

# INLAjoint

Denis Rustand

2023-07-11

## Contents

Datasets for illustrations . . . . .	2
Model 1: Mixed effects regression for a single longitudinal marker . . . . .	2
Model 2: Multiple longitudinal markers with different distributions . . . . .	5
Model 3: Proportional hazards survival model . . . . .	8
Model 4: Longitudinal - survival joint model . . . . .	11
Model 5: Comparison with MCMC . . . . .	15
Model 6: Joint with one longitudinal and competing risks of event . . . . .	18
Model 7: Joint with three longitudinal markers and competing risks of event . . . . .	19
Model 8: Multi-state model . . . . .	21
Model 9: Joint longitudinal and multi-state model . . . . .	22
Model 10: Mixture cure model . . . . .	23
Model 11: Shared frailty model for recurrent events . . . . .	24
Model 12: Joint model for recurrent events and a terminal event . . . . .	25
Model 13: Two-part model for a longitudinal semicontinuous outcome . . . . .	26
Model 14: Two-part joint model for a longitudinal semicontinuous outcome and a terminal event . . . . .	27
Model 15: Two-part joint model for a longitudinal semicontinuous outcome, recurrent events and a terminal event . . . . .	28
Model 16: Joint model for multivariate longitudinal data and competing risks of event (application section of <a href="https://arxiv.org/abs/2203.06256">https://arxiv.org/abs/2203.06256</a> ) . . . . .	30

In this vignette we show how to fit various models with the *joint()* function of the INLAjoint package.

Contact: INLAjoint@gmail.com

## Datasets for illustrations

### Mayo Clinic Primary Biliary Cirrhosis Data (pbc2)

We use the data of the famous randomized clinical trial of Primary Biliary Cholangitis (PBC) patients where 312 PBC patients were followed at the Mayo Clinic between 1974 and 1988 and received either a placebo or D-penicillamine. These data are publicly available in several software including the R package “JM”. During the follow-up, 140 patients died and 29 patients received a liver transplantation which we consider here as a competing event of death. In addition, repeated measures of various longitudinal markers potentially associated with the disease progression were collected. It is used to illustrate multiple models of the vignette (1, 2, 3, 4, 5, 6, 7, 13, 14 and 16).

```
data(pbc2) # dataset
# extract some variable of interest without missing values
Longi <- na.omit(pbc2[, c("id", "years", "status", "drug", "age",
                        "sex", "year", "serBilir", "SGOT", "albumin", "edema",
                        "platelets", "alkaline", "spiders", "ascites")])
Surv <- Longi[c(which(diff(as.numeric(Longi[,which(colnames(Longi)=="id")]))==1),
               length(Longi[,which(colnames(Longi)=="id")])), -c(7:10, 12:16)]
Surv$death <- ifelse(Surv$status=="dead", 1, 0) # competing event 1
Surv$trans <- ifelse(Surv$status=="transplanted", 1, 0) # competing event 2
```

### Bone marrow transplant data (bmt)

Bone marrow transplant study which is available in the package “smcure”, it is used to illustrate the mixture cure model (model 10 of the vignette).

### Kidney catheter data (kidney)

Recurrence times to infection for kidney patients using portable dialysis equipment, available in package “survival”, it is used to illustrate the frailty model (model 11 of the vignette).

### Rehospitalization colorectal cancer (readmission)

Rehospitalization times after surgery in patients diagnosed with colorectal cancer, it is available in the package “frailtypack” and it is used to illustrate the joint frailty model for recurrent events and a terminal event (model 12 of the vignette).

### FFCD 2000-05 phase III metastatic colorectal cancer clinical trial (colorectal and colorectal-Longi)

There are two complementary datasets available in the package “frailtypack”. `colorectal` includes times of new lesions appearance and time of death of 150 randomly chosen patients from the FFCD 2000-05 multicenter clinical trial. `colorectalLongi` contains the corresponding longitudinal measurements of tumors size. It is used to illustrate the two-part joint model for a longitudinal semicontinuous biomarker, recurrent events and a terminal event (model 15 of the vignette).

## Model 1: Mixed effects regression for a single longitudinal marker

This first model shows how to call the *joint()* function for a simple linear mixed effects model for a longitudinal marker, it gives the basic structure of the function. The required arguments are:

- **formLong**: formula for the model with the lme4 structure (including random effects in the formula as: (NAME | ID)).
- **dataLong**: Dataset that must contains the variables given in the formula.
- **id**: Name of the variable for grouping (e.g., individuals).
- **timeVar**: Name of the time variable.
- **family**: Distribution of the outcome (e.g., gaussian, poisson, binomial).

The model structure is given by the following equation:

$$\log(\text{serBilir}_{ij}) = \beta_0 + b_{i0} + \beta_1 \text{year}_{ij} + \beta_2 \text{drug}_i + \varepsilon_{ij} \quad (\text{L1})$$

where  $\beta$  are the fixed effects,  $b_{i0}$  is an individual random intercept and  $\varepsilon_{ij}$  is the residual error term.

```
M1 <- joint(formLong = serBilir ~ year + drug + (1|id),
            dataLong = Longi, id = "id", timeVar = "year",
            family = "lognormal")
```

The summary statistics are available from the **summary** function:

```
summary(M1)

## Longitudinal outcome (lognormal)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept      0.6304 0.0885      0.4568  0.6304    0.8041
## year           0.0843 0.0042      0.0760  0.0843    0.0926
## drugDpenicil   -0.1223 0.1240     -0.3655 -0.1222    0.1209
## Res. err. (variance) 0.2154 0.0077      0.2008  0.2152    0.2309
##
## Random effects variance-covariance (L1)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept 1.1459 0.0969      0.9712  1.1404    1.352
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                               -2802.289                               -2802.287
##
## Deviance Information Criterion: 4805.123
## Widely applicable Bayesian information criterion: 4806.438
## Computation time: 0.91 seconds
```

If one wishes to get the standard deviations instead of variance parameters, it is possible to switch with the **sdcor** argument of the **summary** function:

```
summary(M1, sdcor=TRUE)

## Longitudinal outcome (lognormal)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept      0.6304 0.0885      0.4568  0.6304    0.8041
## year           0.0843 0.0042      0.0760  0.0843    0.0926
## drugDpenicil   -0.1223 0.1240     -0.3655 -0.1222    0.1209
## Res. err. (sd)  0.4640 0.0082      0.4481  0.4639    0.4806
##
## Random effects standard deviation / correlation (L1)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept 1.0695 0.0451      0.9855  1.0679    1.1628
##
```

```
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                                     -2802.289                      -2802.287
##
## Deviance Information Criterion:  4805.123
## Widely applicable Bayesian information criterion:  4806.438
## Computation time: 0.91 seconds
```

The log marginal-likelihood, the Deviance Information Criterion (DIC) and the Widely Applicable Bayesian Information Criterion (WAIC) are provided in the summary statistics.

The `control` argument in the `joint` function has the following components:

- `int.strategy` allows to choose the strategy for the numerical integration used to approximate the marginal posterior distributions of the latent field. Available options are “ccd” (default), “grid” or “eb” (empirical Bayes). The empirical Bayes uses only the mode of the approximations for the integration, which speed up and simplifies computations.
- `priorFixed` allows to set the mean and standard deviation of the Gaussian prior for the fixed effects.
- `priorAssoc` allows to set the mean and standard deviation of the Gaussian prior for the association parameters between the longitudinal and survival submodels.
- `cpo` set to TRUE to compute the Conditional Predictive Ordinate.

An useful function to learn about the priors used in a fitted model is `priors.used`, applied to an object fitted with the `joint` function. The default priors are Gaussian with mean zero and scale 1.

```
priors.used(M1)
```

```
## section=[family]
## tag=[INLA.Data1] component=[lognormal]
##   theta1:
##     parameter=[log precision]
##     prior=[loggamma]
##     param=[1e+00, 5e-05]
## section=[fixed]
## tag=[Intercept_L1] component=[Intercept_L1]
##   beta:
##     parameter=[Intercept_L1]
##     prior=[normal]
##     param=[0.00, 0.01]
## tag=[year_L1] component=[year_L1]
##   beta:
##     parameter=[year_L1]
##     prior=[normal]
##     param=[0.00, 0.01]
## tag=[drugDpenicil_L1] component=[drugDpenicil_L1]
##   beta:
##     parameter=[drugDpenicil_L1]
##     prior=[normal]
##     param=[0.00, 0.01]
## section=[random]
## tag=[IDIntercept_L1] component=[IDIntercept_L1]
##   theta1:
##     parameter=[log precision]
##     prior=[loggamma]
##     param=[1e+00, 5e-05]
## NULL
```

The full list of the arguments is available in the help documentation of the `joint` function which can be accessed by running `?joint`.

## Model 2: Multiple longitudinal markers with different distributions

The following code fits a joint model with 3 longitudinal markers including fixed effects for covariates such as sex, drug and interactions with time. We assume random intercept and random slope for each longitudinal trajectory. Note that the `formLong` argument is now a list of formulas, one for each longitudinal marker and the length of family must match the number of markers.

The model structure is given by the following equation:

$$\begin{cases} \log(\text{serBilir}_{ij}) = \beta_{10} + b_{i10} + (\beta_{11} + b_{i11})\text{year}_{ij} + \beta_{12}\text{drug}_i + \beta_{13}\text{sex}_i + \beta_{14}\text{year}_{ij}\text{drug}_i + \varepsilon_{ij1} & (\text{L1}) \\ \log(E[\text{platelets}_{ij}]) = \beta_{20} + b_{i20} + (\beta_{21} + b_{i21})\text{year}_{ij} + \beta_{22}\text{sex}_i + \beta_{23}\text{drug}_i + \beta_{24}\text{year}_{ij}\text{sex}_i & (\text{L2}) \\ \text{logit}(E[\text{spiders}_{ij}]) = \beta_{30} + b_{i30} + (\beta_{31} + b_{i31})\text{year}_{ij} + \beta_{32}\text{drug}_i & (\text{L3}) \end{cases}$$

```
M2 <- joint(formLong = list(serBilir ~ year * drug + sex + (1+year|id),
                           platelets ~ year * sex + drug + (1+year|id),
                           spiders ~ year + drug + (1+year|id)),
            dataLong = Longi, id = "id", timeVar="year", corLong=TRUE,
            family = c("lognormal", "poisson", "binomial"), control=list(int.strategy="eb"))
summary(M2)
```

```
## Longitudinal outcome (L1, lognormal)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1    0.9197 0.1661    0.5940    0.9197    1.2453
## year_L1         0.1581 0.0187    0.1215    0.1581    0.1947
## drugDpenicil_L1 -0.1702 0.1146   -0.3948   -0.1702    0.0543
## sexfemale_L1    -0.3702 0.1604   -0.6845   -0.3702   -0.0559
## year:drugDpenicil_L1 0.0051 0.0255   -0.0449    0.0051    0.0551
## Res. err. (variance) 0.1090 0.0042    0.1010    0.1089    0.1176
##
## Longitudinal outcome (L2, poisson)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L2     5.4408 0.0677    5.3081    5.4408    5.5735
## year_L2          -0.1240 0.0324   -0.1874   -0.1240   -0.0606
## sexfemale_L2     0.0919 0.0667   -0.0388    0.0919    0.2226
## drugDpenicil_L2 -0.0728 0.0436   -0.1583   -0.0728    0.0127
## year:sexfemale_L2 0.0439 0.0343   -0.0233    0.0439    0.1111
##
## Longitudinal outcome (L3, binomial)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L3    -1.6784 0.3052   -2.2766   -1.6784   -1.0801
## year_L3          0.1867 0.0676    0.0542    0.1867    0.3192
## drugDpenicil_L3 -0.2591 0.4190   -1.0803   -0.2591    0.5621
##
## Random effects variance-covariance
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1     0.9938 0.0872    0.8294    0.9899    1.1779
## year_L1          0.0341 0.0042    0.0264    0.0340    0.0431
## Intercept_L2     0.1502 0.0124    0.1286    0.1497    0.1770
## year_L2          0.0326 0.0046    0.0249    0.0322    0.0424
## Intercept_L3     9.3907 1.8483    6.1016    9.2498   13.4348
## year_L3          0.4801 0.1164    0.2862    0.4654    0.7487
```

```
## Intercept_L1:year_L1      0.0425 0.0128      0.0181 0.0422      0.0683
## Intercept_L1:Intercept_L2 -0.0830 0.0244      -0.1323 -0.0823     -0.0357
## Intercept_L1:year_L2      -0.0229 0.0105      -0.0443 -0.0224     -0.0037
## Intercept_L1:Intercept_L3 1.5313 0.2563      1.0741 1.5083      2.0865
## Intercept_L1:year_L3      0.1078 0.0691      -0.0164 0.1045      0.2576
## year_L1:Intercept_L2      -0.0046 0.0054      -0.0153 -0.0046      0.0065
## year_L1:year_L2          -0.0048 0.0021      -0.0092 -0.0047     -0.0005
## year_L1:Intercept_L3      0.0854 0.0492      -0.0084 0.0846      0.1832
## year_L1:year_L3          0.0577 0.0141      0.0326 0.0572      0.0882
## Intercept_L2:year_L2      -0.0031 0.0040      -0.0115 -0.0030      0.0047
## Intercept_L2:Intercept_L3 -0.3235 0.0861      -0.5036 -0.3173     -0.1738
## Intercept_L2:year_L3      0.0209 0.0269      -0.0280 0.0196      0.0762
## year_L2:Intercept_L3      0.0193 0.0482      -0.0753 0.0190      0.1139
## year_L2:year_L3          -0.0326 0.0183      -0.0745 -0.0306     -0.0014
## Intercept_L3:year_L3      -0.2470 0.2326      -0.7675 -0.2131      0.1333
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                                     -17078.17                                -17057.67
##
## Deviance Information Criterion: 25230.81
## Widely applicable Bayesian information criterion: 44883.59
## Computation time: 16.02 seconds
```

The additional boolean argument `corLong` is set to `TRUE` in order to have correlation between the random effects across the longitudinal markers. Therefore by switching this argument to `TRUE`, instead of having 3 sets of two correlated random effects, we have 1 set of 6 correlated random effects.

