as.factor(
    cmx)+
as.factor(GENDER)+ as.factor(regime_art)+ as.
    factor(second_line)+ as.factor(cd4_grp_1) +
terminal(ppgm),
formula.terminalEvent = ~as.factor(cl_age_art)+ as
    .factor(cmx)+ as.factor(cd4_grp_1),
data = tab, hazard=''Weibull'')
summary(fit.morb)
```

Results for severe morbidity

- significant protective effect of cotrimoxazole prophylaxis (aHR = 0.36)
- children on a PI-based regimen are more likely to develop severe morbidity,
- also for those at more advanced stages of immunodeficiency
- association between the times of severe morbidity and death or loss to follow-up (signif. α and θ)
- different associations with the terminal event

TABLE 3. Determinants of Severe Morbidity Among th HIV-Infected Children on ART, Followed up at the CePF Between 2004 and 2009, Abidjan, Côte d'Ivoire

	aHR	95% CI
For recurrences (severe morbidity)		
Age at ART initiation		
<5 yrs	1	—
≥5 yrs	1.04	0.77 to 1.41
Cotrimoxazole	0.36	0.23 to 0.56
CD4%		
≥25%	1	—
15%–25%	1.57	1.13 to 2.18
<15%	2.53	1.81 to 3.55
Gender: male/female	1.04	0.79 to 1.37
ART regimen		
NNRTI	1	—
PI	1.83	1.35 to 2.47
Second-line treatment	0.68	0.42 to 1.12
For survival		
Age at ART initiation		
<5 yrs	1	—
≥5 yrs	0.85	0.45 to 1.61
Cotrimoxazole	2.52	0.81 to 7.87
CD4%	—	—
≥25%	1	—
15%–25%	2.11	0.76 to 5.87
<15%	8.30	3.18 to 21.65
Variance of random effect (SE)	0.77 (0.11)	
α (SE)*	1.13 (0.33)	

Joint Frailty adjusted model with estimated adjusted hazard ratios (aHRs) and
*When α = 1, the effect of the frailty is identical for the recurrent ev
the terminating event. When α > 1, the recurrent rate and the survival
positively associated

Introduction : prediction

- **After a breast cancer diagnosis**
  $\rightarrow$ single or multiple events
  (recurrences, metastases, death)

Introduction : prediction

- **After a breast cancer diagnosis**
  $\rightarrow$ single or multiple events
  (recurrences, metastases, death)
- **Prediction of death**
  $\rightarrow$ clinical therapeutic decisions, and patient monitoring
  $\rightarrow$ patient information
  $\rightarrow$ trials : defining patient subpopulations

Introduction : prediction

- **After a breast cancer diagnosis**
  $\rightarrow$ single or multiple events
  (recurrences, metastases, death)
- **Prediction of death**
  $\rightarrow$ clinical therapeutic decisions, and patient monitoring
  $\rightarrow$ patient information
  $\rightarrow$ trials : defining patient subpopulations
- **Account for**
  $\rightarrow$ individual characteristics
  $\rightarrow$ tumor characteristics
  $\rightarrow$ previous treatments
  $\rightarrow$ evolution of longitudinal markers (*Rizopoulos, 2011 ; Proust-Lima 2009*)

Introduction : Motivating example

- Cohort of patients with **operable breast cancer**
- Treated in a **comprehensive cancer center** and followed 13.9 years (median)
- **Recurrent events** observed : loco-regional relapses, distant metastases ; until 3 events per patient
- Hypothesis : individual covariates but also **recurrent event process** may improve prediction of death risk

Example 6 : Joint frailty models

```
library(frailtypack)
data(breastc)

joint <- frailtyPenal(formula = Surv(tt0caly,
    tt1caly, event) ~ cluster(groupe2) +
     age1 + age2 + emboln + taille + her2n + rhposn
          + nplusn +
     grade2 + grade3 + terminal(death),
     formula.terminalEvent = ~age1 + age2 + emboln
        + taille + her2n + rhposn + nplusn + grade2
         + grade3,
     data = recurrent, recurrentAG = TRUE, n.knots
        = 4, kappa = c(1e+06, 13000))
```

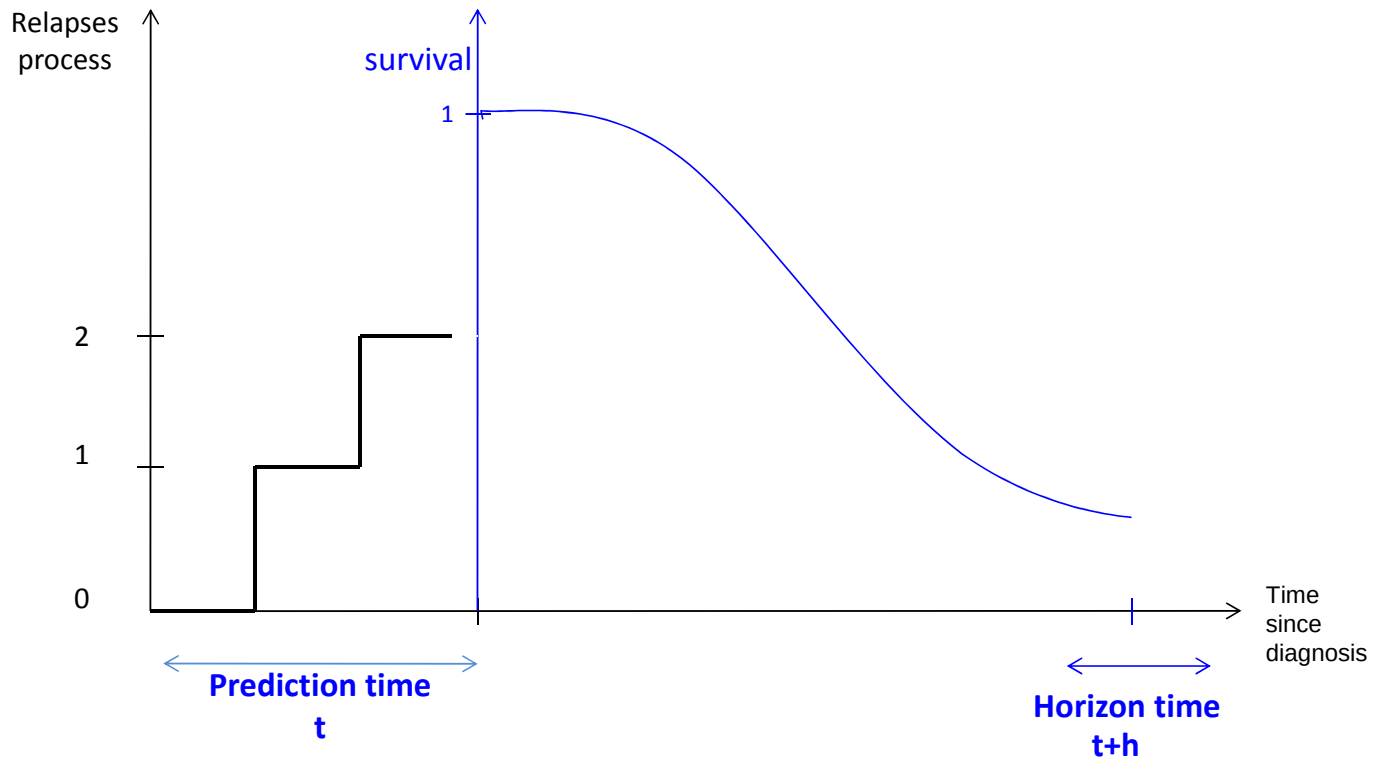
```
  Joint gamma frailty model for recurrent and a terminal event processes
  using a Penalized Likelihood on the hazard function
Recurrences:
-------------
            coef exp(coef) SE coef (H) SE coef (HIH)        z          p
age1    0.162992  1.177028    0.125480      0.125480  1.29895 1.9396e-01
age2    0.933765  2.544069    0.171668      0.171668  5.43936 5.3473e-08
emboln  0.386441  1.471733    0.126220      0.126220  3.06164 2.2012e-03
taille  0.623664  1.865752    0.122879      0.122879  5.07544 3.8660e-07
her2n   0.357323  1.429498    0.168662      0.168662  2.11857 3.4126e-02
rhposn -0.210873  0.809877    0.181728      0.181728 -1.16038 2.4589e-01
nplusn  0.596722  1.816156    0.125068      0.125068  4.77118 1.8315e-06
grade2  0.760003  2.138282    0.163990      0.163990  4.63445 3.5789e-06
grade3  0.795251  2.214997    0.204387      0.204387  3.89091 9.9871e-05
Terminal event:
----------------
            coef exp(coef) SE coef (H) SE coef (HIH)         z          p
age1   -1.035406  0.355082    0.313827      0.313827 -3.299285 9.6931e-04
age2    0.567811  1.764400    0.392589      0.392589  1.446324 1.4809e-01
emboln  1.209753  3.352658    0.318015      0.318015  3.804078 1.4233e-04
taille  1.543038  4.678782    0.281096      0.281096  5.489369 4.0337e-08
her2n   0.270172  1.310190    0.382119      0.382119  0.707037 4.7954e-01
rhposn -1.459270  0.232406    0.434286      0.434286 -3.360155 7.7899e-04
nplusn  1.508557  4.520203    0.316754      0.316754  4.762544 1.9117e-06
grade2  2.078511  7.992562    0.437872      0.437872  4.746848 2.0661e-06
grade3  2.468411 11.803671    0.536216      0.536216  4.603387 4.1567e-06
 Frailty parameters:
   theta (variance of Frailties, w): 1.03638 (SE (H): 0.0648069 ) p = 0
   alpha (w^alpha for terminal event): 4.60777 (SE (H): 0.28441 ) p = 0
   penalized marginal log-likelihood = -3016.24
   Convergence criteria:
   parameters = 4.02e-05 likelihood = 0.000258 gradient = 1.9e-06
   LCV = the approximate likelihood cross-validation criterion
         in the semi parametric case     = 2.03842
   n observations= 1494  n subjects= 1067
   n recurrent events= 427   n terminal events= 330
   number of iterations:  15   Exact number of knots used:  4
   Value of the smoothing parameters: kappa1=1e+06 and kappa2=13000
```

Prognostic joint model

Variable	% of patients	For recurrent events HR	(95%*CI*)	For death HR	(95%*CI*)
- Age					
$]40-55]$ *vs* $]55-84]$	(36.6)	1.18	(0.92-1.51)	0.36	(0.19-0.66)
$[28-40]$ *vs* $]55-84]$	(7.7)	2.54	(1.82-3.56)	1.76	(0.82-3.81)
- P. vasc. invas.	(26.7)	1.47	(1.15-1.88)	3.35	(1.80-6.25)
- Tumor size	(22.7)	1.86	(1.47-2.37)	4.68	(2.70-8.12)
> 20 *vs* ≤ 20 *mm*					
- HER2 positive	(11.2)	1.43	(1.03-1.99)	1.31	(0.62-2.77)
- HR	(83.0)	0.81	(0.57-1.16)	0.23	(0.10-0.54)
(+ vs -)					
- Nodes involv.	(42.3)	1.82	(1.42-2.32)	4.52	(2.43-8.41)
(yes vs no)					
- Grade					
II vs I	(45.7)	2.14	(1.55-2.95)	7.99	(3.39-18.85)
III vs I	(24.6)	2.21	(1.48-3.31)	10.80	(4.13-33.76)
θ			1.04 (se=0.06)		
α			4.61 (se=0.28)		
LCV			2.04		

Prediction using joint frailty models

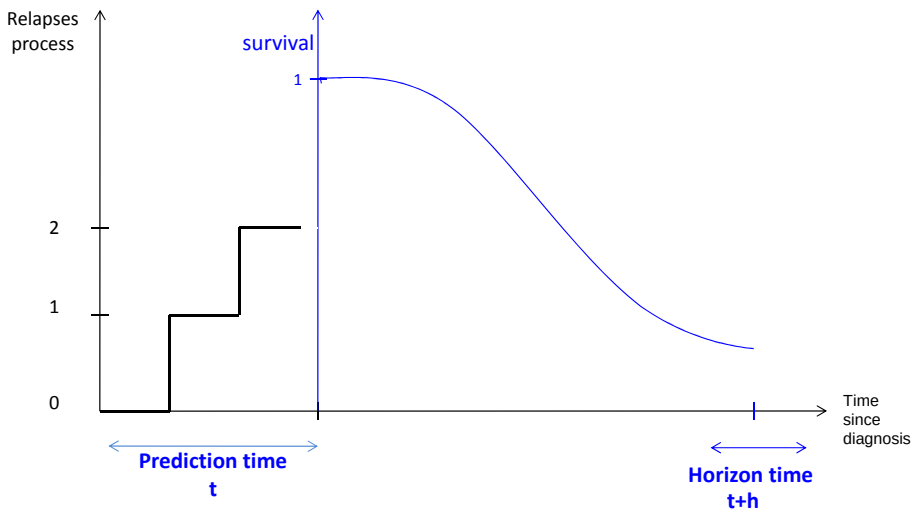
AIM : To predict the risk of death between time t and $t + h$ given the recurrent event process before time t in the context of joint modeling



Dynamic prediction

- Consider a new subject i **free of death at time** t (i.e. $D > t$), for whom we observe j recurrences before t and for whom the vector of covariates Z_{ij}^R and Z_{ij}^D are available at time of prediction
- The history of recurrences for patient i until time t is :

$$\mathcal{H}_i^J(t) = \{N_i^R(t) = J, X_{i1} < \ldots < X_{iJ} \leq t\}$$



Dynamic prediction

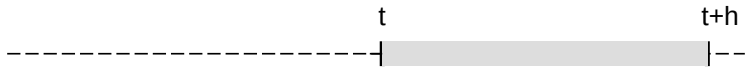
Distinguish **two settings** for the probability of death between t and

Setting 1

Exactly 3 recurrent events before t

Setting 2

Whatever the history of recurrent events before t



Recurrent event

Window of prediction of death

Period where we consider what happens

Period where we do not consider what happens

Dynamic prediction

Setting 1 : with exactly j recurrences before t

$$P^1(t, t+h; \xi) = P(D_i \leq t+h | D_i > t, \mathcal{H}_i^{J,1}(t), Z_{ij}^R, Z_i^D, \xi)$$

$$= \frac{\int_0^\infty [S_i^D(t|Z_i^D, u_i, \xi) - S_i^D(t+h|Z_i^D, u_i, \xi)](u_i)^J S_{i(J+1)}^R(t|Z_{ij}^R, u_i, \xi) g(u_i) \mathrm{d}u_i}{\int_0^\infty S_i^D(t|Z_i^D, u_i, \xi)(u_i)^J S_{i(J+1)}^R(t|Z_{ij}^R, u_i, \xi) g(u_i) \mathrm{d}u_i}$$

and $\mathcal{H}_i^{J,1}(t) = \{N_i^R(t) = J, X_{i1} < \ldots < X_{iJ} \leq t\}$, with $X_{i0} = 0$ and $X_{i(J+1)} > t$

Dynamic prediction

Setting 1 : with exactly j recurrences before t

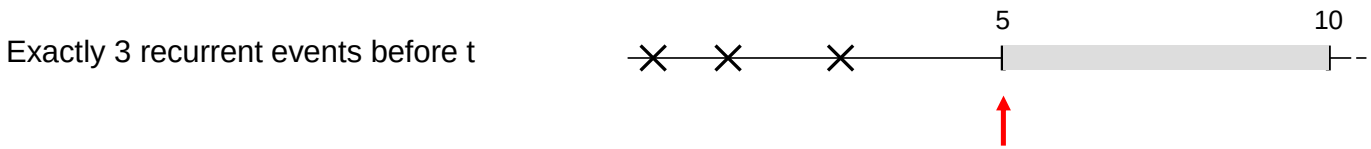
$$P^1(t, t+h; \xi) = P(D_i \leq t+h | D_i > t, \mathcal{H}_i^{J,1}(t), Z_{ij}^R, Z_i^D, \xi)$$

$$= \frac{\int_0^\infty [S_i^D(t|Z_i^D, u_i, \xi) - S_i^D(t+h|Z_i^D, u_i, \xi)](u_i)^J S_{i(J+1)}^R(t|Z_{ij}^R, u_i, \xi) g(u_i) \mathrm{d}u_i}{\int_0^\infty S_i^D(t|Z_i^D, u_i, \xi)(u_i)^J S_{i(J+1)}^R(t|Z_{ij}^R, u_i, \xi) g(u_i) \mathrm{d}u_i}$$

and $\mathcal{H}_i^{J,1}(t) = \{N_i^R(t) = J, X_{i1} < \ldots < X_{iJ} \leq t\}$, with $X_{i0} = 0$ and $X_{i(J+1)} > t$

Example :

Up to now Mrs Martin has developed 3 recurrences of his initial cancer, her probability of dying in the next 5 years is x%



Dynamic prediction

Setting 2 : considering the recurrence history only in the parameters estimation

$$
\begin{aligned}
& P^2(t, t+h; \xi) \\
& = P(D_i \leq t+h | D_i > t, Z_i^D, \xi) \\
& = \frac{\int_0^\infty [S_i^D(t|Z_i^D, u_i, \xi) - S_i^D(t+h|Z_i^D, u_i, \xi)] g(u_i) \mathrm{d}u_i}{\int_0^\infty S_i^D(t|Z_i^D, \xi, u_i) g(u_i) \mathrm{d}u_i}
\end{aligned}
$$

Dynamic prediction

Setting 2 : considering the recurrence history only in the parameters estimation

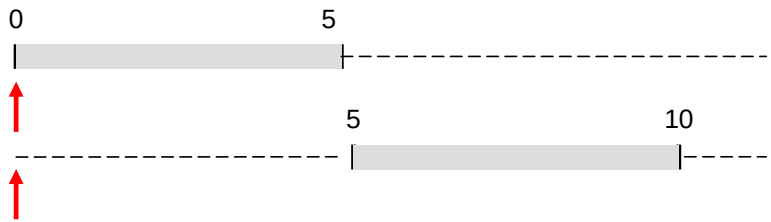
$$
\begin{aligned}
& P^2(t, t+h; \xi) \\
& = P(D_i \leq t+h | D_i > t, Z_i^D, \xi) \\
& = \frac{\int_0^\infty [S_i^D(t|Z_i^D, u_i, \xi) - S_i^D(t+h|Z_i^D, u_i, \xi)] g(u_i) \mathrm{d}u_i}{\int_0^\infty S_i^D(t|Z_i^D, \xi, u_i) g(u_i) \mathrm{d}u_i}
\end{aligned}
$$

Example :

'her probability of dying in the next 5 years is x%'

'if her still alive in 5 years, her probability of dying over the next 5 years will be x%'

Whatever the history of recurrent events before t



Dynamic prediction : variability of the probability estimators

by **Monte Carlo** :

- at each b step (b=1,...,B=1000) :
  $\hat{\xi} = (\widehat{\lambda_0^R(.)}, \widehat{\lambda_0^D(.)}, \hat{\boldsymbol{\beta}}, \hat{\alpha}, \hat{\theta})$ from $\mathcal{MN}(\hat{\xi}, \hat{\Sigma}_\xi)$.
  estimate $P^b(t, t+h; \hat{\xi})$

- Percentile confidence interval : using the 2.5^{th} and the 97.5^{th} percentiles

Dynamic prediction : Error of prediction

Based on a weighted time-dependent Brier Score (IPCW error)

$$Err_{t+h} = \frac{1}{N_t} \sum_{i=1}^{N_t} [I(T_i^D > t+h) - (1 - \hat{P}(t, t+h; \hat{\xi}))]^2 \hat{w}_i(t+h, \hat{G}_N(.))$$

with

$$w_i(t+h, \hat{G}_N(.)) = \frac{I(T_i^D \leq t+h)\delta_i^D}{\hat{G}_N(T_i^D)/\hat{G}_N(t)} + \frac{I(T_i^D > t+h)}{\hat{G}_N(t+h)/\hat{G}_N(t)}$$

T_i^D = observed survival time ; δ_i = event indicator

N_t =patients alive and uncensored at t

$\hat{G}_N(t)$ = KM estimate or adjusted Cox estimate of the censoring distribution

Validated by a 10-fold cross-validation

Brier. Monthly Weather Review 1950 - Gerds et al. Biometrical J 2006

Application

- 1067 patients
- median follow-up : 13.8 years (min=5 months)
- 330 patients died
- 362 patients with recurrent events (mean 0.40), i.e. 427 obsevations (locoregional relapses and distant metastases)

N events	0	1	2	3
Alive	600	114	20	3
Died	105	187	37	1
All	705	301	57	4

with the R package **frailtypack** :
(http ://cran.r-project.org/web/packages/frailtypack/)

Prediction in joint frailty models (1)

4 subjects : age 55 years, no peritumoral vascular invasion, tumor size >20 mm, HER2-, HR+, no lymph node involvment, grade=2

