Web-based Supporting Materials for "Dynamic Prediction for Multiple Repeated Measures and Event Time Data: an application to Parkinson's Disease" by Jue Wang, Sheng Luo, and Liang Li

Table 1: Area under the ROC curve and Brier score (BS) for the DATATOP study.

		Model 1	Model 2	Model 3	Model JM	Cox
t	t'			AUC (95% CI)		
3	9	0.754	0.759	0.761	0.757	0.736
		(0.703, 0.802)	(0.710, 0.805)	(0.713, 0.806)	(0.708, 0.802)	(0.681, 0.786)
	12	0.744	0.744	0.744	0.739	0.725
		(0.702, 0.785)	(0.702, 0.784)	(0.701, 0.785)	(0.696, 0.780)	(0.682, 0.768)
	15	0.744	0.742	0.744	0.726	0.719
		(0.702, 0.783)	(0.702, 0.780)	(0.704, 0.784)	(0.682, 0.768)	(0.677, 0.760)
	18	0.775	0.766	0.772	0.728	0.720
		(0.731, 0.819)	(0.719, 0.811)	(0.723, 0.814)	(0.679, 0.772)	(0.673, 0.765)
6	9	0.789	0.806	0.806	0.770	0.721
		(0.717, 0.851)	(0.739, 0.865)	(0.740, 0.864)	(0.699, 0.834)	(0.642, 0.795)
	12	0.764	0.778	0.775	0.732	0.705
		(0.717, 0.809)	(0.731, 0.821)	(0.729, 0.820)	(0.679, 0.782)	(0.651, 0.757)
	15	0.763	0.771	0.771	0.725	0.697
		(0.716, 0.807)	(0.726, 0.814)	(0.726, 0.814)	(0.678, 0.771)	(0.648, 0.743)
	18	0.786	0.773	0.769	0.726	0.701
		(0.736, 0.832)	(0.726, 0.821)	(0.720, 0.815)	(0.674, 0.776)	(0.650, 0.751)
12	15	0.766	0.787	0.782	0.695	0.647
		(0.701, 0.828)	(0.716, 0.850)	(0.710, 0.849)	(0.623, 0.768)	(0.563, 0.727)
	18	0.758	0.739	0.723	0.700	0.663
		(0.684, 0.824)	(0.671, 0.808)	(0.651, 0.790)	(0.631, 0.774)	(0.592, 0.732)
t	t'			BS (95% CI)		
3	9	0.136	0.138	0.139	0.140	0.139
		(0.116, 0.155)	(0.119, 0.158)	(0.119, 0.158)	(0.120, 0.159)	(0.122, 0.156)
	12	0.204	0.200	0.200	0.203	0.203
		(0.181, 0.226)	(0.180, 0.221)	(0.180, 0.220)	(0.182, 0.225)	(0.184, 0.221)
	15	0.216	0.212	0.211	0.218	0.212
		(0.192, 0.240)	(0.191, 0.234)	(0.191, 0.231)	(0.195, 0.242)	(0.193, 0.232)
	18	0.171	0.163	0.167	0.186	0.185
		(0.148, 0.196)	(0.142, 0.186)	(0.147, 0.188)	(0.162, 0.211)	(0.164, 0.207)
6	9	0.078	0.078	0.078	0.081	0.094
		(0.062, 0.095)	(0.062, 0.095)	(0.062, 0.095)	(0.064, 0.099)	(0.080, 0.108)
	12	0.159	0.154	0.154	0.164	0.173
		(0.137, 0.180)	(0.134, 0.174)	(0.134, 0.174)	(0.143, 0.186)	(0.155, 0.191)
	15	0.183	0.178	0.178	0.194	0.194
		(0.160, 0.207)	(0.157, 0.199)	(0.158, 0.199)	(0.171, 0.218)	(0.176, 0.214)
	18	0.158	0.154	0.159	0.175	0.175
		(0.133, 0.183)	(0.132, 0.178)	(0.138, 0.183)	(0.151, 0.201)	(0.155, 0.197)
12	15	0.108	0.103	0.102	0.124	0.155
		(0.084, 0.133)	(0.082, 0.125)	(0.080, 0.125)	(0.100, 0.150)	(0.136, 0.174)
	18	0.149	0.147	0.153	0.161	0.163
		(0.121, 0.178)	(0.121, 0.174)	(0.126, 0.179)	(0.132, 0.192)	(0.139, 0.187)

Table 2: The outcome-specific parameter estimates for the DATATOP study from Model 1.

