

Analyse de la Survie Nette et de la Mortalité en Excès Des concepts et des méthodes

Roch Giorgi, Nathalie Grafféo

15^{ème} Conférence Francophone d'EPIdémiologie CLINique

EPICLIN2021 / JSCLCC28

28^{èmes} Journées des Statisticiens des Centres de Lutte Contre le Cancer



Plan

- Partie I
 - Contexte
 - Concepts : quelle mesure pour répondre à quelle question ?
 - Méthode d'estimation non-paramétrique de la survie nette
 - Application
- Partie II
 - Méthodes d'estimation paramétrique de la mortalité en excès
 - Application

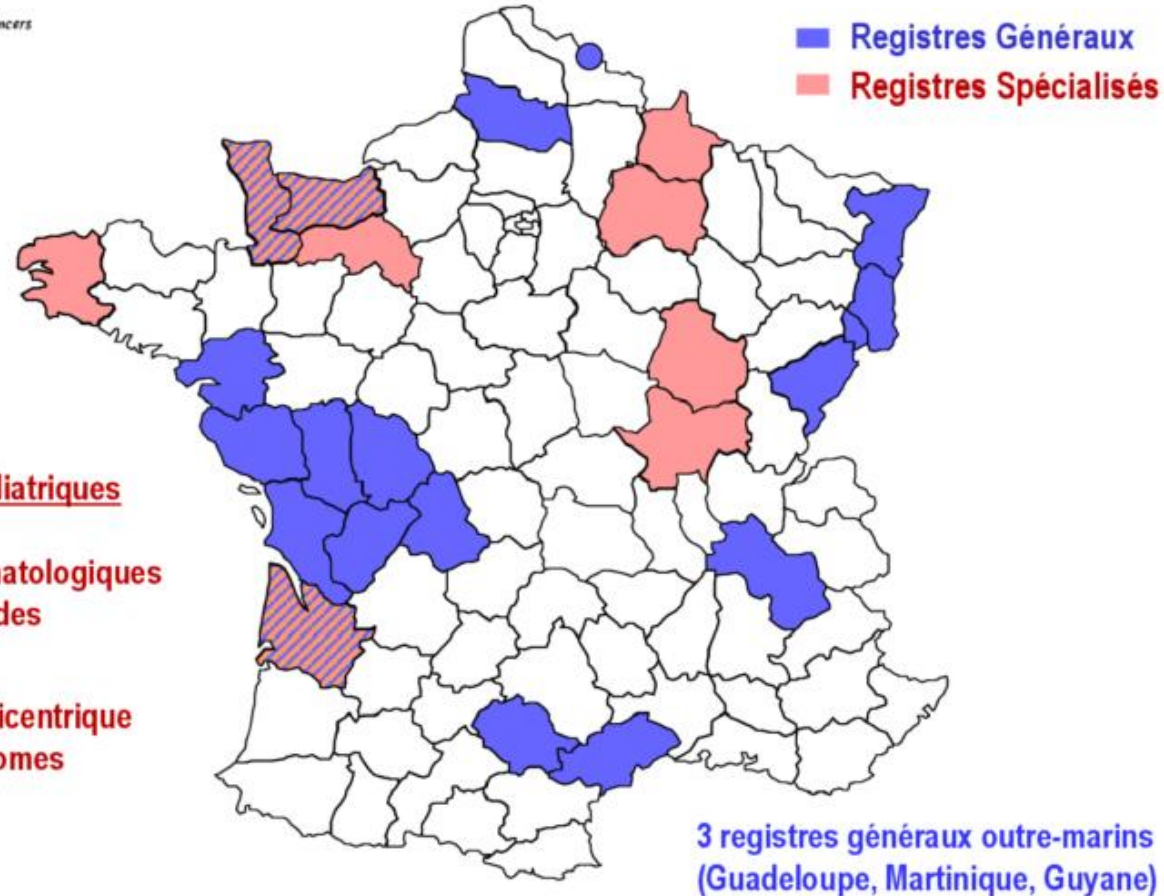
Partie I

Contexte

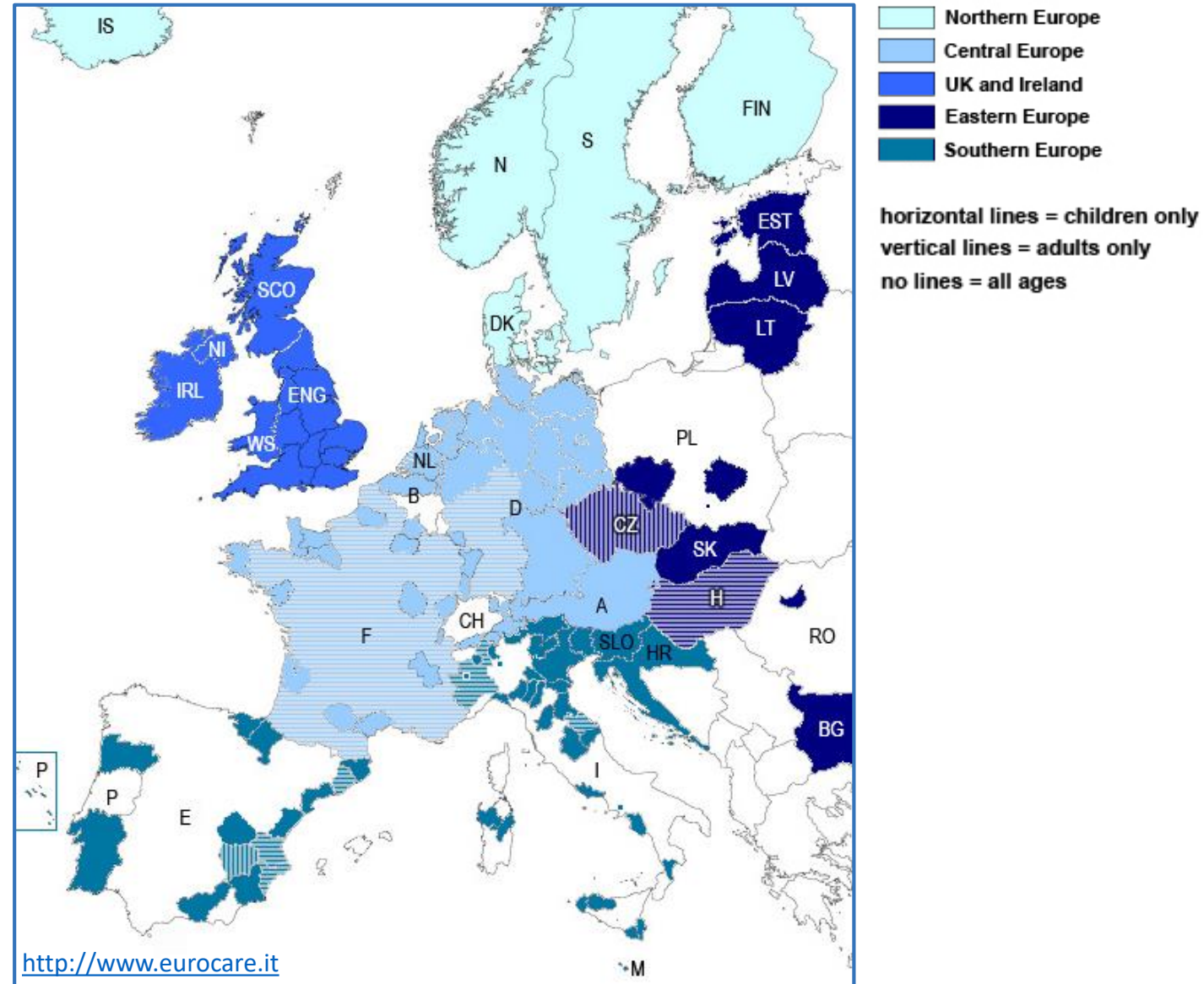
Francim : France-Cancer-Incidence et Mortalité

D'après : [Accueil - Registre des cancers | Loire-Atlantique et Vendée - Association EPIC-PL \(registre-cancers-44-85.fr\)](#)

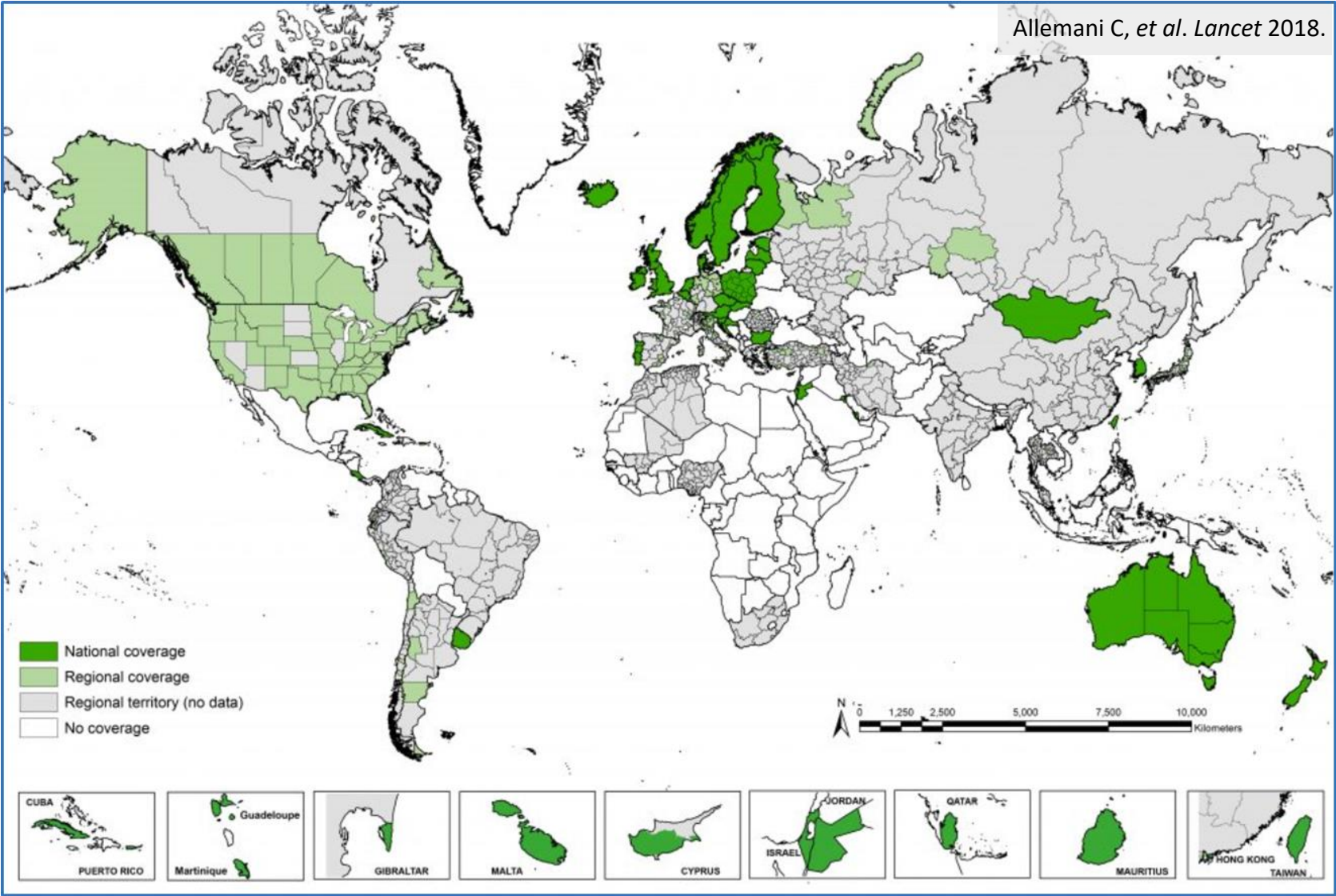
Les Registres de cancers en 2019



EUROCARE-5 : EUROpean CAncer REgistry



CONCORD-3



Données recueillies

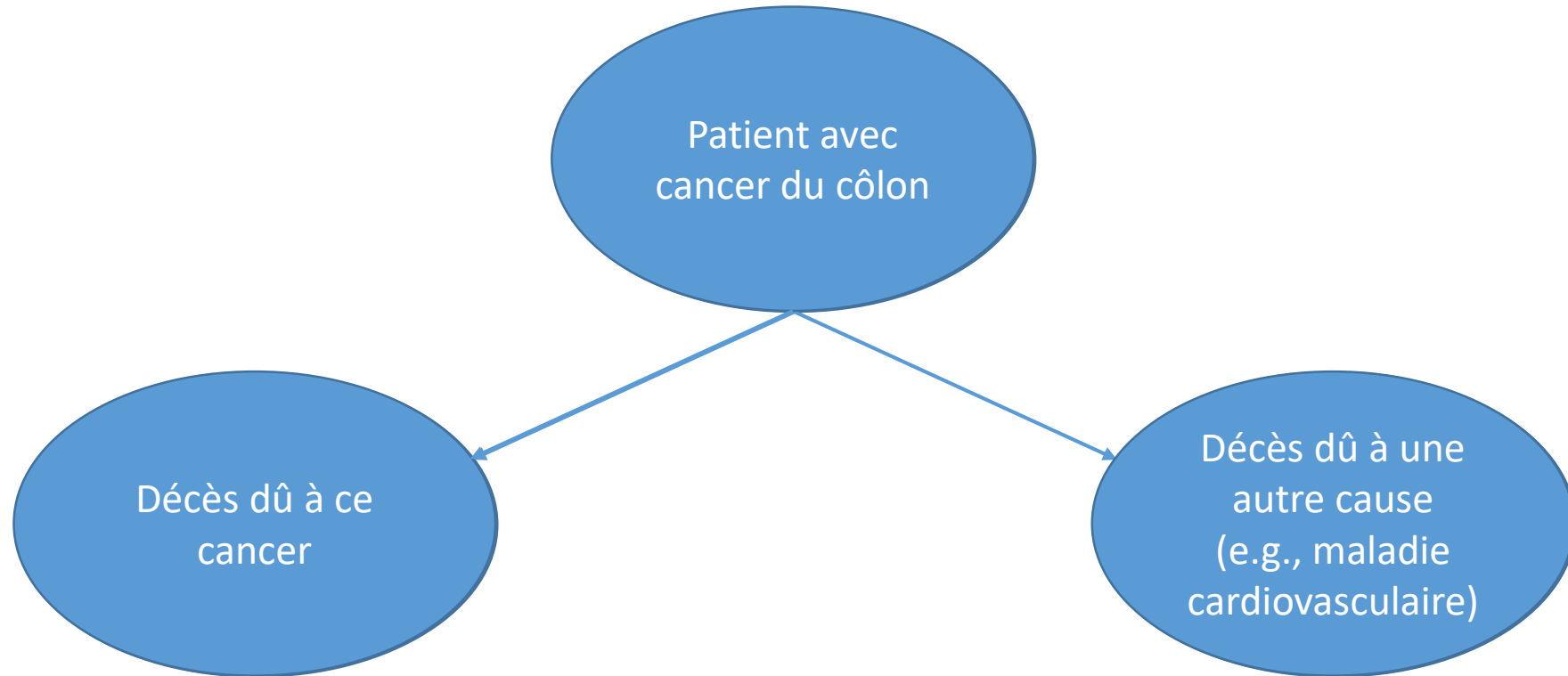
- Collectées en routine (*la plupart du temps*)
 - Date de diagnostic
 - Date de naissance
 - Date de décès (suivi actif)
 - La cause de décès est souvent inconnue ou peu fiable
 - L'attribution de cette cause peut être subjective, avec peu de possibilités d'être comparable entre régions ou pays
 - Sexe
 - Localisation de la tumeur
 - TNM : Taille de la tumeur, invasion ganglionnaire, métastase (*pas toujours*)
- Collectées pour des études spécifiques (études haute résolution)
 - récepteur d'œstrogène, catégorie socio-professionnelle, ...

