INLAjoint

Denis Rustand

2024-03-21

Contents

In this vignette we show how to fit various models with the *joint()* function of the `INLAjoint` package. The package is providing a unified framework to fit various longitudinal, survival and joint models. This vignette illustrates some of these models and thus takes a long time to compute which is why the vignette is not compiled when installing INLAjoint but instead we provide a pre-compiled pdf.

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 contain 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:

$$\log(serBilir_{ij}) = \beta_0 + b_{i0} + \beta_1 year_{ij} + \beta_2 drug_i + \varepsilon_{ij} \qquad \text{(L1)}$$

where β are the fixed effects, b_{i0} is an individual random intercept and ε_{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  0.0886     0.4566   0.6304     0.8043
## year                   0.0843  0.0042     0.0760   0.0843     0.0927
## drugDpenicil          -0.1223  0.1241    -0.3658  -0.1223     0.1212
## Res. err. (variance)   0.2153  0.0076     0.2007   0.2152     0.2308
##
## Random effects variance-covariance (L1)
##             mean      sd 0.025quant 0.5quant 0.975quant
## Intercept 1.1461  0.0966     0.9692   1.1417     1.3487
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                             -2802.283                             -2802.283
##
## Deviance Information Criterion:  4805.105
## Widely applicable Bayesian information criterion:  4806.402
## Computation time: 4.21 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  0.0886     0.4566   0.6304     0.8043
## year             0.0843  0.0042     0.0760   0.0843     0.0927
## drugDpenicil    -0.1223  0.1241    -0.3658  -0.1223     0.1212
## Res. err. (sd)   0.4640  0.0082     0.4480   0.4639     0.4804
##
```

```
## Random effects standard deviation / correlation (L1)
##            mean    sd 0.025quant 0.5quant 0.975quant
## Intercept 1.0696 0.045     0.9845   1.0685     1.1614
##
## log marginal-likelihood (integration)     log marginal-likelihood (Gaussian)
##                             -2802.283                              -2802.283
##
## Deviance Information Criterion:  4805.105
## Widely applicable Bayesian information criterion:  4806.402
## Computation time: 4.21 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 `inla.priors.used`, applied to an object fitted with the `joint` function. For example the default priors for fixed effects are Gaussian with mean zero and variance 100 (i.e., precision 0.01).

```
inla.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]
##       group.theta1:
##           parameter=[logit correlation]
##           prior=[normal]
##           param=[0.0, 0.2]
```

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:

$$
\begin{cases}
\log(serBilir_{ij}) & = \beta_{10} + b_{i10} + (\beta_{11} + b_{i11})year_{ij} + \beta_{12}drug_i + \beta_{13}sex_i + \beta_{14}year_{ij}drug_i + \varepsilon_{ij1} & \text{(L1)} \\
\log(E[platelets_{ij}]) & = \beta_{20} + b_{i20} + (\beta_{21} + b_{i21})year_{ij} + \beta_{22}sex_i + \beta_{23}drug_i + \beta_{24}year_{ij}sex_i & \text{(L2)} \\
\text{logit}(E[spiders_{ij}]) & = \beta_{30} + b_{i30} + (\beta_{31} + b_{i31})year_{ij} + \beta_{32}drug_i & \text{(L3)}
\end{cases}
$$

```r
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           0.9153 0.1662     0.5895   0.9153     1.2411
## year_L1                0.1462 0.0185     0.1099   0.1462     0.1825
## drugDpenicil_L1       -0.1624 0.1149    -0.3876  -0.1624     0.0628
## sexfemale_L1          -0.3695 0.1603    -0.6838  -0.3695    -0.0553
## year:drugDpenicil_L1  -0.0008 0.0254    -0.0505  -0.0008     0.0489
## Res. err. (variance)   0.1082 0.0104     0.0887   0.1080     0.1293
##
## Longitudinal outcome (L2, poisson)
##                         mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L2           5.4439 0.0682     5.3103   5.4439     5.5776
## year_L2               -0.1248 0.0326    -0.1887  -0.1248    -0.0610
## sexfemale_L2           0.0934 0.0671    -0.0382   0.0934     0.2250
## drugDpenicil_L2       -0.0732 0.0440    -0.1593  -0.0732     0.0130
## year:sexfemale_L2      0.0460 0.0346    -0.0217   0.0460     0.1137
##
## Longitudinal outcome (L3, binomial)
##                        mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L3         -1.6874 0.3087    -2.2924  -1.6874    -1.0824
## year_L3               0.1698 0.0698     0.0331   0.1698     0.3065
## drugDpenicil_L3      -0.2558 0.4248    -1.0885  -0.2558     0.5768
```

```
## 
## Random effects variance-covariance
##                              mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1               1.4521  1.5212     0.7707   1.1265     4.2551
## year_L1                    0.0548  0.0594     0.0281   0.0431     0.1540
## Intercept_L2               0.1702  0.0640     0.1247   0.1574     0.2942
## year_L2                    0.0490  0.0442     0.0200   0.0379     0.1441
## Intercept_L3              13.7940 11.3589     5.9127  11.1912    38.1366
## year_L3                    0.5498  0.5604     0.0546   0.3844     2.0162
## Intercept_L1:year_L1       0.0833  0.1622    -0.0998   0.0767     0.3073
## Intercept_L1:Intercept_L2 -0.0735  0.2399    -0.2769  -0.0892     0.2715
## Intercept_L1:year_L2      -0.0660  0.2247    -0.4812  -0.0249     0.0705
## Intercept_L1:Intercept_L3  1.6030  2.2694    -1.3600   1.5649     4.2617
## Intercept_L1:year_L3      -0.0010  0.6541    -1.3538   0.0622     0.8738
## year_L1:Intercept_L2       0.0009  0.0469    -0.0584  -0.0052     0.1091
## year_L1:year_L2           -0.0028  0.0315    -0.0534  -0.0024     0.0470
## year_L1:Intercept_L3       0.2171  0.4894    -0.2191   0.1482     1.0272
## year_L1:year_L3            0.0301  0.1226    -0.2396   0.0324     0.2413
## Intercept_L2:year_L2      -0.0115  0.0388    -0.0891  -0.0060     0.0234
## Intercept_L2:Intercept_L3 -0.4102  0.6709    -1.6299  -0.2986     0.0679
## Intercept_L2:year_L3       0.0352  0.1608    -0.0936  -0.0056     0.4096
## year_L2:Intercept_L3       0.1025  0.4764    -0.3372   0.0402     1.1304
## year_L2:year_L3           -0.0177  0.1259    -0.2693  -0.0108     0.2004
## Intercept_L3:year_L3      -0.6515  2.4366    -6.9137   0.0357     1.5450
## 
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                      -17073.81                                    -17053.31
## 
## Deviance Information Criterion:  26397.27
## Widely applicable Bayesian information criterion:  45563.45
## Computation time: 11.09 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:

```r
summary(M2, sdcor=TRUE)
```

```
## Longitudinal outcome (L1, lognormal)
##                      mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L1       0.9153 0.1662     0.5895   0.9153     1.2411
## year_L1            0.1462 0.0185     0.1099   0.1462     0.1825
## drugDpenicil_L1   -0.1624 0.1149    -0.3876  -0.1624     0.0628
## sexfemale_L1      -0.3695 0.1603    -0.6838  -0.3695    -0.0553
## year:drugDpenicil_L1 -0.0008 0.0254    -0.0505  -0.0008     0.0489
## Res. err. (sd)     0.3285 0.0157     0.2978   0.3286     0.3596
## 
## Longitudinal outcome (L2, poisson)
##                   mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L2    5.4439 0.0682     5.3103   5.4439     5.5776
## year_L2        -0.1248 0.0326    -0.1887  -0.1248    -0.0610
## sexfemale_L2    0.0934 0.0671    -0.0382   0.0934     0.2250
## drugDpenicil_L2 -0.0732 0.0440    -0.1593  -0.0732     0.0130
```

```
## year:sexfemale_L2  0.0460 0.0346    -0.0217   0.0460       0.1137
##
## Longitudinal outcome (L3, binomial)
##                    mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L3    -1.6874 0.3087    -2.2924  -1.6874    -1.0824
## year_L3          0.1698 0.0698     0.0331   0.1698     0.3065
## drugDpenicil_L3 -0.2558 0.4248    -1.0885  -0.2558     0.5768
##
## Random effects standard deviation / correlation
##                               mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L1                1.1418 0.3268     0.8851   1.0564     1.8975
## year_L1                     0.2252 0.0690     0.1653   0.2077     0.4038
## Intercept_L2                0.4083 0.0565     0.3510   0.3959     0.5642
## year_L2                     0.2095 0.0643     0.1373   0.1938     0.3721
## Intercept_L3                3.5866 1.1208     2.4336   3.3448     6.5432
## year_L3                     0.6812 0.3266     0.2377   0.6225     1.5373
## Intercept_L1:year_L1        0.3232 0.2677    -0.3272   0.3670     0.7230
## Intercept_L1:Intercept_L2 -0.2009 0.2298    -0.5835  -0.2167     0.3236
## Intercept_L1:year_L2       -0.1507 0.2564    -0.6972  -0.1351     0.3070
## Intercept_L1:Intercept_L3  0.4003 0.2344    -0.2771   0.4670     0.6613
## Intercept_L1:year_L3        0.1253 0.3890    -0.6711   0.1620     0.7563
## year_L1:Intercept_L2       -0.0741 0.3354    -0.6615  -0.0991     0.6318
## year_L1:year_L2            -0.0502 0.2781    -0.5840  -0.0686     0.6137
## year_L1:Intercept_L3        0.2160 0.2308    -0.3134   0.2189     0.6941
## year_L1:year_L3             0.2570 0.4318    -0.7533   0.3487     0.8449
## Intercept_L2:year_L2       -0.0956 0.1905    -0.5527  -0.0753     0.2344
## Intercept_L2:Intercept_L3 -0.2441 0.1608    -0.6613  -0.2329     0.0513
## Intercept_L2:year_L3        0.0326 0.2417    -0.3180  -0.0287     0.6826
## year_L2:Intercept_L3        0.1159 0.2514    -0.3043   0.0700     0.7591
## year_L2:year_L3            -0.1771 0.4237    -0.8929  -0.1829     0.6624
## Intercept_L3:year_L3       -0.0698 0.4074    -0.8900   0.0216     0.5212
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                       -17073.81                                   -17053.31
##
## Deviance Information Criterion:  26397.27
## Widely applicable Bayesian information criterion:  45563.45
## Computation time: 11.09 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     0.9168 0.1642     0.5949   0.9168     1.2387
## year_L1          0.1579 0.0183     0.1220   0.1579     0.1939
```

