Week 4: Bivariate Joint Model for a Recurrent Event and a Terminal Time-to-Event Outcome

Laurent Briollais (laurent@lunenfeld.ca)

Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital
Dalla Lana School of Public Health, UofT

May 27, 2020

Previous lecture

- Case Study = AIDS data
- Case Study = PBC data
- Extensions of the Bivariate Joint Models
- Presentation of the *R* package JM

Outline of this lecture

- Bivariate Joint Model for a Recurrent and a Terminal Time-to-Event Outcomes in Frailtypack
- Presentation of the *R* package Frailtypack
- Case study: Readmission data
- Joint model for recurrent events and a terminal event with two frailty terms
- Joint model for two clustered survival events

References

- **Joint Models for Longitudinal and Time-to-Event Data With Applications in** *R* from Dimitris Rizopoulos. Chapman & Hall/CRC biostatistics series. CRC Press, Taylor & Francis Group. 2012.

- A. Krol, A. Mauguen, Y. Mazroui, A. Laurent, S. Michiels, V. Rondeau. **Tutorial in Joint Modeling and Prediction: A Statistical Software for Correlated Longitudinal Outcomes, Recurrent Events and a Terminal Event.**

1 Bivariate Joint Model for a Recurrent and a Terminal Time-to-Event Outcomes

Goal

- Model dependence between two time-dependent outcomes, e.g. a terminal event and a recurrent event.
- Study the impact of covariates both on recurrent events and death
- To treat informative censoring by terminal event (e.g. death)

1.1 Notations

T_i^* : true time of terminal event for subject i, $\quad i = 1, \cdots, N$
T_i : observed terminal event time for patient i, i.e., $T_i = \min(C_i; T_i^*)$
δ_i : Event indicator for the terminal event, i.e., $\delta_i = I_{\{T_i = T_i^*\}}$
T_{ij}^* : true time of the jth recurrent event for subject i
T_{ij} : observed time of the jth recurrent event for subject i
$\Rightarrow T_{ij} = \min(T_{ij}^*, C_i, T_i^*)$ and $\delta_{ij} = I_{\{T_{ij} = T_{ij}^*\}}$
$X_{R_{ij}}$ and X_{T_i} : covariate vectors for recurrent and terminal events
r_0 and λ_0 : baseline hazards for risk of recurrent and terminal events

replace longitudinal measurements with recurrent event

1.2 Joint modeling for the risk of recurrent event and terminal event

baseline hazard

hazard for recurrent event $r_{ij}(t|u_i) = u_i r_0(t) \exp(X_{R_{ij}}^T \beta_R)$,

hazard for terminal event $\lambda_i(t|u_i) = u_i^\alpha \lambda_0(t) \exp(X_{T_i}^T \beta_T)$,

mean of distribution = 1 and variance is theta

- Frailty $u_i \sim \Gamma(1/\theta, 1/\theta)$, i.e. $E(u_i) = 1$ and $\text{var}(u_i) = \theta$.

  it can account for...

  - heterogeneity between subjects associated with unobserved prognostic factors
  - within-subject correlation for the dependence between the recurrent events
  - association between recurrent events and terminal event $\Rightarrow$ power term α association between the two processes

  significant of association between the two components if it's significant

1.3 Likelihood

Parameters to estimate: $\xi = (r_0(.), \lambda_0(.), \beta_R^T, \beta_T^T, \theta, \alpha)$

conditioning on the random effects and integrate the random effects

$$L_i(\xi) = \int_{u_i} \prod_{j=1}^{n_i} f(T_{ij}, \delta_{ij}|u_i; \xi) f(T_i, \delta_i|u_i; \xi) f(u_i; \xi) du_i$$

$$= \int_0^{+\infty} \prod_{j=1}^{n_i} \left[r_{ij}(T_{ij}, \delta_{ij}|u_i; \xi)^{\delta_{ij}} S_{ij}(T_{ij}|u_i; \xi) \right] \lambda_i(T_i, \delta_i|u_i; \xi)^{\delta_i} S_i(T_i|u_i; \xi) f(u_i; \xi) du_i$$

$S_i(j)$: conditional survival function for terminal event

$S_{ij}(j)$: conditional survival function for jth recurrent event

n_i : number of recurrent events for individual i

- Marginal log-likelihood (gamma distribution of the frailty)
  likelihood over all individuals - conditioning on random effects

$$l(\xi) = \sum_{i=1}^{N} \Big(\sum_{j=1}^{n_i} \delta_{ij} \log r_{ij}(T_{ij}|u_i) + \delta_i \log \lambda_i(T_i|u_i) - \log \Gamma(1/\theta) - \frac{1}{\theta} \log \theta$$

$$+ \log \int_0^{+\infty} u_i^{(\sum_{j=1}^{n_i} \delta_{ij} + \alpha\delta_i + 1/\theta - 1)} \exp^{-u_i \left(\int_0^{T_i} r_{ij}(t|u_i)dt - u_i^{\alpha} \int_0^{T_i} \lambda_i(t)dt - \frac{u_i}{\theta} du_i \right)}$$

1.4 Estimation that are used to maximize the log-likelihood

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

- Baseline hazard functions: $r_0(t)$ and $\lambda_0(t)$

Let $h_0(t)$ be a baseline hazard function.

- Weibull baseline hazards (parametric)
  - $h_0(t) = (at^{a-1})/b^a$
  - with a > 0 the shape parameter and b > 0 the scale parameter

– Piecewise constant baseline hazards (parametric)