	Mean	SD	95% CI	
For UPDRS				
a_1	17.247	0.341	16.563	17.902
b_1	7.624	0.207	7.251	8.035
For HY				
a_{22}	0.995	0.036	0.927	1.066
a_{23}	4.087	0.079	3.935	4.243
a_{24}	6.340	0.129	6.087	6.593
For SEADL				
a_{31}	-1.462	0.076	-1.610	-1.311
a_{32}	0.583	0.071	0.450	0.720
a_{33}	3.008	0.088	2.838	3.181
a_{34}	3.860	0.096	3.679	4.051
a_{35}	6.020	0.132	5.770	6.283
a_{36}	6.851	0.151	6.558	7.150
a_{37}	8.474	0.203	8.082	8.874
b_3	1.270	0.045	1.187	1.363

Predicted Probability for Ordinal Outcomes

The predicted probability being in each category for outcome HY is presented in Figure 1. For example, Patient 169 had HY measurements equal to 2 at all visits. When only the baseline data are used for prediction (the first plot in upper panels), our model tends to underpredict the disease progression by assigning sizable probabilities to the less severe HY categories 1 and 1.5 even at the end of the study, possibly due to low baseline UPDRS value of 33. After month 3 visit (the second plot in upper panels), our model overpredicts disease progression by assigning abnormally high probability to the severe category 3, possibly due to higher UPDRS values at months 1 and 3. However, using the first 6 or 12 months' data (the last two plots in upper panels), our model has good fit by correctly assigning the largest posterior probability to HY category 2 for all visits from baseline to month 12. Moreover, our model properly assigns higher probabilities to more severe categories 2.5 and 3 and negligible probabilities to less severe categories 1 and 1.5 for visits after month 12, due to the deteriorating UPDRS measure. Similar interpretation can be made to the predicted probability of being in each SEADL category displayed in Figure 2.

