Análisis de datos longitudinales

Grado en Estadística

Tema 2 – Sesión 7

Análisis de Supervivencia

Eventos recurrentes (II)

Juan R González

Departamento de Matemáticas, UAB

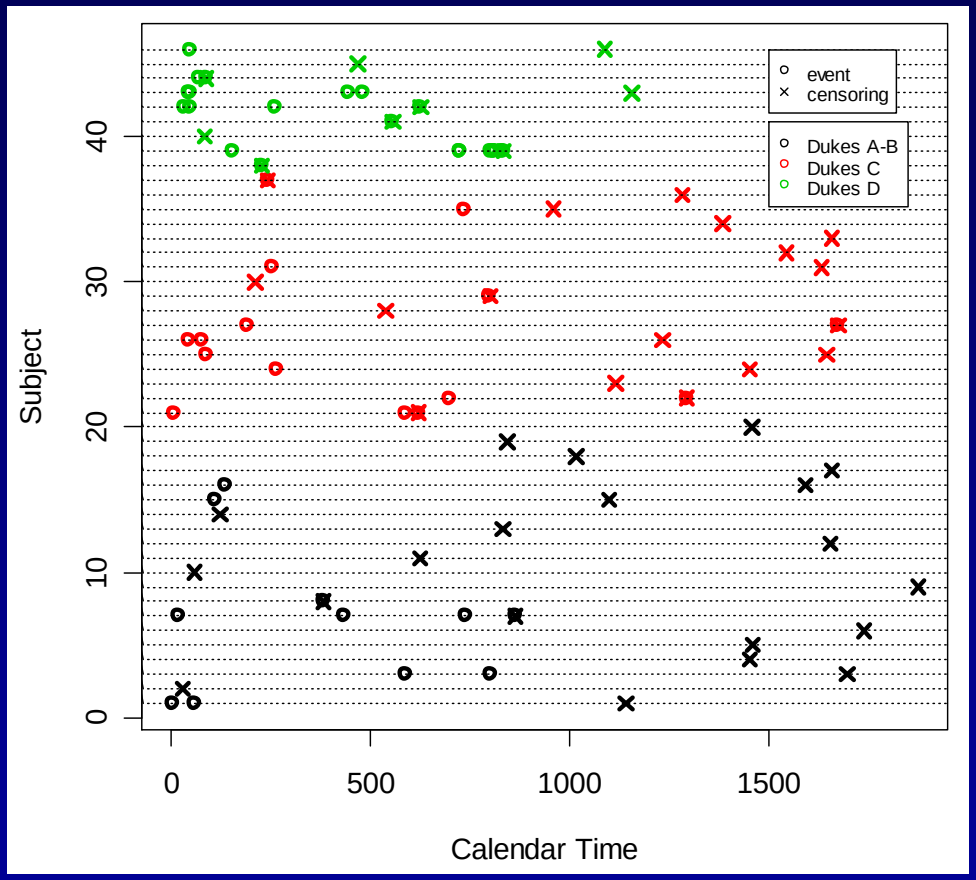
Insituto de Salud Global Barcelona, ISGlobal

Outline of the talk

- Motivation
- Introduction
  - Data with recurrent events
  - Heterogeneity – Frailty Models
- General class of models
  - The general class of models
  - The effective age process
  - EM algorithm
  - Penalized partial likelihood
  - Penalized likelihood approach
  - Application to the re-hospitalization data set
- Dynamic cancer model
  - The effective age process for cancer data
  - Simulation study
  - Application to the non-Hodkin's lymphoma data set

Problem 1: Colorectal cancer hospital readmissions

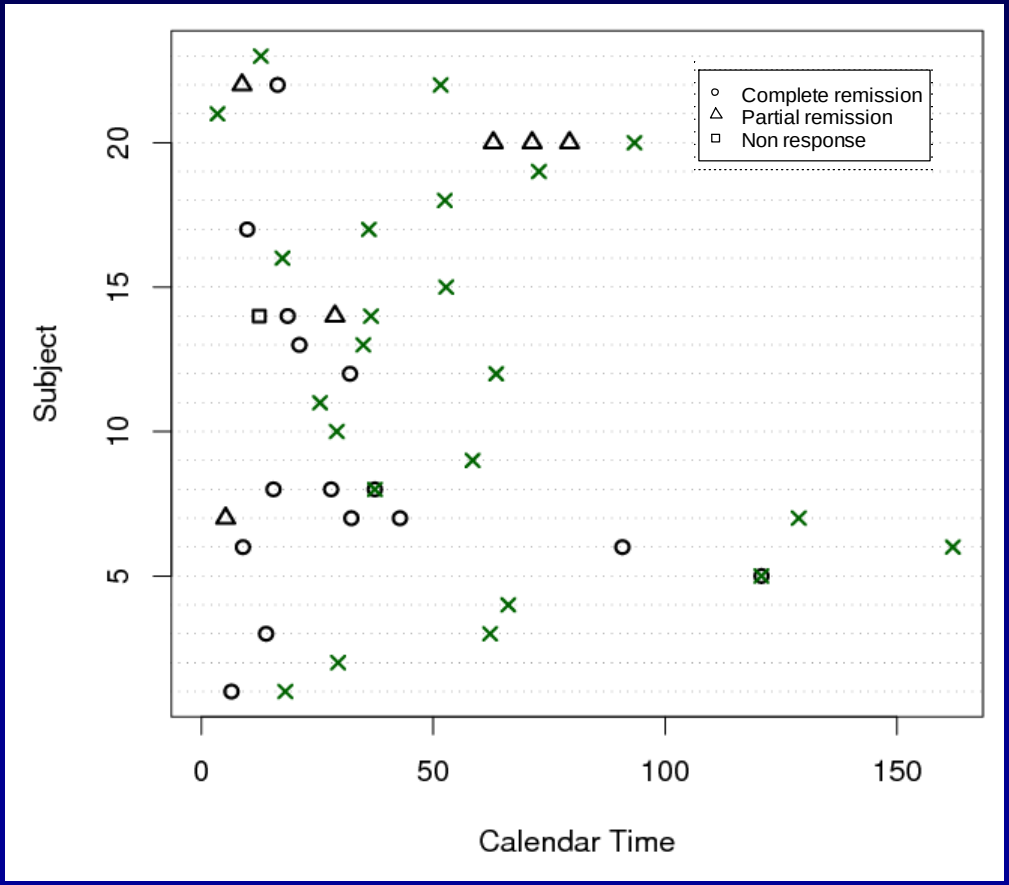
Time to next hospital readmission after surgery in colorectal patients



