

## Research Article

# Improved Landmark Dynamic Prediction Model to Assess Cardiovascular Disease Risk in On-Treatment Blood Pressure Patients: A Simulation Study and Post Hoc Analysis on SPRINT Data

Mehrab Sayadi <sup>1,2</sup>, Najaf Zare <sup>3</sup>, Armin Attar,<sup>4</sup>  
and Seyyed Mohammad Taghi Ayatollahi <sup>2</sup>

<sup>1</sup>Cardiovascular Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

<sup>2</sup>Department of Biostatistics, School of Medicine, Shiraz University of Medical Sciences, Shiraz, Iran

<sup>3</sup>Department of Biostatistics, Infertility Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

<sup>4</sup>Department of Cardiovascular Medicine, TAHA Clinical Trial Group, School of Medicine, Shiraz University of Medical Sciences, Shiraz, Iran

Correspondence should be addressed to Seyyed Mohammad Taghi Ayatollahi; [ayatolahim@sums.ac.ir](mailto:ayatolahim@sums.ac.ir)

Received 16 November 2019; Revised 18 February 2020; Accepted 24 March 2020; Published 23 April 2020

Academic Editor: Momiao Xiong

Copyright © 2020 Mehrab Sayadi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Landmark model (LM) is a dynamic prediction model that uses a longitudinal biomarker in time-to-event data to make prognosis prediction. This study was designed to improve this model and to apply it to assess the cardiovascular risk in on-treatment blood pressure patients. A frailty parameter was used in LM, landmark frailty model (LFM), to account the frailty of the patients and measure the correlation between different landmarks. The proposed model was compared with LM in different scenarios respecting data missing status, sample size (100, 200, and 400), landmarks (6, 12, 24, and 48), and failure percentage (30, 50, and 100%). Bias of parameter estimation and mean square error as well as deviance statistic between models were compared. Additionally, discrimination and calibration capability as the goodness of fit of the model were evaluated using dynamic concordance index (DCI), dynamic prediction error (DPE), and dynamic relative prediction error (DRPE). The proposed model was performed on blood pressure data, obtained from systolic blood pressure intervention trial (SPRINT), in order to calculate the cardiovascular risk. Dynpred, coxme, and coxphw packages in the R.3.4.3 software were used. It was proved that our proposed model, LFM, had a better performance than LM. Parameter estimation in LFM was closer to true values in comparison to that in LM. Deviance statistic showed that there was a statistically significant difference between the two models. In the landmark numbers 6, 12, and 24, the LFM had a higher DCI over time and the three landmarks showed better performance in discrimination. Both DPE and DRPE in LFM were lower in comparison to those in LM over time. It was indicated that LFM had better calibration in comparison to its peer. Moreover, real data showed that the structure of prognostic process was predicted better in LFM than in LM. Accordingly, it is recommended to use the LFM model for assessing cardiovascular risk due to its better performance.

## 1. Introduction

The risk prediction models (RPMs) are used as a diagnostic model to estimate the probability of an event occurrence in a disease or as a prognostic model to estimate probable consequences of a disease. Accurate prediction of a risk is essential in clinical research, and it is the patient's right to be

informed about their disease progress [1]. Recently, RPMs are being used to help clinicians to make the best decision in diagnostic and therapeutic approaches, based on patient's demographics, test results, or disease characteristics [2]. Diagnostic models are usually used for risk classification in patients while prognostic models use time to assess disease progress [3, 4]. Nowadays, more prediction models have been

used in cardiovascular diseases, such as diagnostic models for assessing the risk factors [3]. However, the cardiovascular risk assessment tools are static prediction models that use baseline predictors, but they still have some shortcomings, such as poor prediction [5], for instance, the inability to determine long-term survival of a heart attack patient with previous successful treatment or the inability to decrease the risk of cardiovascular event in a treated hypertensive patient. During the intervention, biomarkers are measured that are potentially informative in order to determine the treatment efficacy [6–9]. In this respect, the risk prediction using longitudinal biomarker is referred to as the dynamic prediction model (DPM), which was introduced by some researchers [10–13]. One DPM model is joint modeling (JM) [14–16], which requires correct determination of biomarker process distribution and event time, but this biomarker distribution is usually unclear. Moreover, its generalization to more than one marker leads to the production of ample computational complexity [17]. The landmark model (LM), another DPM, is used as an appropriate alternative to JM [9, 17–19]. The main advantage of LM is its simplicity, since it requires fewer assumptions compared to JM and might have much more power. In LM, the time is divided into different landmarks and then the simple Cox proportional hazards (PH) model is applied to each landmark for individuals who are still alive until time  $t$  [20, 21]. On the other hand, the biomarker value in each landmark time is considered as a fixed variable; hence, the prediction of risks becomes feasible. A landmark window should be considered to predict survival until time  $sl + w$ , which is called  $t$  horizon (thor).  $w$  is the length of time to predict patient survival as the prediction window.

By analyzing LM, the less frail patients are probably maintained dynamically during the landmark times. On the other hand, the estimated parameters in LM can be affected, if some patients do not follow the specific clinical visit schedule. Also, not considering the correlation between landmarks might affect the risk prediction. Bias in LM probably originates from these neglected issues. In order to improve LM, the frailty parameter was used to present a new model called the landmark frailty model (LFM). Finally, LFM was used to assess the cardiovascular risk in the on-treatment hypertensive patients. To reach this goal, a study with simulation data was designed, and the real blood pressure data which was obtained from systolic blood pressure intervention trial (SPRINT) was analyzed [22, 23].

The rest of this article is organized as follows. Section 2 provides a brief description of the landmark approach as well as the proposed approach. Also, the setting of simulation studies, goodness of fit indices, and real data description are shown in Section 2. We conducted simulation studies to compare LFM with LM in Sections 3. In Section 4, we exhibited our approach with the SPRINT data followed by Section 5, which concludes and discusses simulation and real data.

## 2. Materials and Methods

**2.1. Landmark Approach.** Assuming that  $T_i$  and  $C_i$  are survival time (failure time) and censoring time, then  $T_i^* = \min(T_i, C_i)$  make the observed time.  $X(\cdot)$  represents the vector of covariates, which can be measured once at the beginning of the study. For example, age and gender are measured only

at baseline and they are considered as fixed variables.  $Y(\cdot)$  represents the longitudinal biomarker like systolic or diastolic blood pressure which can be measured for several time intervals. For risk assessment, the Cox (PH) model as the most famous model is defined as

$$h(t) = h_0(t) \exp(X\beta), \quad (1)$$

where  $h(t)$  and  $h_0(t)$  are hazard function at time  $t$  and baseline hazard, respectively. In LM, the time is divided into several landmark times including  $s_1, s_2, \dots, s_l$ . At landmark  $l$  ( $l = 1, 2, \dots, K$ ), subjects who are still at risk are considered for analysis and the remaining individuals will be omitted [9]. At each landmark, longitudinal biomarker value  $Y(s)$  is considered as a fixed variable. Then, a time period, a landmark window is considered to predict survival until time  $sl + w$  which is called  $t$  horizon (thor).  $w$  is the length of time to predict patient survival as prediction window, which is the so-called 3 or 5 years. The Cox PH model in equation (1) is reformulated and the conditional hazard function is estimated by

$$\begin{aligned} h(t | s, Y(S), X, w) \\ = h_{l,0}(t | s, w) \exp(Y(s_l)\alpha_l + X\beta_l), s \leq t \leq s + w. \end{aligned} \quad (2)$$

The model presented in equation (2) is defined as the simple or basic LM. It is used to fit a model to each landmark, and it estimates the specific landmark effect of a biomarker for predicting survival between  $s_l$  and  $t_{\text{thor}}$  where  $h_{l,0}$  is a different baseline hazard in each landmark. We assumed that a longitudinal biomarker,  $Y_i(t_{ij})$ , for subject  $i$  at the time of  $j$  was obtained from the mixed-effect model via the following formula:

$$Y_i(t_{ij}) = Z_i(t_{ij})b_{g_i} + X_i(t_{ij})\beta + \varepsilon_i, \quad (3)$$

where  $Z_i$  and  $X_i$  denote the design vector for random and fixed effect and subscript  $g$  is 0 or 1.