We can also get the standard deviation and correlation of random parameters instead of variance and covariance by adding `sdcor=TRUE` to the summary function call:

```
summary(M2, sdcor=TRUE)

## Longitudinal outcome (L1, lognormal)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1      0.9197 0.1661      0.5940 0.9197      1.2453
## year_L1           0.1581 0.0187      0.1215 0.1581      0.1947
## drugDpenicil_L1   -0.1702 0.1146     -0.3948 -0.1702      0.0543
## sexfemale_L1      -0.3702 0.1604     -0.6845 -0.3702     -0.0559
## year:drugDpenicil_L1 0.0051 0.0255     -0.0449 0.0051      0.0551
## Res. err. (sd)     0.3301 0.0064      0.3178 0.3300      0.3429
##
## Longitudinal outcome (L2, poisson)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L2       5.4408 0.0677      5.3081 5.4408      5.5735
## year_L2            -0.1240 0.0324     -0.1874 -0.1240     -0.0606
## sexfemale_L2       0.0919 0.0667     -0.0388 0.0919      0.2226
## drugDpenicil_L2   -0.0728 0.0436     -0.1583 -0.0728      0.0127
## year:sexfemale_L2  0.0439 0.0343     -0.0233 0.0439      0.1111
##
## Longitudinal outcome (L3, binomial)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L3      -1.6784 0.3052     -2.2766 -1.6784     -1.0801
## year_L3            0.1867 0.0676      0.0542 0.1867      0.3192
## drugDpenicil_L3   -0.2591 0.4190     -1.0803 -0.2591      0.5621
##
## Random effects standard deviation / correlation
```

```
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1      0.9949 0.0421      0.9158  0.9941  1.0780
## year_L1           0.1846 0.0117      0.1626  0.1843  0.2092
## Intercept_L2      0.3872 0.0161      0.3571  0.3864  0.4203
## year_L2           0.1797 0.0124      0.1570  0.1792  0.2048
## Intercept_L3      3.0540 0.3097      2.4852  3.0508  3.6573
## year_L3           0.6868 0.0846      0.5375  0.6809  0.8692
## Intercept_L1:year_L1 0.2343 0.0664      0.1037  0.2354  0.3589
## Intercept_L1:Intercept_L2 -0.2140 0.0596     -0.3304 -0.2138 -0.0979
## Intercept_L1:year_L2 -0.1272 0.0554     -0.2314 -0.1283 -0.0137
## Intercept_L1:Intercept_L3 0.5005 0.0536      0.3889  0.5030  0.5991
## Intercept_L1:year_L3  0.1621 0.0895     -0.0196  0.1618  0.3371
## year_L1:Intercept_L2 -0.0667 0.0740     -0.2185 -0.0653  0.0740
## year_L1:year_L2     -0.1412 0.0601     -0.2554 -0.1426 -0.0235
## year_L1:Intercept_L3  0.1537 0.0846     -0.0113  0.1523  0.3127
## year_L1:year_L3      0.4531 0.0872      0.2544  0.4626  0.6021
## Intercept_L2:year_L2 -0.0439 0.0553     -0.1538 -0.0439  0.0642
## Intercept_L2:Intercept_L3 -0.2766 0.0642     -0.4002 -0.2779 -0.1491
## Intercept_L2:year_L3  0.0755 0.0964     -0.1155  0.0767  0.2616
## year_L2:Intercept_L3  0.0344 0.0836     -0.1283  0.0335  0.1936
## year_L2:year_L3     -0.2550 0.1210     -0.4816 -0.2622 -0.0041
## Intercept_L3:year_L3 -0.1115 0.0982     -0.3133 -0.1079  0.0671
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                               -17078.17                               -17057.67
##
## Deviance Information Criterion: 25230.81
## Widely applicable Bayesian information criterion: 44883.59
## Computation time: 16.02 seconds
```

The link functions between the linear predictors and the longitudinal outcomes are set to default, it is however possible to switch to alternative ones using the `link` argument, e.g., to switch from `logit` to `probit` for the binary marker:

```
M2 <- joint(formLong = list(serBilir ~ year * drug + sex + (1+year|id),
                           platelets ~ year * sex + drug + (1+year|id),
                           spiders ~ year + drug + (1+year|id)),
            dataLong = Longi, id = "id", timeVar="year", corLong=TRUE,
            family = c("lognormal", "poisson", "binomial"),
            link = c("default", "default", "probit"), control=list(int.strategy="eb"))
summary(M2)
```

```
## Longitudinal outcome (L1, lognormal)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1      0.9115 0.1660      0.5861  0.9115  1.2369
## year_L1           0.1536 0.0186      0.1172  0.1536  0.1900
## drugDpenicil_L1   -0.1205 0.1142     -0.3443 -0.1205  0.1033
## sexfemale_L1      -0.3904 0.1601     -0.7041 -0.3904 -0.0766
## year:drugDpenicil_L1 0.0154 0.0254     -0.0344  0.0154  0.0652
## Res. err. (variance) 0.1086 0.0026      0.1036  0.1085  0.1137
##
## Longitudinal outcome (L2, poisson)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L2      5.4421 0.0677      5.3095  5.4421  5.5748
## year_L2           -0.1308 0.0329     -0.1952 -0.1308 -0.0664
```

```

## sexfemale_L2      0.0931 0.0666    -0.0374  0.0931    0.2236
## drugDpenicil_L2   -0.0777 0.0437    -0.1635  -0.0777    0.0080
## year:sexfemale_L2 0.0517 0.0349    -0.0167  0.0517    0.1200
##
## Longitudinal outcome (L3, binomial)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L3   -0.9753 0.1663    -1.3013  -0.9753   -0.6493
## year_L3        -0.0108 0.0491    -0.1070  -0.0108    0.0854
## drugDpenicil_L3 -0.0205 0.2331    -0.4774  -0.0205    0.4363
##
## Random effects variance-covariance
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1     1.0149 0.0895     0.8468  1.0144    1.1946
## year_L1          0.0343 0.0048     0.0259  0.0339    0.0448
## Intercept_L2     0.1506 0.0133     0.1269  0.1497    0.1792
## year_L2          0.0328 0.0047     0.0247  0.0326    0.0427
## Intercept_L3     2.9939 0.5549     2.1067  2.9152    4.2964
## year_L3          0.2790 0.0747     0.1594  0.2672    0.4428
## Intercept_L1:year_L1 0.0458 0.0150     0.0145  0.0460    0.0743
## Intercept_L1:Intercept_L2 -0.0852 0.0264    -0.1376  -0.0847   -0.0345
## Intercept_L1:year_L2 -0.0167 0.0107    -0.0376  -0.0165    0.0035
## Intercept_L1:Intercept_L3 0.8690 0.1563     0.6015  0.8533    1.2189
## Intercept_L1:year_L3  0.1566 0.0703     0.0346  0.1519    0.3061
## year_L1:Intercept_L2 -0.0067 0.0059    -0.0184  -0.0066    0.0052
## year_L1:year_L2     -0.0037 0.0026    -0.0093  -0.0035    0.0013
## year_L1:Intercept_L3  0.0501 0.0320    -0.0104  0.0493    0.1157
## year_L1:year_L3     0.0475 0.0116     0.0289  0.0463    0.0753
## Intercept_L2:year_L2 -0.0034 0.0044    -0.0125  -0.0033    0.0052
## Intercept_L2:Intercept_L3 -0.1989 0.0584    -0.3283  -0.1938   -0.0906
## Intercept_L2:year_L3 -0.0065 0.0233    -0.0571  -0.0057    0.0369
## year_L2:Intercept_L3  0.0153 0.0341    -0.0508  0.0154    0.0825
## year_L2:year_L3     -0.0214 0.0121    -0.0484  -0.0207    0.0002
## Intercept_L3:year_L3  0.0880 0.1194    -0.1185  0.0793    0.3256
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##               -17067.46                      -17046.95
##
## Deviance Information Criterion: 26068.73
## Widely applicable Bayesian information criterion: 46386.51
## Computation time: 15.44 seconds

```

### Model 3: Proportional hazards survival model

Some additional arguments are introduced to fit a survival model:

- **formSurv**: formula for the time-to-event outcome, with the response given as an `inla.surv()` object.
- **dataSurv**: Dataset that must contains the variables given in the formula. When fitting a joint model with a longitudinal component, if `dataSurv` is not provided, the longitudinal dataset is used to get the covariates values included in the time-to-event formula.
- **basRisk**: the baseline risk of event. It can be defined as parametric with either “`exponentialsurv`” for exponential baseline or “`weibullsurv`” for Weibull baseline (note that there are two formulations of the Weibull distribution, see `inla.doc("weibull")` for more details, default is `variant = 0`). Alternatively, there are two options to avoid parametric assumptions on the shape of the baseline risk: “`rw1`” for random walks of order one prior that corresponds to a smooth spline function based on first order



differences. The second option “rw2” assigns a random walk order two prior that corresponds to a smooth spline function based on second order differences. This second option provides a smoother spline compared to order one since the smoothing is then done on the second order. The number of bins that define the intervals for the baseline risk can be specified with `NbasRisk` (default 15 bins).

The model is defined as:

$$\lambda_i(t) = \lambda_0(t) \exp(\gamma_1 \text{drug}_i + \gamma_2 \text{sex}_i + \gamma_3 \text{drug}_i \text{sex}_i)$$

```
DTH <- inla.surv(time = Surv$years, event = Surv$death) # survival outcome
M3 <- joint(formSurv = DTH ~ drug * sex, dataSurv = Surv)
summary(M3)

##
## Survival outcome
##
##               mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance) 0.0102 0.0205      0.0000      0.0033      0.0648
## drugDpenicil            0.1692 0.3907     -0.5970      0.1692      0.9355
## sexfemale              -0.5342 0.3219     -1.1655     -0.5342      0.0971
## drugDpenicil:sexfemale -0.2238 0.4330     -1.0729     -0.2238      0.6254
##
## log marginal-likelihood (integration)      log marginal-likelihood (Gaussian)
##                                     -664.5177                                     -664.5669
##
## Deviance Information Criterion: -1288.312
## Widely applicable Bayesian information criterion: 1297.794
## Computation time: 0.89 seconds
```

Note that the parameter associated to the baseline risk of event “Baseline risk (variance)” corresponds to the variance parameter associated to the random walk to capture the evolution over time. Summary statistics of the random walks baseline risks are available in the “BaselineValues” attribute of an `INLAjoint` summary object.

```
summary(M3)$BaselineValues

## [[1]]
##           time      lower      median      upper
## [1,] 0.0000000 0.05914499 0.1078837 0.1960548
## [2,] 0.9537108 0.05882641 0.1067742 0.1930067
## [3,] 1.9074216 0.06093456 0.1098323 0.1978466
## [4,] 2.8611324 0.06088605 0.1096976 0.1974482
## [5,] 3.8148432 0.06099702 0.1099426 0.1979600
## [6,] 4.7685540 0.06082756 0.1098172 0.1979527
## [7,] 5.7222648 0.06110024 0.1104549 0.1994303
## [8,] 6.6759756 0.06206274 0.1123404 0.2035138
## [9,] 7.6296864 0.06304201 0.1144970 0.2087495
## [10,] 8.5833972 0.06390706 0.1167212 0.2149602
## [11,] 9.5371080 0.06369090 0.1170300 0.2173457
## [12,] 10.4908188 0.06307107 0.1167233 0.2186756
## [13,] 11.4445296 0.06197156 0.1157375 0.2186888
## [14,] 12.3982404 0.06102712 0.1152650 0.2204027
## [15,] 13.3519512 0.06040245 0.1154787 0.2243391
## [16,] 14.3056620 0.05960284 0.1154480 0.2278457
```

It is possible to convert summary statistics to have hazards ratios instead of the mean parameter values in the linear predictor (does not affect baseline parameters).

```
summary(M3, hazr=TRUE)
```

```
##
## Survival outcome
##               exp(mean)      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)    0.0102 0.0205      0.0000  0.0033   0.0648
## drugDpenicil              1.2751 0.5048      0.5551  1.1835   2.5219
## sexfemale                 0.6163 0.1992      0.3140  0.5858   1.0925
## drugDpenicil:sexfemale     0.8753 0.3865      0.3451  0.7988   1.8474
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                                     -664.5177                                -664.5669
##
## Deviance Information Criterion: -1288.312
## Widely applicable Bayesian information criterion: 1297.794
## Computation time: 0.89 seconds
```

