



# Simulation report – prediction of survival probabilities

Ming Yang

July 15, 2015

## 1 Objective

To compare the predictive capabilities of survival probabilities between QRJM and JM using linear mixed model (LMJM) for data from different distribution features.

## 2 Simulation procedure

1. Define different simulation scenarios in terms of the distribution of random error and simulate 30 data sets for each scenario (see below for the specification of scenarios). Each simulated data set has 520 subjects, 500 out of which will be used to fit the model for inference purpose and the rest 20 will be used to make predictions and validation.
2. Fit the data using QRJM and LMJM respectively and save the posterior samples of the model parameters.
3. Validation data preparation: choose a time  $t$  so that all the patients selected to for prediction will only have longitudinal measurements up to this time  $t$ .
4. Make predictions of subject-specific random effects: use saved posterior samples in step 2 and longitudinal measurements from step 3 to predict subject-specific random effects for every subject in the validation samples.
5. Calculate the predictions of survival probabilities for all the subjects in validation data for some time  $u = t + \Delta t$  ( $\Delta t > 0$ ).
6. Summarize the result: make Bland-Altman plots and calculate the MSE and bias for our predictions versus the gold standard, which is calculated from the true simulated values (i.e. the random effects and the parameters).

## 3 Simulation scenarios and results

### 3.1 Scenario one

In this scenario data are generated using JM using QR. Choose the quantile to be **0.25** in the QR model.

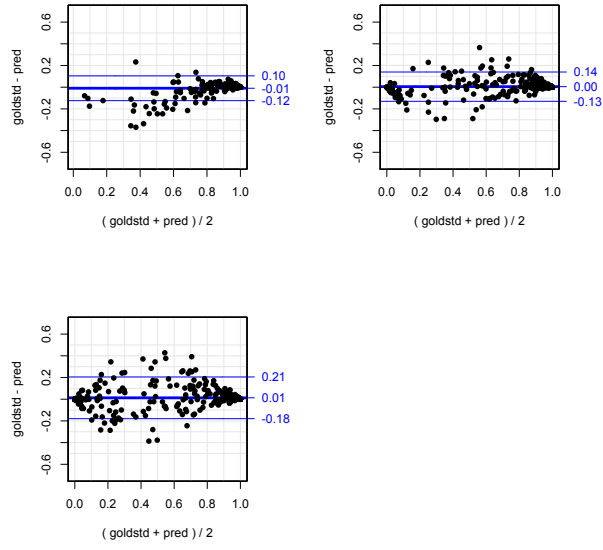


Figure 1: BA plot: data fitted with JM using linear mixed model

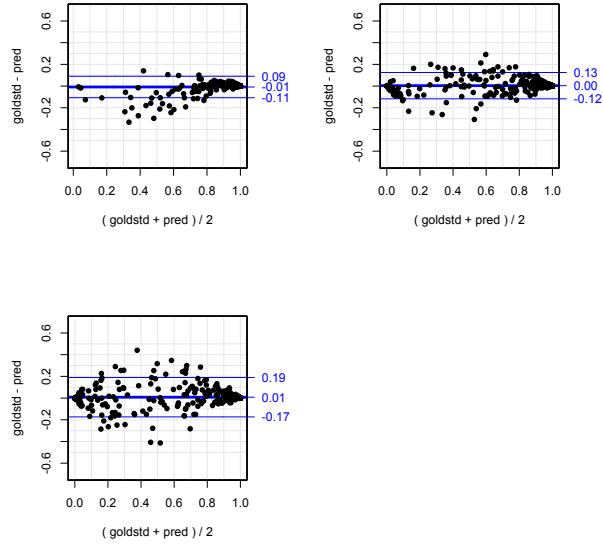


Figure 2: BA plot: data fitted with JM using QR model with  $\tau = 0.25$

### 3.2 Scenario two

In this scenario data are generated using JM using QR. Choose the quantile to be **0.5** in the QR model.