Exemple de données (minimales) disponibles

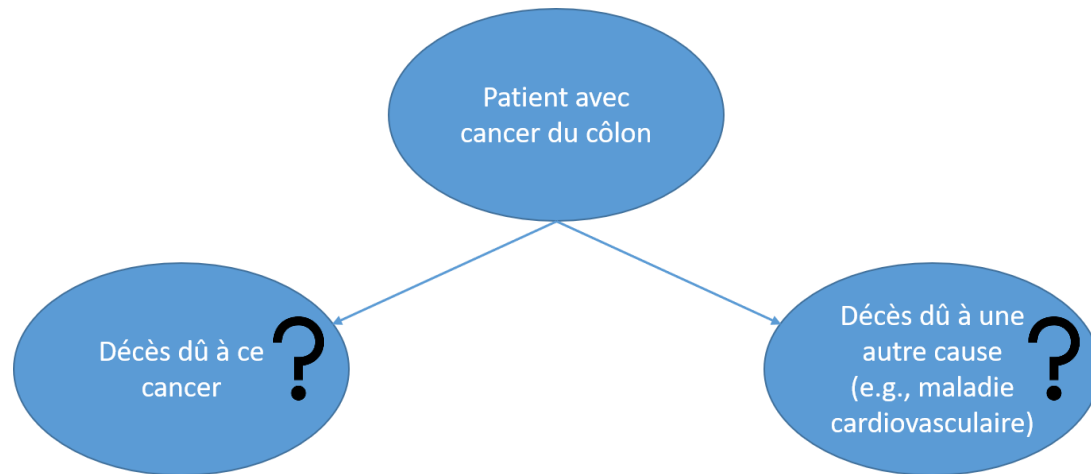
id	Temps	Statut	Age	Sexe	Année diag.
1	1645	0	65	F	2015
2	543	1	76	M	2016
3	2876	0	71	F	2015

- **Id** : identifiant du patient
- **Temps** : temps de suivi (ici, en jours)
- **Statut** : 1 si décédé, 0 sinon
- **Age** : âge au diagnostic
- **Sexe** : sexe (ici, F ou M. D'autres codages sont possibles)
- **Année diag.** : date de diagnostic (ici, année seulement. Précision et codage peuvent différer)

Des risques concurrents



Le cadre de la survie relative : définition



- Cause de décès peu fiable ou inconnue
- Décès dû au cancer étudié ou à une autre cause

👉 Le cadre de la survie relative est basé sur les risques concurrents mais est appliqué dans des études de population où la cause de décès est inconnue

Le cadre de la survie relative : hypothèse 1

Patient avec cancer
du côlon
Mortalité observée
 $\lambda_{0,i}(t|X_i)$

Décès dû à ce cancer
Mortalité en excès
 $\lambda_{E,i}(t|X_i)$

Décès dû à une autre
cause
Mortalité « autres
causes »
 $\lambda_{P,i}(t|X_{Di})$

$$\lambda_{0,i}(t|X_i) = \lambda_{E,i}(t|X_i) + \lambda_{P,i}(t|X_{D,i})$$

Le cadre de la survie relative : hypothèse 2

- Mortalité attendue dans la population générale
- Obtenue à partir des tables de mortalité
- **Tables de mortalité appropriées** : on fait l'hypothèse que ces tables de mortalité reflètent la mortalité due aux autres causes de décès que la maladie étudiée

Décès dû à une autre cause

Mortalité « autres causes »

$$\lambda_{P,i}(t|X_{D,i})$$

Les tables de mortalité (1/3)

- Issues des statistiques officielles nationales
- Stratifiées par des variables démographiques
 - Âge, sexe, année, région, département
 - Ethnie (USA, Nouvelle-Zélande)
 - Indicateur de déprivation (Royaume-Uni)

Les tables de mortalité (2/3)

- Exemple issu du package `survival` dans R

```
survexp.usr[78:82,,, '2004']  
Rate table with dimension(s): age sex race  
, , race = white  
  
      sex  
age      male      female  
77 0.0001476044 0.0001010906  
78 0.0001631797 0.0001127686  
79 0.0001803468 0.0001257690  
80 0.0001992560 0.0001402259  
81 0.0002200673 0.0001563012
```

```
, , race = black  
  
      sex  
age      male      female  
77 0.0001913084 0.0001253706  
78 0.0002061135 0.0001368304  
79 0.0002220203 0.0001493098  
80 0.0002391014 0.0001628950  
81 0.0002574355 0.0001776724
```

Les tables de mortalité (3/3)

- Exemple issu du package survival dans R
- Probabilité de survie de 77 à 78 ans pour un homme blanc ayant 77 ans en 2004 ?

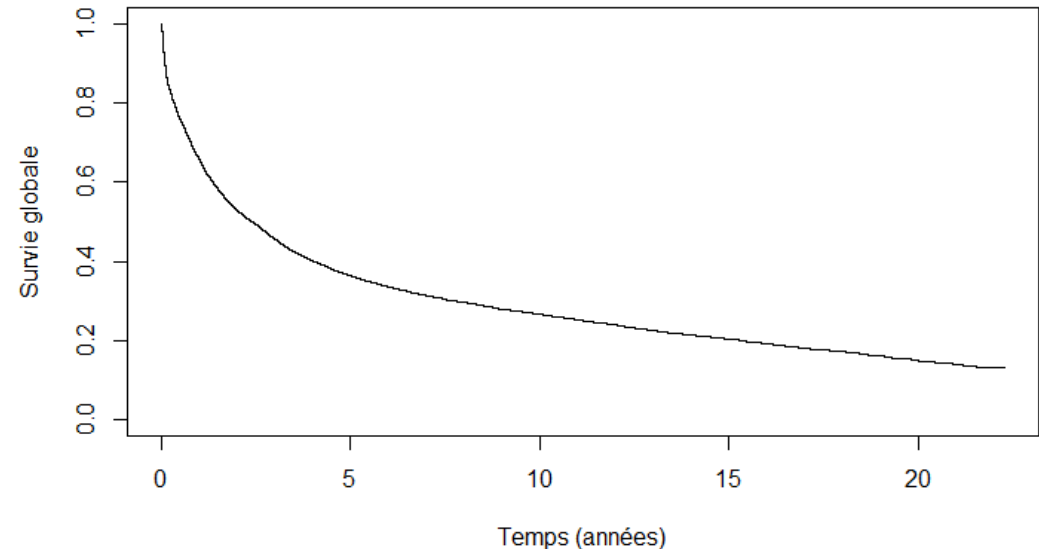
```
dhr <- survexp.usr[78, 'male', 'white', '2004']  
exp(-dhr*365.25)  
0.947515
```


Concepts : quelle mesure pour
répondre à quelle question ?

Quelle est la probabilité de survivre un temps t après le diagnostic ?

Quelle est la probabilité de survivre un temps t après le diagnostic ?

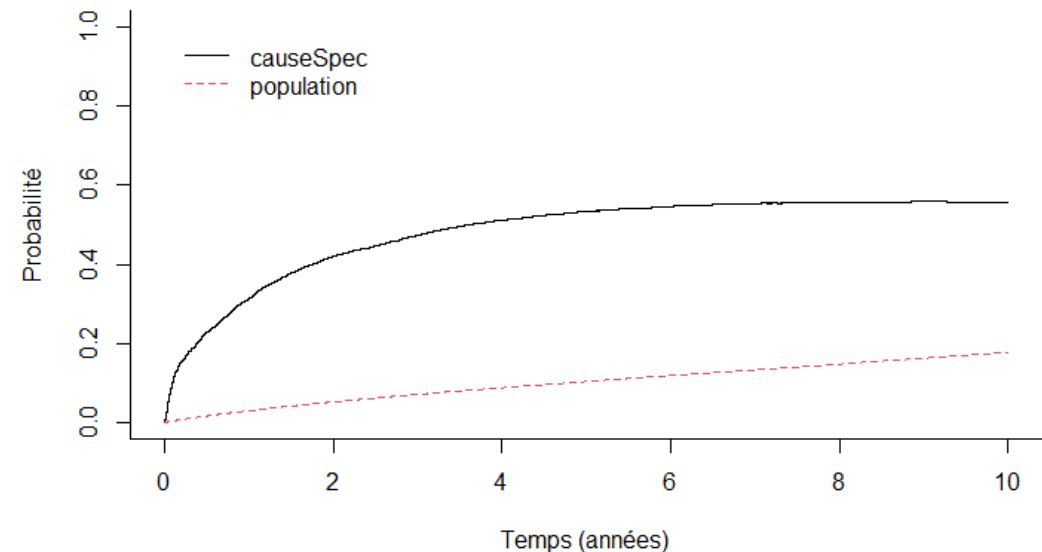
- Survie observée/globale (S_o)
- Notation : T v.a. mesurant le délai du diagnostic au décès
- $S_{o,i}(t) = P(T_i \geq t)$
- $S_{o,i}(t) = \exp\left(-\Lambda_{o,i}(t)\right)$
 $= \exp\left(-\int_0^t \lambda_{o,i}(u) du\right)$



Quelle est la probabilité de décéder du cancer étudié au t après le diagnostic ?

Quelle est la probabilité de décéder du cancer étudié au t après le diagnostic ?

- Fonction d'incidence cumulée (ou probabilité brute)
- $F_K(t) = P(T \leq t, \text{cause} = K)$
- $1 - S_O(t) = F_K(t) + F_{\bar{K}}(t)$
- $F_K(t) = \int_0^t S_O(u-) \lambda_E(u) du$
 - $1 - F_K$ n'est pas une proba. de survie ($\lim_{t \rightarrow \infty} F_K(t) \neq 0$)
 - F_K dépend de Λ_E et Λ_P



Comment comparer la mortalité en excès due au cancer étudié entre plusieurs pays/périodes indépendamment de la mortalité de la population générale ?

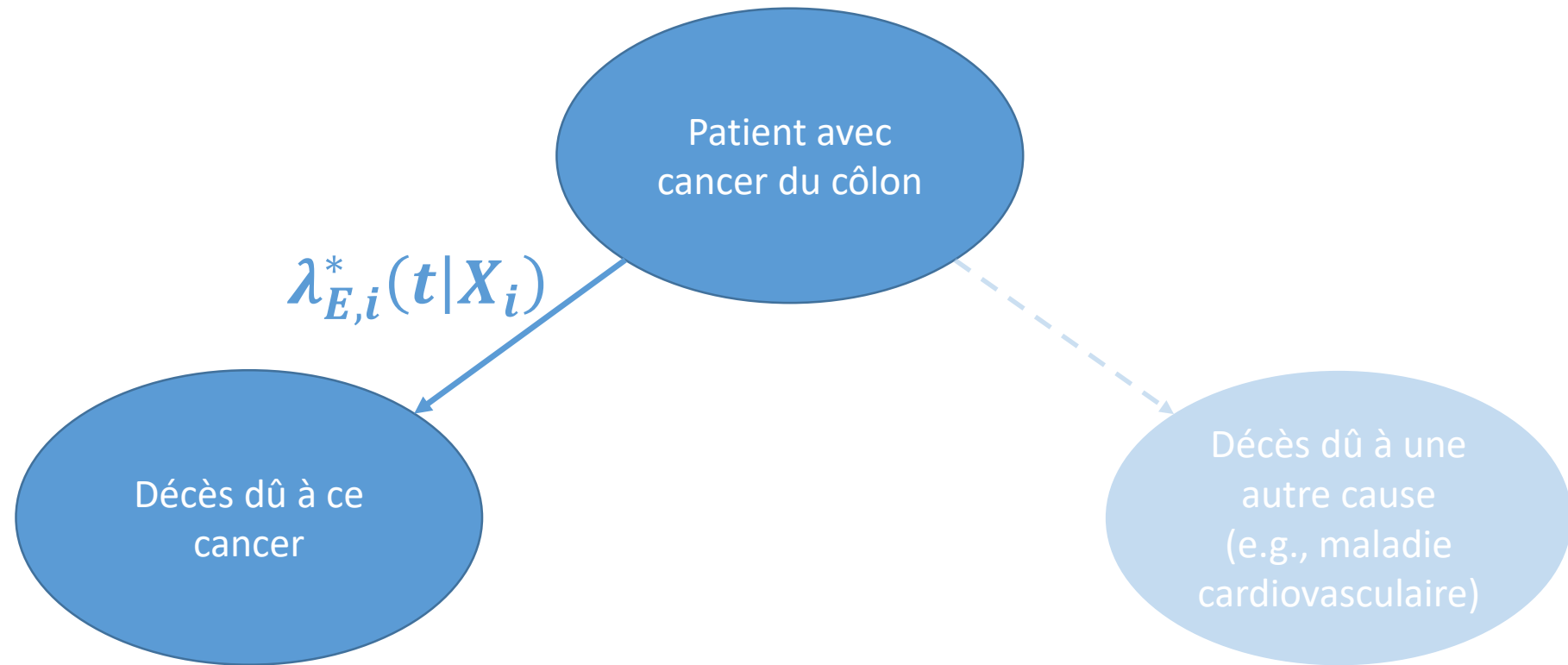
Comment comparer la mortalité en excès due au cancer étudié entre plusieurs pays/périodes indépendamment de la mortalité de la population générale ?

