Submitted to the Annals of Applied Statistics
arXiv: arXiv:0000.0000

DYNAMIC PREDICTION OF DISEASE PROGRESSION FOR LEUKEMIA PATIENTS BY FUNCTIONAL PRINCIPAL COMPONENT ANALYSIS OF LONGITUDINAL EXPRESSION LEVELS OF AN ONCOGENE

By Fangrong Yan[*,†], Xiao Lin[*] and Xuelin Huang[†]

China Pharmaceutical University[*]
and
The University of Texas MD Anderson Cancer Center[†]

Patients' biomarker data are repeated measured over time during their follow-up visits. Statistical models are needed to predict disease progression on the basis of these longitudinal biomarker data. Such predictions must be conducted on a real-time basis so that at any time a new biomarker measurement is obtained, the prediction can be updated immediately to reflect the patient's latest prognosis and further treatment can be initiated as necessary. This is called dynamic prediction. The challenge is that longitudinal biomarker values fluctuate over time, and their changing patterns vary greatly across patients. In this article, we apply functional principal components analysis (FPCA) to longitudinal biomarker data to extract their features, and use these features as covariates in a Cox proportional hazards model to conduct dynamic predictions. Our flexible approach comprehensively characterizes the trajectory patterns of the longitudinal biomarker data. Simulation studies demonstrate its robust performance for dynamic prediction under various scenarios. The proposed method is applied to dynamically predict the risk of disease progression for patients with chronic myeloid leukemia following their treatments with tyrosine kinase inhibitors, with the FPCA method being applied to their longitudinal measurements of *BCR-ABL* gene expression levels during follow-up visits to obtain the changing patterns over time as predictors.

1. Introduction. Precision medicine has been cast as the future of medical care, which has increased interest in prognostic models for many diseases. Examples of such models available in the literature include prognostic models for various types of cancer, such as liver cancer, prostate cancer, and leukemia. However, the majority of prognostic models in the literature provide risk predictions using only a small portion of the recorded information.

MSC 2010 subject classifications: Primary 60K35, 60K35; secondary 60K35
Keywords and phrases: Dynamic prediction, Functional principal component analysis, Longitudinal biomarker, oint modeling, Survival analysis

Patient outcomes are typically measured repeatedly over time, yet only the last one or two measurements are used in prognostic models. An advantage of such a simple model is that it can be easily applied in everyday clinical practice. However, an important limitation is that valuable information is discarded, which, if appropriately used, could offer a better insight into the dynamics of disease progression. In particular, an inherent characteristic of many medical conditions is their dynamic nature. That is, disease severity and the rate of disease progression not only differ from patient to patient but also dynamically change over time for the same patient. It is critical to capture these changing patterns and use this information to predict patients' prognoses and make medical decisions in a real-time fashion. Well-designed statistical methods and software are needed for such dynamic predictions.

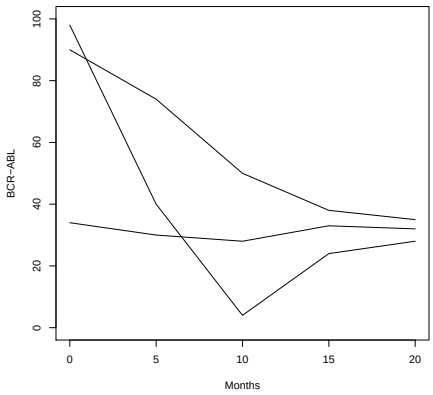

FIG 1. *Three hypothetical examples of changing patterns of BCR-ABL transcript levels over time: always decreasing, flat, decreasing first and then bounce back. These patterns may indicate very different future prognosis, despite similar BCR-ABL levels at the 20th month.*

To construct real-time prediction models for time to next failure event using longitudinal biomarker data, while the current biomarker value is usually an important predictor, quite often the changing pattern of biomarker values over time contains more information and thus has higher predictive power. For example, the transcript level of the gene BCR-ABL is a good indicator of residual disease for chronic myeloid leukemia (CML) patients (Quintas-Cardama et al., 2014). Figure 1 show three patients have similar BCR-ABL transcript levels at 20 months, but their changing patterns before that are quite different. The patient who has always maintain a decreasing

pattern may have the best future outcome (longer time to disease progression). The other patient whose BCR-ABL values decreased initially but had an increase after 10 months may experience disease progression soon (have the worst outcome). The remaining patient has an almost constant BCR-ABL value over time. His/her outcome might be intermediate between the above two scenarios. From these hypothetical examples, we can see that it is important to incorporate biomarker changing patterns into prediction models.

The traditional survival analysis literature has provided many models for estimating the time to an event of interest. However, many of them incorporate baseline covariates only (Zheng, Cai and Feng, 2006; Uno, Cai and Wei, 2007), which means that such models can only be used to predict survival at baseline. Recent work has focused on dynamic prediction of future survival at any time point beyond the baseline (Huang et al., 2016). For this purpose, jointly modeling longitudinal information and survival data has been broadly used. The joint modelling approach usually uses a parametric trajectory model with random effects for longitudinal data, which are used as time-dependent covariates in a Cox proportional hazards model (Wulfsohn and Tsiatis, 1997; Tsiatis and Davidian, 2001; Xu and Zeger, 2001; Song, Davidian and Tsiatis, 2002; Ibrahim, Chen and Sinha, 2004; Huang and Liu, 2007; Liu and Huang, 2009; Rizopoulos and Dimitris, 2011; Rizopoulos and Ghosh, 2011; Rizopoulos and Hatfield, 2013). However, the nature of the longitudinal biomarker trajectory differs in each specific clinical setting. Therefore, it is difficult to identify a satisfactory parametric family to use in modeling longitudinal biomarker data in all situations. Based on this consideration, others have used segmented mixed effect models (Slate and Turnbull, 2000), change point models (Pauler and Finkelstein, 2002) and B-splines (Brown, Ibrahim and DeGruttola, 2005) to characterize longitudinal biomarker trajectories.

Although many non-parametric methods, such as splines and kernel smoothing, have been applied to models for longitudinal biomarkers. They aim to better fit the biomarker trajectories over time and then use the fitted (de-noised) biomarker values to do prediction. In this article, while we can keep these de-noised biomarker values as predictors, we also use functional principal component analysis (FPCA) approach to extract the changing patters (features) of each individual's biomarker trajectory, and then use these features as additional predictors to improve the prediction. The FPCA is employed to characterize the pattern of random trajectories of repeatedly measured biomarkers (Besse and Ramsay, 1986; Rice and Silverman, 1991; Silverman, 1996; James, Hastie and A.,

1999; Yao et al., 2003; Yao, Müller and Wang, 2005; Yao and Lee, 2006; Hall and Wang, 2006; Liu and Yang, 2009; Berkey and Kent, 2009). It attempts to identify the dominant modes of variation in a sample of trajectories around an overall mean trend function. Under this framework, we construct our dynamic prediction models in two steps. We first decompose different patterns of biomarker changes over time, and then use the feature information extracted from this decomposition to make predictions. Simulation studies in Section 4 show that our proposed method has robust performance, reflected by larger area under the curve (AUC) of receiver's operating characteristics (ROC), and smaller residual mean square errors (RMSE), when comparing with some joint models with misspecified biomarker submodels. When comparing with correctly specified joint models, our AUCs and RMSEs are close to them.

