INLAjoint

Denis Rustand

2022-06-16

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

Contact: INLAjoint@gmail.com

Dataset for illustrations

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.

This vignette illustrates how to fit various joint model including multiple longitudinal markers, competing risks of events and multi-state models. The final model illustrated is a joint model for two competing risks of events and 5 longitudinal markers with different distributions as proposed in the application section of the following paper: https://arxiv.org/abs/2203.06256

```
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
```

Model 1: 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:

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

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_L1      0.6304 0.0885     0.4568   0.6304     0.8041
## year_L1           0.0843 0.0042     0.0760   0.0843     0.0926
## drugDpenicil_L1 -0.1223 0.1240    -0.3655  -0.1222     0.1209
## Res. err. (var)  0.2154 0.0077     0.2008   0.2152     0.2309
##
## Random effects variance-covariance (L1)
##                 mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L1 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: 0.84 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_L1      0.6304 0.0885     0.4568   0.6304     0.8041
## year_L1           0.0843 0.0042     0.0760   0.0843     0.0926
## drugDpenicil_L1 -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_L1 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: 0.84 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:

$$\begin{cases} \log(serBilir_{ij}) = \beta_{10} + b_{i10} + (\beta_{11} + b_{i11})year_{ij} + \beta_{12}drug_i + \beta_{13}sex_i + \beta_{14}year_{ij}drug_i + \varepsilon_{ij1} & \text{(L1)} \\ \log(E[platelets_{ij}]) = \beta_{20} + b_{i20} + (\beta_{21} + b_{i21})year_{ij} + \beta_{22}sex_i + \beta_{23}drug_i + \beta_{24}year_{ij}sex_i & \text{(L2)} \\ \text{logit}(E[spiders_{ij}]) = \beta_{30} + b_{i30} + (\beta_{31} + b_{i31})year_{ij} + \beta_{32}drug_i + \beta_{33}year_{ij}drug_i & \text{(L3)} \end{cases}$$

```
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            0.9180 0.1665     0.5917   0.9180     1.2443
## year_L1                 0.1568 0.0186     0.1203   0.1568     0.1932
## drugDpenicil_L1        -0.1621 0.1149    -0.3873  -0.1621     0.0631
## sexfemale_L1           -0.3742 0.1606    -0.6891  -0.3742    -0.0594
## year:drugDpenicil_L1    0.0099 0.0254    -0.0399   0.0099     0.0598
## Res. err. (var)         0.1092 0.0044     0.1009   0.1091     0.1180
##
## Longitudinal outcome (L2, poisson)
##                         mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L2          5.4418 0.0678     5.3089   5.4418     5.5747
## year_L2              -0.1259 0.0322    -0.1890  -0.1259    -0.0629
## sexfemale_L2          0.0914 0.0668    -0.0395   0.0914     0.2223
## drugDpenicil_L2      -0.0742 0.0437    -0.1598  -0.0742     0.0114
## year:sexfemale_L2     0.0483 0.0341    -0.0186   0.0483     0.1152
##
## Longitudinal outcome (L3, binomial)
##                       mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L3    -1.7399 0.3085    -2.3446  -1.7399    -1.1352
## year_L3          0.0518 0.0701    -0.0856   0.0518     0.1892
## drugDpenicil_L3 -0.2670 0.4242    -1.0985  -0.2670     0.5645
##
## Random effects variance-covariance
##                               mean     sd 0.025quant 0.5quant 0.975quant
```

```
## Intercept_L1               1.0003 0.0844     0.8449   0.9962     1.1789
## year_L1                    0.0344 0.0051     0.0257   0.0340     0.0454
## Intercept_L2               0.1481 0.0123     0.1260   0.1473     0.1740
## year_L2                    0.0324 0.0049     0.0238   0.0321     0.0429
## Intercept_L3              10.0301 1.7746     7.1130   9.8520    14.0946
## year_L3                    0.5165 0.1226     0.3179   0.5043     0.7925
## Intercept_L1:year_L1       0.0456 0.0116     0.0242   0.0451     0.0699
## Intercept_L1:Intercept_L2 -0.0830 0.0237    -0.1306  -0.0827    -0.0374
## Intercept_L1:year_L2      -0.0179 0.0107    -0.0398  -0.0175     0.0023
## Intercept_L1:Intercept_L3  1.6201 0.2644     1.1586   1.5990     2.1980
## Intercept_L1:year_L3       0.1277 0.0844    -0.0307   0.1248     0.3078
## year_L1:Intercept_L2      -0.0053 0.0048    -0.0148  -0.0053     0.0042
## year_L1:year_L2           -0.0035 0.0020    -0.0075  -0.0035     0.0005
## year_L1:Intercept_L3       0.0982 0.0522     0.0052   0.0951     0.2113
## year_L1:year_L3            0.0611 0.0155     0.0344   0.0596     0.0947
## Intercept_L2:year_L2      -0.0039 0.0036    -0.0110  -0.0039     0.0030
## Intercept_L2:Intercept_L3 -0.3388 0.0900    -0.5248  -0.3354    -0.1737
## Intercept_L2:year_L3       0.0122 0.0267    -0.0423   0.0123     0.0643
## year_L2:Intercept_L3       0.0164 0.0523    -0.0827   0.0144     0.1248
## year_L2:year_L3           -0.0319 0.0181    -0.0722  -0.0305     0.0002
## Intercept_L3:year_L3      -0.1427 0.3273    -0.8415  -0.1252     0.4700
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                              -17077.99                              -17057.48
##
## Deviance Information Criterion:  25628.79
## Widely applicable Bayesian information criterion:  45405.04
## Computation time: 26.56 seconds
```

The additional boolean argument `corLong` is set to TRUE in order to have correlation between the random effects accross 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           0.9180 0.1665     0.5917   0.9180     1.2443
## year_L1                0.1568 0.0186     0.1203   0.1568     0.1932
## drugDpenicil_L1       -0.1621 0.1149    -0.3873  -0.1621     0.0631
## sexfemale_L1          -0.3742 0.1606    -0.6891  -0.3742    -0.0594
## year:drugDpenicil_L1   0.0099 0.0254    -0.0399   0.0099     0.0598
## Res. err. (sd)         0.3304 0.0066     0.3176   0.3304     0.3435
##
## Longitudinal outcome (L2, poisson)
##                        mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L2         5.4418 0.0678     5.3089   5.4418     5.5747
## year_L2             -0.1259 0.0322    -0.1890  -0.1259    -0.0629
## sexfemale_L2         0.0914 0.0668    -0.0395   0.0914     0.2223
## drugDpenicil_L2     -0.0742 0.0437    -0.1598  -0.0742     0.0114
## year:sexfemale_L2    0.0483 0.0341    -0.0186   0.0483     0.1152
```