According to equation (2), to consider the frailty of patients and the correlation between sequential measurements, LFM is defined as follows:

$$\begin{aligned} h(t | s, Y(s), X, u, w) \\ = h_{l,0}(t | s, w) \exp(Y(s_l)\alpha_l + X\beta_l + u_i), s \leq t \leq s + w, \end{aligned} \quad (4)$$

where  $u_i$  indicates the frailty of patient  $i$ , which follows the multivariate normal distribution with mean 0 and covariance matrix  $\Sigma(\theta)$ . The survival prediction model is related to cumulative hazard function by

$$\begin{aligned} S(t + w | s, Y(s), X, u, w) \\ = \exp(-H(t | s, Y(s), X, u, w)), s \leq t \leq s + w. \end{aligned} \quad (5)$$

To estimate the parameters, Cox partial likelihood is modified via the following integrated (over landmarks) partial log-likelihood (IPL) [24].

$$IPL(\alpha, \beta, u) = \sum_{i=1}^n d_i \sum_l \ln \left( \frac{\exp(Y_i(s_l)\alpha_l + X_i\beta_l + u_i)}{\sum_{j \in R(s_l)} \exp(Y_j(s_l)\alpha_l + X_j\beta_l + u_j)} \right). \tag{6}$$

In this formula,  $d_i$  indicates the risk set,  $d_i = 1$  if subject  $i$  remains until time  $s$  at landmark  $l$ . Otherwise,  $d_i$  is assumed 0.  $R(s_l)$  denotes the risk set at time  $s$  at landmark  $l$ . We used the integrated partial likelihood (IPL\*) by integrating the random effects [25].

$$IPL^*(\alpha, \beta, \theta) = \int IPL(\alpha, \beta, u) f(u) du. \tag{7}$$

By maximizing the IPL\*, maximum likelihood estimators (MLE) for the parameters are provided. In addition, the coxme function in coxme package can only provide an ML estimate. With respect to the complexity of IPL\* calculation, the coxme package uses the Laplace approximation technique. More details are described elsewhere [25, 26].

We can also perform a model for all landmarks by stacking data set defined as super LFM, which considers parameters as if they depended on the time in a smooth fashion. It is formulated as

$$h(t | s, Y(s), X, u, w) = h_0(t | s, w) \exp(Y(s)\alpha(s) + X\beta_s + u_i), \tag{8}$$

where

$$\alpha(s) = \sum_{j=1}^{m_h} \gamma_j f_j(s). \tag{9}$$

Dynpred, coxme, and coxphw packages in the R.3.4.3 software were used for data analysis.

**2.2. Simulation Study Setting.** To assess the application of the models in different aspects, we set up several scenarios with different specificities in terms of sample size ( $n = 100, 200,$  and  $400$ ), number of landmarks (6, 12, 24, and 48), failure rate (30, 50, and 100%), and complete/missing data. In order to perform these models, a dichotomous variable with binomial distribution like treatment effect and continuous covariate with normal distribution like age ( $X_1$  and  $X_2$ , respectively) were considered. The regression coefficient,  $\beta_1$  and  $\beta_2$ , was set at 0.5 and 1.5 as true values, respectively, for  $X_1$  and  $X_2$ . In equation (3), we assume that  $b_g$  has a bivariate normal distribution for random intercept and slope with a mean of 0 and covariance matrix of  $\delta_{11} = 2, \delta_{12} = 0.2, \delta_{22} = 1$ . We also assumed that the individual error term ( $\varepsilon_i$ ) follows a normal distribution with a mean of 0 and variance of 1. Moreover, it is assumed that continuous variable  $Y$  was measured for 10 times for each individual sequentially. The time  $T$  was generated from Weibull distribution [27] as shown in the following:

$$T = \left[ \frac{1}{k} \ln \left( 1 + \frac{(1 + \lambda)(-\ln(v))}{\gamma \exp(X\beta)\varphi\lambda} \right) \right]^{1/(1+\lambda)}. \tag{10}$$

In this equation,  $k = 1.1, \gamma = 0.4, \lambda = 0.01, \varphi = 0.75,$  and  $v$  has a uniform (0, 1) distribution. Moreover, we performed the simple Cox model that just included the baseline data in three different sample sizes and three different failure percentages.

**2.3. Goodness of Fit (GOF) and Prediction Ability Indices.** There are several indices to assess the goodness of fit (GOF) and prediction ability in DPM. We used the standard error, bias of parameter estimation, and the mean square error (MSE), which were obtained from 300 simulation data. To compare LFM with LM, log-likelihood as well as deviance statistic was used. The latter is compared with mixture chi-square value (1.92) that was obtained from  $(1/2)(\chi_0^2 + \chi_1^2)$ . Akaike information criterion (AIC) was used as if smaller AIC implies a better fit. Moreover, the dynamic concordance index (DCI), dynamic prediction error (DPE), and dynamic relative prediction error (DRPE) were used to measure the discrimination and calibration ability. DPE was obtained from the Brier error score formula:

$$\text{Brier error} = \frac{1}{n} \sum_{i=1}^n (d_i(t) - \widehat{S}_i(t)^2). \tag{11}$$

In fact, the Brier score measures the average discrepancies between true event status and predictive values of survival at time  $t$ . Low Brier score of a model indicates the better predictive performance of that model. In this formula,  $d_i$  as the actual observation for subject  $i$  at time  $t$  is an event status, which could be either 1 or 0 (the occurrence or nonoccurrence of an event, respectively).

The predicted survival ( $\widehat{S}_i$ ) is estimated by model LM or LFM [28]. DRPE was calculated from

$$\text{Relative Prediction Error} = \left( 1 - \frac{\text{null model error}}{\text{current model error}} \right). \tag{12}$$

In equation (12), errors are obtained from equation (11) and the null model is a model without any covariates, such as Kaplan-Meier estimate, and the current model is LM or LFM.

**2.4. Real Data.** We used a part of the systolic blood pressure intervention trial (SPRINT) study [29] (National Heart, Lung, and Blood Institute (NHLBI), funded by the National Institutes of Health; ClinicalTrials.gov number NCT01206062) upon a request ID of 4612. In the main study of SPRINT, methods were reported in detail [30]. In summary, in that randomized controlled trial study, 9361 nondiabetic participants with systolic blood pressure (SBP) of equal or more than 130 mmHg were allocated to an intensive treatment (target SBP < 120 mmHg) and standard treatment (target SBP < 140 mmHg) groups. Baseline data, lab data, and repeated measurement of SBP for 21 times in 5 years were collected.

TABLE 1: Simple Cox model results in different simulations.