```
## drugDpenicil_L1        -0.1587 0.1133     -0.3808  -0.1587      0.0634
## sexfemale_L1           -0.3718 0.1583     -0.6820  -0.3718     -0.0616
## year:drugDpenicil_L1    0.0004 0.0251     -0.0488   0.0004      0.0495
## Res. err. (variance)    0.1093 0.0060      0.0981   0.1091      0.1217
##
## Longitudinal outcome (L2, poisson)
##                    mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L2       5.4429 0.0682      5.3093   5.4429      5.5765
## year_L2           -0.1266 0.0323     -0.1899  -0.1266     -0.0633
## sexfemale_L2       0.0920 0.0671     -0.0395   0.0920      0.2235
## drugDpenicil_L2   -0.0756 0.0440     -0.1619  -0.0756      0.0106
## year:sexfemale_L2  0.0473 0.0343     -0.0198   0.0473      0.1144
##
## Longitudinal outcome (L3, binomial)
##                    mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L3      -0.8150 0.1704     -1.1490  -0.8150     -0.4809
## year_L3            0.1929 0.0473      0.1002   0.1929      0.2855
## drugDpenicil_L3   -0.0665 0.2385     -0.5339  -0.0665      0.4008
##
## Random effects variance-covariance
##                           mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1             1.0914 0.1635      0.8715   1.0638      1.4753
## year_L1                  0.0358 0.0075      0.0259   0.0347      0.0500
## Intercept_L2             0.1513 0.0134      0.1264   0.1509      0.1782
## year_L2                  0.0335 0.0053      0.0248   0.0331      0.0458
## Intercept_L3             3.6712 0.9895      2.2054   3.5345      6.0787
## year_L3                  0.2771 0.1549      0.0787   0.2458      0.6636
## Intercept_L1:year_L1     0.0518 0.0357     -0.0081   0.0485      0.1324
## Intercept_L1:Intercept_L2 -0.0742 0.0364    -0.1413  -0.0747     -0.0016
## Intercept_L1:year_L2    -0.0173 0.0243     -0.0694  -0.0164      0.0278
## Intercept_L1:Intercept_L3  1.0064 0.2935    0.5210   0.9724      1.6469
## Intercept_L1:year_L3     0.1723 0.1613     -0.0302   0.1374      0.5728
## year_L1:Intercept_L2    -0.0047 0.0080     -0.0197  -0.0046      0.0103
## year_L1:year_L2         -0.0023 0.0064     -0.0132  -0.0027      0.0113
## year_L1:Intercept_L3     0.0646 0.0587     -0.0264   0.0589      0.1987
## year_L1:year_L3          0.0420 0.0294     -0.0148   0.0405      0.0986
## Intercept_L2:year_L2    -0.0040 0.0073     -0.0184  -0.0038      0.0099
## Intercept_L2:Intercept_L3 -0.1865 0.0695    -0.3283  -0.1845     -0.0588
## Intercept_L2:year_L3     0.0010 0.0252     -0.0346  -0.0040      0.0632
## year_L2:Intercept_L3     0.0099 0.0562     -0.0979   0.0075      0.1313
## year_L2:year_L3         -0.0210 0.0154     -0.0565  -0.0193      0.0071
## Intercept_L3:year_L3     0.0262 0.2552     -0.5367   0.0307      0.5337
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                         -17066.98                                 -17046.48
##
## Deviance Information Criterion:  NA
## Widely applicable Bayesian information criterion:  45281.56
## Computation time: 70.59 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 contain 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:

$$\lambda_i(t) = \lambda_0(t) \, \exp\left(\gamma_1 drug_i + \gamma_2 sex_i + \gamma_3 drug_i sex_i\right)$$

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

```
##
## Survival outcome
##                            mean     sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)  0.0144 0.0327     0.0001   0.0040     0.0969
## drugDpenicil              0.1692 0.3912    -0.5979   0.1692     0.9364
## sexfemale                -0.5340 0.3223    -1.1661  -0.5340     0.0980
## drugDpenicil:sexfemale   -0.2238 0.4335    -1.0739  -0.2238     0.6263
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                              -664.575                              -664.575
##
## Deviance Information Criterion:  1297.789
## Widely applicable Bayesian information criterion:  1297.833
## Computation time: 3.59 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 0.05904699 0.1078777 0.1962539
##  [2,]  0.9537108 0.05864166 0.1066659 0.1930482
##  [3,]  1.9074216 0.06090024 0.1098930 0.1981842
##  [4,]  2.8611324 0.06080306 0.1096956 0.1976802
##  [5,]  3.8148432 0.06088748 0.1099088 0.1981483
##  [6,]  4.7685540 0.06066377 0.1097178 0.1980306
##  [7,]  5.7222648 0.06092226 0.1103300 0.1994620
##  [8,]  6.6759756 0.06193825 0.1122684 0.2036606
##  [9,]  7.6296864 0.06296842 0.1144996 0.2091466
## [10,]  8.5833972 0.06387376 0.1168182 0.2158228
## [11,]  9.5371080 0.06364651 0.1171024 0.2182412
## [12,] 10.4908188 0.06299864 0.1167459 0.2195046
```

```
## [13,] 11.4445296 0.06183079 0.1156758 0.2193106
## [14,] 12.3982404 0.06082085 0.1151637 0.2210703
## [15,] 13.3519512 0.06015912 0.1153870 0.2253772
## [16,] 14.3056620 0.05930220 0.1153515 0.2292188
```

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)    0.0144 0.0327     0.0001   0.0040     0.0969
## drugDpenicil                1.2754 0.5056     0.5546   1.1836     2.5244
## sexfemale                   0.6165 0.1995     0.3138   0.5859     1.0935
## drugDpenicil:sexfemale      0.8755 0.3871     0.3447   0.7988     1.8493
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                           -664.575                                 -664.575
##
## Deviance Information Criterion:  1297.789
## Widely applicable Bayesian information criterion:  1297.833
## Computation time: 3.59 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 = inla.surv(time = years, event = death) ~ drug * sex,
            dataSurv = Surv, basRisk = "weibullsurv")
summary(M3_wei, hazr=T)
```

```
##
## Survival outcome
##                          exp(mean)     sd 0.025quant 0.5quant 0.975quant
## Weibull (shape)             1.0621 0.0946     0.8994   1.0544     1.2707
## Weibull (scale)             0.1026 0.0337     0.0513   0.0976     0.1829
## drugDpenicil                1.3228 0.5307     0.5694   1.2254     2.6363
## sexfemale                   0.6363 0.2094     0.3201   0.6037     1.1383
## drugDpenicil:sexfemale      0.8467 0.3779     0.3303   0.7713     1.7990
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                           -520.6879                               -520.6879
```

```
##
## Deviance Information Criterion:  1024.083
## Widely applicable Bayesian information criterion:  1024.427
## Computation time: 1.87 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:

$$
\begin{cases}
log(serBilir_{ij}) = \eta_i(t_{ij}) + \varepsilon_{ij} & \text{(L1)} \\
\quad = \beta_0 + \beta_1 year_{ij} + (\beta_2 + b_{i2})year_{ij}^2 \\
\quad + (\beta_3 + b_{i3})year_{ij}^3 + \beta_4 drug_i + \beta_5 year_{ij}drug_i \\
\quad + \beta_6 year_{ij}^2 drug_i + \beta_7 year_{ij}^3 drug_i + \varepsilon_{ij} \\
\lambda_{i1}(t) \quad = \lambda_{01}(t)\exp\left(\gamma_1 drug_i + \varphi_1\eta_i(t) + \varphi_2\eta_i'(t)\right) & \text{(S1)}
\end{cases}
$$

where γ denotes fixed effects of the survival part and φ the association parameters.