Results are similar to the coxph function fit from survival:

```
survival::coxph(Surv(years, death) ~ drug * sex, data = Surv)
```

```
## Call:
## survival::coxph(formula = Surv(years, death) ~ drug * sex, data = Surv)
##
##               coef exp(coef) se(coef)      z      p
## drugD-penicil      0.1697    1.1849  0.3975  0.427 0.669
## sexfemale         -0.5299    0.5886  0.3296 -1.608 0.108
## drugD-penicil:sexfemale -0.2219    0.8010  0.4394 -0.505 0.614
##
## Likelihood ratio test=7.99 on 3 df, p=0.04629
## n= 312, number of events= 140
```

While the default baseline risk is set to first order random walk with 15 intervals, we can easily switch to parametric Weibull:

```
M3_wei <- joint(formSurv = DTH ~ drug * sex, dataSurv = Surv, basRisk = "weibullsurv")
summary(M3_wei, hazr=T)
```

```
##
## Survival outcome
##               exp(mean)      sd 0.025quant 0.5quant 0.975quant
## Weibull (shape)      1.0513 0.0715      0.9235  1.0466   1.2058
## Weibull (scale)      0.1029 0.0330      0.0526  0.0980   0.1816
## drugDpenicil        1.3222 0.5305      0.5690  1.2248   2.6354
## sexfemale           0.6361 0.2094      0.3199  0.6036   1.1383
## drugDpenicil:sexfemale 0.8472 0.3783      0.3304  0.7716   1.8003
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                                     -520.4002                                -520.6734
##
## Deviance Information Criterion: 1023.807
## Widely applicable Bayesian information criterion: 1024.084
## Computation time: 0.84 seconds
```

## Model 4: Longitudinal - survival joint model

An additional argument is required to set up the association between the longitudinal and survival parts:

- **assoc:** a character string that specifies the association between the longitudinal and survival components. The available options are “CV” for sharing the current value of the linear predictor, “CS” for the current slope, “CV\_CS” for the current value and the current slope, “SRE” for shared random effects (i.e., sharing the individual deviation from the mean at time  $t$  as defined by the random effects), “SRE\_ind” for shared random effect independent (each random effect’s individual deviation is associated to an association parameter in the survival submodel) and “ ” (empty string) for no association.

The model structure is given by the following equation:

$$\left\{ \begin{array}{l} \log(\text{serBilir}_{ij}) = \eta_i(t_{ij}) + \varepsilon_{ij} \\ \quad = \beta_0 + \beta_1 \text{year}_{ij} + (\beta_2 + b_{i2}) \text{year}_{ij}^2 \\ \quad \quad + (\beta_3 + b_{i3}) \text{year}_{ij}^3 + \beta_4 \text{drug}_i + \beta_5 \text{year}_{ij} \text{drug}_i \\ \quad \quad + \beta_6 \text{year}_{ij}^2 \text{drug}_i + \beta_7 \text{year}_{ij}^3 \text{drug}_i + \varepsilon_{ij} \\ \lambda_{i1}(t) = \lambda_{01}(t) \exp(\gamma_1 \text{drug}_i + \varphi_1 \eta_i(t) + \varphi_2 \eta'_i(t)) \end{array} \right. \quad \begin{array}{l} \text{(L1)} \\ \text{(S1)} \end{array}$$

where  $\gamma$  denotes fixed effects of the survival part and  $\varphi$  the association parameters.

```
# Set up quadratic and cubic functions of time
f1 <- function(x) x^2
f2 <- function(x) x^3

M4 <- joint(formSurv = DTH ~ drug,
            formLong = serBilir ~ (1 + year + f1(year) + f2(year))*drug +
                                (f1(year) + f2(year) |id), family = "lognormal",
            dataLong = Longi, id = "id", timeVar = "year", assoc = "CV_CS",
            basRisk = "rw2", NbasRisk=25, control=list(int.strategy="eb"))
summary(M4)
```

```
## Longitudinal outcome (lognormal)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1      0.6335 0.0529      0.5299  0.6335    0.7371
## year_L1           -0.1102 0.0674     -0.2424 -0.1102    0.0220
## f1year_L1          0.0372 0.0277     -0.0171  0.0372    0.0915
## f2year_L1          0.0212 0.0074      0.0068  0.0212    0.0357
## drugDpenicil_L1    -0.1339 0.0744     -0.2796 -0.1339    0.0119
## year:drugDpenicil_L1 -0.0072 0.0945     -0.1924 -0.0072    0.1781
## f1year:drugDpenicil_L1 0.0161 0.0391     -0.0605  0.0161    0.0926
## f2year:drugDpenicil_L1 -0.0037 0.0103     -0.0239 -0.0037    0.0164
## Res. err. (variance) 0.6415 0.0227      0.5985  0.6410    0.6879
##
## Random effects variance-covariance (L1)
##               mean      sd 0.025quant 0.5quant 0.975quant
## f1year_L1          0.0232 4e-03      0.0167  0.0228    0.0322
## f2year_L1          0.0054 6e-04      0.0044  0.0054    0.0066
## f1year_L1:f2year_L1 -0.0004 4e-04     -0.0013 -0.0004    0.0004
##
## Survival outcome
##               mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S1 0.2025 0.0397      0.1333  0.1998    0.2887
## drugDpenicil_S1       -0.1759 0.2329     -0.6324 -0.1759    0.2806
##
## Association longitudinal - survival
```

```
##           mean      sd 0.025quant 0.5quant 0.975quant
## CV_L1_S1 1.0500 0.0949      0.8575      1.0510      1.2352
## CS_L1_S1 0.7087 0.1967      0.3269      0.7058      1.1114
##
## log marginal-likelihood (integration)      log marginal-likelihood (Gaussian)
##                                     -33378.43                                     -33371.71
##
## Deviance Information Criterion:  -51012.96
## Widely applicable Bayesian information criterion:  -52056.69
## Computation time: 16.84 seconds
```

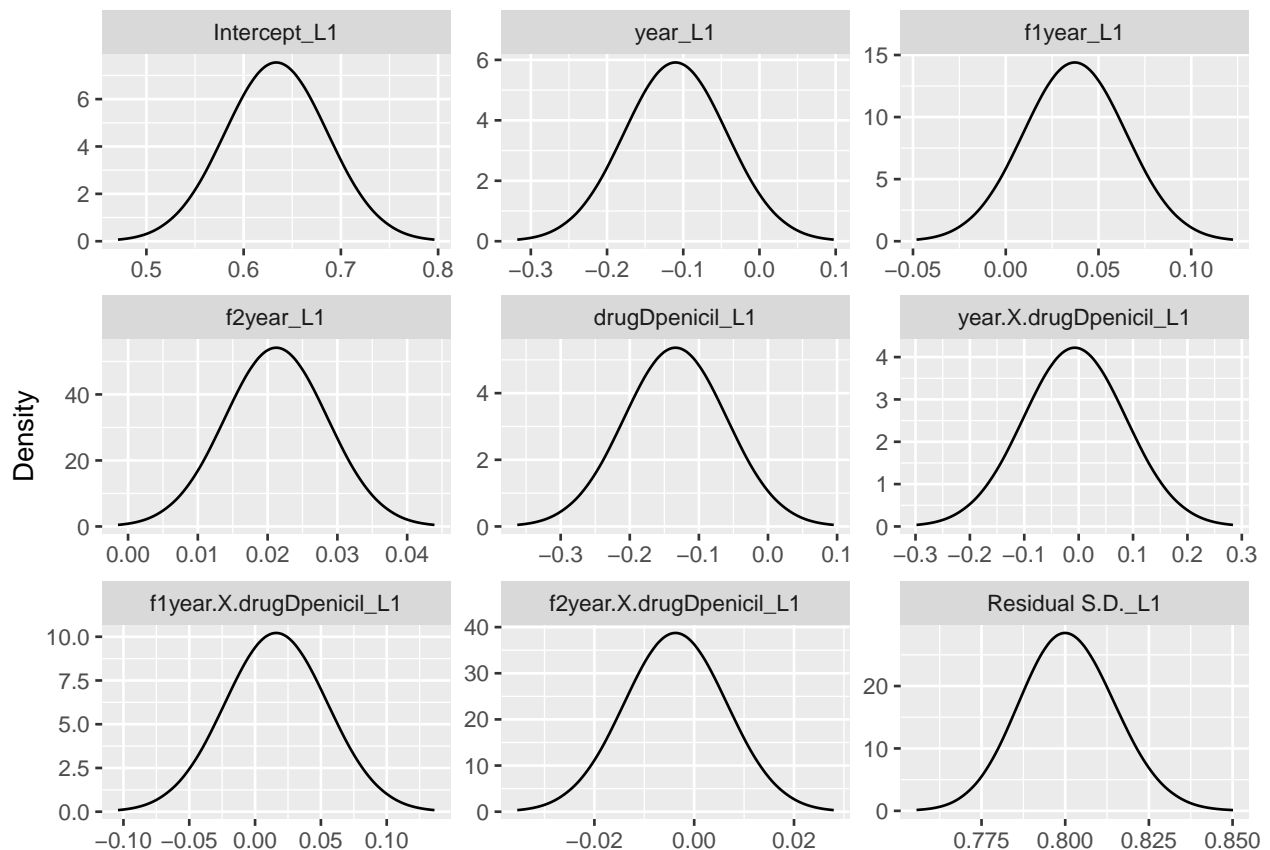
In case some functions of time should be included, they must be set as illustrated in the above example, i.e., create a univariate function of  $x$  named  $f_1, f_2, \dots, f_N$ , and use this function in the formula (where  $N=20$  is the maximum number of functions available at the moment but could be increased if needed.). This is important to be able to compute the value of the linear predictor at any time  $t$ , particularly for the time-dependent association structures. A numerical approximation of the derivative of the function is automatically computed in case the current slope of the linear predictor is shared in the survival submodel.

We can plot the posterior distribution for all the parameters with the `plot` function

```
plotM4 <- plot(M4, sdcor=T)
```

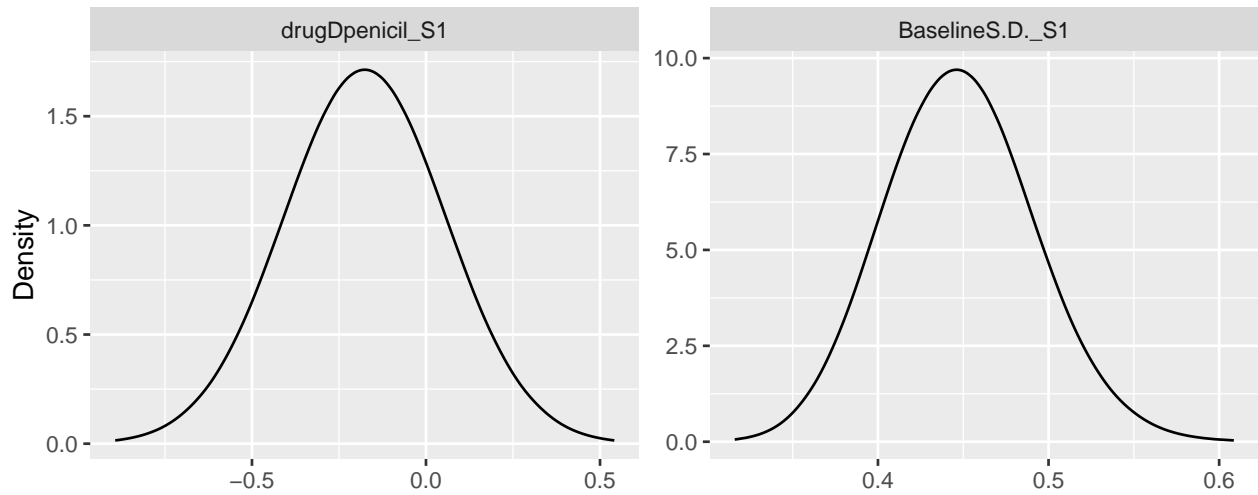
The `plot` function returns multiple plots for each component of the model. First the plots for the longitudinal outcome(s) parameters:

```
plotM4$Outcomes$L1
```



Then the parameters of the survival outcome(s):