$n$	Estimate	Failure = 30%		Failure = 50%		Failure = 100%	
		$\beta_1 = 0.5$	$\beta_2 = 1.5$	$\beta_1 = 0.5$	$\beta_2 = 1.5$	$\beta_1 = 0.5$	$\beta_2 = 1.5$
100	Mean	0.384	1.290	0.384	1.286	0.401	1.103
	SE	0.513	0.368	0.543	0.392	0.318	0.226
	Bias	-0.116	-0.210	-0.115	-0.213	-0.099	-0.397
	MSE	0.677	0.464	0.377	0.304	0.158	0.128
200	Mean	0.402	1.327	0.417	1.298	0.420	1.285
	SE	0.750	0.549	0.246	0.174	0.252	0.180
	Bias	-0.098	-0.173	-0.082	-0.202	-0.08	-0.215
	MSE	0.720	0.692	0.065	0.084	0.080	0.097
400	Mean	0.405	1.287	0.438	1.308	0.431	1.300
	SE	0.320	0.226	0.252	0.180	0.175	0.123
	Bias	-0.095	-0.213	-0.062	-0.191	-0.069	-0.200
	MSE	0.106	0.326	0.368	0.125	0.036	0.059

Heart failure, stroke, myocardial infarction, other acute coronary syndromes, and death from cardiovascular causes were regarded as cardiovascular events. Hence, we designed a case-cohort study from this data, which included Framingham risk factors of age, gender, total cholesterol (TCH) level, high-density lipoprotein cholesterol (HDL) level, and SBP. In our model, treatment effect was added to the abovementioned data. We considered 10 measurements of SBP (baseline, 6, 12, 18, 24, 30, 36, 42, 48, and 54 months). The aim was to determine the dynamic effect of blood pressure on cardiovascular disease risk by comparing LFM with LM. On the other hand, these two models were compared with the simple Cox model by considering only baseline blood pressure data. To compare LFM with LM, we used AIC and deviance criteria. And the deviance criteria were tested using mixture chi-square.

### 3. Results of Simulation

Simple Cox model results are summarized in Table 1, and results of LM and LFM are summarized in Tables 2–4

**3.1. Landmark Models vs. Simple Cox Model.** Both LFM and LM had a better relative performance in comparison to the simple Cox model. As can be seen in all scenarios, bias and MSE of bias were lower in LMs; however, this difference decreased as the sample size increased from 100 to 200-400. Nevertheless, bias did not have basic changes in failure rate.

**3.2. Comparing LFM vs. LM.** The performance of the models was evaluated based on their ability to estimate the true value of the parameters and their ability to classify and predict the actual survival.

**3.2.1. Ability to Estimate the True Value of the Parameters.** Mean of parameter estimation and its SE, bias, and MSE for failure rate (30%) are shown in Table 2. In total, bias and MSE were lower in LFM in comparison with LM. According

to deviance and AIC indices, using 12 landmarks and sample size of 100, there was no statistically significant difference between LFM and LM in both data sets. Deviance was 1.76 in complete data and 1.11 in incomplete data. Deviance index showed that there was a statistically significant difference between the two models at the sample size of 200 and 400 in the complete data. At the sample size of 200, AICs were 1515 and 3636 in LFM and LM, respectively, while at the sample of 400 they were 1521 and 3648.

In these cases, bias and MSE of bias for the two parameters were slightly lower in LFM. In the incomplete data, according to deviance index, there was no significant difference between the two models in all sample sizes, while in the complete data (200-400) statistically significant difference was observed (in both cases, deviance was greater than mixture chi-square and AICs were lower in LFM). In all scenarios with 24 and 40 landmarks, based on deviance and AIC indices, LFM fitted better in comparison to LM. In these cases, the mean estimation of two parameters in the LFM was closer to the true value. This result was more pronounced in the continuous variable. According to the results of failure rate of 50% and 100% summarized in Tables 3 and 4, the superiority of LFM over LM was higher, especially in the model with 12 landmarks.

**3.2.2. Ability to Classify and Predict the Actual Survival.** We used CDI to assess the discrimination ability of the two models that were run with fixed sample size and failure rate in 3 different landmarks (6, 12, and 24), where CDI value is greater than 0.5, indicating that the model had discrimination ability. As illustrated in Figure 1, LFM had better performance. This advantage was more evident by increasing the number of landmarks from 6 to 12-24. On the other hand, the more area under the curve indicated the more accurate model. DPE and DRPE for calibration ability were plotted in Figures 2 and 3. The error rates and relative error rates in LFM were much lower than those in LM, and this became more prominent by increasing landmark numbers.

TABLE 2: LM and LFM results in different scenarios when failure rate is 30%.

L	n	Estimate	Data without missing				Data with missing					
			LM $\beta_1 = 0.5$	LM $\beta_2 = 1.5$	LFM $\beta_1 = 0.5$	LFM $\beta_2 = 1.5$	LM $\beta_1 = 0.5$	LM $\beta_2 = 1.5$	LFM $\beta_1 = 0.5$	LFM $\beta_2 = 1.5$	AIC	D
100		Mean	0.418	1.406	0.422	1.429	0.408	1.421	0.435	1.422		
		SE	0.224	0.182	0.219	0.180	0.339	0.274	0.329	0.280		
		Bias	-0.081	-0.093	-0.077	-0.070	-0.091	-0.078	-0.065	-0.078	241	1.11
		MSE	0.201	0.138	0.210	0.140	0.315	0.680	0.270	0.177		
200		Mean	0.459	1.405	0.465	1.420	0.458	1.402	0.466	1.422		
		SE	0.149	0.125	0.150	0.123	0.217	0.180	0.220	0.183		
		Bias	-0.040	-0.095	-0.034	-0.079	-0.042	-0.097	-0.033	-0.077	621	1.01
		MSE	0.087	-0.064	0.086	0.063	0.107	0.328	0.110	0.323		
400		Mean	0.463	1.400	0.468	1.415	0.458	1.411	0.464	1.430		
		SE	0.103	0.086	0.104	0.085	0.147	0.122	0.150	0.124		
		Bias	-0.036	-0.100	-0.032	-0.085	-0.041	-0.089	-0.035	-0.070	1532	1.79
		MSE	0.039	0.039	0.040	0.038	0.053	-0.049	0.054	0.048		
100		Mean	0.419	1.405	0.449	1.509	0.431	1.421	0.504	1.536		
		SE	0.171	0.144	0.155	0.128	0.232	0.194	0.285	0.203		
		Bias	-0.081	-0.094	-0.051	0.009	-0.069	-0.078	0.004	0.036	433	25.41
		MSE	0.199	0.139	0.234	0.125	0.265	0.174	0.414	0.306		
200		Mean	0.460	1.400	0.495	1.500	0.461	1.399	0.539	1.590		
		SE	0.105	0.024	0.115	0.020	0.152	0.127	0.181	0.152		
		Bias	-0.039	-0.100	-0.005	0.000	-0.038	-0.101	0.038	0.059	1143	52.59
		MSE	0.090	0.066	0.098	0.064	0.110	0.086	0.096	0.120		
400		Mean	0.463	1.401	0.498	1.500	0.456	1.412	0.513	1.579		
		SE	0.079	0.060	0.072	0.020	0.104	0.086	0.122	0.102		
		Bias	-0.037	-0.099	-0.002	0.000	-0.044	-0.088	0.013	0.079	2854	107.85
		MSE	0.039	0.039	0.040	0.036	0.052	0.048	0.069	0.074		

TABLE 2: Continued.