```
##
## Longitudinal outcome (L3, binomial)
##                     mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L3    -1.7399 0.3085    -2.3446  -1.7399    -1.1352
## year_L3          0.0518 0.0701    -0.0856   0.0518     0.1892
## drugDpenicil_L3 -0.2670 0.4242    -1.0985  -0.2670     0.5645
##
## Random effects standard deviation / correlation
##                                mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L1                 0.9991 0.0425     0.9196   0.9977     1.0864
## year_L1                      0.1854 0.0136     0.1604   0.1847     0.2136
## Intercept_L2                 0.3843 0.0159     0.3544   0.3840     0.4165
## year_L2                      0.1798 0.0137     0.1541   0.1792     0.2083
## Intercept_L3                 3.1562 0.2716     2.6692   3.1379     3.7388
## year_L3                      0.7139 0.0837     0.5594   0.7098     0.8893
## Intercept_L1:year_L1         0.2469 0.0563     0.1355   0.2479     0.3546
## Intercept_L1:Intercept_L2 -0.2154 0.0572    -0.3256  -0.2162    -0.0996
## Intercept_L1:year_L2      -0.0983 0.0560    -0.2074  -0.0980     0.0124
## Intercept_L1:Intercept_L3  0.5124 0.0535     0.4009   0.5147     0.6104
## Intercept_L1:year_L3         0.1781 0.1069    -0.0399   0.1803     0.3837
## year_L1:Intercept_L2      -0.0737 0.0671    -0.2014  -0.0750     0.0619
## year_L1:year_L2           -0.1065 0.0597    -0.2191  -0.1075     0.0172
## year_L1:Intercept_L3         0.1670 0.0818     0.0072   0.1659     0.3263
## year_L1:year_L3              0.4611 0.0795     0.2888   0.4666     0.6020
## Intercept_L2:year_L2      -0.0561 0.0492    -0.1506  -0.0561     0.0422
## Intercept_L2:Intercept_L3 -0.2788 0.0670    -0.4025  -0.2805    -0.1424
## Intercept_L2:year_L3         0.0455 0.0924    -0.1411   0.0458     0.2266
## year_L2:Intercept_L3         0.0287 0.0894    -0.1441   0.0281     0.2042
## year_L2:year_L3           -0.2445 0.1202    -0.4718  -0.2486     0.0047
## Intercept_L3:year_L3      -0.0561 0.1337    -0.3164  -0.0554     0.2062
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                              -17077.99                              -17057.48
##
## Deviance Information Criterion:  25628.79
## Widely applicable Bayesian information criterion:  45405.04
## Computation time: 26.56 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          0.3018 0.1646    -0.0208   0.3018     0.6245
```

```
## year_L1              0.0971 0.0184     0.0610   0.0971     0.1332
## drugDpenicil_L1      0.1556 0.1136    -0.0671   0.1556     0.3784
## sexfemale_L1         0.0938 0.1587    -0.2173   0.0938     0.4048
## year:drugDpenicil_L1 0.1040 0.0251     0.0548   0.1040     0.1532
## Res. err. (var)      0.1092 0.0042     0.1010   0.1091     0.1177
##
## Longitudinal outcome (L2, poisson)
##                       mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L2       5.4794 0.0684     5.3453   5.4794     5.6134
## year_L2           -0.0818 0.0323    -0.1451  -0.0818    -0.0184
## sexfemale_L2       0.0586 0.0673    -0.0734   0.0586     0.1905
## drugDpenicil_L2   -0.0977 0.0441    -0.1842  -0.0977    -0.0113
## year:sexfemale_L2  0.0071 0.0343    -0.0601   0.0071     0.0742
##
## Longitudinal outcome (L3, binomial)
##                     mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L3    -1.1762 0.1720    -1.5133  -1.1762    -0.8391
## year_L3          0.1905 0.0472     0.0980   0.1905     0.2830
## drugDpenicil_L3 -1.0360 0.2393    -1.5050  -1.0360    -0.5669
##
## Random effects variance-covariance
##                              mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L1               1.0047 0.0872     0.8493   0.9992     1.1908
## year_L1                    0.0344 0.0045     0.0266   0.0341     0.0442
## Intercept_L2               0.1505 0.0128     0.1268   0.1503     0.1767
## year_L2                    0.0323 0.0047     0.0241   0.0319     0.0423
## Intercept_L3               3.3538 0.5922     2.3987   3.2808     4.7012
## year_L3                    0.2646 0.0751     0.1486   0.2543     0.4433
## Intercept_L1:year_L1       0.0477 0.0150     0.0195   0.0470     0.0792
## Intercept_L1:Intercept_L2 -0.0806 0.0254    -0.1315  -0.0804    -0.0318
## Intercept_L1:year_L2      -0.0209 0.0109    -0.0426  -0.0208     0.0005
## Intercept_L1:Intercept_L3  0.9361 0.1698     0.6379   0.9229     1.3033
## Intercept_L1:year_L3       0.1381 0.0523     0.0419   0.1344     0.2526
## year_L1:Intercept_L2      -0.0063 0.0062    -0.0186  -0.0062     0.0059
## year_L1:year_L2           -0.0041 0.0023    -0.0085  -0.0041     0.0007
## year_L1:Intercept_L3       0.0547 0.0386    -0.0162   0.0529     0.1360
## year_L1:year_L3            0.0483 0.0119     0.0281   0.0471     0.0746
## Intercept_L2:year_L2      -0.0036 0.0035    -0.0105  -0.0035     0.0033
## Intercept_L2:Intercept_L3 -0.2010 0.0569    -0.3211  -0.1984    -0.0961
## Intercept_L2:year_L3      -0.0022 0.0185    -0.0421  -0.0013     0.0319
## year_L2:Intercept_L3       0.0142 0.0320    -0.0489   0.0137     0.0781
## year_L2:year_L3           -0.0252 0.0103    -0.0474  -0.0245    -0.0065
## Intercept_L3:year_L3       0.0503 0.1199    -0.1867   0.0486     0.2907
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                             -17070.94                               -17050.43
##
## Deviance Information Criterion:  26118.08
## Widely applicable Bayesian information criterion:  108134.4
## Computation time: 27.09 seconds
```

Model 3: longitudinal - survival joint model

Some additional arguments are introduced to fit a joint model with a survival component and to set up the association between the longitudinal and survival parts:

- `formSurv`: formula for the time-to-event outcome, with the response given as an inla.surv() object.
- `dataSurv`: Optional, if 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. There are two options: "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. We only propose non-parametric functions for the baseline risk at the moment as it is a flexible approach that avoids parametric assumptions. The number of bins that define the intervals for the baseline risk can be specified with `NbasRisk` (default 15 bins).
- `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:

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

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

```
DTH <- inla.surv(time = Surv$years, event = Surv$death) # survival outcome
f1 <- function(x) x^2
f2 <- function(x) x^3

M3 <- 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(M3)
```

```
## Longitudinal outcome (lognormal)
##                             mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L1              0.5523 0.0529     0.4487   0.5523     0.6559
## year_L1                   0.1988 0.0674     0.0666   0.1988     0.3310
## f1year_L1                -0.0060 0.0278    -0.0605  -0.0060     0.0485
## f2year_L1                 0.0542 0.0074     0.0398   0.0542     0.0686
```

