1 Plan for Simulation Studies

Simulation studies are designed mainly for two aims. First, to check the validity of our proposed fully Bayesian algorithm in estimating the regression coefficients in the JM accurately. Our focus of the simulation results lies on the accuracy (i.e. bias) and the precision (i.e. standard deviation) of our estimations based on the posterior distributions. The second simulation study is planned to check the performance of our proposed estimator of the survival probability given in (??). This can be done by comparing our estimates with the "gold standard", which is calculated based on the true (simulated) value of random effects, fixed effects as well as the parameters.

1.1 Simulation Study 1 (Estimation Accuracy)

The first simulation study is designed to to check the accuracy of our proposed fully Bayesian estimating algorithm, which is the crucial basis for furture prediction performance. Following Farcomeni and Viviani (2014), by varying the values of the association parameters α_1 and α_2 in our model (??), we will have four different settings for simulation study, which are:

1. $(\alpha_1, \alpha_2) = (0,0)$, the two models are indepdent with each other

2. $(\alpha_1, \alpha_2) = (1,0)$, the two models are related only through the observed heterogeneity in some covarites, $\boldsymbol{H}_{it}$ in our model

3. $(\alpha_1, \alpha_2) = (0,1)$, the two models are related only through the unobserved heterogeneity, i.e. the random effects

4. $(\alpha_1, \alpha_2) = (1,1)$, the depdence of the two models is explained by both observed and unobserved heterogeneity

Under different combinations of α_1 and α_2 values, we choose the regression coefficients $\boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\gamma} = (1,1)^\top$, the covriates $\boldsymbol{Z}_{it} = (1,t)^\top$, $\boldsymbol{H}_{it} = (h_{i1}, h_{i2} * t)^\top$, $\boldsymbol{X}_{it} = (1, x_i)^\top$, and $\boldsymbol{W}_i = (w_{i1}, w_{i2})^\top$ with $h_{i1}, h_{i2}, x_i, w_{i1}$ and w_{i2} generated from independent standard normal distributions, and the random effects $\boldsymbol{u}_i$ from bivariate normal with mean 0, standard deviations equal 0.3 and correlation 0.16. We also fix $\sigma = 1$ and vary the quantile τ among $\{0.25, 0.5, 0.75\}$ for the ALD specification when simulating longitudinal data.

To simulate the survival time data, for simplicity, we fix $h_0(s) = 1$ and obtain the survival distribution as

$$S(t|\boldsymbol{u}_i, \boldsymbol{H}_{it}, \boldsymbol{W}_i) = \exp\left\{ -\frac{e^{\alpha_1(\delta_1 H_{i1} + \delta_2 H_{i2t}) + \alpha_2(u_{i1} + u_{i2}t) + \boldsymbol{\gamma}^\top \boldsymbol{W}_i} - e^{\alpha_1 \delta_1 H_{i1} + \alpha_2 u_{i1} + \boldsymbol{\gamma}^\top \boldsymbol{W}_i}}{\alpha_2 u_{i2} + \alpha_1 \delta_2 h_{i2}} \right\}$$

when $\alpha_1 \neq 0$ or $\alpha_2 \neq 0$ and

$$S(t|\boldsymbol{u}_i, \boldsymbol{H}_{it}, \boldsymbol{W}_i) = \exp\{-t e^{\boldsymbol{\gamma}^\top \boldsymbol{W}_i}\}$$

when $\alpha_1 = \alpha_2 = 0$. We then can obtain event time T_i by inverting above survival function after generating n random variates from standard uniform distribution. To obtain a censoring proportion around 25%, we choose the censoring time $C_i/5$ be distributed according to $beta(4,1)$.

To simulate the longitudinal data, we draw them independently from the ALD for the $\tau-$th quantile, centered on

$$\boldsymbol{\beta}^\top \boldsymbol{X}_{it} + \boldsymbol{\delta}^\top \boldsymbol{H}_{it} + \boldsymbol{u}_i^\top \boldsymbol{Z}_{it},$$

and with dispersion parameter σ. We keep maximum six observations for each subject at follow-up time $t = (0, 0.25, 0.5, 0.75, 1, 3)$ respectively, after incorporating the drop-out information.