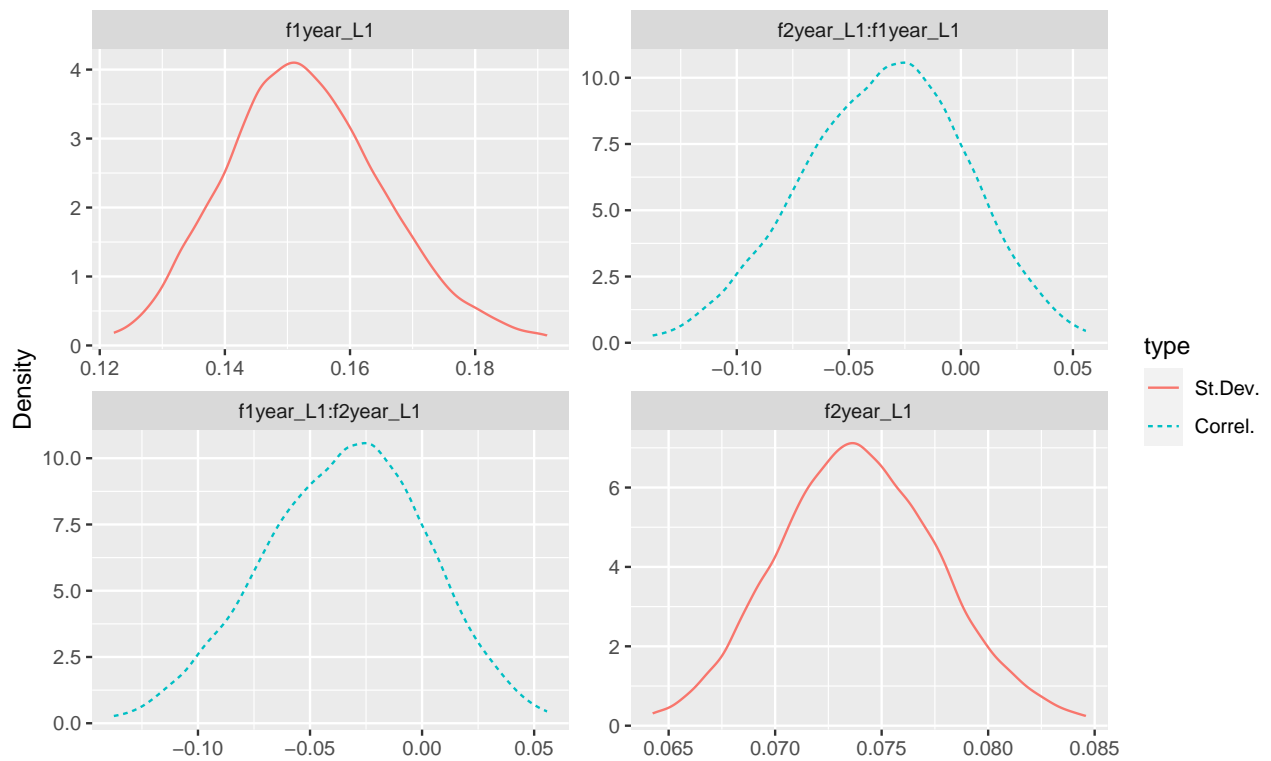
```
plotM4$Outcomes$S1
```



The variance-covariance of the random-effects (converted to standard deviations and correlations when argument `sdcor=TRUE` is added to the call of the plot function):

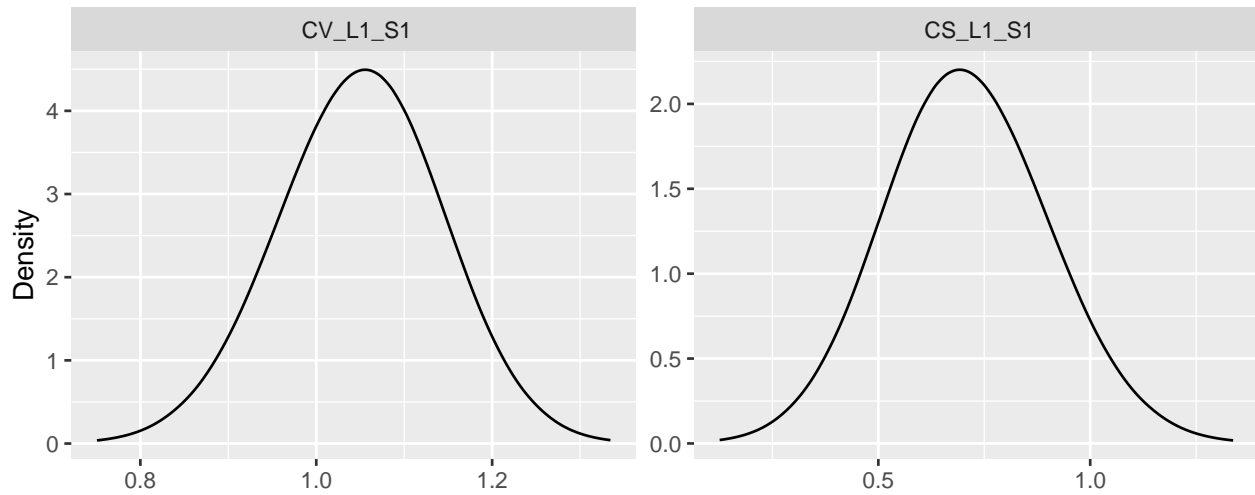
```
plotM4$Covariances
```

```
## $L1
```



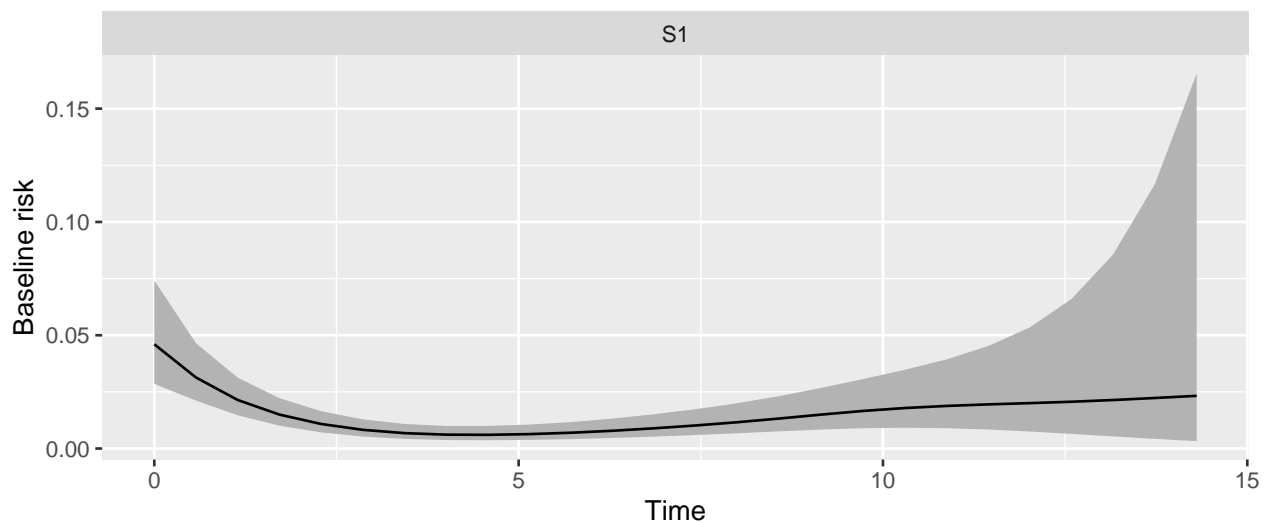
The posterior distributions of the association parameters:

```
plotM4$Associations
```



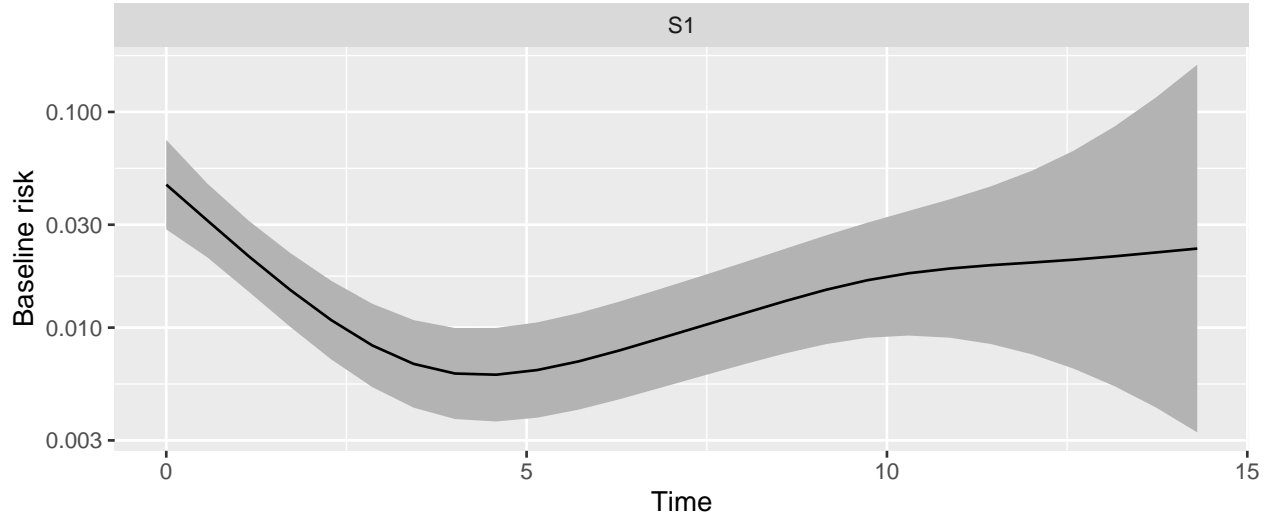
And finally the curve for the baseline risk functions:

```
plotM4$Baseline
```



The model for the baseline risk is a random walk with number of bins given by argument `NbasRisk`, the curve plotted is constant between the bins but converges towards a smooth spline when the number of bins increase. Sometimes the scale for the baseline hazard risk may require to have a `log10` y-axis, this can easily be done using to the `ggplot2` framework. Moreover, the data associated to each plot is available in the object that contains the result of the plot function call (i.e., `PlotM4` in the example).

```
plotM4$Baseline + scale_y_log10()
```



## Model 5: Comparison with MCMC

We can make a comparison of `INLAjoint` with Bayesian estimations with MCMC implemented in alternative R packages such as `JMBayes` (JAGS) or `rstanarm` (Stan).

We propose a comparison for a simple joint model with one longitudinal and one survival component:

$$\begin{cases} \text{albumin}_{ij} = \eta_i(t_{ij}) + \varepsilon_{ij} & (\text{L1}) \\ \lambda_i(t) = \lambda_0(t) \exp(\gamma_1 \text{sex}_i + \gamma_2 \text{drug}_i + \varphi \eta_i(t)) & (\text{S1}) \end{cases}$$

```
# INLAjoint
M5 <- joint(formSurv = DTH ~ sex + drug,
            formLong = albumin ~ (1 + year)*drug + (1 + year | id),
            dataLong = Longi, id = "id", timeVar = "year",
            assoc = "CV", control=list(priorFixed=list(mean=0, prec=0.16,
            mean.intercept=0, prec.intercept=0.16),
            priorAssoc=list(mean=0, prec=0.16)))
summary(M5)
```

Here the prior distributions of the fixed effects and association parameters are changed to have precision 0.16 (i.e., variance 6.25 instead of the default value of 100), in order to match the default prior distributions of `rstanarm` for the fixed effects and association parameters.

```
## Longitudinal outcome (gaussian)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1    3.5439 0.0332    3.4788    3.5439    3.6089
## year_L1        -0.1001 0.0119   -0.1235   -0.1001   -0.0768
## drugDpenicil_L1  0.0090 0.0467   -0.0826    0.0090    0.1005
## year:drugDpenicil_L1 0.0024 0.0165   -0.0301    0.0024    0.0348
## Res. err. (variance) 0.0957 0.0035    0.0890    0.0956    0.1028
##
## Random effects variance-covariance (L1)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1    0.1326 0.0138    0.1077    0.1317    0.1625
## year_L1         0.0115 0.0013    0.0091    0.0114    0.0142
## Intercept_L1:year_L1 -0.0011 0.0038   -0.0089   -0.0010    0.0061
##
```

```

## Survival outcome
##
##               mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S1 0.0424 0.0425      0.0027  0.0293      0.1589
## sexfemale_S1              -0.7152 0.2528      -1.2111 -0.7152      -0.2194
## drugDpenicil_S1           0.0075 0.1949      -0.3747  0.0076      0.3898
##
## Association longitudinal - survival
##               mean      sd 0.025quant 0.5quant 0.975quant
## CV_L1_S1 -3.1278 0.2252      -3.5544 -3.1343      -2.6643
##
## log marginal-likelihood (integration)      log marginal-likelihood (Gaussian)
##               -10406.14                      -10406.64
##
## Deviance Information Criterion: -16253.24
## Widely applicable Bayesian information criterion: -16290.68
## Computation time: 4.98 seconds

# JMBayes
library(JMbayes)
M5JMB_lme <- lme(albumin ~ (1 + year)*drug,
                 random = ~ 1 + year |id, data = Longi)
M5JMB_cox <- coxph(Surv(Surv$years, Surv$death) ~ sex + drug,
                  data = Surv, x = TRUE)
JMpr = list(priorMean.alphas=0, priorTau.alphas = matrix(0.16))
M5JMB <- jointModelBayes(M5JMB_lme, M5JMB_cox, timeVar = "year", priors=JMpr)
# Computation time in the table includes LME + Cox + JM
Summary(M5JMB)

##
## Call:
## jointModelBayes(lmeObject = M4JMB_lme, survObject = M4JMB_cox,
##   timeVar = "year", priors = JMpr)
##
## Data Descriptives:
## Longitudinal Process      Event Process
## Number of Observations: 1866 Number of Events: 140 (44.9%)
## Number of subjects: 312
##
## Joint Model Summary:
## Longitudinal Process: Linear mixed-effects model
## Event Process: Relative risk model with penalized-spline-approximated
##   baseline risk function
## Parameterization: Time-dependent value
##
## LPML      DIC      pD
## -Inf 5842.818 1807.557
##
## Variance Components:
##               StdDev      Corr
## (Intercept)  0.4436 (Intr)
## year         1.9769 0.1275
## Residual     0.3131
##
## Coefficients:
## Longitudinal Process