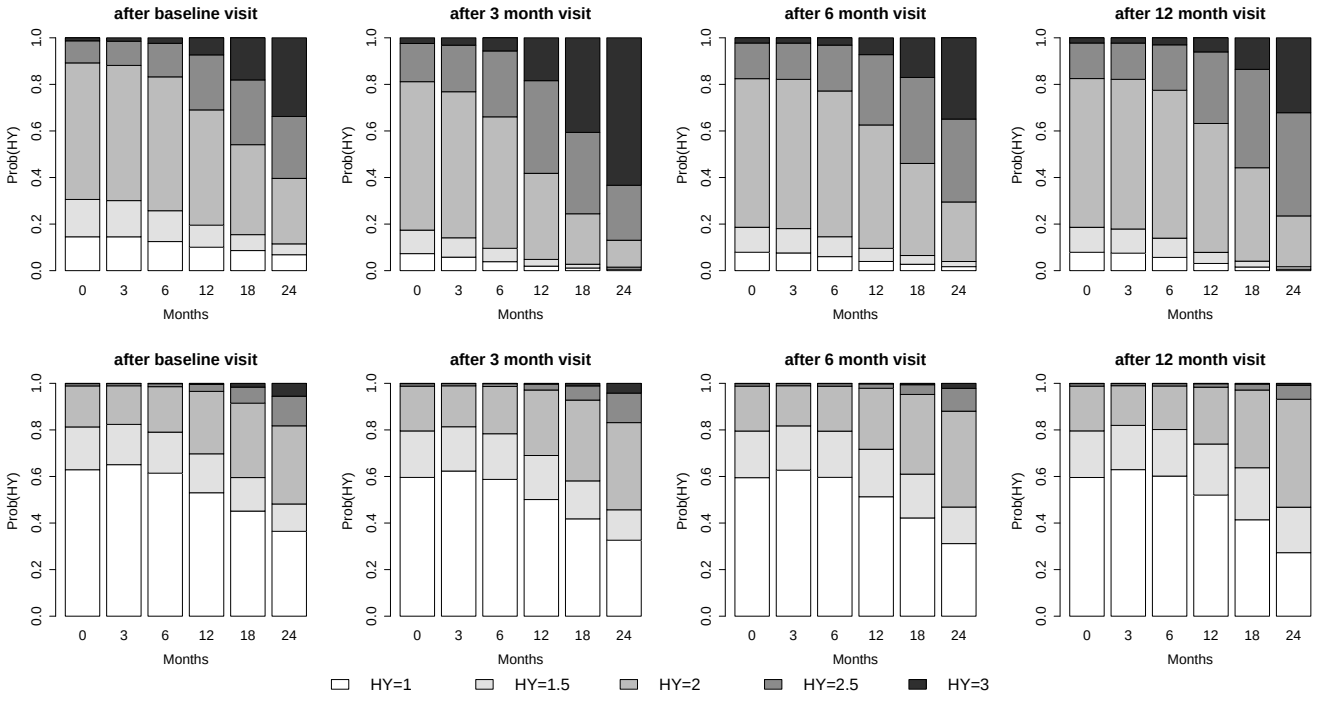

Figure 1: Predicted probability of being in each HY category for Patient 169 (upper panels) and Patient 718 (lower panels). Patient 169 had HY measurements equal to 2 at all 8 visits at months 0, 1, 3, 6, 9, 12, 15, and 16, while Patient 718 had HY measurements equal to 1 at all 9 visits at months 0, 1, 3, 6, 9, 12, 15, and 18.

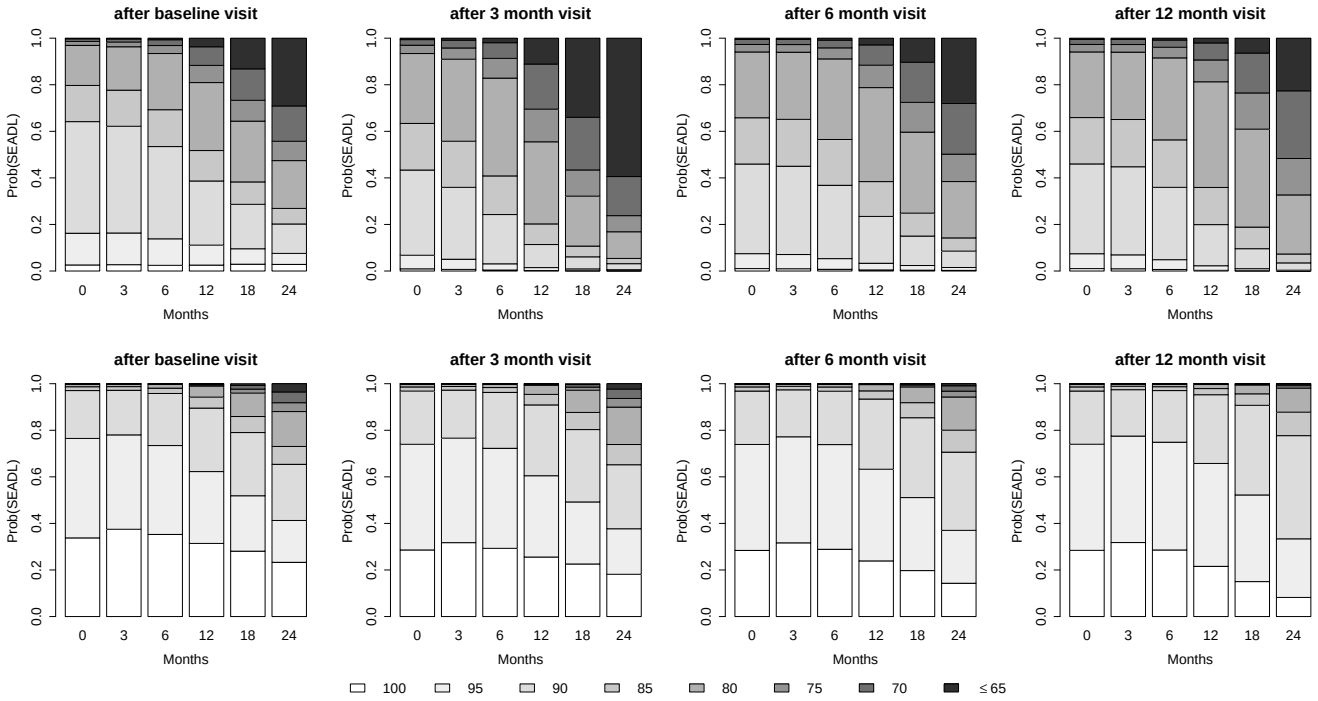

Figure 2: Predicted probability of SEADL to be observed in a given category for Subject 169 (upper panels) and Subject 718 (lower panels). Observed categories of SEADL for Subject 169 in the 8 follow-up visits are 90, 80, 80, 90, 80, 80, 80, 80 and for Subject 718 in the 9 visits are 95, 95, 95, 95, 90, 95, 95, 95, 95.

Table 3: Simulation results using the training dataset.