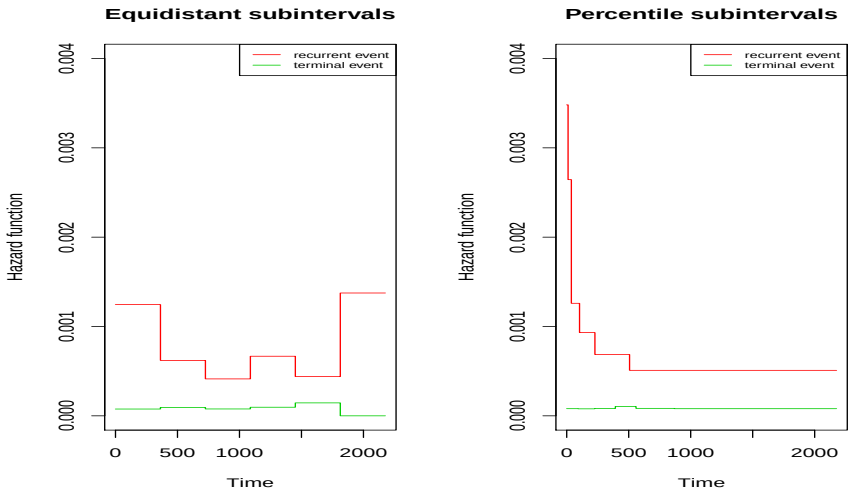
$$h_0(t) = \sum_{i=1}^{n_{int}} I_{\{t\in(t_{i-1},t_i)\}} c_i$$

estimate specific parameter for each interval

The interval $[0, \tau]$ (with the last observed time among N individuals) is divided into n_{int} 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)

– Piecewise constant baseline hazards - Example ($n_{int} = 6$)



– Baseline hazards approximated with splines (semi-parametric)

* cubic M-splines : adapted for the approximation of hazard functions (Ramsay, 1988)

$$h_0(t) = \sum_{i=1}^{m} \zeta_i M_i(.).$$

with $M_i(.)$ cubic M-splines of order 4,
$\zeta = (\zeta_1, \cdots, \zeta_m)$ is the vector of splines coefficients,
$m = T + 2$ with T the number of knots.

  - ∗ Antiderivatives, called I-splines : useful for approximating the cumulative hazard function using the same spline parameters ζ
  - ∗ We can use equidistant or percentiles knots

integral of spline is also a spline function - equal distance and something

1.5 Estimation of the model

- Maximization of log-likelihood : Marquardt algorithm (Marquardt, 1963)
  - Combines the Newton-Raphson and the steepest descent algorithms
  - More stable behavior than the Newton-Raphson algorithm in complex problems while preserving fast convergence
  - Three conditions for convergence
    - ∗ whether the coefficients are stable (difference between two consecutive iterations $< 10^{-4}$)
    - ∗ a condition on the difference between values of the penalized log-likelihood between two iterations ($< 10^{-4}$)
    - ∗ a condition on the gradient of the penalized log-likelihood, whether it is small enough ($< 10^{-4}$)
- Variances of the estimated parameters are directly obtained from the inverse of the Hessian matrix

1.6 Maximum penalized likelihood (MPL) when using splines for baseline hazards

- In medical studies, it is natural to assume a smooth hazard function with low local variations $\rightarrow l(\xi)$ is penalized by a term that has large values for rough functions :

- Square norm of the 2nd derivatives of the baseline hazard functions and smoothing parameters $\kappa > 0 \rightarrow l(\xi)$ smooth estimations privileged

- The penalized log-likelihood is defined as :

$$pl(\xi) = l(\xi) - \kappa_1 \int_0^{+\infty} r_0''^2(t)dt - \kappa_2 \int_0^{+\infty} \lambda_0''^2(t)dt$$

  κ_1 and κ_2 control the trade-off between the fit to the data and the smoothness of the hazard functions.

- Choice of smoothing parameters

  No automatic algorithm. But one can choose :

  – Graphically by assessing the smoothness of several estimated baseline hazard functions and selecting the one which seems the most realistic
  – Using an approximated cross-validation criterion applied to corresponding reduced models (shared frailty model for the recurrent events and the Cox model for the survival) (O'Sullivan, 1988 ; Rondeau et al., 2003)

1.7 Model fit evaluation

– Approximate likelihood cross-validation criterion (LCV)

– Measures the relative goodness of fit among a collection of models

– Lower values indicate a better fit

$$LCV_a = \frac{1}{\sum_{i=1}^{N} n_i}(\text{trace}(H_{pl}^{-1} H_l) - l(.))$$

  - H_{pl} is minus times the converged hessian of the penalized log-likelihood,
  - H_l minus times the converged hessian of the log-likelihood
  - $l(.)$ is the full log-likelihood

- Martingale residuals difference is now we have frality term
assess the model fit
  - Models whether the number of observed events is correctly predicted by a model
  - Based on the counting processes theory
  - The martingale residuals can be expressed by (recurrences)

$$M_i(t) = N_i(t) - \hat{u}_i \int_0^t W_i(s)\hat{r}_i(s)ds$$

  where $W_i(t)$ is equal to 1 if the individual is at risk of the event at time t and 0 otherwise, $N_i(t)$ is number of events until t, and $\hat{r}_i(t) = u_i r_i(t|u_i)$.
- Calculated at T_i, the mean of the martingale residuals at a given time point should be equal to 0

1.8 Time-dependent covariates

- B-splines

$$\hat{\beta}(t) = \sum_{j=-q+1}^{m} \hat{\zeta}_j B_{j,q}(t)$$

  where $B_{j,q}(t)$ is the basis of B-splines calculated using a recurring expression

- quadratic B-splines, i.e., $q = 3$, with a small number of interior knots ($m \leq 5$) ensure stable estimation

- helpful to verify the proportional hazard (PH) assumption using a likelihood ratio (LR) test
- time-dependency tested for both events :

$$H_0 : \beta_R(t) = \beta_R, \quad \beta_T(t) = \beta_T \text{ vs. } H_1 : \beta_R(t) \neq \beta_R, \quad \beta_T(t) \neq \beta_T$$