The rest of this article is organized as follows. In Section 2, we introduce a chronic myeloid leukemia (CML) dataset that motivated this research. In Section 3, we briefly review FPCA for longitudinal biomarkers measured at irregular time intervals and obtain FPCA scores to characterize the trajectory pattern of data observed during the entire span of patient follow-up time. Then, we provide the dynamic prediction based on the FPCA scores. In Section 4, we introduce the formulations of joint modeling and describe simulations we performed to compare our proposed method with some commonly used joint modeling approaches. We illustrate the application of our technique to the CML dataset in Section 5, and provide concluding remarks in Section 6.

2. A motivating example. This article is motivated by a study of CML that focused on the early detection of disease progression (Quintas-Cardama et al., 2014). Up to 95% of patients diagnosed with CML have been found to have a *BCR-ABL* fusion gene. Fusion genes result from the abnormal joining of DNA from two genes (genes *BCR* and *ABL* in this example) as a result of inversion or translocation. Tyrosine kinase inhibitors (TKIs) have been used since year 2000 to stop the expression of *BCR-ABL* in CML patients. After treatment with TKIs, a large fraction of patients will achieve some level of good response, defined by improved clinical symptoms and reduced *BCR-ABL* expression levels. These patients then take the TKI drugs daily for life (until they become resistant to the drug), and have regular follow-up visits. Residual evidence of CML can be represented by the transcript level of *BCR-ABL*. The best outcome for patients is to achieve major molecular response (MMR), which is defined as the *BCR-ABL* transcript level standardized by the international scale is less than 0.1%. This

value is simply denoted as 0.1, and similarly done thereafter in this article for all the *BCR-ABL* values (i.e., the % sign is removed). Some patients may never achieve MMR, but they are still free of clinical symptoms of CML, and do not need any additional treatments. Their disease may remain under control for many years, with an almost constant low level of *BCR-ABL* expression. However, CML may progress with increased levels of *BCR-ABL*. The common clinical practice has been to wait until patients show symptoms of disease progression to start new treatments. However, for many patients, their *BCR-ABL* transcript levels increase before clinical symptoms of disease progression appear. Thus, it will be helpful to use this biomarker to predict the time to disease progression so that physicians can initiate new treatments early to prevent it.

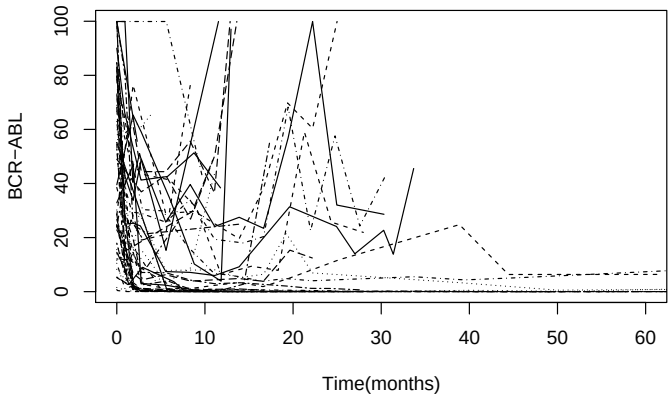

FIG 2. *Plots of BCR-ABL trajectories for 50 randomly selected subjects.*

Imatinib and dasatnib are first- and second-generation TKIs. The study under consideration was a randomized trial that used second-line TKI therapy to treat 670 patients with CML in a dose optimization phase, and to compare different dose schedules of dasatinib in patients with chronic phase CML who had become resistant to imatinib therapy. A 6-year update of this study showed similar efficacy results across the 4 dose schedules tested. In this article, we do not consider the comparison between the 4 dose schedules, but focus on the dynamic prediction of disease progression using longitudinal *BCR-ABL* transcript levels. All patients were followed every 3 months in the first year, every 6 months in the second year, and annually thereafter. The

transcript levels of *BCR-ABL* were measured by polymerase chain reaction during these follow-up visits. For illustration purpose, Figure 2 shows the *BCR-ABL* trajectories for 50 randomly selected patients. These trajectories have different changing patterns over time, which may be important predictors for time to disease progression. Moreover, due to the bumpy shapes of the *BCR-ABL* trajectories, these changing patterns may not be easily characterized by some simple summary statistics, such as changing slopes calculated from the raw data over a specific time interval. This motivates us to use a more systematic approach, namely, functional principal component analysis approach to extract "features" from individual biomarker trajectories, and then use these "features" to make predictions of time to disease progression.

3. Method. In Section 3.1, we decompose the biomarker trajectories into some "feature" functions, which we then use in Section 3.2 to predict the time to disease progression, the event of interest.

3.1. *Functional principal component analysis (FPCA).* FPCA has emerged as a powerful approach for modeling noisy and irregularly measured longitudinal data. Similar to the way in which principal component analysis extracts features from multivariate random vectors, FPCA extracts features from random functional data observed over time. Here, we model the n individuals' biomarker trajectories as independent realizations from a square integrable stochastic process $L^2[0,U]$ on time interval $[0,U]$, where U is the maximum follow-up time.

For subject $i = 1, 2, \cdots, n$, let Y_{ij} be the observed biomarker at random times U_{ij} for $j = 1, \cdots, m_i$, where m_i is the number of observations from the ith subject. Denote by $Z_i(t)$ the biomarker trajectory of subject i that is free of measurement errors. $Z_i(t)$ are often not directly observable, but have to be reconstructed from noisy observations. We write this formula as

$$Y_{ij} = Z_i(U_{ij}) + \varepsilon_{ij}, U_{ij} \in [0,U], \tag{3.1}$$

where ε_{ij} are independent measurement error terms with $E(\varepsilon_{ij}) = 0$ and $Var(\varepsilon_{ij}) = \sigma^2$. Here, Z_i, U_{ij}, and ε_{ij} are mutually independent. The observation time points U_{ij} can be either the same across individuals (regular time intervals) or differently and irregularly spaced for each individual.

Let $Z_i(t), i = 1, \cdots, n$ be n independent realizations of the same square-integrable stochastic process $Z(t)$, which has the mean function $E[Z_i(u)|T_i \geq u] = \mu(u)$ and covariance function $E[\{Z_i(u) - \mu(u)\} \times \{Z_i(v) - \mu(v)\}|T_i \geq u, v] = G(u,v)$, for $u, v \in [0,U]$, where T_i is the survival time for subject i.

Here, $G(u,v)$ is symmetric about u and v, non-negative definite. According to Mercer's theorem (Leng and Müller, 2006), there exists a square integrable orthonormal basis $\{\rho_i(u), 0 \le u \le U, i = 1, \cdots, \infty\}$ (eigenfunctions) and $\{\lambda_i, i = 1, \cdots, \infty\}$ (eigenvalues) such that

$$G(u,v) = \sum_{k=1}^{\infty} \lambda_k \rho_k(u) \rho_k(v), \tag{3.2}$$

where $\rho_k(v)$ is the orthonormal eigenfunction in $L^2[0,U]$ corresponding to the eigenvalue λ_k for $\lambda_1 \ge \lambda_2 \ge \cdots > 0$. This decomposition provides a basic tool to describe the distribution of the random trajectories Z_i. We use the Karhunen-Loeve decomposition (Yao, Müller and Wang, 2005), which represents the mean of a random curve $Z_i(t)$ (biomarker trajectory for subject i), as

$$Z_i(u) = \mu(u) + \sum_{k=1}^{\infty} \gamma_{ik}\, \rho_k(u), i = 1, \cdots, n, \tag{3.3}$$

where $\gamma_{ik} = \int_0^U \{Z_i(t) - \mu(t)\}\rho_k(t)dt$ is the $k-$th FPCA score of random trajectory $Z_i(t)$, $0 \le t \le U$. Since $\rho_k(t)$ and $\rho_j(t)$, $0 \le t \le U$, are orthogonal for $j \ne k$, the random variables γ_{ik}, $1 \le k < \infty$, are not correlated with each other, with $E(\gamma_{ik}) = 0$ and $Var(\gamma_{ik}) = \lambda_k$. A good approximation of the equation (3.3) usually can be achieved by using only the first few components of the above decomposition. That is to say,

$$Z_i(u) \approx \mu(u) + \sum_{k=1}^{K} \gamma_{ik} \rho_k(u), i = 1, \cdots, n. \tag{3.4}$$

The choice of K can be based on the fraction of variance as explained by (Yao, Müller and Wang, 2005), or some information criterion, which we introduce later.

