Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 51 (3) (2022), 900 – 913
DOI : 10.15672/hujms.884326

Research Article

The joint modeling approach with a simulation study for evaluating the association between the trajectory of serum albumin levels and mortality in peritoneal dialysis patients

Merve Basol*[1], Dincer Goksuluk[1], Murat Hayri Sipahioglu[2], Ergun Karaagaoglu[3]

[1] *Department of Biostatistics, Erciyes University, Kayseri 38280, Turkey*
[2] *Department of Nephrology, Erciyes University, Kayseri 38280, Turkey*
[3] *Department of Biostatistics, Hacettepe University, Ankara 06100, Turkey*

Abstract

We aimed to study the association between mortality and trajectory of serum albumin levels (g/dL) in peritoneal dialysis patients via a joint modeling approach. Joint modeling is a statistical method used to evaluate the relationship between longitudinal and time-to-event processes by fitting both sub-models simultaneously. A comprehensive simulation study was conducted to evaluate model performances and generalize the findings to more general scenarios. Model performances and prediction accuracies were evaluated using the time-dependent ROC area under the curve (AUC) and Brier score (BS). According to the real-life dataset results, the trajectory of serum albumin levels was inversely associated with mortality increasing the risk of death 2.21 times (p=0.003). The simulation results showed that the model performances increased with sample size. However, the model complexity had increased as more repeated measurements were taken from patients and resulted in lower prediction accuracy unless the sample size was increased. In conclusion, using the trajectory of risk predictors rather than baseline (or averaged) values provided better predictive accuracy and prevented biased results. Finally, the study design (e.g., number of samples and repeated measurements) should be carefully defined since it played an important role in model performances.

Mathematics Subject Classification (2020). 92B15, 62N02, 62P10

Keywords. joint modeling, longitudinal model, survival model, personalized medicine, simulation

*Corresponding Author.
Email addresses: basolmerve21@gmail.com (M. Basol), dincer.goksuluk@gmail.com (D. Goksuluk), mhsipahioglu@erciyes.edu.tr (M. Sipahioglu), ekaraaga@gmail.com (E. Karaagaoglu)
Received: 25.02.2021; Accepted: 24.03.2022

1. Introduction

Recently, in follow-up clinical studies, personalized medicine became popular in the monitoring and treating of chronic diseases since it considered subject-specific features rather than the overall characteristics of a study group [17]. From a statistical perspective, subject-specific statistical models might be preferred instead of marginal (i.e., overall) models when there is significant variability between patients. These models are crucial in personalized medicine because they focus on the subjects' characteristics and aim to obtain patient-specific estimations [5]. Figure 1 shows an example of overall and personalized trajectories of serum albumin levels in a follow-up study. Figure 1a shows the overall trend in serum albumin levels for each subset (alive and dead) of peritoneal dialysis (PD) patients. While serum albumin levels in the censored group are constant in time and higher than dead patient, it increases and decreases in time for dead patients. However, the overall trajectory of serum albumin levels is not representative for all patients because its pattern is very different between patients. Figure 1b shows the trajectories of serum albumin levels for randomly selected patients in each subgroup.

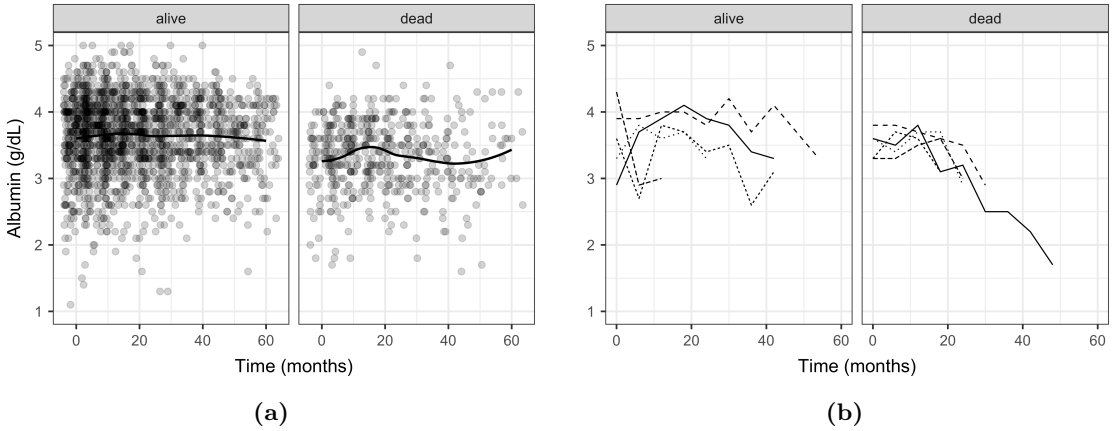


Figure 1. Trajectory of serum albumin levels (a) all patients, (b) randomly selected 5 patients in each group.

One purpose of a follow-up study is to analyze the survival data of patients and explore possible risk factors associated with mortality. In such studies, the survival and longitudinal data are collected simultaneously from patients. However, these two processes are generally analyzed independently of each other. This strategy ignores the relationship between two processes (i.e., longitudinal and time-to-event) and leads to information loss and bias because each submodel evaluates a portion of collected information [22]. Therefore, the relationship between mortality and a longitudinal biomarker should be analyzed simultaneously, and the longitudinal nature of a biomarker should be considered while analyzing time-to-event data. In the literature, Sir David Cox [3] published the first study that aimed to evaluate two processes jointly. In this study, the longitudinal data was introduced into the survival sub-model as a time-dependent predictor of mortality. This modeling strategy assumes that the longitudinal covariate was measured without error, which was not realistic. Schluchter [25] considered a joint model approach which consists of two sub-models: (i) a mixed effect model for the longitudinal data and (ii) a parametric survival model for the time-to-event data, and the model parameters were simultaneously estimated. However, this approach required distributional assumptions in model fitting, e.g., specifying an underlying distribution for baseline hazard, survival outcome, etc. Faucett and Thomas [6] extended Schluchter's proposal by setting baseline hazard function unspecified, unlike the parametric approach. There is an increasing interest in

joint modeling in clinical studies due to its several advantages, these are: (i) dealing with the missing data that arose from both processes (i.e., miss the appointment or censoring, etc.), (ii) providing subject-specific and dynamic risk predictions; hence, approaching to survival data analysis from the perspective of personalized medicine [18, 21], (iii) and finally requiring fewer samples and achieving higher power compared to time-to-event data [12]. There are some review articles for more details about joint modeling [8, 16, 28, 30].