Are there any differences between tumor stages?
How do we consider the heterogeneity?
What is the impact of number of events?

Problem 2: Non-Hodgkin lymphomas

Time to next hospital relapse after diagnosis in cancer patients



What is the effect of the response to the treatment on the time of next relapse?

In biomedical settings …

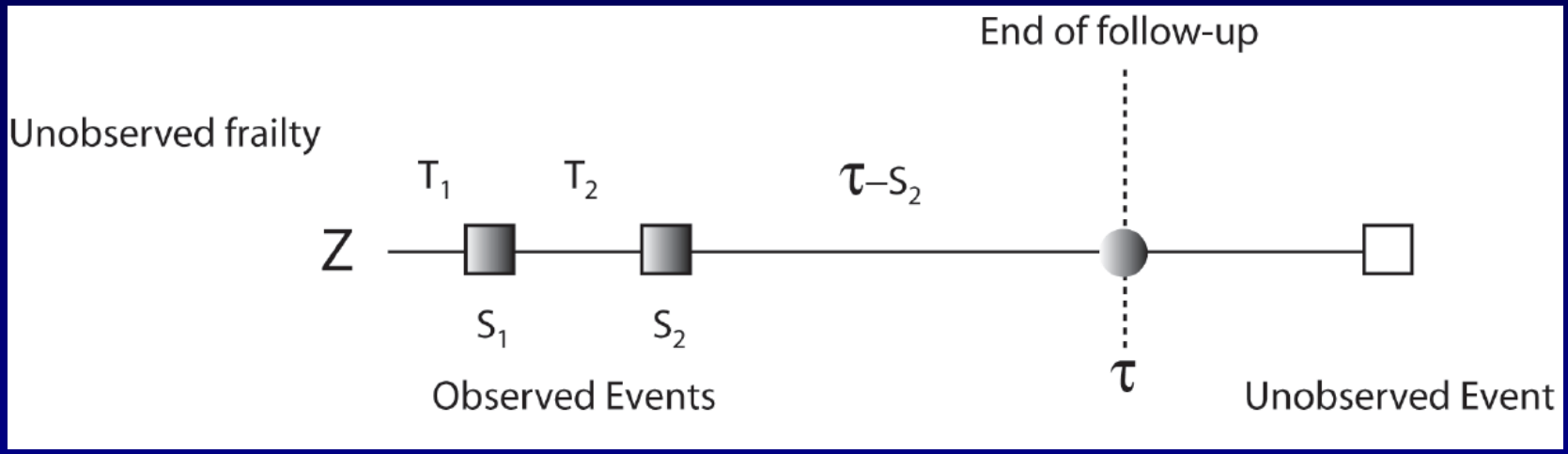
- MacLaughlin (2002), in the context of non-indolent lymphomas, pointed out that:

  "it is necessary a model designed specifically for relapsing patients"

- Duchateau et al. (2003) stated that:

  "Little attention, however, has been given to the timescale that is used for subsequent events and the interpretation attached to it."

