`BeQut`, an R-package for the **B**ayesian **e**stimation of mixed effect **Qu**an**t**ile regression models using `JAGS`

Rencontres R, Vannes, 2024

A.Barbieri, C. Tzourio and H. Jacqmin-Gadda

Univ. Bordeaux, Inserm U1219 - Biostatistic Team, France

Motivations to `BeQut` package

- Clinical motivations: explore the link between blood pressure evolution and risk to cerebro-and cardiovascular disease
  - Joint models for longitudinal and time-to-event data
    - Combining a linear mixed effect model and a proportional hazard model
    - Association structures based on features from the conditional expectation

- Interests in using quantile regression rather than classical regression based on the conditional expectation
  - Explore the entire distribution of interest
  - Interest lies in the tails of distribution

- There was no package for quantile regression joint models
  - `lqmm` package for estimating linear quantile mixed models[1]
  - Some works about quantile regression joint models[2] but no existing package

[1] Geraci, 2014;
[2] Farcomeni and Viviani, 2015; Yang et al., 2019;

Linear quantile regression

Koenker and Bassett, 1978

- General framework

$$Y_i \quad = \quad \mathbf{x}_i^\top \beta \; + \varepsilon_i \; , \quad i = 1, \ldots, n$$

with

- Y_i the response variable;
- β the p-vector of unknown parameters;
- $\mathbf{x}_i$ the p-length covariate vector including intercept;
- ε_i the random residual error.

Linear quantile regression

Koenker and Bassett, 1978

- General framework

$$Y_i = \mathbf{x}_i^\top \boldsymbol{\beta}_\tau + \varepsilon_{i\tau}, \quad i = 1, \ldots, n$$

with

  - Y_i the response variable;
  - $\boldsymbol{\beta}$ the p-vector of unknown parameters;
  - $\mathbf{x}_i$ the p-length covariate vector including intercept;
  - ε_i the random residual error.

- Objective: to fit the conditional quantile of order $\tau \in]0, 1[$ of the distribution of Y_i given X_i denoted by $q_\tau (Y_i|X_i) = x_i^\top \boldsymbol{\beta}_\tau$

- Optimization problem
  - Based on the quantile loss function $\rho_\tau (u) = u (\tau - \mathbb{1}_{u<0})$

$$\widehat{\beta}_\tau = \underset{\beta \in \mathbb{R}^p}{\operatorname{argmin}} \sum_{i=1}^{n} \rho_\tau \left(y_i - \mathbf{x}_i^\top \boldsymbol{\beta}\right). \qquad (1)$$

Asymmetric Laplace distribution

Koenker and Machado, 1999; Yu and Moyeed, 2001; Yu and Zhang, 2005.

- Inference naturally based on Asymmetric Laplace (AL) distribution
  - Denoted by $Y \sim \mathcal{AL}(\mu, \sigma, \tau)$ with $\mu \in \mathbb{R}$ a location parameter, $\sigma \in \mathbb{R}^+$ a scale parameter and $\tau \in]0, 1[$ a skewness parameter

$$f_Y(y \,|\mu\,, \sigma, \tau) = \frac{\tau(1-\tau)}{\sigma} \exp\left\{-\frac{1}{\sigma} \rho_\tau(y - \mu)\right\},$$

- Why this distribution?
  - Direct link between μ and τ

$$\Pr(Y \leq \mu) = \int_{-\infty}^{\mu} f_Y(y|\mu, \sigma, \tau)\, dy = \tau$$

Asymmetric Laplace distribution

Koenker and Machado, 1999; Yu and Moyeed, 2001; Yu and Zhang, 2005.