```



```
##               Value Std.Err Std.Dev   2.5% 97.5%      P
## (Intercept)      3.5691  0.0010  0.0393  3.4921 3.6455 <0.001
## year             -0.2904  0.0048  0.1654 -0.6202 0.0428  0.083
## drugD-penicil    -0.0013  0.0015  0.0550 -0.1081 0.1072  0.976
## year:drugD-penicil 0.1400  0.0063  0.2351 -0.3361 0.6002  0.561
##
## Event Process
##               Value Std.Err Std.Dev   2.5% 97.5%      P
## sexfemale        -0.7576  0.0241  0.3374 -1.4324 -0.1215  0.028
## drugD-penicil    -0.0091  0.0118  0.2418 -0.4821  0.4494  0.988
## Assoct           -3.4237  0.0962  0.3077 -4.0608 -2.8758 <0.001
## tauBs            350.9354 29.3848 257.8804 47.5459 1031.0303    NA
##
## MCMC summary:
## iterations: 20000
## adapt: 3000
## burn-in: 3000
## thinning: 10
## time: 1.7 min
```

```
# rstanarm
library(rstanarm)
library(survival)
options(mc.cores = parallel::detectCores())
M5rstanarm <- stan_jm(
  formulaLong = list(albumin ~ (1 + year)*drug + (1 + year | id)),
  formulaEvent = Surv(years, death) ~ sex + drug,
  dataLong = Longi, dataEvent = Surv,
  time_var = "year",
  priorLong_intercept = normal(0, 2.5, autoscale=TRUE),
  priorLong = normal(0, 2.5),
  priorEvent_assoc = normal(0, 2.5),
  seed = 12345)
summary(M5rstanarm)
```

```
## stan_jm
## formula (Long1): albumin ~ (1 + year) * drug + (1 + year | id)
## family (Long1): gaussian [identity]
## formula (Event): Surv(years, death) ~ sex + drug
## baseline hazard: bs
## assoc:          etavalue (Long1)
## -----
##
## Longitudinal submodel: albumin
##               Median MAD_SD
## (Intercept)      3.532  0.032
## year             -0.088  0.008
## drugD-penicil      0.012  0.046
## year:drugD-penicil -0.002  0.010
## sigma            0.314  0.006
##
## Event submodel:
##               Median MAD_SD exp(Median)
## (Intercept)      10.229    1.233 27681.664
## sexfemale        -0.792    0.262    0.453
```

```
## drugD-penicil      -0.003      0.202      0.997
## Long1|etavalue     -3.697      0.361      0.025
## b-splines-coef1    -0.170      0.488      NA
## b-splines-coef2    -0.516      0.491      NA
## b-splines-coef3     0.122      0.441      NA
## b-splines-coef4    -2.821      0.807      NA
## b-splines-coef5     0.669      1.199      NA
## b-splines-coef6    -3.378      1.587      NA
##
## Group-level error terms:
## Groups Name          Std.Dev. Corr
## id      Long1|(Intercept) 0.36480
##          Long1|year      0.05028 0.01
## Num. levels: id 312
##
## Sample avg. posterior predictive distribution
## of longitudinal outcomes:
##           Median MAD_SD
## Long1|mean_PPD 3.412 0.010
##
## -----
## For info on the priors used see help('prior_summary.stanreg').
```

Package	INLAjoint	JMbayes	rstanarm
algorithm	INLA	JAGS MCMC	Stan MCMC
comp. time	5 sec.	109 sec.	596 sec.

A more detailed comparison between INLA and MCEM and MCMC is available at <https://arxiv.org/abs/2203.06256> and a comparison between INLA and Levenberg-Marquardt algorithm (Newton-Raphson like that performs MLE) is available at <https://arxiv.org/abs/2010.13704>

## Model 6: Joint with one longitudinal and competing risks of event

In order to account for competing risks of event, the `formSurv` argument is given as a list with one element for each risk submodel. Moreover, the `basRisk` argument must be a vector with the same number of elements as the number of survival submodels.

The model structure is given by the following equation:

$$\left\{ \begin{array}{l} \log(\text{serBilir}_{ij}) = \eta_i(t_{ij}) + \varepsilon_{ij} \\ \quad = \beta_0 + b_{i0} + (\beta_1 + b_{i1})\text{year}_{ij} + \beta_2\text{drug}_i + \beta_3\text{sex}_i \\ \quad \quad + \beta_4\text{year}_{ij}\text{drug}_i + \beta_5\text{year}_{ij}\text{sex}_i + \varepsilon_{ij} \\ \lambda_{i1}^{\text{death}}(t) = \lambda_{01}(t) \exp(\gamma_{11}\text{sex}_i + \gamma_{12}\text{drug}_i + \varphi_{11}(b_{i0} + b_{i1}t)) \\ \lambda_{i2}^{\text{transpl.}}(t) = \lambda_{02}(t) \exp(\gamma_{21}\text{edema\_no}_i + \gamma_{22}\text{edema\_de}_i + \gamma_{23}\text{sex}_i \\ \quad \quad + \gamma_{24}\text{edema\_no}_i\text{sex}_i + \gamma_{25}\text{edema\_de}_i\text{sex}_i + \varphi_{21}b_{i0} + \varphi_{22}b_{i1}) \end{array} \right. \quad \begin{array}{l} \text{(L1)} \\ \text{(S1)} \\ \text{(S2)} \end{array}$$

```
# set up competing time-to-event outcome
TSP <- inla.surv(time = Surv$years, event = Surv$trans)

M6 <- joint(formLong = serBilir ~ year * (drug + sex) + (1+year|id), dataLong = Longi,
            formSurv = list(DTH ~ sex + drug,
                            TSP ~ edema * sex),
            id = "id", timeVar = "year", family = "lognormal", basRisk = c("rw1", "rw1"),
```

```

    assoc = c("SRE", "SRE_ind"), control=list(int.strategy="eb"))
summary(M6)

```

```

## Longitudinal outcome (lognormal)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1      0.8065 0.1773      0.4591  0.8065      1.1540
## year_L1           0.1832 0.0370      0.1107  0.1832      0.2558
## drugDpenicil_L1   -0.1276 0.1080     -0.3393 -0.1276      0.0841
## sexfemale_L1      -0.2855 0.1767     -0.6319 -0.2855      0.0609
## year:drugDpenicil_L1 0.0097 0.0233     -0.0359  0.0097      0.0553
## year:sexfemale_L1 -0.0379 0.0362     -0.1089 -0.0379      0.0330
## Res. err. (variance) 0.1090 0.0042      0.1009  0.1089      0.1175
##
## Random effects variance-covariance (L1)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1      0.9939 0.0846      0.8344  0.9914      1.1694
## year_L1           0.0345 0.0045      0.0270  0.0342      0.0440
## Intercept_L1:year_L1 0.0538 0.0174      0.0221  0.0528      0.0907
##
## Survival outcome (S1)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S1 0.1956 0.1086      0.0611  0.1706      0.4791
## sexfemale_S1       -0.8211 0.3456     -1.4985 -0.8211     -0.1438
## drugDpenicil_S1    -0.1023 0.2415     -0.5755 -0.1023      0.3710
##
## Survival outcome (S2)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S2      0.3827 0.2057      0.1113  0.3413      0.9030
## edemaedema.no.diuretics_S2     -0.5510 1.2256     -2.9532 -0.5510      1.8511
## edemaedema.despite.diuretics_S2  0.3955 1.2395     -2.0338  0.3955      2.8248
## sexfemale_S2           0.1121 0.8046     -1.4648  0.1121      1.6891
## edemaedema.no.diuretics:sexfemale_S2  0.8141 1.2976     -1.7291  0.8141      3.3573
## edemaedema.despite.diuretics:sexfemale_S2 -1.4382 1.3897     -4.1619 -1.4382      1.2856
##
## Association longitudinal - survival
##               mean      sd 0.025quant 0.5quant 0.975quant
## SRE_L1_S1          1.1519 0.0840      0.9909  1.1502      1.3219
## SRE_Intercept_L1_S2 1.0714 0.2049      0.6594  1.0745      1.4674
## SRE_year_L1_S2      1.7483 0.7998      0.1443  1.7587      3.2989
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##               -12088.67                      -12080.11
##
## Deviance Information Criterion: 18754.82
## Widely applicable Bayesian information criterion: -13897.77
## Computation time: 11.02 seconds

```

## Model 7: Joint with three longitudinal markers and competing risks of event

When multiple longitudinal submodels and survival submodels are included, the arguments `formSurv` and `formLong` are both given as lists. The `assoc` parameter should then be a list with one element for each longitudinal submodel and each element is a vector for the association with each survival submodel.

The model structure is given by the following equation:

$$\begin{cases}
\log(\text{serBilir}_{ij}) = \eta_{i1}(t_{ij}) + \varepsilon_{ij1} = \beta_{10} + b_{i10} + (\beta_{11} + b_{i11})\text{year}_{ij} + \beta_{12}\text{drug}_i + \beta_{13}\text{sex}_i & (L1) \\
\quad \quad \quad + \beta_{14}\text{year}_{ij}\text{drug}_i + \varepsilon_{ij1} \\
\log(E[\text{platelets}_{ij}]) = \eta_{i2}(t_{ij}) = \beta_{20} + b_{i20} + (\beta_{21} + b_{i21})\text{year}_{ij} + \beta_{22}\text{sex}_i + \beta_{23}\text{drug}_i + \beta_{24}\text{year}_{ij}\text{sex}_i & (L2) \\
\logit(E[\text{spiders}_{ij}]) = \eta_{i3}(t_{ij}) = \beta_{30} + b_{i30} + (\beta_{31} + b_{i31})\text{year}_{ij} + \beta_{32}\text{drug}_i + \beta_{33}\text{year}_{ij}\text{drug}_i & (L3) \\
\lambda_{i1}(t) = \lambda_{01}(t) \exp(\gamma_{11}\text{drug}_i + \varphi_{11}\eta_{i1}(t) + \varphi_{12}(b_{i20} + b_{i21}t) + \varphi_{13}\eta_{i3}(t) + \varphi_{14}\eta'_{i3}(t)) & (S1) \\
\lambda_{i2}(t) = \lambda_{02}(t) \exp(\gamma_{21}\text{drug}_i + \varphi_{21}\eta_{i1}(t) + \varphi_{22}\eta'_{i3}(t)) & (S2)
\end{cases}$$

```

M7 <- joint(formLong = list(serBilir ~ year * drug + sex + (1|id),
                           platelets ~ year + f1(year) + drug + sex + (1|id),
                           albumin ~ year + f1(year) + f2(year) + drug + (1|id)),
            formSurv = list(DTH ~ drug,
                           TSP ~ drug),
            dataLong = Longi, id = "id", corLong=TRUE, timeVar = "year",
            family = c("lognormal", "poisson", "gaussian"), basRisk = c("rw1", "rw1"),
            assoc = list(c("CV", "CV"), c("SRE", ""), c("CV_CS", "CS")),
            control=list(int.strategy="eb"))
summary(M7)

```

```

## Longitudinal outcome (L1, lognormal)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1      1.0065 0.1693      0.6747  1.0065  1.3383
## year_L1           0.0814 0.0060      0.0696  0.0814  0.0931
## drugDpenicil_L1   -0.1643 0.1245     -0.4083 -0.1643  0.0796
## sexfemale_L1      -0.3930 0.1599     -0.7063 -0.3930 -0.0797
## year:drugDpenicil_L1 0.0150 0.0083     -0.0014 0.0150  0.0313
## Res. err. (variance) 0.2147 0.0076      0.2003  0.2145  0.2302
##
## Longitudinal outcome (L2, poisson)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L2      5.4156 0.0679      5.2825  5.4156  5.5487
## year_L2           -0.0572 0.0016     -0.0603 -0.0572 -0.0541
## f1year_L2          0.0018 0.0002      0.0015  0.0018  0.0021
## drugDpenicil_L2   -0.0726 0.0449     -0.1607 -0.0726  0.0155
## sexfemale_L2      0.1061 0.0664     -0.0240  0.1061  0.2363
##
## Longitudinal outcome (L3, gaussian)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L3      3.5178 0.0343      3.4505  3.5178  3.5851
## year_L3           -0.0596 0.0154     -0.0899 -0.0596 -0.0294
## f1year_L3         -0.0068 0.0034     -0.0135 -0.0068 -0.0001
## f2year_L3          0.0005 0.0002      0.0002  0.0005  0.0009
## drugDpenicil_L3    0.0090 0.0454     -0.0800  0.0090  0.0980
## Res. err. (variance) 0.1114 0.0039      0.1039  0.1113  0.1193
##
## Random effects variance-covariance
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1      1.1407 0.0976      0.9653  1.1344  1.3421
## Intercept_L2      0.1583 0.0128      0.1351  0.1576  0.1850
## Intercept_L3      0.1365 0.0133      0.1118  0.1358  0.1645
## Intercept_L1:Intercept_L2 -0.1126 0.0272     -0.1697 -0.1123 -0.0613
## Intercept_L1:Intercept_L3 -0.2471 0.0290     -0.3082 -0.2458 -0.1956
## Intercept_L2:Intercept_L3  0.0515 0.0098      0.0340  0.0509  0.0715
##

```

```
## Survival outcome (S1)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S1 0.0429 0.0480      0.0025  0.0276    0.1765
## drugDpenicil_S1          -0.0017 0.1832      -0.3608 -0.0017    0.3574
##
## Survival outcome (S2)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S2 0.1472 0.2719      0.0048  0.0615    0.8544
## drugDpenicil_S2          -0.4464 0.3746      -1.1807 -0.4464    0.2879
##
## Association longitudinal - survival
##               mean      sd 0.025quant 0.5quant 0.975quant
## CV_L1_S1      1.0389 0.1102      0.8154  1.0413    1.2496
## CV_L1_S2      1.0676 0.1850      0.7040  1.0672    1.4336
## SRE_L2_S1     -0.3663 0.2183     -0.7966 -0.3663    0.0643
## CV_L3_S1     -2.3857 0.3308     -3.0584 -2.3775   -1.7556
## CS_L3_S1     -0.0451 2.5044     -4.9861 -0.0551    4.9498
## CS_L3_S2      0.2687 3.7482     -7.1309  0.2732    7.6461
##
## log marginal-likelihood (integration)      log marginal-likelihood (Gaussian)
##               -56519.34                      -56504.35
##
## Deviance Information Criterion: -42314.91
## Widely applicable Bayesian information criterion: -6043.845
## Computation time: 31.15 seconds
```

The longitudinal markers are assumed correlated but it is also possible to set `corLong` to `FALSE` to have independent random effects across markers and reduce the number of covariance parameters.