- Survie nette individuelle : $S_{N,i}(t) = \exp\left(-\int_0^t \lambda_{E,i}(u) du\right)$
- Ne dépend que de $\lambda_{E,i}$ (et pas de $\lambda_{P,i}$)
- $S_{N,i}(t) = \exp\left(-\int_0^t \left(\lambda_{O,i}(u) - \lambda_{P,i}(u)\right) du\right) = \frac{S_{O,i}(t)}{S_{P,i}(t)}$
- Survie nette populationnelle (échantillon) : $S_N(t) = \frac{1}{n} \sum_{i=1}^n \frac{S_{O,i}(t)}{S_{P,i}(t)}$

Comment comparer la mortalité en excès due au cancer étudié entre plusieurs pays/périodes indépendamment de la mortalité de la population générale ?

- **Interprétation 1** : la survie nette populationnelle est la *moyenne des ratios individuels* de la survie observée du patient sur la survie issue des tables de mortalité d'un individu avec les mêmes caractéristiques démographiques

Comment comparer la mortalité en excès due au cancer étudié entre plusieurs pays/périodes indépendamment de la mortalité de la population générale ?




Comment comparer la mortalité en excès due au cancer étudié entre plusieurs pays/périodes indépendamment de la mortalité de la population générale ?

- **Interprétation 1** : la survie nette populationnelle est la *moyenne des ratios individuels* de la survie du patient sur la survie issue des tables de mortalité d'un individu avec les mêmes caractéristiques démographiques
- **Interprétation 2** : la survie nette est la survie que l'on observerait, *dans un monde hypothétique*, où la seule cause possible de décès serait le cancer étudié

Comment comparer la survie observée des patients à celle d'un groupe sans ce cancer ayant la même structure démographique ?

Comment comparer la survie observée des patients à celle d'un groupe sans ce cancer ayant la même structure démographique ?

- Ratio de survie relative (S_R)
- $S_R(t) = \frac{S_O(t)}{S_P(t)}$ défini au niveau populationnel (échantillon)
- On compare la survie globale à la survie de la population générale pour une population ayant les mêmes caractéristiques démographiques


$$S_R(t) = \frac{\frac{1}{n} \sum_{i=1}^n S_{O,i}(t)}{\frac{1}{n} \sum_{i=1}^n S_{P,i}(t)} \neq \frac{1}{n} \sum_{i=1}^n \frac{S_{O,i}(t)}{S_{P,i}(t)} = S_N(t)$$



Extrait de Perme, Maja Pohar, Jacques Estève, and Bernard Rachet. "Analysing population-based cancer survival—settling the controversies." *BMC cancer* 16.1 (2016): 1-8.

"We mimic real data to create two data sets A and B with same demographic distributions, and assume that

- (i) the cohorts were diagnosed in two different calendar years (1990 and 2000) and thus had a different general population mortality rate λ_{Pi} ;
- (ii) but the true, underlying cancer prognosis did not improve between periods A and B, i.e. the cancer-specific hazards λ_{Ci} are the same. The data sets are large (20000 patients each) to enable clear distinction between the random variation and true differences."

Extrait de Perme, Maja Pohar, Jacques Estève, and Bernard Rachet. "Analysing population-based cancer survival—settling the controversies." *BMC cancer* 16.1 (2016): 1-8.

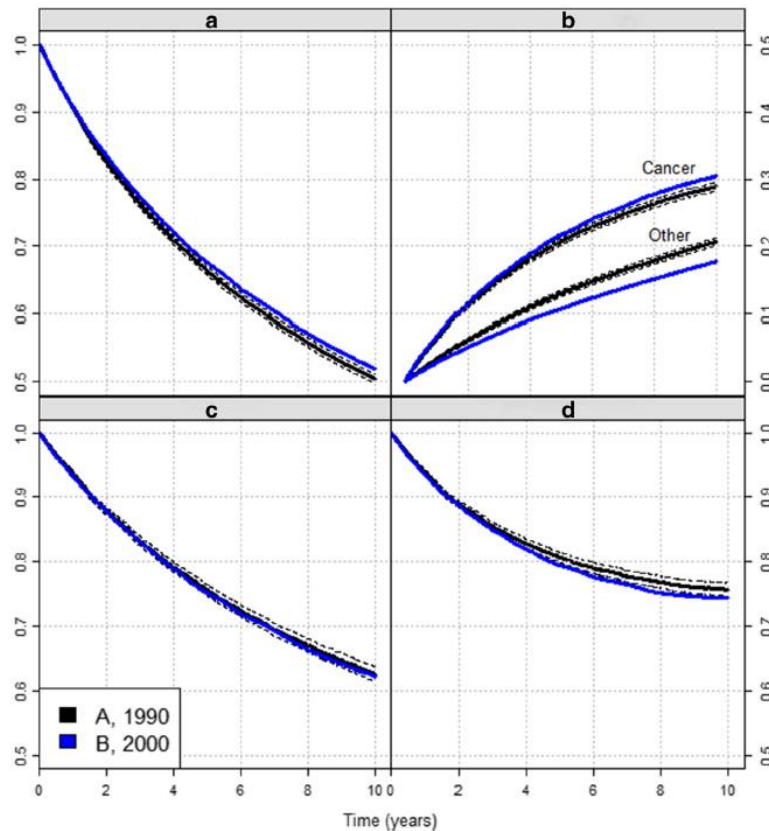


Fig. 2 Comparing two cohorts with the different measures. Colon cancer survival (simulated data): a) Overall survival, b) Crude mortality, c) Net survival, d) Relative survival ratio. Black curves: cohort A, diagnosed 1990; blue curves: cohort B, diagnosed 2000. Dashed lines denote the 95% confidence intervals for cohort A

- Données générées de sorte que
 - $\lambda_{E,1990} = \lambda_{E,2000}$
 - Diminution de la mortalité « autres causes » entre 1990 et 2000
- **Interprétation du cadran a ?**
- **Interprétation du cadran b ?**
- **Interprétation du cadran d ?**

Méthodes d'estimation non-paramétriques de la survie nette

Notions de processus

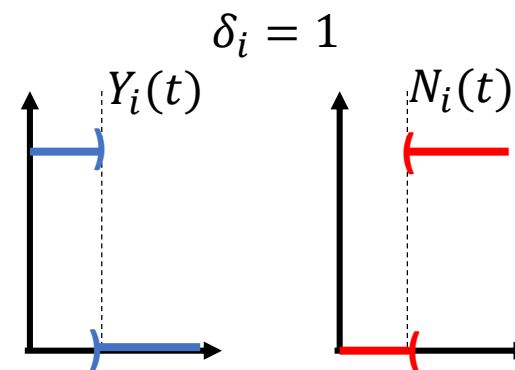
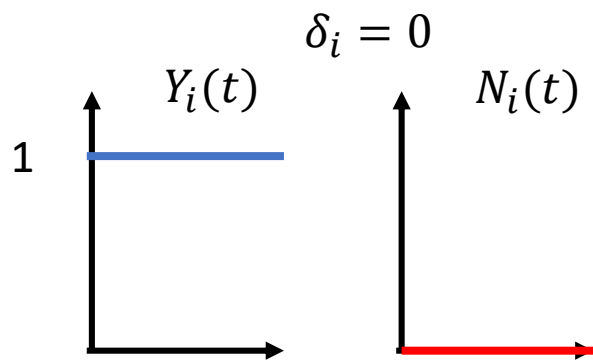
- Nombre à risque (« risk set ») :

$$Y(t) = \sum_{i=1}^n Y_i(t), \text{ où } Y_i(t) = \mathbb{1}(T_i \geq t, C_i \geq t)$$

- Nombre de décès :

$$N(t) = \sum_{i=1}^n N_i(t), \text{ où } N_i(t) = \mathbb{1}(T_i \leq t, T_i \leq C_i)$$

- Indicatrice de décès : $\delta_i = \mathbb{1}(T_i \leq C_i)$



Survie observée

- $S_o(t) = P(T > t) = \exp\left(-\int_0^t \lambda_o(u) du\right)$
- Estimateur de Kaplan-Meier :

$$\widehat{S}_o(t) = \prod_{u \leq t} \left(1 - \frac{dN(u)}{Y(u)}\right)$$

- Dans R, package `survival`

```
survfit(Surv(time, status) ~ 1, data = mydata)
```

- Estimateur de Nelson-Aalen :

$$\widehat{\Lambda}_o(t) = \int_0^t \frac{dN(u)}{Y(u)}$$

Estimateur Pohar-Perme de la survie nette (1/4)

- $S_N(t) = \exp\left(-\int_0^t \lambda_E(u) du\right)$
- Analogie avec Nelson-Aalen ?

$$\widehat{\Lambda}_E(t) = \int_0^t \frac{dN_E(u)}{Y(u)} \quad ?$$

- Problèmes :
 - Calcul de dN_E en l'absence de la cause de décès ?
 - Comment corriger le problème de censure informative ?

Estimateur Pohar-Perme de la survie nette (2/4)

- **Estimation de dN_E en l'absence de la cause de décès ?**
- Intuitivement, différence entre le nombre de décès observés et le nombre de décès que l'on attendrait due à la mortalité populationnelle
- $E(dN_{P,i}(t)|Y_i(t)) = Y_i(t)\lambda_{P,i}(t)dt$
- $d\widehat{N}_E(t) = \sum_{i=1}^n Y_i(t)dN_i(t) - \sum_{i=1}^n Y_i(t)\lambda_{P,i}(t)dt$

Estimateur Pohar-Perme de la survie nette (3/4)

- **Comment corriger le problème de censure informative ?**
- La mortalité due au cancer étudié et la mortalité « autres causes » partagent l'influence des mêmes variables démographiques (e.g. âge)
- Procédé de pondération IPW (Inverse Probability Weighting) pour corriger ceux qui sortent « prématurément » de l'étude du fait d'un décès « autres causes »

$$- N^w(u) = \sum_{i=1}^n N_i^w$$

$$- Y^w(u) = \sum_{i=1}^n Y_i^w$$

$$- \text{Avec } dN_i^w(u) = \frac{dN_i(u)}{S_{P,i}(u)} \text{ et } Y_i^w = \frac{Y_i(u)}{S_{P,i}(u)}$$

Estimateur Pohar-Perme de la survie nette (4/4)

- L'estimateur Pohar-Perme est un estimateur consistant de la survie nette

$$\widehat{\Lambda}_E(t) = \int_0^t \frac{dN^w(u)}{Y^w(u)} - \int_0^t \frac{\sum_{i=1}^n Y_i^w(u) \lambda_{P,i}(u) du}{Y^w(u)}$$

- Dans R, package `relsurv`

```
rs.surv(Surv(time, status)~1, data = mydata,  
        ratetable = slopop, method = ``pohar-perme``)
```

Comparaison de courbes de survie globale entre $K \geq 2$ groupes sur $[0, T]$

- $(H_0): \forall t \in [0, T], \Lambda_1(t) = \Lambda_2(t) = \dots = \Lambda_K(t)$
- Distribution asymptotique sous $(H_0): \chi^2(K - 1)$
- Test stratifié en cas d'hétérogénéité

- Dans R, package `survival`

```
survdiff(Surv(time, status) ~ X1, data = mydata)
# si besoin d'un test stratifié sur sur une variable X2
survdiff(Surv(time, status) ~ X1 + strata(X2), data = mydata)
```

Comparaison de courbes de survie nette entre $K \geq 2$ groupes sur $[0, T]$

- (H_0) : $\forall t \in [0, T], \Lambda_{E_1}(t) = \Lambda_{E_2}(t) = \dots = \Lambda_{E_K}(t)$
- Distribution asymptotique sous (H_0) : $\chi^2(K - 1)$
- Test stratifié en cas d'hétérogénéité
- Dans R, package `relsurv`

```
rs.diff(Surv(time, status) ~ X1, data = mydata,  
        ratetable = slopop)  
  
# si besoin d'un test stratifié sur sur une variable X2  
rs.diff(Surv(time, status) ~ X1 + strata(X2), data = mydata,  
        ratetable = slopop)
```

Application : jeu de données colrec

Exemple adapté à partir de Perme, Maja Pohar, and Klemen Pavlic.
"Nonparametric relative survival analysis with the R package relsurv."
Journal of Statistical Software 87.1 (2018): 1-27.