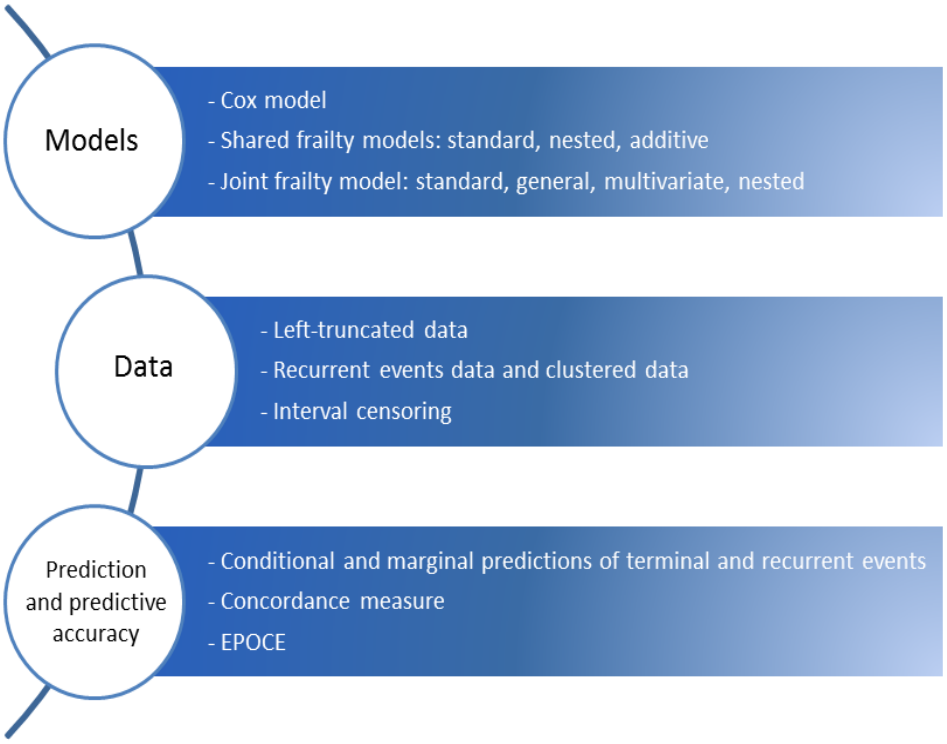
The LR statistic has a χ^2 distribution of degree $2 \times (m + q - 1)$

1.9 Available softwares

Limited software, to my knowledge :

- SAS: proc nlimixed
- *R* packages:
  - joint.Cox : MPL estimation under the copula-based joint Cox models for time to clustered events (progression) and a terminal event
  - frailtypack

2 *R* package *frailtypack*



(Rondeau and Gonzalez, 2005 ; Rondeau et al., 2012 ; Krol et al., 2017)

2.1 General features

* In the package frailtypack there are four different functions for the estimation of joint models, one for each model.

* The joint models for recurrent events and a terminal event (and models for two clustered survival outcomes) as well as the general joint frailty models are estimated with function frailtyPenal.

* The joint models for two recurrent events and a terminal event are estimated with multivPenal.

* The estimation of joint models for a longitudinal outcome and a terminal event is performed with longiPenal.

* Finally, function trivPenal estimates trivariate joint models.

* All these functions make calls to compiled Fortran codes programmed for computation and optimization of the log-likelihood.

2.2 *R* function *frailtyPenal*

```
frailtyPenal(formula, formula.terminalEvent, data, recurrentAG = FALSE,
cross.validation = FALSE, jointGeneral, n.knots, kappa, maxit = 300,
hazard = "Splines", nb.int, RandDist = "Gamma", betaknots = 1,
betaorder = 3, initialize = TRUE, init.B, init.Theta, init.Alpha, Alpha,
init.Ksi, Ksi, init.Eta, LIMparam = 1e-3, LIMlogl = 1e-3, LIMderiv = 1e-3,
print.times = TRUE)
```

* The argument formula is a two-sided formula for a survival object Surv from the survival package (Therneau 2017) and it represents the recurrent event process (the first survival outcome for the joint models in case of clustered data) with the combination of covariates on the right-hand side, the indication of a grouping variable (with term cluster(group)) and the indication of the variable for the terminal event (e.g., terminal(death)).

* It should be noted that the function cluster(x) is different from that included in the package survival. In both cases it is used for the identification of the correlated groups but in frailtypack it indicates the application of the frailty model and in survival, a GEE (generalized estimating equations) approach is used, without random effects.

* The argument formula.terminalEvent requires the combination of covariates related to the terminal event on the right-hand side.

* The logical argument recurrentAG indicates whether the calendar timescale for recurrent events or clustered data with the counting process approach of Andersen and Gill (1982) (TRUE) or the gap timescale (FALSE; by default) is to be used.

* The argument for the cross-validation cross.validation is not yet implemented for the joint models; thus its logical value must be FALSE. The smoothing parameters κ for a joint model can be chosen by first fitting suitable shared frailty and Cox models with

the cross-validation method.

* The general joint frailty models can be estimated if argument jointGeneral is TRUE. In this case, the additional gamma frailty term is assumed and the parameter is not considered. These models can be applied only with the Gamma distribution for the random effects.

* The type of approximation of the baseline hazard functions is defined by argument hazard. Options are "Splines" for semi-parametric functions using equidistant intervals, "Splines-per" using percentile intervals, "Piecewise-equi" and "Piecewise-per" for piecewise constant functions using equidistant and percentile intervals, respectively and "Weibull" for the parametric Weibull baseline hazard functions.

* If either "Splines" or "Splines-per" is chosen for the baseline hazard functions, arguments kappa for the positive smoothing parameters and n.knots should be given with the number of knots chosen between 4 and 20 which corresponds to n.knots+2 splines functions for approximation of the baseline hazard functions (the same number for hazard functions for both outcomes).

* If "Percentile-equi" or "Percentile-per" is chosen for the approximation, argument nb.int should be given with a 2-element vector of numbers of time intervals (1 to 20) for the two baseline hazard functions of the model.

* Argument RandDist represents the type of the random effect distribution, either "Gamma" for the Gamma distribution or "LogN" for the normal distribution (log-normal joint model). If it is assumed that in the joint model that the parameter α equal to zero, the argument Alpha should be set to "None".

* In case of time dependent covariates, the arguments betaknots

and betaorder are used for the number of inner knots used for the estimation of B-splines (1, by default) and the order of B-splines (3 for quadratic B-splines, by default), respectively.

* The rest of the arguments are allocated for the optimization algorithm. Argument maxit declares the maximum number of iterations for the Marquardt algorithm.

* Arguments init.B, init.Theta, init.Eta and init.Alpha are vectors of initial values for regression coefficients, variances of the random effects and for the α parameter (by default, 0.5 is set for all the parameters).