- Inference naturally based on Asymmetric Laplace (AL) distribution
  - Denoted by $Y \sim \mathcal{AL}(\mu, \sigma, \tau)$ with $\mu \in \mathbb{R}$ a location parameter, $\sigma \in \mathbb{R}^+$ a scale parameter and $\tau \in]0,1[$ a skewness parameter

$$L(\mathbf{y}; \theta) = \prod_{i=1}^{n} f_Y(y_i | \mu_i, \sigma, \tau) = \left(\frac{\tau(1-\tau)}{\sigma}\right)^n \exp\left\{-\frac{1}{\sigma}\sum_{i=1}^{n} \rho_\tau(y_i - \mu_i)\right\},$$

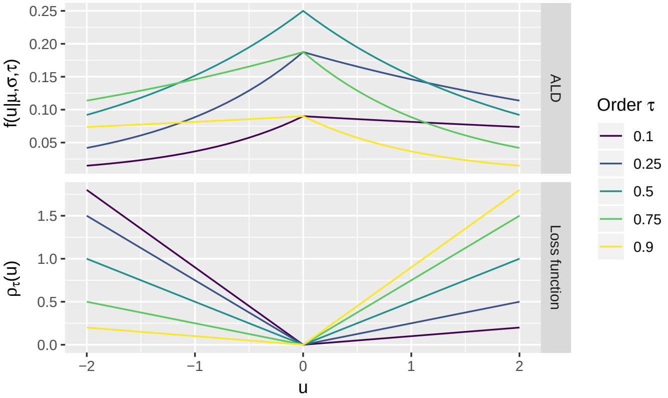
- Why this distribution?
  - Direct link between μ and τ

$$\Pr(Y \leq \mu) = \int_{-\infty}^{\mu} f_Y(y|\mu, \sigma, \tau)\, dy = \tau$$

  - Get $\widehat{\beta}_\tau$ by maximizing the likelihood $L(\mathbf{y}; \theta)$ is equivalent to solving (1) with $\theta = (\boldsymbol{\beta}_\tau{}^\top, \sigma)^\top$ the vector of parameters given $\mu_i = \mathbf{x}_i^\top \boldsymbol{\beta}_\tau$

$\mathcal{AL}$ distribution and quantile regression lost function

- The $\mathcal{AL}$ density distribution $f\left(u|\mu=0,\sigma=1,\tau\right)$
- The quantile loss function: $\rho_\tau\left(u\right)=u\left(\tau-\mathbb{1}_{u<0}\right)$
  - ρ_τ corresponds to the absolute cost function when $\tau=0.5$ (i.e. median);



Alternative of the $\mathcal{AL}$ distribution

Kotz et al., 2012

- Alternative based on Gaussian distribution mixture
  - Given $Y|\mu,\sigma,\tau \sim \mathcal{AL}(\mu,\sigma,\tau)$, an alternative (joint) distribution is from

$$f_Y(y|\mu,\sigma,\tau) = \int_{\mathbb{R}^+} f_{Y|W}(y|W=w,\mu,\sigma,\tau)\, f_W(w|\sigma)\, dw$$

where

$$\begin{cases} Y|W=w,\mu,\sigma,\tau & \sim \mathcal{N}\Big(\mu + c_1(\tau)\, w,\ c_2(\tau)\,\sigma w\Big) \\ W|\sigma & \sim \mathcal{E}xp\left(\frac{1}{\sigma}\right) \end{cases}$$

given

- ⋆ $c_1(\tau) = \frac{1-2\tau}{\tau(1-\tau)}$ and $c_2(\tau) = \frac{2}{\tau(1-\tau)}$ two constants
- ⋆ W a random variable exponentially distributed such as $\mathbb{E}(W) = \sigma$

\+ Classical continuous density functions

\- W is a latent random variable specific for each measurement

Estimation with BeQut

Some details about the estimation

- Use the rewriting of the asymmetrical Laplace distribution
- Based on Bayesian inference
  - Considering Gaussian and Inverse-Gamma vague priors for $\boldsymbol{\beta}_\tau$ and σ, respectively

$$\pi\left(\boldsymbol{\beta}_\tau, \sigma, \boldsymbol{w} | \boldsymbol{y}\right) \propto \prod_{i=1}^{n} f\left(y_i | w_{i,\tau}, \boldsymbol{\beta}_\tau, \sigma\right) f\left(w_{i,\tau} | \sigma\right) \pi\left(\boldsymbol{\beta}_\tau\right) \pi\left(\sigma\right)$$