The value of γ_{ik} measures the similarity between $Z_i(t) - \mu(t)$, the deviation of individual curve $Z_i(t)$ from the population mean, and the kth eigenfunction $\rho_k(u)$. The above FPCA framework for functional data is a flexible method for capturing the trajectories of longitudinal biomarker data. It is analogous to the representation of random vectors in multivariate analysis by principal components, by which a random vector can be represented as a linear combination of the orthonormal basis defined by the eigenvectors of its covariance matrix.

We use the principal analysis by conditional estimation (PACE) algorithm (Yao, Müller and Wang, 2005) to estimate the mean function $\mu(u)$,

covariance function $G(u,v)$, eigenfunction $\rho_k(t)$ and FPCA scores γ_{ik} from the entire set of observed data $\{Y_{ij}, i = 1, \cdots, n, j = 1, \cdots, m_i\}$. The PACE method has been shown to be versatile and powerful when applied to sparse and irregularly measured longitudinal data contaminated with measurement errors. Briefly, the PACE method carries out FPCA as follows using the data $\{(U_{ij}, Y_{ij}), i = 1, \cdots, n, j = 1, \cdots, m_i\}$. First, we estimate $\hat{\mu}(u)$ of the mean function $\mu(u)$, which is obtained by a one-dimensional kernel smoother, such as a local linear smoother. Second, the estimate covariance $\hat{G}(u,v)$, given $u, v \in [0, U]$, is obtained by a two-dimensional kernel smoother with all pairwise products $\{Y_{ij} - \hat{\mu}(t_{ij})\}\{Y_{il} - \hat{\mu}(t_{il})\}$ for $j \neq l$ as the response and (t_{ij}, t_{il}) as the predictors. Details about how smoothing parameters were chosen can be found in Yao, Müller and Wang (2005) and Dai et al. (2016). The smoothing techniques in these two steps are conducted over all the subjects who are still at-risk at time u. Third, the estimates of eigenfunctions and eigenvalues correspond to the solution $\hat{\rho}_k$, $\hat{\lambda}_k$ of the equation

$$\int_U \hat{G}(u,v)\hat{\rho}_k(u)du = \hat{\lambda}_k\hat{\rho}_k(v), k \geq 1, \tag{3.5}$$

where the $\hat{\rho}_k$ are subject to $\int_U \hat{\rho}_k^2(u)du = 1$ and $\int_U \hat{\rho}_k(u)\hat{\rho}_l(u)du = 0$ for $l \neq k$.

Based on the above results, we can estimate $\gamma_{ik} = \int (Z_i(t) - \mu(t))\rho_k dt$ by integration. Yao, Müller and Wang (2005) provides an alternative method to avoid numerical integration. Their PACE method takes measurement errors and sparse measurements into account by assuming γ_{ik} and ε_{ij} to be mutually independent and predicting the random effects γ_{ik} based on its conditional expectation: $\tilde{\gamma}_{ik} = E(\gamma_{ik}|Y_i)$. Predictions for γ_{ik} are then obtained by plugging in estimates of the parameters from the entire dataset, borrowing information from all subjects. Specifically, let $Y_i = (Y_{i1}, \cdots, Y_{im_i})', \mu_i = (\mu(t_{i1}), \cdots, \mu(t_{im_i}))'$, and $\rho_{ik} = (\rho_k(t_{i1}), \cdots, \rho_k(t_{im_i}))'$. Write $Z_i = (Z_{i1}, \cdots, Z_{im_i})'$, then let $\Sigma_{Y_i} = \text{cov}(Y_i, Y_i) = \text{cov}(Z_i, Z_i) + \sigma^2 I_{m_i}$. That is to say, the (j,l) entry of the $m_i \times m_i$ matrix Σ_{Y_i} is $(\Sigma_{Y_i})_{j,l} = G(t_{ij}, t_{il}) + \sigma^2\delta_{jl}$ with $\delta_{jl} = 1$ if $j = l$, and 0 otherwise. Note that the diagonal terms $(j = l)$ have an σ^2 term. That is why the above $\hat{G}(u,v)$ in (3.5) was obtained without including those terms $\{Y_{ij} - \hat{\mu}(t_{ij})\}\{Y_{il} - \hat{\mu}(t_{il})\}$ with $j = l$ in the computaion. On the other hand, applying local linear smoother to these terms with $j = l$, an estimator $\hat{V}(t)$ for $G(t,t) + \sigma^2$ is obtained. Consequently, σ^2 can be estimated by

$$\hat{\sigma^2} = \frac{2}{U}\int_{U/4}^{3U/4} \{\hat{V}(t) - \hat{G}(t,t)\}dt \tag{3.6}$$

where the interval $[\frac{U}{4}, \frac{3U}{4}]$ is used to mitigate boundary effects. Staniswalis and Lee (1998) showed in their Theorem 2 that under certain regularity conditions, the estimator for σ^2 is consistent. Following the above process, we can obtain

$$\hat{\gamma}_{ik} = \hat{E}(\gamma_{ik}|Y_i) = \hat{\lambda}_k \hat{\rho}_{ik}' \hat{\Sigma}_{Y_i}^{-1}(Y_i - \hat{\mu}_i). \tag{3.7}$$

To choose K, which is the number of eigenfunctions needed to provide a reasonable approximation for the infinite-dimensional process, we may use the cross-validation score based on the leave-one-out prediction error (Rice and Silverman, 1991). Let $\hat{\mu}^{(-i)}$ and $\hat{\rho}_k^{(-i)}$ be the estimated mean and eigenfunctions after removing the data for the ith subject. Then we choose K so as to minimize the cross-validation score based on the squared prediction error,

$$CV(K) = \sum_{i=1}^{n} \sum_{j=1}^{m_i} \{Y_{ij} - \hat{Y}_i^{(-i)}(t_{ij})\}^2, \tag{3.8}$$

where $\hat{Y}_i^{(-i)}$ is the predicted curve for the ith subject, computed after removing the data for this subject, that is, $\hat{Y}_i^{(-i)}(t) = \hat{\mu}^{(-i)}(t) + \sum_{k=1}^{K} \hat{\gamma}_{ik} \hat{\rho}_k^{(-i)}(t)$, where $\hat{\gamma}_{ik}$ is estimated by (3.7).