```
## drugDpenicil_L1        -0.1332 0.0744    -0.2790  -0.1332     0.0127
## year:drugDpenicil_L1   -0.4340 0.0945    -0.6192  -0.4340    -0.2489
## f1year:drugDpenicil_L1  0.1342 0.0392     0.0574   0.1342     0.2110
## f2year:drugDpenicil_L1 -0.0347 0.0103    -0.0548  -0.0347    -0.0145
## Res. err. (var)         0.6421 0.0234     0.5977   0.6416     0.6897
##
## Random effects variance-covariance (L1)
##                         mean     sd 0.025quant 0.5quant 0.975quant
## f1year_L1            0.0234 0.0038     0.0168   0.0231     0.0313
## f2year_L1            0.0054 0.0007     0.0042   0.0054     0.0069
## f1year_L1:f2year_L1 -0.0008 0.0004    -0.0017  -0.0007     0.0001
##
## Survival outcome
##                     mean     sd 0.025quant 0.5quant 0.975quant
## Baseline (var)   0.2998 0.1016     0.1492   0.2836     0.5455
## Intercept_S1    -6.3524 0.2444    -6.8314  -6.3524    -5.8734
## drugDpenicil_S1 -1.9457 0.2392    -2.4146  -1.9457    -1.4768
##
## Association longitudinal - survival
##            mean     sd 0.025quant 0.5quant 0.975quant
## CV_L1_S1 1.1762 0.1315     0.9414   1.1675     1.4559
## CS_L1_S1 0.6628 0.2102     0.2223   0.6733     1.0476
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                              -33381.62                              -33374.89
##
## Deviance Information Criterion:  -5.333665e+17
## Widely applicable Bayesian information criterion:  676091484270
## Computation time: 26.26 seconds
```

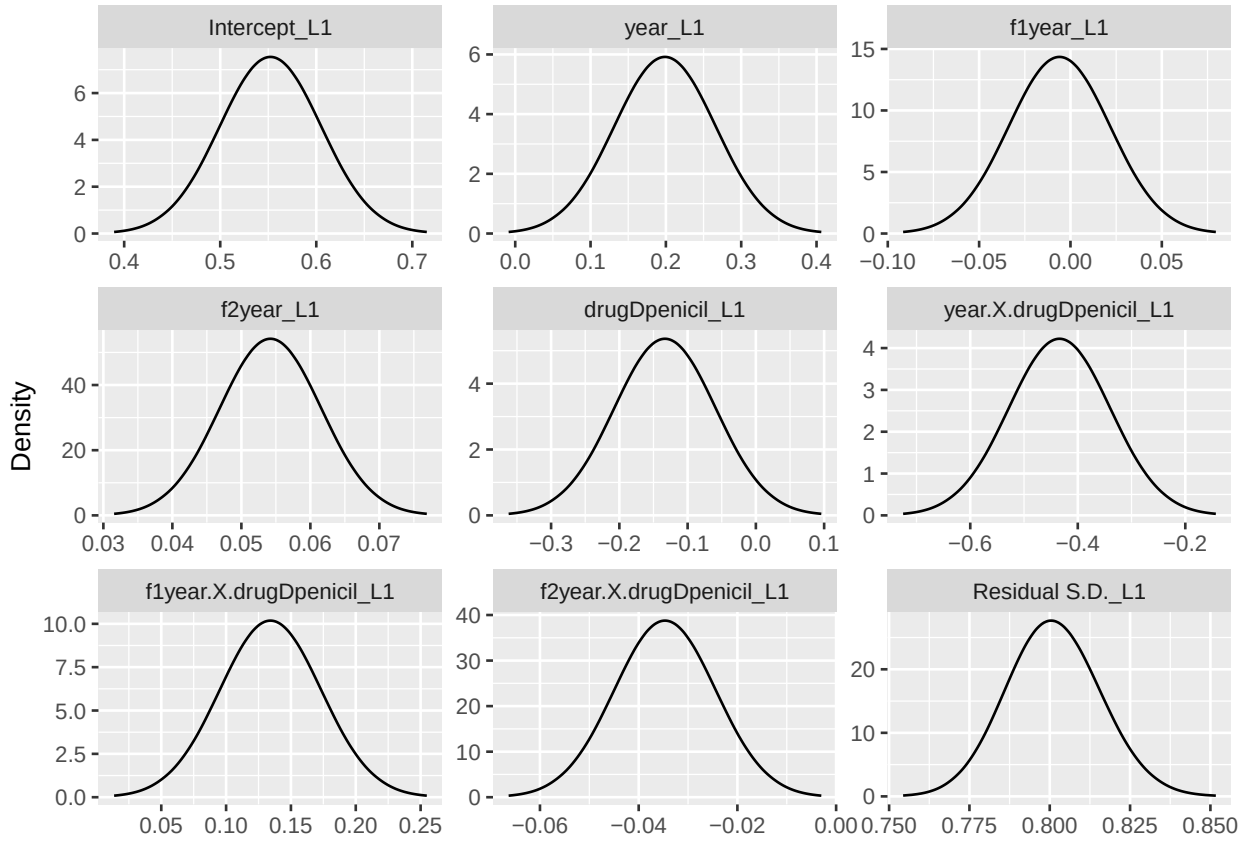
In case some functions of time should be included, they must be set as illustrated in the above example ; i.e., create a univariate function of x named f1, f2, . . . , fN, and use this function in the formula. This is important to be able to compute the value of the linear predictor at any time t, particularly 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

```
plotM3 <- plot(M3, sdcor=T)
```

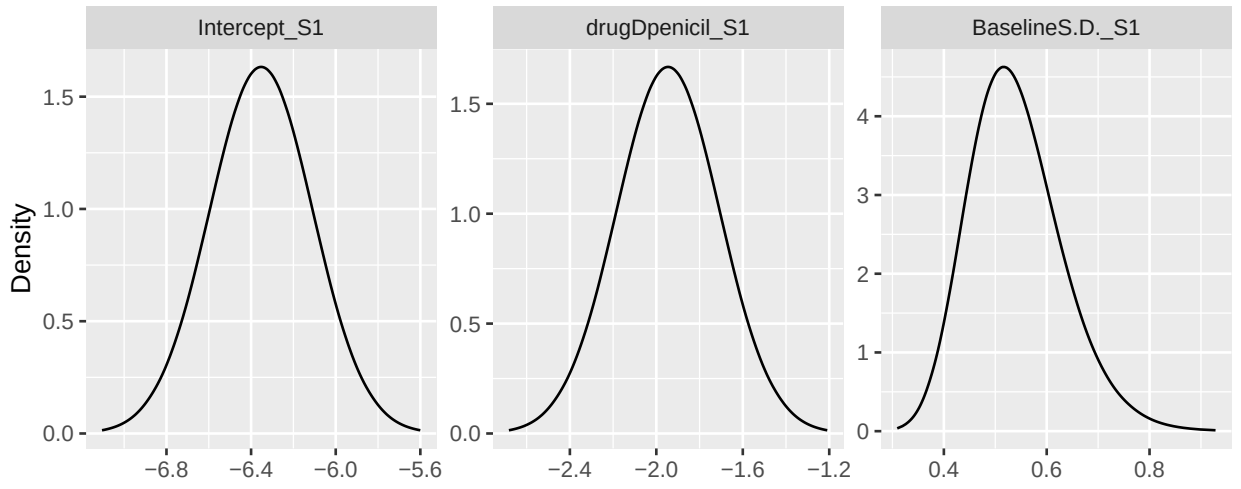
The plot function returns multiple plots for each component of the model. First the plots for the longitudinal outcome(s) parameters:

```
plotM3$Outcomes$L1
```



Then the parameters of the survival outcome(s):

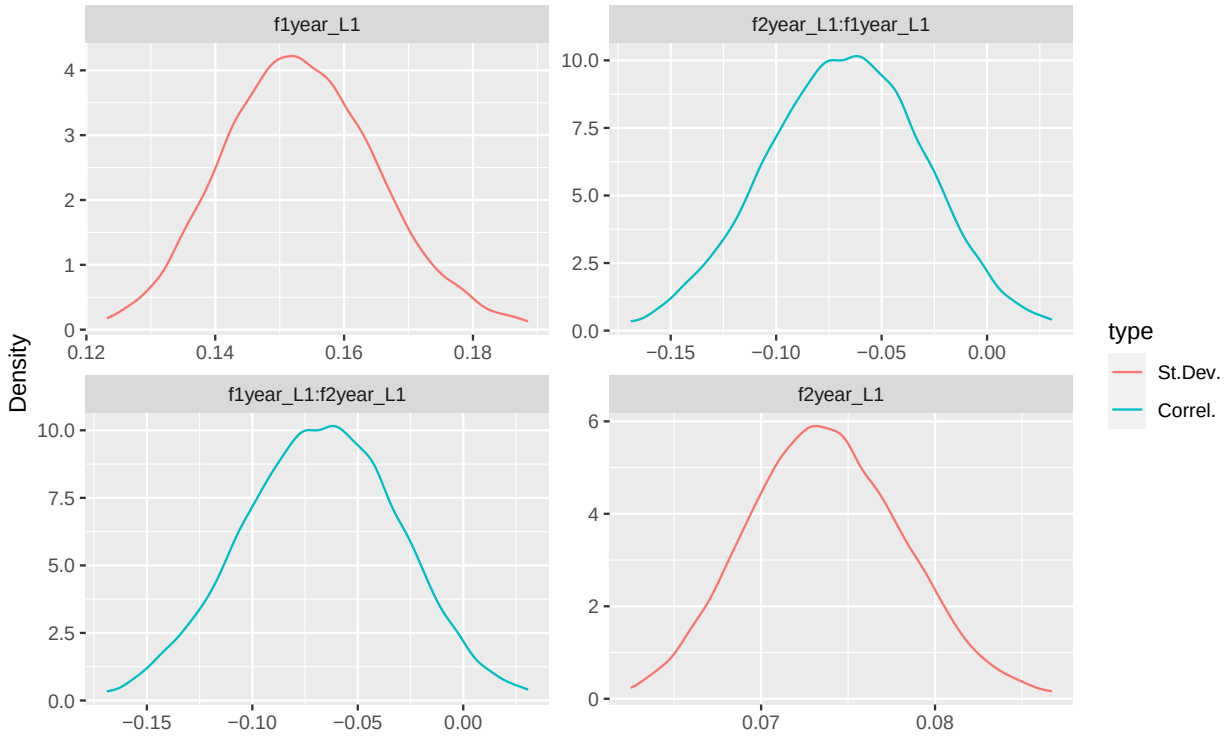
```
plotM3$Outcomes$S1
```



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):

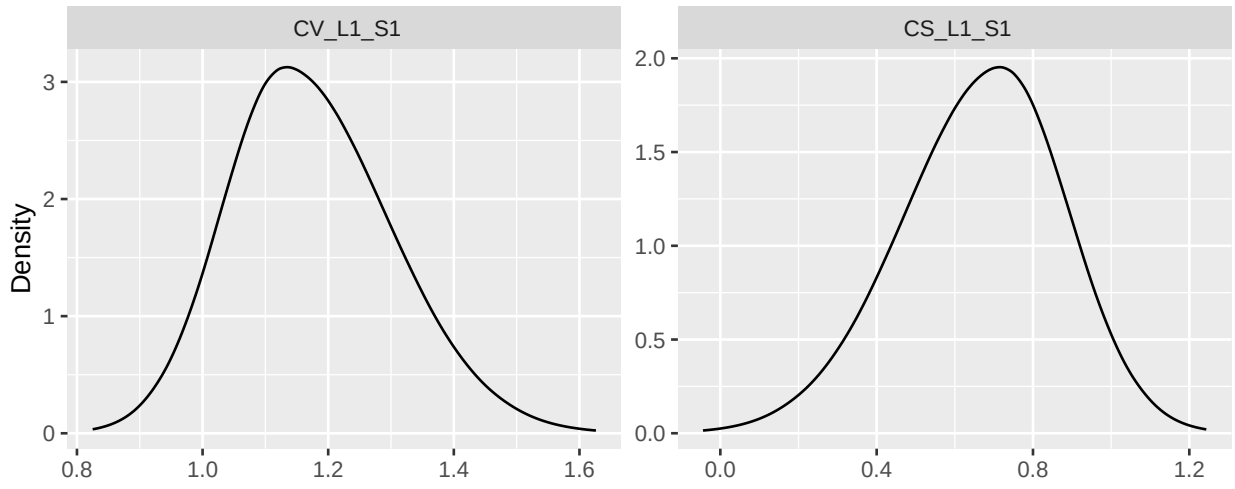
```
plotM3$Covariances
```

```
## $L1
```



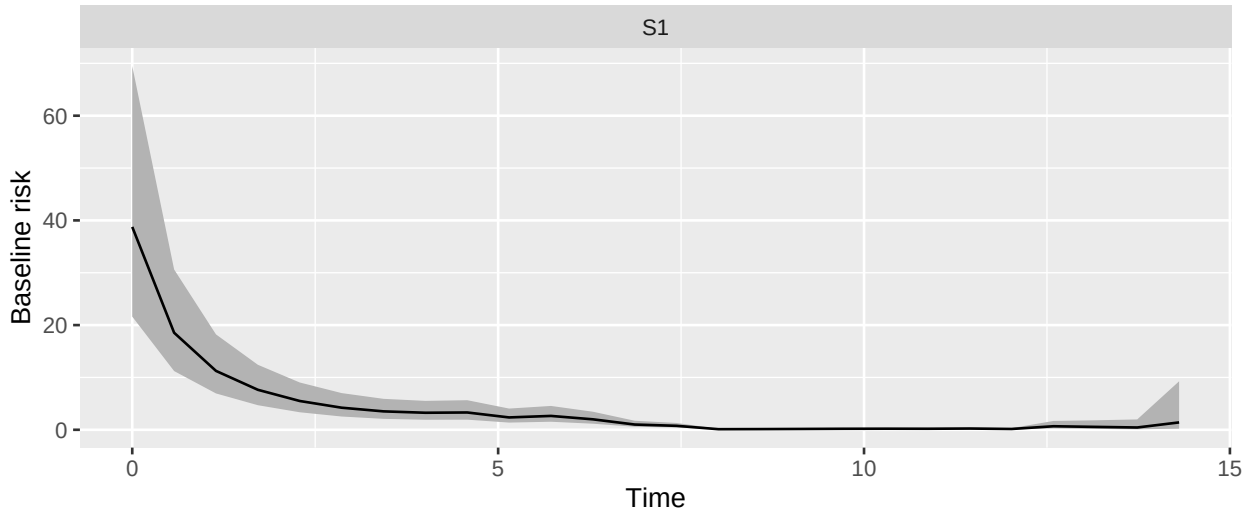
The posterior distributions of the association parameters:

```
plotM3$Associations
```



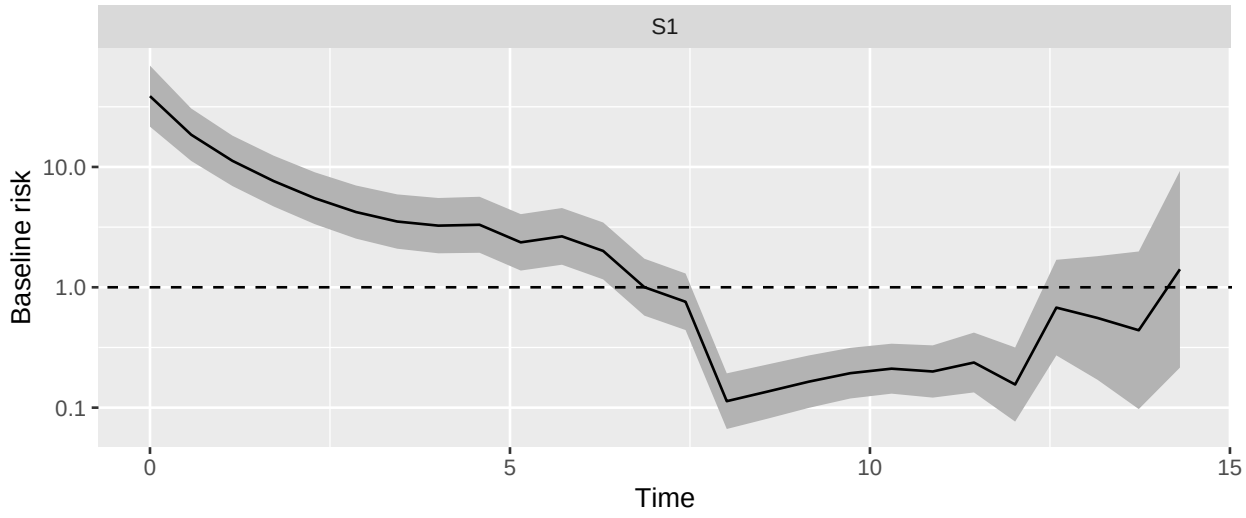
And finally the curve for the baseline risk functions:

```
plotM3$Baseline
```



The model for the baseline risk is a random walk with number of bins given by argument `NbasRisk`, the curve plotted is linear 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., `PlotM3` in the example).

```
plotM3$Baseline + scale_y_log10() + geom_abline(slope=0, intercept=0, linetype='dashed')
```



Model 4: 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:

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

```
# INLAjoint
M4 <- 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(M4)
```

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     3.4787   3.5439     3.6090
## year_L1               -0.1003 0.0119    -0.1237  -0.1003    -0.0770
## drugDpenicil_L1        0.0087 0.0468    -0.0830   0.0087     0.1004
## year:drugDpenicil_L1   0.0027 0.0165    -0.0298   0.0026     0.0351
## Res. err. (var)        0.0957 0.0036     0.0890   0.0957     0.1030
##
## Random effects variance-covariance (L1)
##                          mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L1           0.1325 0.0138     0.1075   0.1318     0.1613
## year_L1                0.0114 0.0014     0.0090   0.0114     0.0144
## Intercept_L1:year_L1 -0.0003 0.0038    -0.0078  -0.0002     0.0069
##
## Survival outcome
##                     mean     sd 0.025quant 0.5quant 0.975quant
## Baseline (var)    0.0259 0.0334     0.0013   0.0150     0.1189
## Intercept_S1      7.1930 0.3951     6.4401   7.1863     7.9784
## sexfemale_S1     -0.7161 0.2524    -1.2111  -0.7161    -0.2213
## drugDpenicil_S1   0.0051 0.1946    -0.3766   0.0052     0.3868
##
## Association longitudinal - survival
##             mean   sd 0.025quant 0.5quant 0.975quant
## CV_L1_S1 -3.1619 0.25    -3.6807  -3.1522    -2.6963
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                              -10409.50                               -10409.72
##
## Deviance Information Criterion:  -12331.69
## Widely applicable Bayesian information criterion:  32443.11
## Computation time: 6.66 seconds
```

```
# JMBayes
library(JMbayes)
M4JMB_lme <- lme(albumin ~ (1 + year)*drug,
```

```
                  random = ~ 1 + year |id, data = Longi)
M4JMB_cox <- coxph(Surv(Surv$years, Surv$death) ~ sex + drug,
                   data = Surv, x = TRUE)
JMpr = list(priorMean.alphas=0, priorTau.alphas = matrix(0.16))
M4JMB <- jointModelBayes(M4JMB_lme, M4JMB_cox, timeVar = "year", priors=JMpr)
# Computation time in the table includes LME + Cox + JM
Summary(M4JMB)
```

```
##
## 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())
M4rstanarm <- 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(M4rstanarm)
```

```
## 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	7 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 5: 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:

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

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

M5 <- 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(M5)
```

```
## Longitudinal outcome (lognormal)
##                          mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L1           0.7705 0.1780     0.4217   0.7705     1.1194
## year_L1                0.1851 0.0366     0.1134   0.1851     0.2568
## drugDpenicil_L1       -0.1200 0.1072    -0.3301  -0.1200     0.0901
## sexfemale_L1          -0.2548 0.1774    -0.6026  -0.2548     0.0930
## year:drugDpenicil_L1   0.0138 0.0223    -0.0299   0.0138     0.0574
## year:sexfemale_L1     -0.0396 0.0359    -0.1100  -0.0396     0.0309
## Res. err. (var)        0.1092 0.0042     0.1011   0.1091     0.1178
##
## Random effects variance-covariance (L1)
##                         mean     sd 0.025quant 0.5quant 0.975quant
```

```
## Intercept_L1         0.9998 0.0851     0.8465   0.9946     1.1784
## year_L1              0.0343 0.0042     0.0268   0.0340     0.0434
## Intercept_L1:year_L1 0.0494 0.0169     0.0165   0.0491     0.0825
##
## Survival outcome (S1)
##                       mean     sd 0.025quant 0.5quant 0.975quant
## Baseline_S1 (var)  0.1849 0.0979     0.0569   0.1646     0.4344
## Intercept_S1      -1.6132 0.3843    -2.3665  -1.6132    -0.8599
## sexfemale_S1      -0.7762 0.3476    -1.4574  -0.7762    -0.0949
## drugDpenicil_S1   -0.0668 0.2351    -0.5275  -0.0668     0.3939
##
## Survival outcome (S2)
##                                          mean     sd 0.025quant 0.5quant 0.975quant
## Baseline_S2 (var)                      0.4268 0.2430     0.1450   0.3643     1.0756
## Intercept_S2                          -5.0677 1.0106    -7.0484  -5.0677    -3.0870
## edemaedema.no.diuretics_S2             0.0581 1.3664    -2.6200   0.0581     2.7362
## edemaedema.despite.diuretics_S2        1.0470 1.3778    -1.6534   1.0470     3.7474
## sexfemale_S2                           0.8444 1.0195    -1.1537   0.8444     2.8425
## edemaedema.no.diuretics:sexfemale_S2  -0.1324 1.4338    -2.9425  -0.1324     2.6777
## edemaedema.despite.diuretics:sexfemale_S2 -2.4152 1.5154    -5.3854  -2.4152     0.5550
##
## Association longitudinal - survival
##                        mean     sd 0.025quant 0.5quant 0.975quant
## SRE_L1_S1            1.1339 0.0837     0.9658   1.1352     1.2956
## SRE_Intercept_L1_S2 1.0047 0.2176     0.5586   1.0113     1.4156
## SRE_year_L1_S2      5.9478 1.0544     3.6841   6.0348     7.7885
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                              -12095.77                              -12087.21
##
## Deviance Information Criterion:  -13266.68
## Widely applicable Bayesian information criterion:  -13903.94
## Computation time: 30.14 seconds
```