- Use the JAGS software to generate posterior samples of parameters
  - Via `rjags` and `jagsUI` packages from R
  - Need data and the model specification (likelihood and prior) in a .txt file[3]
  - Convergence assessed using Gelman-Rubin criterion and parameter trace plots

[3] CRAN comment: be careful not to write to the user's computer;

Example of `lqm` using `wave` data (1)

Estimation function

- Illustration of `lqm` function on `wave` data
  - Environmental measurements around Bordeaux from CANDHIS database and data from InfoClimat website

```
#---- Use data
data(wave, package = "BeQut")

#---- Fit regression model for the first quartile
lqm_025 <- lqm(formula = h110d ~ vent_vit_moy,
               tau = 0.25,
               data = wave,
               n.chains = 3,
               n.iter = 1500,
               n.burnin = 500,
               n.thin = 1,
               n.adapt = NULL,
               save_jagsUI = TRUE,
               parallel = FALSE)
```

Example of lqm using wave data (2)

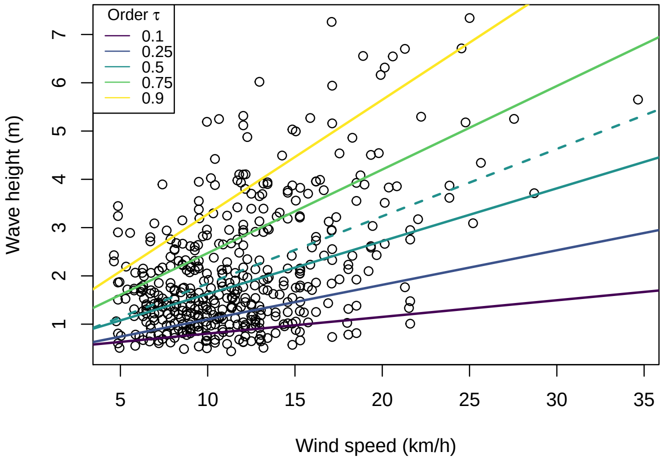
Summary

```
> #---- Summary of output
> summary(lqm_025)

#-- Statistical model: Linear quantile model
     - Quantile order:  0.25
     - Number of observations:  453

#-- Estimation of linear predictor parameters and their CI bounds:
                  Value       2.5%     97.5%     Rhat
(Intercept)  0.35652367 0.06220305 0.6436728 1.030245
vent_vit_moy 0.07300024 0.04679899 0.1004022 1.043209

#-- Estimation of 'sigma' parameter associated with AL distribution:
         Value      2.5%     97.5%     Rhat
sigma 0.298787 0.2736124 0.3266074 1.001231
```

Example of `lqm` using `wave` data (3)

Plots

Linear tend



Linear quantile mixed model, `lqmm`

`lqmm` function to fit longitudinal data

$$Y_i(t_{ij}) = \underbrace{\mathbf{x}_i(t_{ij})^\top \boldsymbol{\beta}_\tau + \mathbf{u}_i(t_{ij})^\top \mathbf{b}_{i,\tau}}_{=q_{i,\tau}(t_{ij}|\mathbf{b}_{i,\tau})} + \varepsilon_{i,\tau}(t_{ij})$$

with

- $Y_i(t_{ij})$ is the response variable for repeated measure $j(j = 1, \dots, n_i)$ of the subject $i(i = 1, \dots, n)$
- $q_{i,\tau}(t_{ij}|\mathbf{b}_{i,\tau})$ is the individual quantile with order τ for subject i at time t_{ij}
- $\boldsymbol{\beta}_\tau$ is the vector of fixed parameters
- $\mathbf{b}_{i,\tau}$ is the subject-specific random effect vector, with $\mathbf{b}_{i,\tau} \sim \mathcal{N}(\mathbf{0}, \Sigma_b)$
- $\varepsilon_{i,\tau}(t_{ij}) \sim \mathcal{AL}(0, \sigma, \tau)$ is the residual error