```
# construction of the dataframe for prediction
datapred <- data.frame(tt1.cal=0,event=0,subject=0,age1=0,
    emboln=0,taille=0,her2n=0,rhposn=0,nplusn=0,grade2=0,
    grade3=0)

# subject 1: one relapse at 1
datapred[1,] <- c(1,1,1,0,0,0,0,1,0,1,0)
# subject 2: one relapse at 2.5
datapred[2,] <- c(2.5,1,2,0,0,0,0,1,0,1,0)
# subject 3: one relapse at 4.9
datapred[3,] <- c(4.9,1,3,0,0,0,0,1,0,1,0)

# subject 4: first relapse at 1
datapred[4,] <- c(1,1,4,0,0,0,0,1,0,1,0)
# subject 4: second relapse at 2
datapred[5,] <- c(2,1,4,0,0,0,0,1,0,1,0)
# subject 4: censoring at 3
datapred[5,] <- c(3,0,4,0,0,0,0,1,0,1,0)
# ...
```

Prediction in joint frailty models (2)

Prediction between 5 and 10 or between 5 and 15 for each subject given relapses

```
pred <- prediction(joint,datapred,5,c(5,10))
```

with 'joint' the fit of the joint model
with 'datapred' the dataframe for prediction

Prediction values - between 5 and 10 years

Recurrence history	Risk of death between 5 and 10 years	
	$P^1(5, 10; \hat{\xi})$	$P^2(5, 10; \hat{\xi})$
No recurrence	10.8 (4.2)	12.7 (4.5)
One recurrence		
$X_{i1} = 1$	30.3 (8.9)	12.7 (4.5)
$X_{i1} = 2.5$	30.3 (8.9)	12.7 (4.5)
$X_{i1} = 4.9$	30.3 (8.9)	12.7 (4.5)
Two recurrences		
$X_{i1} = 1, X_{i2} = 2$	50.6 (11.4)	12.7 (4.5)
$X_{i1} = 2, X_{i2} = 4$	50.6 (11.4)	12.7 (4.5)
$X_{i1} = 4, X_{i2} = 4.9$	50.6 (11.4)	12.7 (4.5)
Three recurrences		
$X_{i1} = 1, X_{i2} = 2, X_{i3} = 3$	67.4 (11.9)	12.7 (4.5)
$X_{i1} = 1, X_{i2} = 2.5, X_{i3} = 4.9$	67.4 (11.9)	12.7 (4.5)
$X_{i1} = 3, X_{i2} = 4, X_{i3} = 4.9$	67.4 (11.9)	12.7 (4.5)

Prediction values - between 5 and 10 years

Recurrence history	Risk of death between 5 and 10 years	
	$P^1(5, 10; \hat{\xi})$	$P^2(5, 10; \hat{\xi})$
No recurrence	10.8 (4.2)	12.7 (4.5)
One recurrence		
$X_{i1} = 1$	30.3 (8.9)	12.7 (4.5)
$X_{i1} = 2.5$	30.3 (8.9)	12.7 (4.5)
$X_{i1} = 4.9$	30.3 (8.9)	12.7 (4.5)
Two recurrences		
$X_{i1} = 1, X_{i2} = 2$	50.6 (11.4)	12.7 (4.5)
$X_{i1} = 2, X_{i2} = 4$	50.6 (11.4)	12.7 (4.5)
$X_{i1} = 4, X_{i2} = 4.9$	50.6 (11.4)	12.7 (4.5)
Three recurrences		
$X_{i1} = 1, X_{i2} = 2, X_{i3} = 3$	67.4 (11.9)	12.7 (4.5)
$X_{i1} = 1, X_{i2} = 2.5, X_{i3} = 4.9$	67.4 (11.9)	12.7 (4.5)
$X_{i1} = 3, X_{i2} = 4, X_{i3} = 4.9$	67.4 (11.9)	12.7 (4.5)

Prediction values - between 5 and 10 years

Recurrence history	Risk of death between 5 and 10 years	
	$P^1(5, 10; \hat{\xi})$	$P^2(5, 10; \hat{\xi})$
No recurrence	10.8 (4.2)	12.7 (4.5)
One recurrence		
$X_{i1} = 1$	30.3 (8.9)	12.7 (4.5)
$X_{i1} = 2.5$	30.3 (8.9)	12.7 (4.5)
$X_{i1} = 4.9$	30.3 (8.9)	12.7 (4.5)
Two recurrences		
$X_{i1} = 1, X_{i2} = 2$	50.6 (11.4)	12.7 (4.5)
$X_{i1} = 2, X_{i2} = 4$	50.6 (11.4)	12.7 (4.5)
$X_{i1} = 4, X_{i2} = 4.9$	50.6 (11.4)	12.7 (4.5)
Three recurrences		
$X_{i1} = 1, X_{i2} = 2, X_{i3} = 3$	67.4 (11.9)	12.7 (4.5)
$X_{i1} = 1, X_{i2} = 2.5, X_{i3} = 4.9$	67.4 (11.9)	12.7 (4.5)
$X_{i1} = 3, X_{i2} = 4, X_{i3} = 4.9$	67.4 (11.9)	12.7 (4.5)

Prediction values - **between 5 and 15 years**

Recurrence history	Risk of death between 5 and 15 years	
	$P^1(5, 15; \hat{\xi})$	$P^2(5, 15; \hat{\xi})$
No recurrence	22.7 (4.8)	25.6 (4.7)
One recurrence		
$X_{i1} = 1$	53.0 (6.9)	25.6 (4.7)
$X_{i1} = 2.5$	53.0 (6.9)	25.6 (4.7)
$X_{i1} = 4.9$	53.0 (6.9)	25.6 (4.7)
Two recurrences		
$X_{i1} = 1, X_{i2} = 2$	75.6 (6.0)	25.6 (4.7)
$X_{i1} = 2, X_{i2} = 4$	75.6 (6.0)	25.6 (4.7)
$X_{i1} = 4, X_{i2} = 4.9$	75.6 (6.0)	25.6 (4.7)
Three recurrences		
$X_{i1} = 1, X_{i2} = 2, X_{i3} = 3$	88.4 (4.1)	25.6 (4.7)
$X_{i1} = 1, X_{i2} = 2.5, X_{i3} = 4.9$	88.4 (4.1)	25.6 (4.7)
$X_{i1} = 3, X_{i2} = 4, X_{i3} = 4.9$	88.4 (4.1)	25.6 (4.7)

Death prediction for 2 particular cases : n407 and n506

with the same caracteristics at time t $= 0$:
between 40 and 55 years ; no peritumoral vasc. invasion ; tumor size $\leq$ 20 mm ; HER2 - ; RH $+$; no lymph node involv. ; grade I

```
#death prediction of 2 subjects:n 407 and n 506
datapred <- recurrent[recurrent$subject %in% c
    (407,506),]
```

Death prediction for 2 particular cases : n407 and n506

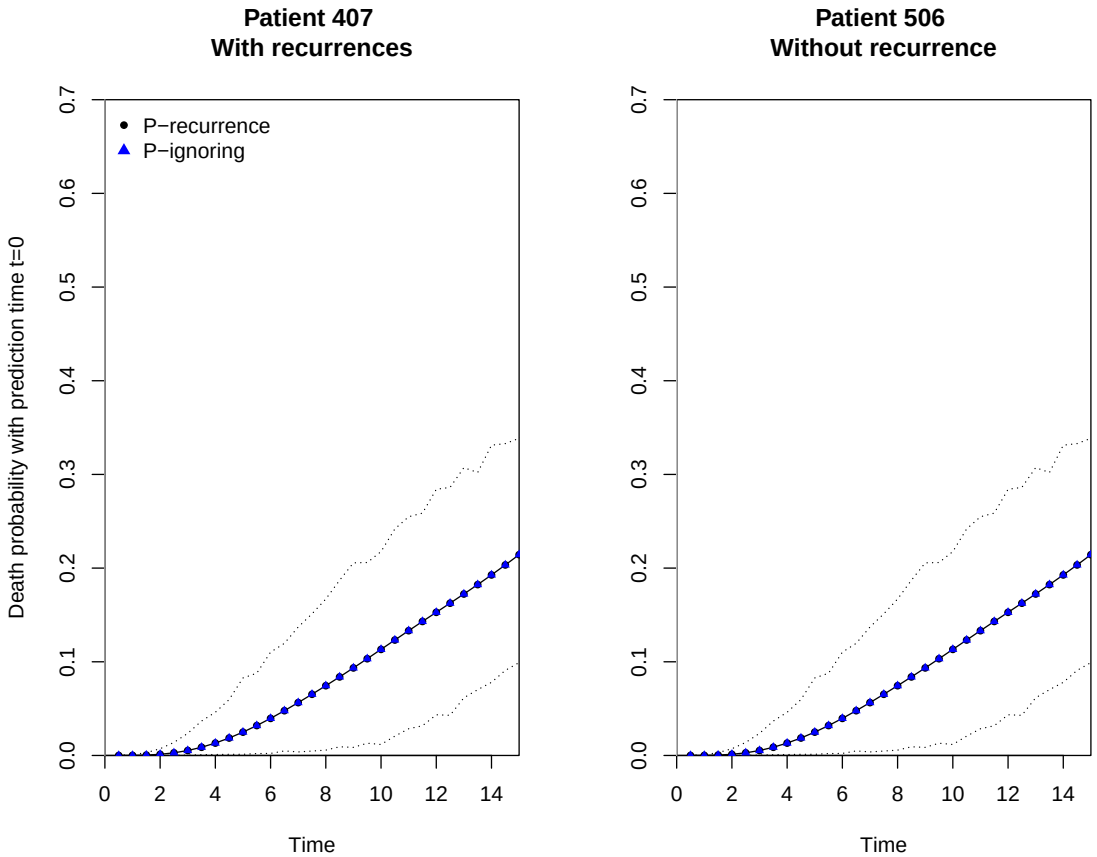
with the same caracteristics at time t = 0 :
between 40 and 55 years ; no peritumoral vasc. invasion ; tumor size $\leq$ 20 mm ; HER2 - ; RH + ; no lymph node involv. ; grade I
but, with a different history of relapses