In this study, we aimed to extend the survival model to repeated (i.e., time-dependent biomarkers) measurements using the joint modeling approach to explore the relationship between the trajectory of serum albumin level and mortality in peritoneal dialysis (PD) patients. We evaluated the performances of built joint models via a comprehensive simulation study and on a real-life dataset. We considered several factors in the simulation study to generalize the findings under different scenarios (e.g., sample size and length of the longitudinal period). The predictive ability of built models was evaluated using the time-dependent ROC area under the curve (AUC) and the Brier score (BS). The current study is, to the authors best knowledge, the first study that examined the joint modeling approach on PD patients in Turkey, and among few studies published worldwide [2, 9, 13, 31].

2. Material and methods

2.1. Joint modeling approach

Let T_i and C_i denote the true event time and right censoring time of the i-th patient, respectively. The observed event time is defined as $\widetilde{T}_i = min(T_i, C_i)$ and the main event as $\delta_i = I(T_i < C_i)$ where the indicator function $I(.)$ equals 1 when observed event time is equal to the true event time and 0 otherwise, i.e., the event is observed before censoring time. The follow-up time is generally modeled by using the Cox proportional hazard model [22] as given in Equation (2.1).

$$h_i\left(t \mid Y_i^*(t), w_i\right) = h_0\left(t\right) \exp\left(\gamma^T w_i + \alpha y_i^*(t)\right) \tag{2.1}$$

Here, w_i is(are) a baseline covariate(s) with regression coefficient γ and is independent of the time. $\{Y_i^*(t) = y_i(s), 0 \leq s < t\}$ is the predicted history of longitudinal outcome up to time point t, and $y_i^*(t)$ is the predicted value of longitudinal outcome included in the model as a time-dependent predictor of the event. The regression coefficient, α, corresponds to the relationship between longitudinal response and survival times. If the α is not statistically significant ($p > 0.05$), that means there is no relationship between longitudinal response and time-to-event data, then both processes can be analyzed separately. The baseline hazard function, $h_0(t)$, can be used as unspecified, parametric (Weibull, log-normal, etc.) or more flexible methods such as piecewise-constant and B-spline basis functions for cubic splines [22].

Longitudinal data are generally fitted to mixed-effects models [10] as in Equation (2.2).

$$y_i(t) = y_i^*(t) + \varepsilon_i(t) = X_i(t)\beta + Z_i(t)b_i + \varepsilon_i(t) \tag{2.2}$$

Here, $y_i(t)$ is the longitudinal outcome for the i-th subject at time t. $X_i(t)$ and $Z_i(t)$ are the vectors of fixed and random effect with regression parameters β and b_i, respectively. $\varepsilon_i(t)$ is a time-dependent error term that is assumed to be normally distributed with mean zero and a constant variance σ^2. It is also independent of random effects, which are assumed to follow a multivariate normal distribution with mean zero and inter-subject variance-covariance matrix Σ.

As seen in Equation (2.1), the relationship between two processes is based on the current value of longitudinal biomarker at time t. However, the current value approach may not be sufficient in practice for explaining the relationship between longitudinal and survival processes. Andrinopoulou et al. [1] indicated that selecting the correct parameterization is

crucial in model fitting and provides unbiased results. There are several parameterization alternatives to the association structure specified in Equation (2.1). These alternatives [22] are given in Equations (2.3)–(2.5).

Time-dependent slopes parameterization.

$$\begin{aligned} h_i(t) =& h_0(t) \exp\left(\gamma^T w_i + \alpha_1 y_i^*(t) + \alpha_2 y_i^{*'}(t)\right) \\ y_i^{*'}(t) = \frac{d}{dt}, y_i^*(t) =& \frac{d}{dt}\left(x_i^T(t)\beta + z_i^T(t) b_i\right) \end{aligned} \tag{2.3}$$

Cumulative effects parameterization.

$$h_i(t) = h_0(t) \exp\left(\gamma^T w_i + \alpha \int_0^t y_i^*(s)\, ds\right) \tag{2.4}$$

Random effects parameterization.

$$h_i(t) = h_0(t) \exp\left(\gamma^T w_i + \alpha^T b_i\right) \tag{2.5}$$

2.2. Parameter estimation of joint modeling

Two popular approaches are available to jointly estimate model parameters, i.e., maximum likelihood estimation (MLE) [11, 34] and the Bayesian approach [6]. These two methods should not be confused with a two-stage estimation technique [24, 33] that sequentially estimates model parameters for each sub-model. In this study, we used an MLE approach and estimated model parameters from the joint likelihood function. The log-likelihood function of the model for the i-th patient is given in Equation (2.6) [22].

$$\begin{aligned} \log p(T_i, \delta_i, y_i; \theta) &= \log \int p(T_i, \delta_i, y_i, b_i; \theta)\, db_i \\ &= \log \int p(T_i, \delta_i \mid b_i; \theta_t)\, (y_i \mid b_i; \theta_y)\, p(b_i; \theta_b)\, db_i \end{aligned} \tag{2.6}$$

Let $\theta = \left(\theta_t^T, \theta_y^T, \theta_b^T\right)^T$ denote the full parameter vector of the joint model. Here, θ_t^T and θ_y^T represent parameter vectors of survival and longitudinal sub-model, respectively. θ_b^T is the parameter of the random-effects covariance matrix. $p(b_i; \theta_b)$ represents the probability density function of random effects, which is assumed to be distributed multivariate normal with mean zero and inter-subject variance-covariance matrix Σ. The likelihood of joint longitudinal part, with random effects in Equation (2.6), is defined as Equation (2.7), clearly.

$$\begin{aligned} p(y_i \mid b_i; \theta_y)\, p(b_i; \theta_b) &= \prod_j p(y_i(t_{ij}) \mid b_i; \theta_y)\, p(b_i; \theta_y) \\ &= (2\pi\sigma^2)^{-n_i/2} \exp\left(-\|y_i - X_i\beta - Z_i b_i\|^2 / 2\sigma^2\right) \\ &\quad \times (2\pi)^{-q_b/2} \det(\Sigma)^{-1/2} \exp\left(-b_i^T \Sigma^{-1} b_i / 2\right) \end{aligned} \tag{2.7}$$