Jeu de données (1/3)

```
# Chargement du package utilisé
library(relsurv)
# Données colrec incluses dans le package relsurv
str(colrec)
'data.frame':    5971 obs. of  7 variables:
 $ sex      : int   1 2 1 2 2 1 2 1 2 2 ...
 $ age      : num  23004 12082 24277 29256 30416 ...
 $ diag     : 'date' num   12656 13388 12711 13971 12997
 ..
 $ time     : num   16 504 22 3998 9 ...
 $ stat     : int   0 0 0 0 0 0 0 0 0 ...
 $ stage    : Factor w/ 4 levels "1","2","3","99": 1 3 3
1 4 2 1 1 1 1 ...
 $ site     : Factor w/ 2 levels "colon","rectum": 2 2 1
1 1 1 2 2 2 2 ...
```

Jeu de données (2/3)

```
colrec[1:3, ]
```

```
  sex  age  diag  time  stat  stage  site
1   1 23004 12656   16    0     1 rectum
2   2 12082 13388  504    0     3 rectum
3   1 24277 12711   22    0     3  colon
```

```
as.Date(colrec[1,"diag"]) # pour retrouver la date de
diagnostic
```

```
"1994-08-26"
```

Jeu de données (3/3)

```
mydata <- colrec
# pour enlever les patients avec date de diag au 29/02
mydata$diag2 <- as.Date(mydata$diag)
id <- which(format(mydata$diag2, '%m')=="02" &
            format(mydata$diag2, '%d')=="29")
mydata <- mydata[-id, ]
# pour enlever les patients pour lesquels le stade est
manquant
mydata <- mydata[mydata$stage != 99, ]
dim(mydata)
[1] 5575      8
```

Avant de commencer : utilisation de la table de mortalité dans `reksurv`

- Utilisation de la table de mortalité slovène `slopop` (incluse dans le package `reksurv`)
- Examiner cette table (`str`, `head`, ...)
- Que fait l'option `rmap` dans `rs.surv` (`?rs.surv`)

The slopop : a ratetable object

```
head(slopop[,c("1998", "1999", "2000"),])  
Rate table with dimension(s): age year sex  
, , sex = male  
  
   year  
age  1998      1999      2000  
0 1.702793e-05 1.394407e-05 1.664226e-05  
1 1.396694e-06 2.054209e-06 8.762741e-07  
2 1.122777e-06 8.488862e-07 0.000000e+00  
3 8.214987e-07 0.000000e+00 2.738055e-07  
4 1.615848e-06 1.095386e-06 0.000000e+00  
5 5.202539e-07 2.738055e-07 2.738055e-07
```

```
, , sex = female  
  
   year  
age  1998      1999      2000  
0 1.111108e-05 1.179843e-05 1.229343e-05  
1 1.807623e-06 3.011876e-07 3.011876e-07  
2 8.762741e-07 3.011876e-07 6.024083e-07  
3 2.738055e-07 6.024083e-07 3.011876e-07  
4 0.000000e+00 2.738055e-07 0.000000e+00  
5 5.750232e-07 2.738055e-07 1.698036e-06
```

The `ratetable` format and the `rmap` option

- 3 dimensions dans `slopop` : `age`, `sex` et `year`

`rmap`: an optional list to be used if the variables are not organized and named in the same way as in the `ratetable` object. See details below.

Details

NOTE: The follow-up time must be specified in days. The `ratetable` being used may have different variable names and formats than the user's data set, this is dealt with by the `rmap` argument. For example, if `age` is in years in the data set but in days in the `ratetable` object, `age=age*365.241` should be used. The calendar year can be in any date format (`date`, `Date` and `POSIXt` are allowed), the date formats in the `ratetable` and in the data may differ.

Créer une variable groupe selon que l'âge soit inférieur à 70 ans (ou supérieur)

```
# 2 groupes de patients suivant l'âge
d1 <- subset(mydata, mydata$age < 70 * 365.241)
d1$gp <- 1
d2 <- subset(mydata, mydata$age >= 70 * 365.241)
d2$gp <- 2
mydata <- rbind(d1, d2)
```

Censure administrative à 5 ans

```
ind <- which(mydata$time > 365.241 * 5)
mydata$time[ind] <- 365.241 * 5
mydata$stat[ind] <- 0
```


Est-il raisonnable d'estimer la survie nette à 5 ans ?

Etude de la décroissance dans le temps de la taille de l'échantillon due à la mortalité populationnelle (?nessie)

```
breaks <- c(0, seq(from = 45, to = 90, by = 5), Inf)
mydata$agegr <- cut(mydata$age / 365.241, breaks)
nessie(Surv(time, stat) ~ gp + agegr,
      data = mydata[mydata$age / 365.241 > 50, ],
      ratetable = slopop,
      times = seq(0, 5, 1),
      rmap = list(age = age, sex = sex, year = diag))
```

Etude de la décroissance dans le temps de la taille de l'échantillon due à la mortalité populationnelle (?nessie)

	0	1	2	3	4	5	c.exp.surv
gp=1, agegr (50, 55]	383	380.0	376.9	373.6	370.2	366.5	28.0
gp=1, agegr (55, 60]	576	569.3	562.2	554.7	546.9	538.8	23.5
gp=1, agegr (60, 65]	848	831.9	815.2	798.1	780.3	761.9	18.8
gp=1, agegr (65, 70]	1026	997.2	967.2	936.2	904.2	871.5	14.8
gp=2, agegr (70, 75]	955	915.5	875.0	833.5	791.2	747.9	11.3
gp=2, agegr (75, 80]	672	630.3	588.2	545.6	502.5	459.0	8.3
gp=2, agegr (80, 85]	411	366.8	324.5	284.1	245.6	209.8	6.0
gp=2, agegr (85, 90]	198	165.4	136.1	110.2	87.7	68.9	4.2
gp=2, agegr (90, Inf]	33	24.9	18.4	13.3	9.3	6.3	2.9

Sélection des patients de moins de 80 ans

```
mydata <- mydata[mydata$age <= 80 * 365.241, ]  
summary(mydata[c(4,6,7,10)])
```

	time	stage	site	agegr
Min.	: 1	1 : 781	colon :2912	(65,70]:1026
1st Qu.:	281	2 :2953	rectum:2021	(70,75]: 955
Median	:1112	3 :1199		(60,65]: 848
Mean	:1053	99: 0		(75,80]: 672
3rd Qu.:	1826			(55,60]: 576
Max.	:1826			(50,55]: 383
				(Other): 473

1/ Survie globale à 5 ans ?

2/ Probas. brutes de décéder à 5 ans (K et \bar{K})

3/ Lien ?

```
so_fit <- survfit(Surv(time, stat)~gp, data = mydata)
crudep_fit <- cmp.rel( Surv(time, stat) ~ gp,
                     data = mydata, ratetable = slopop,
                     rmap = list(age = age, sex = sex, year = diag))
crd.K <- summary(crudep_fit, times=5, scale = 365.241)$est[c(1,3)]
crd.other <- summary(crudep_fit, times=5)$est[c(2,4)]
(total <- crd.K + crd.other)
[1] 0.5498830 0.6871543
1-summary(so_fit, times=3652.41)$surv
[1] 0.5498830 0.6871543
```

Survie nette à 1, 2 et 5 ans ?

```
sn_fit <- rs.surv(Surv(time, stat) ~ gp,  
                 data = mydata,  
                 ratetable = slopop,  
                 method = "pohar-perme",  
                 add.times = c(1, 2, 5) * 365.241,  
                 rmap = list(age = age, sex = sex, year = diag))  
  
summary(sn_fit, times = c(1, 2, 5) * 365.241)
```

Survie nette à 1, 2 et 5 ans ?

```
summary(sn_fit, times = c(1, 2, 5) * 365.241)
```

gp=1

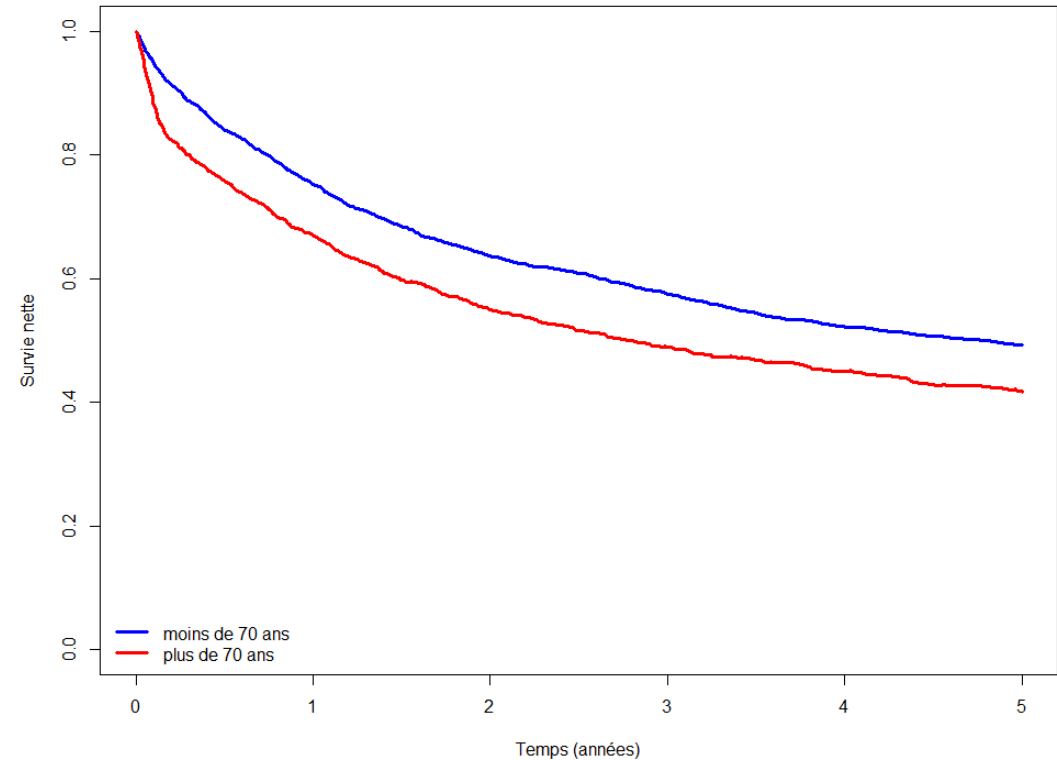
time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
365	2445	858	0.753	0.00776		0.738		0.768
730	2029	415	0.636	0.00877		0.619		0.653
1826	1486	543	0.493	0.00954		0.475		0.512

gp=2

time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
365	1038	589	0.671	0.0125		0.647		0.696
730	806	232	0.549	0.0138		0.523		0.577
1826	509	297	0.417	0.0155		0.387		0.448

Survie nette : évolution

```
plot(sn_fit,  
     col=c("blue", "red"),  
     lwd=3,  
     xscale = 365.241,  
     xlab = "Temps  
(années)",  
     ylab = "Survie nette")
```



Comparaison des distributions de survie nette : le test de type log-rank

```
lrn <- rs.diff(Surv(time, stat) ~ gp,  
              data = mydata, ratetable = slopop,  
              rmap = list(age = age, sex = sex, year = diag))
```

```
lrn
```

```
Value of test statistic: 26.66999
```

```
Degrees of freedom: 1
```

```
P value: 2.413411e-07
```

Comparaison des distributions de survie nette : le test de type log-rank stratifié sur le stade

```
lrn.strt <- rs.diff(Surv(time, stat) ~ gp + strata(stage),  
                   data = mydata, ratetable = slopop,  
                   rmap = list(age = age, sex = sex, year = diag))
```

```
lrn.strt
```

Value of test statistic: 51.58615

Degrees of freedom: 1

P value: 6.852297e-13

Partie II

Méthodes d'estimation paramétrique de la mortalité en excès

Des Méthodes Non-Paramétriques aux Modèles Paramétriques

- Intérêt pour estimer l'effet de covariables sur la mortalité en excès
 - Démographiques, cliniques,...
- Intérêt (plus tardivement) pour estimer la dynamique de la mortalité en excès
- Modèles de régression de « type » modèle de Cox (*survie globale*)

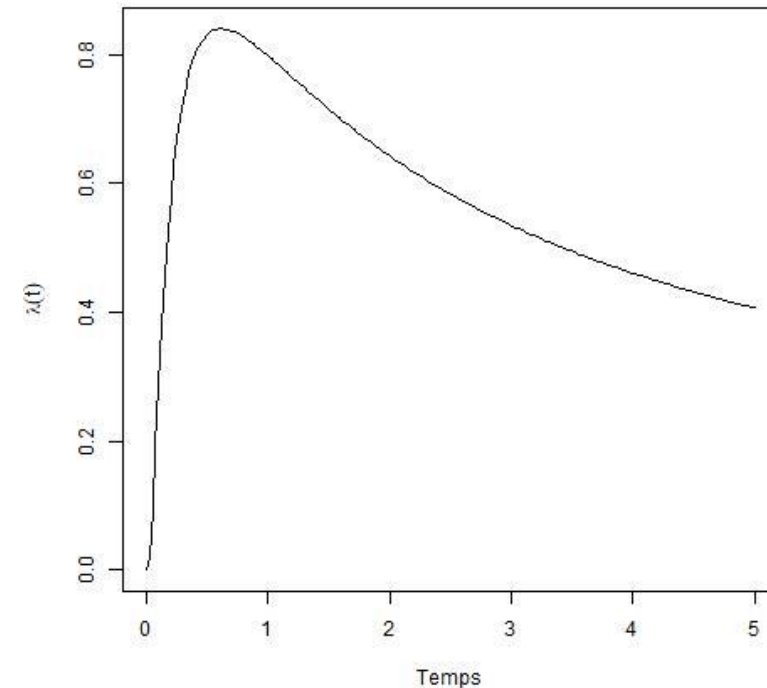
Modèle de Régression de Cox

(1)

- Repose sur la modélisation du taux auquel les individus présentent l'événement d'intérêt

$$\lambda(t) = \lim_{h \rightarrow 0} \frac{P(t \leq T < t + h | T \geq t)}{h}$$

Taux instantané



- Repose sur la modélisation du taux auquel les individus présentent l'événement d'intérêt
- Conditionnellement à leurs covariables z_i , $\lambda(t|z_i)$
- Modèle semi-paramétrique
 - Effet multiplicatif de covariables (*exponentiel*) sur un taux de mortalité de base

$$\lambda(t|z_i) = \lambda_0(t) \exp(\beta z_i)$$

- β vecteur des paramètres associés au vecteur de covariables z
- $\lambda_0(t)$ taux de mortalité de base

Modèle de Régression de Mortalité en Excès

- Repose sur la décomposition du taux de mortalité observé

$$\lambda_O(t|z_i) = \lambda_E(t|z_i) + \lambda_P(t|z_{D_i})$$

- Patient i ayant un vecteur de covariables z_i
- λ_E : mortalité en excès
- λ_P : pour un patient i diagnostiqué à l'âge a_i en l'année y_i , $\lambda_P(t|z_{D_i})$ est obtenu en prenant dans la table de mortalité du même pays (ou à un niveau géographique plus fin, si disponible) le taux pour un individu d'âge $a_i + t$ et à l'année $y_i + t$