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


M4 <- joint(formSurv = inla.surv(years, death) ~ drug,
            formLong = serBilir ~ (1 + year + f1(year) + f2(year))*drug +
                       (f1(year) + f2(year) |id), family = "lognormal",
            dataLong = Longi, dataSurv=Surv, 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            0.6331 0.0529     0.5295   0.6331     0.7367
## year_L1                -0.1097 0.0674    -0.2419  -0.1097     0.0225
## f1year_L1               0.0374 0.0278    -0.0170   0.0374     0.0919
## f2year_L1               0.0210 0.0074     0.0065   0.0210     0.0354
## drugDpenicil_L1        -0.1337 0.0744    -0.2795  -0.1337     0.0120
## year:drugDpenicil_L1   -0.0073 0.0945    -0.1926  -0.0073     0.1779
## f1year:drugDpenicil_L1  0.0162 0.0392    -0.0606   0.0162     0.0930
## f2year:drugDpenicil_L1 -0.0037 0.0103    -0.0239  -0.0037     0.0165
## Res. err. (variance)    0.6429 0.0252     0.5958   0.6420     0.6948
##
## Random effects variance-covariance (L1)
##                       mean     sd 0.025quant 0.5quant 0.975quant
## f1year_L1           0.0226 0.0042     0.0155   0.0223     0.0315
## f2year_L1           0.0056 0.0006     0.0046   0.0056     0.0068
## f1year_L1:f2year_L1 -0.0006 0.0003    -0.0012  -0.0006     0.0000
##
## Survival outcome
```

```
##                                mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S1   0.2780 0.0968      0.1437   0.2590     0.5205
## drugDpenicil_S1              -0.1779 0.2337     -0.6360  -0.1779     0.2802
##
## Association longitudinal - survival
##            mean      sd 0.025quant 0.5quant 0.975quant
## CV_L1_S1 1.0506 0.1143     0.8068   1.0567     1.2544
## CS_L1_S1 0.6795 0.1659     0.3550   0.6789     1.0081
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                          -33379.43                               -33372.71
##
## Deviance Information Criterion:  -51084.93
## Widely applicable Bayesian information criterion:  -52226.49
## Computation time: 18.19 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 f1, f2, ..., fN, 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
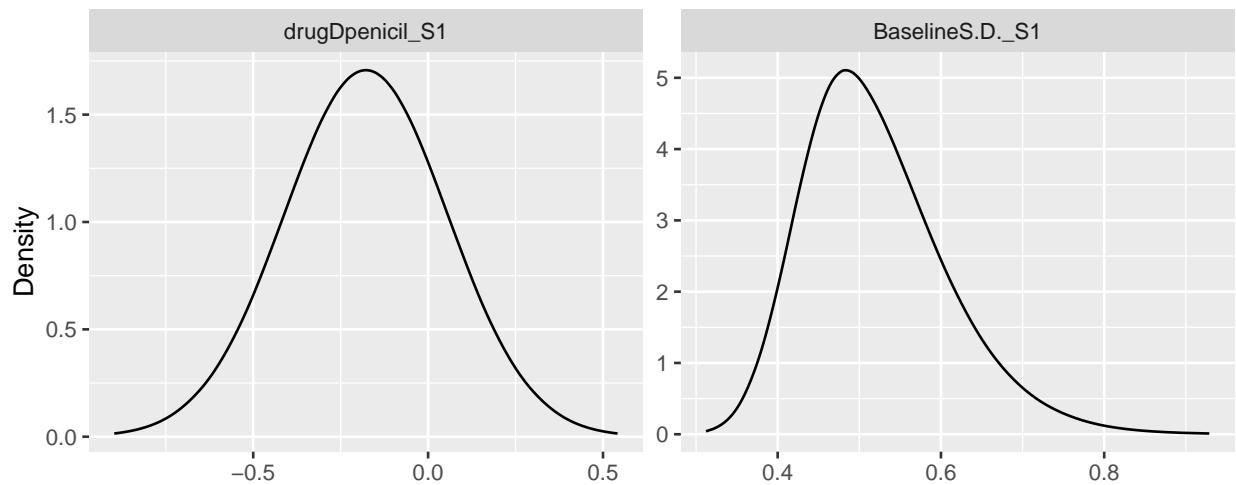
```
plotM4 <- plot(M4, sdcor=T)
```

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

```
plotM4$Outcomes$L1
```
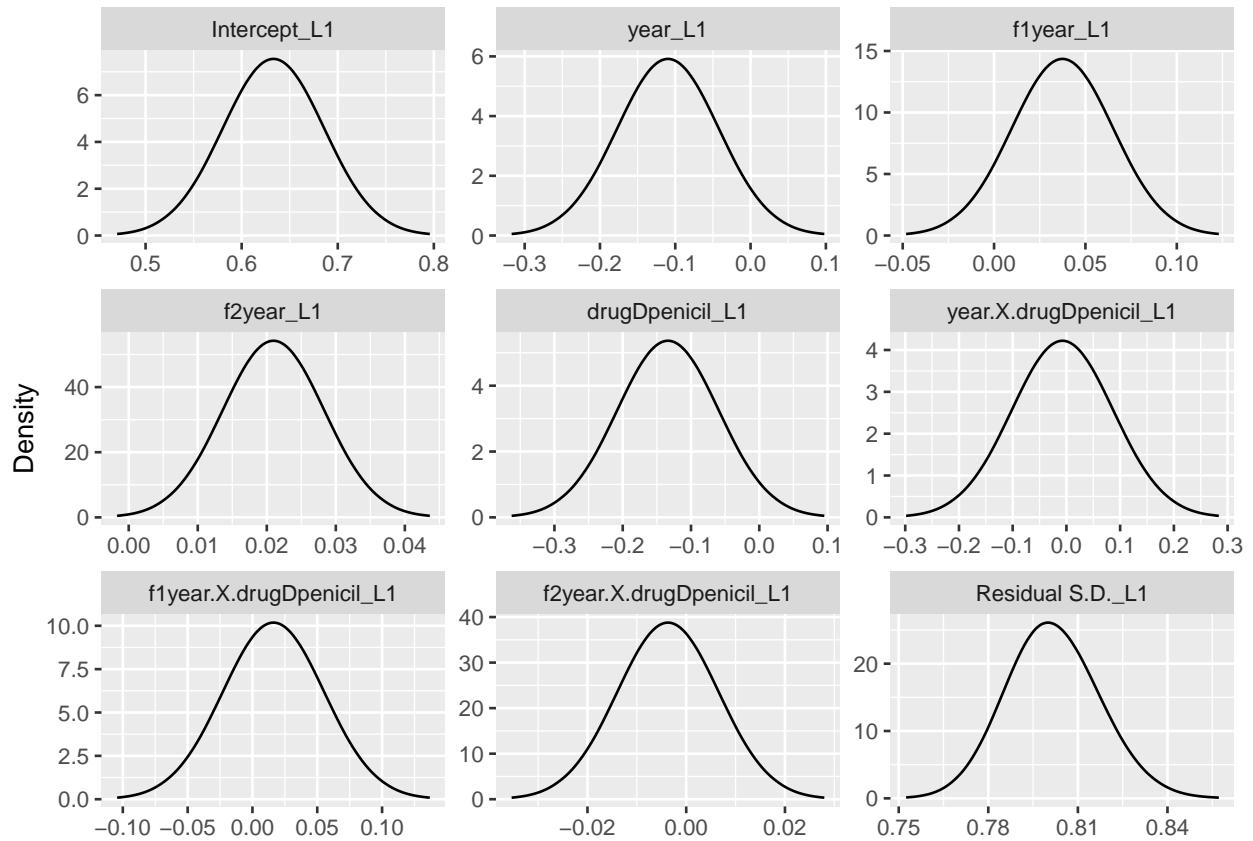
Then the parameters of the survival outcome(s):

```
plotM4$Outcomes$S1
```
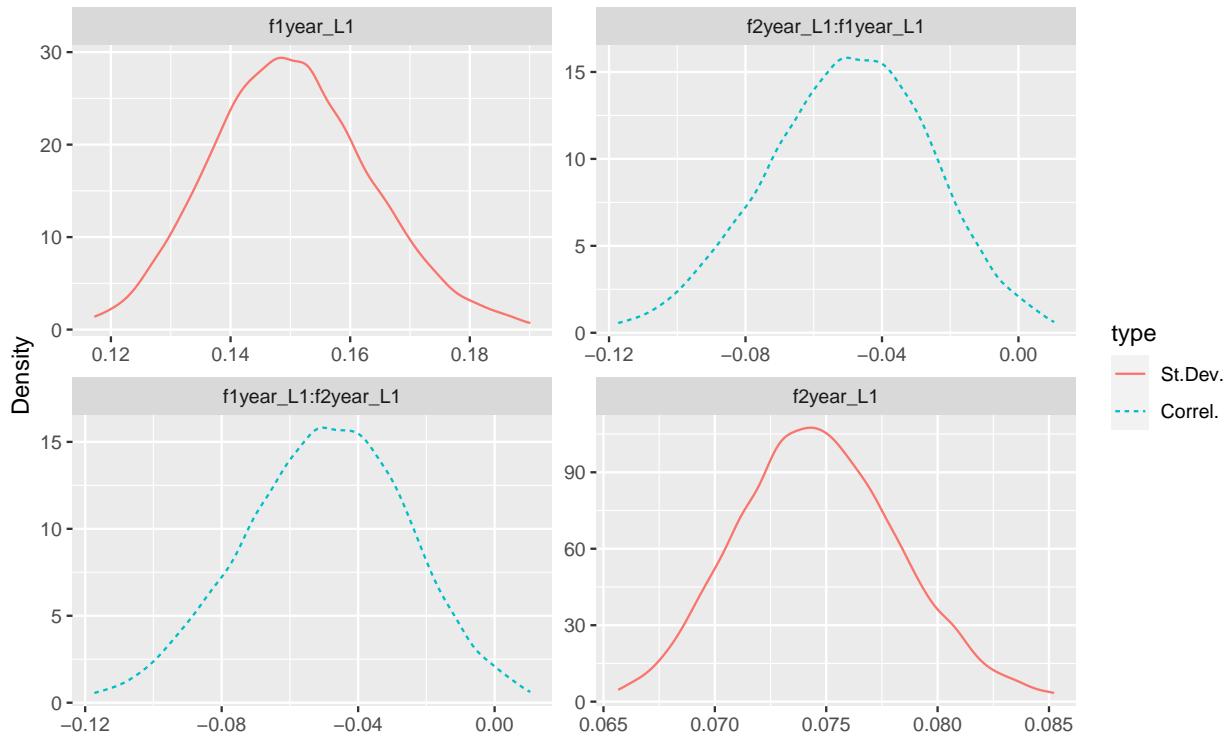