	BIAS	SD	CP	RMSE
For the latent disease severity				
$\beta_0 = -1$	0.007	0.114	0.945	0.114
$\beta_1 = -0.2$	−0.010	0.118	0.970	0.118
$\beta_2 = 0.8$	0.003	0.022	0.970	0.022
$\beta_3 = -0.2$	−0.001	0.015	0.940	0.015
$\sigma_1 = 1.5$	0.009	0.060	0.950	0.060
$\sigma_2 = 0.15$	0.000	0.007	0.960	0.007
$\rho = 0.4$	−0.003	0.048	0.935	0.048
For the survival process				
$\gamma = -0.12$	−0.001	0.007	0.950	0.007
$\nu = 0.75$	0.005	0.044	0.930	0.044
For the first outcome (continuous)				
$a_1 = 15$	−0.035	0.471	0.955	0.471
$b_1 = 7$	−0.024	0.183	0.960	0.184
$\sigma_\varepsilon = 5$	−0.000	0.099	0.960	0.099
For the second outcome (ordinal)				
$a_{22} = 1$	0.004	0.066	0.925	0.066
$a_{23} = 2$	0.014	0.089	0.930	0.090
$a_{24} = 4$	0.028	0.124	0.940	0.127
$a_{25} = 5$	0.040	0.148	0.920	0.153
$a_{26} = 6$	0.038	0.169	0.915	0.173
For the third outcome (ordinal)				
$a_{31} = -1$	0.004	0.106	0.950	0.106
$a_{32} = 1$	0.001	0.110	0.940	0.110
$a_{33} = 3$	0.011	0.131	0.950	0.132
$a_{34} = 4$	0.012	0.144	0.960	0.144
$a_{35} = 6$	0.023	0.194	0.930	0.195
$a_{36} = 8$	0.022	0.232	0.950	0.233
$b_3 = 1.2$	−0.000	0.040	0.965	0.040