Here, q_b is the length of the random effects vector and $\|.\|$ is the Euclidean norm. The probability density function of the survival part, which is conditional on random effects, is defined as in Equation (2.8).

$$
\begin{aligned}
p\left(T_i, \delta_i \mid b_i; \theta_t\right) &= h_i\left(T_i \mid Y_i^*(T_i); \theta_t\right)^{\delta_i} \; S_i\left(T_i \mid Y_i^*(T_i); \theta_t\right) \\
&= \left[h_0(T_i) \exp(\gamma^T w_i + \alpha y_i^*(T_i))\right]^{\delta_i} \\
&\quad \times \exp\left(-\int_0^{T_i} h_0(s) \exp(\gamma^T w_i + \alpha y_i^*(s) ds\right)
\end{aligned} \tag{2.8}
$$

Maximization of the likelihood function is done using the expectation-maximization (EM) algorithm [34] or Newton-Raphson algorithm [14]. The computational complexity of integrals can be overcome using 7-point or 15-point Gauss-Kronrod rule [22], quadrature rule [11, 34], Monte Carlo sampling, or Laplace approximations [22].

2.3. Measuring predictive accuracy

Joint modeling plays an important role in the development of personalized medicine [17, 23]. It provides dynamic risk predictions during the follow-up period, which helps the physicians to decide about patients treatment, such as quitting or continuing a new drug, collecting additional data, etc. The dynamically predicted survivals in future time points are obtained using all the longitudinal information collected up to time point, i.e., $y_i(s) = \{y_{ij}; 0 \leq t_{ij} \leq s\}, s \geq 0$. Hence, the conditional survival probability of the i-th subject at horizon $s + t$ is obtained as in Equation (2.9) [22].

$$
\pi_i\left(s + t \mid s\right) = P\left(\tilde{T}_i \geq s + t \mid \tilde{T}_i > s, y_i\left(s\right); \theta\right), \quad t > 0 \tag{2.9}
$$

The conditional survival probability can be predicted using Monte Carlo simulation or numerical integrations [18, 21]. In this study, we used the Monte Carlo approach to obtain predicted survivals along with 95% confidence intervals. Predictive accuracy of dynamic predictions was evaluated using time-dependent AUC and Brier score (BS) as *discrimination* and *calibration* measures, respectively.

Brier score, also known as prediction error, is the difference between actual status and predicted survival probability. Using the longitudinal data up to time point s, the BS of the fitted joint model at horizon time $s + t$ is calculated as [26]

$$
\widehat{BS}\left(s + t \mid s\right) = \frac{1}{\sum_{i=1}^{N} I\left(\tilde{T}_i > t\right)} \sum_{i=1}^{N} \widehat{W}_i\left(s, t\right) \left(\widehat{G}_i(s, t) - \widehat{\pi}_i\left(s + t \mid s\right)\right)^2, \tag{2.10}
$$

where $\widehat{G}_i(s, t) = I(\tilde{T}_i > s + t)$ is an indicator function of surviving status that equals 1 if observed event time is larger than $s + t$ and 0 otherwise. Note that the true status of a patient at time point $s + t$ is unknown if this patient was censored within the interval $(s, s + t]$. Hence, an *Inverse Probability of Censoring Weights* (IPCW) is used to adjust the calibration of the fitted model for censored patients [7]. Censored patients were assumed to be alive with a probability of $\widehat{\pi}_i\left(\tilde{T}_i + t \mid \tilde{T}_i\right)$ and dead with a probability of $1 - \widehat{\pi}_i\left(\tilde{T}_i + t \mid \tilde{T}_i\right)$. Hence, censored patients contributed to both dead and alive subsets with corresponding weights. For dead and alive patients, on the other hand, the weights equal 1.

The time-dependent ROC AUC is used to measure how well the fitted model discriminates a randomly selected pair with a low (j-th) and high risk (i-th) of death in the interval $(s, s + t]$. This measure was extended to dynamic predictions [21] and adjusted for censored patients using IPCW [4] as similar to BS. The estimated AUC is obtained as in Equation (2.11).

$$\widehat{\mathrm{AUC}}\left(s+t \mid s\right)=\frac{\sum_{i=1}^{N}\sum_{j=1}^{N} I\left(\widehat{\pi}_i(s+t \mid s)<\widehat{\pi}_j(s+t \mid s)\right)\widehat{G}_j(s,t)(1-\widehat{G}_i(s,t))\widehat{W}_i(s,t)\widehat{W}_j(s,t)}{\sum_{i=1}^{N}\sum_{j=1}^{N}\widehat{G}_j(s,t)(1-\widehat{G}_i(s,t))\widehat{W}_i(s,t)\widehat{W}_j(s,t)} \tag{2.11}$$

The higher values of AUC and the lower values of BS indicated better predictive accuracy of the fitted joint model.

2.4. Real-life dataset

We used the data from the previously published paper by [27]. This dataset was retrospectively collected from medical records. It consisted of 511 patients undergoing PD between the years 1995 and 2007 at Erciyes University Nephrology Department. PD is a frequent and developing therapy for end-stage renal diseases (ESRD). In Turkey, five-year survival rates in adults were estimated as 68.8% by [27] and as 65% by [29]. These values were higher than other regions reported by [15]. Four hundred seventeen patients were included in the study according to exclusion criteria originally described in [27]. Patients having no measured data in longitudinal response were also excluded. In this study, patients' data including demographic (age at onset, body mass index, gender, cause of ESRD, existence of comorbid diseases, dialysis history, the transport property of peritoneal membrane (low or high), number of illnesses, and peritonitis rate) and clinical/biochemical variables (serum albumin levels) were collected from medical records. Data were collected from each patient every two months. However, data at 6-month intervals up to 5-years of follow-up were considered in the statistical analyses. Patients that died during the PD process or 3-months after transferring to HD were considered as PD-related deaths and censored otherwise.

2.5. Simulation scenarios

In addition to real data application, we conducted a comprehensive simulation study to generalize the performance of joint models under different scenarios. In this context, we generated data using all the combinations of

- Sample size n = 100, 400, and 800 as small, medium and large, and
- Length of longitudinal period as 60 and 120 months.