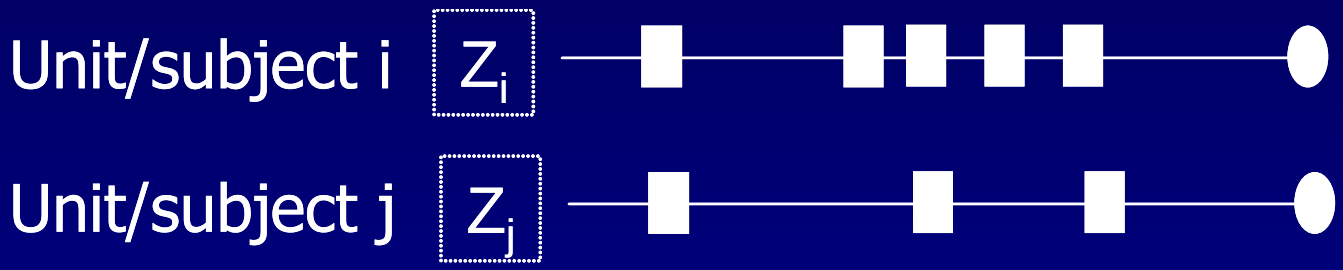
Data with recurrent events



- $T_1, T_2, \ldots$ = inter-event or gap times
- $S_1, S_2, \ldots$ = calendar times of event occurrences
- τ = end of observation period
- Z = unobserved frailty variable

- Accrued History: $F = \{F(s) : 0 \leq s\}$
- $N(s)$ = number of events in $[0, s]$
- $Y(s)$ = at-risk indicator at time s
- $X(s)$ = covariate vector, possibly time-dependent

Heterogeneity



- Biomedical data (unobserved variables, uncontrolled variables, non-measurable variables: genetic susceptibility, …)
- There exists a random variable Z with known distribution. If we condition to Z=z the inter-occurrence times are i.i.d.

Frailty Models

- Cox model must satisfy baseline proportional hazards assumption
  - What if have a mixture of hazards in the population or heterogeneity of the hazards?
- Frailty models are basically random effect survival models
- The random effects or frailty, described unexplained heterogeneity, the influence of unobserved risk factors in the model
- The concept of frailty may be used to explanin the unaccounted for heterogeneity which leads to the differential survival patterns of members of a population (Keiding et al., 1997, Vaupel et al., 1997)

Frailty model: $\lambda(s|x) = Z\ \lambda_0(s|x)\ exp(\beta'X)$

- The term Z, called frailty, varies from individual to individual and is not observable
- The distribution of z of the population, G(z), must be specified
- Hazard function is non-negative, hence, z, must be defined using a positive distribution (gamma, log-normal)
- Another way of looking at fratilty model formulation is:

$$\lambda(t|x) = \lambda_0(t|x)\ exp(\beta'X + e)$$

con $e=log(z)$

Example: unobserved variable

- Consider a situation with the following population:

Group	Proportion of Population	Hazard Rate with placebo	Hazard Rate with Drug A
1	40%	1	.5
2	40%	2	1
3	20%	10	8

- Obviously Drug A is effective for the entire population.
- But what happens in the Cox Model if the group is unobservable?

Example: unobserved variable

- Let's perform a simulation study

- Simulate k measurements for 100 people according to the previous table probabilities and randomly assign them to treatment or placebo

- For each person, have k simulate # of incidents within period of length 1. At the end of period of length 1 right-censoring occurs.

COX MODEL

```
 coef exp(coef) se(coef)     z    p
treat -0.206     0.814    0.135 -1.53 0.13
Likelihood ratio test=2.37  on 1 df, p=0.124
```

FRAILTY COX MODEL

```
coxph(Surv(time,status)~ treat + frailty(id), datos)
               coef   se(coef) se2   Chisq  DF   p
treat        -0.275 0.198    0.136   1.94  1.0 1.6e-01
frailty(id)                          117.60 44.3 1.4e-08

Variance of random effect= 0.462  I-likelihood = -1201.4
Degrees of freedom for terms=  0.5 44.3
Likelihood ratio test=131   on 44.7 df, p=2.37e-10   n=
332117.44 47.5 7.5e-08
```

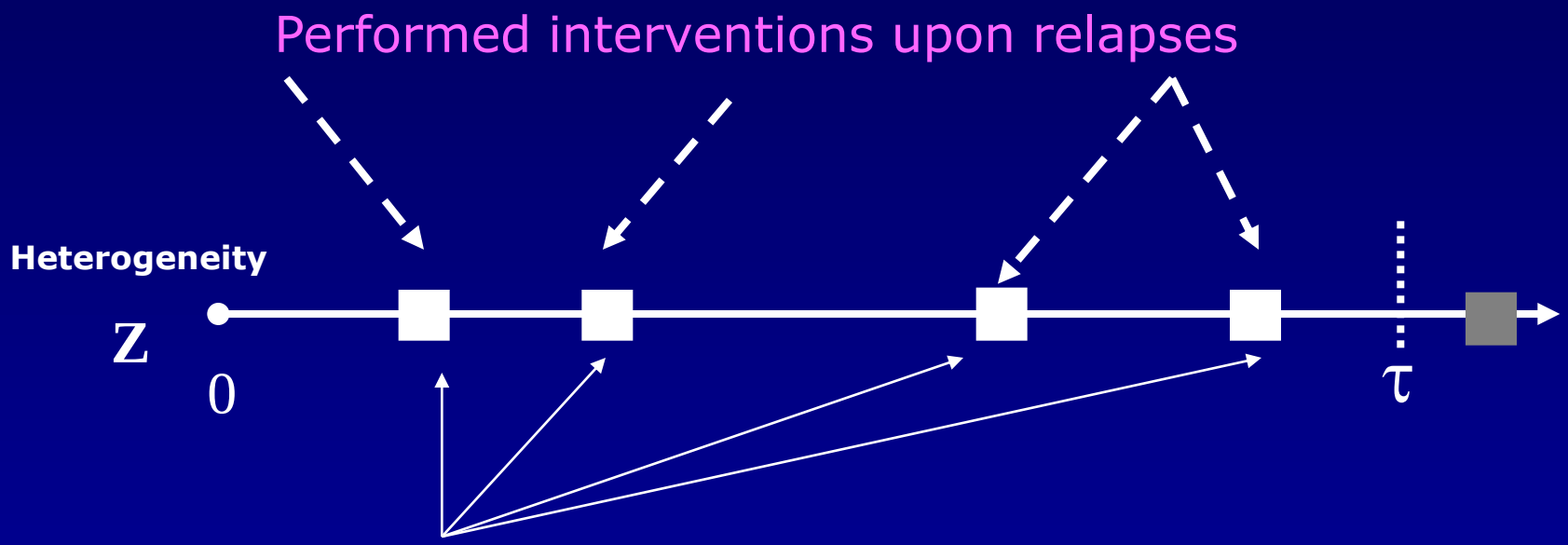
Frailty model: $\lambda(s|x) = Z\,\lambda_0(s|x)\exp(\beta'X)$