```
> datapred

tt0 tt1.cal event subject age1 age2 emboln taille her2n rhposn nplusn grade2 grade3

0.000000  1.735797     1     407    1     0    0    0    0    1      0      1      0
1.735797  3.028063     1     407    1     0    0    0    0    1      0      1      0
3.028063  8.224504     1     407    1     0    0    0    0    1      0      1      0
8.224504 10.830938     0     407    1     0    0    0    0    1      0      1      0
0.000000 18.869268     0     506    1     0    0    0    0    1      0      1      0
```

Death prediction for 2 particular cases

Baseline prediction :

```
pred <- prediction(joint,datapred,t=0,window=seq(0.1,15,0.1))
plot(pred)
```

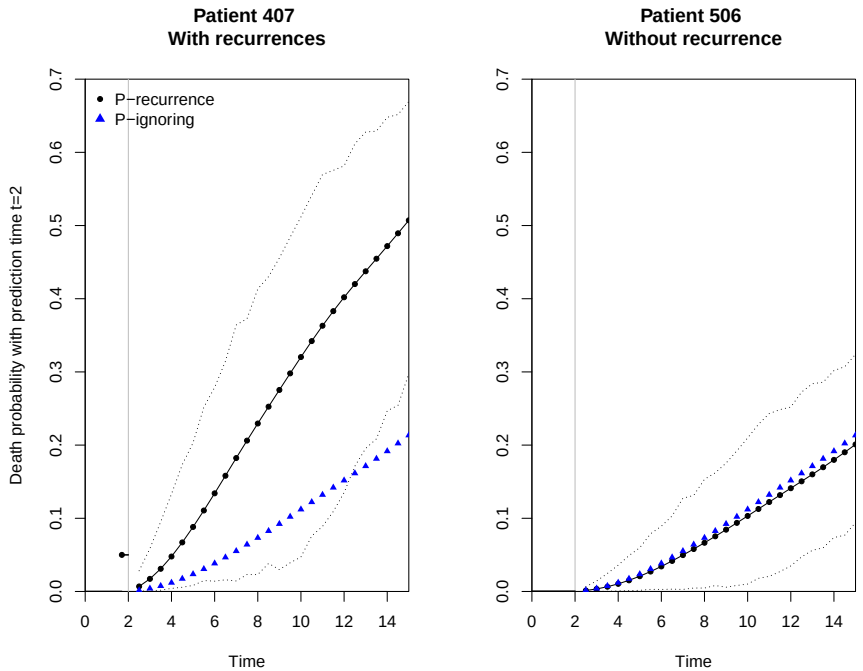


Death prediction for 2 particular cases

Prediction time t=2 years

```
pred <- prediction(joint,datapred,t=2,window=seq(0.1,13,0.1))
plot(pred)
```

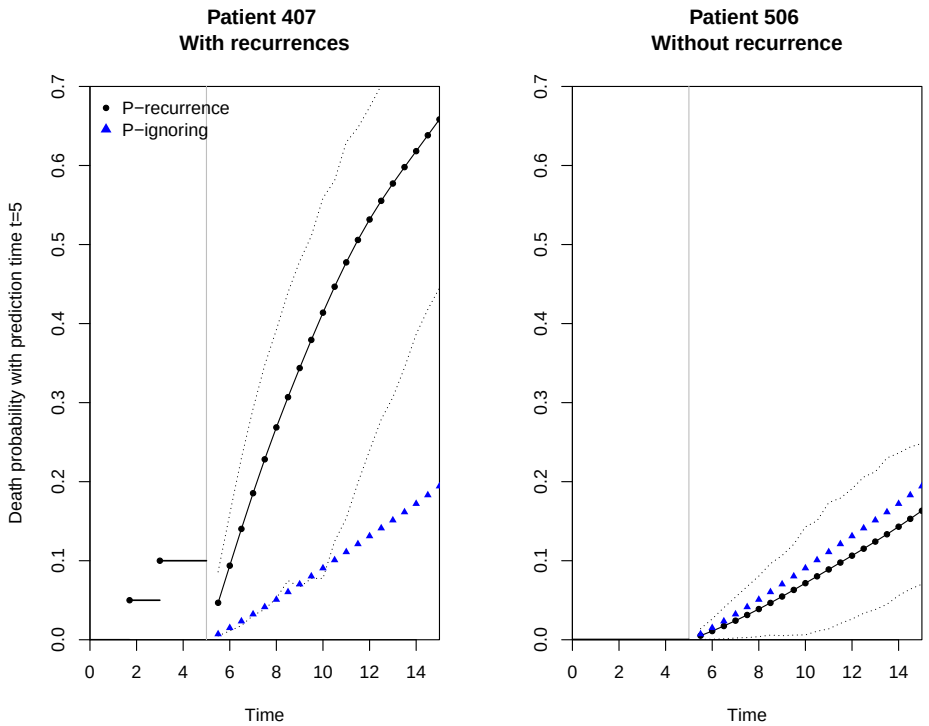
each y point corresponds to the prediction of death between 2 years and x years



Death prediction for 2 particular cases

Prediction time t=5 years

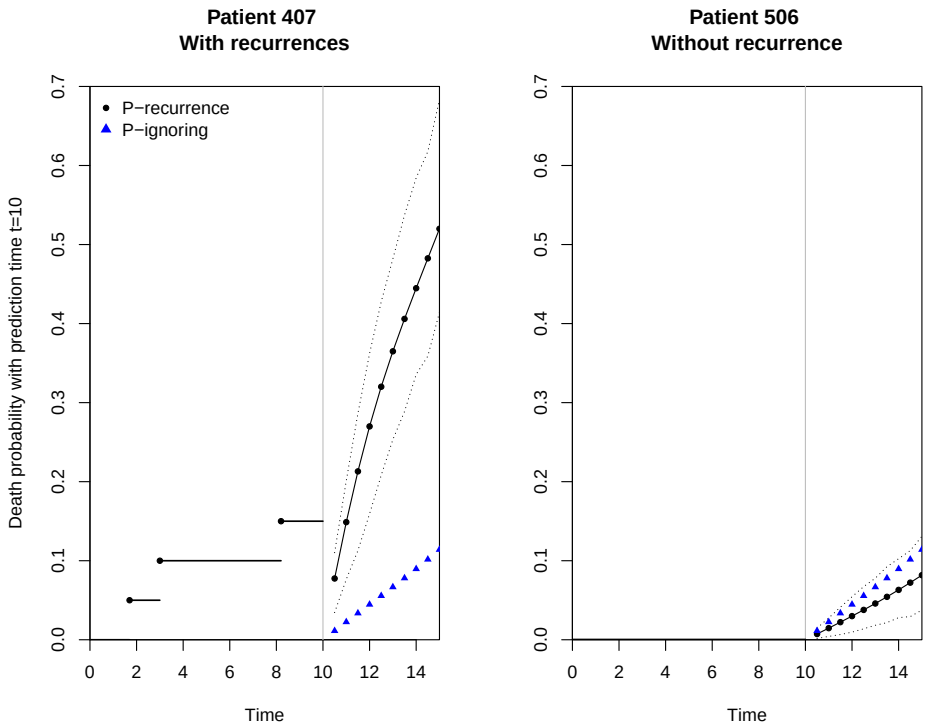
```
pred <- prediction(joint,datapred,t=5,window=seq(0.1,10,0.1))
plot(pred)
```



Death prediction for 2 particular cases

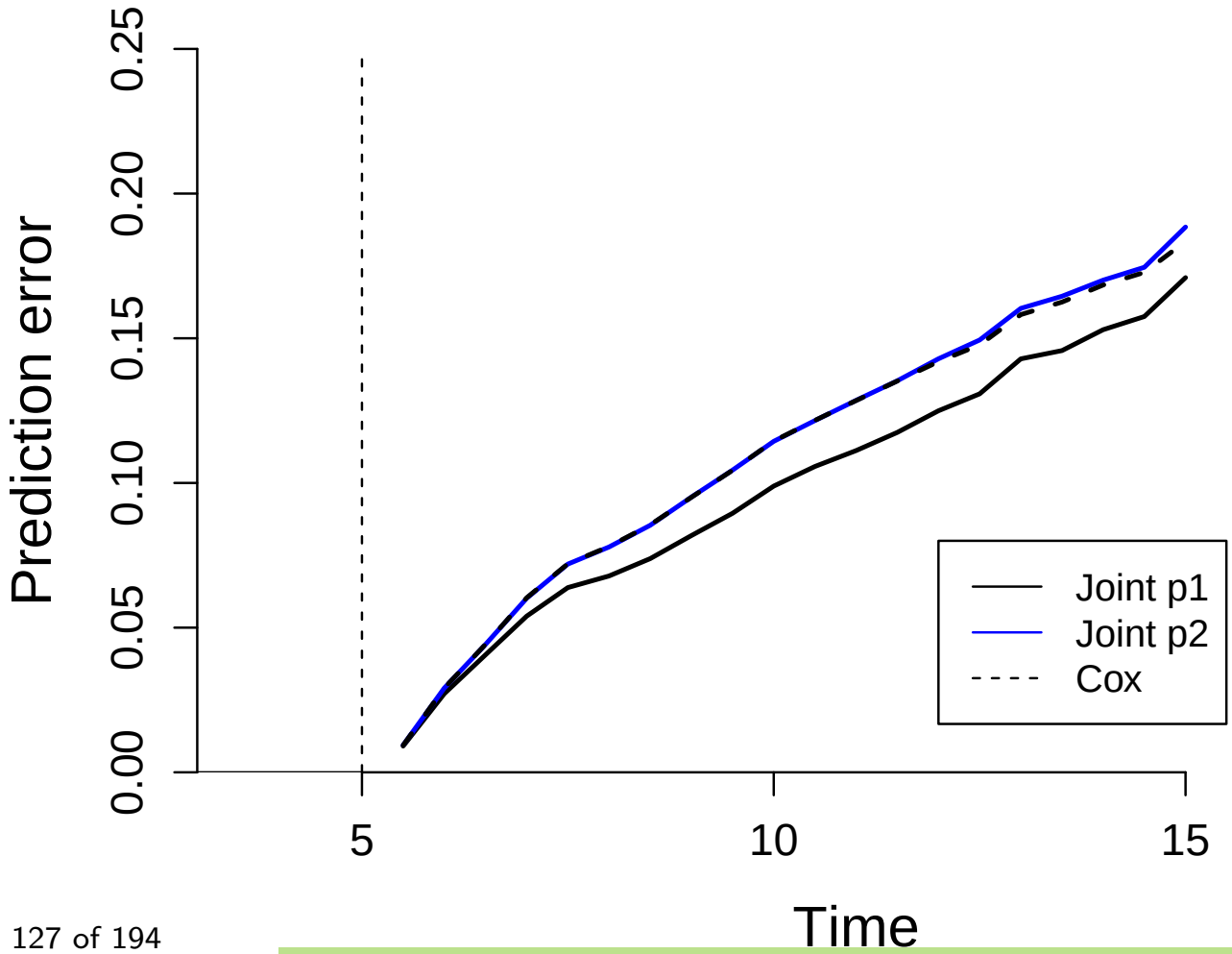
Prediction time t=10 years