Alternatively, we may use an adapted Akaike information criterion (AIC). A pseudo-Gaussian log-likelihood $\hat{L}$ can be defined as the sum of the contributions from all subjects, treating the estimated FPCA scores $\hat{\gamma}_{ik}$ as normally distributed with variance $\hat{\sigma^2}$ and independent across both i and k, as below,

$$\hat{L} = \sum_{i=1}^{n} \left\{ -\frac{m_i}{2} \log(2\pi\hat{\sigma^2}) - \frac{1}{2\hat{\sigma^2}} \sum_{j=1}^{m_i} \left(Y_{ij} - \hat{Z}_i^K(t_{ij}) \right)^2 \right\}. \tag{3.9}$$

Then let $AIC = -\hat{L} + K$. It has been showed that this AIC is computationally more efficient and achieves results that are similar to those obtained by cross-validation (Yao, Müller and Wang, 2005).

3.2. *Survival analysis with longitudinal biomarker data.* The FPCA scores γ_{ik} can be estimated from the observation $\{Y_{i1}, \cdots, Y_{i,m_i}\}$. In this section, we show how to use those FPCA scores in the survival analysis. Assume that the infinite-dimensional covariate trajectories $Z_i(t)$ under consideration are well approximated by the projection onto the function space

spanned by K eigenfunctions. The estimated trajectory $Z_i(t)$ for the ith subject, using the first K eigenfunctions, is given by

$$\hat{Z}_i^K(t) = \hat{\mu}(t) + \sum_{k=1}^{K} \hat{\gamma}_{ik}\hat{\rho}_k(t), t \in [0, U]. \tag{3.10}$$

The number of eigenfunctions, K, can be chosen by cross-validation based on the AIC or leave-one-out prediction error. Given the estimates $\hat{\mu}(t)$ and $\hat{\rho}_k(k = 1, \cdots, K)$, the various FPCA scores $\hat{\gamma}_{ik}$ result in different trajectory patterns. Therefore, the FPCA scores $\hat{\gamma}_{ik}$ can be used as covariates in modeling the relationship between the survival time and the patterns of the trajectories.

Let T_i and C_i denote the event and censoring times, respectively, and assume C_i is independent of biomarker measurements. Rather than observe T_i for all i, we observe only $V_i = min(T_i, C_i)$ and $\Delta_i = I(T_i \leq C_i)$. Let X_i be a $q-$dimensional vector of the baseline covariates and let $Z_i(t)$, $t \geq 0$, be the longitudinal biomarker trajectory for subject i. Our consideration of the baseline covariates for simplicity does not alter the general insights we highlight in the next section. The following approach is commonly used for dynamic prediction. For subject i, assume a Cox proportional hazards model (Cox, 1972) that specifies $h_i(t)$, the hazard function for T_i as

$$h_i(t|X_i, Z_i(t)) = h_0(t) \exp\{\theta' X_i + \alpha Z_i(t)\}, \tag{3.11}$$

where $h_0(t)$ is an arbitrary non-negative function, and θ and α are unknown parameters.

The above approach uses the biomarker values measured at time t only. It may use historical biomarker values in some ad hoc way, such as letting $Z(t)$ be the biomarker value change or changing rate from the previous observation. However, these approaches may not be sufficient to fully capture the longitudinal biomarker information. In many situations, the biomarker trajectory features (changing patterns) are more important than the current biomarker value or recent changing magnitude or slope, in terms of predicting a future event. In the decomposition, equation (3.3), each $\hat{\rho}(u)$, $u \in [0, U]$ may be viewed as a changing pattern, and $\hat{\gamma}_{ik}$ describes how strongly the data from subject i follow this pattern. Our idea is to use $\hat{\gamma}_{ik}$, $k = 1, \cdots, K$, as predictors. With this preparation, we conduct dynamic prediction at any time t using the following model,

$$h_i(t|X_i, \hat{\gamma}_i) = h_0(t) \exp\{\theta' X_i + \beta' \hat{\gamma}_i\}, \tag{3.12}$$

where $\beta = (\beta_1, \cdots, \beta_K)'$ are the regression coefficients for the K estimated FPCA score vector $\hat{\gamma}_i = (\hat{\gamma}_{i1}, \cdots, \hat{\gamma}_{iK})'$, and $\theta = (\hat{\theta}_1, \cdots, \hat{\theta}_q)'$ are the regression coefficients for the baseline covariates.

In the above, for simplicity, we use the Cox(1972) proportional hazards model for prediction. However, in practice, it is important to test whether the proportional hazards assumption holds. A few different approaches have been proposed to test this assumption. For a categorical covariate, we may plot the estimated cumulative hazard functions for different levels of this variable, and check whether they are parallel of each other. For both continuous and categorical covariates, we may add to the model their interactions with a function of time, such as $\log(t)$. If some of these interactions terms turn out to be statistically significant, the proportional hazards assumption is violated. In this case, those significant interaction terms can be added to the model to improve its goodness-of-fit (Cox, 1972). Lin, Wei and Ying (1993) proposed to check the Cox model assumption with cumulative sums of martingale-based residuals. Their method has been implemented in SAS Proc Phreg (SAS Institute Inc., Cary, NC, USA). Grambsch and Therneau (1994) provided a diagnosis test based on weighted residuals, which can be done by the cox.zph function in R (https://www.r-project.org/). Other approaches include Lin, Zhang and Davidian (2006), Grant, Chen and May (2014), among others.

3.3. *Dynamic individualized predictions.* To apply the above dynamic prediction method to an existing dataset, we first obtain estimated FPCA scores $(\hat{\gamma}_{i1}, \cdots, \hat{\gamma}_{iK})$ using the longitudinal biomarker data, and then estimate θ, β, $h_0(t)$ using the baseline covariates, survival information and FPCA scores. Then for a new subject (not in the dataset) with baseline covariate X_{n+1} and biomarker measurements $Y_{n+1} = (Y_{n+1,1}, \cdots, Y_{n+1,m_{n+1}})$ at time points $U_{n+1,1}, \cdots, U_{n+1,m_{n+1}} \leq U$, we use the following formula to compute the FPCA scores for this subject:

$$\begin{aligned} \hat{\gamma}_{n+1,k} &= \hat{E}(\gamma_{n+1,k}|Y_{n+1}) \\ &= \hat{\lambda}_k \hat{\rho}'_{n+1,k} \hat{\Sigma}^{-1}_{Y_{n+1}} (Y_{n+1} - \hat{\mu}_{n+1}),\ k = 1, \cdots, K, \end{aligned} \tag{3.13}$$

where $\hat{\rho}'_{n+1,k}$, $\hat{\Sigma}_{Y_{n+1}}$ and $\hat{\mu}_{n+1}$ are computed similarly as done for $\hat{\rho}'_{i,k}$, $\hat{\Sigma}_{Y_i}$ and $\hat{\mu}_i$, $1 \leq i \leq n$, in the preparation of equation (3.7). Then the prediction for this new subject can proceed as follows. At any time t after the last biomarker observation, i.e., $t \geq U_{n+1,m_{n+1}}$, the predicted future survival

distribution can be written as

$$\begin{aligned}&\Pr(T_{n+1} \geq t+u | T_{n+1} > t, X_{n+1}, \hat{\gamma}_{n+1}) \\ &= \left\{ \frac{\hat{S}_0(t+u)}{\hat{S}_0(t)} \right\}^{\exp\{\hat{\theta}' X_{n+1} + \hat{\beta}' \hat{\gamma}_{n+1}\}},\end{aligned} \tag{3.14}$$

where $S_0(t) = exp\{-\int_0^t h_0(u)du\}$ and $\hat{S}_0(t)$ is its Breslow estimator (Breslow, 1972) resulting from model (3.12). Note this prediction is dynamic, which means it can be updated at any time, as soon as the subject $n+1$ has new biomarker measurements. This is to simply replace the biomarker vector Y_{n+1} by its updated version, and then update equations (3.13) and (3.14) accordingly.