- Univariate frailty (per subject basis)
- Multivariate frailty (grouping factor)
  - families
  - siblings
  - countries/regions
  - …
- Repeated measurements (per subject basis)
  - Recurrent events
- Repeated measurements/grouping (2 nested groups)

Frailty model: $\lambda(s|x) = Z\,\lambda_0(s|x)\exp(\beta'X)$

- EM algorithm has some drawbacks
  - do not provide direct estimates of the variance of the parameters
  - jackknife estimation (time-consuming)
- Penalized likelihood inference
  - Penalized partial likelihood (Therneau et al.'s, 2003)
    - the penalty incorporates a regression coefficient
    - do not provide estimates of variance of frailty variance
  - Penalized full likelihood (Rondeau et al.'s, 2003)
    - gives an smooth estimates of baseline hazard function
    - obtains an estimation of variance of frailty variance
    - smoothing parameter

General class of models



Event dependence: accumulating events

Covariate vector: $\mathbf{X}(s) = (X_1(s), X_2(s), \ldots, X_q(s)^t$

Peña and Hollander (2004)

Intensity process

$$\lambda_i(s|\mathbf{Z},\mathbf{X}_i) = Z_i\,\lambda_0[\mathcal{E}_i(s)]\,\rho[N_i^{\dagger}(s-);\alpha]\,\psi[\beta^{\mathrm{t}}\mathbf{X}_i(s)]$$

Frailty: **Heterogeneity**

Baseline hazard

Models the **intervention** after events (Effective Age)

Models the event dependence (α^k):

Models the effect of **covariates**

Peña and Hollander (2004)

Effective age process

- Model I: Patient always achieves complete remission after each treatment (perfect repair). Gap time formulation **(superficial bladder cancer)**

$$\mathcal{E}_i(s) = s - S_{N\dagger}(s-)$$

- Model II: Patient does not (always) improve after each treatment, e.g. stable disease after each treatment (minimal repair). Calendar time formulation

$$\mathcal{E}_i(s) = s$$

Inference for the general class of models

- Model without frailties
  - Partial Likelihood process

$$L_p(s|\alpha,\beta) = \prod_{i=1}^{n} \prod_{j=1}^{N_i^{\dagger}(s)} \left[\frac{\rho(j-1;\alpha)\,\psi[\beta^{\mathrm{t}}\mathbf{X}_i(S_{ij})]}{S_0[s,\mathcal{E}_i(S_{ij})|\alpha,\beta]} \right]^{\Delta N_i^{\dagger}(S_{ij})}$$

  - α and β are estimated using Newton-Raphson iterative procedure

- Model with frailties
  - EM algorithm Penalized
  - Penalized partial likelihood
  - Penalized full likelihood

EM algorithm

Assuming $\rho(.;k)=\alpha^k$, the intensity for the general class of models becomes

$$\lambda_i(s|Z_i,\mathbf{X}_i) = Z_i\,\lambda_0[\mathcal{E}_i(s)]\,\alpha^{N_i^{\dagger}(s-)}\exp[\beta'\mathbf{X}_i]$$

- Frailties $Z_1, Z_2, \ldots, Z_n$ are unobserved and are IID Gamma(ξ,ξ)
- Unknown parameters: (α,β,ξ, $\lambda(\cdot)$)
- Use of the EM-type algorithm (Dempster, et al;Nielsen, et al), with frailties as missing observations.
- Standard errors via Jackknife

EM algorithm (Peña, Slate, González, 2006)

Step 0: (Initialization) Seed values $\hat{\xi}, \hat{\alpha}, \hat{\beta}$; no-frailty estimator $\hat{\Lambda}_0$.