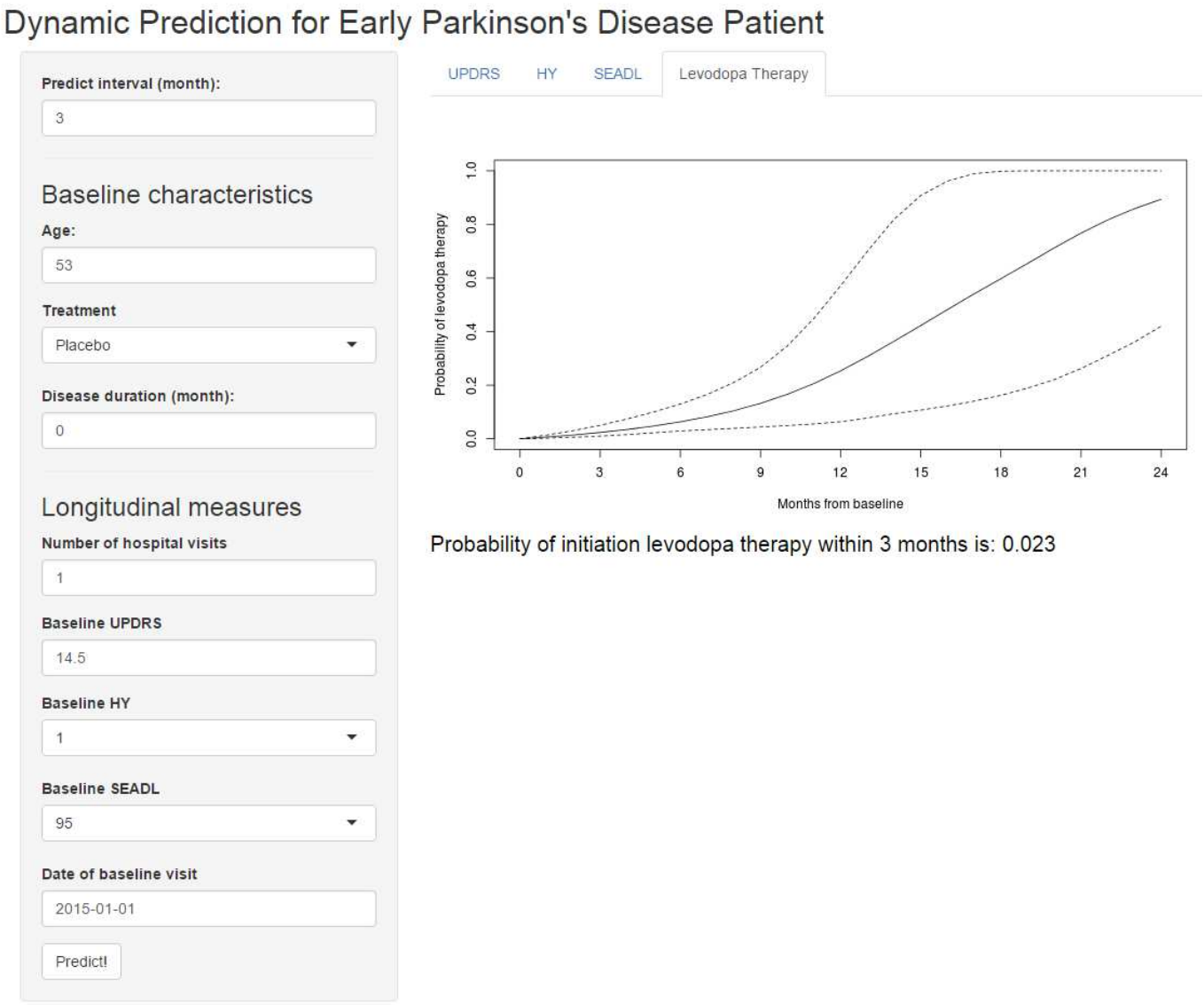

Figure 3: A screenshot of the web-based calculator for prediction.

Table 4: Area under the ROC curve (AUC) for the simulation study.

t	t'	Model 1	Model JM	Cox	True AUC
3	9	0.922	0.909	0.892	0.934
	12	0.920	0.908	0.875	0.943
	15	0.915	0.903	0.853	0.952
	18	0.907	0.896	0.830	0.959
6	9	0.926	0.911	0.883	0.930
	12	0.930	0.915	0.868	0.940
	15	0.932	0.916	0.847	0.950
	18	0.930	0.914	0.825	0.958