3 Case study and R code examples

3.1 Readmission data set

* Rehospitalization study of patients after a surgery and diagnosed with colorectal cancer (Gonzalez et al. 2005 ; Rondeau et al. 2012)

* Information on times (in days) of successive hospitalizations and death (or last registered time of follow-up for right-censored patients) counting from date of surgery

* Patients characteristics : type of treatment, sex, Dukes' tumoral stage, comorbidity Charlson's index and survival status

* Study includes 403 patients with 861 rehospitalization events in total

* 112 (28%) patients died during the study

* Justification for a joint frailty model for this data ?

```
> data("readmission")
> head(readmission)
  id enum t.start t.stop time event      chemo    sex dukes charlson death
1  1    1       0     24   24     1    Treated Female     D        3     0
2  1    2      24    457  433     1    Treated Female     D        0     0
3  1    3     457   1037  580     0    Treated Female     D        0     0
4  2    1       0    489  489     1 NonTreated   Male     C        0     0
5  2    2     489   1182  693     0 NonTreated   Male     C        0     0
6  3    1       0     15   15     1 NonTreated   Male     C        3     0
```

dukes: C, D: very aggressive cancer. A/B is less aggressive cancer.
death: died or not during the follow-up time
by default, gap time is implented
something needs to be true in order for to use the calender time

3.2 Implementation of joint frailty model

– We fit a model that includes Dukes' stage, Charlson's index, sex and treatment as covariates for both hospitalizations and death. The frailties are from the Gamma distribution (default option) and the baseline hazard functions are approximated by splines with 8 knots and the smoothing parameters for the penalized log-likelihood are $2.11e+8$ for the recurrent part and $9.53e+11$ for the terminal part.

```
# load the package
library(frailtypack)
# load the data
data(readmission)

# fit a joint frailty model using gap time and
   splines for baseline hazards
modJoint.gap <- frailtyPenal(Surv(time, event) ~
  cluster(id) + dukes + charlson + sex + chemo +
  terminal(death), formula.terminalEvent = ~ dukes
  charlson + sex + chemo, data = readmission,
  n.knots = 8, kappa = c(2.11e+08, 9.53e+11))
```

cluster is individual event
n.knots for the spline, 8 for terminal event and 8 for recurrent event

– Output

gap time is the default - for interpretation

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


  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
dukesC        0.344028  1.410617    0.165718      0.165718  2.07598 3.7896e-02
dukesD        1.260629  3.527639    0.206285      0.206285  6.11111 9.8938e-10
charlson1-2   0.410971  1.508282    0.256589      0.256589  1.60167 1.0923e-01
charlson3     0.413175  1.511609    0.137136      0.137136  3.01289 2.5878e-03
sexFemale    -0.537464  0.584228    0.141488      0.141488 -3.79865 1.4549e-04
chemoTreated -0.154436  0.856899    0.147137      0.147137 -1.04961 2.9390e-01

           chisq df global p
dukes    41.6554  2 9.01e-10
charlson 11.6428  2 2.96e-03
```

baselines

higher stage increases the chance of hospitalization

female has lower chance of hospitalization

```
Terminal event:
----------------
                  coef exp(coef) SE coef (H) SE coef (HIH)        z          p
dukesC        1.318393  3.737410    0.367781      0.367781 3.584724 3.3743e-04
dukesD        3.185326 24.175173    0.432573      0.432573 7.363668 1.7897e-13
charlson1-2   0.507129  1.660516    0.634718      0.634718 0.798983 4.2430e-01
charlson3     1.274563  3.577139    0.257553      0.257553 4.948735 7.4698e-07
sexFemale    -0.234725  0.790789    0.230204      0.230204 -1.019638 3.0790e-01
chemoTreated  1.096094  2.992456    0.258081      0.258081 4.247090 2.1657e-05

           chisq df global p
dukes    61.1967  2 5.14e-14
charlson 24.4911  2 4.81e-06

 Frailty parameters:
   theta (variance of Frailties, w): 0.738406 (SE (H): 0.105178 ) p = 1.1049e-12
   alpha (w^alpha for terminal event): 0.861979 (SE (H): 0.250164 ) p = 0.00056968

   penalized marginal log-likelihood = -4133.31
   Convergence criteria:
   parameters = 1.18e-05 likelihood = 4.88e-05 gradient = 1.21e-07

   LCV = the approximate likelihood cross-validation criterion
         in the semi parametric case     = 4.84003

   n observations= 861  n subjects= 403
   n recurrent events= 458
   n terminal events= 109
   n censored events= 403
   number of iterations:  9

   Exact number of knots used:  8
   Value of the smoothing parameters:  2.11e+08 9.53e+11
```

higher duke stage - higher stage - higher risk of dying

higher Charlson - higher risk of dying

having chemo increases the risk of dying, it's very significant

recurrent events are correlated within individuals

alpha - recurrent event and terminal event is very correlated - motivation for JM

lower critaya??? - better fit

– With this model, it is found that chemotherapy is a prognostic factor only on death with a positive association (HR = 2.99, p $<$ 0.001). Both Charlson's index (Index 3 vs. Index 0) and Dukes stage (Stage C and D vs. Stages A, B) are positively related to the recurrent and terminal events.

– The recurrent event (hospitalizations) is strongly associated to the terminal event (death) with $\alpha = 0.86$ and $p = 0.00057$.

– To verify whether the model predicts correctly the number of observed events, we represent the martingale residuals for both events against the follow-up time. These residuals in a well adjusted model should have a mean equal to 0 and thus a smoothing curve added to a graph should be approximately overlapping with the horizontal line $y = 0$.