Step 1: (E-step) Compute $\hat{Z}_i = E(Z_i | \text{data}, \hat{\xi}, \hat{\alpha}, \hat{\beta}, \hat{\Lambda}_0)$

Step 2: (M-step 1) New estimate of $\Lambda_0(\cdot)$ Form: analogous to the no-frailty case with $\hat{Z}_i$'s

Step 3: (M-step 2) New estimates of α and β.

Step 4: (M-step 3) New estimate of ξ; maximize marginal likelihood for ξ.

Step 5: Check for convergence.

Penalized partial likelihood

Assuming $\rho(.;k)=\alpha^k$, the intensity for the general class of models becomes

$$\lambda_i(s|Z_i,\mathbf{X}_i) = Z_i\,\lambda_0[\mathcal{E}_i(s)]\,\alpha^{N_i^\dagger(s-)}\exp[\beta'\mathbf{X}_i]$$

With the parameterization $Z_i=\exp(z_i)$

$$\lambda_i(s|z_i,\mathbf{X}_i) = \lambda_0[\mathcal{E}_i(s)]\,\alpha^{N_i^\dagger(s-)}\exp[\beta'\mathbf{X}_i + z'\mathbf{M}_i]$$

Where M is a matrix of n indicator variables such as $M_{ij}=1$ when observation I is a re-occurrence of individual j and 0 otherwise

Instead of maximize the partial likelihood we maximize the penalized partial likelihood

$$pl_P(s|\alpha,\beta,z) = l_P(s|\alpha,\beta,z) - g(z;\nu)$$

where g is a penalty function to restrict the values of z and ν is a constant

Penalized partial likelihood (cont.)

$$pl_P(s|\alpha,\beta,z) = l_P(s|\alpha,\beta,z) - g(z;\nu)$$

- To estimate α, β and z we may use a Newton-Raphson (N-R) procedure.
  - The scores for α, and β are similar to those obtained for the general class of models wit z as an offset
  - The score for z is similarly developed
  - The Hessian is now

$$H = H(\alpha,\beta,w) = \mathcal{I} + \begin{pmatrix} 0 & 0 \\ 0 & g'' \end{pmatrix}$$

- The algorithm consists in an inner and outer loop
  - For a fixed ν N-R is used to solve the penalized likelihood
    - Usually a few steps (5 or 6)
    - The information matrix has many parameters
  - The outer loop chooses ν to maximize the profile likelihood

Penalized full likelihood

- To penalize the full log-likelihood

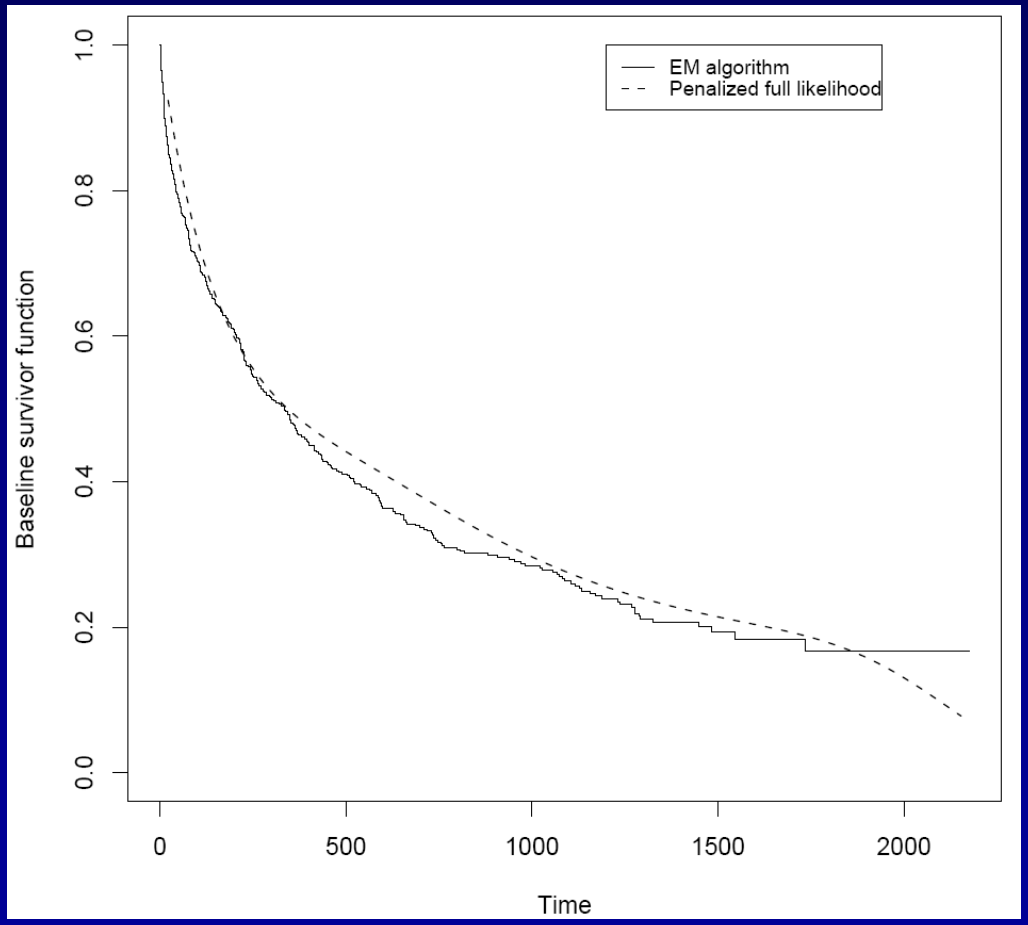
$$pl_F(\alpha, \beta, \Lambda_0(\cdot), z) = l_F(\alpha, \beta, \Lambda_0(\cdot), z) - \kappa \int_0^\infty (\lambda_0'')^2(t)dt$$