$L$	$n$	Estimate	Data without missing				Data with missing			
			$\beta_1 = 0.5$	$\beta_2 = 1.5$	AIC	D	$\beta_1 = 0.5$	$\beta_2 = 1.5$	AIC	D
100		Mean	0.421	1.409	0.462	1.545	0.431	1.420	0.534	1.558
		SE	0.127	0.106	0.145	0.123	0.191	0.159	0.251	0.213
		Bias	-0.078	-0.090	-0.038	0.045	-0.069	-0.079	0.034	0.058
		MSE	0.200	0.137	0.246	0.160	0.261	-0.174	0.469	0.395
200		Mean	0.460	1.399	0.504	1.544	0.460	1.401	0.556	1.564
		SE	0.097	0.083	0.086	0.072	0.125	0.105	0.155	0.132
		Bias	-0.040	-0.101	0.004	0.044	-0.040	-0.099	0.056	0.064
		MSE	0.105	0.074	0.091	0.064	-0.111	0.080	0.168	0.161
400		Mean	0.465	1.403	0.513	1.541	0.458	1.413	0.538	1.563
		SE	0.060	0.050	0.067	0.050	0.086	0.072	0.105	0.088
		Bias	-0.035	-0.096	0.013	0.041	-0.042	-0.068	0.038	0.063
		MSE	0.048	0.040	0.040	0.040	0.054	0.048	0.078	0.099
40		Mean								
		SE								
		Bias								
		MSE								

$L$  = landmarks;  $n$  = sample size; LM = landmark model; LFM = landmark frailty model; AIC = Akaike information criterion;  $D$  = deviance; SE = standard error; MSE = mean square error.



TABLE 3: Continued.

$L$	$n$	Estimate	Data without missing						Data with missing										
			$\beta_1 = 0.5$	$\beta_2 = 1.5$	LM	AIC	$\beta_1 = 0.5$	$\beta_2 = 1.5$	LFM	$D$	AIC	$\beta_1 = 0.5$	$\beta_2 = 1.5$	LM	AIC	$D$			
100		Mean	0.438	1.377		0.486	1.520		0.438	1.403		0.535	1.568		0.438	1.403		0.535	1.568
		SE	0.096	0.080		0.110	0.093		0.143	0.135		0.185	0.158		0.143	0.135		0.185	0.158
		Bias	-0.062	-0.123		-0.014	0.020		-0.062	-0.097		0.035	0.068		-0.062	-0.097		0.035	0.068
200		MSE	0.095	0.083		0.121	0.095		0.128	0.101		0.237	0.250		0.128	0.101		0.237	0.250
		Mean	0.452	1.407		0.500	1.525		0.455	1.413		0.552	1.561		0.455	1.413		0.552	1.561
		SE	0.077	0.055		0.074	0.024		0.096	0.080		0.120	0.103		0.096	0.080		0.120	0.103
400		Bias	-0.048	-0.092		0.000	0.025		-0.045	-0.087		0.052	0.061		-0.045	-0.087		0.052	0.061
		MSE	0.043	0.040		0.051	0.045		0.053	0.053		0.086	0.128		0.053	0.053		0.086	0.128
		Mean	0.457	1.402		0.503	1.539		0.458	1.409		0.509	1.558		0.458	1.409		0.509	1.558
40		SE	0.046	0.038		0.051	0.040		0.066	0.055		0.045	0.038		0.066	0.055		0.045	0.038
		Bias	-0.043	-0.098		0.003	0.039		-0.042	-0.091		0.009	0.058		-0.042	-0.091		0.009	0.058
		MSE	0.022	0.026		0.025	0.025		0.029	0.031		0.049	0.028		0.029	0.031		0.049	0.028

$L$  = landmarks;  $n$  = sample size; LM = landmark model; LFM = landmark frailty model; AIC = Akaike information criterion;  $D$  = deviance; SE = standard error; MSE = mean square error.



TABLE 4: LM and LFM results in different scenarios when failure rate is 100%.

L	n	Estimate	Data without missing						Data with missing					
			LM			LFM			LM			LFM		
			$\beta_1 = 0.5$	$\beta_2 = 1.5$	D	AIC	$\beta_1 = 0.5$	$\beta_2 = 1.5$	D	AIC	$\beta_1 = 0.5$	$\beta_2 = 1.5$	D	AIC
100	Mean	0.390	1.197		0.395	1.216		0.383	1.205		0.392	1.232		
	SE	0.099	0.083		0.102	0.085		0.147	0.123		0.150	0.120		
	Bias	-0.110	-0.303	1785	-0.105	-0.284	1776	-0.117	-0.295	707	-0.108	-0.268	703	
	MSE	0.045	0.035		0.046	0.034		0.064	0.046		0.064	0.045		
12	Mean	0.474	1.394		0.480	1.411		0.480	1.390		0.488	1.421		
	SE	0.081	0.066		0.080	0.068		0.115	0.096		0.118	0.099		
	Bias	-0.026	-0.106	5147	-0.020	-0.089	5135	-0.020	-0.110	2114	-0.012	-0.079	2109	
	MSE	0.023	0.027		0.023	0.025		0.029	0.031		0.029	0.028		
400	Mean	0.464	1.393		0.468	1.408		0.465	1.397		0.472	1.420		
	SE	0.055	0.046		0.057	0.047		0.080	0.067		0.082	0.068		
	Bias	-0.036	-0.107	12176	-0.032	-0.091	12155	-0.036	-0.103	5156	-0.028	-0.080	5145	
	MSE	0.012	0.018		0.011	0.015		0.015	0.021		0.015	0.017		
100	Mean	0.388	1.195		0.410	1.275		0.382	1.204		0.432	1.382		
	SE	0.070	0.058		0.077	0.065		0.104	0.087		0.124	0.105		
	Bias	-0.112	-0.305	3567	-0.090	-0.225	3455	-0.117	-0.295	1408	-0.067	-0.118	1324	
	MSE	0.045	0.039		0.054	0.036		0.062	0.046		0.088	0.088		
24	Mean	0.475	1.395		0.500	1.478		0.480	1.398		0.546	1.477		
	SE	0.056	0.048		0.061	0.051		0.082	0.068		0.095	0.080		
	Bias	-0.025	-0.105	10297	0.000	-0.022	10071	-0.020	-0.102	4228	0.046	-0.023	4056	
	MSE	0.022	0.028		0.027	0.025		0.029	0.031		0.044	0.043		
400	Mean	0.464	1.391		0.486	1.465		0.464	1.397		0.514	1.537		
	SE	0.039	0.032		0.042	0.035		0.056	0.047		0.065	0.054		
	Bias	-0.036	-0.109	24354	-0.014	-0.035	23904	-0.036	-0.103	10313	0.014	0.037	9964	
	MSE	0.012	0.018		0.012	0.011		0.015	0.021		0.022	0.020		