The variance-covariance of the random-effects (converted to standard deviations and correlations when argument `sdcor=TRUE` is added to the call of the plot function):
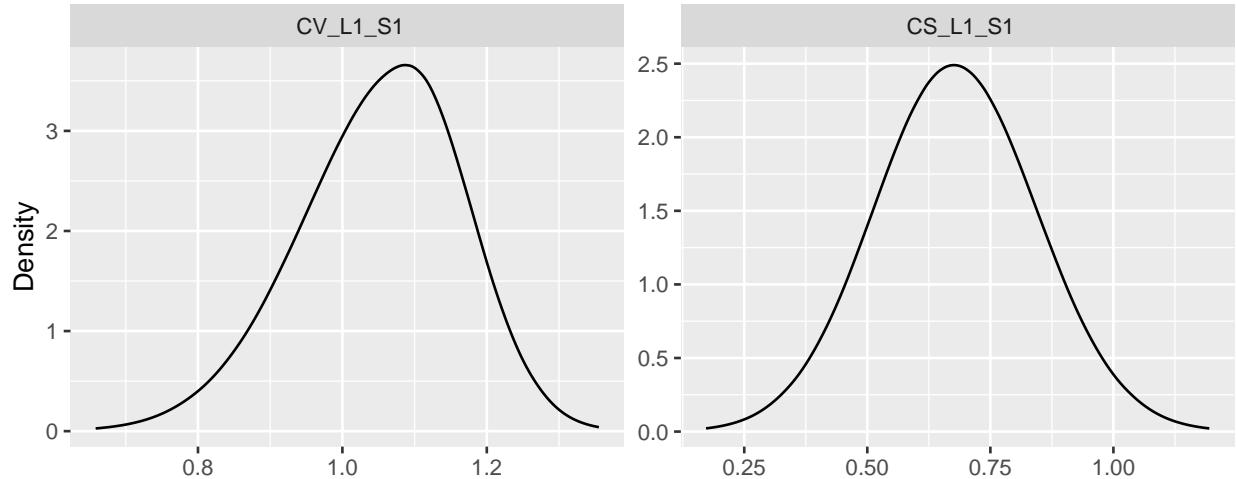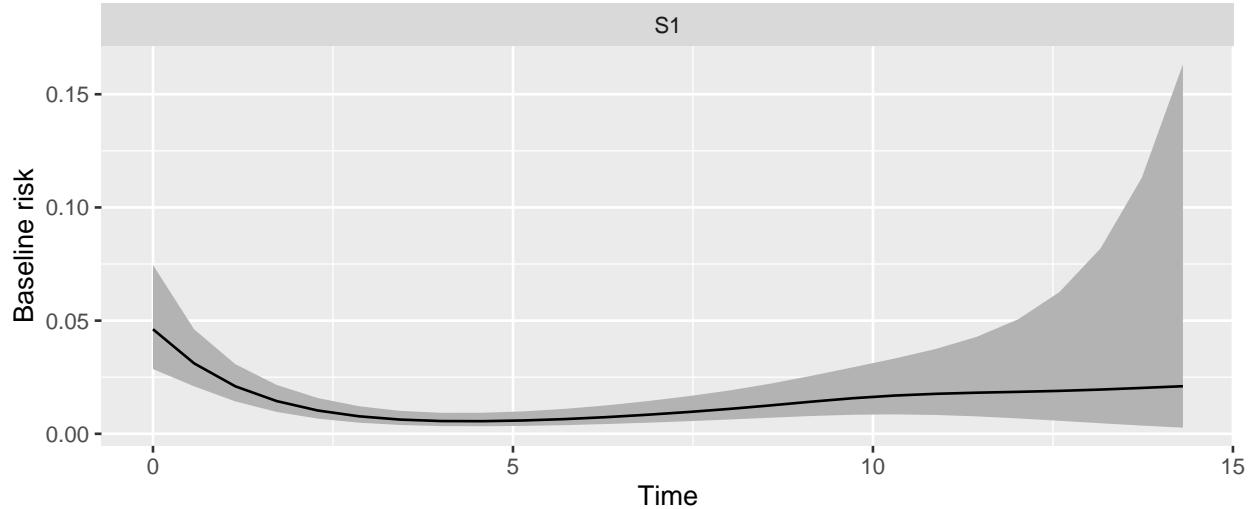
```
plotM4$Covariances
```

`## $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()
```
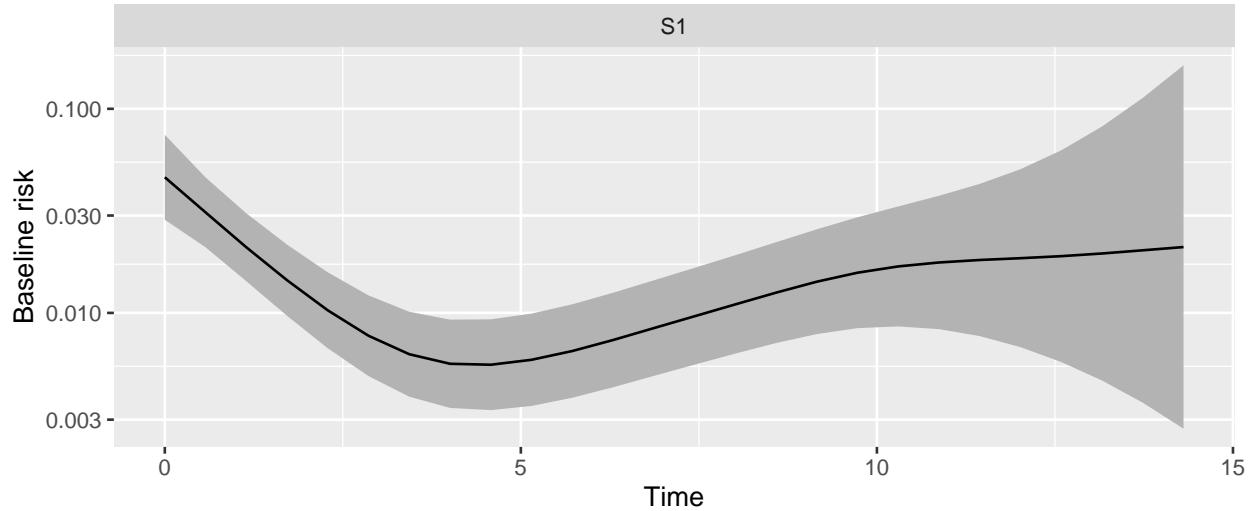

Model 5: Comparison with MCMC

We can make a comparison of `INLAjoint` with Bayesian estimations with MCMC implemented in alternative R packages such as `JMBayes2` (C++ implementation of MCMC) or `rstanarm` (Stan).

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

$$
\begin{cases}
albumin_{ij} = \eta_i(t_{ij}) + \varepsilon_{ij} & \text{(L1)} \\
\quad\quad = \beta_0 + b_{i0} + (\beta_1 + b_{i1})year_{ij} + \beta_2 drug_i + \beta_3 year_{ij} drug_i + \varepsilon_{ij} \\
\lambda_i(t) \quad = \lambda_0(t)\exp\left(\gamma_1 sex_i + \gamma_2 drug_i + \varphi\eta_i(t)\right) & \text{(S1)}
\end{cases}
$$

```r
# INLAjoint
M5 <- joint(formSurv = inla.surv(years, death) ~ sex + drug,
            formLong = albumin ~ (1 + year)*drug + (1 + year |id),
            dataLong = Longi, dataSurv=Surv, 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            3.5440 0.0334     3.4785   3.5440     3.6094
## year_L1                -0.1002 0.0120    -0.1237  -0.1002    -0.0768
## drugDpenicil_L1         0.0085 0.0469    -0.0835   0.0086     0.1006
## year:drugDpenicil_L1    0.0023 0.0166    -0.0302   0.0023     0.0347
## Res. err. (variance)    0.0956 0.0037     0.0885   0.0956     0.1029
##
## Random effects variance-covariance (L1)
##                           mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1            0.1387 0.0209     0.1053   0.1365     0.1864
## year_L1                 0.0115 0.0014     0.0090   0.0114     0.0144
## Intercept_L1:year_L1   -0.0024 0.0054    -0.0142  -0.0019     0.0074
##
## Survival outcome
##                              mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S1 0.2874 1.3466    0.0000   0.0137     2.3883
## sexfemale_S1               -0.7118 0.2507   -1.2035  -0.7117    -0.2201
## drugDpenicil_S1             0.0090 0.1934   -0.3702   0.0090     0.3882
##
## Association longitudinal - survival
##              mean     sd 0.025quant 0.5quant 0.975quant
## CV_L1_S1 -3.0117 0.4515    -3.8478   -3.028     -2.072
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                          -10406.06                            -10406.06
##
## Deviance Information Criterion:  -16205.89
## Widely applicable Bayesian information criterion:  -15962.15
## Computation time: 6.84 seconds
```