By using FPCA scores, we achieve the following advantages. First, a biomarker trajectory model is not assumed, which is usually difficult to specify, and a misspecified trajectory model may lead to biased predictions. Second, we obtain the FPCA scores using only the observed biomarker values. There is no need for all subjects to have biomarker measurements made at the same post-baseline time point. Third, the proposed method can be used to continuously conduct and update predictive analyses over time.

4. Simulations. We conducted simulation studies under several scenarios to compare our proposed method to joint modeling approaches with parametric models for longitudinal biomarker trajectories, and assessed the advantages and disadvantages of these approaches.

4.1. *Model specification.* Most approaches for jointly modeling time-to-event and longitudinal data are based on the Cox proportional hazards model with time-dependent covariates. Focusing on normally distributed longitudinal outcomes, we use a linear mixed-effects model to generate the subject-specific longitudinal trajectories. Namely, we have

$$\begin{aligned}&Y_i(t) = Z_i(t) + \varepsilon_i(t), \\ &\varepsilon_i(t) \sim N(0, \sigma^2)\end{aligned}$$

where $Z_i(t)$ denotes the true (unobserved) value of the longitudinal biomarker data without error at time t, and $Y_i(t)$ is a measurement of $Z_i(t)$ with error $\varepsilon_i(t)$. Then, we assume the hazard function of subject i is

$$h_i(t|Z_i(t)) = h_0(t) \exp\{\alpha Z_i(t)\}, \tag{4.1}$$

where $h_0(t) = \lambda t^{\lambda-1} \exp(\eta)$, a Weibull baseline hazard function with $\lambda = 2, \eta = -5$, and the association parameter $\alpha = 0.5$. For the longitudinal

process, we consider four different scenarios, including linear and nonlinear mixed-effects models, to capture the variation of biomarkers for an individual subject, as follows.

$$\text{I: Linear Model} \quad Z_i(t) = a + bt + b_{i1} + b_{i2}t,$$
$$\text{II: Exponential Model} \quad Z_i(t) = c\exp(at) + b_{i1} + b_{i2}t,$$
$$\text{III: Quadratic Model} \quad Z_i(t) = a(t-b)^2 + c + b_{i1} + b_{i2}t$$
$$\text{IV: Piecewise Function Model}$$
$$Z_i(t) = \begin{cases} W_i\, exp(-a_i t) + b_{i1} + b_{i2}t, & t \le 2, \\ d_i + b_{i1} + b_{i2}t, & 2 < t \le 5, \\ d_i + c_i(t-5)^2 + b_{i1} + b_{i2}t, & t > 5. \end{cases}$$

In scenario I, a linear longitudinal trajectory is described with $a = 1, b = -2$. In scenario II, an exponential trajectory is represented with $a = 0.1, c = 3$. Similarly, scenario III uses a quadratic model to generate a nonlinear longitudinal trajectory with $a = 0.2, b = 3, c = 1$. Scenario IV defines a piecewise function trajectory, with $W_i \sim N(3, 0.1^2), a_i \sim N(1, 0.1^2), c_i \sim N(0.3, 0.1^2), d_i \sim N(0.3, 0.1^2)$. All trajectories considered above use random effect terms $b_i = (b_{i1}, b_{i2})' \sim N(0, D)$, with $D = \begin{pmatrix} 0.4 & 0.1 \\ 0.1 & 0.2 \end{pmatrix}$. For simplicity, we generate longitudinal data on irregular time points $t = 0$ and $t = j + \epsilon_{ij}$, $j = 1, 2, \cdots, 10$ and $\epsilon_{ij} \sim N(0, 0.1^2)$ independent across all i and j. Considering the existence of measurement error, we also set $\varepsilon_i(t) \sim N(0, 0.6^2)$, independent of each other across i and t. We simulated the censoring times from a uniform distribution in $(0, t_{max})$, with t_{max} set to result in about 25% censoring in each scenario. Using the models and parameter settings above, we generated four scenarios of datasets. For each scenario, we simulated 100 datasets with sample sizes $n = 400$.

The next step after generating the data for the four scenarios is to analyze these datasets with dynamic prediction methods. For each scenario, we considered four dynamic prediction methods: our proposed method, FPCA, which uses a functional form to capture the variation of the longitudinal biomarker data, and three approaches that use a joint modeling framework, JML, JME and JMQ. JML uses a linear model as a fixed-effect term in a longitudinal data submodel; JME uses an exponential formula to model the main trend of the longitudinal data; and JMQ employs a quadratic expression to capture the non-monotonic longitudinal trajectories. The R package JM (Rizopoulos, 2010) was used to fit these joint models.

4.2. *Measures to assess predictive performance.* We use two measures to evaluate the predictive performance of our proposed method. The first

TABLE 1

Simulation results: Mean (standard deviation) of the root mean squared prediction errors (RMSEs) by the four methods (JML, JME, JMQ, FPCA) in each scenario of longitudinal biomarker values as a function of observation time (Linear, Exponential, Quadratic, Piecewise function models).

RMSE	Linear	Exponential	Quadratic	Piecewise
JML	0.09642	0.09831	0.15694	0.14949
	(0.00537)	(0.01112)	(0.00462)	(0.01674)
JME	0.09698	0.09656	0.17599	0.15863
	(0.00537)	(0.01100)	(0.01819)	(0.02641)
JMQ	0.09677	0.09840	0.15617	0.23850
	(0.00526)	(0.01115)	(0.01840)	(0.06086)
FPCA	0.08537	0.07100	0.07218	0.08724
	(0.00785)	(0.00948)	(0.00938)	(0.01975)

is the root mean squared errors (RMSEs) between the predicted survival probabilities and their true values (which are known in simulation studies). The second is the area under the receiver operating characteristic curve (AUC). To compute AUC, we focus on a time interval of medical relevance $(t, t + \Delta t)$. Let $\pi_i(t + \Delta t)$ represents the survival probability for subject i at $t + \Delta t$. It has been proposed that the AUC can be computed as follows (Harrell, Lee and Mark, 1996; Heagerty and Zheng, 2005; Antolini, Boracchi and Biganzoli, 2005). For a randomly chosen pair of subjects $\{i, j\}$ who have both provided measurements up to time t,

$$\text{AUC}(t, \Delta t) = Pr[\pi_i(t + \Delta t|t) < \pi_j(t + \Delta t|t) \Big| \{T_i^* \in (t, t + \Delta t]\} \cap \{T_j^* > t + \Delta t\}], \tag{4.2}$$

That is, the AUC can be computed as a concordance measure between predictions and observed events.

4.3. *Analysis and results.* In each scenario, we used all observations for 400 subjects to fit the four dynamic predictive models (FPCA, JML, JME, and JMQ). The JML, JME and JMQ are the true models for scenarios I, II and III respectively. The FPCA approach used the FPCA scores obtained as in (3.7) as covariates to form the hazard function in equation (3.12). The FPCA scores represent the changing patterns of the longitudinal observations throughout the time intervals.

We used the RMSE to assess calibration. The mean RMSEs and their standard deviations over all subjects from the 100 datasets are shown in Table 1, for predictions conducted at $t = 4$ for survival probability at $t+\Delta t = 6$.

TABLE 2

Simulation results: Mean (standard deviation) of the area under the receiver's operating characteristics curve (AUCs) by the four methods (JML, JME, JMQ, FPCA) in each scenario of longitudinal biomarker values as a function of observation time (Linear, Exponential, Quadratic, Piecewise function models).