Model 6: 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:

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

```
M6 <- 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(M6)
```

```
## Longitudinal outcome (L1, lognormal)
##                           mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L1            1.0017 0.1697     0.6690   1.0017     1.3344
## year_L1                 0.0815 0.0060     0.0697   0.0815     0.0932
## drugDpenicil_L1        -0.1644 0.1245    -0.4084  -0.1644     0.0796
## sexfemale_L1           -0.3881 0.1604    -0.7025  -0.3881    -0.0736
## year:drugDpenicil_L1    0.0149 0.0084    -0.0015   0.0149     0.0312
## Res. err. (var)         0.2148 0.0076     0.2002   0.2147     0.2301
##
## Longitudinal outcome (L2, poisson)
##                      mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L2       5.4149 0.0677     5.2823   5.4149     5.5475
## 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 -0.0724 0.0449    -0.1603  -0.0724     0.0155
## sexfemale_L2       0.1068 0.0661    -0.0228   0.1068     0.2363
##
## Longitudinal outcome (L3, gaussian)
##                      mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L3       3.5191 0.0343     3.4517   3.5191     3.5864
## year_L3           -0.0623 0.0153    -0.0924  -0.0623    -0.0322
## f1year_L3         -0.0061 0.0034    -0.0128  -0.0061     0.0006
## f2year_L3          0.0005 0.0002     0.0001   0.0005     0.0009
## drugDpenicil_L3    0.0088 0.0455    -0.0803   0.0088     0.0979
## Res. err. (var)    0.1115 0.0039     0.1039   0.1115     0.1194
##
## Random effects variance-covariance
##                                mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L1                 1.1419 0.0988     0.9630   1.1372     1.3469
## Intercept_L2                 0.1573 0.0128     0.1341   0.1566     0.1846
## Intercept_L3                 0.1361 0.0134     0.1125   0.1352     0.1651
## Intercept_L1:Intercept_L2 -0.1117 0.0259    -0.1654  -0.1107    -0.0631
## Intercept_L1:Intercept_L3 -0.2487 0.0300    -0.3128  -0.2476    -0.1942
## Intercept_L2:Intercept_L3  0.0521 0.0099     0.0335   0.0518     0.0725
##
## Survival outcome (S1)
##                        mean     sd 0.025quant 0.5quant 0.975quant
## Baseline_S1 (var)    0.0142 0.0133     0.0008   0.0102     0.0498
## Intercept_S1         4.5400 0.1459     4.2541   4.5400     4.8260
## drugDpenicil_S1     -0.0102 0.1844    -0.3717  -0.0102     0.3513
##
## Survival outcome (S2)
##                        mean     sd 0.025quant 0.5quant 0.975quant
## Baseline_S2 (var)    0.0063 0.0101     0.0000   0.0025     0.0349
## Intercept_S2        -5.1632 0.2507    -5.6545  -5.1632    -4.6719
## drugDpenicil_S2     -0.3781 0.3785    -1.1199  -0.3781     0.3636
```

```
##
## Association longitudinal - survival
##               mean     sd 0.025quant 0.5quant 0.975quant
## CV_L1_S1    0.9733 0.0935     0.7716   0.9793     1.1384
## CV_L1_S2    1.0515 0.1744     0.6841   1.0601     1.3709
## SRE_L2_S1 -0.3547 0.2103    -0.7885  -0.3474     0.0397
## CV_L3_S1  -2.6018 0.3619    -3.3057  -2.6057    -1.8780
## CS_L3_S1   1.4442 3.9169    -6.0383   1.3544     9.3856
## CS_L3_S2   0.2464 3.7198    -7.3977   0.3701     7.2430
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                              -56524.59                               -56509.60
##
## Deviance Information Criterion:  -42356.61
## Widely applicable Bayesian information criterion:  -6756.457
## Computation time: 88.41 seconds
```

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

Model 7: 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).

```
# 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

M7 <- joint(formSurv=list(E12 ~ X, E13 ~ X, E23 ~ X),
            basRisk = c("rw2", "rw2", "rw2"), dataSurv = SurvMS)
summary(M7)
```

```
##
## Survival outcome (S1)
##                       mean     sd 0.025quant 0.5quant 0.975quant
## Baseline_S1 (var)   0.4320 0.2521     0.1225   0.3735     1.0908
## Intercept_S1       -1.7983 0.4084    -2.6113  -1.7941    -1.0088
## X_S1                0.3067 0.0875     0.1354   0.3066     0.4786
##
## Survival outcome (S2)
##                       mean     sd 0.025quant 0.5quant 0.975quant
```

```
## Baseline_S2 (var)  0.2867 0.2819     0.0217   0.2014     1.0613
## Intercept_S2      -4.9414 1.1050    -7.3275  -4.8601    -2.9914
## X_S2               0.6671 0.1152     0.4423   0.6667     0.8941
##
## Survival outcome (S3)
##                       mean     sd 0.025quant 0.5quant 0.975quant
## Baseline_S3 (var)  0.0218 0.0507     0.0001   0.0057     0.1495
## Intercept_S3      -2.4669 0.5926    -3.6315  -2.4662    -1.3066
## X_S3              -0.1452 0.0915    -0.3246  -0.1452     0.0341
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                              -272.9119                              -272.7587
##
## Deviance Information Criterion:  387.2479
## Widely applicable Bayesian information criterion:  386.4479
## Computation time: 1.23 seconds
```

Model 8: 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
M8 <- joint(formSurv=list(E12 ~ X, E13 ~ X, E23 ~ X),
            formLong=list(y ~ time + X + (1+time|id)),
            basRisk = c("rw2", "rw2", "rw2"), timeVar = "time",
            assoc = list(c("CV", "CV", "CV")), id="id",
            dataSurv = SurvMS, dataLong = LongMS)
summary(M8)
```

```
## Longitudinal outcome (gaussian)
##                     mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L1    -0.4715 0.2198    -0.9023  -0.4717    -0.0399
## time_L1          0.1008 0.1202    -0.1344   0.1005     0.3373
## X_L1             1.5168 0.0363     1.4455   1.5168     1.5880
## Res. err. (var)  1.5515 0.1406     1.2978   1.5436     1.8502
##
## Random effects variance-covariance (L1)
##                          mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L1          0.2600 0.2721     0.0555   0.1877     0.9027
## time_L1               0.4007 0.1060     0.2341   0.3876     0.6473
## Intercept_L1:time_L1 -0.0264 0.1439    -0.3895   0.0055     0.1604
##
## Survival outcome (S1)
##                       mean     sd 0.025quant 0.5quant 0.975quant
## Baseline_S1 (var)  0.0445 0.0622     0.0029   0.0243     0.2152
## Intercept_S1       0.9145 0.3760     0.1766   0.9146     1.6517
```