```r
# JMBayes2
library(JMbayes2)
M5JMB2_lme <- lme(albumin ~ (1 + year)*drug,
                  random = ~ 1 + year |id, data = Longi)
M5JMB2_cox <- coxph(Surv(Surv$years, Surv$death) ~ sex + drug,
                    data = Surv, x = TRUE)
JMpr = list(mean_alphas=list(0), Tau_alphas = list(matrix(0.16)))
M5JMB2 <- try(jm(M5JMB2_cox, list(M5JMB2_lme), time_var = "year", priors=JMpr))
# Computation time in the table includes LME + Cox + JM
summary(M5JMB2)
```

```
## 
## Call:
## jm(Surv_object = M5JMB2_cox, Mixed_objects = list(M5JMB2_lme),
##     time_var = "year", priors = JMpr)
## 
## Data Descriptives:
## Number of Groups: 312        Number of events: 140 (44.9%)
## Number of Observations:
##   albumin: 1866
## 
##               DIC      WAIC       LPML
## marginal    2653.459 2642.044 -1323.7902
## conditional 1619.250 1452.281  -948.2341
## 
## Random-effects covariance matrix:
## 
##        StdDev   Corr
## (Intr) 0.3600 (Intr)
## year   0.0489 0.0080
## 
## Survival Outcome:
##                  Mean  StDev    2.5%    97.5%      P   Rhat
## sexfemale      -0.8018 0.3296 -1.4253 -0.1461 0.0140 1.0287
## drugD-penicil   0.0008 0.2050 -0.3991  0.4121 0.9969 1.0097
## value(albumin) -3.7520 0.3421 -4.4894 -3.1107 0.0000 1.0456
## 
## Longitudinal Outcome: albumin (family = gaussian, link = identity)
##                    Mean  StDev    2.5%    97.5%      P   Rhat
## (Intercept)       3.5324 0.0332  3.4665  3.5970 0.0000 1.0034
## year             -0.0879 0.0078 -0.1034 -0.0728 0.0000 1.0223
## drugD-penicil     0.0127 0.0462 -0.0797  0.1007 0.7793 0.9998
## year:drugD-penicil -0.0017 0.0095 -0.0209  0.0166 0.8602 1.0030
## sigma             0.3150 0.0059  0.3039  0.3267 0.0000 1.0008
## 
## MCMC summary:
## chains: 3
## iterations per chain: 3500
## burn-in per chain: 500
## thinning: 1
## time: 28.8 min
```

```r
# 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)        3.534  0.034
## year              -0.089  0.008
## drugD-penicil      0.011  0.048
## year:drugD-penicil -0.002  0.010
## sigma              0.314  0.006
##
## Event submodel:
##                   Median  MAD_SD   exp(Median)
## (Intercept)       10.232   1.273 27778.574
## sexfemale         -0.782   0.265     0.458
## drugD-penicil     -0.005   0.195     0.995
## Long1|etavalue    -3.692   0.370     0.025
## b-splines-coef1   -0.166   0.482        NA
## b-splines-coef2   -0.519   0.491        NA
## b-splines-coef3    0.116   0.426        NA
## b-splines-coef4   -2.830   0.798        NA
## b-splines-coef5    0.663   1.165        NA
## b-splines-coef6   -3.385   1.588        NA
##
## Group-level error terms:
##  Groups Name            Std.Dev. Corr
##  id     Long1|(Intercept) 0.36448
##         Long1|year        0.05082  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	JMbayes2	rstanarm
algorithm	INLA	MCMC	MCMC
comp. time	7 sec.	1736 sec.	596 sec.

A more detailed comparison between INLA, 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:

$$
\begin{cases}
\log(serBilir_{ij}) = \eta_i(t_{ij}) + \varepsilon_{ij} & \text{(L1)} \\
\qquad\qquad = \beta_0 + b_{i0} + (\beta_1 + b_{i1})year_{ij} + \beta_2 drug_i + \beta_3 sex_i \\
\qquad\qquad\quad + \beta_4 year_{ij}drug_i + \beta_5 year_{ij}sex_i + \varepsilon_{ij} \\
\lambda_{i1}^{death}(t) \quad = \lambda_{01}(t)\exp\left(\gamma_{11}sex_i + \gamma_{12}drug_i + \varphi_{11}(b_{i0} + b_{i1}t)\right) & \text{(S1)} \\
\lambda_{i2}^{transpl.}(t) = \lambda_{02}(t)\exp\left(\gamma_{21}edema\_no_i + \gamma_{22}edema\_de_i + \gamma_{23}sex_i \right. \\
\qquad\qquad\quad \left. + \gamma_{24}edema\_no_isex_i + \gamma_{25}edema\_de_isex_i + \varphi_{21}b_{i0} + \varphi_{22}b_{i1}\right) & \text{(S2)}
\end{cases}
$$

```
M6 <- joint(formLong = serBilir ~ year * (drug + sex) + (1+year|id), dataLong = Longi,
         formSurv = list(inla.surv(years, death) ~ sex + drug,
                       inla.surv(years, trans) ~ edema * sex),
         id = "id", timeVar = "year", family = "lognormal", basRisk = c("rw1", "rw1"),
         assoc = c("SRE", "SRE_ind"), dataSurv=Surv, control=list(int.strategy="eb"))
summary(M6)
```

```
## Longitudinal outcome (lognormal)
##                        mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1          0.8067 0.1773    0.4593   0.8067    1.1541
## year_L1               0.1843 0.0369    0.1119   0.1843    0.2566
## drugDpenicil_L1      -0.1281 0.1080   -0.3397  -0.1281    0.0835
## sexfemale_L1         -0.2855 0.1767   -0.6318  -0.2855    0.0609
## year:drugDpenicil_L1  0.0106 0.0232   -0.0349   0.0106    0.0562
## year:sexfemale_L1    -0.0358 0.0361   -0.1066  -0.0358    0.0350
## Res. err. (variance)  0.1090 0.0042    0.1009   0.1089    0.1176
##
## Random effects variance-covariance (L1)
##                       mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1          1.0009 0.0887    0.8457   0.9955    1.1880
## year_L1               0.0348 0.0046    0.0268   0.0346    0.0449
## Intercept_L1:year_L1  0.0548 0.0200    0.0164   0.0543    0.0956
##
## Survival outcome (S1)
##                            mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S1  0.1753 0.0830    0.0640   0.1587    0.3852
## sexfemale_S1                -0.8188 0.3459   -1.4967  -0.8188   -0.1409
## drugDpenicil_S1             -0.1034 0.2417   -0.5772  -0.1034    0.3703
##
## Survival outcome (S2)
##                                        mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S2             0.4062 0.3703    0.0576   0.2961    1.4305
## edemaedema.no.diuretics_S2             -0.5474 1.2258   -2.9499  -0.5474    1.8551
## edemaedema.despite.diuretics_S2         0.4170 1.2398   -2.0130   0.4170    2.8470
## sexfemale_S2                            0.1130 0.8046   -1.4640   0.1130    1.6901
## edemaedema.no.diuretics:sexfemale_S2    0.8182 1.2977   -1.7253   0.8182    3.3617
## edemaedema.despite.diuretics:sexfemale_S2 -1.4554 1.3900  -4.1797  -1.4554    1.2690
##
## Association longitudinal - survival
##                        mean      sd 0.025quant 0.5quant 0.975quant
```

```
## SRE_L1_S1              1.1381 0.0858    0.9661   1.1392      1.3039
## SRE_Intercept_L1_S2 1.0795 0.2481    0.5798   1.0833      1.5565
## SRE_year_L1_S2          1.6757 0.8844   -0.0224   1.6613      3.4592
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                              -12088.09                          -12079.53
##
## Deviance Information Criterion:  -13254.8
## Widely applicable Bayesian information criterion:  -13898.78
## Computation time: 14.65 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:

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

```r
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(inla.surv(years, death) ~ sex + drug,
                            inla.surv(years, trans) ~ edema * sex),
        dataLong = Longi, dataSurv=Surv, 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           1.0056 0.1693     0.6739   1.0056      1.3374
## year_L1                0.0814 0.0060     0.0696   0.0814      0.0932
## drugDpenicil_L1       -0.1645 0.1248    -0.4092  -0.1645      0.0802
## sexfemale_L1          -0.3920 0.1596    -0.7049  -0.3920     -0.0791
## year:drugDpenicil_L1   0.0150 0.0083    -0.0014   0.0150      0.0313
## Res. err. (variance)   0.2147 0.0076     0.2001   0.2146      0.2300
##
## Longitudinal outcome (L2, poisson)
##                    mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L2      5.4173 0.0677     5.2845   5.4173      5.5501
## year_L2          -0.0573 0.0016    -0.0604  -0.0573     -0.0542
## f1year_L2         0.0018 0.0002     0.0015   0.0018      0.0021
## drugDpenicil_L2  -0.0726 0.0449    -0.1606  -0.0726      0.0154
## sexfemale_L2      0.1047 0.0662    -0.0250   0.1047      0.2344
##
```