```
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")
plot(aggregate(readmission$t.stop, by = list(readmission$id),
        FUN = max)[2][, 1], modJoint.gap$martingaledeath.res, ylab = "",
        xlab = "time", main = "Death", ylim = c(-2, 2))
lines(lowess(aggregate(readmission$t.stop, by = list(readmission$id),
        FUN = max)[2][, 1], modJoint.gap$martingaledeath.res, f = 1), lwd = 3,
        col = "grey")
```

– For the rehospitalization process the mean of residuals is approximately 0 with the smooth curve close to the line y = 0, but in case of death this tendency is deviated by relatively higher values for short follow-up times. This may suggest, that the model may have underestimated the number of deaths in the early follow-up period.

– The identified individuals of which the residuals result in non-zero mean, have short intervals between their rehospitalization and death

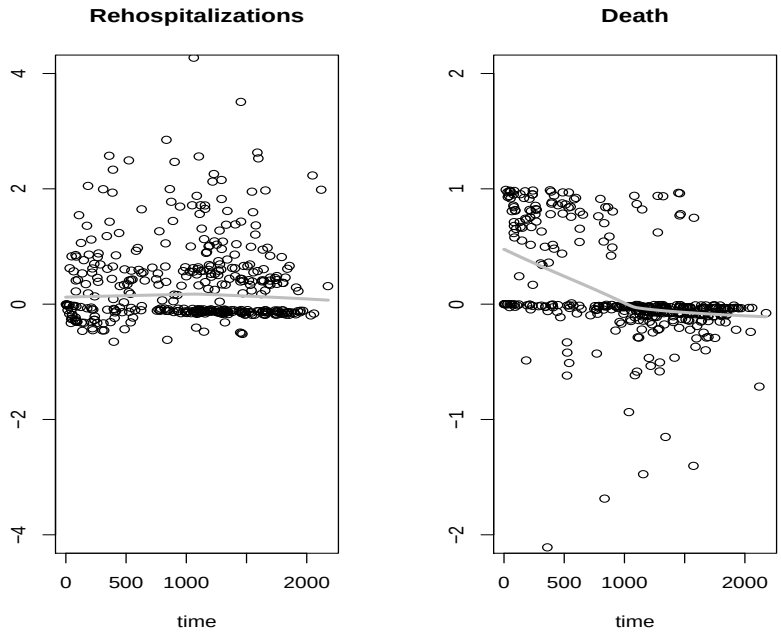


Figure 1: Martingale residuals for rehospitalizations and death against the follow-up time (in days). The grey line corresponds to a smooth curve obtained with lowess.

(1 day). Indeed, the removal of these individuals (50 patients) results in residuals with the mean close to 0 all along the follow-up period (plot not shown here).

to have a good fit, residuals need to close to 0.
see residuals over time
very good fit for recurrent event
deviation from 0 for people with a short time here for terminal event
time is between last hospitalization to time-to-death
model for terminal event is not very specified
may be because of heterogeity in the sample in the terminal event - died early in the hospitalization

– The package *frailtypack* provides also the estimation of the random effects. The vector frailty.pred from a jointPenal object contains the individual empirical Bayes estimates. They can be graphically represented for each individual with additional information on number of events (point size) to identify the outlying data.

```
plot(1:403, modJoint.gap$frailty.pred, xlab = "Id of patients",
        ylab = "Frailty predictions for each patient", type = "p", axes = FALSE,
        cex = as.vector(table(readmission$id)), pch = 1, ylim = c(-0.1, 7),
        xlim = c(-2, 420))
axis(1, round(seq(0, 403, length = 10), digit = 0))
axis(2, round(seq(0, 7, length = 10), digit = 1))
```

two individuals that are seem very sensitive

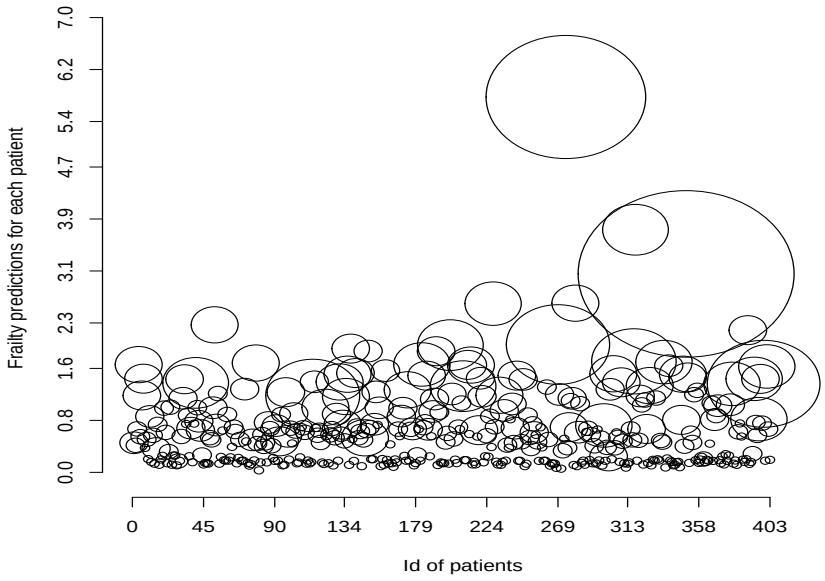


Figure 2: Individual prediction of the frailties. The size of points corresponds to number of an individuals recurrent events.

– Figure 4 shows the values of frailty prediction for each patient with an association to the number of events (the bigger the point, the greater the number of rehospitalizations). The frailties tend to have bigger values if the number of events of a given individual is high.

From the plot it can be noticed that there is an outlying frailty suggesting verification of the follow-up of the concerned individual.

– To find the optimal number of knots, we first fitted the model with a small number of knots (n.knots = 4) and increased the number of knots until the graph of the baseline hazard functions was not changing importantly anymore. The smoothing parameters are obtained from a shared frailty and Cox model with respectively recurrent and terminal event as the outcome using the cross-validation method.