```
pred <- prediction(joint,datapred,t=10,window=seq(0.1,5,0.1))
plot(pred)
```



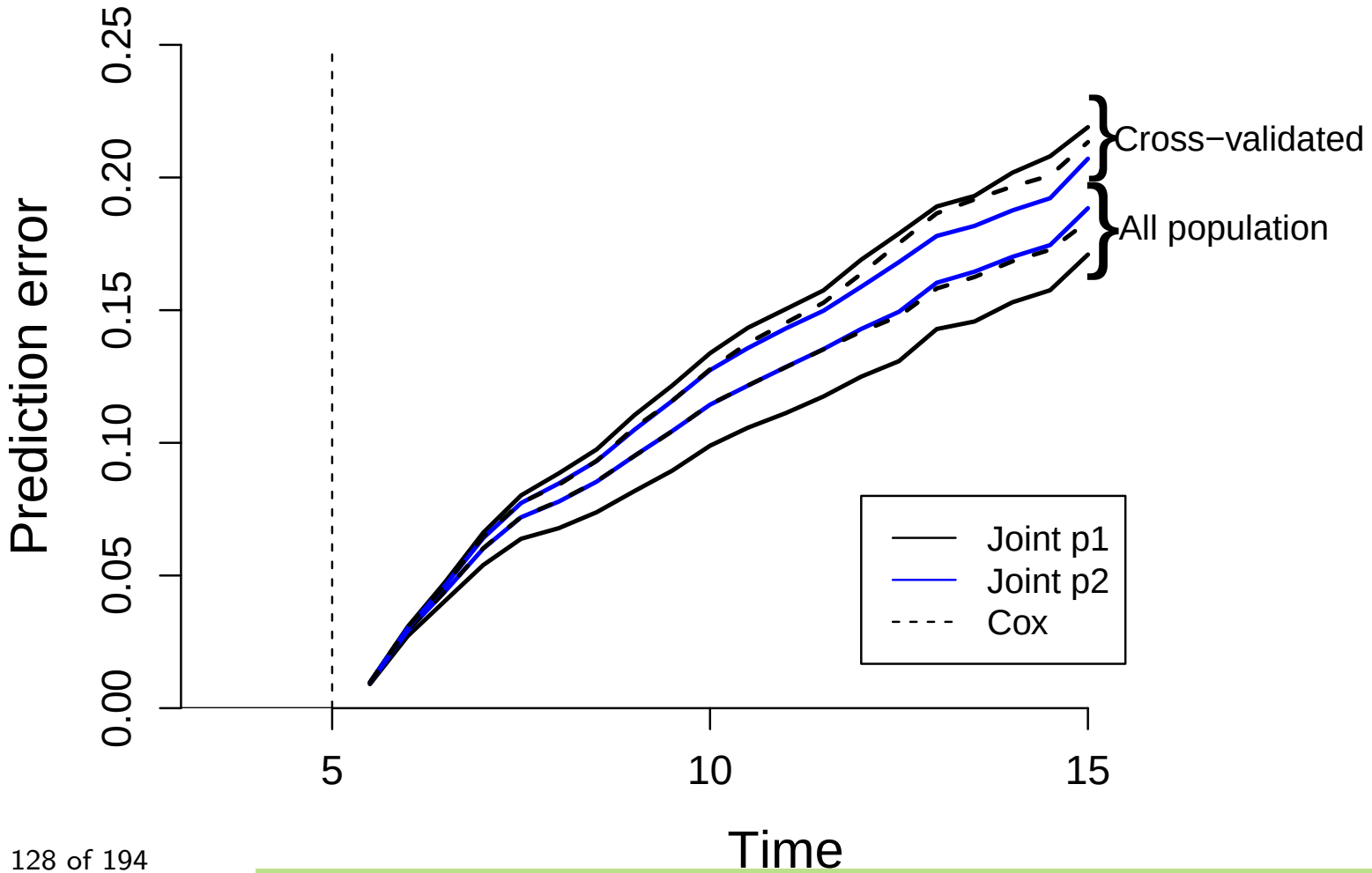
Death prediction error

Prediction at 5 years (949 patients alive)



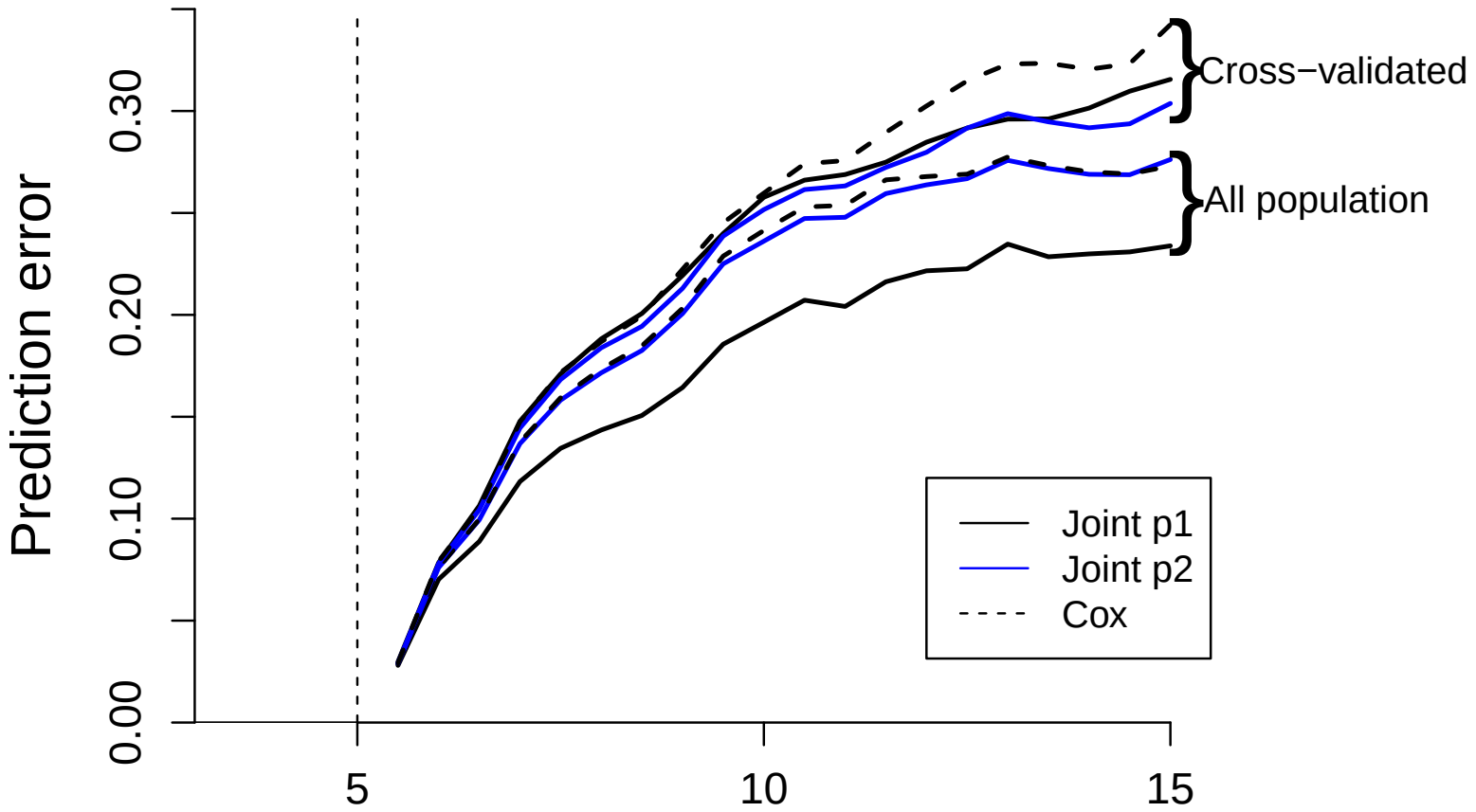
Prediction error

Prediction at 5 years (949 patients alive), with 10-fold cross-validation



Prediction error

Prediction at 5 years (267 patients alive with recurrence), with 10-fold cross-validation



Conclusion

- Recurrent event process seems interesting to predict the risk of death, in framework of joint models
- Dynamic prediction : updated with new events
- Joint modeling gives better results than Cox model with lower prediction error
- However, the 10-fold cross-validation suggests a higher risk of over-fitting
- Conditional prediction possible, but interest is limited (ex : for a specific subject from the study)

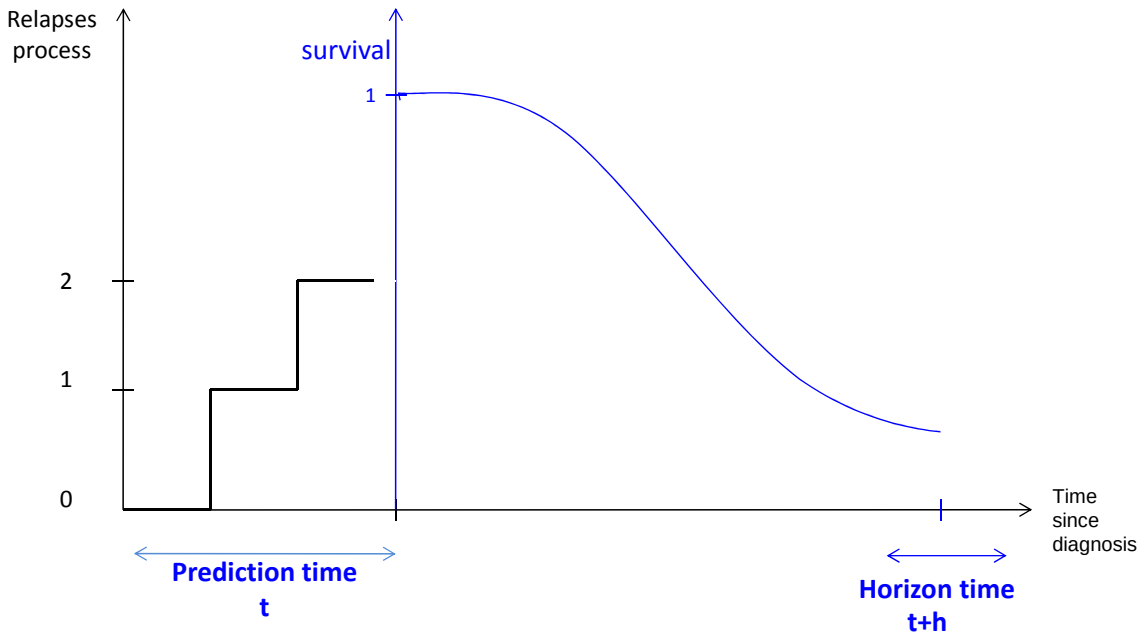
Conclusion

- Other aspects :
  - Independent external validation (to avoid over-optimistic validation results)
  - To study the prediction of the risk of events (relapse) along with the risk of death
  - Prediction using alternative models (landmark approach, additive frailty models ...)

Prediction of a new recurrent event using joint frailty models

AIM : To predict the **risk of a new recurrent event** between time t and $t + h$ given the recurrent event process before time t in the context of joint modeling

ex : predict the risk of a third relapse

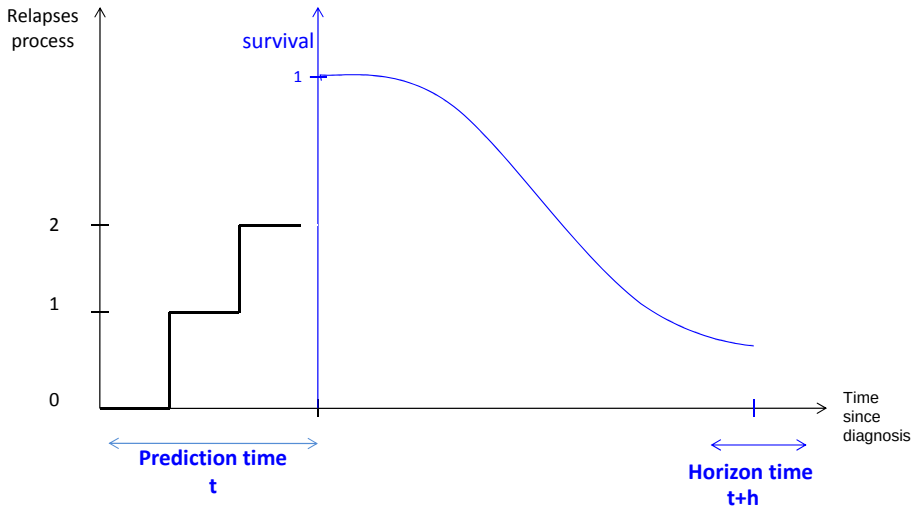


Dynamic prediction of a new recurrent event

- Consider a new subject i **free of death at time** t (i.e. $D > t$), for whom we observe exactly J recurrences before t and for whom the vector of covariates Z_{ij}^R and Z_{ij}^D are available at time of prediction
- The history of recurrences for patient i until time t is :

$$\mathcal{H}_i^J(t) = \{N_i^R(t) = J, X_{i1} < ... < X_{ij} \leq t\}$$

with $X_{i0} = 0, N_i^R(t)$ is the number of observed recurrent events before t



Dynamic prediction of a new recurrent event

$$\mathbb{P}(t, t+w, \xi) = \mathbb{P}(X_{i(J+1)} \leq t+w | X_{i(J+1)} > t, D_i > t, \mathcal{H}_i^J(t), Z_{ij}^R, Z_i^D, \xi)$$

Dynamic prediction of a new recurrent event

$$\mathbb{P}(t, t+w, \xi) = \mathbb{P}(X_{i(J+1)} \leq t+w | X_{i(J+1)} > t, D_i > t, \mathcal{H}_i^J(t), Z_{ij}^R, Z_i^D, \xi)$$

$$= \int_0^\infty \mathbb{P}(X_{i(J+1)} \leq t+w | X_{i(J+1)} > t, D_i > t, \mathcal{H}_i^J(t), Z_{ij}^R, Z_i^D, u_i, \xi)$$

$$\times g(u_i | X_{i(J+1)}^R > t, D_i > t, Z_{ij}^R, Z_i^D, \xi) \mathrm{d}u_i$$

Dynamic prediction of a new recurrent event

$$\mathbb{P}(t, t+w, \xi) = \mathbb{P}(X_{i(J+1)} \leq t+w | X_{i(J+1)} > t, D_i > t, \mathcal{H}_i^J(t), Z_{ij}^R, Z_i^D, \xi)$$

$$= \int_0^\infty \mathbb{P}(X_{i(J+1)} \leq t+w | X_{i(J+1)} > t, D_i > t, \mathcal{H}_i^J(t), Z_{ij}^R, Z_i^D, u_i, \xi)$$

$$\times g(u_i | X_{i(J+1)}^R > t, D_i > t, Z_{ij}^R, Z_i^D, \xi) \mathrm{d}u_i$$

$$= \frac{\int_0^\infty [S_{i(J+1)}^R(t|Z_{ij}^R, u_i, \xi) - S_{i(J+1)}^R(t+w|Z_{ij}^R, u_i, \xi)] \cdot S_i^D(t|Z_i^D, u_i, \xi) \cdot (u_i)^J \cdot S_{i(J+1)}^R(X_{iJ}|Z_{ij}^R, u_i, \xi) \cdot g(u_i) \mathrm{d}u_i}{\int_0^\infty S_{i(J+1)}^R(t|Z_{ij}^R, u_i, \xi) \cdot S_i^D(t|Z_i^D, u_i, \xi) \cdot (u_i)^J \cdot S_{i(J+1)}^R(X_{iJ}|Z_{ij}^R, u_i, \xi) \cdot g(u_i) \mathrm{d}u_i}$$

Dynamic prediction of a new recurrent event

Implementation with R :

```
#-- prediction of relapse between 100 and 100+w
    given relapses
(with confidence intervals)
pred.joint <- prediction(joi,datapredj,t=100,
    window=seq(50,1500,50),
event = 'Recurrent',MC.sample=100)
plot(pred.joint,conf.bands=TRUE)
# each y-value of the plot corresponds to the
    prediction between [100,x]
```

Dynamic prediction of a new recurrent event

Implementation with R : (for recurrent and death prediction)

```
#-- prediction of relapse and death between 100
    and 100+w given relapses
(with confidence intervals)
pred.joint <- prediction(joi,datapredj,t=100,
    window=seq(50,1500,50),
event = 'Both',MC.sample=100)
plot(pred.joint,conf.bands=TRUE)
# each y-value of the plot corresponds to the
    prediction between [100,x]
```

Extension to a more general joint frailty models : with two independent frailty terms

In the standard frailty model, the frailty term u_i reflects :

- the intra-subject correlation for the recurrent event
- but also, the association between the recurrent and the terminal events

Aim of this more general joint frailty model :
to distinguish the origin of dependence (with two frailties)

Extension to a more general joint frailty models : with two independent frailty terms

The model with two independent frailty terms u_i and v_i :

$$\begin{cases} r_{ij}(t|u_i) = \mathbf{u_i v_i} r_0(t) \exp(\mathbf{X}_{Rij}^\top \beta_R) = u_i v_i r_{ij}(t) & \text{(recurrent event)} \\ \lambda_i(t|u_i) = \mathbf{u_i} \lambda_0(t) \exp(\mathbf{X}_{Ti}^\top \beta_T) = u_i \lambda_i(t) & \text{(terminal event)} \end{cases}, \tag{2}$$

$v_i \sim \Gamma(\frac{1}{\eta}, \frac{1}{\eta})$ $(\eta > 0)$ specific to the recurrent event rate

$u_i \sim \Gamma(\frac{1}{\theta}, \frac{1}{\theta})$ $(\theta > 0)$ specific to the association between the processes

- high variance η : strong dependence between the recurrent events
- high variance θ : recurrent and terminal events are strongly dependent

Mazroui et al. 2012, Stat in Medicine

Extension to a more general joint frailty models : with two independent frailty terms

Implementation with R :