- Contribution d'un patient qui décède en t_i ($\delta_i = 1$)
 - $L_i(\beta|t_i, \delta_i) = \lambda_o(t_i)S_o(t_i)$
 - $\log L_i(\beta|t_i, \delta_i) = \log(\lambda_o(t_i)) - \Lambda_o(t_i)$
- Contribution d'un patient censuré en t_i ($\delta_i = 0$)
 - $L_i(\beta|t_i, \delta_i) = S_o(t_i)$
 - $\log L_i(\beta|t_i, \delta_i) = -\Lambda_o(t_i)$
- Contribution d'un patient qui décède selon δ_i
 - $L_i(\beta|t_i, \delta_i) = \lambda_o(t_i)^{\delta_i}S_o(t_i)$
 - $\log L_i(\beta|t_i, \delta_i) = \delta_i \log(\lambda_o(t_i)) - \Lambda_o(t_i)$

- Vraisemblance : produit des vraisemblances individuelles
 - $L(\beta|\mathbf{z}) = \prod_i L_i(\beta|t_i, \delta_i) = \prod_i \lambda_o(t_i)^{\delta_i} S_o(t_i)$
 - $\log L(\beta|\mathbf{z}) = \sum_i \delta_i \log(\lambda_o(t_i)) - \Lambda_o(t_i)$
- Puisque $\lambda_o(t_i|z_i) = \lambda_E(t_i|z_i) + \lambda_P(t_i|z_{D_i})$
 - $\log L(\beta|\mathbf{z}) = \sum_i \left(\delta_i \log[\lambda_E(t_i|z_i) + \lambda_P(t_i|z_{D_i})] - \Lambda_E(t_i|z_i) + \Lambda_P(t_i|z_{D_i}) \right)$
- Puisque $\Lambda_P(t_i|z_{D_i}) \perp \beta$, on maximise
 - $\sum_i \left(\delta_i \log[\lambda_E(t_i|z_i) + \lambda_P(t_i|z_{D_i})] - \Lambda_E(t_i|z_i) \right)$

Modèle de Mortalité en Excès d'Estève *et coll.* (1)

$$\lambda_E(t|z_i) = \lambda_{E,0}(t) \exp(\beta z_i)$$

- Avec $\lambda_{E,0}(t) = \sum_{j=1}^J \tau_j I_j(t)$: taux de base en excès
où $I_j(t) = 1$ si $\tau_{j-1} < t < \tau_j$ ($\mathbf{z} = 0$)

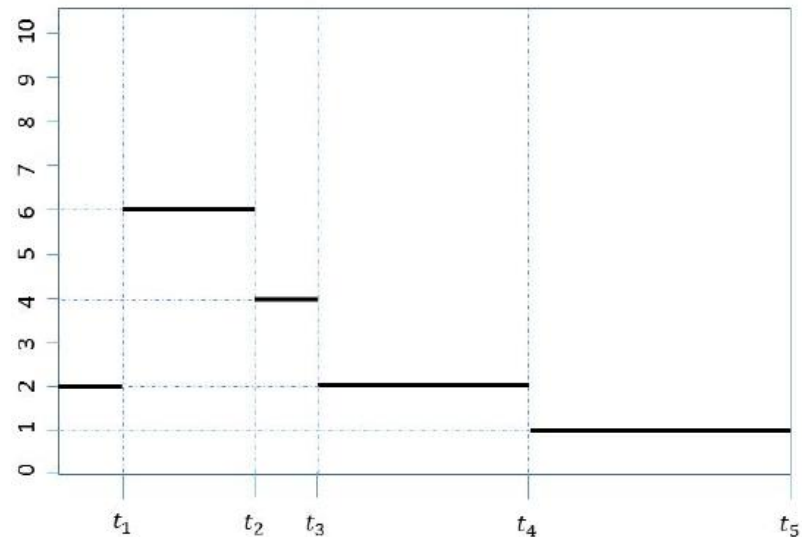
Modèle de Mortalité en Excès d'Estève *et coll.*

(2)

$$\lambda_E(t|z_i) = \lambda_{E,0}(t) \exp(\beta z_i)$$

- Avec $\lambda_{E,0}(t) = \sum_{j=1}^J \tau_j I_j(t)$: taux de base en excès

où $I_j(t) = 1$ si $\tau_{j-1} < t < \tau_j$ ($z = 0$)



$$\lambda_{E,0}(t) = 2 \quad (\forall t \text{ entre } 0 \text{ et } t_1)$$

$$\lambda_{E,0}(t) = 6 \quad (\forall t \text{ entre } t_1 \text{ et } t_2)$$

$$\lambda_{E,0}(t) = 4 \quad (\forall t \text{ entre } t_2 \text{ et } t_3)$$

$$\lambda_{E,0}(t) = 2 \quad (\forall t \text{ entre } t_3 \text{ et } t_4)$$

$$\lambda_{E,0}(t) = 1 \quad (\forall t \text{ entre } t_4 \text{ et } t_5)$$

Fonction constante par morceaux (exemple fictif)

Modèle de Mortalité en Excès d'Estève *et coll.* (3)

$$\lambda_E(t|z_i) = \lambda_{E,0}(t) \exp(\beta z_i)$$

- Avec $\lambda_{E,0}(t) = \sum_{j=1}^J \tau_j I_j(t)$: taux de base en excès
où $I_j(t) = 1$ si $\tau_{j-1} < t < \tau_j$ ($\mathbf{z} = 0$)
- Et $\exp(\beta z_i)$: taux relatif de mortalité en excès
- Estimation des β et des τ ($\lambda_{E,0}$) par maximisation de la vraisemblance totale

Interprétation des Paramètres β

- Pour 2 individus i, j ayant pour covariables z_i et z_j

$$\frac{\lambda_E(t|z_i)}{\lambda_E(t|z_j)} = \frac{\lambda_{E,0}(t)\exp(\beta z_i)}{\lambda_{E,0}(t)\exp(\beta z_j)} = \exp(\beta(z_i - z_j)) = TRME(t) = TRME$$

TRME : Taux Relatif de Mortalité en Excès

Interprétation des Paramètres β

- Pour 2 individus i, j ayant pour covariables z_i et z_j

$$\frac{\lambda_E(t|z_i)}{\lambda_E(t|z_j)} = \frac{\lambda_{E,0}(t)\exp(\beta z_i)}{\lambda_{E,0}(t)\exp(\beta z_j)} = \exp(\beta(z_i - z_j)) = TRME(t) = TRME$$

TRME : Taux Relatif de Mortalité en Excès

- z binaire (sexe, codé en 0 si féminin et 1 si masculin)
 - $TRME = \exp(\beta(z = 1 - z = 0)) = \exp(\beta)$
 - **Hypothèse des risques proportionnels**

Interprétation des Paramètres β

- Pour 2 individus i, j ayant pour covariables z_i et z_j

$$\frac{\lambda_E(t|z_i)}{\lambda_E(t|z_j)} = \frac{\lambda_{E,0}(t)\exp(\beta z_i)}{\lambda_{E,0}(t)\exp(\beta z_j)} = \exp(\beta(z_i - z_j)) = TRME(t) = TRME$$

TRME : Taux Relatif de Mortalité en Excès

- z binaire (sexe, codé en 0 si féminin et 1 si masculin)
 - $TRME = \exp(\beta(z = 1 - z = 0)) = \exp(\beta)$
 - **Hypothèse des risques proportionnels**
- z continu (âge, en années en prenant par exemple ces 2 valeurs)
 - $TRME = \exp(\beta(z = 25 - z = 24)) = \exp(\beta)$
 - **Hypothèse de log-linéarité**
- Interprétation comme n'importe quel « RR » par rapport à 1

Modèle de Mortalité en Excès d'Estève *et coll.* *Modèle à Taux de Mortalité en Excès Proportionnels*

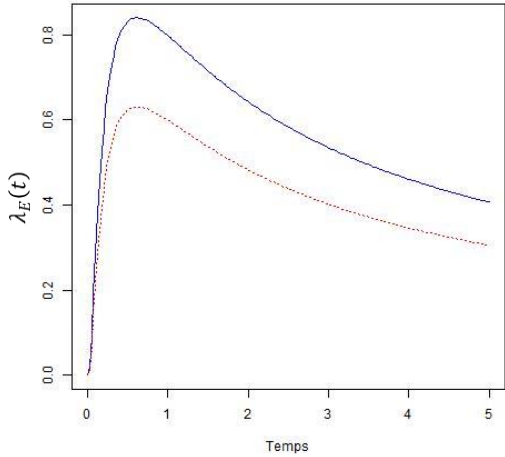
- Pour 2 individus i, j ayant pour covariables z_i et z_j

$$\frac{\lambda_E(t|z_i)}{\lambda_E(t|z_j)} = \frac{\lambda_{E,0}(t)\exp(\beta z_i)}{\lambda_{E,0}(t)\exp(\beta z_j)} = \exp(\beta(z_i - z_j)) = TRME(t) = TRME$$

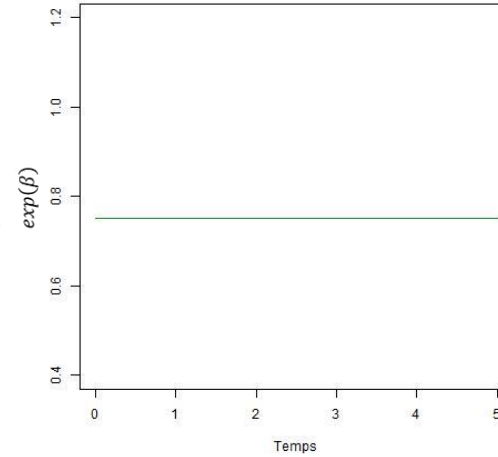
TRME : Taux Relatif de Mortalité en Excès

- Par conséquence, les fonctions du taux de mortalité en excès sont proportionnelles, leur rapport est constant et ne dépend pas de t
 - Hypothèse à vérifier au vue des données
- Rapport des fonctions du taux de mortalité en excès est un « risque relatif » en excès à l'instant t des sujets de caractéristique z_i par rapport à ceux de caractéristique z_j

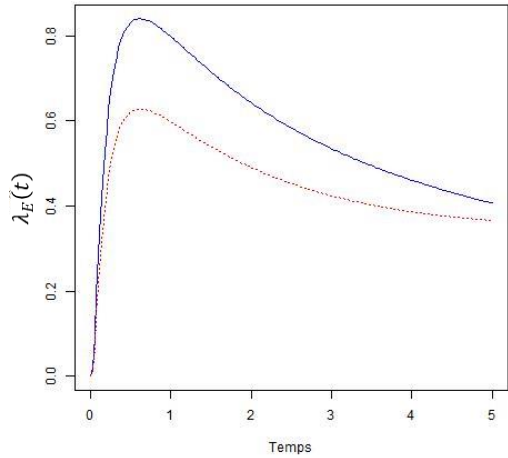
Modèle à Taux de Mortalité en Excès Proportionnels



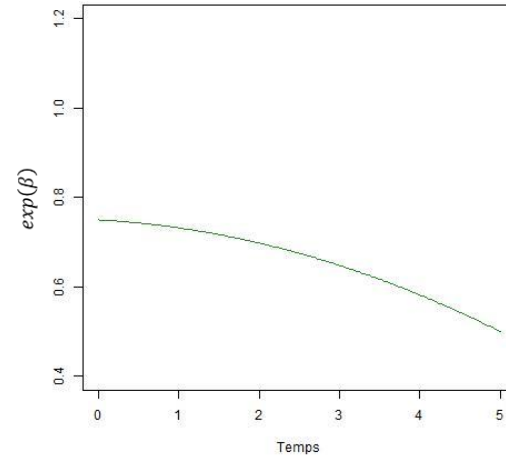
$$\frac{\lambda_E(t|z_1)}{\lambda_E(t|z_2)} = \frac{\lambda_{E,0}(t)\exp(\beta z_1)}{\lambda_{E,0}(t)\exp(\beta z_2)} = \exp(\beta)$$



Modèle à Taux de Mortalité en Excès non Proportionnels



$$\lambda_E(t|z_1)$$
$$\lambda_E(t|z_2)$$



$$\frac{\lambda_{E,0}(t)\exp(\beta z_1)}{\lambda_{E,0}(t)\exp(\beta z_2)} = \exp(\beta(t))$$

Modèle de Mortalité en Excès d'Estève *et coll.* *Modèle log-linéaire sur l'effet des covariables*

- Pour 2 individus i, j ayant pour covariables z_i et z_j

$$\frac{\lambda_E(t|z_i)}{\lambda_E(t|z_j)} = \frac{\lambda_{E,0}(t)\exp(\beta z_i)}{\lambda_{E,0}(t)\exp(\beta z_j)} = \exp(\beta(z_i - z_j)) = TRME(t) = TRME$$

TREM : Taux Relatif de Mortalité en Excès

$$\ln\left(\frac{\lambda_{E,0}(t)\exp(\beta z_i)}{\lambda_{E,0}(t)\exp(\beta z_j)}\right) = \beta(z_i - z_j)$$

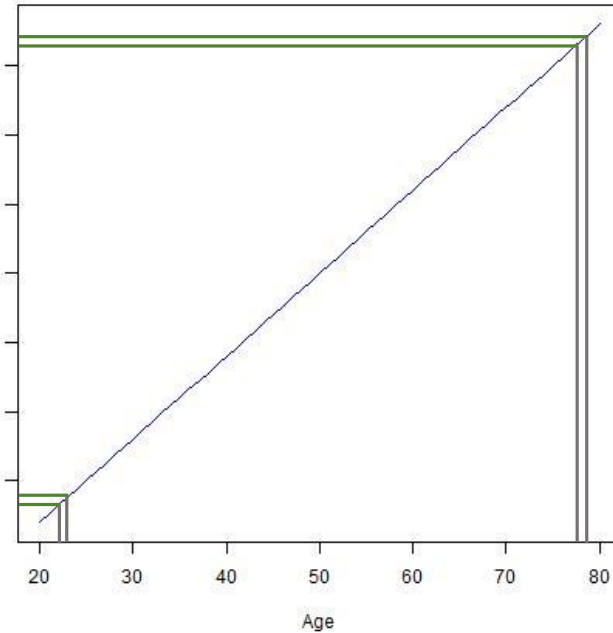
- Par conséquence, le logarithme du rapport des taux de mortalité en excès est une fonction linéaire des covariables
 - Hypothèse à vérifier au vue des données

Effet Log-Linéaire sur une Covariable

(1)