- κ is a positive smoothing parameter which control the trade-off between the data fit and the smoothness of the function
- λ_o is approximated by a linear combination of m cubic M-splines

$$\tilde{\lambda}(.) = \sum_{i=1}^{m} \eta_j M_j(.)$$

- We need to estimate η as well as α and β
- In that case, neither the score nor the hessian have a simple analytical form: Marquart algorithm
- Smoothing parameter: cross-validation
- This method provides an estimation of variance of frailty variance

Application to hospital readmission data set

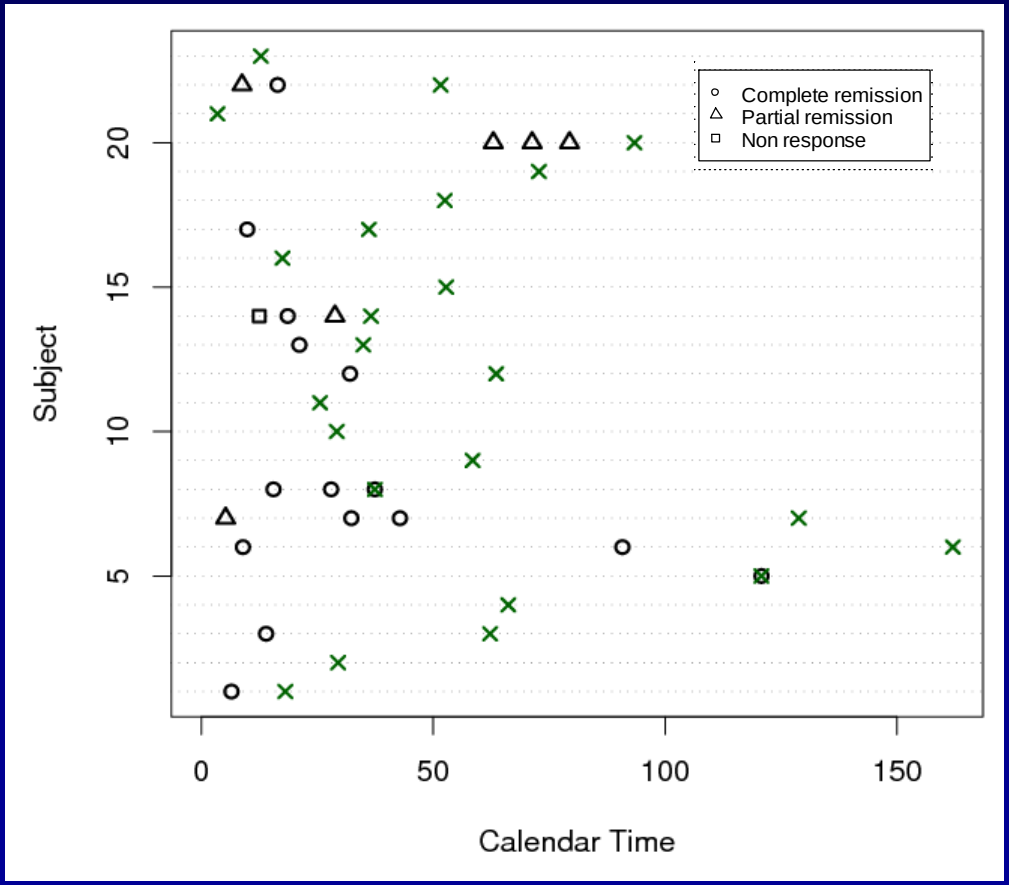


Application to hospital readmission data set

	Penalized approach		EM approach
	Partial Likelihood	Full Likelihood	Jackknife
Covariate	HR (95%CI)	HR (95%CI)	HR (95%CI95)
Gender			
Female	1	1	1
Male	1.56 (1.17-2.09)	1.63 (1.25-2.13)	1.56 (1.20-2.03)
Dukes stage			
A-B	1	1	1
C	1.49 (1.11-2.00)	1.51 (1.14-2.01)	1.49 (1.08-2.05)
D	3.06 (2.14-4.39)	3.41 (2.39-4.86)	3.06 (2.04-4.58)
log N(s-) α	1.08 (0.97-1.19)	1.14 (1.03-1.24)	1.08 (0.81-1.35)
Frailty ν	0.40	0.58	0.40
(SE ν)	(NA)	(0.15)	(NA)
ξ	2.50	1.72	2.50
κ		3.36×10^{11}	