```
plot(modJoint.gap)
```

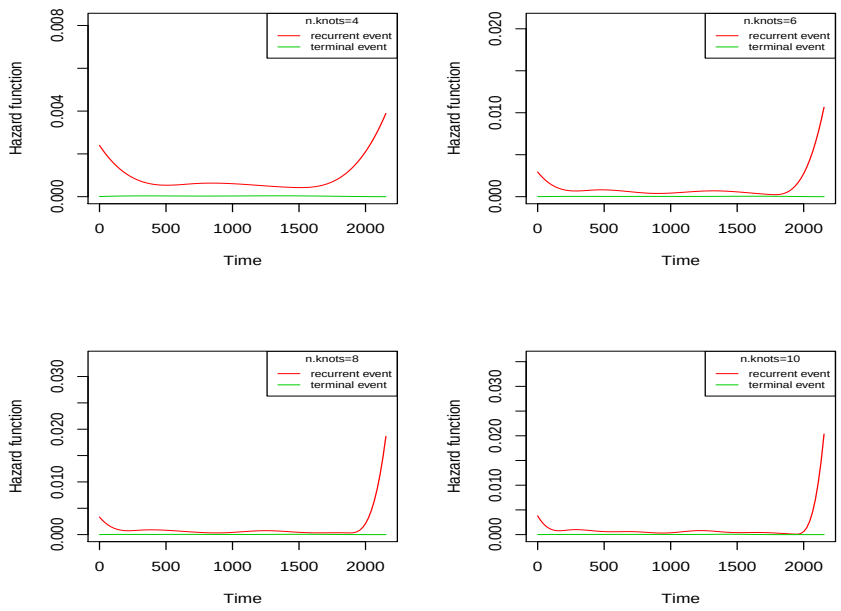


Figure 3: Changing the number of knots

shape changes when knots changed
you want a function that's relatively stable
between 8 and 9: shapes become stable
use 8 knots is probably enough
computing the SCV-kartaya???
determine how many knots you want to have for the spline function

– The smoothing parameter(s) κ can be chosen by using the cross.validation argument from a shared frailty model with a single time to event outcome.

you have to estimate kappa for shared frailty model for the recurrent event and terminal event separately

```
CoxMod.rec <- frailtyPenal(Surv(time, event) ~ cluster(id) + dukes +
                             charlson + sex + chemo, data = readmission,
                             n.knots = 8, kappa=5000, cross.validation =T)
CoxMod.rec
```

```
> CoxMod.rec
Call:
frailtyPenal(formula = Surv(time, event) ~ cluster(id) + dukes +
    charlson + sex + chemo, data = readmission, cross.validation = T,
    n.knots = 8, kappa = 5000)


  Shared Gamma Frailty model parameter estimates
  using a Penalized Likelihood on the hazard function

                  coef exp(coef) SE coef (H) SE coef (HIH)        z          p
dukesC         0.297710  1.346771    0.160616      0.160616  1.85355 6.3804e-02
dukesD         1.059439  2.884751    0.194424      0.194424  5.44911 5.0623e-08
charlson1-2    0.455345  1.576717    0.259515      0.259515  1.75460 7.9328e-02
charlson3      0.415539  1.515187    0.136950      0.136950  3.03423 2.4115e-03
sexFemale     -0.533690  0.586437    0.138639      0.138639 -3.84948 1.1837e-04
chemoTreated -0.198651  0.819836    0.142920      0.142920 -1.38994 1.6455e-01

           chisq df global p
dukes    30.4469  2 2.45e-07
charlson 11.2275  2 3.65e-03

    Frailty parameter, Theta: 0.671573 (SE (H): 0.141443 ) p = 1.0272e-06

      penalized marginal log-likelihood = -3250.34
      Convergence criteria:
      parameters = 4.56e-05 likelihood = 0.000249 gradient = 8.17e-08

      LCV = the approximate likelihood cross-validation criterion
            in the semi parametrical case     = 3.79458

      n= 861
      n events= 458  n groups= 403
      number of iterations:  7

      Exact number of knots used:  8
```

```
Best smoothing parameter estimated by
an approximated Cross validation:  2434216478, DoF:  9.25
```

– Comparison of gamma vs. log-normal joint frailty model using the approximate likelihood cross-validation criterion (LCV)

there's no clusters here
when I have two kappas, then fit them in the something something

```
modJoint.gap.logN <- frailtyPenal(Surv(time, event) ~ cluster(id) + dukes +
                  charlson + sex + chemo + terminal(death),
                  formula.terminalEvent = ~ dukes + charlson + sex + chemo,
                  data = readmission, n.knots = 8,
                  kappa = c(2.11e+08, 9.53e+11),
                  RandDist = 'LogN')

modJoint.gap.logN$LCV

modJoint.gap$LCV
```

```
> modJoint.gap.logN$LCV
[1] 0.03771733
> modJoint.gap$LCV
[1] 0.03701317
```

it seems that gamma is better because it gives a lower LCV value

– Time-varying coefficients: Using function timedep in a formula of frailtyPenal, the time-dependent coefficients can be estimated using B-splines of order q (option betaorder) with m interior knots (option betaknots).

time-dependent covariates for the terminal event

– In the example of the readmission dataset we are interested in verifying whether the variable sex has a time-varying effect on both recurrent and terminal events. Thus, we fit a model equivalent to modJoint.gap but with time dependent effects assuming quadratic B-splines and 3 interior knots