- Some details on the distributions

$$\begin{cases} Y_{ij}|w_{ij}, \mathbf{b}_i & \sim \quad \mathcal{N}\Big(q_{ij,\tau} + c_1(\tau)\, w_{ij},\ c_2(\tau)\,\sigma w_{ij}\Big) \\ W_{ij}|\sigma_\tau & \sim \quad \mathcal{E}xp\left(\frac{1}{\sigma}\right) \quad \text{with} \quad \mathbb{E}(W_{ij}) = \sigma \\ \mathbf{b}_i|\Sigma_b & \sim \quad \mathcal{N}(0, \Sigma_b) \end{cases}$$

Quantile regression joint model, `qrjm`

`qrjm` function to fit both longitudinal and time-to-event data

- Joint model combining lqmm and proportional hazard model

$$\begin{cases} Y_i(t) = \mathbf{x}_i(t)^\top \boldsymbol{\beta}_\tau + \mathbf{u}_i(t)^\top \mathbf{b}_{i,\tau} + \varepsilon_{i,\tau}(t) \\ h_i(t|\mathbf{b}_{i,\tau}) = h_0(t)\exp\left\{z_i^\top \alpha_z + \alpha_q q_{i,\tau}(t|\mathbf{b}_{i,\tau})\right\} \end{cases},$$

with

- $h_0(t)$ is the Weibull baseline hazard function
- z_i the vector of baseline time-independent-covariates associated with parameter vector α_z
- α_q is the parameter quantifying the impact of the current quantile on the risk to develop the event of interest
- the survival function defined by

$$S(t|\mathbf{b}_i) = \exp\left\{-\int_0^t h(u|\mathbf{b}_i)\,du\right\}.$$

Estimation with BeQut

Some details

- Bayesian estimation using JAGS

$$\pi\left(\boldsymbol{\theta}_\tau, \boldsymbol{b}, \boldsymbol{w} | \boldsymbol{y}, \boldsymbol{t}, \boldsymbol{\delta}\right) \propto \prod_{i=1}^{n} S\left(t_i | \mathbf{b}_{i,\tau}, \boldsymbol{\theta}_\tau\right) h\left(t_i | \mathbf{b}_{i,\tau}, \boldsymbol{\theta}_\tau\right)^{\delta_i}$$
$$\times \prod_{j=1}^{n_i} f\left(y_{ij} | \mathbf{b}_{i,\tau}, w_{ij}, \boldsymbol{\theta}_\tau\right) f\left(w_{ij} | \boldsymbol{\theta}_\tau\right) \phi\left(\mathbf{b}_{i,\tau} | \Sigma_b\right) \pi\left(\boldsymbol{\theta}_\tau\right)$$

  - Use Gaussian, Inverse-Gamma and Inverse-Wishart (data-driven) vague priors
  - Difference between lqmm and qrjm is the addition of the survival contribution
  - Survival contribution
    - t_i is the minimum between the event time t_i^e and the censoring time t_i^c
    - $\delta_i = \mathbb{1}\left\{t_i^e \leq t_i^c\right\}$ the censoring indicator
    - Survival function approximated using Gauss-Kronrod method
    - Distributions for survival analysis do not defined in JAGS: use the zeros-trick

Illustration using *aids* data (1)