```
## X_S1              -0.6641 0.1120    -0.8841  -0.6638    -0.4455
##
## Survival outcome (S2)
##                       mean     sd 0.025quant 0.5quant 0.975quant
## Baseline_S2 (var)  0.0377 0.0704     0.0007   0.0150     0.2233
## Intercept_S2      -2.4468 0.8119    -4.0406  -2.4468    -0.8537
## X_S2              -0.2038 0.1475    -0.4929  -0.2032     0.0827
##
## Survival outcome (S3)
##                       mean     sd 0.025quant 0.5quant 0.975quant
## Baseline_S3 (var)  0.0261 0.0500     0.0002   0.0091     0.1609
## Intercept_S3      -2.8648 0.5968    -4.0347  -2.8650    -1.6937
## X_S3              -0.4411 0.1047    -0.6463  -0.4411    -0.2356
##
## Association longitudinal - survival
##            mean     sd 0.025quant 0.5quant 0.975quant
## CV_L1_S1 0.6337 0.1565     0.3211   0.6353     0.9380
## CV_L1_S2 0.5661 0.2042     0.1504   0.5709     0.9560
## CV_L1_S3 0.1558 0.0788    -0.0363   0.1558     0.2792
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                              -2433.993                               -2435.065
##
## Deviance Information Criterion:  -1705.902
## Widely applicable Bayesian information criterion:  -1818.489
## Computation time: 5.11 seconds
```

Model 9: model from application section of https://arxiv.org/abs/2203.06256

The model structure is given by the following equation:

$$
\begin{cases}
\log(serBilir_{ij}) &= \eta_{i1}(t_{ij}) + \varepsilon_{ij1} \\
&= (\beta_{10} + b_{i10}) + \beta_{11}X_i + (\beta_{12} + b_{i11})\mathrm{NS}_1(t_{ij}) + (\beta_{13} + b_{i12})\mathrm{NS}_2(t_{ij}) \\
&\quad +(\beta_{14} + b_{i13})\mathrm{NS}_3(t_{ij}) + \beta_{15}X_i\mathrm{NS}_1(t_{ij}) + \beta_{16}X_i\mathrm{NS}_2(t_{ij}) + \beta_{17}X_i\mathrm{NS}_3(t_{ij}) + \varepsilon_{ij1} \\
\log(SGOT_{ij}) &= \eta_{i2}(t_{ij}) + \varepsilon_{ij2} \\
&= (\beta_{20} + b_{i20}) + \beta_{21}X_i + (\beta_{22} + b_{i21})\mathrm{NS}_1(t_{ij}) + (\beta_{23} + b_{i22})\mathrm{NS}_2(t_{ij}) \\
&\quad +(\beta_{24} + b_{i23})\mathrm{NS}_3(t_{ij}) + \beta_{25}X_i\mathrm{NS}_1(t_{ij}) + \beta_{26}X_i\mathrm{NS}_2(t_{ij}) + \beta_{27}X_i\mathrm{NS}_3(t_{ij}) + \varepsilon_{ij2} \\
albumin_{ij} &= \eta_{i3}(t_{ij}) + \varepsilon_{ij3} \\
&= (\beta_{30} + b_{i30}) + \beta_{31}X_i + (\beta_{32} + b_{i31})t + \beta_{33}X_it + \varepsilon_{ij3} \\
\log(\mathrm{E}[platelets_{ij}]) &= \eta_{i4}(t_{ij}) \\
&= (\beta_{40} + b_{i40}) + \beta_{41}X_i + (\beta_{42} + b_{i41})\mathrm{NS}_1(t_{ij}) + (\beta_{43} + b_{i42})\mathrm{NS}_2(t_{ij}) \\
&\quad +(\beta_{44} + b_{i43})\mathrm{NS}_3(t_{ij}) + \beta_{45}X_i\mathrm{NS}_1(t_{ij}) + \beta_{46}X_i\mathrm{NS}_2(t_{ij}) + \beta_{47}X_i\mathrm{NS}_3(t_{ij}) \\
\mathrm{logit}(\mathrm{E}[spiders_{ij}]) &= \eta_{i5}(t_{ij}) \\
&= (\beta_{50} + b_{i50}) + \beta_{51}X_i + (\beta_{52} + b_{i51})t + \beta_{53}X_it \\
\lambda_{i1}(t) &= \lambda_{01}(t)\,\exp\left(\gamma_1 X_i + \varphi_1\eta_{i1}(t) + \varphi_3\eta'_{i1}(t) + \varphi_4\eta_{i2}(t)\right. \\
&\quad \left. + \varphi_5\eta_{i3}(t) + \varphi_7\eta_{i4}(t) + \varphi_9\eta_{i5}(t)\right) \\
\lambda_{i2}(t) &= \lambda_{02}(t)\,\exp\left(\gamma_2 X_i + \varphi_2\eta_{i1}(t) + \varphi_6\eta_{i3}(t) + \varphi_8\eta_{i4}(t)\right)
\end{cases}
\tag{L1}\tag{L2}\tag{L3}\tag{L4}\tag{L5}\tag{S1}\tag{S2}
$$

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.

```
# 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]

M9 <- joint(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)),
            formSurv = list(DTH ~ drug, TSP ~ drug),
            dataLong = Longi, id = "id", timeVar = "year", basRisk = c("rw2", "rw1"),
            family = c("lognormal", "lognormal", "gaussian", "poisson", "binomial"),
            assoc = list(c("CV_CS", "CV"),  c("CV", ""), c("CV", "CV"),
                         c("CV", "CV"), c("CV", "")),
            control=list(priorFixed=list(mean=0, prec=1,
                         mean.intercept=0, prec.intercept=1),
                         priorAssoc=list(mean=0, prec=1),
                         priorRandom=list(r=10, R=1), int.strategy="eb"))
summary(M9)
```

```
## Longitudinal outcome (L1, lognormal)
##                             mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L1              0.5836 0.0808     0.4252   0.5836     0.7420
## f1year_L1                 1.0639 0.1424     0.7848   1.0639     1.3430
## f2year_L1                 1.5986 0.1646     1.2761   1.5986     1.9212
## f3year_L1                 1.4710 0.2052     1.0689   1.4710     1.8732
## drugDpenicil_L1          -0.0955 0.1133    -0.3176  -0.0955     0.1267
## f1year:drugDpenicil_L1    0.1581 0.1972    -0.2284   0.1581     0.5446
## f2year:drugDpenicil_L1 -0.1645 0.2290    -0.6133  -0.1645     0.2843
## f3year:drugDpenicil_L1 -0.1132 0.2843    -0.6704  -0.1132     0.4440
## Res. err. (var)           0.0779 0.0023     0.0735   0.0779     0.0824
##
## Random effects variance-covariance (L1)
##                              mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L1              1.0580 0.1659     0.8488   1.0291     1.4401
## f1year_L1                 1.5976 0.3073     1.0889   1.5662     2.2901
## f2year_L1                 2.2067 0.4578     1.5529   2.1237     3.3457
## f3year_L1                 1.5252 0.4165     0.8686   1.4736     2.4961
## Intercept_L1:f1year_L1 0.3233 0.1843    -0.0110   0.3141     0.7140
## Intercept_L1:f2year_L1 0.5399 0.2256     0.1835   0.5149     1.0662
## Intercept_L1:f3year_L1 0.4301 0.2808    -0.0374   0.4011     1.0695
## f1year_L1:f2year_L1       1.6844 0.2972     1.1717   1.6631     2.3217
## f1year_L1:f3year_L1       0.7412 0.3204     0.1104   0.7406     1.3790
## f2year_L1:f3year_L1       1.1095 0.4238     0.4590   1.0502     2.1086
##
## Longitudinal outcome (L2, lognormal)
##                             mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L2              4.7479 0.0368     4.6759   4.7479     4.8200
```