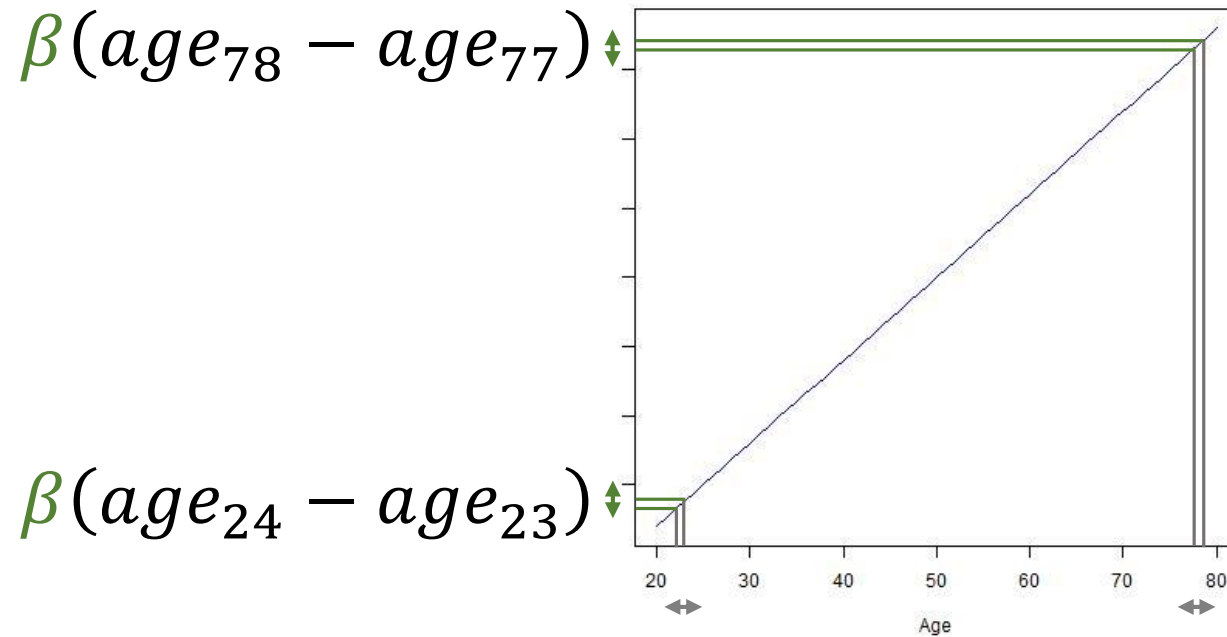
$$\beta(\text{age}_{78} - \text{age}_{77})$$

$$\beta(\text{age}_{24} - \text{age}_{23})$$

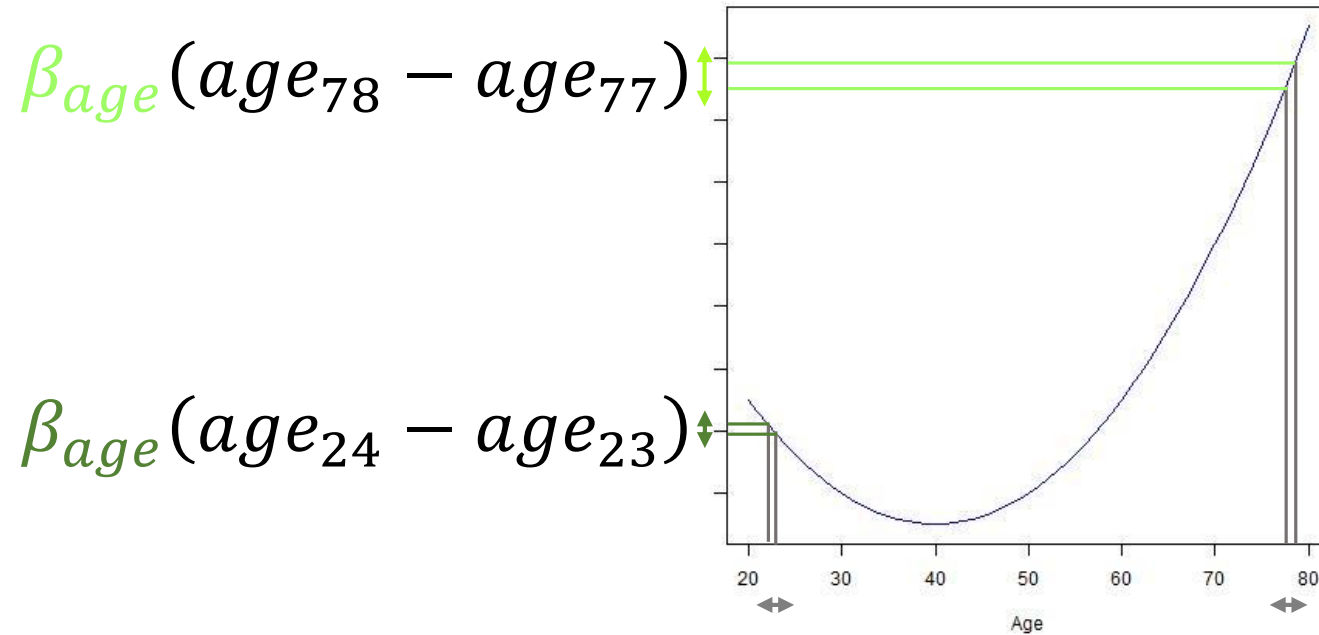


Effet Log-Linéaire sur une Covariable

(2)



Effet Non Log-Linéaire sur une Covariable



Tests des Paramètres β

Une seule (nominale), plusieurs variables

- Hypothèse nulle $H_0: \beta = \beta^{(0)}$ (dimension de $\beta = p$)

- Test

- Rapport de vraisemblance

- $2[L(\hat{\beta}) - L(\beta^{(0)})]$

- Wald

- $(\hat{\beta} - \beta^{(0)})^t I(\hat{\beta})^{-1} (\hat{\beta} - \beta^{(0)})$

- Score

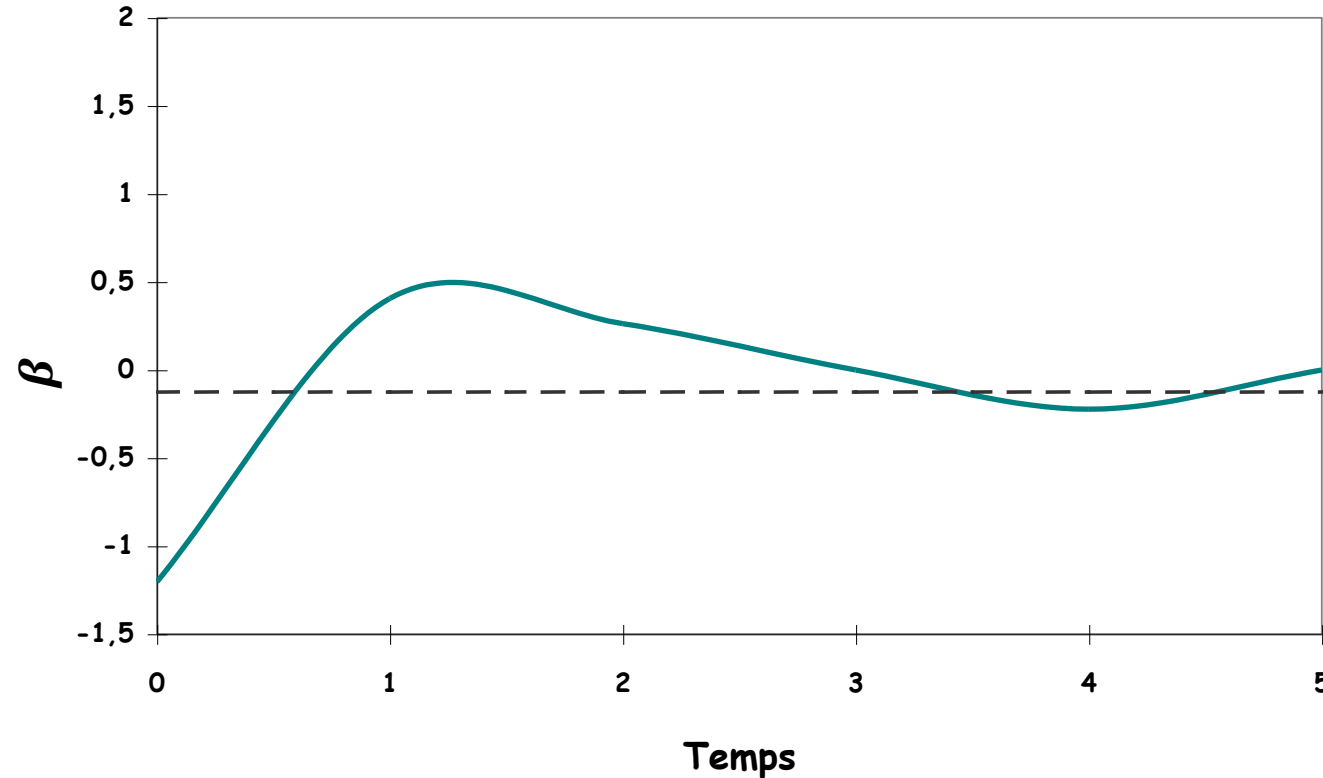
- $U^t(\beta^{(0)}) I(\beta^{(0)})^{-1} U(\beta^{(0)})$

où I est la matrice d'information de Fisher

} $\xrightarrow{H_0} \chi^2$ à p degrés de liberté

Estimation avec Hypothèse TME Proportionnels vs non Proportionnels

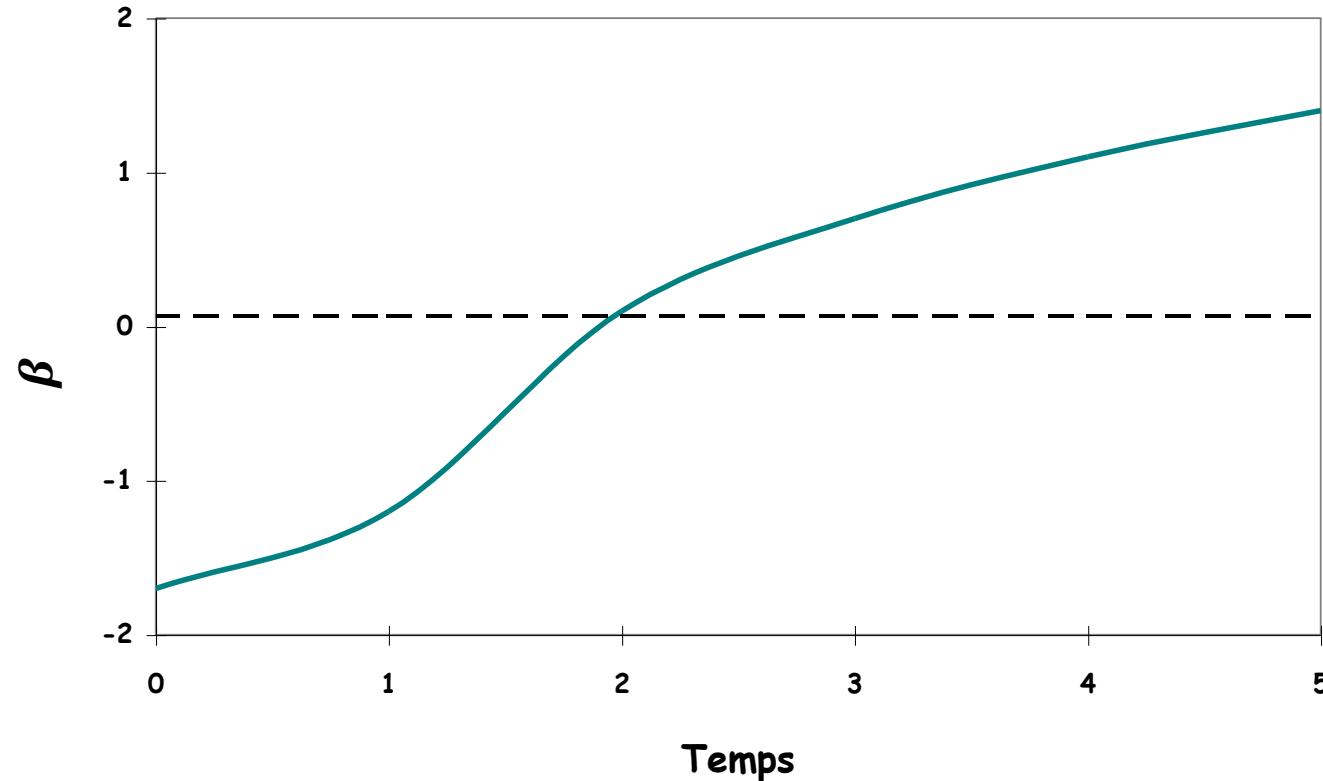
(1)



► Estimation d'un effet moyen ◀

Estimation avec Hypothèse TME Proportionnels vs non Proportionnels

(2)



► Conclusion à tort à l'absence d'effet ◀

Modèle de Mortalité en Excès de Giorgi *et coll.* (1)

$$\lambda_E(t|z_i) = \lambda_{E,0}(t) \exp(\beta(t)z_i)$$

$$\lambda_E(t|z_i) = \sum_{j=1}^J \tau_j I_j(t) \exp(\beta z_i)$$

Estève *et coll.*

- Avec $\lambda_{E,0}(t)$ et $\beta(t)$ modélisés par des fonction B-Splines

Fonctions B-Splines

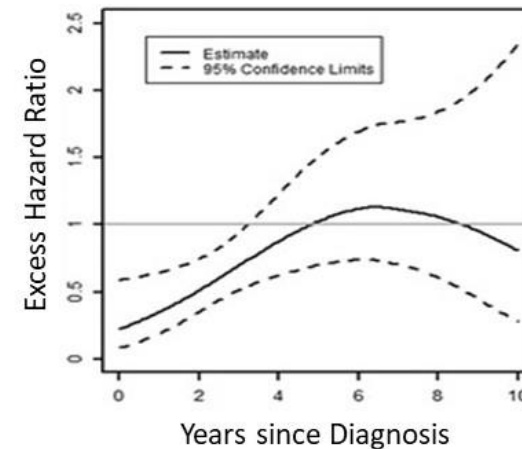
(1)