Abrams et al, 1994

- Data available in several packages such as `joineR`, `JMbayes` concerning HIV patients
  - Longitudinal marker: CD4 lymphocyte count (5 time points)
  - Event: death (188 observed, 279 censoring)

```
#---- Load data
data(aids, package = "joineR")
#---- Fit quantile regression joint model for the first quartile
aids_qrjm_Q3 <- qrjm(formFixed = CD4 ~ obstime,
                     formRandom = ~ obstime,
                     formGroup = ~ id,
                     formSurv = Surv(time,death)~drug+gender+prevOI+AZT,
                     survMod = "weibull",
                     param = "value",
                     timeVar= "obstime",
                     tau = 0.75,
                     data = aids,
                     n.chains = 3,
                     n.iter = 45000,
                     n.thin = 5)
```

Illustration using *aids* data (2)

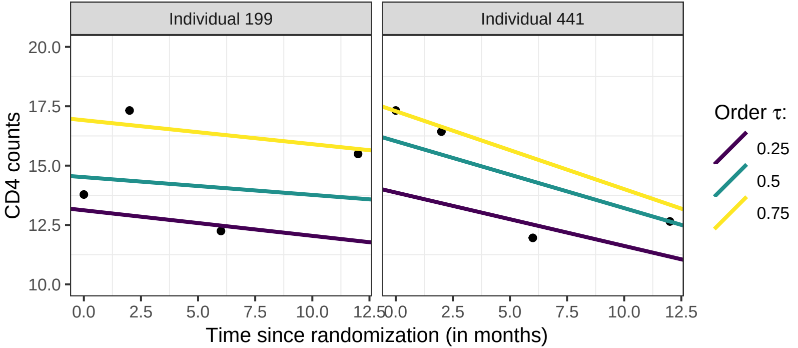
Summary

```
> summary(aids_qrjm_Q3)
#-- Statistical model: Quantile regression joint model
     - Association structure: Current quantile value of longitudinal process
     - Survival baseline risk function: Weibull distribution
     - Quantile order(s):  0.75
     - Number of observations:  1405
     - Number of statistic units (e.g. subject):  467
     - Number of observed events:  188

#-- Estimation of longitudinal regression parameters and their credible interval bounds:
                 Value       2.5%     97.5%     Rhat
(Intercept)  8.1987660  7.8704164  8.587420 1.077274
obstime     -0.1825353 -0.2124628 -0.152474 1.012812

#-- Estimation of 'sigma' parameter associated with asymmetric Laplace distribution:
          Value      2.5%    97.5%     Rhat
sigma 0.4220492 0.3919794 0.454098 1.000413

#-- (Co)variance matrix of the random-effect(s):
            (Intercept)     obstime
(Intercept)  24.3821626 -0.15381445
obstime      -0.1538145  0.03822103

#-- Estimation of survival models:
                  Value         2.5%       97.5%     Rhat
shape        1.29988842  1.12349389  1.4813052 1.001099
(Intercept) -3.55940887 -4.41639070 -2.7214120 1.002357
drugddI      0.29870655  0.01594646  0.5843527 1.001131
gendermale  -0.35184236 -0.79985255  0.1412264 1.002351
prevOIAIDS   0.76058839  0.32260612  1.2199628 1.000013
AZTfailure   0.09181464 -0.21034365  0.4028616 1.000081
alpha.assoc -0.21242456 -0.27766882 -0.1523698 1.000978
```

Focus on longitudinal submodel

Predictions of individual quantile over time

Use independent estimations of 3 quantile regression joint models



Discussion

- Quantile regression provides an interesting alternative to the classical approach
  - More robust against outliers
  - Explore the entire distribution of interest

- `BeQut` to fit mixed effect quantile regression models using `JAGS`
  - Estimation procedure validated by simulations
  - `BeQut` showed good estimation performance in comparison to other packages
  - Quantile regression joint models applied on data showed some convergence issues

- But, development in progress...
  - Improve convergence for quantile regression joint models
  - Provide a selection criterion for choosing the best model/quantile
  - Make more fexible *BeQut* package