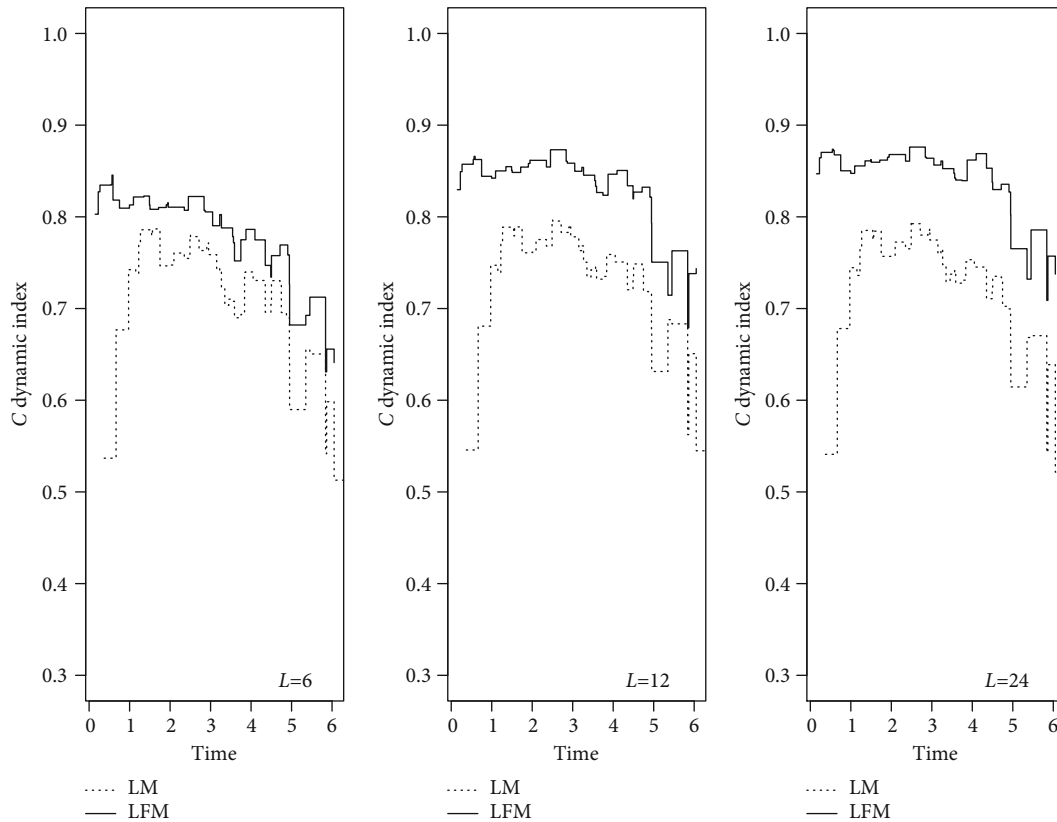


FIGURE 1: Simulated C dynamic index for landmark model (LM) and landmark frailty model (LFM): landmarks = 6, 12, and 24; sample size = 200; and failure rate = 50%. The higher values of C index indicate more discrimination ability.

#### 4. Results of Real Data

Results of real data are summarized in Table 5 and Figure 4. The adjusted hazard ratio of variables and its confidence interval (CI) are provided. Furthermore, AIC and deviance index were extracted to assess the models. While in both LFM and LM the SBP was highly significant, it had no significant impact on cardiovascular events in simple Cox ( $p$  value = 0.258). However, LFM fitted better since it had an AIC equal to 5437 while it was 6385 in LM. Also, the deviance between two models was 559.7 ( $p < 0.001$ ). After adjusting the treatment effect as well as baseline risk factor effect (Figure 4), it was shown that while SBP was decreasing over time, the hazard ratio (HR) was decreasing in line with SBP in both models. However, it is noteworthy that this reduction was more in the LFM. On the contrary, HR was constant over time in the Cox model. As the blood pressure decreases, the 3-year survival prediction increases. LFM predicts higher survival than LM and the simple Cox model (Figure 5).

#### 5. Discussion

5.1. Discussion of Simulation Data. DPM includes time-dependent marker information during follow-up in order to improve personal survival prediction probabilities. At any follow-up, time-updated marker value can be used to generate a dynamic prediction [10–12]. These models are

essential to identify high-risk individuals and timely clinical decision-making. Recently, LM as DPM was extensively investigated by researchers [9, 24]. Some of them used LM in different aspects of survival data such as competing risk and cure data [14, 20]. However, they paid little attention to individuals' frailty and regularity of visits as well as correlation between different landmarks. On the other hand, LM can be affected by the way how landmarks are selected and number of landmarks. Ignoring these issues might lead to an estimation error. Hence, we proposed a modified LM which used the frail parameter as LFM. Indeed, in LM, individuals who experienced the intended event or being censored at a defined landmark time are considered in data analysis. Frailty plays a critical role in those who are retained and repeated dynamically in sequential landmarks due to their low frailty. However, in our proposed model, considering the frailty of patients was included in the analysis; hence, we were able to overcome the abovementioned problems. Simulations showed that both LFM and LM had a relative advantage over the simple Cox model. This was confirmed by various criteria such as bias and MSE. Bias and MSE in dichotomous and continuous variables were higher in the simple Cox model, which is in line with other studies [20, 31]. In the simple Cox model, as the sample size increased, the estimation error slightly decreased and the estimates were closer to the true values. But the simple Cox model was still behind the LMs. LFM and LM were compared regarding different

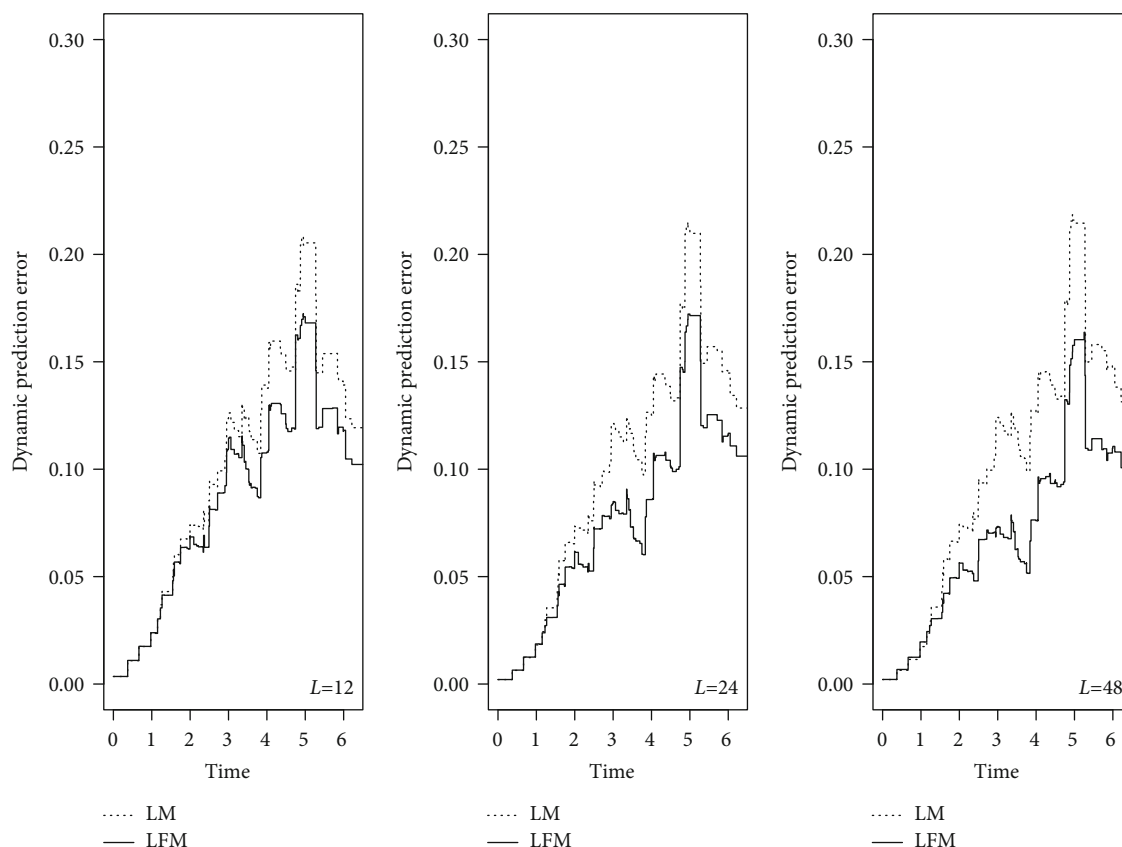


FIGURE 2: Simulated dynamic prediction error for landmark model (LM) and landmark frailty model (LFM): landmarks = 6, 12, and 24; sample size = 200; and failure rate = 50%. The lower values of prediction error indicate more calibration.

sample sizes, different number of landmarks, different failure rate, and diverse data structure. Generally, the superiority of the LFM over LM was confirmed in the present study. This conformation was very clear in a large sample size and higher landmark number in both complete and incomplete data.