- Fonctions polynomiales par morceaux d'ordre q (polynôme de degré $q - 1$)
 - Flexibilité augmente avec l'ordre
- Point de jonction de deux polynômes contiguës au niveau de points appelés nœuds
 - Nombre de nœuds contrôle la flexibilité
- Chaque morceau (polynôme) relie le morceau suivant de manière lisse et homogène

Fonctions B-Splines

(1)

- Fonctions polynomiales par morceaux d'ordre q (polynôme de degré $q - 1$)
 - Flexibilité augmente avec l'ordre
- Point de jonction de deux polynômes contiguës au niveau de points appelés nœuds
 - Nombre de nœuds contrôle la flexibilité
- Chaque morceau (polynôme) relie le morceau suivant de manière lisse et homogène

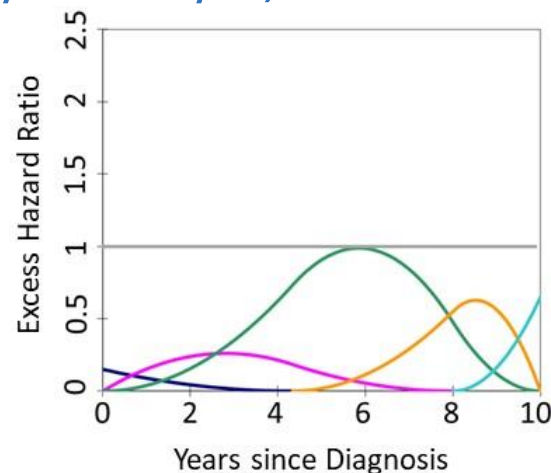


Fonctions B-Splines

(1)

- Fonctions polynomiales par morceaux d'ordre q (polynôme de degré $q - 1$)
 - Flexibilité augmente avec l'ordre
- Point de jonction de deux polynômes contiguës au niveau de points appelés nœuds
 - Nombre de nœuds contrôle la flexibilité
- Chaque morceau (polynôme) relie le morceau suivant de manière lisse et homogène

B-Splines quadratique, 2 nœuds intérieurs

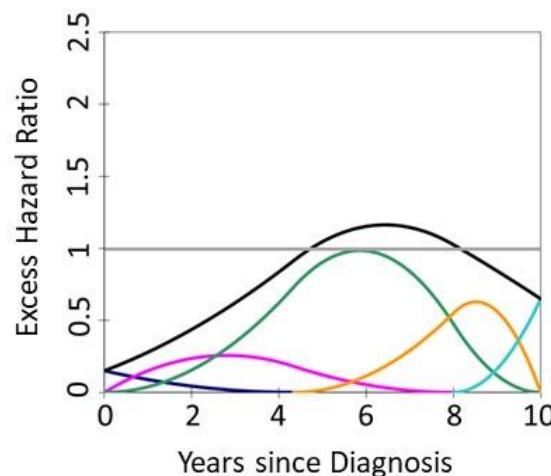
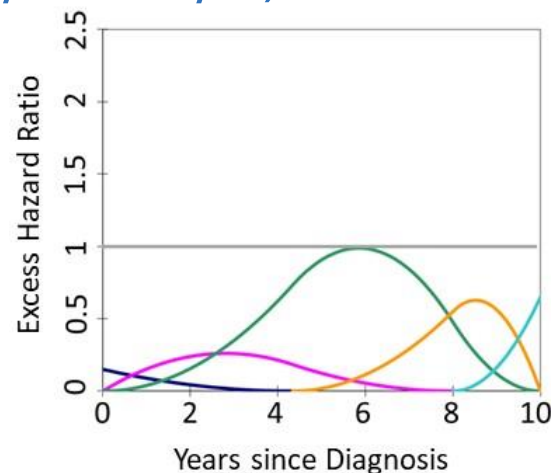


Fonctions B-Splines

(1)

- Fonctions polynomiales par morceaux d'ordre q (polynôme de degré $q - 1$)
 - Flexibilité augmente avec l'ordre
- Point de jonction de deux polynômes contiguës au niveau de points appelés nœuds
 - Nombre de nœuds contrôle la flexibilité
- Chaque morceau (polynôme) relie le morceau suivant de manière lisse et homogène

B-Splines quadratique, 2 nœuds intérieurs

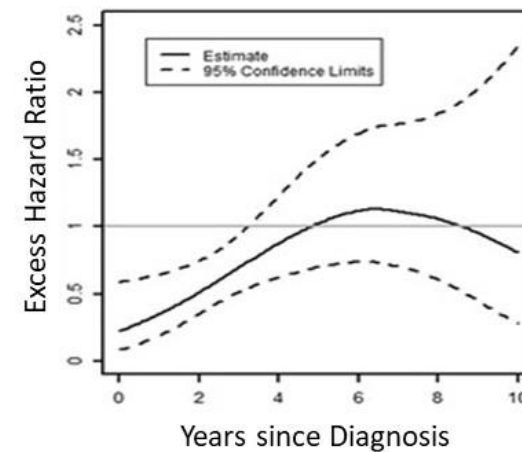
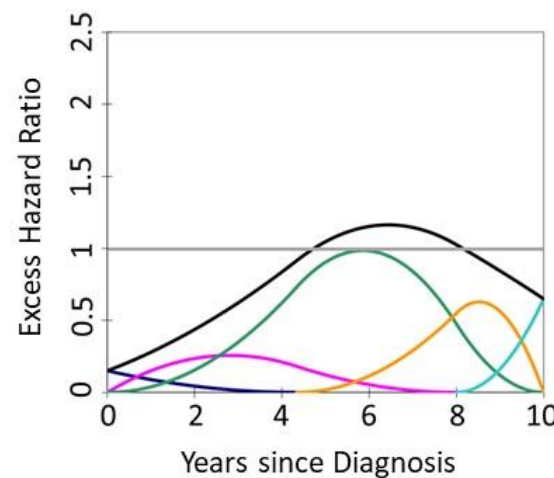
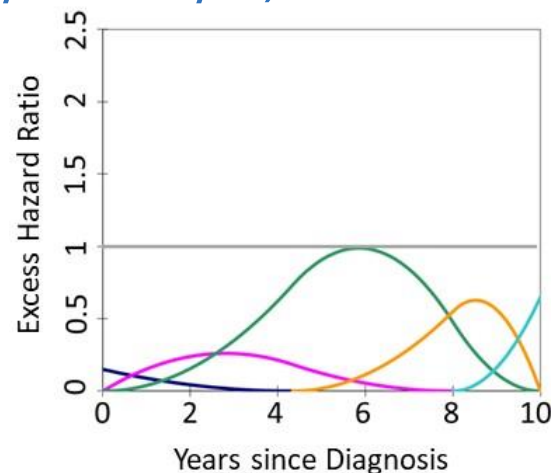


Fonctions B-Splines

(1)

- Fonctions polynomiales par morceaux d'ordre q (polynôme de degré $q - 1$)
 - Flexibilité augmente avec l'ordre
- Point de jonction de deux polynômes contiguës au niveau de points appelés nœuds
 - Nombre de nœuds contrôle la flexibilité
- Chaque morceau (polynôme) relie le morceau suivant de manière lisse et homogène

B-Splines quadratique, 2 nœuds intérieurs



- Modélisation souple et flexible des variations de l'effet au cours du temps
- Estimation non plus basée sur des *a priori* mais sur la modélisation des données observées
- Avantage
 - Estimation « fine », non biaisée
 - Rend bien compte de l'effet tel qu'il est estimé
- Inconvénient – Difficulté relative
 - Interprétation non résumée à l'interprétation d'un seul paramètre

Fonctions B-Splines

Quadratique, 2 nœuds intérieurs

(3)

$$\beta_i(t) = \sum_{j=-(q-1)}^m \alpha_{i,j} B_{j,q}(t) \quad i = 1, \dots, P$$

- Fonctions B-Splines quadratiques (ordre $q = 3$)
- Deux nœuds intérieurs ($m = 2$)
 - Localisation basée sur les quantiles de la distribution des temps de décès observés, connaissances épidémiologiques/clinique,...
- 5 fonctions de bases pour l'espace de régression des splines
 - $5P$ paramètres à estimer par coefficients associés aux covariables

Modèle de Mortalité en Excès de Giorgi *et coll.* (2)

$$\lambda_E(t|z_i) = \lambda_{E,0}(t) \exp(\beta(t)z_i)$$

$$\lambda_E(t|z_i) = \sum_{j=1}^J \tau_j I_j(t) \exp(\beta z_i)$$

Estève *et coll.*

- Avec $\lambda_{E,0,i}(t)$ et $\beta(t)$ modélisés par des fonction B-Splines

$$\lambda_E(t|z_i) = \sum_{j=-2}^2 \nu_j B_{j,3}(t) \exp\left(\sum_{p=1}^P \sum_{j=-2}^2 \alpha_{p,j} B_{j,3}(t) z_i\right)$$

- Estimation des paramètres ν et α par maximisation de la vraisemblance totale

Tests des Paramètres α

- Tests du rapport de vraisemblance (TRV)
 - Pas d'association
 - $\beta_p(t) = 0$: TRV avec $(q + m)P$ degrés de liberté
 - Proportionnalité des taux
 - $\beta_p(t) = C$: TRV avec $((q + m) - 1)P$ degrés de liberté

Modèle de Mortalité en Excès de Giorgi *et coll.* (3)

$$\lambda_E(t|z_i) = \lambda_{E,0}(t) \exp(\beta(t)z_i)$$

$$\lambda_E(t|z_i) = \sum_{j=1}^J \tau_j I_j(t) \exp(\beta z_i)$$

Estève *et coll.*

$$\lambda_E(t|z_i) = \sum_{j=-2}^2 \nu_j B_{j,3}(t) \exp\left(\sum_{p=1}^P \sum_{j=-2}^2 \alpha_{p,j} B_{j,3}(t) z_i\right)$$

$$\lambda_E(t|z_i) = \sum_{j=-2}^2 \nu_j B_{j,3}(t) \exp\left[\left(\sum_{p=1}^{P-l} \sum_{j=-2}^2 \alpha_{p,j} B_{j,3}(t) z_i\right) + \left(\sum_{h=P-l+1}^P \beta z_i\right)\right]$$

Taux de base en excès :

modélisation souple et flexible de ces variations au cours du **temps**

Taux relatif de mortalité en excès :

modélisation souple et flexible d'effets au cours du **temps**

modélisation constante d'effets au cours du temps

Application : jeu de données colrec

Exemples adaptés de ceux donnés dans le package xhaz :

Juste Goungounga, Nathalie Grafféo, Roch Giorgi.

Excess hazard modelling taking into account inappropriate expected mortality rates.

Références

- Cox DR. Regression models and life tables. *JRSS B* 1972; 34: 187-220.
- Esteve J, Benhamou E, Croasdale M, Raymond L. Relative survival and the estimation of net survival: elements for further discussion. *Stat Med* 1990; 9: 529-538.
- Giorgi R, Abrahamowicz M, Quantin C, Bolard P, Esteve J, Gouvernet J, Faivre J. A relative survival regression model using B-splines functions. *Stat Med* 2003; 22:2767-84.
- Grafféo N, Castell F, Belot A, Giorgi R. A log-rank-type test to compare net survival distributions. *Biometrics* 2016; 72(3):760-769.
- Perme MP, Stare J, Estève J. On estimation in relative survival. *Biometrics* 2012; 68(1):113-120.
- Perme MP, Estève J, Rachet B. Analysing population-based cancer survival - settling the controversies. *BMC Cancer* 2016; 16(1):1-8.
- Perme MP, Pavlic K. Nonparametric relative survival analysis with the R package relsurv. *Journal of Statistical Software* 2018; 87(1):1-27.

Pour aller plus loin ?

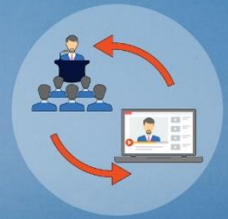
- Modèles paramétriques flexibles
 - Méthodes non pénalisées, pénalisées
- Outils diagnostic
 - Ecart à la proportionnalité, à la log-linéarité
- Gestion des données manquantes
- Gestion des événements multiples
 - Événements répétés, risques semi-compétitifs, modèles multi-états
- Modélisation en présence de tables de mortalité insuffisamment stratifiée
- Extensions des modèles de mortalité en excès en situation non-populationnelle
 - Recherche clinique, santé au travail



15^{ème} Conférence Francophone d'**EPI**démiologie **CLIN**ique

EPICLIN2021 / JSCLCC28

28^{èmes} Journées des Statisticiens des Centres de Lutte Contre le Cancer



**8-11
Juin 2021**



Annexe 1

Ratio de Survie Relative

Ratio de survie relative

(1)

- $S_R(t) = \frac{S_O(t)}{S_P(t)}$

- Ederer I

$$\int_0^t \frac{dN(u)}{Y(u)} - \int_0^t \frac{\sum_{i=1}^n S_{P,i}(u) \lambda_{P,i}(u)}{\sum_{i=1}^n S_{P,i}(u)} du$$

- On peut montrer que

$$\lambda_P(u) = \frac{E(S_{P,1}(u) \lambda_{P,1}(u))}{E(S_{P,1}(u))}$$

- D'où Ederer I estimateur consistant de $\Lambda_0(t) - \Lambda_P(t)$ ($\neq \Lambda_E(t)$)

Ratio de survie relative

(2)

- $S_R(t) = \frac{S_O(t)}{S_P(t)} = \frac{\exp(-\Lambda_O(t))}{\exp(-\Lambda_P(t))} = \exp(\Lambda_P(t) - \Lambda_O(t))$
- Or Ederer I estimateur consistant de $\Lambda_O(t) - \Lambda_P(t)$
⇒ Ederer I estimateur consistant du ratio de survie relative
- Dans R, package `relsurv`

```
rs.surv(Surv(time, status)~1, data = mydata,  
        ratetable = slopop, method = ``ederer1``)
```

Ratio de survie relative

(3)

```
# Ratio de survie relative à 1, 2, et 5 ans ?
rsr_fit <- rs.surv(Surv(time, stat) ~ gp, data = mydata,
                 ratetable = slopop,
                 method = "ederer1",
                 add.times = c(1, 5, 10) * 365.241,
                 rmap = list(age = age, sex = sex, year = diag))

summary(rsr_fit, times = c(1, 2, 5) * 365.241)
```