References:

- Abrams D et al. A comparative trial of didanosine or zalcitabine after treatment with zidovudine in patients with human immunode
  ciency virus infection. *New England Journal of Medicine*, 330(10):657-662, 1994.
- Farcomeni A and Viviani S. Longitudinal quantile regression in the presence of informative dropout through longitudinal survival joint modeling. *Statistics in Medicine*, 34(7):1199-1213, 2015.
- Marco Geraci. Linear quantile mixed models: The lqmm package for laplace quantile regression. *Journal of Statistical Software*, 57(13):1-29, 2014.
- Roger Koenker. Quantile Regression. Cambridge University Press, 2005.
- Kozumi H and Kobayashi G. Gibbs sampling methods for bayesian quantile regression. 81(11):1565-1578, 2011.
- Rizopoulos D. The r package jmbayes for
  tting joint models for longitudinal and time-to-event data using mcmc. *Journal of Statistical Software*, Articles, 72(7):1-46, 2016.
- Yang M, Luo S, DeSantis S. Bayesian quantile regression joint models: Inference and dynamic predictions. *Statistical Methods in Medical Research*, 28(8):2524-2537, 2019.

Model specification

Example with *lqm*

```
model{
  # constants
  c1 <- (1-2*tau)/(tau*(1-tau))
  c2 <- 2/(tau*(1-tau))
  # Likelihood
  for (i in 1:I){
    y[i] ~ dnorm(mu[i], prec[i])
    w[i] ~ dexp(1/sigma)
    prec[i] <- 1/(sigma*c2*w[i])
    mu[i] <- inprod(beta[1:ncX], X[i, 1:ncX]) + c1*w[i]
  }#end of i loop
  # priors for parameters
  for(p in 1:ncX){
    beta[p] ~ dnorm(0, 0.001)
  }
  sigma ~ dgamma(0.001, 0.001)
}
```

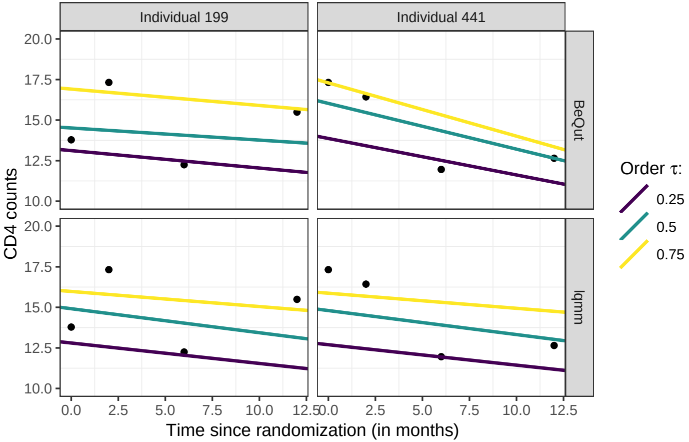
Comparison with `quantreg` and `lqmm` packages