```
joint <- frailtyPenal(formula = Surv(tt0caly,
    tt1caly, event) ~ cluster(groupe2) +
     age1 + age2 + emboln + taille + her2n + rhposn
          + nplusn +
     grade2 + grade3 + terminal(death),
     formula.terminalEvent = ~age1 + age2 + emboln
        + taille + her2n + rhposn + nplusn + grade2
         + grade3,
     data = recurrent, recurrentAG = TRUE, n.knots
        = 4, kappa = c(1e+06, 13000),jointGeneral=
        TRUE)
```

Extensions of frailty models with time-varying effects of covariates

For Cox, shared or joint frailty models
With a linear combination of B-splines with coefficients ζ of order q with m interior knots

$$\beta(\mathbf{t}) = \sum_{\mathbf{j}=-\mathbf{q}+\mathbf{1}}^{\mathbf{m}} \zeta_{\mathbf{j}} \mathbf{B}_{\mathbf{j},\mathbf{q}}(\mathbf{t})$$

2 tests :

- Proportional hazard assumption ? $H0 : \beta(t) = \beta$
  LRT statistic $\sim \chi^2$ of degree $m + q - 1$
- Significant association ? $H0 : \beta(t) = 0$
  LRT statistic $\sim \chi^2$ of degree $m + q$

(Mazroui et al. 2015 LIDA)

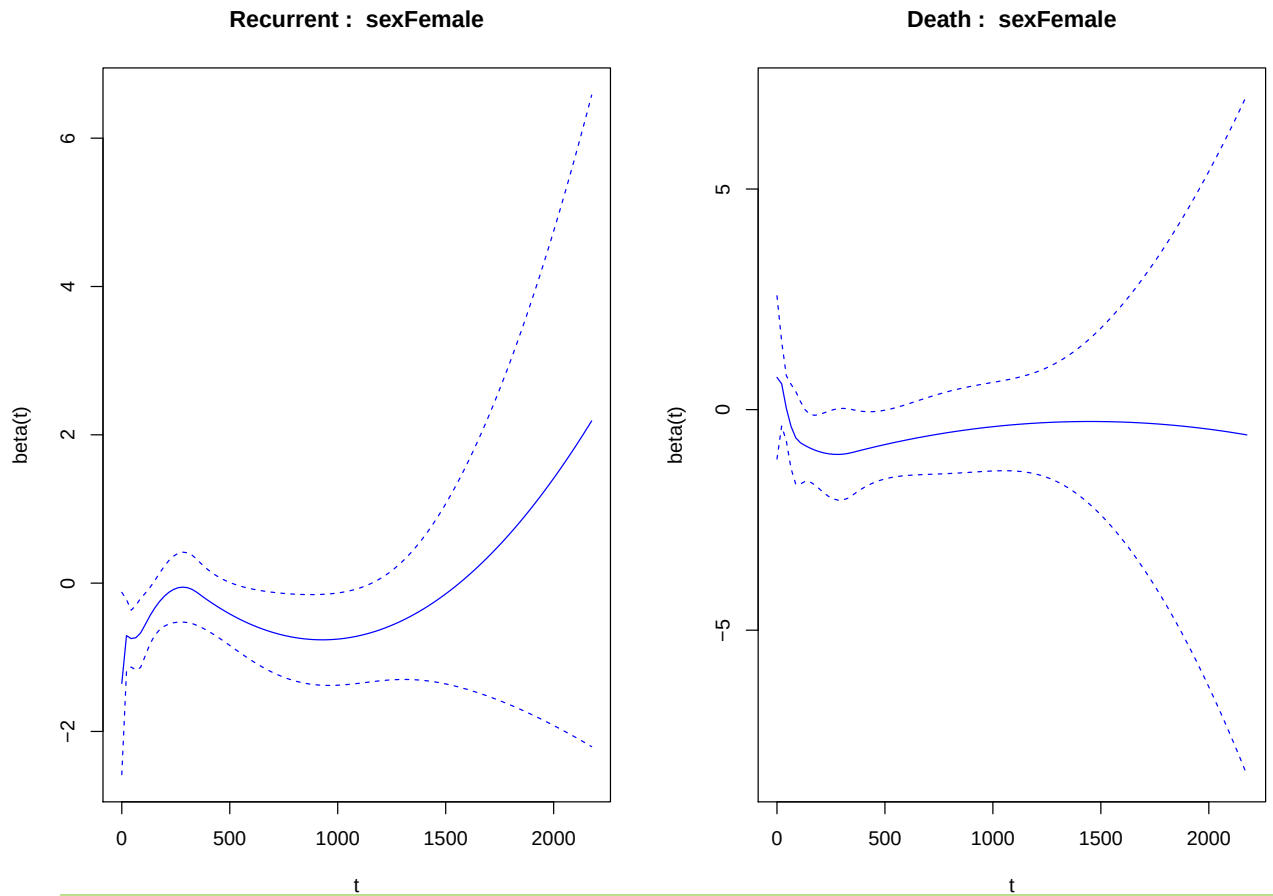
Extensions of frailty models with time-varying effects of covariates

Example : on readmission dataset
time-dependent coefficients can be estimated using B-splines of order q (option **betaorder**) with m interior knots (option **betaknots**).

```
R> modJoint.gap.timedep <- frailtyPenal(Surv(time,
    event) ~ cluster(id) +
    dukes + charlson + timedep(sex) + chemo +
       terminal(death),
    formula.terminalEvent = ~ dukes + charlson +
       timedep(sex) + chemo,
    data = readmission, n.knots = 8, kappa = c
       (2.11e+08, 9.53e+11),
    betaorder = 3, betaknots = 3)
```

Extensions of frailty models with time-varying effects of covariates

Example : on readmission dataset



Extensions of frailty models : Multivariate frailty model for 2 types of recurrent events and death

$$
\begin{aligned}
r_i^{(1)}(t|u_i, v_i) &= r_0^{(1)}(t)\exp(\beta_1' Z_i(t) + u_i) && \text{(rec. of type 1)} \\
r_i^{(2)}(t|u_i, v_i) &= r_0^{(2)}(t)\exp(\beta_2' Z_i(t) + v_i) && \text{(rec. of type 2)} \\
\lambda_i(t|u_i, v_i) &= \lambda_0(t)\exp(\beta_3' Z_i(t) + \alpha_1 u_i + \alpha_2 v_i) && \text{(death)}
\end{aligned}
$$

With two Gaussian and correlated random effects u_i, v_i :

$(u_i, v_i)^T \sim \mathcal{N}(0, \Sigma_{uv})$, with $\Sigma_{uv} = \begin{pmatrix} \theta_1 & \rho\sqrt{\theta_1\theta_2} \\ \rho\sqrt{\theta_1\theta_2} & \theta_2 \end{pmatrix}$

Multivariate frailty model for 2 types of recurrent events and death

- R : frailtypack *(Mazroui, Biom J 2013)*