```
## Longitudinal outcome (L3, gaussian)
##                          mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L3           3.5175 0.0344     3.4501   3.5175     3.5849
## year_L3               -0.0590 0.0154    -0.0892  -0.0590    -0.0289
## 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        0.0091 0.0455    -0.0800   0.0091     0.0983
## 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             1.1400 0.0954     0.9612   1.1380     1.3395
## Intercept_L2             0.1574 0.0128     0.1344   0.1567     0.1835
## Intercept_L3             0.1356 0.0136     0.1111   0.1348     0.1643
## Intercept_L1:Intercept_L2 -0.1118 0.0252    -0.1619  -0.1114    -0.0665
## Intercept_L1:Intercept_L3 -0.2470 0.0289    -0.3052  -0.2453    -0.1932
## Intercept_L2:Intercept_L3  0.0498 0.0097     0.0315   0.0495     0.0695
##
## Survival outcome (S1)
##                             mean     sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S1  0.0172 0.0150     0.0022   0.0129     0.0581
## drugDpenicil_S1             -0.0040 0.1834    -0.3635  -0.0040     0.3555
##
## Survival outcome (S2)
##                             mean     sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S2  0.2245 0.2163     0.0313   0.1583     0.8253
## drugDpenicil_S2             -0.4659 0.3752    -1.2012  -0.4659     0.2695
##
## Association longitudinal - survival
##            mean     sd 0.025quant 0.5quant 0.975quant
## CV_L1_S1   1.0428 0.1140     0.8245   1.0408     1.2733
## CV_L1_S2   1.1170 0.2028     0.7085   1.1200     1.5071
## SRE_L2_S1 -0.4041 0.2206    -0.8415  -0.4031     0.0271
## CV_L3_S1  -2.3628 0.3196    -2.9747  -2.3685    -1.7166
## CS_L3_S1   0.0575 1.8449    -3.5975   0.0654     3.6667
## CS_L3_S2  -0.9921 1.2840    -3.8407  -0.8796     1.1320
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                          -56520.94                              -56505.95
##
## Deviance Information Criterion:  -40773.79
## Widely applicable Bayesian information criterion:  -654.0701
## Computation time: 60.2 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:

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

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 0.4227 0.2421     0.1232   0.3671     1.0540
## X_S1                        0.3013 0.0867     0.1315   0.3012     0.4717
##
## Survival outcome (S2)
##                              mean     sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S2 0.5020 0.3511     0.0900   0.4171     1.4221
## X_S2                        0.5029 0.0925     0.3246   0.5018     0.6872
##
## 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)
##                            -239.8306                             -239.8306
##
## Deviance Information Criterion:  377.6674
## Widely applicable Bayesian information criterion:  377.3037
## Computation time: 2.73 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:

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

```
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          -0.4840 0.2215    -0.9177  -0.4842    -0.0490
## time_L1                0.0582 0.1219    -0.1801   0.0580     0.2981
## X_L1                   1.5243 0.0368     1.4520   1.5243     1.5964
## Res. err. (variance)   1.5239 0.1402     1.2674   1.5176     1.8182
##
## Random effects variance-covariance (L1)
##                          mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L1           0.2337 0.1751     0.0594   0.1897     0.6505
## time_L1                0.3972 0.1156     0.2127   0.3840     0.6707
## Intercept_L1:time_L1   0.0404 0.1242    -0.2221   0.0443     0.2791
##
## Survival outcome (S1)
##                               mean     sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S1 0.2032 0.1783     0.0303   0.1512     0.6938
## X_S1                        -0.3946 0.1202    -0.6295  -0.3945    -0.1597
##
## Survival outcome (S2)
##                               mean     sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S2 0.6103 0.4163     0.1429   0.5023     1.7266
## X_S2                        0.2348 0.1620    -0.0632   0.2390     0.5284
##
## Survival outcome (S3)
##                            mean     sd 0.025quant 0.5quant 0.975quant
## Exponential (rate)_S3    0.1660 0.0727     0.0658   0.1518     0.3484
## X_S3                    -0.2864 0.1058    -0.4955  -0.2857    -0.0812
##
## Association longitudinal - survival
##            mean     sd 0.025quant 0.5quant 0.975quant
## CV_L1_S1 0.4733 0.1602     0.1577   0.4733     0.7885
## CV_L1_S2 0.2261 0.2692    -0.2754   0.2170     0.7828
## CV_L1_S3 0.1245 0.1046    -0.0726   0.1216     0.3387
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                             -2396.415                             -2396.415
##
## Deviance Information Criterion:  -1738.366
## Widely applicable Bayesian information criterion:  -1852.169
## Computation time: 9.61 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 $t \to \infty$.

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).

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

```
##
## Survival outcome
##                      mean     sd 0.025quant 0.5quant 0.975quant
## Int(cure)         -0.9970 0.3204    -1.6306  -0.9960    -0.3691
## TRT(cure)         -0.3908 0.4570    -1.3019  -0.3869     0.4974
## Weibull (shape)    1.0903 0.0857     0.9415   1.0838     1.2781
## Weibull (scale)    0.0027 0.0009     0.0013   0.0026     0.0045
## TRT                0.7037 0.2554     0.2027   0.7038     1.2042
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                            -487.1772                             -487.1772
##
## Deviance Information Criterion:  954.4916
## Widely applicable Bayesian information criterion:  955.5733
## Computation time: 0.55 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)
M11 <- joint(formSurv=inla.surv(time, status) ~ sex + age + (1|id),
             id="id", dataSurv = kidney)

summary(M11)
```

```
##
## Survival outcome
##                           mean     sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance) 0.0292 0.0844     0.0000   0.0052     0.2205
## sex                     -1.4569 0.3775    -2.2086  -1.4532    -0.7261
## age                      0.0054 0.0112    -0.0165   0.0054     0.0276
##
```

```
## Frailty term variance
##               mean      sd 0.025quant 0.5quant 0.975quant
## IDIntercept 0.4098 0.2362      0.105   0.3605     1.0109
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                              -210.339                              -210.339
##
## Deviance Information Criterion:  379.1391
## Widely applicable Bayesian information criterion:  379.4611
## Computation time: 2.34 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 φ_1 is scaling the frailty term in the terminal event submodel.

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

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

```
data(readmission, package="frailtypack")
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  0.0444 0.0136    0.0243   0.0422     0.0773
## chemoTreated_S1             -0.0396 0.1411    -0.3159  -0.0397     0.2375
##
## Frailty term variance (S1)
##                   mean      sd 0.025quant 0.5quant 0.975quant
## IDIntercept_S1  0.8138 0.1993     0.4793    0.796     1.2586
##
## Survival outcome (S2)
##                              mean      sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S2 0.0051 0.0017    0.0021   0.0050     0.0084
## chemoTreated_S2             0.3118 0.2664    -0.2110   0.3118     0.8341
##
## Association survival - survival
##                     mean      sd 0.025quant 0.5quant 0.975quant
## IDIntercept_S1_S2 1.8984 0.1753     1.5719   1.8925     2.2613
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                            -2480.481                             -2480.481
##
## Deviance Information Criterion:  4756.36
## Widely applicable Bayesian information criterion:  4745.209
## Computation time: 3.62 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:

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

```
# 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             3.3664 0.2591     2.8617   3.3652     3.8779
## year_L1                 -0.3774 0.0664    -0.5076  -0.3774    -0.2471
## drugDpenicil_L1         -0.2075 0.3559    -0.9057  -0.2075     0.4905
## year:drugDpenicil_L1     0.0372 0.0926    -0.1443   0.0372     0.2188
##
## Longitudinal outcome (L2, gaussian)
##                            mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L2             3.6029 0.0274     3.5492   3.6029     3.6565
## year_L2                 -0.0582 0.0115    -0.0807  -0.0582    -0.0357
## drugDpenicil_L2          0.0141 0.0386    -0.0616   0.0141     0.0899
## year:drugDpenicil_L2     0.0078 0.0163    -0.0241   0.0078     0.0397
## Res. err. (variance)     0.0717 0.0029     0.0660   0.0717     0.0775
##
## Random effects variance-covariance
##                              mean      sd 0.025quant 0.5quant 0.975quant
## Intercept_L1               4.3631 1.3084     2.4168   4.1739     7.4443
## year_L1                    0.1762 0.0521     0.1004   0.1683     0.2996
## Intercept_L2               0.0834 0.0099     0.0658   0.0828     0.1041
## year_L2                    0.0093 0.0011     0.0074   0.0092     0.0118
## Intercept_L1:year_L1      -0.0265 0.2011    -0.4989  -0.0041     0.3065
## Intercept_L1:Intercept_L2  0.4766 0.0920     0.3145   0.4704     0.6702
## Intercept_L1:year_L2      -0.0408 0.0167    -0.0779  -0.0394    -0.0133
## year_L1:Intercept_L2       0.0227 0.0210    -0.0188   0.0225     0.0645
## year_L1:year_L2            0.0029 0.0035    -0.0031   0.0026     0.0106
## Intercept_L2:year_L2      -0.0076 0.0022    -0.0124  -0.0074    -0.0039
##
## log marginal-likelihood (integration)   log marginal-likelihood (Gaussian)
##                            -1270.014                            -1270.014
##
## Deviance Information Criterion:  1495.943
## Widely applicable Bayesian information criterion:  2498.889
## Computation time: 9.29 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:

$$
\begin{cases}
\text{Logit}(Prob(albumin_{ij} > 0)) = \beta_{10} + b_{i10} + (\beta_{11} + b_{i11})year_{ij} + \beta_{12}drug_i + \beta_{13}year_{ij}drug_i & \text{(L1)} \\
\text{E}[albumin_{ij}|albumin_{ij} > 0] = \beta_{20} + b_{i20} + (\beta_{21} + b_{i21})year_{ij} + \beta_{22}drug_i + \beta_{23}year_{ij}drug_i + \varepsilon_{ij} & \text{(L2)} \\
\lambda_i(t) = \lambda_0(t) \exp\left(\gamma_1 drug_i + \varphi_1 b_{i10} + \varphi_2 b_{i11} + \varphi_3 b_{i20} + \varphi_4 b_{i21}\right) & \text{(S1)}
\end{cases}
$$

```
M14 <- joint(formLong = list(binary ~ year * drug +  (1+year|id),
                        albumin ~ year * drug + (1+year|id)),
          formSurv = inla.surv(years, death) ~ drug, dataSurv=Surv,
          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          3.5140 0.2707     2.9866   3.5127     4.0485
## year_L1              -0.5104 0.0692    -0.6463  -0.5103    -0.3749
## drugDpenicil_L1      -0.1853 0.3716    -0.9142  -0.1853     0.5434
## year:drugDpenicil_L1  0.0584 0.0959    -0.1294   0.0584     0.2465
##
## Longitudinal outcome (L2, gaussian)
##                         mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L2          3.5985 0.0276     3.5444   3.5986     3.6526
## year_L2              -0.0634 0.0115    -0.0861  -0.0634    -0.0408
## drugDpenicil_L2       0.0156 0.0389    -0.0607   0.0156     0.0920
## year:drugDpenicil_L2  0.0099 0.0163    -0.0221   0.0099     0.0419
## Res. err. (variance)  0.0724 0.0030     0.0667   0.0723     0.0785
##
## Random effects variance-covariance
##                            mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L1             4.6448 1.4913     2.5184   4.3948     8.2700
## year_L1                  0.1577 0.0548     0.0774   0.1490     0.2819
## Intercept_L2             0.0808 0.0111     0.0610   0.0801     0.1049
## year_L2                  0.0090 0.0014     0.0066   0.0089     0.0120
## Intercept_L1:year_L1     0.2286 0.1428    -0.0679   0.2303     0.4972
## Intercept_L1:Intercept_L2  0.4672 0.0766     0.3311   0.4604     0.6351
## Intercept_L1:year_L2    -0.0125 0.0154    -0.0381  -0.0137     0.0228
## year_L1:Intercept_L2     0.0389 0.0209    -0.0002   0.0379     0.0800
## year_L1:year_L2          0.0044 0.0033    -0.0010   0.0040     0.0118
## Intercept_L2:year_L2    -0.0049 0.0019    -0.0088  -0.0048    -0.0013
##
## Survival outcome
##                            mean     sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S1  0.2283 0.1686     0.0496   0.1820     0.6867
## drugDpenicil_S1             -0.0836 0.2540    -0.5819  -0.0835     0.4144
##
## Association longitudinal - survival
##                       mean     sd 0.025quant 0.5quant 0.975quant
## SRE_Intercept_L1_S1 -0.5261 0.1258    -0.7739  -0.5260    -0.2787
## SRE_year_L1_S1      -1.9051 0.4216    -2.7229  -1.9093    -1.0628
## SRE_Intercept_L2_S1 -0.2324 0.6730    -1.5744  -0.2266     1.0756
## SRE_year_L2_S1      -0.5692 0.8762    -2.2719  -0.5767     1.1778
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                          -1869.17                               -1869.17
##
## Deviance Information Criterion:  2571.514
## Widely applicable Bayesian information criterion:  3600.515
## Computation time: 19.73 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 (i.e., SLD of target lesions) in a colorectal cancer clinical trial, the recurrent events are new tumors and the terminal event is death.

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

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

```
data(colorectal, package="frailtypack")
data(colorectalLongi, package="frailtypack")
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     7.0054 1.0966     4.8566   7.0049     9.1572
## year_L1          0.2544 1.0324    -1.7701   0.2544     2.2789
## treatmentC_L1   -2.0337 1.1736    -4.3348  -2.0337     0.2678
## year:treatmentC_L1 -1.0449 1.0621  -3.1275  -1.0449     1.0378
##
## Longitudinal outcome (L2, lognormal)
```

```
##                          mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L2           2.1263 0.0845     1.9605   2.1263     2.2922
## year_L2               -0.1364 0.0785    -0.2906  -0.1364     0.0175
## treatmentC_L2          0.0542 0.1212    -0.1835   0.0542     0.2920
## year:treatmentC_L2    -0.3174 0.1118    -0.5368  -0.3173    -0.0983
## Res. err. (variance)   0.1646 0.0094     0.1470   0.1643     0.1838
##
## Random effects variance-covariance
##                           mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L1            2.7479 1.5273     0.9246   2.3766     6.9109
## Intercept_L2            0.4925 0.0731     0.3707   0.4846     0.6549
## year_L2                 0.1858 0.0532     0.1039   0.1794     0.3101
## Intercept_L1:Intercept_L2  0.6985 0.2789     0.1019   0.6969     1.2496
## Intercept_L1:year_L2   -0.1929 0.1695    -0.5601  -0.1801     0.1010
## Intercept_L2:year_L2   -0.0836 0.0515    -0.1918  -0.0795     0.0089
##
## Survival outcome (S1)
##                             mean     sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S1  0.1391 0.1381     0.0213   0.0960     0.5233
## treatmentC_S1              -0.3621 0.2466    -0.8464  -0.3619     0.1211
##
## Frailty term variance (S1)
##                 mean     sd 0.025quant 0.5quant 0.975quant
## IDIntercept_S1 1.0458 0.6859     0.2524    0.874     2.8707
##
## Survival outcome (S2)
##                             mean     sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S2  0.0676 0.0625     0.0111   0.0485     0.2415
## treatmentC_S2              -0.1512 0.3344    -0.8073  -0.1511     0.5043
##
## Association longitudinal - survival
##                        mean     sd 0.025quant 0.5quant 0.975quant
## SRE_Intercept_L1_S1   0.3177 0.2408    -0.1497   0.3155     0.7984
## SRE_Intercept_L1_S2   0.5747 0.3033    -0.0115   0.5709     1.1828
## SRE_Intercept_L2_S1  -0.1468 0.4808    -1.1001  -0.1445     0.7930
## SRE_year_L2_S1       -0.0910 0.5912    -1.2246  -0.1011     1.1028
## SRE_Intercept_L2_S2   0.3910 0.5588    -0.7293   0.3979     1.4708
## SRE_year_L2_S2       -0.3440 0.5751    -1.4362  -0.3571     0.8271
##
## Association survival - survival
##                   mean     sd 0.025quant 0.5quant 0.975quant
## IDIntercept_S1_S2 1.2469 0.2601     0.7291   1.2488     1.7533
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                          -3504.581                             -3504.581
##
## Deviance Information Criterion:  6394.141
## Widely applicable Bayesian information criterion:  6414.054
## Computation time: 15.74 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(serBilir_{ij}) &= \eta_{i1}(t_{ij}) + \varepsilon_{ij1} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \text{(L1)}\\
&= (\beta_{10} + b_{i10}) + \beta_{11}X_i + (\beta_{12} + b_{i11})\mathrm{NS}_1(t_{ij}) + (\beta_{13} + b_{i12})\mathrm{NS}_2(t_{ij})\\
&\quad + (\beta_{14} + b_{i13})\mathrm{NS}_3(t_{ij}) + \beta_{15}X_i\mathrm{NS}_1(t_{ij}) + \beta_{16}X_i\mathrm{NS}_2(t_{ij}) + \beta_{17}X_i\mathrm{NS}_3(t_{ij}) + \varepsilon_{ij1}\\
\log(SGOT_{ij}) &= \eta_{i2}(t_{ij}) + \varepsilon_{ij2} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \text{(L2)}\\
&= (\beta_{20} + b_{i20}) + \beta_{21}X_i + (\beta_{22} + b_{i21})\mathrm{NS}_1(t_{ij}) + (\beta_{23} + b_{i22})\mathrm{NS}_2(t_{ij})\\
&\quad + (\beta_{24} + b_{i23})\mathrm{NS}_3(t_{ij}) + \beta_{25}X_i\mathrm{NS}_1(t_{ij}) + \beta_{26}X_i\mathrm{NS}_2(t_{ij}) + \beta_{27}X_i\mathrm{NS}_3(t_{ij}) + \varepsilon_{ij2}\\
albumin_{ij} &= \eta_{i3}(t_{ij}) + \varepsilon_{ij3} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \text{(L3)}\\
&= (\beta_{30} + b_{i30}) + \beta_{31}X_i + (\beta_{32} + b_{i31})t + \beta_{33}X_it + \varepsilon_{ij3}\\
\log(\mathrm{E}[platelets_{ij}]) &= \eta_{i4}(t_{ij}) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \text{(L4)}\\
&= (\beta_{40} + b_{i40}) + \beta_{41}X_i + (\beta_{42} + b_{i41})\mathrm{NS}_1(t_{ij}) + (\beta_{43} + b_{i42})\mathrm{NS}_2(t_{ij})\\
&\quad + (\beta_{44} + b_{i43})\mathrm{NS}_3(t_{ij}) + \beta_{45}X_i\mathrm{NS}_1(t_{ij}) + \beta_{46}X_i\mathrm{NS}_2(t_{ij}) + \beta_{47}X_i\mathrm{NS}_3(t_{ij})\\
\mathrm{logit}(\mathrm{E}[spiders_{ij}]) &= \eta_{i5}(t_{ij}) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \text{(L5)}\\
&= (\beta_{50} + b_{i50}) + \beta_{51}X_i + (\beta_{52} + b_{i51})t + \beta_{53}X_it\\
\lambda_{i1}(t) &= \lambda_{01}(t) \exp\left(\gamma_1 X_i + \varphi_1\eta_{i1}(t) + \varphi_3\eta_{i1}'(t) + \varphi_4\eta_{i2}(t)\right. \qquad\qquad\qquad \text{(S1)}\\
&\quad \left. + \varphi_5\eta_{i3}(t) + \varphi_7\eta_{i4}(t) + \varphi_9\eta_{i5}(t)\right)\\
\lambda_{i2}(t) &= \lambda_{02}(t) \exp\left(\gamma_2 X_i + \varphi_2\eta_{i1}(t) + \varphi_6\eta_{i3}(t) + \varphi_8\eta_{i4}(t)\right) \qquad\qquad\qquad\; \text{(S2)}
\end{cases}
$$