## Model 8: Multi-state model

For the multi-state model, we need to define a survival model for each transition. Here for example in the case of illness-death we have three possible transitions: 1->2, 1->3 and 2->3.

The model structure is given by the following equation:

$$\begin{cases} \lambda_{i,12}(t) = \lambda_{0,12}(t) \exp(\gamma_{12}X_i)(S1) \\ \lambda_{i,13}(t) = \lambda_{0,13}(t) \exp(\gamma_{13}X_i)(S2) \\ \lambda_{i,23}(t) = \lambda_{0,23}(t) \exp(\gamma_{23}X_i)(S3) \end{cases}$$

Note that the baseline risk function is transition-specific (one baseline risk for each survival model). The dataset for this application is simulated and included in `INLAjoint`.

```
# set up outcomes for each transition
data(SurvMS) # load small simulated dataset for multi-state
E12 <- inla.surv(time = SurvMS[[1]]$Tstop, event = SurvMS[[1]]$status) # transition 1->2
E13 <- inla.surv(time = SurvMS[[2]]$Tstop, event = SurvMS[[2]]$status) # transition 1->3
E23 <- inla.surv(time = SurvMS[[3]]$Tstop, truncation=SurvMS[[3]]$Tstart,
                 event =SurvMS[[3]]$status) # transition 2->3

M8 <- joint(formSurv=list(E12 ~ X, E13 ~ X, E23 ~ X),
            basRisk = c("rw2", "rw1", "exponentialsurv"), dataSurv = SurvMS)
summary(M8)

##
## Survival outcome (S1)
```

```
##               mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S1 0.4233 0.2440      0.1229  0.3669    1.0604
## X_S1                0.3013 0.0867      0.1315  0.3011    0.4717
##
## Survival outcome (S2)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S2 0.5017 0.3498      0.0908  0.4173    1.4181
## X_S2                0.5028 0.0924      0.3246  0.5017    0.6870
##
## Survival outcome (S3)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Exponential (rate)_S3  0.0779 0.0328      0.0321  0.0716    0.1597
## X_S3                -0.0513 0.0897     -0.2272 -0.0513    0.1246
##
## log marginal-likelihood (integration)      log marginal-likelihood (Gaussian)
##               -239.8301                      -239.8301
##
## Deviance Information Criterion: 377.675
## Widely applicable Bayesian information criterion: 377.3267
## Computation time: 1.15 seconds
```

## Model 9: Joint longitudinal and multi-state model

We can extend the previous model to joint longitudinal and multi-state. The model structure is given by the following equation:

$$\begin{cases} Y_{ij} = \eta_i(t_{ij}) + \varepsilon_{ij} = \beta_0 + b_{i0} + (\beta_1 + b_{i1})time_{ij} + \beta_{12}X_i + \varepsilon_{ij} & (L1) \\ \lambda_{i,12}(t) = \lambda_{0,12}(t) \exp(\gamma_{12}X_i + \varphi_{12}\eta_i(t)) & (S1) \\ \lambda_{i,13}(t) = \lambda_{0,13}(t) \exp(\gamma_{13}X_i + \varphi_{13}\eta_i(t)) & (S2) \\ \lambda_{i,23}(t) = \lambda_{0,23}(t) \exp(\gamma_{23}X_i + \varphi_{23}\eta_i(t)) & (S3) \end{cases}$$

```
data(LongMS) # load longitudinal data for joint longitudinal and multi-state
M9 <- joint(formSurv=list(E12 ~ X, E13 ~ X, E23 ~ X),
            formLong=list(y ~ time + X + (1+time|id)),
            basRisk = c("rw2", "rw1", "exponentialsurv"), timeVar = "time",
            assoc = list(c("CV", "CV", "CV")), id="id",
            dataSurv = SurvMS, dataLong = LongMS)
summary(M9)
```

```
## Longitudinal outcome (gaussian)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1    -0.4853 0.2207     -0.9178 -0.4854   -0.0519
## time_L1         0.0583 0.1218     -0.1799  0.0581    0.2980
## X_L1            1.5245 0.0367      1.4524  1.5245    1.5965
## Res. err. (variance) 1.5212 0.1387      1.2659  1.5157    1.8107
##
## Random effects variance-covariance (L1)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1    0.1917 0.1127      0.0576  0.1659    0.4742
## time_L1         0.3911 0.1118      0.2151  0.3766    0.6413
## Intercept_L1:time_L1 0.0388 0.0815     -0.1464  0.0445    0.1791
##
## Survival outcome (S1)
##               mean      sd 0.025quant 0.5quant 0.975quant
```

```
## Baseline risk (variance)_S1  0.1913 0.1419      0.0366  0.1535      0.5739
## X_S1                        -0.3888 0.1215      -0.6273  -0.3885      -0.1522
##
## Survival outcome (S2)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S2 0.6115 0.4635      0.1182  0.4854      1.8675
## X_S2                      0.2371 0.1630      -0.0752  0.2389      0.5440
##
## Survival outcome (S3)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Exponential (rate)_S3      0.1666 0.0729      0.0661  0.1523      0.3496
## X_S3                      -0.2844 0.1034      -0.4883  -0.2838      -0.0834
##
## Association longitudinal - survival
##               mean      sd 0.025quant 0.5quant 0.975quant
## CV_L1_S1 0.4811 0.1688      0.1560  0.4787      0.8205
## CV_L1_S2 0.2121 0.2738      -0.3048  0.2049      0.7725
## CV_L1_S3 0.1167 0.0886      -0.0533  0.1152      0.2953
##
## log marginal-likelihood (integration)      log marginal-likelihood (Gaussian)
##               -2396.858                      -2396.858
##
## Deviance Information Criterion: -1738.338
## Widely applicable Bayesian information criterion: -1852.247
## Computation time: 5.49 seconds
```

## Model 10: Mixture cure model

The mixture cure model is used to analyze survival data with a cure fraction. The cured patients are subject to no excess risk and the uncured patients are subject to excess risk modeled using a parametric survival distribution. Let  $p(C=1)$  denote the probability of being cured and  $\lambda_i(t|C=0)$  the survival for the part of the population that is not cured, the mixture cure model is defined by:

$$\begin{cases} \text{Logit}(p(C=1)) = \beta_0 + \beta_1 X_i & (\text{probability of being cured}) \\ \lambda_i(t|C=0) = \lambda_0(t) \exp(\gamma X_i) & (\text{survival for non-cured}) \end{cases}$$

*Note: the survival function for cured patients is always equal to 1 and the limit of the population survival  $S(t)$  converges to  $p(C=1)$  when  $t \rightarrow \infty$ .*

The mixture cure model can be added to any survival model and can deal with left, right and interval censoring. However, only parametric baseline risks can be defined at the moment for mixture cure models and the covariates that affect the probability of being cured can be associated to fixed effects only (i.e., no random effects in the logistic part).

```
library("smcure")
data("bmt")
MC_DTH <- inla.surv(time = bmt$Time, event = bmt$Status,
                   cure=cbind("Int"=1, "TRT"=bmt$TRT))
M10 <- joint(formSurv=MC_DTH ~ TRT, basRisk = "weibullsurv", dataSurv = bmt)
summary(M10)
```

```
##
## Survival outcome
##               mean      sd 0.025quant 0.5quant 0.975quant
## Int(cure)      -1.0346 0.3482      -1.7214  -1.0343      -0.3494
```

```
## TRT(cure)          -0.4220 0.5232      -1.4262 -0.4323      0.6348
## Weibull (shape)    1.0819 0.0953      0.9210 1.0726      1.2954
## Weibull (scale)    0.0029 0.0010      0.0013 0.0028      0.0050
## TRT                0.6968 0.2567      0.1930 0.6969      1.1998
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                                         -486.6769                                -486.8218
##
## Deviance Information Criterion: 954.5002
## Widely applicable Bayesian information criterion: 956.0407
## Computation time: 0.79 seconds
```

## Model 11: Shared frailty model for recurrent events

The frailty model can fit recurrent events with a gaussian distribution for the frailty term (i.e., lognormal frailty model).

$$\lambda_i(t) = \lambda_0(t) \exp(\gamma_1 Sex_i + \gamma_2 Age_i + \omega_i)$$

$$\omega_i \sim \mathcal{N}(\mu, \sigma^2)$$

```
library(survival)
rec <- inla.surv(time = kidney$time, event = kidney$status)
M11 <- joint(formSurv=rec ~ sex + age + (1|id), id="id", dataSurv = kidney)
summary(M11)

##
## Survival outcome
##
##               mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance) 0.0106 0.0178      0.0000  0.0040  0.0607
## sex                    -1.4552 0.3768      -2.2057 -1.4515 -0.7256
## age                     0.0056 0.0112      -0.0162  0.0056  0.0277
##
## Frailty term variance
##               mean      sd 0.025quant 0.5quant 0.975quant
## IDIntercept 0.3532 0.1818      0.0922  0.3247  0.7845
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                                         -209.6719                                -210.1179
##
## Deviance Information Criterion: -345.7255
## Widely applicable Bayesian information criterion: 379.1821
## Computation time: 0.85 seconds
```

We can compare the results with the coxph function from survival package:

```
M11s <- coxph(Surv(time, status) ~ sex + age + frailty.gaussian(id), data = kidney)
summary(M11s)
```

```
## Call:
## coxph(formula = Surv(time, status) ~ sex + age + frailty.gaussian(id),
##       data = kidney)
##
##      n= 76, number of events= 58
##
##               coef      se(coef) se2      Chisq DF      p
```



```
## sex                -1.410822 0.44518  0.315038 10.04  1.00 0.0015
## age                0.004713 0.01248  0.008557  0.14  1.00 0.7100
## frailty.gaussian(id)                26.54 14.73 0.0290
##
##      exp(coef) exp(-coef) lower .95 upper .95
## sex      0.2439   4.0993    0.1019   0.5837
## age      1.0047    0.9953    0.9805   1.0296
##
## Iterations: 6 outer, 39 Newton-Raphson
##      Variance of random effect= 0.5691225
## Degrees of freedom for terms=  0.5  0.5 14.7
## Concordance= 0.82 (se = 0.032 )
## Likelihood ratio test= 47.55 on 15.7 df,  p=5e-05
```

## Model 12: Joint model for recurrent events and a terminal event

The joint frailty model deals with recurrent events with a frailty model and shares the frailty term in a proportional hazards model for a terminal event. The parameter  $\varphi_1$  is scaling the frailty term in the terminal event submodel.

$$\begin{cases} \lambda_{i1}(t) = \lambda_{01}(t) \exp(\gamma_{11} \text{Chemo}_i + \omega_i) & (S1) \\ \lambda_{i2}(t) = \lambda_{02}(t) \exp(\gamma_{12} \text{Chemo}_i + \varphi_1 \omega_i) & (S2) \end{cases}$$

$$\omega_i \sim \log \mathcal{N}(\mu, \sigma^2)$$

```
library(frailtypack)
data(readmission)
terminalData <- readmission[readmission$event==0,]
terminalEVENT <- inla.surv(time = terminalData$t.stop, event = terminalData$death)
recurrentEVENT <- inla.surv(time = readmission$t.stop, event = readmission$event)
M12 <- joint(formSurv=list(recurrentEVENT ~ chemo + (1|id),
                          terminalEVENT ~ chemo), id="id",
             basRisk=c("rw1", "rw1"), assocSurv=TRUE,
             dataSurv = list(readmission,terminalData))
summary(M12)
```

```
##
## Survival outcome (S1)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S1  0.0415 0.0165    0.0170  0.0390    0.0812
## chemoTreated_S1             -0.0406 0.1407    -0.3163 -0.0407    0.2355
##
## Frailty term variance (S1)
##               mean      sd 0.025quant 0.5quant 0.975quant
## IDIntercept_S1 0.7969 0.1566    0.5202  0.7879    1.1327
##
## Survival outcome (S2)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S2 0.0075 0.0028    0.0031  0.0071    0.0139
## chemoTreated_S2             0.3114 0.2652    -0.2086  0.3113    0.8318
##
## Association survival - survival
##               mean      sd 0.025quant 0.5quant 0.975quant
## IDIntercept_S1_S2 1.9277 0.282    1.4217  1.9102    2.5263
```

```
##
## log marginal-likelihood (integration)      log marginal-likelihood (Gaussian)
##                                     -2479.894                                -2479.962
##
## Deviance Information Criterion:  -4303.778
## Widely applicable Bayesian information criterion:  4745.661
## Computation time: 3.87 seconds
```

The parameter `assocSurv` indicates that the frailty term from the recurrent event model is shared in the terminal event model.