To the best of our knowledge, this is the first study to investigate the effect of number of landmarks on accuracy of the results. Wright et al. [31] performed LM with 20 landmarks, and others have empirically found that 20 to 100 landmarks are appropriate [24]. However, in the case of large sample size, data become too large, and it took too much time to run the programs. There was no significant difference between LM and LFM in small sample size and low number of landmarks, which was the result of checking the deviance and AIC indices between the two models. Both models performed better by increasing failure rate. Although no statistical comparison was made, in each case, LFM was more appropriately fitted with fewer estimation errors. In most times, the discrimination ability of LFM was more than LM since DCI was more than 70% in LFM while this index was lower in LM. The difference between DCIs increased by the implementation of increased number of landmarks. Also, evaluation of collaboration ability (DPE and DRPE) showed that LFM had a better performance than LM.

**5.2. Discussion of SPRINT Data.** Hypertension is not only recognized as a major cardiovascular risk factor but also

has a significant impact on the occurrence of events followed by therapeutic interventions [32]. In this study, we showed that considering hypertension as a dynamic risk factor had a basic role for obtaining real estimation of successful treatment in cardiovascular diseases. Hence, it should be considered dynamically during the course of treatment and not only at the time of admission as a primary measure that it was emphasized by other similar studies [9, 33]. SPRINT data confirmed the simulation results, which contained repeated measurements of SBP as a single longitudinal biomarker. As mentioned in the previous section, more than one biomarker could be used in LMs. By only using the baseline blood pressure data (simple Cox model), the role of SBP was hindered due to dominance of treatment effect; hence, it was not recognized as a cardiovascular risk factor. This result is in line with the results of our previous study [22] and the same result was obtained from the study carried out by Group S.R. [29]. In both landmark models, as SBP decreased after treatment, the risky effect of SBP was also reduced. While HR in the simple Cox model is close to 1 and constant over time, HR in the two landmark models was close to 3 at the beginning of the study and then it relatively decreased by decreasing blood pressure over time, although the intensity of the significant reduction was higher in our proposed model. On the other hand, in the simple Cox model, the effect of intensive treatment was 38% ( $1-1/0.720$ ) in comparison with standard treatment, while in LFM and

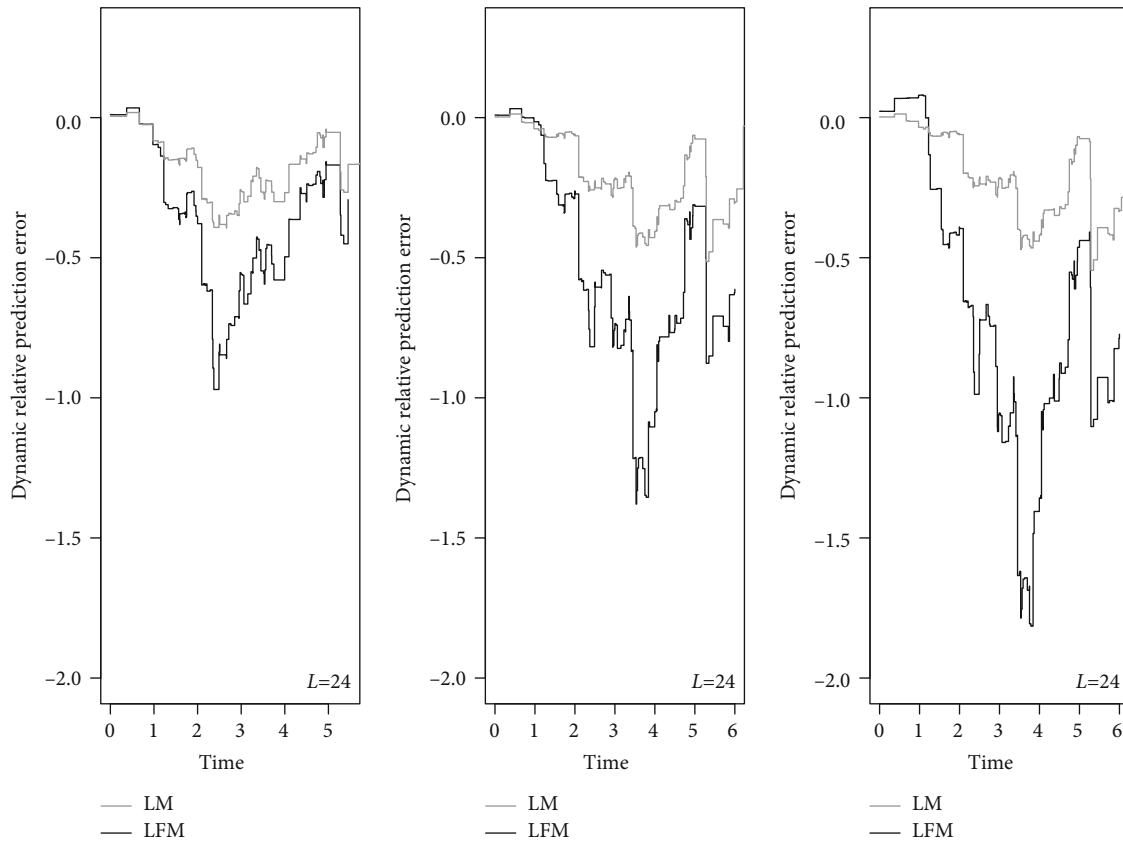


FIGURE 3: Simulated dynamic relative prediction error for landmark model (LM) and landmark frailty model (LFM): landmarks = 6, 12, and 24; sample size = 200; and failure rate = 50%. The figures show which model has been able to reduce the error more.

TABLE 5: Static and dynamic effect of SBP on cardiovascular event.

Variables	Simple Cox		LM		LFM	
	HR	<i>p</i> value	HR	<i>p</i> value	HR	<i>p</i> value
Age at enrolment day, yr	1.051	<0.001	1.018	<0.001	1.022	<0.001
Gender (female)	0.691	<0.001	1.019	0.164	0.843	0.291
TCH (mmol/L)	1.001	<0.001	1.001	0.163	1.000	0.582
HDL-C (mmol/L)	0.986	<0.001	0.991	0.001	0.987	0.023
Current smoker	1.820	<0.001	1.178	0.040	1.271	0.045
Treatment, intensive	0.720	<0.001	0.482	<0.001	0.475	<0.001
SBP (mmHg)	1.001	0.258	1.083	<0.001	1.109	<0.001
SBP*time	—	—	0.996	<0.001	0.839	<0.001
SBP*time <sup>2</sup>	—	—	0.966	<0.001	0.895	<0.001

LM = landmark model; LFM = landmark frailty model; HR = hazard ratio; TCH = total cholesterol; HDL-C = high-density lipoprotein; SBP = systolic blood pressure.