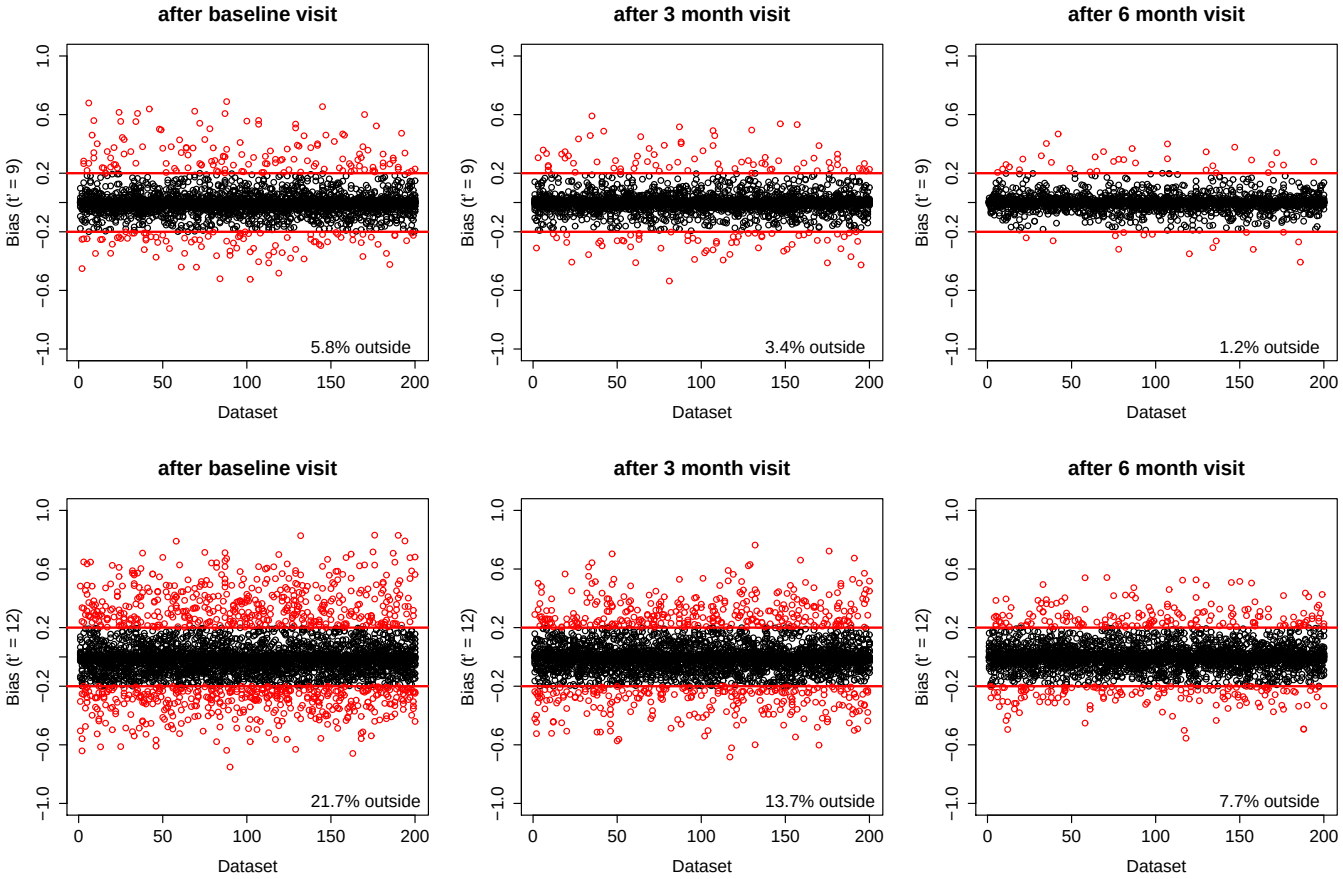

Figure 4: Bias between the predicted failure probability $\widehat{\pi}_i(t'|\boldsymbol{y}_i^{\{t\}}, \boldsymbol{X}_i^{\{t\}})$ with true failure probability when $t' = 9$ (upper panels) and $t' = 12$ (lower panels) for 20 randomly selected subjects from each of the 200 simulation datasets.

Stan code for the simulation study

```
data {
  int<lower=0> N_train; // Number of subjects in training data
  int<lower=0> obs; // Number of observations
  int subject[obs]; // Subject ID
  int<lower=0> K_ordi; // number of ordinal outcomes
  real Y_conti[obs];
  int<lower=0> Y_ordi[obs, K_ordi];
  int<lower=0> n_ordi; // Number of categories for ordinal outcomes
  vector[2] zero;
  real<lower=0> time[obs];
  int<lower=0> treat[obs];
  int<lower=0> treat_pts[N_train];
  int<lower=0, upper=100> age_pts[N_train];
  real tee[N_train]; // Survival time
  int<lower=0> event[N_train]; // Censoring indicator
}
parameters {
  vector<lower=-10, upper=10>[2] beta0;
  vector<lower=-10, upper=10>[2] beta1;
  vector[2] U[N_train];
  real<lower=0> var1;
  real<lower=0> var2;
  real<lower=-1, upper=1> rho;
  real<lower=0> var_conti;
  real gamma;
  real nu;
  real h0;
  real a_conti;
  real<lower=0> b_conti;
  real a_ordi_temp;
  real<lower=0> b_ordi_temp;
  vector<lower=0>[n_ordi-2] delta[K_ordi];
}
transformed parameters {
  real<lower=0> sig1;
  real<lower=0> sig2;
  cov_matrix[2] Sigma_U;
  real<lower=0> sd_conti;
  vector[n_ordi-1] a_ordi[K_ordi];
  vector<lower=0>[K_ordi] b_ordi;
  real theta[obs];
  real mu_conti[obs];
  real<lower=0, upper=1> psi[obs, K_ordi, n_ordi];
  vector<lower=0, upper=1>[n_ordi] prob_y[obs, K_ordi];

  // construct the latent variable theta
  for (i in 1:obs)
    theta[i] <- beta0[1] + beta0[2]*treat[i] + U[subject[i], 1] +
      (beta1[1] + beta1[2]*treat[i] + U[subject[i], 2])*time[i];

  // construct the means for the continuous variables
  for (i in 1:obs)
    mu_conti[i] <- a_conti + b_conti*theta[i];

  // construct the probability vector for the remaining ordinal variables
  a_ordi[1, 1] <- 0;
  for (l in 2:(n_ordi-1)) a_ordi[1, l] <- a_ordi[1, l-1] + delta[1, l-1] ;
  for (k in 2:K_ordi) {
```

```
    a_ordi[k, 1] <- a_ordi_temp;
    for (l in 2:(n_ordi-1)) a_ordi[k, l] <- a_ordi[k, l-1] + delta[k, l-1];
  }
  b_ordi[1] <- 1;
  for (k in 2:K_ordi) b_ordi[k] <- b_ordi_temp;

  for (i in 1:obs) {
    for (k in 1:K_ordi) {
      for (l in 1:(n_ordi-1)) {
        psi[i, k, l] <- inv_logit(a_ordi[k, l] - b_ordi[k]*theta[i]);
      }
      psi[i, k, n_ordi] <- 1;

      prob_y[i, k, 1] <- psi[i, k, 1];
      for (l in 2:n_ordi) {prob_y[i, k, l] <- psi[i, k, l] - psi[i, k, l-1];}
    }
  }

  sd_conti <- sqrt(var_conti);
  sig1 <- sqrt(var1);
  sig2 <- sqrt(var2);

  // construct the variance-covariance matrix
  Sigma_U[1,1] <- sig1*sig1;
  Sigma_U[1,2] <- rho*sig1*sig2;
  Sigma_U[2,1] <- Sigma_U[1,2];
  Sigma_U[2,2] <- sig2*sig2;
}
model {
  real h[N_train];
  real S[N_train];
  real LL[N_train];

  Y_conti ~ normal(mu_conti, sd_conti);
  for (i in 1:obs) {
    for (k in 1:K_ordi) {
      Y_ordi[i, k] ~ categorical(prob_y[i, k]);
    }
  }

  // construct random effects
  U ~ multi_normal(zero, Sigma_U);

  // construct survival part
  for (i in 1:N_train) {
    h[i] <- exp(gamma*age_pts[i] + nu*(beta0[1] + beta0[2]*treat_pts[i] + U[i, 1] +
                                       (beta1[1] + beta1[2]*treat_pts[i] + U[i, 2])*tee[i]))*h0;
    S[i] <- exp(-h0*exp(gamma*age_pts[i]+nu*(beta0[1]+beta0[2]*treat_pts[i]+U[i, 1])) *
                  (exp(nu*(beta1[1]+beta1[2]*treat_pts[i]+U[i, 2])*tee[i])-1) / (nu*(beta1[1]+beta1[2]*treat_pts[i]+U[i, 2])));
    LL[i] <- log(pow(h[i],event[i])*S[i]);  // event=1 for event; 0 for censored
  }
  increment_log_prob(LL);

  // construct the priors
  beta0 ~ normal(0, 10);
  beta1 ~ normal(0, 10);
  var1 ~ inv_gamma(0.01, 0.01);
  var2 ~ inv_gamma(0.01, 0.01);
  rho ~ uniform(-1, 1);
  var_conti ~ inv_gamma(0.01, 0.01);
```

```
  h0 ~ gamma(0.01, 0.01);
  nu ~ normal(0, 10);
  gamma ~ normal(0, 10);

  for (i in 1:(n_ordi-2)) delta[1, i] ~ normal(0, 10) T[0,] ;
  for (k in 2:K_ordi) {
    b_ordi_temp ~ uniform(0, 10);
    a_ordi_temp ~ normal(0, 10);
    for (i in 1:(n_ordi-2)) delta[k, i] ~ normal(0, 10) T[0,] ;
  }
}
```