### Model 13: Two-part model for a longitudinal semicontinuous outcome

The two-part model is used to fit a semicontinuous outcome, usually zero-inflated but a point mass other than zero can also be handled. It includes a logistic mixed effects model for a binary outcome (zero vs. positive) and a linear mixed effects model for the positive-only values (can be a Poisson model for count or any other distribution available). We illustrate with an example based on PBC by making a fake zero-inflated outcome based on the albumin outcome.

The model structure is given by:

$$\begin{cases} \text{Logit}(\text{Prob}(\text{albumin}_{ij} > 0)) = \beta_{10} + b_{i10} + (\beta_{11} + b_{i11})\text{year}_{ij} + \beta_{12}\text{drug}_i + \beta_{13}\text{year}_{ij}\text{drug}_i & (\text{L1}) \\ E[\text{albumin}_{ij} | \text{albumin}_{ij} > 0] = \beta_{20} + b_{i20} + (\beta_{21} + b_{i21})\text{year}_{ij} + \beta_{22}\text{drug}_i + \beta_{23}\text{year}_{ij}\text{drug}_i + \varepsilon_{ij} & (\text{L2}) \end{cases}$$

```
# make a zero-inflated continuous outcome for illustration
Longi$binary <- ifelse(Longi$albumin<3, 0, 1)
# dataset with only positives for the continuous part
LongiPositive <- Longi[which(Longi$albumin>=3),]
M13 <- joint(formLong = list(binary ~ year * drug + (1+year|id),
                             albumin ~ year * drug + (1+year|id)),
             dataLong = list(Longi, LongiPositive), timeVar="year",
             id = "id", family = c("binomial", "gaussian"), corLong=TRUE)
summary(M13)
```

```
## Longitudinal outcome (L1, binomial)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1      3.3114 0.2494      2.8243   3.3106   3.8026
## year_L1           -0.3834 0.0644     -0.5096  -0.3834  -0.2571
## drugDpenicil_L1    -0.1950 0.3450     -0.8718  -0.1950   0.4817
## year:drugDpenicil_L1 0.0387 0.0899     -0.1377   0.0387   0.2151
##
## Longitudinal outcome (L2, gaussian)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L2      3.6019 0.0275      3.5478   3.6019   3.6559
## year_L2           -0.0632 0.0116     -0.0859  -0.0632  -0.0405
## drugDpenicil_L2     0.0138 0.0389     -0.0625   0.0138   0.0901
## year:drugDpenicil_L2 0.0060 0.0163     -0.0260   0.0060   0.0380
## Res. err. (variance) 0.0725 0.0032      0.0667   0.0722   0.0794
##
## Random effects variance-covariance
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1      3.4038 0.7410      2.1409   3.3422   5.0018
## year_L1           0.1056 0.0292      0.0591   0.1027   0.1719
## Intercept_L2      0.0803 0.0100      0.0622   0.0798   0.1004
```

```
## year_L2                0.0093 0.0011      0.0073  0.0092    0.0117
## Intercept_L1:year_L1    0.1333 0.0968    -0.0641  0.1349    0.3230
## Intercept_L1:Intercept_L2 0.4131 0.0620      0.3016  0.4102    0.5464
## Intercept_L1:year_L2    -0.0184 0.0160    -0.0486 -0.0190    0.0125
## year_L1:Intercept_L2    0.0213 0.0174    -0.0099  0.0203    0.0585
## year_L1:year_L2         0.0039 0.0021    -0.0004  0.0039    0.0082
## Intercept_L2:year_L2    -0.0043 0.0032    -0.0095 -0.0046    0.0027
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                                     -1273.642                                -1271.871
##
## Deviance Information Criterion: 1423.544
## Widely applicable Bayesian information criterion: 2361.693
## Computation time: 9.79 seconds
```

## Model 14: Two-part joint model for a longitudinal semicontinuous outcome and a terminal event

We can fit a joint two-part model by adding a survival component to the previous model. The association is based on shared random effects (each random effect is shared in the survival submodel and associated to a scaling parameter). Simulation studies for this model are available at <https://arxiv.org/abs/2010.13704>

The model structure is given by:

$$\begin{cases} \text{Logit}(\text{Prob}(\text{albumin}_{ij} > 0)) = \beta_{10} + b_{i10} + (\beta_{11} + b_{i11})\text{year}_{ij} + \beta_{12}\text{drug}_i + \beta_{13}\text{year}_{ij}\text{drug}_i & (\text{L1}) \\ E[\text{albumin}_{ij} | \text{albumin}_{ij} > 0] = \beta_{20} + b_{i20} + (\beta_{21} + b_{i21})\text{year}_{ij} + \beta_{22}\text{drug}_i + \beta_{23}\text{year}_{ij}\text{drug}_i + \varepsilon_{ij} & (\text{L2}) \\ \lambda_i(t) = \lambda_0(t) \exp(\gamma_1\text{drug}_i + \varphi_1b_{i10} + \varphi_2b_{i11} + \varphi_3b_{i20} + \varphi_4b_{i21}) & (\text{S1}) \end{cases}$$

```
M14 <- joint(formLong = list(binary ~ year * drug + (1+year|id),
                             albumin ~ year * drug + (1+year|id)),
             formSurv = DTH ~ drug,
             dataLong = list(Longi, LongiPositive), timeVar="year",
             id = "id", corLong=TRUE, assoc = c("SRE_ind", "SRE_ind"),
             family = c("binomial", "gaussian"))
summary(M14)
```

```
## Longitudinal outcome (L1, binomial)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1    3.4915 0.2651    2.9737   3.4908   4.0136
## year_L1        -0.5123 0.0700    -0.6495  -0.5123  -0.3749
## drugDpenicil_L1 -0.1814 0.3668    -0.9008  -0.1814   0.5382
## year:drugDpenicil_L1 0.0595 0.0973    -0.1312   0.0595   0.2504
##
## Longitudinal outcome (L2, gaussian)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L2    3.5971 0.0276    3.5429   3.5971   3.6513
## year_L2        -0.0669 0.0116    -0.0896  -0.0669  -0.0442
## drugDpenicil_L2  0.0148 0.0390    -0.0617   0.0148   0.0913
## year:drugDpenicil_L2 0.0070 0.0164    -0.0251   0.0070   0.0391
## Res. err. (variance) 0.0726 0.0025    0.0681   0.0725   0.0780
##
## Random effects variance-covariance
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1    4.6070 1.1443    2.8078   4.4506   7.1028
```

```

## year_L1                0.1622 0.0457      0.0892  0.1571    0.2644
## Intercept_L2           0.0798 0.0093      0.0628  0.0795    0.0994
## year_L2                0.0094 0.0013      0.0073  0.0093    0.0121
## Intercept_L1:year_L1   0.3032 0.1242      0.0383  0.3060    0.5401
## Intercept_L1:Intercept_L2 0.4699 0.0785      0.3346  0.4653    0.6389
## Intercept_L1:year_L2   -0.0154 0.0147     -0.0459 -0.0152    0.0129
## year_L1:Intercept_L2   0.0441 0.0163      0.0139  0.0433    0.0780
## year_L1:year_L2        0.0041 0.0028     -0.0009  0.0039    0.0103
## Intercept_L2:year_L2   -0.0054 0.0018     -0.0092 -0.0053   -0.0023
##
## Survival outcome
##               mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S1 0.1932 0.0880      0.0709  0.1770    0.4121
## drugDpenicil_S1          -0.0838 0.2565     -0.5870 -0.0838    0.4191
##
## Association longitudinal - survival
##               mean      sd 0.025quant 0.5quant 0.975quant
## SRE_Intercept_L1_S1 -0.5395 0.0791     -0.6932 -0.5406   -0.3783
## SRE_year_L1_S1      -1.9231 0.2674     -2.4749 -1.9178   -1.4098
## SRE_Intercept_L2_S1 -0.1060 0.3918     -0.9436 -0.0912    0.6212
## SRE_year_L2_S1      -0.6478 0.7016     -2.0640 -0.6377    0.7134
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##               -1865.935                      -1867.986
##
## Deviance Information Criterion: -1811.088
## Widely applicable Bayesian information criterion: 3453.113
## Computation time: 26.39 seconds

```

## Model 15: Two-part joint model for a longitudinal semicontinuous outcome, recurrent events and a terminal event

The model includes a two-part model for a semicontinuous longitudinal biomarker, a frailty model for recurrent events and a proportional hazards model for a terminal event. The random effects from the two mixed effects submodels for the longitudinal semicontinuous outcome are shared and scaled in the frailty model for recurrent events and the proportional hazards model for a terminal event. The frailty term is shared and scaled in the proportional hazards model. We use the `colorectal` and `colorectalLongi` datasets where the longitudinal semicontinuous biomarker is the tumor size in a colorectal cancer clinical trial, the recurrent events are new tumors and the terminal event is death.

$$\begin{cases}
\text{Logit}(\text{Prob}(Y_{ij} > 0)) = \beta_{10} + b_{i10} + \beta_{11}\text{year}_{ij} + \beta_{12}\text{drug}_i + \beta_{13}\text{year}_{ij}\text{drug}_i & (\text{L1}) \\
\text{E}[Y_{ij}|Y_{ij} > 0] = \beta_{20} + b_{i20} + (\beta_{21} + b_{i21})\text{year}_{ij} + \beta_{22}\text{drug}_i + \beta_{23}\text{year}_{ij}\text{drug}_i + \varepsilon_{ij} & (\text{L2}) \\
\lambda_{i1}(t) = \lambda_{01}(t) \exp(\gamma_1 \text{Chemo}_i + \varphi_{11}b_{i10} + \varphi_{12}b_{i20} + \varphi_{13}b_{i21} + \omega_i) & (\text{S1}) \\
\lambda_{i2}(t) = \lambda_{02}(t) \exp(\gamma_2 \text{Chemo}_i + \varphi_{21}b_{i10} + \varphi_{22}b_{i20} + \varphi_{23}b_{i21} + \varphi_{24}\omega_i) & (\text{S2})
\end{cases}$$

$$\omega_i \sim \mathcal{N}(\mu, \sigma^2)$$

```

library(frailtypack)
data(colorectal)
data(colorectalLongi)
summary(colorectalLongi)
# reverse Box-Cox transformation to have original
# tumor size measurements (zero inflated and right skewed)

```

```

# because tumor.size in the dataset is transformed
colorectalLongi$CONTINUOUS <- round((colorectalLongi$tumor.size*0.3+1)^(1/0.3), 5)
# positive value indicator (binary outcome)
colorectalLongi$BINARY <- ifelse(colorectalLongi$CONTINUOUS==0,0,1)
# extract terminal event data
colorectalSurv <- subset(colorectal, new.lesions == 0)
# extract longitudinal positive data
colorectalLongiPositive <- colorectalLongi[colorectalLongi$CONTINUOUS>0,]

RECURRENT <- inla.surv(time=colorectal$gap.time, event=colorectal$new.lesions)
TERMINAL <- inla.surv(time=colorectalSurv$time1, event=colorectalSurv$state)

M15 <- joint(formSurv=list(RECURRENT ~ treatment + (1|id),
                           TERMINAL ~ treatment),
             formLong=list(BINARY ~ year * treatment + (1|id),
                           CONTINUOUS ~ year * treatment + (1+year|id)),
             dataSurv = list(colorectal, colorectalSurv),
             dataLong = list(colorectalLongi, colorectalLongiPositive),
             id="id", timeVar="year", basRisk=c("rw2", "rw2"), corLong=TRUE,
             family=c("binomial", "lognormal"), assocSurv=TRUE,
             assoc=list(c("SRE_ind", "SRE_ind"), c("SRE_ind", "SRE_ind")))
summary(M15)

```

```

## Longitudinal outcome (L1, binomial)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1    6.9773 1.0975     4.8259   6.9769   9.1309
## year_L1         0.2340 1.0371    -1.8000   0.2340   2.2679
## treatment_L1    -2.0599 1.1735    -4.3612  -2.0599   0.2415
## year:treatment_L1 -1.0489 1.0668    -3.1409  -1.0489   1.0433
##
## Longitudinal outcome (L2, lognormal)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L2     2.1260 0.0845     1.9602   2.1260   2.2918
## year_L2          -0.1384 0.0783    -0.2921  -0.1384   0.0150
## treatment_L2     0.0541 0.1212    -0.1836   0.0541   0.2918
## year:treatment_L2 -0.3172 0.1114    -0.5359  -0.3171  -0.0988
## Res. err. (variance) 0.1638 0.0084     0.1479   0.1637   0.1808
##
## Random effects variance-covariance
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1     2.1800 1.0141     0.8179   1.9671   4.7563
## Intercept_L2     0.4865 0.0681     0.3749   0.4809   0.6331
## year_L2          0.1798 0.0488     0.1067   0.1730   0.2944
## Intercept_L1:Intercept_L2 0.7233 0.2092     0.3821   0.6972   1.2269
## Intercept_L1:year_L2    -0.1900 0.1133    -0.4414  -0.1774  -0.0045
## Intercept_L2:year_L2    -0.0907 0.0445    -0.1880  -0.0859  -0.0137
##
## Survival outcome (S1)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S1 0.1633 0.0817     0.0596   0.1449   0.3752
## treatment_S1       -0.3371 0.2333    -0.7949  -0.3371   0.1204
##
## Frailty term variance (S1)
##               mean      sd 0.025quant 0.5quant 0.975quant