AUC	Linear	Exponential	Quadratic	Piecewise
JML	0.6634	0.7052	0.5057	0.5753
	(0.09681)	(0.11995)	(0.05239)	(0.05877)
JME	0.6520	0.7070	0.5677	0.5783
	(0.09673)	(0.11469)	(0.05454)	(0.05306)
JMQ	0.6542	0.6971	0.5712	0.5409
	(0.09874)	(0.12827)	(0.04857)	(0.06424)
FPCA	0.6908	0.8613	0.6182	0.8487
	(0.09931)	(0.07412)	(0.05600)	(0.04022)

We used the AUC described above to similarly assess the discrimination ability of the four approaches. For each simulated dataset and based on each joint model, we estimated AUC$(t = 2, \Delta t = 6)$ using all 400 subjects. The means and standard deviations of AUCs from the 100 simulated datasets are shown in Table 2. From Tables 1 and 2, we observe that for all scenarios, FPCA has robust predictive performance, in terms of both calibration and discrimination, outperforming joint models with parametric biomarker models. The JM package had not implemented the computation for conditional survival probabilities based on Cox proportional hazards models with time-dependent covariates, so piecewise constant hazard functions were assumed. This may have affected the performance of joint models for prediction.

FPCA is somewhat time-consuming compared to the other methods. Using a personal computer (RAM 12G, CPU 3.4GHz) for the above simulation study, it takes about 6 hours using FPCA, while it takes about 1 hour for each of JML, JME and JMQ.

5. Application. We return to the CML dataset described in Section 2. This was a study of 670 patients diagnosed with CML and enrolled in a trial to receive dasatinib. Only 567 of the patients had *BCR-ABL* measurements taken both before and after the dasatinib treatment, and thus were included in our data analysis for the prediction model. Figure 3 shows their progression-free survival distribution estimated by the Kaplan-Meier method (Kaplan and Meier, 1958) without using the information from their longitudinal *BCR-ABL* measurements. We are interested in predicting progression-free survival probabilities for patients each time after they obtain their updated *BCR-ABL* measurements during follow-up visits.

Although patients are supposed to have follow-up visits at 3, 6, 9, 12, 18

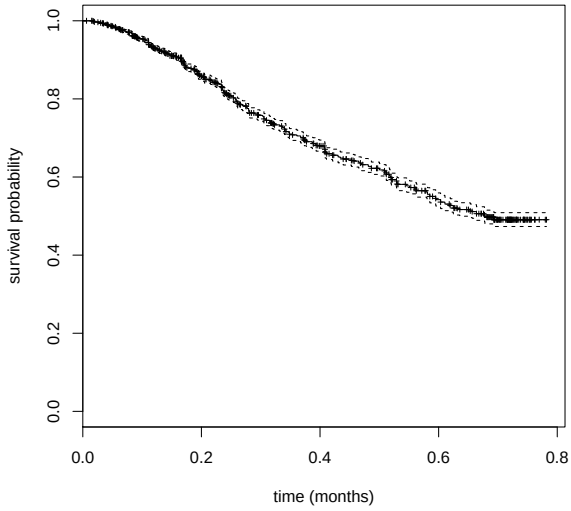

FIG 3. *Kaplna-Meier estimator for time to disease progression.*

and 24 months, and every year thereafter, in reality, their visiting times are irregular due to various constraints, and they may miss some of their scheduled visits. Therefore, patients and physicians would like to have updated predictions at any time immediately after new *BCR-ABL* measurements are available, rather than on just a few discrete time points. In order to handle irregular time intervals and sparse observation data by the FPCA method introduced in Section 3, we decompose the *BCR-ABL* trajectories as

$$Y_{ij} = \mu(U_{ij}) + \sum_{k=1}^{\infty} \gamma_{ik}\rho_k(U_{ij}) + \varepsilon_{ij}, U_{ij} \in [0, U]. \tag{5.1}$$

The PACE method was employed to estimate the mean function $\mu(U_{ij})$, eigenvalue λ_k and eigenfunction $\rho_k(U_{ij})$. With the resulting estimators $\hat{\mu}(U_{ij})$, $\hat{\lambda}_k$, $\hat{\rho}_k(U_{ij})$ available, FPCA scores were obtained by equation (3.7). All calculations mentioned above can be completed in R with the package "fpca". The cubic spline smoothed mean function of *BCR-ABL* is plotted in Figure 4. By adjusting parameter grids in R function fpca.mle, we obtain more smooth eigenfunctions and mean function. The pattern of the mean function demonstrates that most patients experience a sharp decrease at the beginning of treatment (within about 15 months), then a slight decline

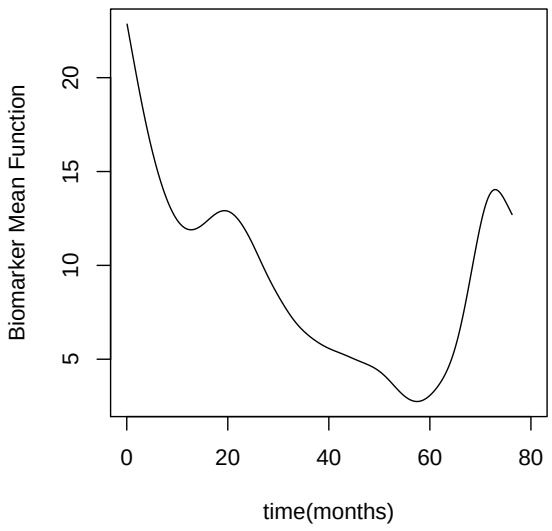

FIG 4. *Mean function of BCR-ABL with functional principle component analysis*

followed. After 60 months, the mean *BCR-ABL* level increases substantially.

By applying the FPCA techniques, we obtain eigenfunctions, which are basis functions to form biomarker trajectories, as defined in equations (3.2-3.4). The number of eigenfunctions was chosen using the cross-validation score. Three eigenfunctions were chosen and shown in the bottom panels of Figure 5. Based on the mean and eigenfunctions, the fitted biomarker trajectories for three randomly selected subjects are shown in the top panels of Figure 5.

The FPCA scores represent the changing pattern of *BCR-ABL* levels over the whole time interval. We used them to model the hazard function as below,

$$h_i(t) = h_0(t) \exp\{\beta_1 \hat{\gamma}_{i1} + \beta_2 \hat{\gamma}_{i2} + \beta_3 \hat{\gamma}_{i3}\}. \tag{5.2}$$

To check the proportional hazards assumption, the interaction terms $\gamma_{i1} \log(t)$, $\gamma_{i2} \log(t)$ and $\gamma_{i3} \log(t)$ are added to the model in (5.2). The parameter estimation of this extended model shows that γ_{i3} and $\gamma_{i3} log(t)$ are not significant, with P-values 0.3984 and 0.3576, respectively, while the remaining four terms are all significant with P-values less than 0.01. This