Four hundred data sets were generated in each scenario and equally split into two parts, 200 samples each, as train and test sets. Generated data were fitted to two joint models, named true ($\mathcal{M}_1$) and wrong ($\mathcal{M}_2$) models. The true model included the predictors given in Table 2. The wrong model, on the other hand, included a random intercept LME sub-model and a Cox proportional hazard sub-model, and adjusted for gender which was found insignificant predictor in real-life data analysis. The model performances in the simulation study were evaluated using the differences in AUCs and BS in two models as ($\mathcal{M}_1$ $\mathcal{M}_2$). In joint modeling, we used an R package [20] which was specifically developed for joint modeling of longitudinal and survival processes.

2.6. Availability of computer programs and source codes

We performed real-life data analysis and simulations in the R programming language environment [19]. We mainly used R packages **JM** [20] and **ggplot2** [32] to fit a Joint Model and draw figures, respectively. All the source codes are freely available through the GitHub[†] network.

[†]https://github.com/basolmerve/JMArticleSupplementary

3. Results

During the follow-up period, 86 (20.6%) of 417 patients died due to PD-related causes. The median follow-up period of the study was 30 months (range: 3 to 137). We presented the descriptive statistics for demographic and clinical variables in Table 1 for total and survival subgroups. Males were higher in the entire group (%57.1); however, the distribution of gender was similar in dead and censored subgroups. When the causes of ESRD were examined, we found that 145 (34.8%) patients had diabetes, followed by 62 (14.9%) patients with hypertension and 38 (9.1%) patients with glomerulonephritis. Fifty (12.0%) patients received hemodialysis (HD) before PD. The average age of starting PD was 45.92 $\pm$ 14.33 years.

Table 1. Biochemical, clinical and demographic findings of study group (n=417).

Characteristic	Total (n: 417)	Dead (n: 86)	Censored (n: 331)
Age	45.92 ± 14.33	50.3 ± 13.76	44.78 ± 14.27
BMI	23.61 ± 4.09	24.1 ± 4.15	23.49 ± 4.08
Gender (female)	179 (42.9)	36 (41.9)	143 (32.0)
Cause of ESRD†			
Diabetes mellitus (DM)	145 (34.8)	39 (45.4)	106 (32.0)
Glomerulonephritis	38 (9.1)	6 (10.5)	32 (9.7)
Hypertension	62 (14.9)	9 (10.5)	53 (16.0)
PKD	19 (4.6)	6 (7.0)	13 (3.9)
Other	41 (9.8)	9 (10.5)	32 (9.7)
Unknown	112 (27.1)	16 (18.6)	96 (29.0)
Comorbidity			
Cardiovascular disease	92 (22.1)	35 (40.7)	57 (17.2)
Lung disease	13 (3.1)	6 (7.0)	7 (2.1)
Hepatitis	60 (14.4)	27 (31.4)	33 (9.9)
PD history (present)	50 (12.0)	20 (23.3)	30 (9.1)
TPPM (High)	208 (49.9)	45 (52.3)	163 (49.2)
Number of illness	1 [0 – 5]	2 [0 – 5]	1 [0 – 4]
Peritonitis rate (episodes/patient-year)	0.32 [0 – 5.33]	0.57 [0 – 3]	0.25 [0 – 5.33]

* Summarized using mean $\pm$ standard deviation, frequency (percentage) or median [minimum, maximum] where appropriate.

PKD: Polycystic kidney disease, **BMI**: Body Mass Index(kg/m^2), **PD**: Peritoneal Dialysis, **TPPM**: Transport Property of Peritoneal Membrane, **ESRD**: end stage renal diseases

† Patients might have more than one disease causing ESRD.

The longitudinal sub-model was fitted to serum albumin trajectories using a random slope and random intercept linear mixed model as given in Equation (3.1). We determined the explanatory variables associated with the longitudinal trajectory of serum albumin level using a univariate mixed-effect model. Finally, we adjusted the serum albumin trajectories for significant confounders, i.e., baseline age (p<0.001), the transport property of peritoneal membrane (TPPM) (p<0.001), and peritonitis rate (PR) (p=0.012) (Table (2).

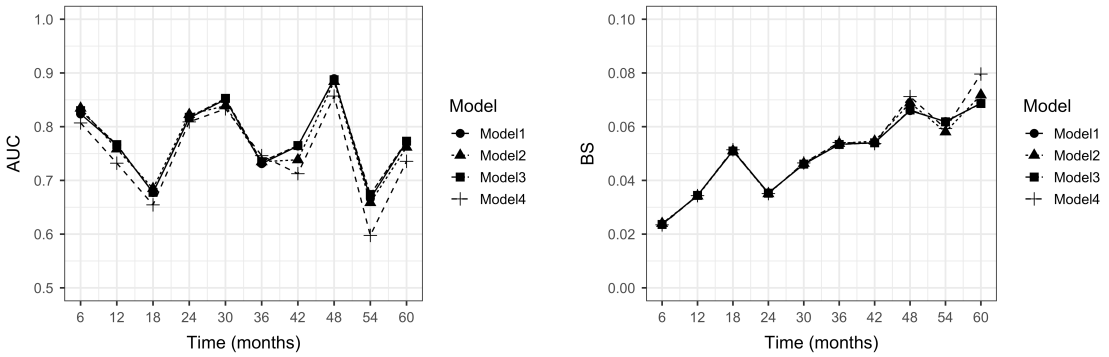
$$\begin{aligned} \mathrm{ALB}_i(t) &= \beta_0 + \beta_1 \mathrm{age} + \beta_2 \mathrm{TPPM} + \beta_3 \mathrm{PR} + \beta_4 \mathrm{time} \\ &\quad + b_{i0} + b_{i1} \mathrm{time} + \epsilon_i(t) \\ &= y_i^*(t) + \epsilon_i(t) \end{aligned} \tag{3.1}$$

In the survival sub-model, the explanatory variables were specified using a univariate Cox proportional hazard regression model. The hazard of death was adjusted for age onset (p<0.001), history of PD (p=0.031), number of illness/comorbid diseases (NI) (p<0.001), and peritonitis rate (PR) (p<0.001) (Table (2). Finally, the fitted trajectory of serum albumin was included in the survival sub-model considering different parameterization as Equation (3.2).