```

```

## IDIntercept_S1 0.7532 0.1763      0.4697   0.7317      1.1605
##
## Survival outcome (S2)
##
##               mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S2 0.4203 0.1371      0.2299   0.3931      0.7635
## treatment_S2              -0.1513 0.3022      -0.7443  -0.1512      0.4412
##
## Association longitudinal - survival
##
##               mean      sd 0.025quant 0.5quant 0.975quant
## SRE_Intercept_L1_S1 0.2598 0.1292      0.0035   0.2600      0.5147
## SRE_Intercept_L1_S2 0.3931 0.1414      0.0943   0.3983      0.6557
## SRE_Intercept_L2_S1 -0.0818 0.1996     -0.4571  -0.0880      0.3333
## SRE_year_L2_S1      -0.1948 0.2578     -0.7301  -0.1862      0.2895
## SRE_Intercept_L2_S2 0.4849 0.2539      0.0166   0.4748      1.0205
## SRE_year_L2_S2      -0.4104 0.3409     -1.0798  -0.4119      0.2673
##
## Association survival - survival
##
##               mean      sd 0.025quant 0.5quant 0.975quant
## IDIntercept_S1_S2 1.2487 0.2078      0.8575   1.242      1.6758
##
## log marginal-likelihood (integration)      log marginal-likelihood (Gaussian)
##
##               -3494.331                      -3495.563
##
## Deviance Information Criterion: -5800.672
## Widely applicable Bayesian information criterion: 6430.379
## Computation time: 19.06 seconds

```

## Model 16: Joint model for multivariate longitudinal data and competing risks of event (application section of <https://arxiv.org/abs/2203.06256>)

The model structure is given by the following equation:

$$\begin{cases}
 \log(\text{serBilir}_{ij}) = \eta_{i1}(t_{ij}) + \varepsilon_{ij1} & (L1) \\
 \log(\text{SGOT}_{ij}) = \eta_{i2}(t_{ij}) + \varepsilon_{ij2} & (L2) \\
 \text{albumin}_{ij} = \eta_{i3}(t_{ij}) + \varepsilon_{ij3} & (L3) \\
 \log(E[\text{platelets}_{ij}]) = \eta_{i4}(t_{ij}) & (L4) \\
 \text{logit}(E[\text{spiders}_{ij}]) = \eta_{i5}(t_{ij}) & (L5) \\
 \lambda_{i1}(t) = \lambda_{01}(t) \exp(\gamma_1 X_i + \varphi_1 \eta_{i1}(t) + \varphi_3 \eta'_{i1}(t) + \varphi_4 \eta_{i2}(t) & (S1) \\
 \quad + \varphi_5 \eta_{i3}(t) + \varphi_7 \eta_{i4}(t) + \varphi_9 \eta_{i5}(t)) \\
 \lambda_{i2}(t) = \lambda_{02}(t) \exp(\gamma_2 X_i + \varphi_2 \eta_{i1}(t) + \varphi_6 \eta_{i3}(t) + \varphi_8 \eta_{i4}(t)) & (S2)
 \end{cases}$$

where  $\text{NS}_1(t)$ ,  $\text{NS}_2(t)$ ,  $\text{NS}_3(t)$  are the natural cubic splines with internal knots at 1 and 4 years. We assume independent random effects between longitudinal markers.

```

# set up natural cubic splines for longitudinal markers's trajectories
Nsplines <- ns(Longi$year, knots=c(1,4))
f1 <- function(x) predict(Nsplines, x)[,1]
f2 <- function(x) predict(Nsplines, x)[,2]
f3 <- function(x) predict(Nsplines, x)[,3]

M16 <-joint(formSurv = list(DTH ~ drug, TSP ~ drug),
            formLong = list(serBilir ~ (1 + f1(year) + f2(year) + f3(year)) * drug +
                            (1 + f1(year) + f2(year) + f3(year) | id),
                            SGOT ~ (1 + f1(year) + f2(year) + f3(year)) * drug +
                            (1 + f1(year) + f2(year) + f3(year) | id),
                            albumin ~ (1 + year) * drug + (1 + year | id),
                            platelets ~ (1 + f1(year) + f2(year) + f3(year)) * drug +
                            (1 + f1(year) + f2(year) + f3(year) | id),
                            spiders ~ (1 + year) * drug + (1 + year | id)),
            dataLong = Longi, id = "id", timeVar = "year",
            family = c("lognormal", "lognormal", "gaussian", "poisson", "binomial"),
            basRisk = c("rw2", "rw1"), NbasRisk = 15, assoc = list(c("CV_CS", "CV"),
                                                                    c("CV", ""), c("CV", "CV"), c("CV", "CV"), c("CV", "")),
            control=list(int.strategy="eb", cfg=T))
summary(M16)

```

```

## Longitudinal outcome (L1, lognormal)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1      0.5975 0.0815      0.4378  0.5975  0.7571
## f1year_L1         1.1150 0.1476      0.8257  1.1150  1.4044
## f2year_L1         1.7678 0.1771      1.4207  1.7678  2.1149
## f3year_L1         1.7147 0.2230      1.2777  1.7147  2.1517
## drugDpenicil_L1   -0.1095 0.1145     -0.3339 -0.1095  0.1148
## f1year:drugDpenicil_L1 0.1051 0.2058     -0.2983  0.1051  0.5086
## f2year:drugDpenicil_L1 -0.3009 0.2501     -0.7911 -0.3009  0.1892
## f3year:drugDpenicil_L1 -0.2742 0.3138     -0.8892 -0.2742  0.3408
## Res. err. (variance) 0.0779 0.0021      0.0738  0.0779  0.0821
##
## Random effects variance-covariance (L1)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1      1.0373 0.1474      0.8306  1.0119  1.4039
## f1year_L1         1.5957 0.2643      1.1428  1.5671  2.2003
## f2year_L1         2.4456 0.5174      1.7036  2.3523  3.7494
## f3year_L1         1.7928 0.5537      0.9787  1.7049  3.0861
## Intercept_L1:f1year_L1 0.3195 0.1574      0.0159  0.3218  0.6333
## Intercept_L1:f2year_L1 0.6316 0.2472      0.2085  0.6044  1.2115
## Intercept_L1:f3year_L1 0.6308 0.3010      0.1735  0.5839  1.3813
## f1year_L1:f2year_L1  1.7652 0.2967      1.2806  1.7339  2.4232
## f1year_L1:f3year_L1  0.8869 0.3174      0.3334  0.8673  1.5624
## f2year_L1:f3year_L1  1.4199 0.5185      0.6505  1.3300  2.6320
##
## Longitudinal outcome (L2, lognormal)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L2      4.7541 0.0367      4.6821  4.7541  4.8261
## f1year_L2         -0.1349 0.0779     -0.2875 -0.1349  0.0178
## f2year_L2          0.0706 0.0937     -0.1131  0.0706  0.2542
## f3year_L2         -0.0227 0.1256     -0.2688 -0.0227  0.2235
## drugDpenicil_L2   -0.0856 0.0516     -0.1869 -0.0856  0.0156

```

```

## f1year:drugDpenicil_L2  0.1015 0.1083    -0.1108   0.1015    0.3138
## f2year:drugDpenicil_L2 -0.1897 0.1311    -0.4466  -0.1897    0.0673
## f3year:drugDpenicil_L2  0.0449 0.1737    -0.2956   0.0449    0.3853
## Res. err. (variance)    0.0682 0.0014     0.0655   0.0682    0.0709
##
## Random effects variance-covariance (L2)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L2      0.1671 0.0217     0.1328   0.1643   0.2187
## f1year_L2         0.2061 0.0463     0.1329   0.2010   0.3127
## f2year_L2         0.3160 0.0872     0.1967   0.3020   0.5217
## f3year_L2         0.2164 0.0739     0.1092   0.2047   0.3880
## Intercept_L2:f1year_L2 -0.0158 0.0237    -0.0659  -0.0145   0.0261
## Intercept_L2:f2year_L2  0.0193 0.0340    -0.0541   0.0212   0.0848
## Intercept_L2:f3year_L2 -0.0092 0.0373    -0.0885  -0.0086   0.0659
## f1year_L2:f2year_L2    0.1677 0.0402     0.0951   0.1654   0.2556
## f1year_L2:f3year_L2    0.0886 0.0406     0.0129   0.0865   0.1776
## f2year_L2:f3year_L2    0.1524 0.0707     0.0558   0.1397   0.3193
##
## Longitudinal outcome (L3, gaussian)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L3      3.5495 0.0327     3.4855   3.5495   3.6135
## year_L3           -0.0984 0.0121    -0.1221  -0.0984  -0.0748
## drugDpenicil_L3    -0.0002 0.0460    -0.0905  -0.0002   0.0900
## year:drugDpenicil_L3  0.0043 0.0169    -0.0289   0.0043   0.0374
## Res. err. (variance) 0.0963 0.0034     0.0897   0.0962   0.1032
##
## Random effects variance-covariance (L3)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L3      0.1262 0.0133     0.1024   0.1256   0.1542
## year_L3           0.0109 0.0014     0.0085   0.0109   0.0137
## Intercept_L3:year_L3 -0.0016 0.0024    -0.0066  -0.0015   0.0029
##
## Longitudinal outcome (L4, poisson)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L4      5.5204 0.0307     5.4602   5.5204   5.5807
## f1year_L4         -0.1885 0.1151    -0.4141  -0.1885   0.0371
## f2year_L4         -0.9484 0.2424    -1.4235  -0.9484  -0.4733
## f3year_L4         -1.2717 0.4480    -2.1499  -1.2717  -0.3935
## drugDpenicil_L4   -0.0585 0.0432    -0.1432  -0.0585   0.0262
## f1year:drugDpenicil_L4  0.1957 0.1624    -0.1225   0.1957   0.5139
## f2year:drugDpenicil_L4 -0.4363 0.3450    -1.1126  -0.4363   0.2399
## f3year:drugDpenicil_L4 -0.5519 0.6366    -1.7997  -0.5519   0.6959
##
## Random effects variance-covariance (L4)
##               mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L4      0.1461 0.0124     0.1237   0.1456   0.1710
## f1year_L4         1.3119 0.2214     0.9565   1.2868   1.8080
## f2year_L4         5.0368 1.1012     3.2088   4.9107   7.5142
## f3year_L4        16.3988 3.8762     9.8944  15.9810  25.1049
## Intercept_L4:f1year_L4 -0.0349 0.0367    -0.1091  -0.0347   0.0351
## Intercept_L4:f2year_L4 -0.0861 0.0774    -0.2429  -0.0855   0.0675
## Intercept_L4:f3year_L4 -0.0669 0.1454    -0.3635  -0.0665   0.2294
## f1year_L4:f2year_L4   -1.8336 0.4449    -2.8479  -1.7832  -1.0959
## f1year_L4:f3year_L4   -3.7078 0.8424    -5.5910  -3.6012  -2.3068

```



```

## f2year_L4:f3year_L4      8.6741 2.0495      5.2473  8.4709      13.3437
##
## Longitudinal outcome (L5, binomial)
##           mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L5      -1.3928 0.2393      -1.8619  -1.3928      -0.9237
## year_L5           0.1464 0.0599       0.0290   0.1464       0.2639
## drugDpenicil_L5    -0.1038 0.3398      -0.7697  -0.1038       0.5622
## year:drugDpenicil_L5 -0.0299 0.0850      -0.1965  -0.0299       0.1367
##
## Random effects variance-covariance (L5)
##           mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L5       5.9232 1.2328       4.0816   5.7595       8.8806
## year_L5            0.1363 0.0325       0.0853   0.1325       0.2107
## Intercept_L5:year_L5 -0.2553 0.1620      -0.6425  -0.2338      -0.0073
##
## Survival outcome (S1)
##           mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S1 0.0382 0.0184       0.0131   0.0348       0.0840
## drugDpenicil_S1      -0.0739 0.1922      -0.4506  -0.0739       0.3027
##
## Survival outcome (S2)
##           mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S2 0.0555 0.0315       0.0166   0.0483       0.1377
## drugDpenicil_S2      -0.4202 0.3813      -1.1675  -0.4202       0.3270
##
## Association longitudinal - survival
##           mean      sd 0.025quant 0.5quant 0.975quant
## CV_L1_S1  1.2177 0.0795       1.0544   1.2191       1.3713
## CS_L1_S1  0.8936 0.3041       0.3007   0.8879       1.5281
## CV_L1_S2  1.1539 0.0955       0.9568   1.1546       1.3459
## CV_L2_S1 -0.3455 0.1629      -0.6604  -0.3489      -0.0070
## CV_L3_S1 -1.8168 0.1777      -2.1713  -1.8165      -1.4640
## CV_L3_S2 -1.1405 0.2408      -1.6103  -1.1413      -0.6305
## CV_L4_S1 -0.6268 0.0765      -0.7890  -0.6262      -0.4743
## CV_L4_S2 -0.2741 0.1554      -0.6002  -0.2740       0.0305
## CV_L5_S1 -0.0053 0.0221      -0.0518  -0.0052       0.0390
##
## log marginal-likelihood (integration)      log marginal-likelihood (Gaussian)
##           -79593.38                        -79547.15
##
## Deviance Information Criterion: -65911.44
## Widely applicable Bayesian information criterion: -59308.9
## Computation time: 276.21 seconds

```