```
## f1year_L2              -0.1354 0.0782    -0.2888  -0.1354     0.0179
## f2year_L2               0.0827 0.0935    -0.1004   0.0827     0.2659
## f3year_L2              -0.0074 0.1260    -0.2544  -0.0074     0.2397
## drugDpenicil_L2        -0.0794 0.0516    -0.1807  -0.0794     0.0218
## f1year:drugDpenicil_L2  0.1050 0.1087    -0.1080   0.1050     0.3180
## f2year:drugDpenicil_L2 -0.1901 0.1303    -0.4455  -0.1901     0.0653
## f3year:drugDpenicil_L2  0.0427 0.1738    -0.2979   0.0427     0.3833
## Res. err. (var)         0.0682 0.0021     0.0643   0.0682     0.0725
##
## Random effects variance-covariance (L2)
##                             mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L2            0.1655 0.0197     0.1323   0.1638     0.2091
## f1year_L2               0.2101 0.0480     0.1323   0.2048     0.3199
## f2year_L2               0.3065 0.0789     0.1933   0.2934     0.5001
## f3year_L2               0.2273 0.0762     0.1127   0.2157     0.4085
## Intercept_L2:f1year_L2 -0.0173 0.0239    -0.0662  -0.0169     0.0290
## Intercept_L2:f2year_L2  0.0219 0.0302    -0.0397   0.0220     0.0808
## Intercept_L2:f3year_L2 -0.0053 0.0325    -0.0701  -0.0054     0.0608
## f1year_L2:f2year_L2     0.1752 0.0393     0.1083   0.1718     0.2632
## f1year_L2:f3year_L2     0.0869 0.0369     0.0184   0.0850     0.1657
## f2year_L2:f3year_L2     0.1484 0.0673     0.0493   0.1372     0.3100
##
## Longitudinal outcome (L3, gaussian)
##                           mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L3          3.5463 0.0327     3.4822   3.5463     3.6103
## year_L3              -0.1023 0.0121    -0.1260  -0.1023    -0.0786
## drugDpenicil_L3       0.0025 0.0460    -0.0878   0.0025     0.0927
## year:drugDpenicil_L3  0.0027 0.0169    -0.0305   0.0027     0.0359
## Res. err. (var)       0.0959 0.0026     0.0909   0.0959     0.1011
##
## Random effects variance-covariance (L3)
##                           mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L3          0.1262 0.0132     0.1023   0.1257     0.1541
## year_L3               0.0112 0.0013     0.0088   0.0111     0.0140
## Intercept_L3:year_L3 -0.0014 0.0023    -0.0059  -0.0014     0.0033
##
## Longitudinal outcome (L4, poisson)
##                           mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L4           5.5145 0.0307     5.4543   5.5145     5.5746
## f1year_L4             -0.2124 0.1032    -0.4147  -0.2124    -0.0102
## f2year_L4             -0.8701 0.1996    -1.2612  -0.8701    -0.4790
## f3year_L4             -1.1329 0.3647    -1.8478  -1.1329    -0.4181
## drugDpenicil_L4       -0.0531 0.0431    -0.1376  -0.0531     0.0314
## f1year:drugDpenicil_L4  0.1885 0.1416    -0.0891   0.1885     0.4661
## f2year:drugDpenicil_L4 -0.4286 0.2700    -0.9579  -0.4286     0.1006
## f3year:drugDpenicil_L4 -0.5330 0.4910    -1.4954  -0.5330     0.4293
##
## Random effects variance-covariance (L4)
##                           mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L4           0.1454 0.0126     0.1225   0.1448     0.1718
## f1year_L4              1.2925 0.2229     0.9311   1.2671     1.7906
## f2year_L4              4.8087 1.0124     3.1647   4.6903     7.1357
## f3year_L4             15.6405 3.5897     9.8345  15.2080    23.7960
## Intercept_L4:f1year_L4 -0.0319 0.0364    -0.1041  -0.0318     0.0395
```

```
## Intercept_L4:f2year_L4 -0.0881 0.0767    -0.2477  -0.0874     0.0585
## Intercept_L4:f3year_L4 -0.0696 0.1431    -0.3679  -0.0677     0.2093
## f1year_L4:f2year_L4    -1.7596 0.4268    -2.7365  -1.7117    -1.0616
## f1year_L4:f3year_L4    -3.5749 0.8140    -5.4130  -3.4860    -2.2449
## f2year_L4:f3year_L4     8.2598 1.8903     5.1801   8.0463    12.6111
##
## Longitudinal outcome (L5, binomial)
##                           mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L5         -1.3207 0.2257    -1.7631  -1.3207    -0.8783
## year_L5               0.1430 0.0601     0.0253   0.1430     0.2607
## drugDpenicil_L5      -0.1666 0.3162    -0.7864  -0.1666     0.4532
## year:drugDpenicil_L5 -0.0246 0.0850    -0.1912  -0.0246     0.1420
##
## Random effects variance-covariance (L5)
##                           mean     sd 0.025quant 0.5quant 0.975quant
## Intercept_L5          5.7226 1.1255     3.9132   5.6041     8.2786
## year_L5               0.1475 0.0369     0.0885   0.1430     0.2306
## Intercept_L5:year_L5 -0.1967 0.1645    -0.5828  -0.1754     0.0621
##
## Survival outcome (S1)
##                      mean     sd 0.025quant 0.5quant 0.975quant
## Baseline_S1 (var) 0.0193 0.0051     0.0110   0.0187     0.0309
## Intercept_S1      2.0926 0.1809     1.7380   2.0926     2.4473
## drugDpenicil_S1   0.0159 0.1848    -0.3464   0.0159     0.3781
##
## Survival outcome (S2)
##                       mean     sd 0.025quant 0.5quant 0.975quant
## Baseline_S2 (var)  0.0251 0.0120     0.0091   0.0227     0.0554
## Intercept_S2      -0.2508 0.2474    -0.7357  -0.2508     0.2340
## drugDpenicil_S2   -0.3942 0.3507    -1.0815  -0.3942     0.2931
##
## Association longitudinal - survival
##             mean     sd 0.025quant 0.5quant 0.975quant
## CV_L1_S1  1.1272 0.0747     0.9764   1.1278     1.2736
## CS_L1_S1  0.9743 0.2381     0.4972   0.9721     1.4745
## CV_L1_S2  1.0766 0.1286     0.8157   1.0785     1.3265
## CV_L2_S1  0.0424 0.1641    -0.2846   0.0429     0.3661
## CV_L3_S1 -1.6279 0.1246    -1.8902  -1.6273    -1.3778
## CV_L3_S2 -1.0844 0.3522    -1.7878  -1.0815    -0.3965
## CV_L4_S1 -0.3452 0.1432    -0.6281  -0.3452    -0.0623
## CV_L4_S2 -0.3530 0.1821    -0.7085  -0.3551     0.0150
## CV_L5_S1  0.0038 0.0358    -0.0683   0.0043     0.0730
##
## log marginal-likelihood (integration)    log marginal-likelihood (Gaussian)
##                              -79549.49                               -79503.26
##
## Deviance Information Criterion:  -1528603
## Widely applicable Bayesian information criterion:  121971.2
## Computation time: 599.23 seconds
```