$$\begin{aligned}
&\text{Model 1}: h_i(t) = h_0(t)\exp\left(\gamma_1 \text{age} + \gamma_2 \text{PD history} + \gamma_3 \text{NI} + \gamma_4 \text{PR} + \alpha_1 y_i^*(t)\right)\\
&\text{Model 2}: h_i(t) = h_0(t)\exp\left(\gamma_1 \text{age} + \gamma_2 \text{PD history} + \gamma_3 \text{NI} + \gamma_4 \text{Pr} + \alpha_1 y_i^*(t) + \alpha_2 y_i^{*'}(t)\right)\\
&\text{Model 3}: h_i(t) = h_0(t)\exp\left(\gamma_1 \text{age} + \gamma_2 \text{PD history} + \gamma_3 \text{NI} + \gamma_4 \text{PR} + \alpha_1 \int_0^t y_i^*(s)\,ds\right)\\
&\text{Model 4}: h_i(t) = h_0(t)\exp\left(\gamma_1 \text{age} + \gamma_2 \text{PD history} + \gamma_3 \text{NI} + \gamma_4 \text{PR} + \alpha_1 b_i\right) \qquad (3.2)
\end{aligned}$$

Joint modeling results are presented in Table 2. The model parameters were iteratively estimated using the EM algorithm. The iterations history was given in the Appendix section with Supplementary Figures A1 and A2. We observed that survival times were not associated with subject-specific slopes of serum albumin level according to Model 4 (p=0.201). Model 1 3 show that changes in the adjusted serum albumin levels were negatively and significantly associated with the risk of death (p=0.003, p<0.001, and p=0.005, respectively). A unit decrease in serum albumin levels at a time point t resulted in 2.21 times higher risk of death (95% CI: 1.30 to 3.74) for Model 1. Baseline age, HD history, number of diseases, and peritonitis rate were positively associated with the risk of death. Moreover, serum albumin levels were negatively affected by baseline age of PD, transportation characteristics of the membrane, and peritonitis rate for all models.

When we evaluated the overall predictive performance of all models, we could say that Model 1 was better according to deviance information criteria (DIC). We also compared the dynamic predictive performance of all models using a 5-fold and 10-repeat cross-validation. Cross-validated AUC and BS were given in Figure 2. The horizon time (t) was six months regarding the clinical significance. For example, the AUC value estimated at the 48-th month corresponded to the prediction accuracy of the model at the 54-th month. All models except Model 4 performed similarly with AUC values between 0.65 and 0.90 in time (Figure 2a). We could say that there was an increasing trend in the discrimination performance, AUC, as more data were available from patients. However, decreasing the sample size (or increasing the failure event cumulative incidence) in time resulted in increased prediction error (Figure 2b).



(a) **(b)**

Figure 2. Dynamic prediction performance of the models (5-fold and 10-repeat cross-validation) – (a) time-dependent AUC, (b) Brier Score.

Table 2. The results of joint models with different parameterizations ($n = 417$).

	Model 1				**Model 2**			
Longitudinal Sub-model								
		%95 C.I.				**%95 C.I.**		
	Coefficient	**Lower**	**Upper**	**p**	**Coefficient**	**Lower**	**Upper**	**p**
Age	−0.006	−0.009	−0.003	0.001	−0.006	−0.009	−0.003	< 0.001
TPPM (High)	−0.353	−0.439	−0.267	< 0.001	−0.350	−0.436	−0.264	< 0.001
PR	−0.069	−0.129	−0.010	0.023	−0.070	−0.129	−0.010	0.002
σ		0.352				0.351		
Survival Sub-model								
		%95 C.I.				**%95 C.I.**		
	Coefficient	**Lower**	**Upper**	**p**	**Coefficient**	**Lower**	**Upper**	**p**
Age	0.031	0.014	0.049	0.001	0.039	0.020	0.056	< 0.001
PD history (Yes)	0.877	0.337	1.418	0.002	0.824	0.273	1.376	0.003
NI	0.439	0.129	0.689	0.001	0.499	0.244	0.756	< 0.001
PR	0.608	0.267	0.949	0.001	0.616	0.260	0.970	< 0.001
α_1	−0.791	−1.319	−0.262	0.003	−0.981	−1.547	−0.414	< 0.001
α_2					0.439	0.071	0.807	0.019
BIC		3818.852				3819.002		

	Model 3				**Model 4**			
Longitudinal sub-model								
		%95 C.I.				**%95 C.I.**		
	Coefficient	**Lower**	**Upper**	**p**	**Coeficient**	**Lower**	**Upper**	**p**
Age	−0.006	−0.009	−0.003	< 0.001	−0.006	−0.011	−0.003	< 0.001
TPPM (High)	−0.354	−0.439	−0.268	< 0.001	−0.349	−0.356	−0.179	< 0.001
PR	−0.069	−0.129	−0.010	0.023	−0.070	−0.311	−0.087	0.002
σ		0.352				0.352		
Survival sub-model								
		%95 C.I.				**%95 C.I.**		
	Coefficient	**Lower**	**Upper**	**p**	**Coefficient**	**Lower**	**Upper**	**p**
Age	0.032	0.015	0.050	< 0.001	0.038	0.021	0.068	< 0.001
PD history (Yes)	0.880	0.340	1.420	0.001	0.896	0.106	0.641	< 0.001
NI	0.448	0.199	0.697	< 0.001	0.518	0.854	0.668	< 0.001
PR	0.613	0.272	0.952	< 0.001	0.672	−1.180	−0.129	< 0.001
α_1	−1.435	−2.434	−0.437	0.005	0.247	−0.847	1.253	0.201
BIC		3819.831				3826.744		

TPPM: Transport Property of Peritoneal Membrane, **PR**: Peritonitis rate, **PD**: Peritoneal dialiysis
NI: Number of diseases, **BIC**: Bayesian information criteria, **C.I.**: Confidence intervals

We evaluated the effect of sample size and the number of repeated measurements through simulation work. Since models 1 – 3 had similar predictive performance, we conducted a simulation study using the model parameters from Model 1, which had less model complexity than other models. Figures 3 and 4 show the model performances obtained from the simulation study (sample sizes in the columns and number of replications in the rows). The discrimination ability of the true model ($\mathcal{M}_1$) was generally better than the wrong model ($\mathcal{M}_2$) and increased as more samples were available. Although increasing sample size contributed to the model performances, the true model performed slightly better than the wrong model after 84-months when the length of the longitudinal period was 120 months (Figure 3). Calibration results showed that the amount of prediction error of the true model was not affected by the sample size (first row in Figure 4). However,

it was slightly lower when the length of the longitudinal process was 60 months. There was a very slight difference between true and wrong models ($\mathcal{M}_1$ $\mathcal{M}_2$) when the sample size was small and the length of the follow-up period was 120 months. Both models had a similar prediction error after the 84-th month likely the discrimination ability of the fitted models.