Full Conditionals

For illustration purpose, we assume that there are one continuous outcome (denoted by $y_{i1}(t)$) and two ordinal outcomes (denoted by $y_{i2}(t)$ and $y_{i3}(t)$, respectively), while model (3) is formulated as $\theta_i(t) = \boldsymbol{X}_i(t)\boldsymbol{\beta} + \boldsymbol{Z}_i(t)\boldsymbol{u}_i$. Assuming non-informative prior distribution for the parameter vector $\boldsymbol{\Theta}$, denoted by $f(\boldsymbol{\Theta})$, the joint likelihood is

$$
\begin{aligned}
& L(\boldsymbol{\Theta}; \cdot) = p(\boldsymbol{y}|\boldsymbol{u})p(\boldsymbol{u})f(\boldsymbol{\Theta}) \\
\propto\ & \prod_{i=1}^{I} \left\{ \prod_{j=1}^{J_i} p\big[Y_{i1}(t_{ij}) = y_{i1}(t_{ij})\big] p\big[Y_{i2}(t_{ij}) = y_{i2}(t_{ij})\big] p\big[Y_{i3}(t_{ij}) = y_{i3}(t_{ij})\big] \right\} \big\{ h_i(t_i)^{\delta_i} S_i(t_i) \big\} p(\boldsymbol{u}_i) \\
=\ & \prod_{i=1}^{I} L_{y_1} L_{y_2} L_{y_3} L_S \cdot p(\boldsymbol{u}_i),
\end{aligned}
$$