```
    modMultiv.weib <- multivPenal(Surv(TIMEGAP,
        INDICREC)~cluster(PATIENT)+v1+v2+
event2(INDICMETA)+terminal(INDICDEATH),formula.
    Event2=~v1+v2+v3,
formula.terminalEvent=~v1,data=dataMultiv,hazard
    =''Weibull'')
print(modMultiv.weib)
```

Joint frailty models for 2 clustered time to events

References for joint frailty models

- Liu et al. (2004). Shared frailty models for recurrent events and a terminal event. **Biometrics**

- Rondeau et al. (2007). Joint frailty models for recurring events and death using maximum penalized likelihood estimation : application on cancer events. **Biostatistics**

- Mauguen et al. (2013). Dynamic prediction of risk of death using history of cancer recurrences in joint frailty models **Stat Med**

- Gerds et al. (2006). Consistent estimation of the expected brier score in general survival models with right-censored event times. **Biometrical J**

- Proust-Lima et al. (2009). Development and validation of a dynamic prognostic tool for prostate cancer recurrence using repeated measures of posttreatment PSA : a joint modeling approach. **Biostatistics**

- Rizopoulos et al. (2011) Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. **Biometrics**

- Brier (1950). **Monthly weather review**

- Rondeau V, and Gonzalez, JR. (2012) FRAILTYPACK : An R package for the analysis of correlated data with frailty models using the penalized likelihood estimation. **JSS.**

http ://cran.r-project.org/web/packages/**frailtypack**/
http ://cran.r-project.org/web/packages/pec/

PART 4 :

Joint models for a longitudinal biomarker an a terminal event

- Joint models
- Prediction using joint models

Joint models for longitudinal data and survival event(s)

Introduction - follow-up studies :

- **Repeated** (correlated) evaluations of the same measure in each subject over time :
  - a biomarker on a patient (e.g. PSA measurements, CD4 cell counts, cholesterol level)
  - longitudinal studies allow to investigate :
    - how means differ **at specific time points**, e.g. at the end of the study (cross-sectional effect)
    - how between means **change over time** (longitudinal effect)

Study case : Tumor evaluation in clinical trials

Context :

- Continuously increasing number of cancer clinical trials for treatment evaluation $\rightarrow$ necessity of a "common language"
- Some history
  - 1979 - WHO criteria
  - 2000, 2009 (v1.1) - RECIST(Response Evaluation Criteria in Solid Tumors)
  - 2009 - irRC (Immune Related Response Criteria)

RECIST criteria

http://www.irrecist.com/recist/recist-in-practice/02.html

- Target lesions
  - Unidimensional size, max 2 lesions per organ and up to 5 total
  - Progression : $>$ 20% increase over smallest sum observed ($>$ 5 mm absolute increase)
- Appearance of new lesions $\rightarrow$ global progression
- Unequivocal progression of non-target lesions $\rightarrow$ global progression

4 categories (Complete Response, Partial Response, Progressive Disease, Stable Disease)

$\Rightarrow$ dichotomization : response or no response / progression or no progression

Study case : Tumor evaluations in clinical trials

Objective of the study :
To study the **link** between the tumor size evolution and the risk of death, and to evaluate the **predictive accuracy** of the longitudinal tumor size on the OS.

Research questions :

- Which prognostic factors are linked to the biomarker and/or to the survival event ?
- What is the association between longitudinal measurements and the risk of an outcome of interest ?
- Can we use the longitudinal outcome to predict the event ?

Joint models for longitudinal data and survival event(s)

Issues :

- The repeated biomarker and the terminal event can be associated ?
- It is not recommended to use the longitudinal biomarker as a time-dependent covariate, because
  - its value is affected by the survival process (endogenous variable)
  - a terminal event can stop the evolution of the biomarker
  - the biomarker is measured with measurement errors and not observed at the failure times.

↪ **Solution : joint models**

(Rizopoulos et al. 2012)

Joint models for longitudinal data and survival event(s)

Main objectives of the joint model :

- **incorporating a longitudinal endogenous** time-dependent covariate measured with error to the survival model
- considering **informative censoring** for the longitudinal process
- analyzing **strength and structure of the association** between the survival and longitudinal processes

Joint models for longitudinal data and survival event(s)

Notations :

For individual i ($i = 1, \ldots, N$) we observe :

- l_i measurements of **longitudinal biomarker** (ex : sum of the longest diameters, SLD) : $y_i(t_{ik})$ for $k = 1, \ldots, l_i$
- **Observed time to terminal event** (death) : $T_i = \min(C_i, T_i^*)$
- **True time to terminal event** (death) : T_i^*
- $\delta_i^T = I_{\{T_i^* = T_i\}}$, event indicator,i.e., equals 1 for true events

Joint models for longitudinal data and survival event(s)

The joint model : in two parts

$$\begin{cases} Y_i(t) = m_i(t) + \epsilon_i(t) = \mathbf{X}_{Li}(t)^\top \beta_L + \mathbf{Z}_i(t)^\top \mathbf{b}_i + \epsilon_i(t) & \text{(biomarker)} \\ \lambda_i(t|\mathbf{b}_i) = \lambda_0(t) \ \exp(\mathbf{X}_{Ti}{}^\top \beta_T + h(\mathbf{b}_i, \beta_L, \mathbf{Z}_i(t), \mathbf{X}_{Li}(t))^\top \eta_T) & \text{(death)} \end{cases}$$

Joint models for longitudinal data and survival event(s)

The joint model : in two parts

$$\begin{cases} Y_i(t) = m_i(t) + \epsilon_i(t) = \mathbf{X}_{Li}(t)^\top \beta_L + \mathbf{Z}_i(t)^\top \mathbf{b}_i + \epsilon_i(t) & \text{(biomarker)} \\ \lambda_i(t|\mathbf{b}_i) = \lambda_0(t) \ \exp(\mathbf{X}_{Ti}{}^\top \beta_T + h(\mathbf{b}_i, \beta_L, \mathbf{Z}_i(t), \mathbf{X}_{Li}(t))^\top \eta_T) & \text{(death)} \end{cases}$$

Where

- $\mathbf{X}_{Li}(t)$ and $\mathbf{X}_{Ti}$ are vectors of **fixed effects covariates** (β_L and β_T their coefficients)

Joint models for longitudinal data and survival event(s)

The joint model : in two parts

$$\begin{cases} Y_i(t) = m_i(t) + \epsilon_i(t) = \mathbf{X}_{Li}(t)^\top \beta_L + \mathbf{Z}_i(t)^\top \mathbf{b}_i + \epsilon_i(t) & \text{(biomarker)} \\ \lambda_i(t|\mathbf{b}_i) = \lambda_0(t)\ \exp(\mathbf{X}_{Ti}{}^\top \beta_T + h(\mathbf{b}_i, \beta_L, \mathbf{Z}_i(t), \mathbf{X}_{Li}(t))^\top \eta_T) & \text{(death)} \end{cases}$$

Where

- $\mathbf{X}_{Li}(t)$ and $\mathbf{X}_{Ti}$ are vectors of fixed effects covariates (β_L and β_T their coefficients)
- $\mathbf{b}_i$ the vector of **random effects** $\mathbf{b}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{B}_1)$

Joint models for longitudinal data and survival event(s)

The joint model : in two parts

$$\begin{cases} Y_i(t) = m_i(t) + \epsilon_i(t) = \mathbf{X}_{Li}(t)^\top \beta_L + \mathbf{Z}_i(t)^\top \mathbf{b}_i + \boldsymbol{\epsilon_i(t)} & \text{(biomarker)} \\ \lambda_i(t|\mathbf{b}_i) = \lambda_0(t)\ \exp(\mathbf{X}_{Ti}{}^\top \beta_T + h(\mathbf{b}_i, \beta_L, \mathbf{Z}_i(t), \mathbf{X}_{Li}(t))^\top \eta_T) & \text{(death)} \end{cases}$$

Where

- $\mathbf{X}_{Li}(t)$ and $\mathbf{X}_{Ti}$ are vectors of fixed effects covariates (β_L and β_T their coefficients)
- b_i the vector of random effects $\mathbf{b}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{B}_1)$
- $\boldsymbol{\epsilon_i(\cdot)}$, **measurements errors**, *iid* normally distributed with mean 0 and variance σ^2_ϵ

Joint models for longitudinal data and survival event(s)

$$\begin{cases} Y_i(t) = m_i(t) + \epsilon_i(t) = \mathbf{X}_{Li}(t)^\top \beta_L + \mathbf{Z}_i(t)^\top \mathbf{b}_i + \epsilon_i(t) & \text{(biomarker)} \\ \lambda_i(t|\mathbf{b}_i) = \lambda_0(t)\ \exp(\mathbf{X}_{Ti}{}^\top \beta_T + \mathbf{h}(\mathbf{b}_i, \beta_L, \mathbf{Z}_i(t), \mathbf{X}_{Li}(t))^\top \eta_T) & \text{(death)} \end{cases}$$

The link between the two processes (or the two sub-models) :

- by **random effects** b_i
- and **link functions h(.)** and their coefficients η_T (for the association strength)
- $h(.)$ can be :
  - directly b_i
  - the biomarker's current level $m_i(t)$
  - and/or the slope $\partial m_i(t)/\partial t$
  - Structure of association chosen a priori

Study case : Tumor evaluations in clinical trials

Other Research questions :

- Does the tumor size **and the appearance of new lesions** enable better prediction of OS ?

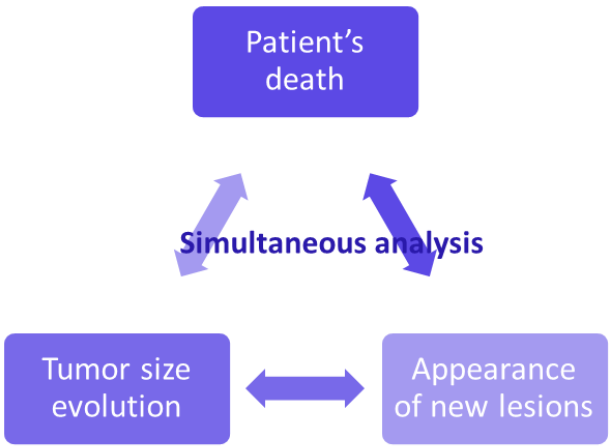
Solution : Work on a trivariate joint model

- for tumor size evolution
- for appearance of new lesions
- for a terminal event (OS)

Joint models for longitudinal data, recurrent events and survival event(s)

For individual i $(i = 1, \ldots, N)$ we observe :

- **Longitudinal biomarker** : $y_i(t_{ik})$
- **Recurrences** : $T_{ij} = \min(T_{ij}^*, C_i, T_i^*)$ and $\delta_{ij} = I_{\{T_{ij}^* = T_{ij}\}}$
- **Death** : $T_i = \min(C_i, T_i^*)$ and $\delta_i = I_{\{T_i^* = T_i\}}$



Joint model for longitudinal data, recurrent events and a terminal event

System of linear mixed-effects model and two hazard functions :

$$\begin{cases} Y_i(t) = \mathbf{X}_{Li}(t)^\top \boldsymbol{\beta}_L + \mathbf{Z}_i(t)^\top \mathbf{b}_i + \epsilon_i(t) & \text{(Biomarker)} \\ r_{ij}(t|v_i, \mathbf{b}_i) = r_0(t) \exp\left(v_i + \mathbf{X}_{Rij}^\top \boldsymbol{\beta}_R + g(\mathbf{b}_i, t)^\top \boldsymbol{\eta}_R\right) & \text{(Recurrences)} \\ \lambda_i(t|v_i, \mathbf{b}_i) = \lambda_0(t) \exp\left(\alpha v_i + \mathbf{X}_{Ti}^\top \boldsymbol{\beta}_T + h(\mathbf{b}_i, t)^\top \boldsymbol{\eta}_T\right) & \text{(Death)} \end{cases}$$

- $u_i = (\mathbf{b}_i^T, v_i)^T \sim \mathcal{N}(\mathbf{0}, \mathbf{B})$ with $\mathbf{B} = \begin{pmatrix} \mathbf{B_1} & \mathbf{0} \\ \mathbf{0} & \sigma_v^2 \end{pmatrix}$
- measurement errors *iid*, $\epsilon_i(t_{ik}) \sim \mathcal{N}(0, \sigma_\epsilon^2)$
- $g(\mathbf{b}_i, t)$ and $h(\mathbf{b}_i, t)$ - link functions
- $r_0(t)$, $\lambda_0(t)$ - baseline hazard functions

Joint model for longitudinal data, recurrent events and a terminal event

System of linear mixed-effects model and two hazard functions :

$$\begin{cases} Y_i(t) = \mathbf{X}_{Li}(t)^\top \boldsymbol{\beta}_L + \mathbf{Z}_i(t)^\top \mathbf{b}_i + \epsilon_i(t) & \text{(Biomarker)} \\ r_{ij}(t|v_i, \mathbf{b}_i) = r_0(t) \exp\left(v_i + \mathbf{X}_{Rij}^\top \boldsymbol{\beta}_R + g(\mathbf{b}_i, t)^\top \boldsymbol{\eta}_R\right) & \text{(Recurrences)} \\ \lambda_i(t|v_i, \mathbf{b}_i) = \lambda_0(t) \exp\left(\alpha v_i + \mathbf{X}_{Ti}^\top \boldsymbol{\beta}_T + h(\mathbf{b}_i, t)^\top \boldsymbol{\eta}_T\right) & \text{(Death)} \end{cases}$$

- $u_i = (\mathbf{b}_i^T, v_i)^T \sim \mathcal{N}(\mathbf{0}, \mathbf{B})$ with $\mathbf{B} = \begin{pmatrix} \mathbf{B_1} & \mathbf{0} \\ \mathbf{0} & \sigma_v^2 \end{pmatrix}$
- measurement errors *iid*, $\epsilon_i(t_{ik}) \sim \mathcal{N}(0, \sigma_\epsilon^2)$
- $g(\mathbf{b}_i, t)$ and $h(\mathbf{b}_i, t)$ - link functions
- $r_0(t)$, $\lambda_0(t)$ - baseline hazard functions

Estimation

- Joint marginal likelihood

$$L_i(\boldsymbol{\theta}) = \int_{\mathbf{u}_i} \prod_{k=1}^{n_i} f_{y|\mathbf{u}_i}(Y_i(t_{ik})|\mathbf{u}_i;\boldsymbol{\theta}) \prod_{j=1}^{r_i} f_{T^r|\mathbf{u}_i}(T_{ij}, \delta_{ij}|\mathbf{u}_i;\boldsymbol{\theta}) \cdot f_{T^t|\mathbf{u}_i}(T_i, \delta_i|\mathbf{u}_i;\boldsymbol{\theta}) f_{\mathbf{u}_i}(\mathbf{u}_i;\boldsymbol{\theta}) \mathrm{d}\mathbf{u}_i$$

  - l_i - number of biomarker measurement of individual i,
    n_i - number of recurrent events of individual i
  - Parameters to estimate
    $\boldsymbol{\theta} = (\boldsymbol{\beta}_L^\top, \boldsymbol{\beta}_R^\top, \boldsymbol{\beta}_T^\top, \boldsymbol{\eta}_R^\top, \boldsymbol{\eta}_T^\top, \alpha, r_0(\cdot), \lambda_0(\cdot), \mathbf{B}, \sigma_\epsilon)^\top$

- **Penalized maximum likelihood** estimation using Marquardt algorithm
- Baseline hazard functions approximation using **splines** : smooth estimation
- Integrals approximated using **Gauss-Hermite quadrature** : approach of iterated integrals and Genz algorithm (HRMSYM Fortran subroutine)

Goodness of fit

- for verification model assumptions
- in the context of survival data (recurrent and terminal) : Martingale residuals
- in the context of longitudinal data : residuals conditionnal on random effects or marginal resiuals

Goodness of fit for survival data

Martingale residuals :
the difference between the number of events of subject i until t and the Breslow estimator of the cumulative hazard function of t

$$M_i(t) = N_i(t) - \widehat{\mathbf{u}}_i \int_0^t W_i(s)\widehat{\zeta_i^{(p)}}(s)ds,$$

where, $W_i(t)$ is equal to 1 if the individual is at risk of the event at time t and 0 otherwise
$N_i(t)$ be the counting process of the event of type p (recurrent or terminal)
$\mathbf{u}_i$: random effects
$\mathbf{u}_i\zeta_i^{(p)}(t) = \mathbf{u}_i\zeta_0^{(p)}(t)\exp(\mathbf{X}_{pi}(t)^\top\beta_p)$process's intensity
$\hookrightarrow$ graphical visualisation (around zero)

Goodness of fit for longitudinal data

Raw residuals for checking homoscedasticity of the variances :

- marginal residuals averaged on the population level
  $\mathbf{R}_i^{(m)} = \mathbf{y}_i - \mathbf{X}_{Li}^\top \hat{\boldsymbol{\beta}}_L$
- conditional residuals, subject-specific
  $\mathbf{R}_i^{(c)} = \mathbf{y}_i - \mathbf{X}_{Li}^\top \hat{\boldsymbol{\beta}}_L - \mathbf{Z}_i^\top \hat{\mathbf{b}}_i.$

Goodness of fit for longitudinal data

Cholesky residuals (or decorrelated residuals) for checking normality assumption and detection of outlying observations :

$$\mathbf{R}_i^{(m)*} = \widehat{\mathbf{U}_i^{(m)}}\mathbf{R}_i^{(m)}, \qquad \mathbf{R}_i^{(c)*} = \widehat{\mathbf{U}_i^{(c)}}\mathbf{R}_i^{(c)}$$

where the raw residuals are multiplied by the upper-triangular matrices ($\widehat{\mathbf{U}_i^{(m)}}$ and $\widehat{\mathbf{U}_i^{(c)}}$) obtained by the Cholesky decomposition of the variance-covariance matrices

Goodness of fit $\hat{u}_i$

In the calculation of the residuals we need to estimate $\hat{u}_i$ using the formula for the **posterior probability function** :

$$f(\mathbf{u}_i|\mathbf{\Theta}_i;\widehat{\boldsymbol{\xi}}) = \frac{f(\mathbf{\Theta}_i|\mathbf{u}_i;\widehat{\boldsymbol{\xi}})f(\mathbf{u}_i;\widehat{\boldsymbol{\xi}})}{f(\mathbf{\Theta}_i;\widehat{\boldsymbol{\xi}})} \propto f(\mathbf{\Theta}_i|\mathbf{u}_i;\widehat{\boldsymbol{\xi}})f(\mathbf{u}_i;\widehat{\boldsymbol{\xi}}).$$

For the joint models, this expression does not have an analytical solution and the numerical computation is applied that finds such $\mathbf{u}_i$ that maximizes $f(\mathbf{u}_i|\mathbf{\Theta}_i;\widehat{\boldsymbol{\xi}})$:

$$\widehat{\mathbf{u}_i} = \arg\max_{\mathbf{u}_i} f(\mathbf{u}_i|\mathbf{\Theta}_i;\widehat{\boldsymbol{\xi}}),$$

Goodness of fit with Frailtypack : **martingale.res** and **martingaledeath.res**

```
plot(aggregate(readmission$t.stop, by = list(
    readmission$id),FUN = max)[2][ ,1],modJoint.gap
    $martingale.res, ylab='',xlab='time',main='
    Rehospitalizations',ylim=c(-4,4))