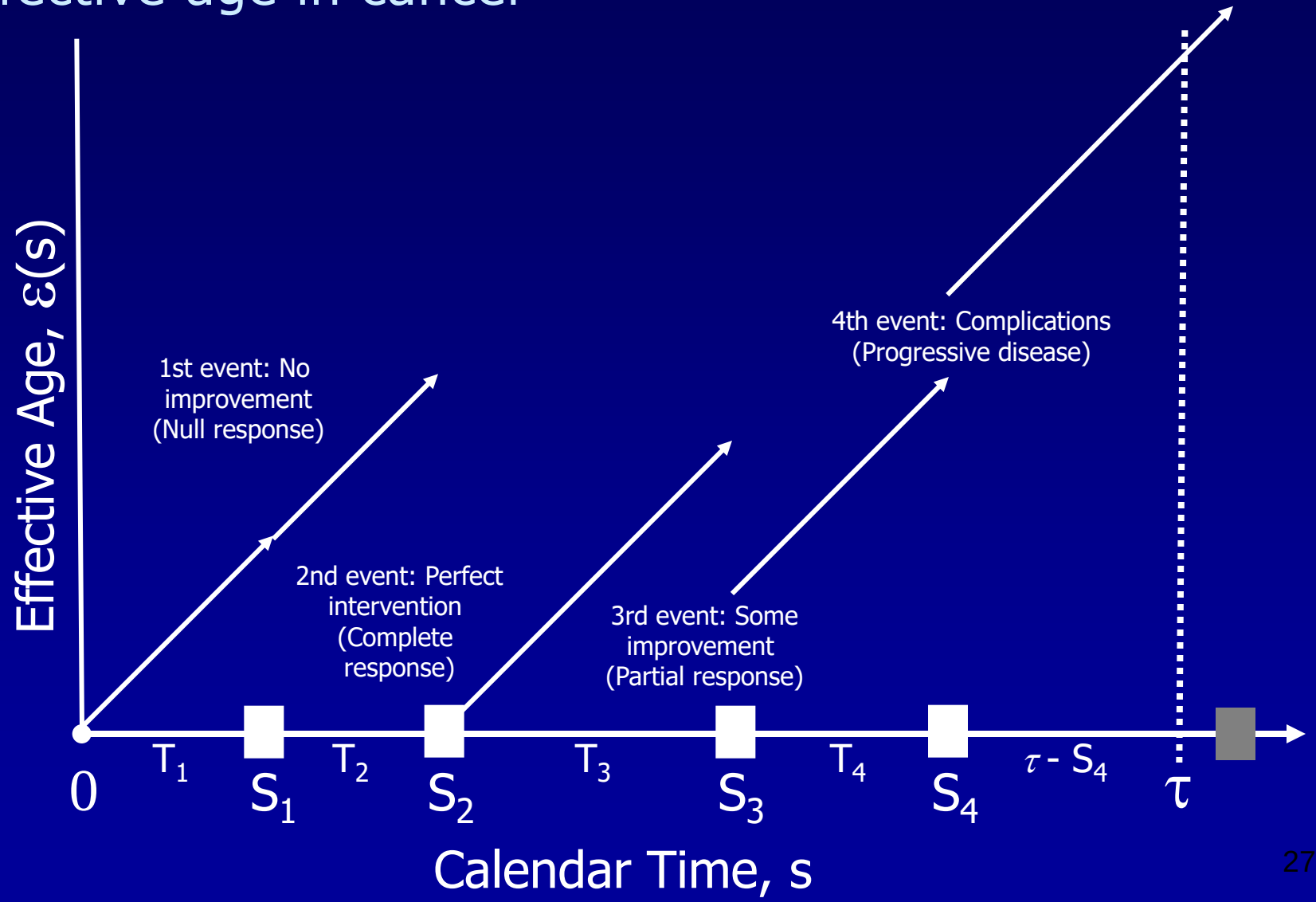
Example 2: Non-Hodgkin lymphomas

Time to next hospital relapse after diagnosis in cancer patients



What is the effect of the response to the treatment on the time of next relapse?

Effective age in cancer



Possible effective ages in cancer data

- Model I: Patient always achieves complete remission after each treatment (perfect repair). Gap time formulation (superficial bladder cancer)

$$\mathcal{E}_i(s) = s - S_{N^\dagger(s-)}$$

- Model II: Patient does not (always) improve after each treatment, e.g. stable disease after each treatment (minimal repair). Calendar time formulation

$$\mathcal{E}_i(s) = s$$

Possible effective ages in cancer data

- Model III: Patient achieves null, partial or complete response in different recurrences