```
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)
```

```
pdf("Timedep.pdf")
modJoint.gap.timedep
dev.off()
```

overtime, increase the variability.
not big variation in the beta coefficients
more beta coefficients over time
it depends on you that
if use a TDV
or use a statistical test

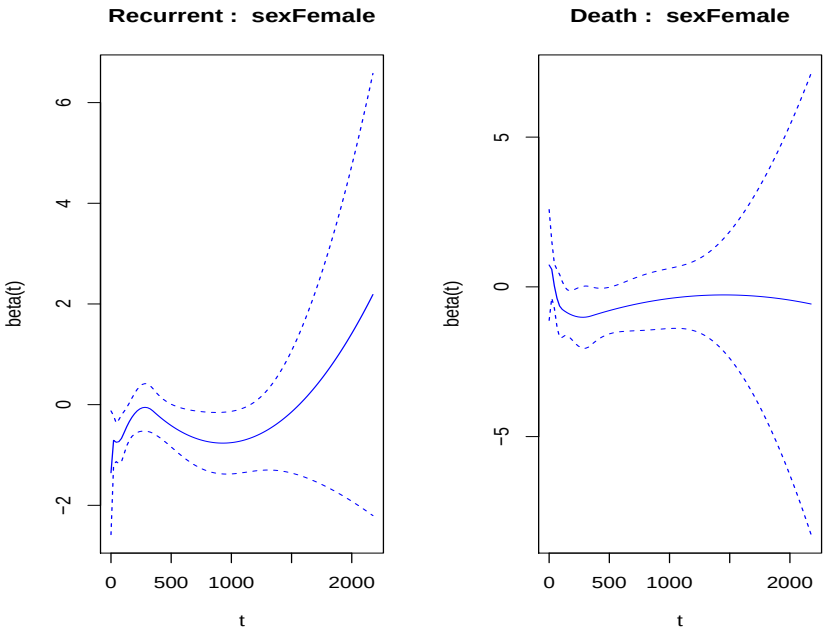


Figure 4: Plots of sex as time-dependent covariate effect

– For rehospitalizations, we found firstly a protective effect for females $\beta(t) < 0$ and later an increased risk $\beta(t) > 0$. For death, at the beginning, the effect of sex was weakening the risk but shortly became non-influential $\beta(t)$ around 0.

– The PH assumption for the variable sex can be checked using the LR test. We compare two models: modJoint.gap (related to the null hypothesis that the effect is constant in time) and modJoint.gap.timedep (related to the alternative hypothesis that the effect varies with time).

4 Joint model for recurrent events and a terminal event with two frailty terms

– In the simple JM for recurrent and terminal events, the frailty term u_i reflects the inter- and intra-subject correlation for the recurrent event as well as the association between the recurrent and the terminal events. In order to distinguish the origin of dependence, two independent frailty terms (u_i, v_i) can be considered (Mazroui et al. 2012).

– Model formulation

$$\begin{cases} r_{ij}(t|u_i, v_i) = u_i v_i r_0(t) \exp(X_{R_{ij}}^T \beta_R), \\ \\ \lambda_i(t|u_i) = u_i \lambda_0(t) \exp(X_{T_i}^T \beta_T), \end{cases}$$

,

where frailty $u_i \sim \Gamma(1/\theta, 1/\theta)$ is specific to the association between the processes and $v_i \sim \Gamma(1/\eta, 1/\eta)$ is specific to the recurrent event rate.

– The variance of the frailty terms represents the heterogeneity in the data, associated with unobserved covariates.

– Moreover, a high value of the variance η indicates a strong dependence between the recurrent events and a high value of the variance θ indicates that the recurrent and the terminal events are strongly dependent.

– To estimate the general frailty model, the argument jointGeneral must be equal to TRUE in function frailtyPenal.

– Example:

```
modJoint.general <- frailtyPenal(Surv(time, event) ~ cluster(id) +
                dukes + charlson + sex + chemo + terminal(death),
                formula.terminalEvent = ~ dukes + charlson + sex + chemo,
                data = readmission, jointGeneral = TRUE, n.knots = 8,
                kappa = c(2.11e+08, 9.53e+11))

modJoint.general

  General 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)
dukesC       "0.339621710391044" "1.40441621400241" "0.160851644444906" "0.16084780389202"
dukesD       "1.27715168089166"  "3.58640992369933" "0.197274065264374" "0.197271454206402"
charlson1-2  "0.4059703215963"   "1.50075801169495" "0.251439348862587" "0.251439139818686"
charlson3    "0.407039611463208" "1.50236361530438" "0.135365777401148" "0.135365624773996"
sexFemale    "-0.526105261016811" "0.590901902472797" "0.138288207320059" "0.138287448902685"
chemoTreated "-0.15254487006031"  "0.858520371231308" "0.142728992293761" "0.142725278387212"
             z                   p
dukesC       "2.11139719188491"  "0.034738"
dukesD       "6.47399686917846"  "9.5444e-11"
charlson1-2  "1.61458547929252"  "0.1064"
charlson3    "3.0069609858404"   "0.0026387"
sexFemale    "-3.80441160683481" "0.00014214"
chemoTreated "-1.06877283731077" "0.28517"

           chisq df global p
dukes    46.3706  2 8.53e-11
charlson 11.6487  2 2.95e-03

Terminal event:
----------------
             coef                exp(coef)           SE coef (H)          SE coef (HIH)
dukesC       "1.34235990376546"  "3.82806672329046"  "0.348756953342793" "0.348700207902842"
dukesD       "3.28870225512253"  "26.8080510609116"  "0.382644984113067" "0.382581613176839"
charlson1-2  "0.506463445601929" "1.65941220385971"  "0.639042994901726" "0.639037144439215"
charlson3    "1.27216198963833"  "3.56855941676669"  "0.260132543367247" "0.26012586451991"
sexFemale    "-0.242655274324886" "0.784541918903732" "0.23361009701543"  "0.233603788043318"
chemoTreated "1.12506300749455"  "3.08041093177862"  "0.248323731753921" "0.248280049560608"
             z                   p
dukesC       "3.84898391530006"  "0.00011861"
dukesD       "8.59465664431852"  "< 1e-16"
charlson1-2  "0.792534226401801" "0.42805"
charlson3    "4.89043767139249"  "1.0061e-06"
sexFemale    "-1.03871911969994" "0.29894"
chemoTreated "4.53063023637805"  "5.8808e-06"

         chisq              df  global p
dukes    "83.3358239012546" "2" "< 1e-16"
charlson "23.920442696996"  "2" "6.39e-06"
```

```
Frailty parameters:
  theta (variance of u, association between recurrences and terminal event): 0.678514
  (SE (H): 0.126084 )    p = 3.6943e-08
  eta (variance of v, intra-subject correlation): 0.00500415 (SE (H): 0.000352111 ) p = < 1e

  penalized marginal log-likelihood = -3892.94
  Convergence criteria:
  parameters = 7.57e-07 likelihood = 5.03e-06 gradient = 1.51e-11

  LCV = the approximate likelihood cross-validation criterion
        in the semi parametric case     = 4.5609

  n observations= 861  n subjects= 403
  n recurrent events= 458
  n terminal events= 109
  n censored events= 403
  number of iterations:  12