LM it was 110% and 107%, respectively. This means that the protective effect of intensive treatment is highly exhibited in our model in comparison with simple Cox and LM. Thereby, predictability of the 3-year dynamic survival is higher in LFM. Other studies that worked on blood pressure have considered landmarks separately while we used a model that landmarks were considered continuously [21, 33, 34]. This study showed that landmark models can be used to help clinicians to make better decision for diagnosis and treatment.

Landmark models, especially our proposed model, are useful for risk assessment, when the data is not complete or regular, similar to our data.

### 6. Conclusion

In this study, we provided a modified LM, which considered the frailty of the patients as well as the correlation between the landmarks. Our approach can be fitted better in the sense

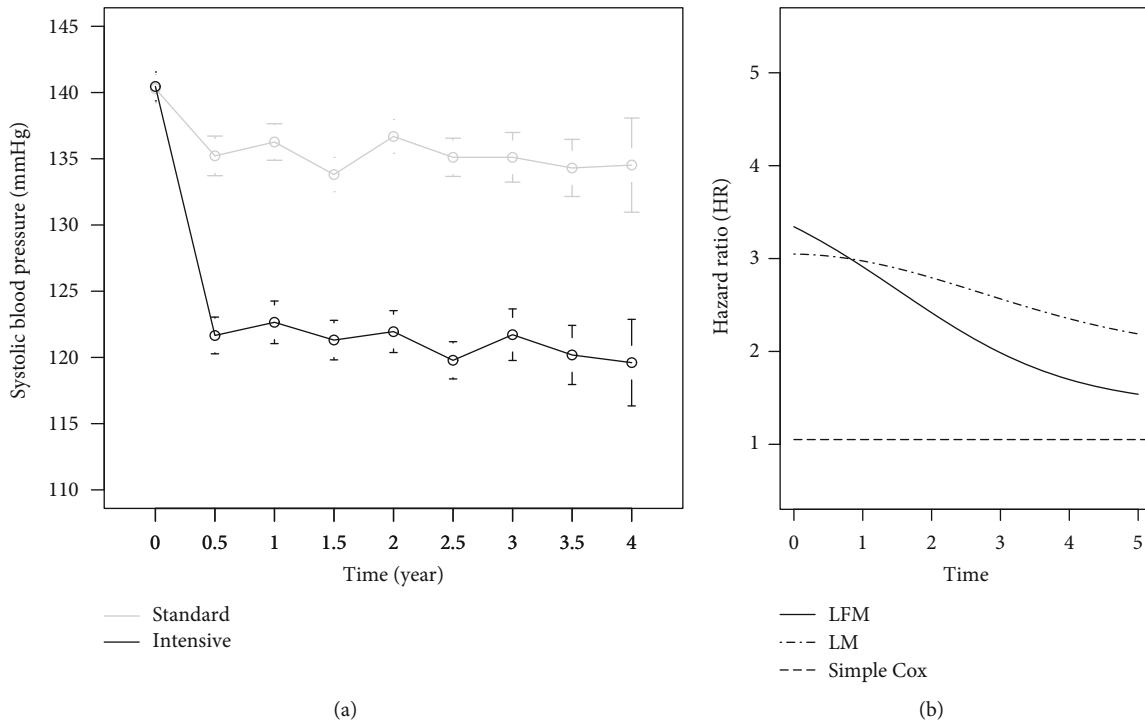


FIGURE 4: (a) Systolic blood pressure (SBP) in the two treatment groups over time. Target of SBP in the intensive group and the standard group was less than 120 mmHg and 140 mmHg, respectively. (b) Hazard ratio prediction of systolic blood pressure over time adjusted for treatment effect and other covariates. It is shown that as blood pressure decreases over time due to treatment effect, the risk also decreases. This reduction is more in LFM. Also, the hazard ratio in the simple Cox model (included just baseline SBP) is fixed over time. AIC = 5437 and 6385 in LFM and LM, respectively, also deviance = 559.7 ( $p < 0.001$ ).

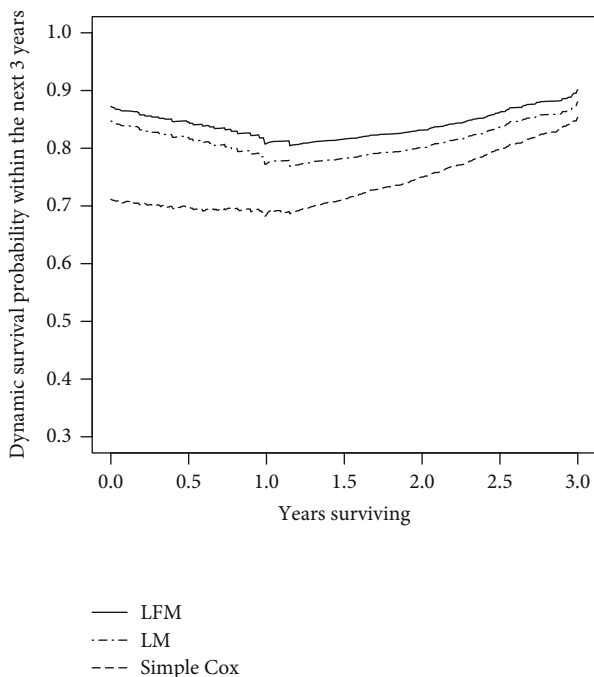


FIGURE 5: Dynamic prediction of survival within window ( $w = 3$ ) by adjusting the covariates in LFM, LM, and simple Cox model.

that it has a better GOF, improved real data analysis, and more optimized cardiovascular risk assessment.

### Data Availability

The data used to support the findings of this study were supplied by [National Heart, Lung, and Blood Institute (NHLBI),] under license and so cannot be made freely available. Requests for access to these data should be made to {National Heart, Lung, and Blood Institute (NHLBI), Funded by the National Institutes of Health; ClinicalTrials.gov number, NCT01206062}. We have a request ID of 4612.

### Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

### Acknowledgments

This work was extracted from the Ph.D. dissertation of Mr. M. Sayadi [35], and it was supported by Shiraz University of Medical Sciences, Shiraz, Iran (Grant no. 97-16792). The authors would like to thank Mr. I. Razeghian-Jahromi and Ms. Z. Hadi for their valuable suggestions. The authors wish to thank Mr. H. Argasi at the Research Consultation Center (RCC) of Shiraz University of Medical Sciences for his invaluable assistance in editing this manuscript.