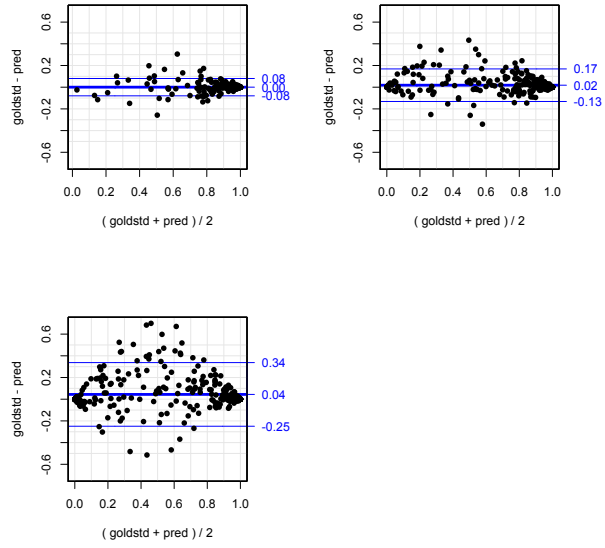


Figure 3: BA plot: data fitted with JM using linear mixed model

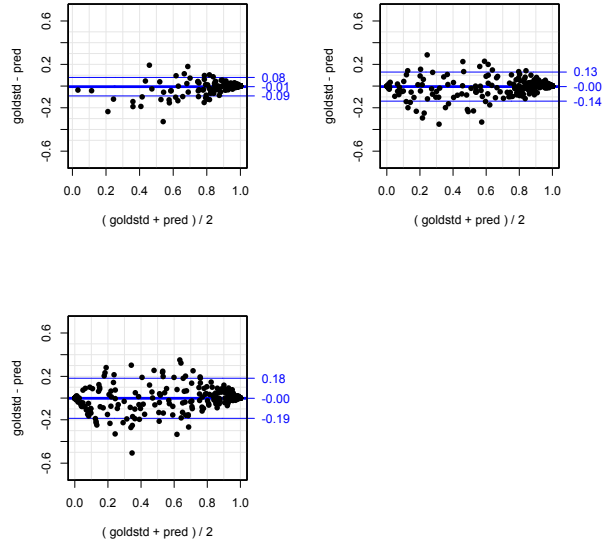


Figure 4: BA plot: data fitted with JM using QR model with  $\tau = 0.5$

### 3.3 Scenario three

In this scenario data are generated using JM using LM, i.e. random error follows standard normal distribution.

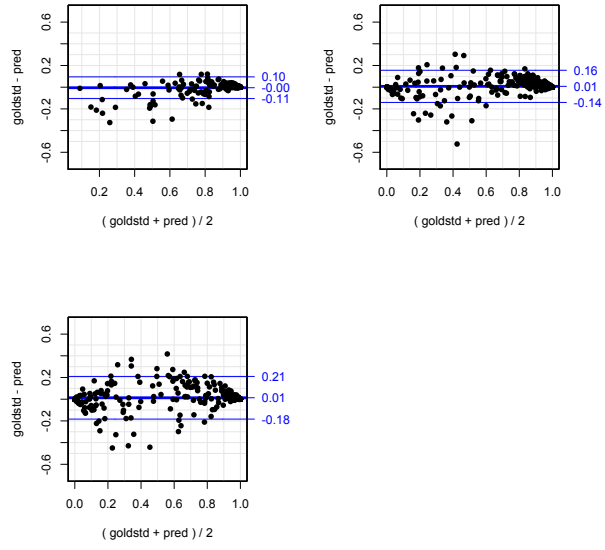


Figure 5: BA plot: data fitted with JM using linear mixed model

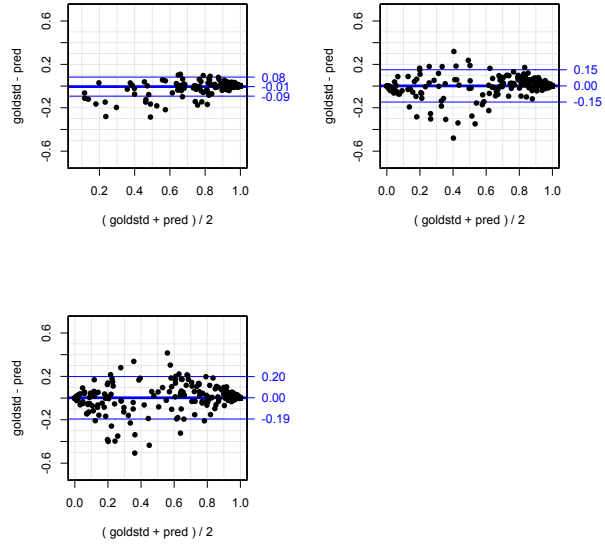


Figure 6: BA plot: data fitted with JM using QR model with  $\tau = 0.5$

### 3.4 Summary table

Table 1: MSE and bias of the predictions of survival probabilities from two models ( $\Delta t_1 > \Delta t_2 > \Delta t_3$ )

Scenario		LMJM		QRJM	
		MSE	Bias	MSE	Bias
1	$\Delta t_1$	0.003	0.010	0.003	0.009
	$\Delta t_2$	0.005	-0.005	0.004	-0.004
	$\Delta t_3$	0.009	-0.013	0.008	-0.008
2	$\Delta t_1$	0.002	-0.000	0.002	0.006
	$\Delta t_2$	0.006	-0.018	0.005	0.005
	$\Delta t_3$	0.023	-0.044	0.009	0.002
3	$\Delta t_1$	0.003	0.005	0.002	0.005
	$\Delta t_2$	0.006	-0.008	0.006	-0.002
	$\Delta t_3$	0.010	-0.013	0.010	-0.003

### 3.5 Comments

1. In both model setting, the accuracy (in terms of MSE and bias) of predictions decreases as  $\Delta t$  increases, which makes sense as it's more difficult to accurately predict the survival for longer time in the future since there are more variabilities and uncertainty.
2. QRJM and LMJM perform closely in predicting the survival probabilities when  $\Delta t$  is smaller but when  $\Delta t$  increases QRJM will outperform LMJM based on the BA plots, bias and MSE.