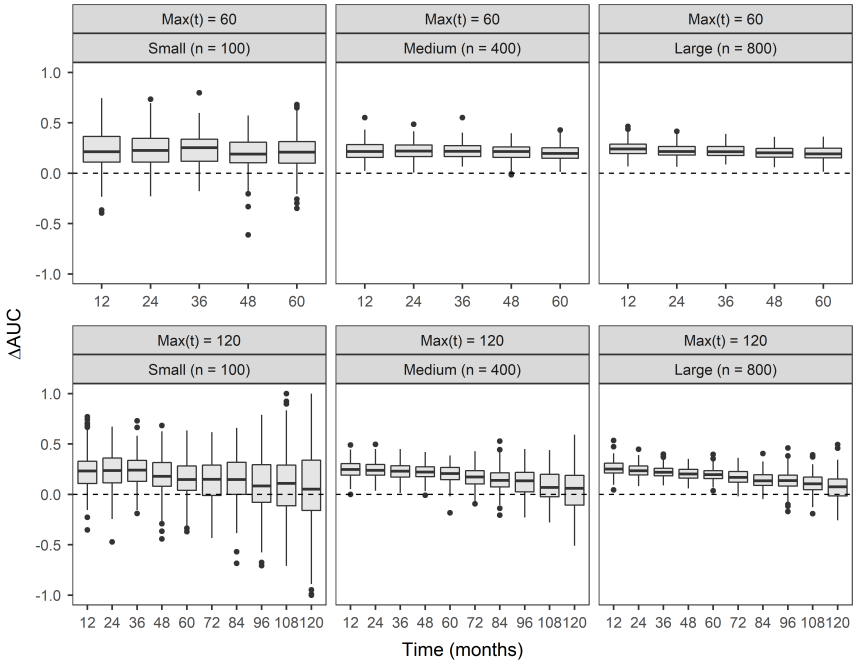


Figure 3. Simulation results Difference in AUCs from true ($\mathcal{M}_1$) and wrong ($\mathcal{M}_2$) models, i.e. ($\mathcal{M}_1$ $\mathcal{M}_2$).

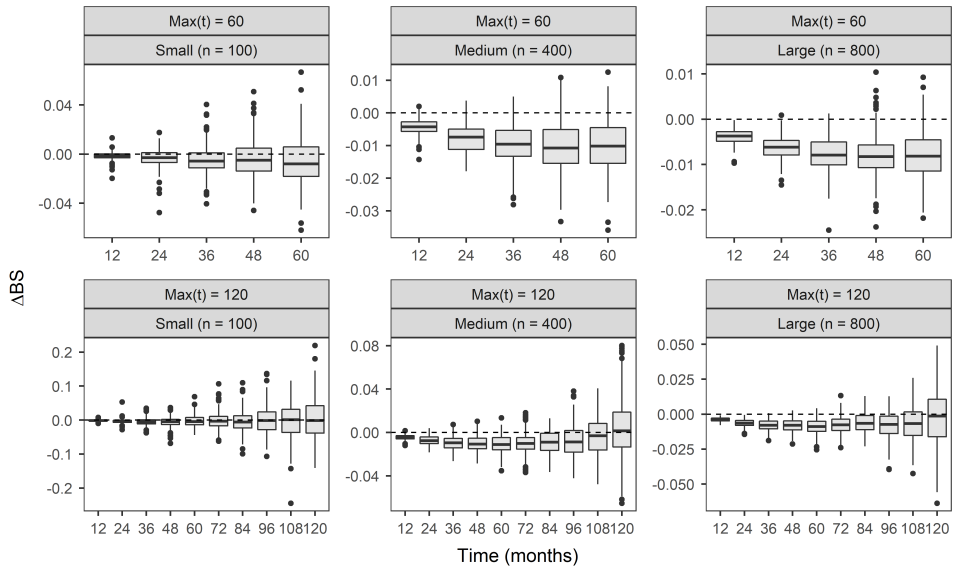


Figure 4. Simulation results Difference in Brier Score from true ($\mathcal{M}_1$) and wrong ($\mathcal{M}_2$) models, i.e. ($\mathcal{M}_1$ $\mathcal{M}_2$).

4. Discussion

Modeling time-to-event data via Cox proportional hazard model and/or Kaplan-Meier method is a simple and easy way when time-dependent covariates do not exist. In such cases, baseline or averaged values of multiple measurements are preferred as the predictor of the main event (i.e., mortality). However, this strategy might be appropriate when not the changes in a biomarker but its averaged or baseline values are related to the main event. Otherwise, omitting the longitudinal nature of biomarkers may lead to information loss and incorrect conclusions when the longitudinal and time-to-event data are collected simultaneously. Joint modeling is one of the appropriate methodologies that can be used in such situations. It models longitudinal and time-to-event data simultaneously and takes repeated measurements of a biomarker into account while fitting time-to-event data to Cox proportional hazard model.

In this study, we studied the applicability of the joint modeling approach on a real-life dataset and generalized the findings through a comprehensive simulation study. We were able to adjust longitudinal and time-to-event responses using common or different predictors through sub-models of the joint model. An important finding was that the averaged serum albumin levels were not associated with mortality; however, its change in time was significantly and reversely associated with mortality increasing the risk of death 2.21 times (95% CI: 1.30 3.74). Results were similar in the previously published studies [13, 31]. This finding clearly showed the importance of using repeated measurements in the risk estimations rather than baseline or averaged values.

Joint modeling may have a crucial role in personalized medicine, in which treatment will be assigned to a patient while monitoring chronic diseases. We fitted the longitudinal data to a linear mixed-effect model and obtained patient-specific model parameters. It enabled us to consider patients individually and to get risk predictions based on a patients characteristics. Serum albumin changes were monitored during the follow-up, and survival predictions in future time points were dynamically estimated using the longitudinal data at hand. Dynamic predictions helped us in two ways: (i) it was possible to see the trend in the longitudinal biomarkers and detect instant increases and/or decreases, and (ii) risk predictions in future time points were dynamically updated during the follow-up period. Therefore, personalized and dynamic risk predictions might help physicians decide about patients treatment, such as quitting or continuing a new drug, collecting additional data, etc. This study showed the importance of using longitudinal biomarkers and accounting for patient variability. It was clear from the findings that risk predictions were significantly changed when the longitudinal data were preferred instead of baseline or averaged values.