$$A_0 = 0, A_j = (1 - \psi_j)(A_{j-1} + T_j)$$

$$\psi_{i \geq 1} = \{0, 0.5, 1\}$$

NR PR CR

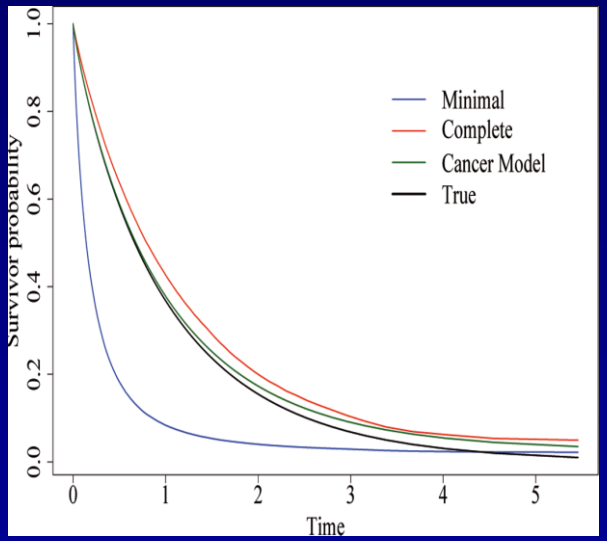
$$\mathcal{E}(s) = A_{N^\dagger(s-)} + \left(s - S_{N^\dagger(s-)}\right)$$

Simulation study

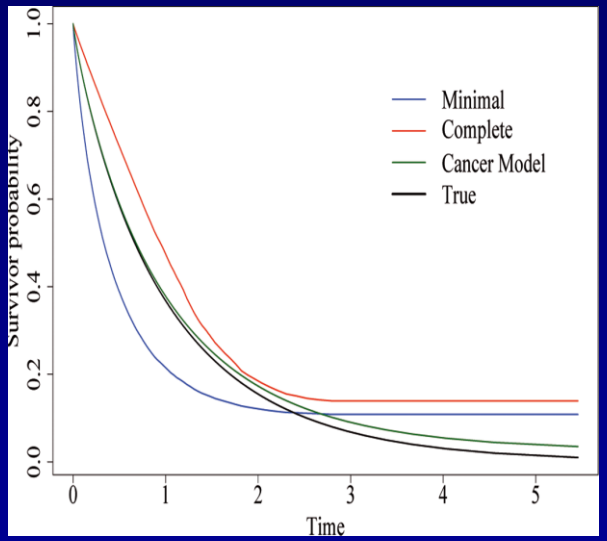
- Design
  - Different sample sizes
  - Censoring variables
  - Baseline hazard function
  - Impact of accumulating events
  - Effect of covariates
  - Heterogeneity
  - Effective age: patient achieves complete, partial or null response depending on the vector ψ taking values in the set {1=CR, 0.5=PR, 0=NR}: {(.8,.1,.1), (.3,.5,.2), (.1,.2,.7)}
- Aim
  - The effect of sample size
  - Bias and variance of the estimators
  - The performance of the estimators ot the λ_0 in terms of bias and RMSE
  - The consequences of miss-specifying the effective age function

Simulation study. Miss-specifying effective age, λ_0 estimates

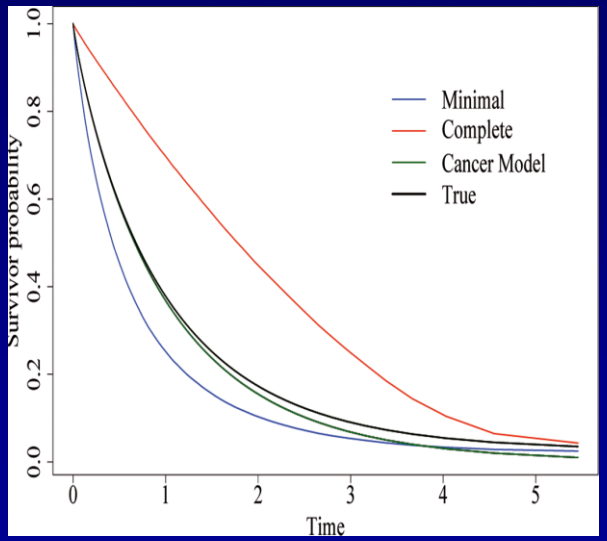
$P(\psi)=(0.8,0.1,0.1)$



$P(\psi)=(0.3,0.5,0.2)$



$P(\psi)=(0.1,0.2,0.7)$



Application to non-Hodgkin's lymphoma data set

	Minimal[a]	Perfect[b]	Cancer[c]
α	0.72 (.37)	0.90 (.15)	0.68 (0.21)
Frailty ξ	2.24	∞	1.36
ν	0.45	8.97×10^{-8}	0.73
Lesions			
Single	1	1	1
Localized	2.48 (0.93-6.60)	2.42 (0.90-6.50)	2.59 (0.88-7.62)
>1 nodal site	4.47 (1.64-12.15)	3.26 (1.20-8.87)	4.55 (1.22-16.93)
Generalized	3.24 (0.70-14.95)	2.69 (0.59-12.31)	3.09 (1.01-9.37)

[a] Effective Age is $\mathcal{E}(s) = s$.
[b] Effective Age is $\mathcal{E}(s) = s - S_{N^\dagger(s-)}$.
[c] Effective Age is $A_{N^\dagger(s-)} + \left(s - S_{N^\dagger(s-)}\right)$.

Dynamic cancer model

Application to non-Hodgkin's lymphoma data set. Single lesions (solid lines) versus multiple lesions (dotted lines)