where

$$
\begin{aligned}
L_{y_1} &= \prod_{j=1}^{J_i} \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left\{ -\frac{\big[y_{i1}(t_{ij}) - a_1 - b_1\theta_i(t_{ij})\big]^2}{2\sigma_\varepsilon^2} \right\}, \\
L_{y_k} &= \prod_{j=1}^{J_i}\prod_{l=1}^{n_k} p\big[Y_{ik}(t_{ij}) = l\big]^{I[Y_{ik}(t_{ij})=l]} \\
&= \prod_{j=1}^{J_i}\prod_{l=1}^{n_k} \Big\{ p\big[Y_{ik}(t_{ij}) \le l|\theta_i(t_{ij})\big] - p\big[Y_{ik}(t_{ij}) \le l-1|\theta_i(t_{ij})\big] \Big\}^{I[Y_{ik}(t_{ij})=l]} \\
&= \prod_{j=1}^{J_i} \Bigg[\Big\{ 1 - \operatorname{expit}\big[a_{k(n_k-1)} - b_k\theta_i(t_{ij})\big] \Big\}^{I[Y_{ik}(t_{ij})=n_k]} \\
&\qquad \cdot \prod_{l=2}^{n_k-1} \Big\{ \operatorname{expit}\big[a_{kl} - b_k\theta_i(t_{ij})\big] - \operatorname{expit}\big[a_{k(l-1)} - b_k\theta_i(t_{ij})\big] \Big\}^{I[Y_{ik}(t_{ij})=l]} \\
&\qquad \cdot \Big\{ \operatorname{expit}\big[a_{k1} - b_k\theta_i(t_{ij})\big] \Big\}^{I[Y_{ik}(t_{ij})=1]} \Bigg], \ k = 2, 3, \\
L_S &= \Big\{ h_0(t_i) \exp\left[\boldsymbol{W}_i\boldsymbol{\gamma} + \nu\theta_i(t_i)\right] \Big\}^{\delta_i} \exp\left[-\int_0^{t_i} h_0(s) \exp\left[\boldsymbol{W}_i\boldsymbol{\gamma} + \nu\theta_i(s)\right] ds \right], \\
p(\boldsymbol{u}_i) &= \frac{1}{2\pi\sqrt{|\boldsymbol{\Sigma}|}} \exp\left[-\frac{1}{2}\boldsymbol{u}_i'\boldsymbol{\Sigma}^{-1}\boldsymbol{u}_i \right], \\
\operatorname{expit}(\cdot) &= \frac{\exp(\cdot)}{1+\exp(\cdot)}.
\end{aligned}
$$

The full conditionals of all parameters are

1. $f(a_1|\text{others}) \propto N\left(\frac{\sum_{i=1}^{I}\sum_{j=1}^{J_i}\left[y_{i1}(t_{ij}) - b_1\theta_i(t_{ij})\right]}{N_T}, \frac{\sigma_\varepsilon^2}{N_T} \right)$;
2. $f(b_1|\text{others}) \propto N\left(\frac{\sum_{i=1}^{I}\sum_{j=1}^{J_i}\left[y_{i1}(t_{ij}) - a_1\right]\theta_i(t_{ij})}{\sum_{i=1}^{I}\sum_{j=1}^{J_i}\theta_i(t_{ij})^2}, \frac{\sigma_\varepsilon^2}{\sum_{i=1}^{I}\sum_{j=1}^{J_i}\theta_i(t_{ij})^2} \right)$;
3. $f(\frac{1}{\sigma_\varepsilon^2}|\text{others}) \propto \text{Gamma}\left(\frac{N_T}{2} + 1, \frac{\sum_{i=1}^{I}\sum_{j=1}^{J_i}\left[y_{i1}(t_{ij}) - a_1 - b_1\theta_i(t_{ij})\right]^2}{2} \right)$;
4. $[\boldsymbol{a}_2, b_2|\text{others}] \propto \prod_{i=1}^{I} L_{y_2}$;
5. $[\boldsymbol{a}_3, b_3|\text{others}] \propto \prod_{i=1}^{I} L_{y_3}$;

6. $[\boldsymbol{\beta}|\text{others}] \propto \prod_{i=1}^{I} L_{y_1} L_{y_2} L_{y_3} L_S$;

7. $[\boldsymbol{\gamma}, \nu|\text{others}] \propto \prod_{i=1}^{I} L_S$;

8. $[\boldsymbol{u}_i|\text{others}] \propto \left\{ \prod_{j=1}^{J_i} p\big[Y_{i1}(t_{ij}) = y_{i1}(t_{ij})\big] p\big[Y_{i2}(t_{ij}) = y_{i2}(t_{ij})\big] p\big[Y_{i3}(t_{ij}) = y_{i3}(t_{ij})\big] \right\} L_S \cdot p(\boldsymbol{u}_i)$;

9. $[\boldsymbol{\Sigma}|\text{others}] \propto \prod_{i=1}^{I} p(\boldsymbol{u}_i)$,

where $N_T = \sum_{i=1}^{I} J_i$.