It is crucial to generalize findings since real-life data analysis may not reflect the model performances in a wider perspective. This study provided comprehensive simulation results to show the model performances under different scenarios. It also revealed two significant findings from the simulation study. First, as expected, model performances increased, and the variability in the estimations was decreased as the sample size increased. Second, the model performances decreased as the number of repeated measurements increased. It might be due to the increasing complexity of the model and the decreasing sample size in time due to censoring and death. However, increasing the sample size along with the number of repeated measurements provided better model performances. Therefore, it is possible to say that increasing the number of repeats without increasing the sample size will not (or barely) contribute to the model performances. For small samples, using fewer repeats might be reasonable to fit less complicated joint models. When the follow-up period is long, and the biomarker is measured at many time points, one should include more samples in the study to better fit the underlying longitudinal structure or use a subset of repeated measurements to decrease model complexity.

In conclusion, the joint modeling approach provided better and patient-specific risk predictions. It used a comprehensive modeling strategy by considering the longitudinal and time-to-event data simultaneously. Therefore, we may say that it was a flexible and interpretable model. Furthermore, using subject-specific risk predictions might be beneficial for monitoring patients and taking patient-specific actions during the follow-up period.

References

[1] E.R. Andrinopoulou and D. Rizopoulos, *Bayesian shrinkage approach for a joint model of longitudinal and survival outcomes assuming different association structures*, Stat. Med. **35** (26), 4813-4823, 2016.

[2] M. Basol, D. Goksuluk, M.H. Sipahioglu and E. Karaagaoglu, *Effect of serum albumin changes on mortality in patients with peritoneal dialysis: a joint modeling approach and personalized dynamic risk predictions*, Biomed Res. Int. **2021**, 6612464, 2021.

[3] D.R. Cox, *Regression models and life tables (with discussion)*, J. R. Stat. Soc. Ser. B. Stat. Methodol. **34**, 187-200, 1972.

[4] S. Desmee, F. Mentre, C. Veyrat-Follet, B. Sebastien and J. Guedj, *Nonlinear joint models for individual dynamic prediction of risk of death using Hamiltonian Monte Carlo: application to metastatic prostate cancer*, BBMC Med. Res. Methodol. **17** (105), 1-12, 2017.

[5] R.M. Elashoff, G. Li and N. Li, *Joint Modelling of Longitudinal and Time-to-Event Data*, CRC Press. Taylor & Francis Group, 2017.

[6] C. Faucett and D. Thomas, *Simultaneously modelling censored survival data and repeatedly measured covariates: A Gibbs sampling approach*, Stat. Med. **15** (15), 1663-1685, 1996.

[7] T. Gerds and M. Schumacher, *Consistent estimation of the expected Brier score in general survival models with right-censored event times*, Biom. J. **48** (6), 1029-1040, 2006.

[8] A.L. Gould, M.E. Boye, M.J. Crowther, J.G. Ibrahim, G. Quartey, S. Micallef and F.Y. Bois, *Joint modeling of survival and longitudinal non-survival data: current methods and issues. Report of the DIA Bayesian joint modeling working group*, Stat. Med. **34** (14), 2181-2195, 2015.

[9] I. Guler, C. Faes, C. Cadarso-Suarez, L. Teixeira, A. Rodrigues and D. Mendonca, *Two-stage model for multivariate longitudinal and survival data with application to nephrology research*, Biom. J. **59** (6), 1204-1220, 2017.

[10] D. Hedeker and R.D. Gibbons, *Longitudinal Data Analysis*, John Wiley & Sons Inc, 2006.

[11] R. Henderson, P. Diggle and A. Dobson, *Joint modelling of longitudinal measurements and event time data*, Biostatistics **1** (4), 465-448, 2000.

[12] J.G. Ibrahim, H. Chu and L.M. Chen, *Basic concepts and methods for joint models of longitudinal and survival data*, Kidney Res Clin Pract. **28** (16), 2796-2801, 2010.

[13] M. Khoshhali, I. Kazemi, S.M. Hoseeini and S. Sierafian, *Predicting tree-year clinical outcomes using the baseline and trajectories of serum albumin in patients on peritoneal dialysis*, Iran Red Crescent Med J **19** (10), 2017.

[14] K. Lange, *Optimization*, Springer-Verlag, 2004.

[15] P.K.T. Li, K.M. Chow, M.W.M. Van de Luijtgaarden, D.W. Johnson, K.J. Jager, R. Mehrotra, S. Naicker, R. Pecoits-Filho, X.Q. Yu and N. Lameire, *Changes in the worldwide epidemiology of peritoneal dialysis*, Nat. Rev. Nephrol. **13** (2), 90-103, 2016.

[16] L.M. McCrink, A.H. Marshall and K.J. Cairns, *Advances in joint modelling: a review of recent developments with application to the survival of end stage renal disease patients*, Int. Stat. Rev. **81** (2), 249-269, 2013.