where $\mathrm{NS}_1(t), \mathrm{NS}_2(t), \mathrm{NS}_3(t)$ are the natural cubic splines with internal knots at 1 and 4 years. We assume independent random effects between longitudinal markers. For computational convenience, here we fix the variance and covariance of all the random effects, so the model can run in a reasonable computation time on any computer. This is done by the use of the control argument 'initVC' to set initial values for variance and covariance terms (which can also be replaced by 'initSD' to set initial values for standard deviations and correlations instead) and the control argument 'fixRE' which is a boolean that allows to fix the values of random effects variances and covariances (or standard deviations and correlations). Noe that removing 'initVC' and 'initRE' from control arguments will estimate the full model.

```r
# 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]


initVC <- list(c(1.04, 1.58, 2.43, 1.80, 0.32, 0.63, 0.63, 1.76, 0.89, 1.42),
               c(0.17, 0.21, 0.31, 0.22, -0.02, 0.02, 0.01, 0.17, 0.09, 0.15),
               c(0.13, 0.01, 0),
               c(0.15, 1.31, 4.99, 16.32, -0.04, -0.08, -0.06, -1.83, -3.70, 8.62),
               c(5.89, 0.14, -0.25))
fixRE <- list(T,T,T,T,T)


M16 <-joint(formSurv = list(inla.surv(years, death) ~ drug,
                       inla.surv(years, trans) ~ 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, dataSurv=Surv, 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", initVC=initVC, fixRE=fixRE))
summary(M16)
```

```
## Longitudinal outcome (L1, lognormal)
##                             mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L1              0.5928 0.1952     0.2103   0.5928     0.9754
## f1year_L1                 0.7642 0.0938     0.5803   0.7642     0.9481
## f2year_L1                 1.4084 0.1624     1.0901   1.4084     1.7267
## f3year_L1                 1.2461 0.1861     0.8813   1.2461     1.6110
## drugDpenicil_L1          -0.1130 0.2743    -0.6506  -0.1130     0.4245
## f1year:drugDpenicil_L1    0.2056 0.1304    -0.0499   0.2056     0.4612
## f2year:drugDpenicil_L1   -0.1470 0.2290    -0.5960  -0.1470     0.3019
## f3year:drugDpenicil_L1   -0.1018 0.2595    -0.6105  -0.1018     0.4068
## Res. err. (variance)      0.0908 0.0048     0.0828   0.0902     0.1016
##
## Random effects variance-covariance (L1)
##            Intercept_L1 f1year_L1 f2year_L1 f3year_L1
## Intercept_L1        1.04      0.32      0.63      0.63
## f1year_L1           0.32      1.58      1.76      0.89
## f2year_L1           0.63      1.76      2.43      1.42
## f3year_L1           0.63      0.89      1.42      1.80
##
## Longitudinal outcome (L2, lognormal)
##                             mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L2              4.7520 0.0477     4.6585   4.7520     4.8455
## f1year_L2                -0.1509 0.0703    -0.2887  -0.1509    -0.0130
## f2year_L2                 0.1167 0.0934    -0.0664   0.1167     0.2997
## f3year_L2                 0.0556 0.1239    -0.1872   0.0556     0.2983
## drugDpenicil_L2          -0.0842 0.0670    -0.2156  -0.0842     0.0471
## f1year:drugDpenicil_L2    0.1129 0.0979    -0.0789   0.1129     0.3047
## f2year:drugDpenicil_L2   -0.2457 0.1310    -0.5024  -0.2457     0.0109
## f3year:drugDpenicil_L2   -0.0492 0.1711    -0.3845  -0.0492     0.2861
## Res. err. (variance)      0.0692 0.0046     0.0605   0.0691     0.0784
##
## Random effects variance-covariance (L2)
##            Intercept_L2 f1year_L2 f2year_L2 f3year_L2
## Intercept_L2        0.17     -0.02      0.02      0.01
## f1year_L2          -0.02      0.21      0.17      0.09
## f2year_L2           0.02      0.17      0.31      0.15
## f3year_L2           0.01      0.09      0.15      0.22
##
## Longitudinal outcome (L3, gaussian)
##                             mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L3              3.5494 0.0332     3.4844   3.5494     3.6145
## year_L3                  -0.1015 0.0116    -0.1242  -0.1015    -0.0788
## drugDpenicil_L3           0.0000 0.0468    -0.0916   0.0000     0.0917
## year:drugDpenicil_L3      0.0034 0.0162    -0.0284   0.0034     0.0353
## Res. err. (variance)      0.0958 0.0050     0.0865   0.0955     0.1062
##
## Random effects variance-covariance (L3)
##             Intercept_L3 year_L3
## Intercept_L3         0.13    0.00
## year_L3              0.00    0.01
##
```

```
## Longitudinal outcome (L4, poisson)
##                         mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L4           5.5195 0.0579     5.4060   5.5195     5.6330
## f1year_L4             -0.1246 0.1411    -0.4013  -0.1246     0.1520
## f2year_L4             -0.9997 0.2087    -1.4088  -0.9997    -0.5906
## f3year_L4             -1.3772 0.3815    -2.1248  -1.3772    -0.6295
## drugDpenicil_L4       -0.0585 0.0814    -0.2180  -0.0585     0.1009
## f1year:drugDpenicil_L4  0.2157 0.1986   -0.1736   0.2157     0.6050
## f2year:drugDpenicil_L4 -0.3868 0.2963   -0.9675  -0.3868     0.1939
## f3year:drugDpenicil_L4 -0.4530 0.5402   -1.5117  -0.4530     0.6057
##
## Random effects variance-covariance (L4)
##            Intercept_L4 f1year_L4 f2year_L4 f3year_L4
## Intercept_L4       0.15     -0.04     -0.08     -0.06
## f1year_L4         -0.04      1.31     -1.83     -3.70
## f2year_L4         -0.08     -1.83      4.99      8.62
## f3year_L4         -0.06     -3.70      8.62     16.32
##
## Longitudinal outcome (L5, binomial)
##                      mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L5       -1.4119 0.2463    -1.8947  -1.4119    -0.9292
## year_L5             0.1486 0.0606     0.0298   0.1486     0.2674
## drugDpenicil_L5    -0.1086 0.3496    -0.7938  -0.1086     0.5765
## year:drugDpenicil_L5 -0.0288 0.0860  -0.1973  -0.0288     0.1397
##
## Random effects variance-covariance (L5)
##             Intercept_L5 year_L5
## Intercept_L5        5.89   -0.25
## year_L5            -0.25    0.14
##
## Survival outcome (S1)
##                          mean     sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S1  0.0220 0.0326   0.0006   0.0111     0.1124
## drugDpenicil_S1         -0.1169 0.1871    -0.4836  -0.1169     0.2497
##
## Survival outcome (S2)
##                          mean     sd 0.025quant 0.5quant 0.975quant
## Baseline risk (variance)_S2  0.1856 0.3375   0.0056   0.0784     1.0692
## drugDpenicil_S2         -0.4280 0.3763    -1.1656  -0.4280     0.3096
##
## Association longitudinal - survival
##            mean     sd 0.025quant 0.5quant 0.975quant
## CV_L1_S1  1.3604 0.1467     1.0765   1.3587     1.6542
## CS_L1_S1 -0.0950 1.0995    -2.1977  -0.1155     2.1309
## CV_L1_S2  1.0568 0.2298     0.5915   1.0608     1.4973
## CV_L2_S1 -0.3504 0.2890    -0.9396  -0.3437     0.1979
## CV_L3_S1 -1.7794 0.3288    -2.4526  -1.7709    -1.1584
## CV_L3_S2 -0.8825 0.3860    -1.6701  -0.8732    -0.1513
## CV_L4_S1 -0.5757 0.6134    -1.7196  -0.5959     0.6927
## CV_L4_S2 -0.0814 0.4831    -1.0487  -0.0759     0.8536
## CV_L5_S1 -0.0091 0.0655    -0.1393  -0.0087     0.1188
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                        -79986.31                                  -79973.16
```

```
##
## Deviance Information Criterion:  -64827.96
## Widely applicable Bayesian information criterion:  -58501.67
## Computation time: 94.08 seconds
```