lines(lowess(aggregate(readmission$t.stop,by=list(
    readmission$id),FUN = max)[2][ ,1], modJoint.
    gap$martingale.res, f = 1), lwd=3,col='grey')
```

Goodness of fit with Frailtypack

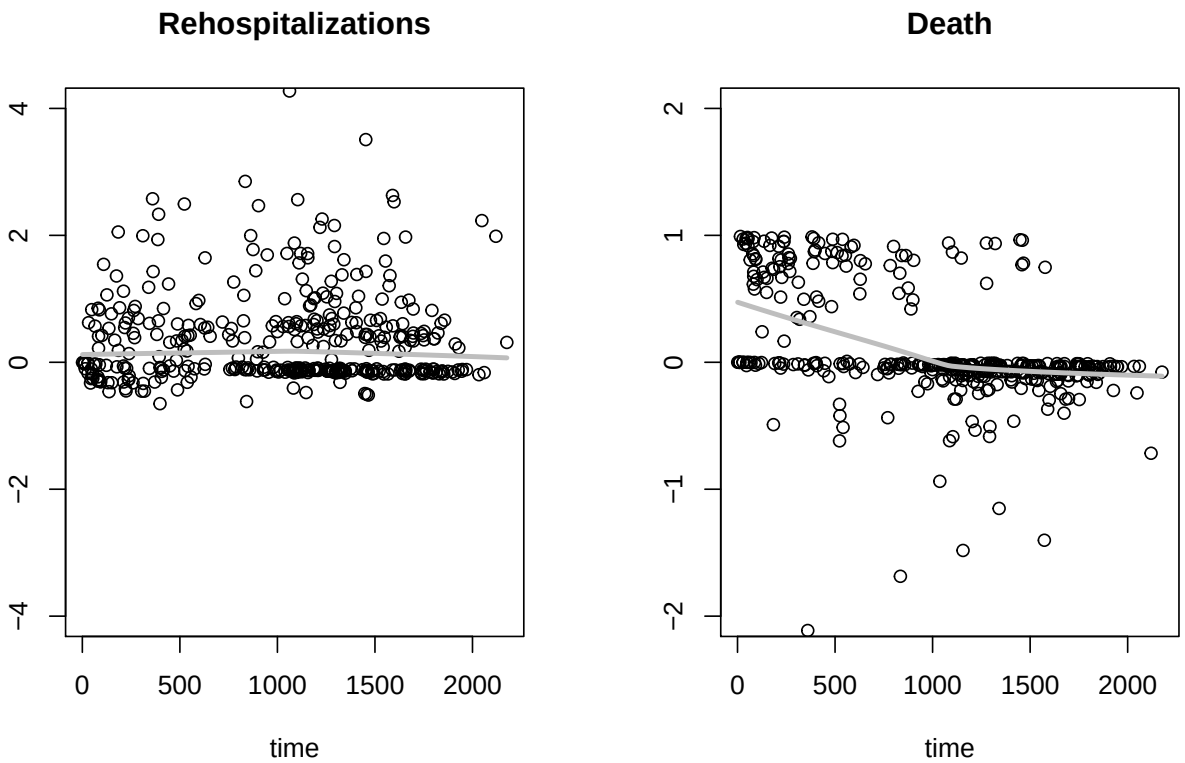
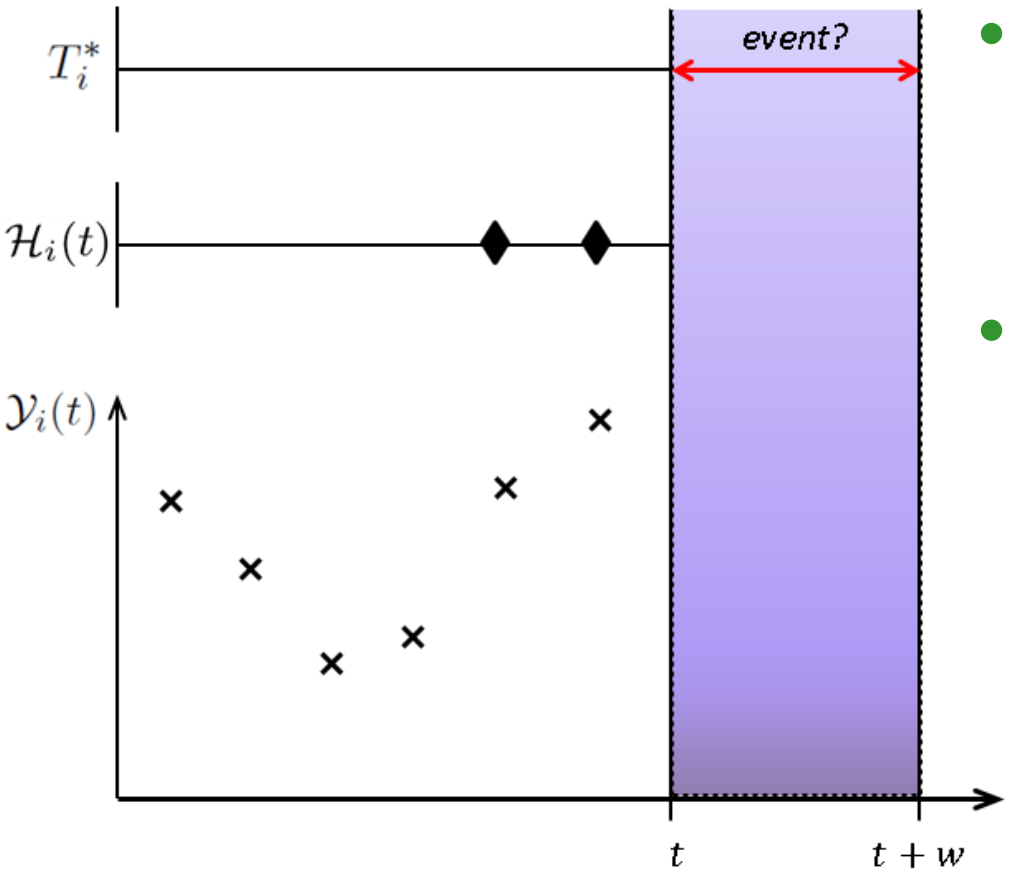


FIGURE: Martingale residuals for rehospitalizations and death against the follow-up time (in days). The grey line corresponds to a smooth curve obtained with lowess.

Dynamic predictions



- $\mathcal{H}_i(t)$ - history of recurrences of individual i until t
  $\mathcal{Y}_i(t)$ - history of the biomarker of individual i until t
- Predicted probability of terminal event T_i^* in a horizon $[t, t+w]$

$$\mathbb{P}(T_i^* \leq t+w | T_i^* > t, \mathcal{F}_i(t), X_i; \theta)$$

$\mathcal{F}_i(t) = \mathcal{H}_i(t)$,
$\mathcal{F}_i(t) = \mathcal{Y}_i(t)$
or $\mathcal{F}_i(t) = \{\mathcal{H}_i(t), \mathcal{Y}_i(t)\}$

Use two measures of predictive abilities for an internal validation

- Expected Prognostic Observed Cross-Entropy (EPOCE) *Commenges et al.*, 2012
  in a time window, "the lower the better"
  - Evaluation of conditional density of the event given the individual history
  - Internal validation : approximate cross-validated estimator CVPOL_a
- Brier score (with cross validation)
  - The inverse probability of censoring weighted error estimator (data-based Brier score) *Gerds and Schumacher*, 2006
  - comparison of predictions and actual observed events

EPOCE (Expected Prognostic Observed Cross-Entropy)

Commenges et al., 2012

- Risk of a estimator of a joint distribution, based on information theory and adjusted for right-censored data
- Approximated estimator (*Cross-Validated Prognostic Observed Log-Likelihood*)

$$CVPOL_a(s) = -\frac{1}{N_s}\sum_{i=1}^{N_s} 1_{\{T_i^* > s\}} l_{T_i|\mathcal{F}_i(s), T_i^* > s} + N \ \mathrm{Trace}(H^{-1}K_s)$$

l - conditional log-likelihood,
N_s - number of subjects still at risk at s,
$\mathcal{F}_i(s)$ - ith individual's history until s,
H - hessian of joint log-likelihood,
K_s - product of the gradients of the contributions to respectively the joint log-likelihood and the conditional log-likelihood

- Model comparison
  - in a time window, "the lower the better"

Brier Score

(*Gerds and Schumacher, 2006 ; Mauguen et al., 2013*)

- Inverse probability of censoring weighted error estimator

$$\hat{\mathrm{BS}}(t,w) = \frac{1}{N_t}\sum_{i=1}^{N_t}\left[I_{\{T_i^*>t+w\}} - (1 - P_{[t,t+w]}(\hat{\theta}))\right]^2 \omega_i$$

N_t - number of subjects at risk of the event at time t
$P_{[t,t+w]}(\hat{\theta}) = \mathbb{P}(T_i^* \leq t+w | T_i^* > t, \mathcal{F}_i(t), X_i; \hat{\theta})$

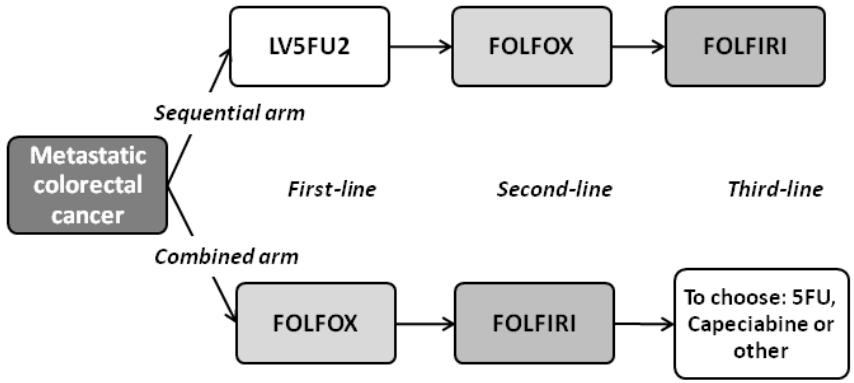
$$\omega_i = \frac{I_{\{T_i^* \leq t+w\}}\delta_i}{\hat{G}(T_i^*)/\hat{G}(t)} + \frac{I_{\{T_i^*>t+w\}}}{\hat{G}(t+w)/\hat{G}(t)}$$

$\hat{G}(t)$ - Kaplan-Meier estimate of the survival function of the censoring distribution at t

- Internal validation : 10-fold cross-validation

Clinical trial FFCD 2000-05

- Follow-up :
  - Phase III randomized multi-center clinical trial (53 centers in France), 407 patients



  - Tumor evaluation every 8 weeks, max 4 target lesions in 2 dimensions
  - Progression defined with the WHO criteria : more than 25% increase of one or more lesions observed and/or appearance of new lesions (on the best response obtained)
  - Change of line : progression, unacceptable toxicity, decision of investigator

Clinical trial FFCD 2000-05

- Objectives :
  - Which of **longitudinal biomarker**, **times of appearance of new lesions** or **times of progression** provide the most accurate **prediction** the **overall survival** ?
  - To identify the prognostic factors on the outcomes of interest
  - To evaluate the treatment effect

Data

- Biomarker definition : sum of the longest diameters

$$SLD_{ij} = \sum_{k=1}^{n_{ij}} d_{ijk}, \quad j = 0, 1, \ldots, n_i, \quad i = 1, \ldots, 407$$

$n_i \in \{0, 1, ..., 17\}$ - number of visits of individual i, $n_{ij} \in \{1, 2, 3, 4\}$ - number of target lesions measured during visit j, d_{ijk} - max diameter of lesion k measured during visit j of individual i

Data preparation

Biomarker transformation : in case of violation of the normality assumption

Popular classes of transformations :

- Box-Cox with parameter λ

$$y_i'(t_{ik}) = \begin{cases} (y_i(t_{ik})^\lambda - 1)/\lambda, & \lambda \neq 0 \\ \log(y_i(t_{ik})), & \lambda = 0 \end{cases}$$

- Logarithmic transformation with parameter α

$$y_i'(t_{ik}) = \log(y_i(t_{ik}) + \alpha)$$

$\hookrightarrow$ best values for λ or α obtained by profile likelihood over a grid of different values

Data : FFCD 2000-05

N=402 patients analyzed (53 centers in France)

- Observed : 6.18 tumor size measurements per patient
- 1.05 appearance of new lesions per patient
- 1.82 progression per patient
- 321 deaths

Application with R FRAILTYPACK

```
  library(frailtypack)