Ratio de survie relative

(4)

```
summary(rsr_fit, times = c(1, 5, 10) * 365.241)
```

gp=1

time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
365	2445	858	0.753	0.00776		0.738		0.768
730	2029	415	0.637	0.00877		0.620		0.654
1826	1486	543	0.496	0.00953		0.477		0.515

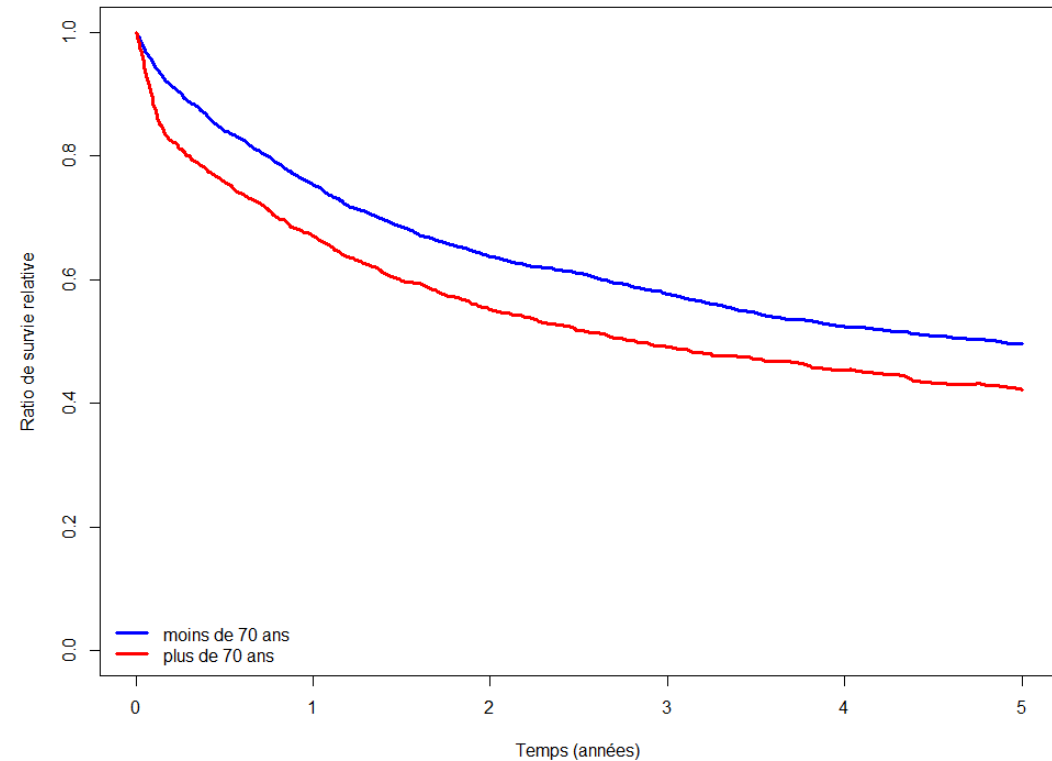
gp=2

time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
365	1038	589	0.671	0.0125		0.647		0.696
730	806	232	0.551	0.0138		0.524		0.578
1826	509	297	0.422	0.0155		0.392		0.453

Ratio de survie relative

(5)

```
plot(rsr_fit,  
     col=c("blue", "red"),  
     lwd=3,  
     xscale = 365.241,  
     xlab = "Temps  
(années)",  
     ylab = "Ratio de  
survie relative")
```



Annexe 2

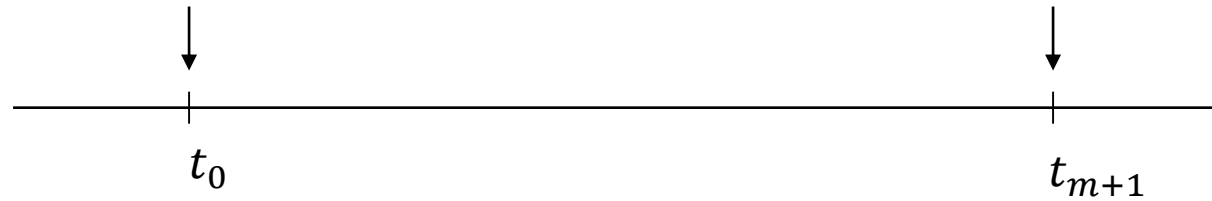
Fonctions B-Splines

Fonctions B-Splines

(1)

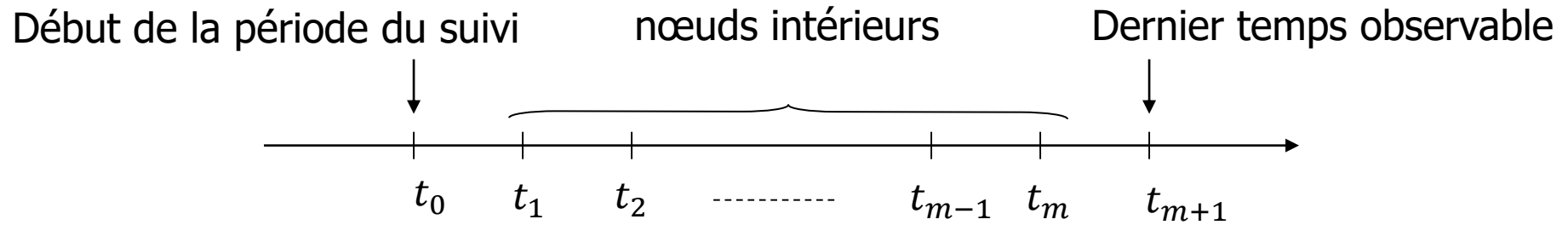
Début de la période du suivi

Dernier temps observable



Fonctions B-Splines

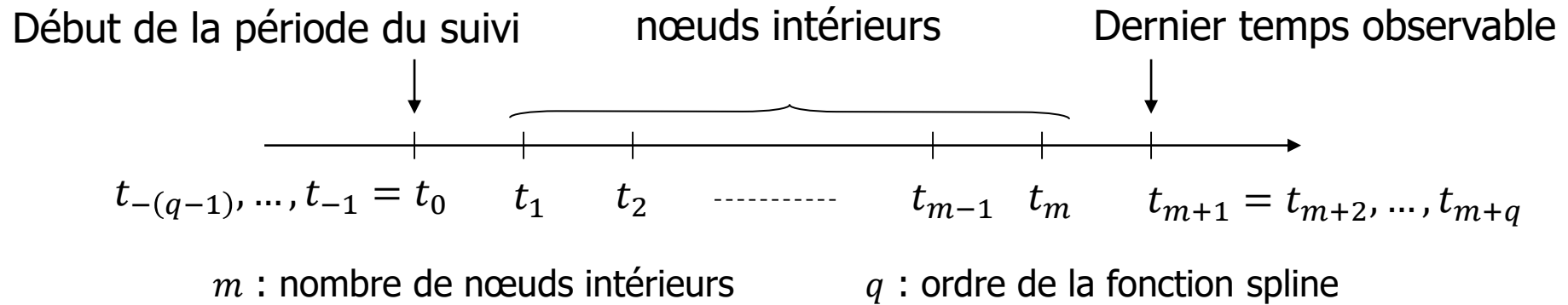
(1)



m : nombre de nœuds intérieurs

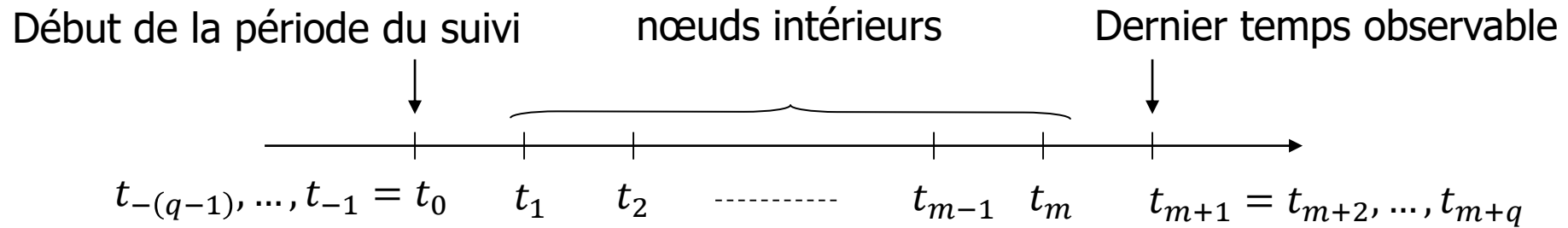
Fonctions B-Splines

(1)



Fonctions B-Splines

(1)



m : nombre de nœuds intérieurs

q : ordre de la fonction spline

$$B_{j,q}(t) = \frac{t - t_j}{t_{j+q-1} - t_j} B_{j,q-1}(t) + \frac{t_{j+q} - t}{t_{j+q} - t_{j+1}} B_{j+1,q-1}(t)$$

$$\left| \begin{array}{l} j = -(q-1), \dots, m \\ B_{j,1}(t) = \begin{cases} 1 & \text{si } t \in [t_j, t_{j+1}[\\ 0 & \text{sinon} \end{cases} \end{array} \right.$$

Fonctions B-Splines

Quadratique, 2 nœuds intérieurs

(2)

$$B_{j,q}(t) = \frac{t - t_j}{t_{j+q-1} - t_j} B_{j,q-1}(t) + \frac{t_{j+q} - t}{t_{j+q} - t_{j+1}} B_{j+1,q-1}(t)$$

$m = 2$: nombre de nœuds intérieurs $q = 3$: ordre de la fonction spline

Avec $j = -2$ et $q = 3$:

$$B_{-2,3}(t) = \frac{t - t_{-2}}{t_{-2+3-1} - t_{-2}} B_{-2,3-1}(t) + \frac{t_{-2+3} - t}{t_{-2+3} - t_{-2+1}} B_{-2+1,3-1}(t)$$

$$B_{-2,2}(t) = \frac{t - t_{-2}}{t_{-1} - t_{-2}} B_{-2,1}(t) + \frac{t_0 - t}{t_0 - t_{-1}} B_{-1,1}(t)$$

$$B_{-1,2}(t) = \frac{t - t_{-1}}{t_0 - t_{-1}} B_{-1,1}(t) + \frac{t_1 - t}{t_1 - t_0} B_{0,1}(t)$$

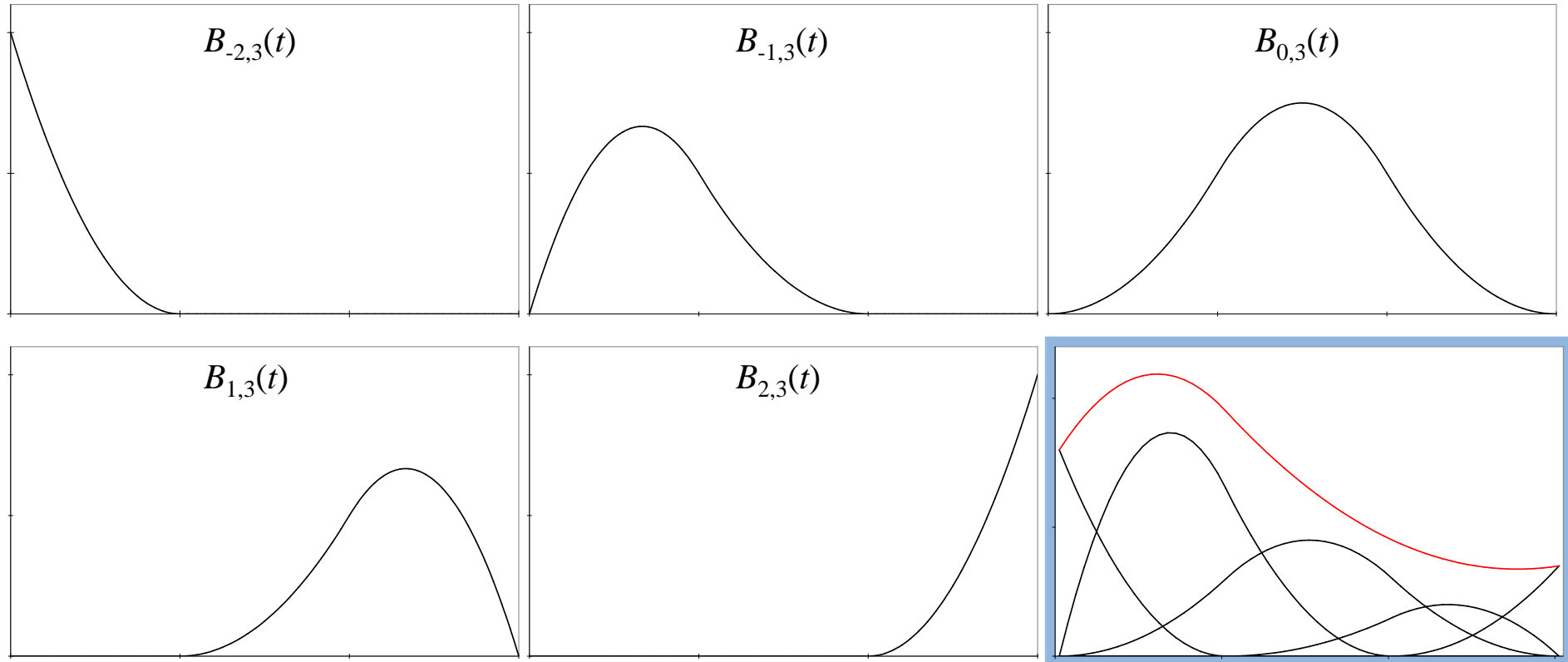
$$B_{-2,3}(t) = \frac{t_1 - t}{t_1 - t_{-1}} \times \frac{t_1 - t}{t_1 - t_0} B_{0,1}(t)$$

$$B_{-2,3}(t) = 1 - \frac{2}{t_1} t + \frac{1}{t_1^2} t^2 \text{ si } t \in [t_0, t_1[, 0 \text{ sinon}$$

Fonctions B-Splines

Quadratique, 2 nœuds intérieurs

(3)



$$\beta(t) = \sum_{j=-(q-1)}^m \alpha_j B_{j,q}(t), t \in (t_0, t_{m+1})$$