  Exact number of knots used:  8
  Value of the smoothing parameters:  2.11e+08 9.53e+11
```

signicant recurrrent event between individuals

- The estimated variance θ of the frailty u_i associating recurrent events and death indicates strong relationship between the processes ($\hat{\theta} = 0.68, p < 0.001$).

- Moreover, the estimate of η implies small but significant dependence between the recurrent event gap times explained by the frailty v_i, ($\hat{\eta} = 0.005, p < 0.001$).

- This information complements the inference from the standard joint frailty model because it separates the correlation linked to the recurrent events with correlation between the recurrent events and the terminal event.

5 Joint model for two clustered survival events

- An increasing number of studies favor the presence of clustered data, especially multi-center or multi-trial studies.

– The clustering creates some heterogeneity between subjects, which needs to be accounted for.

– We let the index j ($j = 1, \cdots, n_i$) represents a subject from cluster i ($i = 1, \cdots, N$) and the cluster-specific frailty term u_i is shared by the subjects of a given group.

– Model formulation

$$\begin{cases} r_{ij}(t|u_i) = u_i r_0(t) \exp(X_{R_{ij}}^T \beta_R), & \text{(time to event 1)} \\ \\ \lambda_{ij}(t|u_i) = u_i^\alpha \lambda_0(t) \exp(X_{T_{ij}}^T \beta_T), & \text{(time to event 2)} \end{cases}$$

where the frailty terms u_i to be iid Gamma distributed.

– The events can be chosen arbitrarily but it is assumed that the event 2 impedes the process of the event 1. Usually the event 2 is death of a patient and the other is an event of interest (e.g., surrogate end-point for OS) such as time to tumor progression or progression-free survival.

– The interest of using the joint model for the clustered data stems from the fact that it considers the dependency between the survival processes and assumes that event 2 is a competitive event for event 1. The frailty term u_i is common for a given group and represents the clustered association between the processes (at the cluster level) as well as the intra-cluster correlation.

– The joint models for clustered survival data can be estimated using again the frailtyPenal function.

– A dataset should include information on two survival outcomes for

individuals from several groups.

- This model is presented using the readmission dataset with artificially created clusters on individuals.
- The first survival event will be the first observed rehospitalization and the second event, death.

- The framework of semi-competing risks is used here, thus individuals follow-up stops at time of the rehospitalization, death or in case when none of these events are observed, the censoring time.

- We consider 6 clusters defined by a new variable group:

```
readmission <- transform(readmission, group = id %% 6 + 1 )
```

```
> readmission
     id enum t.start t.stop time event      chemo    sex dukes charlson death group
1     1    1       0     24   24     1    Treated Female     D        3     0     2
2     1    2      24    457  433     1    Treated Female     D        0     0     2
3     1    3     457   1037  580     0    Treated Female     D        0     0     2
4     2    1       0    489  489     1 NonTreated   Male     C        0     0     3
5     2    2     489   1182  693     0 NonTreated   Male     C        0     0     3
```

```
readm_cluster <- subset(readmission,
              (t.start == 0 & event == 1) | event == 0)
```

```
> head(readm_cluster, 20)
   id enum t.start t.stop time event      chemo    sex dukes charlson death group
1   1    1       0     24   24     1    Treated Female     D        3     0     2
3   1    3     457   1037  580     0    Treated Female     D        0     0     2
4   2    1       0    489  489     1 NonTreated   Male     C        0     0     3
5   2    2     489   1182  693     0 NonTreated   Male     C        0     0     3
6   3    1       0     15   15     1 NonTreated   Male     C        3     0     4
7   3    2      15    783  768     0 NonTreated   Male     C        3     1     4
8   4    1       0    163  163     1    Treated Female   A-B        0     0     5
12  4    5     686   2048 1362     0    Treated Female   A-B        0     0     5
```

- The new dataset *readm_cluster* includes clusters with 97 to 107 individuals per group. For definition of a clustered joint model two inner functions are required in frailtyPenal, num.id for the individual level and cluster for the groups:

```
joi.clus <- frailtyPenal(Surv(t.start, t.stop, event) ~ cluster(group) +
            num.id(id) + dukes + sex + chemo + terminal(death),
            formula.terminalEvent = ~ dukes + sex + chemo, data = readm_cluster,
            n.knots = 8, kappa = c(1.e+10, 1.e+10), recurrentAG = TRUE,
            Alpha = "None")

  For clustered data
  Joint gamma frailty model for a survival and a terminal event processes
  using a Penalized Likelihood on the hazard function

Survival event:
-------------
                  coef exp(coef) SE coef (H) SE coef (HIH)         z       p
dukesC        0.144170  1.155081    0.182691      0.183399  0.789147 0.43003
dukesD        0.330643  1.391863    0.186025      0.186142  1.777413 0.07550
sexFemale    -0.164064  0.848688    0.229262      0.229274 -0.715617 0.47423
chemoTreated  0.172778  1.188602    0.186471      0.186569  0.926566 0.35415

       chisq df global p
dukes 3.6713  2     0.16

Terminal event:
----------------
                     coef exp(coef) SE coef (H) SE coef (HIH)          z       p
dukesC       -0.06263165  0.939289    0.224199      0.223825 -0.2793568 0.77997
dukesD        0.03960691  1.040402    0.263444      0.263467  0.1503428 0.88049
sexFemale     0.01330086  1.013390    0.200137      0.199732  0.0664589 0.94701
chemoTreated -0.00669263  0.993330    0.200873      0.200778 -0.0333177 0.97342

         chisq df global p
dukes 0.157446  2    0.924

 Frailty parameters:
   theta (variance of Frailties, w): 0.0073641 (SE (H): 0.00412441 ) p = 0.037091
   alpha is fixed (=1)

   penalized marginal log-likelihood = -1985.25
   Convergence criteria:
   parameters = 2.91e-05 likelihood = 0.000537 gradient = 2.26e-07

   LCV = the approximate likelihood cross-validation criterion
         in the semi parametric case     = 4.99787

   n observations= 403  n subjects= 403  n groups= 403
   n events= 134
   n terminal events= 109
```

```
n censored events= 269
number of iterations:  24

Exact number of knots used:  8
Value of the smoothing parameters:  1e+10 1e+10
```

– the estimates of prognostic factors for both types of events are obtained.

– The estimate of the variance of the frailty term indicates whether, at the cluster level, the processes are associated with each other and measures the heterogeneity between individuals (intracluster correlation).

– In the given example the estimate of the variance θ is significantly different from 0 (p value = 0.037), thus there is a positive association between the risk of hospitalizations and death via the non-observed frailty.

– In case of the joint frailty models for clustered data it should be noted that sufficient amount of information must be provided, i.e., number of observations per cluster. Otherwise, given the complexity of the model, convergence might not be attained.

– The parameter α is assumed to be equal to 1 as these models are defined in the framework of semi-competing risks and not of recurrent events.