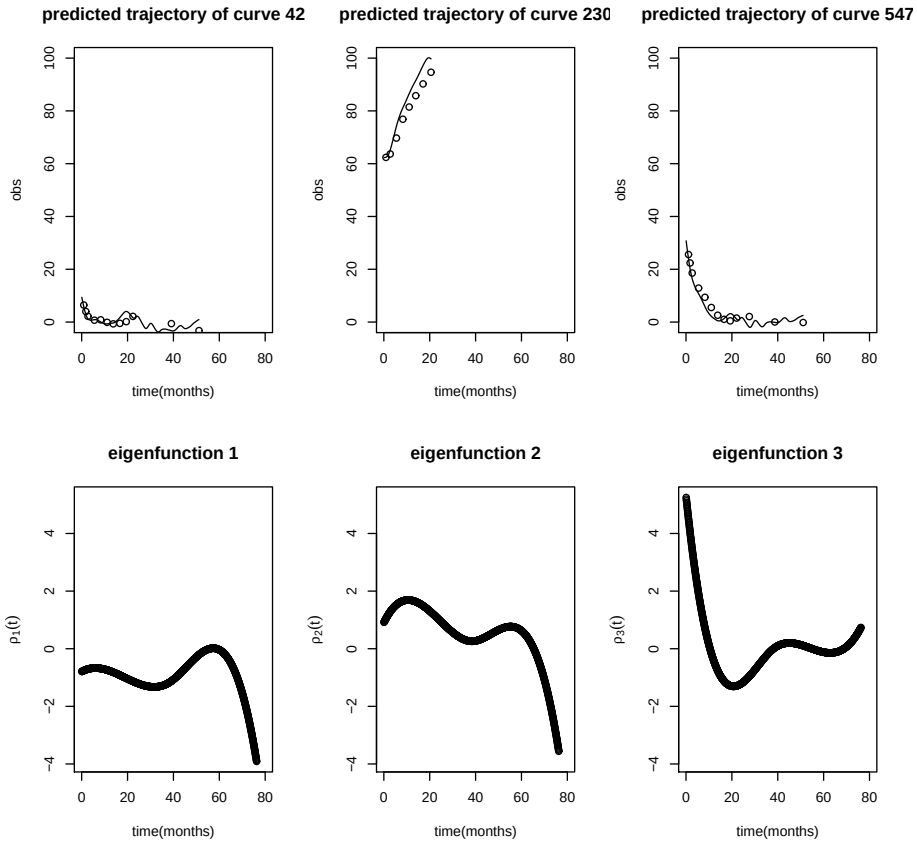

FIG 5. *Predicted trajectories with observed measurements (dots) and three eigenfunctions recovered by this method*

means the proportional hazards assumption is violated. So, we further improved model (5.2) by including two principal components γ_{i1}, γ_{i2} and their interaction terms with $log(t)$, as below,

$$h_i(t) = h_0(t)exp\{\beta_1\hat{\gamma}_{i1} + \beta_2\hat{\gamma}_{i2} + \beta_3\hat{\gamma}_{i1}log(t) + \beta_4\hat{\gamma}_{i2}log(t)\}. \tag{5.3}$$

Based on this model and equation (3.14), dynamic prediction can be performed to provide future survival rate estimation at time time point t. In order to keep consistent with simulation, here we also used FPCA, JML, JME and JMQ to fit joint models. Figure 6 plots the estimated survival curves of a specific patient based on these different prediction methods. It can be seen from this figure that the biomarker trajectory of this patient shows a

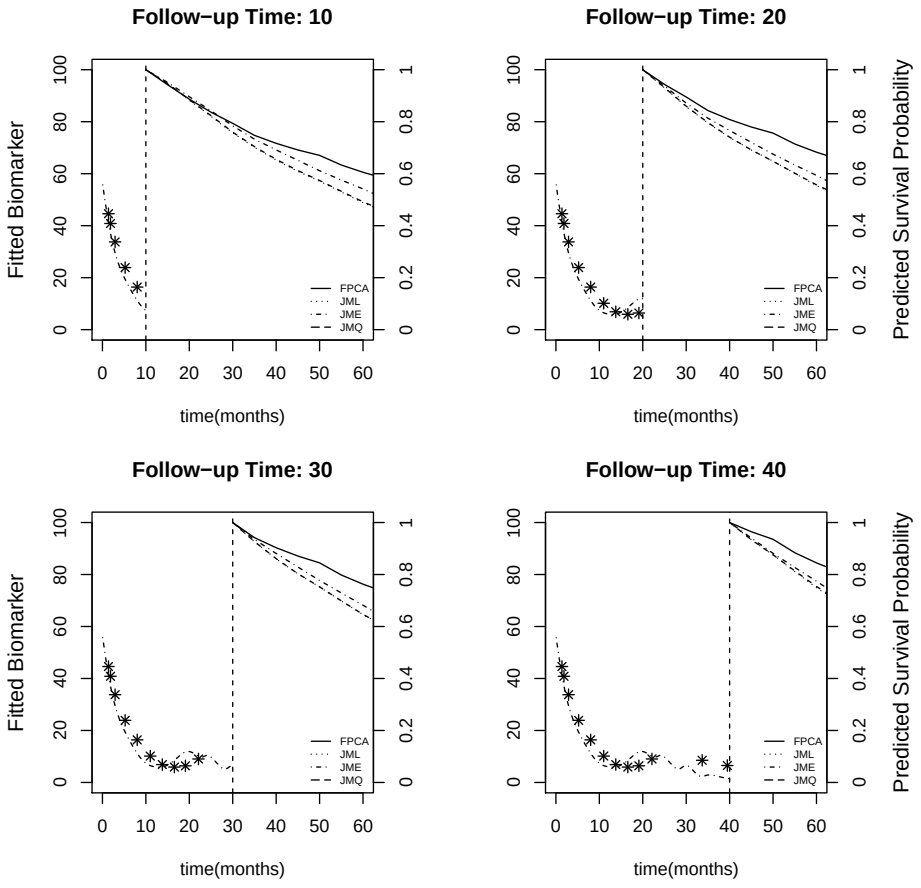

FIG 6. *Dynamic predictions for one patient by four models. Each panel shows the survival probabilities conditional on the longitudinal measurements up to the time point conducting the predictions.*

non-linear decreasing trend, where linear, exponential and quadratic longitudinal submodles may not be suitable to fit biomarker values. This may be the reason why JML, JME and JMQ show similar suboptimal performance. In contrast, FPCA uses a non-parametric method to extract dominant information from longitudinal biomarker, which may give more accurate prediction probabilities than the joint modeling approaches. Certainly these predictions for a particular subject are presented more for illustration, rather than the comparison of the performance between different methods.

The prediction performances of the aforementioned four methods are compared below by their time-dependent AUC curves. That is, at a time point

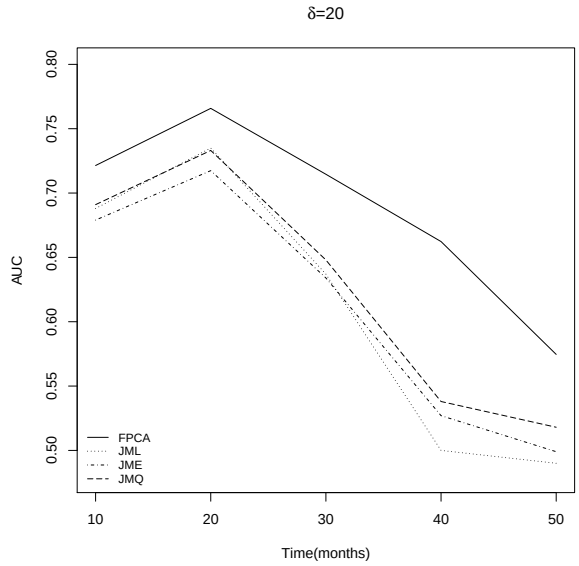

FIG 7. *Time-dependent AUC curves for dynamic predictions conducted at time t (the x-axis), for survival outcome at $t + \Delta t$, with $\Delta t = 20$ months: Comparison of prediction performance by four methods of joint modelling longitudinal and survival data, with longitudinal data modelled by functional principal component analysis (FPCA, solid line), linear model (JML, dot line), quadratic model (JMQ, dash line) or exponential model (JME, dot-dash line).*

t, we conduct prediction of disease-progression by time $t+\Delta t$, with $\Delta t = 20$ months. Then we compute the AUC of this prediction using the method defined in equation (4.2). This process is repeated for series of different time points t, and the results are plotted in Figure 7. We can see that our proposed FPCA method has the highest AUC levels over time, outperforms the joint models using linear (JML), quadratic (JMQ) and exponential (JME) submodels for longitudinal data. This may be explained by the observation that, in this data set, the biomarkers do not follow a linear, quadratic or exponential model. This can be reflected in the mean function plot in Figure 4. On the other hand, the proposed FPCA method is flexible to fit biomarker trajectories of all kinds of different shapes. This is a great advantage of the proposed method since the biomarker trajectories in real world are always much more complicated than the commonly used parametric forms.