## References

- [1] K. G. M. Moons, D. G. Altman, Y. Vergouwe, and P. Royston, "Prognosis and prognostic research: application and impact of prognostic models in clinical practice," *BMJ*, vol. 338, no. jun04 2, p. b606, 2009.
- [2] J. M. T. Hendriksen, G. J. Geersing, K. G. M. Moons, and J. A. H. de Groot, "Diagnostic and prognostic prediction models," *Journal of Thrombosis and Haemostasis*, vol. 11, pp. 129–141, 2013.
- [3] M. Sayadi, M. Zibaenezhad, and S. M. Taghi Ayatollahi, "Simple prediction of type 2 diabetes mellitus via decision tree modeling," *International Cardiovascular Research Journal*, vol. 11, no. 2, pp. 71–76, 2017.
- [4] N. R. Cook, "Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve," *Clinical chemistry*, vol. 54, no. 1, pp. 17–23, 2008.
- [5] S. van Dieren, J. W. J. Beulens, A. P. Kengne et al., "Prediction models for the risk of cardiovascular disease in patients with type 2 diabetes: a systematic review," *Heart*, vol. 98, no. 5, pp. 360–369, 2012.
- [6] W. Bouwmeester, J. W. R. Twisk, T. H. Kappen, W. A. van Klei, K. G. M. Moons, and Y. Vergouwe, "Prediction models for clustered data: comparison of a random intercept and standard regression model," *BMC medical research methodology*, vol. 13, no. 1, 2013.
- [7] G. S. Collins, S. Mallett, O. Omar, and L.-M. Yu, "Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting," *BMC medicine*, vol. 9, no. 1, article 447, 2011.
- [8] S. Mallett, P. Royston, S. Dutton, R. Waters, and D. G. Altman, "Reporting methods in studies developing prognostic models in cancer: a review," *BMC medicine*, vol. 8, no. 1, 2010.
- [9] L. Parast, L. Tian, and T. Cai, "Landmark estimation of survival and treatment effect in a randomized clinical trial," *Journal of the American Statistical Association*, vol. 109, no. 505, pp. 384–394, 2014.
- [10] H. van Houwelingen and H. Putter, *Dynamic Prediction in Clinical Survival Analysis*, CRC Press., 2011.
- [11] C. Proust-Lima and J. M. Taylor, "Development and validation of a dynamic prognostic tool for prostate cancer recurrence using repeated measures of posttreatment PSA: a joint modeling approach," *Biostatistics*, vol. 10, no. 3, pp. 535–549, 2009.
- [12] J. Wang, S. Luo, and L. Li, "Dynamic prediction for multiple repeated measures and event time data: an application to Parkinson's disease," *The annals of applied statistics*, vol. 11, no. 3, pp. 1787–1809, 2017.
- [13] G. L. Hickey, S. W. Grant, C. Caiado et al., "Dynamic prediction modeling approaches for cardiac surgery," *Circulation: Cardiovascular Quality and Outcomes*, vol. 6, no. 6, pp. 649–658, 2013.
- [14] E.-R. Andrinopoulou, D. Rizopoulos, J. J. Takkenberg, and E. Lesaffre, "Combined dynamic predictions using joint models of two longitudinal outcomes and competing risk data," *Statistical methods in medical research*, vol. 26, no. 4, pp. 1787–1801, 2017.
- [15] N. Roustaei, S. M. T. Ayatollahi, and N. Zare, "A proposed approach for joint modeling of the longitudinal and time-to-event data in heterogeneous populations: an application to HIV/AIDS's disease," *BioMed research international*, vol. 2018, pp. 1–13, 2018.
- [16] J. Z. Musoro, R. B. Geskus, and A. H. Zwinderman, "A joint model for repeated events of different types and multiple longitudinal outcomes with application to a follow-up study of patients after kidney transplant," *Biometrical Journal*, vol. 57, no. 2, pp. 185–200, 2015.
- [17] K. Suresh, J. M. Taylor, D. E. Spratt, S. Daignault, and A. Tsodikov, "Comparison of joint modeling and landmarking for dynamic prediction under an illness-death model," *Biometrical Journal*, vol. 59, no. 6, pp. 1277–1300, 2017.
- [18] L. Ferrer, H. Putter, and C. Proust-Lima, "Individual dynamic predictions using landmarking and joint modelling: validation of estimators and robustness assessment," *Statistical methods in medical research*, vol. 28, no. 12, pp. 3649–3666, 2018.
- [19] J. R. Anderson, K. C. Cain, and R. D. Gelber, "Analysis of survival by tumor response," *Journal of Clinical Oncology*, vol. 1, no. 11, pp. 710–719, 1983.
- [20] H. Shi and G. Yin, "Landmark cure rate models with time-dependent covariates," *Statistical methods in medical research*, vol. 26, no. 5, pp. 2042–2054, 2017.
- [21] U. Dafni, "Landmark analysis at the 25-year landmark point," *Circulation: Cardiovascular Quality and Outcomes*, vol. 4, no. 3, pp. 363–371, 2011.
- [22] A. Attar, M. Sayadi, and M. Jannati, "Effect of intensive blood pressure lowering on cardiovascular outcomes based on cardiovascular risk: a secondary analysis of the SPRINT trial," *European journal of preventive cardiology*, vol. 26, no. 3, pp. 238–245, 2019.
- [23] A. Attar and M. Sayadi, "Effect of chronic kidney disease on cardiovascular events An Epidemiological Aspect from SPRINT Trial," *Iranian journal of kidney diseases*, vol. 13, no. 5, pp. 328–336, 2019.
- [24] H. C. Van Houwelingen, "Dynamic prediction by landmarking in event history analysis," *Scandinavian Journal of Statistics*, vol. 34, no. 1, pp. 70–85, 2007.
- [25] T. Therneau, "Mixed effects Cox models," *CRAN repository*, 2015.
- [26] T. M. Therneau and P. M. Grambsch, "The Cox model," *Modeling survival data: extending the Cox model*, pp. 39–77, 2000.
- [27] P. C. Austin, "Generating survival times to simulate Cox proportional hazards models with time-varying covariates," *Statistics in Medicine*, vol. 31, no. 29, pp. 3946–3958, 2012.
- [28] E. Graf, C. Schmoor, W. Sauerbrei, and M. Schumacher, "Assessment and comparison of prognostic classification schemes for survival data," *Statistics in Medicine*, vol. 18, no. 17–18, pp. 2529–2545, 1999.
- [29] SPRINT Research Group, J. T. Wright Jr., J. D. Williamson et al., "A randomized trial of intensive versus standard blood-pressure control," *New England Journal of Medicine*, vol. 373, no. 22, pp. 2103–2116, 2015.
- [30] B. Vipler, C. Magee, and K. Douglas, "A randomized trial of intensive versus standard blood-pressure control," *The New England journal of medicine*, vol. 374, no. 23, pp. 2290–2295, 2016.
- [31] J. Musoro, G. Struijk, R. Geskus, I. Ten Berge, and A. Zwinderman, "Dynamic prediction of recurrent events data by landmarking with application to a follow-up study of patients after kidney transplant," *Statistical methods in medical research*, vol. 27, no. 3, pp. 832–845, 2018.
- [32] M. J. Zibaenezhad, S. S. Mohammadi, M. Sayadi, S. Khorshidi, E. Bahramali, and I. Razeghian-Jahromi, "The impact of diabetes mellitus and hypertension on clinical outcomes in a

population of Iranian patients who underwent percutaneous coronary intervention: a retrospective cohort study,” *The Journal of Clinical Hypertension*, vol. 21, no. 11, pp. 1647–1653, 2019.

- [33] S. Teramukai, Y. Okuda, S. Miyazaki, R. Kawamori, M. Shirayama, and T. Teramoto, “Dynamic prediction model and risk assessment chart for cardiovascular disease based on on-treatment blood pressure and baseline risk factors,” *Hypertension Research*, vol. 39, no. 2, pp. 113–118, 2016.
- [34] M. Hamoen, Y. Vergouwe, A. H. Wijga et al., “Dynamic prediction of childhood high blood pressure in a population-based birth cohort: a model development study,” *BMJ Open*, vol. 8, no. 11, article e023912, 2018.
- [35] M. Sayadi, “Optimization of dynamic prediction models in survival clinical data analysis, in Biostatistics,” Shiraz University of Medical Science, Shiraz, Iran, 2018.