# Trivariate joint model for longitudinal data,
    recurrent events and a terminal event
 trivPenal(formula, formula.terminalEvent,
 formula.LongitudinalData, data,  data.Longi,
 random, id, intercept = TRUE,
link = ''Random-effects'', left.censoring = FALSE,
recurrentAG=FALSE, n.knots, kappa, maxit=300,
hazard=''Splines'',
init.B, init.Random, init.Eta, init.Alpha,
method.GH=''Standard'', n.nodes, LIMparam=1e-3,
LIMlogl=1e-3, LIMderiv=1e-3, print.times=TRUE)
```

Results of the trivariate model

Covariate	Biomarker : SLD Est. (SE)	p-value	New lesions HR (95% CI)	Death HR (95% CI)
Intercept	2.81 (0.28)	<0.001	-	-
Time	−0.29 (0.12)	0.012	-	-
Treatement (C/S)	−0.17 (0.14)	0.25	0.99 (0.77-1.28)	1.12 (0.66-1.91)
Treatement (C/S)×Time	−0.40 (0.15)	0.008	-	-
Age (60-69/<60 years)	0.22 (0.17)	0.20	0.80 (0.59-1.10)	1.10 (0.57-2.12)
Age (≥70/<60 years)	0.02 (0.16)	0.92	0.90 (0.66-1.21)	1.58 (0.84-2.98)
Sex (Women/Men)	0.27 (0.14)	0.05	0.90 (0.96-1.17)	1.06 (0.60-1.84)
Baseline WHO PS (1/0)	−0.11 (0.15)	0.46	1.18 (0.89-1.56)	1.87 (1.03-3.40)
Baseline WHO PS (2/0)	0.47 (0.21)	0.024	2.32 (1.53-3.51)	16.03 (6.66-38.59)
.....				

- significant decreasing value of SLD with time (-0.29) and **decreasing with** time more pronounced for the combination arm (-0.40)
- strong effect of **WHO performance** status 2 on the risk of death, new lesions and on repeated tumor size (larger tumor size)
- no significant associations with **gender and age**
- **smaller centers** had an increase risk of death

Results of the trivariate model

Association parameters (with the random effects)

Parameters*	Est. (SE)	p-value	**B Matrix**	Est. (SE)	p-value
η_{r1} (intercept)	0.18 (0.06)	0.005	Var(Interc.)	1.41 (0.06)	<0.001
η_{r2} (slope)	−0.07 (0.13)	0.58	Var(slope)	0.71 (0.07)	<0.001
η_{t1} (intercept)	0.96 (0.14)	<0.001	cov(Interc.,slope)	−0.15 (0.09)	0.131
η_{t2} (slope)	−0.07 (0.33)	0.83	σ_v^2	0.58 (0.09)	<0.001
α	2.74 (0.32)	<0.001			
σ_ϵ^2	1.26 (0.02)	<0.001			

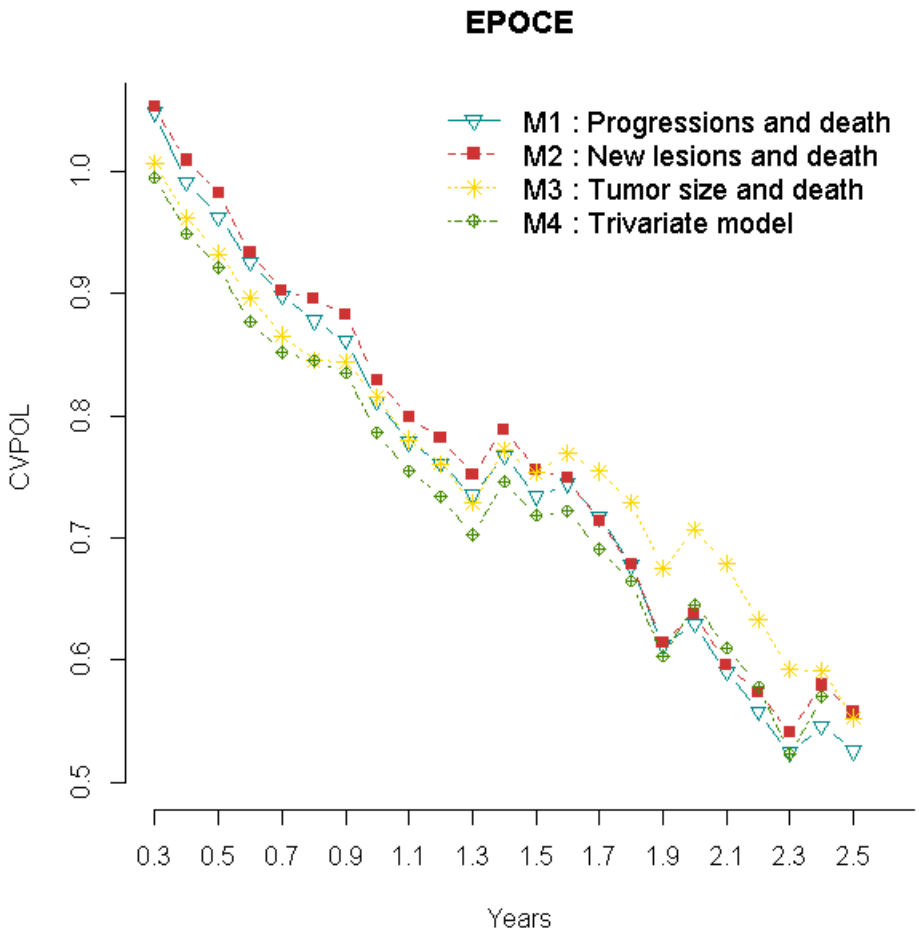
* η_{R1}, η_{R2} - link parameters (biomarker and recurrences)
η_{T1}, η_{T2} - link parameters (biomarker and death)
σ_v^2 - frailty variance, α - frailty power, σ_ϵ^2 - variance of meas. errors

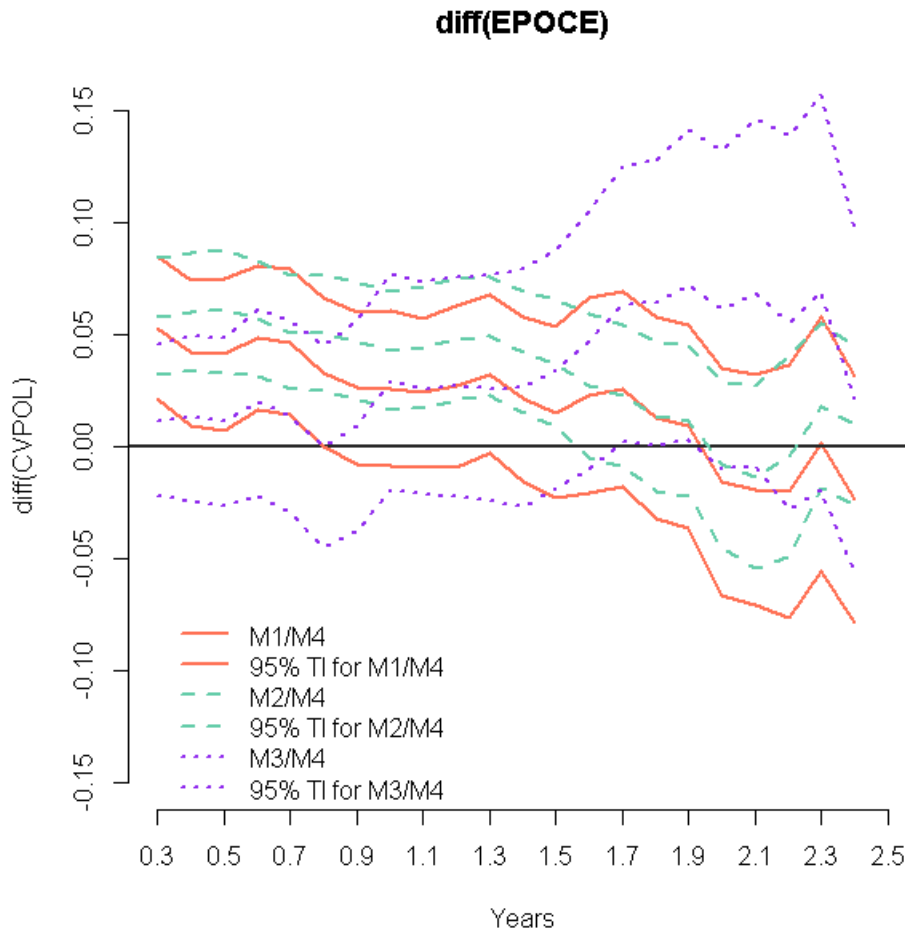
- strong positive association between tumor size and new lesions (0.18)
- strong positive association between tumor size and deaths (0.96)
- strong positive association between appearances of new lesions and death (0.58 and 2.74))

Comparison with the alternative models - predictive ability

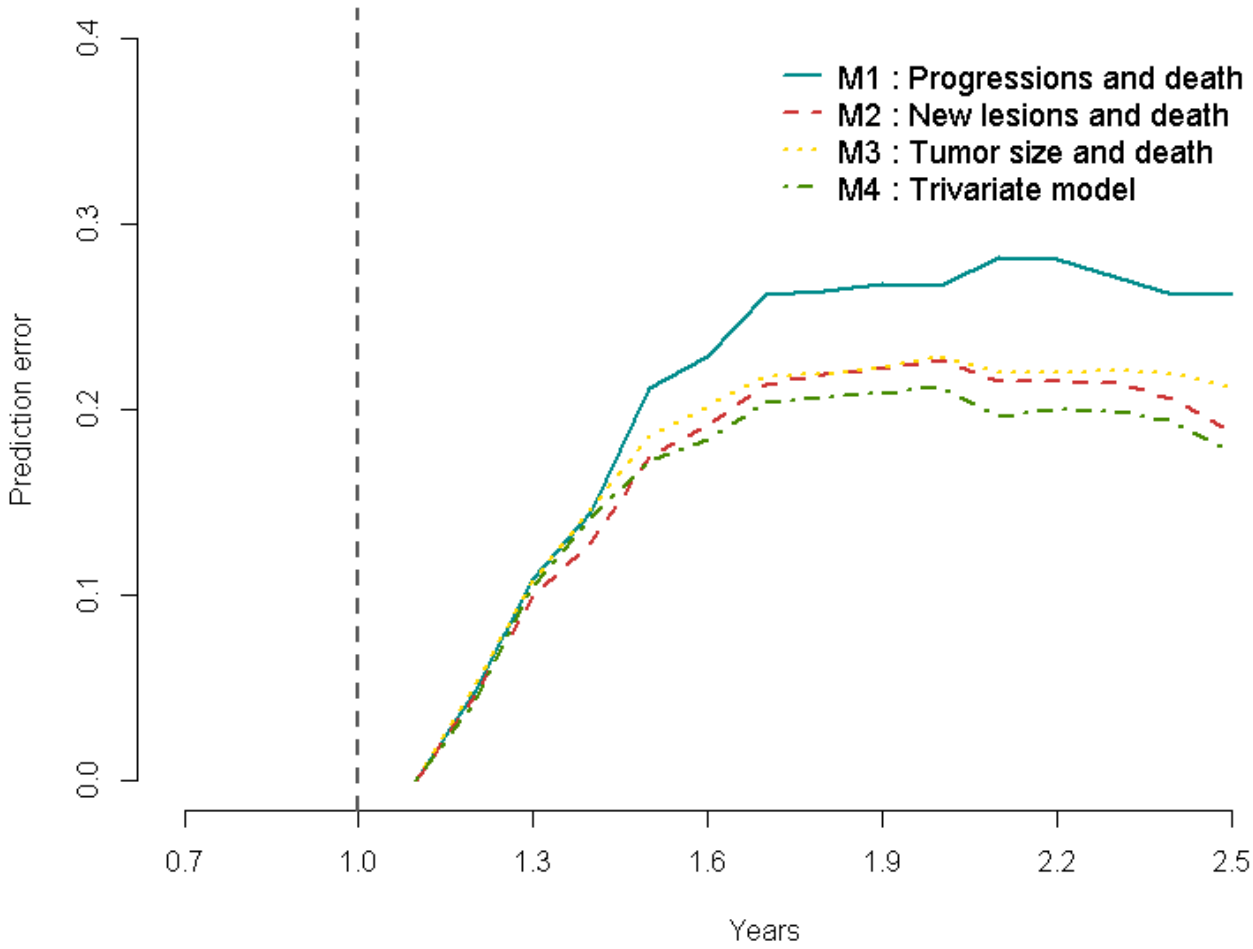
- Comparison of the models in terms of the **predictive ability** of the overall survival
  - Joint modelling of times of progression and time of death (M1)
  - Joint modelling of times of appearance of new lesions and time of death (M2)
  - Joint modelling of tumor size (SLD) and time of death (M3)
  - Joint modelling of tumor size (SLD), times of appearance of new lesions and time of death (M4)
- Measures of predictive ability using internal validation
  - **Brier score** (10-fold cross-validation)
  - **EPOCE** (CVPOL - approximated cross-validation)

Results - EPOCE





Results - Brier score



Implementation with Frailtypack

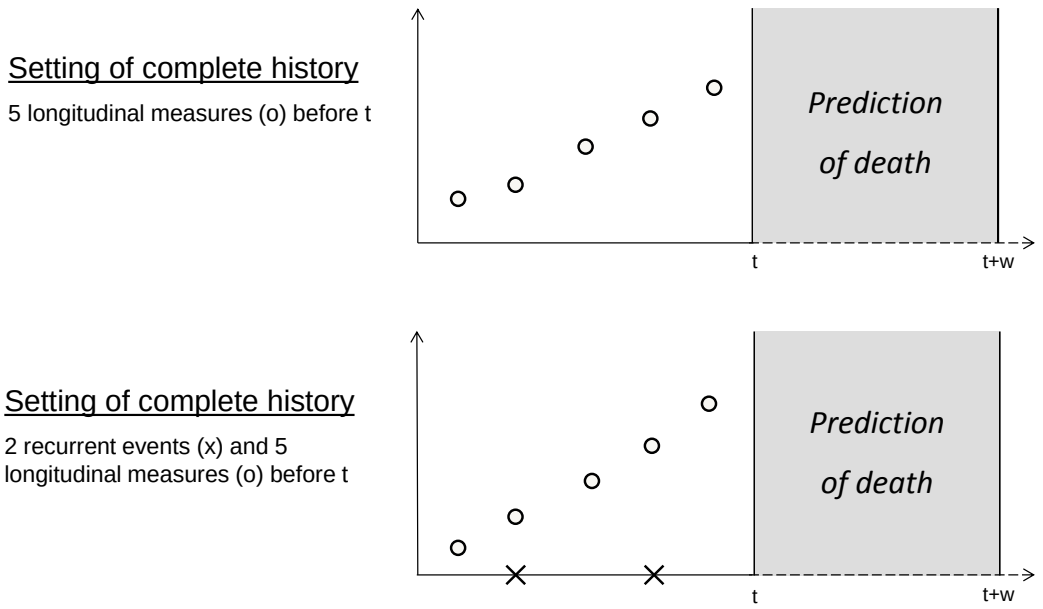


FIGURE: The possible prediction settings including the longitudinal data and considering the whole information available. The top setting correspond to the bivariate Model and the bottom graphic to the trivariate model.

Implementation with Frailtypack : prediction

```
# prediction on a TRIVARIATE JOINT model
#-- construction of the dataframe for predictions
#-- history of recurrences and terminal event
datapredj <- data.frame(time0=0,time1=0,
new.lesions=0,id=0,treatment=0,age=0,
who.PS=0,prev.resection=0)
datapredj$treatment<-as.factor(datapredj$treatmen)
levels(datapredj$treatment)<-1:2
datapredj$age<-as.factor(datapredj$age)
levels(datapredj$age)<-1:3
...
datapredj[1,]<-c(0,0.4,1,1,2,1,1,1)
datapredj[2,]<-c(0.4,1.2,1,1,2,1,1,1)
datapredj[3,]<-c(0,0.5,1,2,2,1,1,1)
```

Implementation with Frailtypack : prediction

```
# prediction on a TRIVARIATE JOINT model
#-- construction of the dataframe for predictions
#-- history of the biomarker observations
datapredj_longi<-data.frame(id=0,year=0,tumor.size
    =0,treatment=0,age=0,who.PS=0,prev.resection=0)
datapredj_longi$treatment<-as.factor(datapredj_
    longi$treatment)
levels(datapredj_longi$treatment)<-1:2
datapredj_longi$age<-as.factor(datapredj_longi$age
    )
levels(datapredj_longi$age)<-1:3
...
# patient 1: increasing tumor size
datapredj_longi[1,] <- c(1, 0,1.2 ,2,1,1,1)
datapredj_longi[2,] <- c(1,0.3,1.4,2,1,1,1)
datapredj_longi[3,] <- c(1,0.6,1.9,2,1,1,1)
```

Implementation with Frailtypack : prediction

```
#computation of the model (can be long)
model.spli.RE.cal <-trivPenal(Surv(time0, time1,
    new.lesions) ~ cluster(id)
+ age + treatment + who.PS +  terminal(state),
formula.terminalEvent =~ age + treatment + who.PS
    + prev.resection,tumor.size~year*treatment +
    age + who.PS, data=colorectal,
data.Longi=colorectalLongi,random=c('1','year'),id
    ='id',link='Random-effects',left.censoring
    =-3.33, recurrentAG=TRUE,n.knots=6,kappa=c
    (0.01,2),method.GH='Pseudo-adaptive',
n.nodes=7,init.B=c(-0.07,-0.13,-0.16,-0.17,0.42,#
    recurrent events covarates
-0.23,-0.1,-0.09,-0.12,0.8,-0.23,#terminal event
    covariates
3.02,-0.30,0.05,-0.63,-0.02,-0.29,0.11,0.74))#
    biomarker covariates
```

Implementation with Frailtypack : prediction

```
#-- prediction of death between 1 year and 1+2
    given history of the biomarker and recurrences
pred.jointTri0 <- prediction(model.spli.RE.cal,
    datapredj,
datapredj_longi, t = 1, window = 2)
print(pred.jointTri0)
#-- prediction of death between 1 year and 1+w
    given history of the biomarker and recurrences
pred.jointTri <- prediction(model.spli.RE.cal,
    datapredj,
datapredj_longi, t = 1, window = seq(0.5, 2.5,
    0.2), MC.sample = 100)
plot(pred.jointTri, conf.bands = TRUE)
# each y-value of the plot corresponds to the
    prediction between [1,x]
```

Conclusion

- Advantages of using joint models for simultaneous analysis of prognostic factors
- Comparison of joint models of different types in terms of predictive accuracy
- Proposition of a new trivariate joint model
- FFCD 2000-05 : Improvement of predictive abilities using tumor size and appearance of new lesions

References (book) on frailty models

- **Therneau** T, Grambsch P. Modeling survival data : extending the Cox model, 2000 *Springer-Verlag New York*
- **Hougaard** P. Analysis of Multivariate Survival Data, 2000 *Springer-Verlag New-York.*
- **Cook** RJ, Lawless JF. The statistical analysis of recurrent events. *Statistics for biology and health,* 2007, Springer.
- **Duchateau**, L. and Janssen, P. The frailty model, 2008, Springer.
- **Wienke** A. Frailty models in survival analysis, 2010, Chapman & Hall/CRC Biostatistics series.
- **Rondeau** V, Mazroui, Y. and Gonzalez, J.R. FRAILTYPACK : An R package for the analysis of correlated survival data with frailty models using penalized likelihood estimation, Journal of Statistical Software, 2012.
- *http ://cran.r-project.org/web/packages/frailtypack/*

References (book) on Joint Models

- **Rizopoulos** D. Joint Models for Longitudinal and Time-to-Event Data : With Applications in R, 2012 *Chapman and Hall/CRC.*
- **Commenges**, D. and Jacqmin-Gadda, H. and Amadou, A. and Joly, P. and Liquet, B. and Proust-Lima, C. and Rondeau, V. and Thiebaut, R.,2015 Dynamical biostatistical models, 2012 *Chapman and Hall/CRC.*
- **Elashoff**, RM, Gang Li, Ning Li, 2016, Joint modeling of longitudinal and time-to-event data *Chapman and Hall/CRC.*