Estimation results on `wave` data

		Parameters			
		Intercept		Slope	
Order	**Package**	Coefficient	$IC_{95\%}$	Coefficient	$IC_{95\%}$
$\tau = 0.1$	`quantreg`	0.41	[0.25 ; 0.66]	0.04	[0.01 ; 0.05]
	`lqmm`	0.44	[0.16 ; 0.71]	0.04	[0.01 ; 0.06]
	`BeQut`	0.45	[0.26 ; 0.63]	0.04	[0.02 ; 0.05]
$\tau = 0.25$	`quantreg`	0.39	[-0.03 ; 0.84]	0.07	[0.03 ; 0.11]
	`lqmm`	0.44	[0.21 ; 0.67]	0.07	[0.04 ; 0.09]
	`BeQut`	0.35	[0.06 ; 0.64]	0.07	[0.05 ; 0.10]
$\tau = 0.5$	`quantreg`	0.55	[0.32 ; 0.74]	0.11	[0.10 ; 0.13]
	`lqmm`	0.44	[0.18 ; 0.71]	0.12	[0.10 ; 0.14]
	`BeQut`	0.52	[0.27 ; 0.73]	0.11	[0.09 ; 0.13]
$\tau = 0.75$	`quantreg`	0.74	[0.51 ; 1.07]	0.17	[0.14 ; 0.20]
	`lqmm`	0.54	[0.24 ; 0.84]	0.19	[0.15 ; 0.22]
	`BeQut`	0.75	[0.48 ; 1.04]	0.17	[0.15 ; 0.20]
$\tau = 0.9$	`quantreg`	0.89	[0.41 ; 1.59]	0.24	[0.17 ; 0.29]
	`lqmm`	0.46	[0.17 ; 0.75]	0.27	[0.24 ; 0.31]
	`BeQut`	0.87	[0.50 ; 1.29]	0.24	[0.21 ; 0.27]
$\mathbb{E}(Y)$	`stat` (`lm` function)	0.45	[0.15 ; 0.74]	0.14	[0.12 ; 0.16]

Comparison between *lqmm* and *BeQut* packages

Prediction of individual conditional quantile



Estimation of joint models on `aids` data

Table: Posterior mean of the parameters and their 95% credible interval for joint model using 40000 draws

Parameters	**BeQut** $\tau = 0.25$	**BeQut** $\tau = 0.5$	**BeQut** $\tau = 0.75$	**JMbayes**
Survival				
Shape	1.28 [1.11; 1.46]	1.30 [1.13; 1.48]	1.30 [1.11; 1.49]	-
Intercept	-3.65 [-4.47;-2.85]	-3.61 [-4.47;-2.77]	-3.55 [-4.47;-2.74]	-
Drug_{ddI}	0.33 [-0.03; 0.63]	0.30 [-0.01; 0.60]	0.32 [-0.01; 0.61]	0.33 [-0.03; 0.70]
$\text{Gender}_{\text{male}}$	-0.36 [-0.83; 0.16]	-0.34 [-0.82; 0.21]	-0.32 [-0.76; 0.32]	-0.39 [-0.95; 0.18]
$\text{PrevOI}_{\text{AIDS}}$	0.64 [0.20; 1.11]	0.66 [0.21; 1.19]	0.69 [0.23; 1.17]	0.66 [0.17; 1.18]
$\text{AZT}_{\text{failure}}$	0.10 [-0.24; 0.43]	0.11 [-0.22; 0.47]	0.10 [-0.25; 0.44]	0.06 [-0.31; 0.45]
Shared association				
Current value	-0.28 [-0.37; -0.21]	-0.24 [-0.32;-0.18]	-0.21 [-0.28;-0.15]	-0.27 [-0.36;-0.19]
Longitudinal				
Intercept	6.03 [5.56 ; 6.44]	7.11 [6.74; 7.66]	8.22 [7.77; 8.64]	7.20 [6.78; 7.62]
Slope	-0.18 [-0.21;-0.15]	-0.17 [-0.20;-0.15]	-0.19 [-0.22;-0.16]	-0.22 [-0.28;-0.16]
Σ_0	19.34 [16.86; 22.24]	21.37 [18.66;24.64]	24.35 [21.32; 27.94]	21.04 [18.28;24.29]
Σ_1	0.03 [0.02; 0.04]	0.02 [0.01; 0.03]	0.04 [0.02; 0.05]	0.33 [0.28; 0.38]
Σ_{01}	-0.06 [-0.18; 0.07]	-0.05 [0.18; 0.06]	-0.15 [-0.30;-0.005]	-0.04 [-0.40; 0.19]
σ	0.44 [0.41; 0.47]	0.59 [0.55; 0.63]	0.42 [0.39; 0.46]	2.91 [1.61; 1.81]