[17] L.C. Proust and P. Blanche, *Dynamic Predictions*, Wiley StatsRef-Statistics Reference Online, 2016.
[18] L.C. Proust and J. Taylor, *Development and validation of a dynamic prognostic tool for prostate cancer recurrence using repeated measures of posttreatment PSA: a joint modeling approach*, Biostatistics **10** (3), 535-549, 2009.
[19] R Core Team, *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, 2021. URL https://www.R-project.org/.
[20] D. Rizopoulos, *JM: an R package for the joint modelling of longitudinal and time-to-event data*, J. Stat. Softw. **35** (9), 1-33, 2010.
[21] D. Rizopoulos, *Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data*, Biometrics **67** (3), 819 - 829,2011.
[22] D. Rizopoulos, *Joint Models for Longitudinal and Time-to-Event Data with Applications in R.*, Chapman and Hall/CRC Biostatistics Series: Boca Raton, 2012.
[23] D. Rizopoulos, J.M. Taylor, J.V. Rosmalen, E.W. Steyerberg and J.J.M. Takkenberg, *Personalized screening intervals for biomarkers using joint models for longitudinal and survival data*, Biostatistics **17** (1), 149-164, 2016.
[24] A. Sayers, J. Heron, A.D.A.C. Smith, C. Macdonald-Wallis, M.S. Gilthorpe, F. Steele and K. Tilling, *Joint modelling compared with two stage methods for analysing longitudinal data and prospective outcomes: A simulation study of childhood growth and BP*, Stat. Methods Med. Res. **26** (1), 437-452, 2017.
[25] M.D. Schluchter, *Methods for the analysis of informatively censored longitudinal data*, Stat. Med. **11** (14-15), 1861-1870, 1992.
[26] R. Schoop, E. Graf and M. Schumacher, *Quantifying the predictive performance of prognostic models for censored survival data with time-dependent covariates*, Biometrics **64** (2), 603-610, 2008.
[27] M.H. Sipahioglu, A. Aybal, A. Unal, B. Tokgoz, O. Oymak and C. Utas, *Patient and technique survival and factors affecting mortality on peritoneal dialysis in Turkey: 12 years experience in a single center*, Perit Dial Int. **28** (3), 238-245, 2008.
[28] I. Sousa, *A review on joint modelling of longitudinal measurements and time-to-event*, Revstat Stat. J. **9** (1), 57-81, 2011.
[29] N. Tekkarismaz and D. Torun, *Long-term clinical outcomes of peritoneal dialysis patients: 9-year experience of a single centre in Turkey*, Turk. J. Med. Sci. **50** (2), 386-397, 2020.
[30] A.A. Tsiatis, *Joint modeling of longitudinal and time-to-event data: an overview*, SStatist. Sinica **14** (3), 809-834, 2004.
[31] X. Wang, Q. Han, T. Wang and W. Tang, *Serum albumin changes and mortality risk of peritoneal dialysis patients*, Int Urol Nephrol **52** (3), 565-571, 2020.
[32] H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*, Springer-Verlag, New York, 2016.
[33] L. Wu, W. Liu, G.Y. Yi and Y. Huang, *Analysis of longitudinal and survival data: Joint modeling, inference methods, and issues*, J. Probab. Stat. **2012**, 640153, 2012.
[34] M. Wulfsohn and A.A. Tsiatis, *A joint model for survival and longitudinal data measured with error*, Biometrics **53** (1), 330-339, 1997.

APPENDIX

We provided the convergence information for Model 1 in Table 2. The iteration histories of model parameters of the longitudinal and survival sub-models were given in Figures A1 and A2, respectively. The EM estimations were converged at iteration 48.

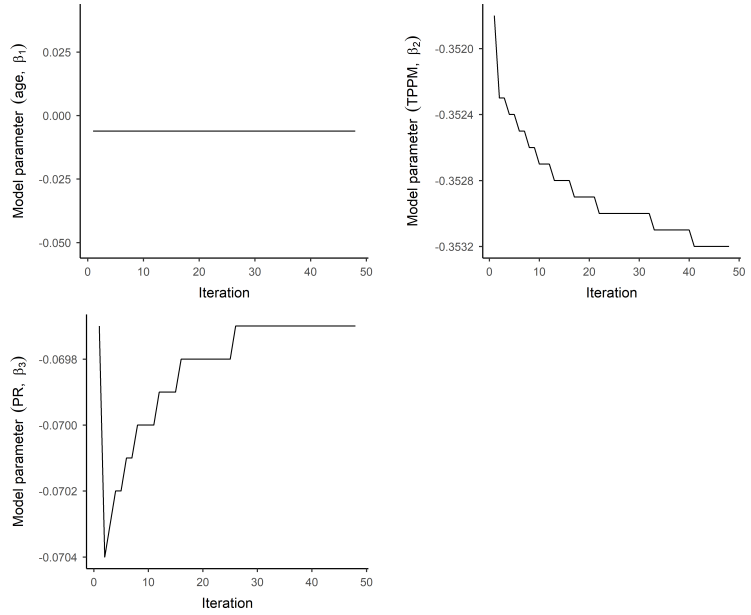


Figure A1. Iteration history of model parameters – Linear mixed effect sub-model of Model 1 given in Table 2.

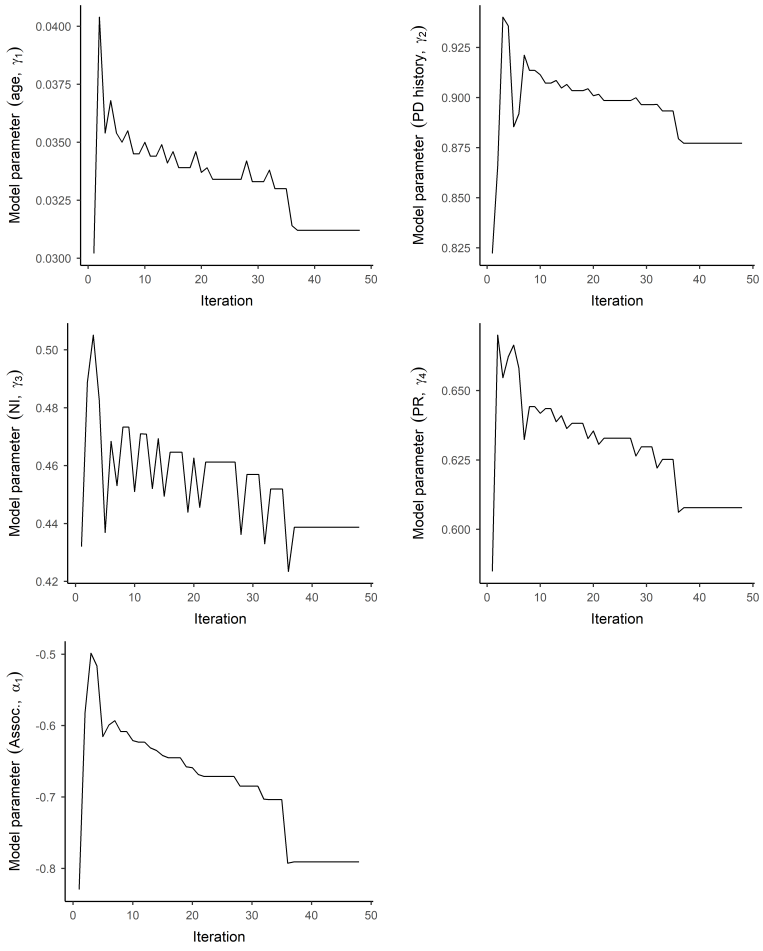


Figure A2. Iteration history of model parameters and log-likelihood of the model – Survival sub-model of Model 1 given in Table 2.