τ=0.25

n	α_1	α_2	β_1 bias	s.d.	β_2 bias	s.d.	δ_1 bias	s.d.	δ_2 bias	s.d.	γ_1 bias	s.d.	γ_2 bias	s.d.	α_1 bias	s.d.	α_2 bias	s.d.
500	0	0																
500	1	0																
500	0	1																
500	1	1																

τ=0.5

n	α_1	α_2	β_1 bias	s.d.	β_2 bias	s.d.	δ_1 bias	s.d.	δ_2 bias	s.d.	γ_1 bias	s.d.	γ_2 bias	s.d.	α_1 bias	s.d.	α_2 bias	s.d.
500	0	0																
500	1	0																
500	0	1																
500	1	1																

τ=0.75

n	α_1	α_2	β_1 bias	s.d.	β_2 bias	s.d.	δ_1 bias	s.d.	δ_2 bias	s.d.	γ_1 bias	s.d.	γ_2 bias	s.d.	α_1 bias	s.d.	α_2 bias	s.d.
500	0	0																
500	1	0																
500	0	1																
500	1	1																

Table 1: Bias and standard error of the parameter estimates from proposed fully Bayesian estimating method

1.2 Simulation Study 2 (Prediction Performance)

In this simulation study, we will use the true simulated random effects, covariates and coefficients to calculate the survival probability for subject i as the "gold standard", which is given by

$$\frac{S_i[m|\mathcal{M}_i(m, u_i, \boldsymbol{\theta}; \boldsymbol{\theta})]}{S_i[t|\mathcal{M}_i(t, u_i, \boldsymbol{\theta}; \boldsymbol{\theta})]}.$$

We then use Algorithm (**??**) to computed the predicted survival probability for the same subject. By doing this for many different selected subjects, we will be able to evaluate how "close" the predicted values from our proposed model are to the true values by using Bland-Altman plot (Bland and Altman, 1986), which is a plotting method used to assess the agreement of the results between two measurement methods. We will do this simulation study by varying the follow-up time, i.e. t, in longitudinal process as well as the prediction time interval, i.e. Δt, in the model specification to mimic what is happening in the real-world studies.
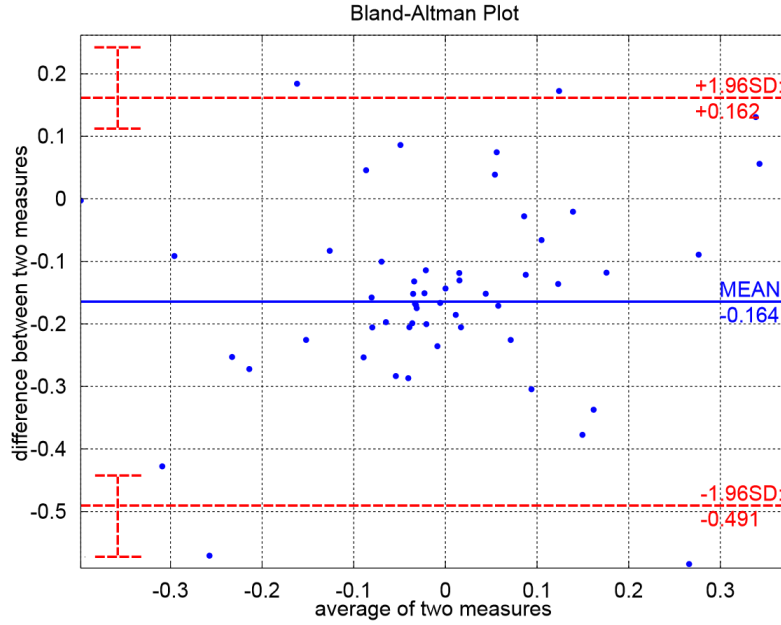

Figure 1: An example of Bland-Altman plot (from Wikipedia.org)

Table 2: Summary table of comparing the predictive results from proposed method with gold standard

	$\Delta t = 2$ bias(lower, upper)	$\Delta t = 4$ bias(lower, upper)	$\Delta t = 6$ bias(lower, upper)
$t=2$			
$t=4$			
$t=8$			
$t=16$			

References

J Martin Bland and DouglasG Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet*, 327(8476):307–310, 1986.

Alessio Farcomeni and Sara Viviani. Longitudinal quantile regression in the presence of informative dropout through longitudinal–survival joint modeling. *Statistics in medicine*, 2014.