For the Cox proportional hazards model in (5.2), we attempted to include some baseline covariates such as patient age, race and sex. However, these

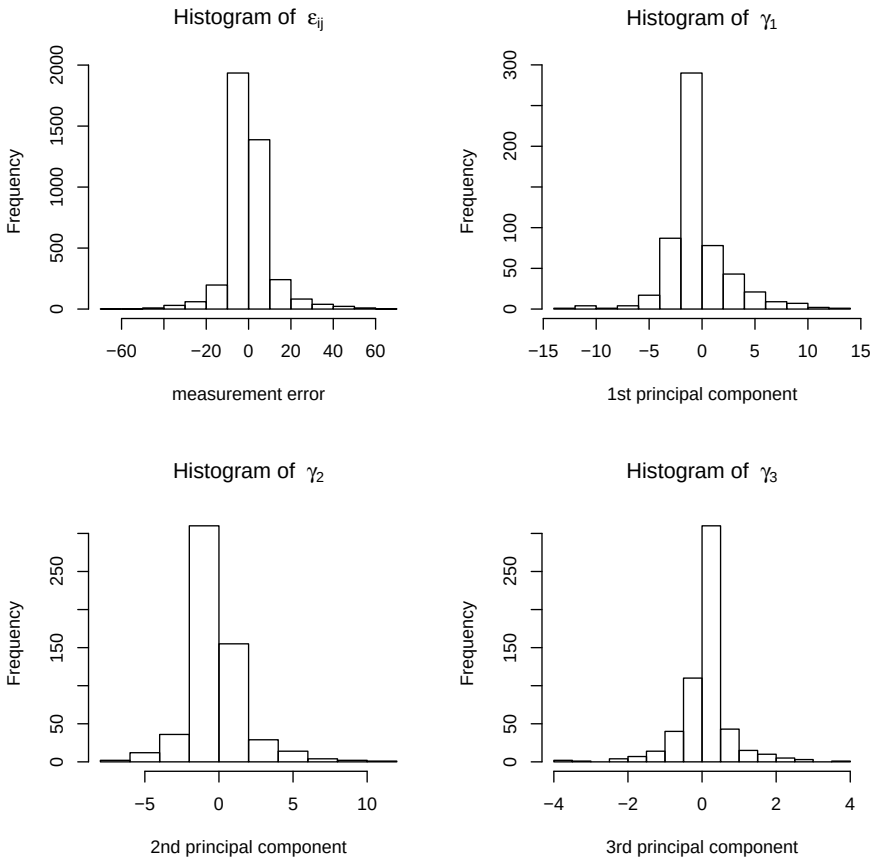

FIG 8. *Histograms to check normality assumptions*

variables do not have much predictive power. Prediction at $t = 40$ months with $\Delta t = 20$ months by a Cox model with only these three variables results a low AUC value of 0.53. Adding the FPCA terms γ_1, γ_2 and γ_3 improves the AUC to 0.62.

Yao, Müller and Wang (2005) imposed normal distribution assumptions for γ_{ik} and ε_{ik} in deriving the FPC scores Histogram plots for γ_{ik} and ε_{ij} show that both normal distribution assumption are satisfied (Figure 8).

In summary, accurate prediction of the risks of disease progression can help physicians and patients make better treatment decisions. For example, they may decide that at any time when the risk of disease progression in the next year (or six months) is greater than 20%, then they should start

a prevention therapy. Our proposed method can also help identify patients who are at elevated risks of disease progression, and remind them to have more frequent follow-up visits.

6. Discussion. The relationship between longitudinal biomarker data and clinical outcomes is important in precision medicine. Conventional survival analysis fits the Cox proportional hazards model, using biomarker data measured at a specific time point as covariates. However, it is often the pattern of covariate values that predicts a patient's survival time. The aim of this article is to capture the trajectory pattern of longitudinal biomarker data within the collection period and use this summary information as predictors to conduct real-time dynamic prediction of the time to a specific event.

The simulation studies show that our proposed method can characterize the trajectory patterns of the covariates and have more robust performance than other joint modeling approaches that use parametric biomarker models. Our proposed method does not need to specify a model for the longitudinal biomarker, and thus avoids the bias caused by the misspecification of such a model. It performs well under various scenarios, demonstrating that it is a versatile and robust approach for dynamic prediction.

A limitation of all principal component analysis approaches is that they are conducted independent of the outcomes, and thus the resulting order of the principal components may not indicate the order of their predictive power. With this consideration, it warrants further research to explore alternative approaches, such as using supervised functional principal component analysis for dynamic prediction.

We have assumed independence between censoring time and biomarker measurement. This assumption is reasonable for most types of administrative censoring, such as when censoring happen at the end of the study. If many subjects drop out of the study, then this assumption may not hold. To handle such a situation, we need to make changes to both the estimation of functional principal components and the survival modelling. It is important to consider such scenarios, which warrants further research.

Acknowledgements. The research of F.Y. was supported in part by the National Science Foundation of China grants No.11171065, No.81130068, Fundamental Research Funds for the Central Universities No. JKPZ2013015, No.JKQZ2013026, and supported by the State Scholarship Fund of China Scholarship Council. The research of X.H. was supported by the US National Science Foundation grant DMS-1612965, National Institutes of Health grants U54 CA096300, U01 CA152958 and 5P50

CA100632.

References.

Antolini, L., Boracchi, P. and Biganzoli, E. (2005). A time-dependent discrimination index for survival data. *Statistics in Medicine* **24** 3927-3944.

Berkey, C. and Kent, R. J. (2009). Longitudinal principal components and non-linear regression models of early childhood growth. *Annals of Human Biology* **10** 523-536(14).

Besse, P. and Ramsay, J. O. (1986). Principal components analysis of sampled functions. *Psychometrika* **51** 285-311.

Breslow, N. E. (1972). Discussion of the paper by D R Cox cited below. *Journal of the Royal Statistical Society, Series B* **34** 187–220.

Brown, E., Ibrahim, J. and DeGruttola, V. (2005). A Flexible B-Spline Model for Multiple Longitudinal Biomarkers and Survival. *Biometrics* **61(1)** 64-73.

Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B* **34** 187–220.

Dai, X., Hadjipantelis, P. Z., Ji, H., Mueller, H. G. and Wang, J. L. (2016). *Functional Data Analysis and Empirical Dynamics*. https://cran.r-project.org/web/packages/fdapace/fdapace.pdf.

Grambsch, P. and Therneau, T. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* **81** 515-26.

Grant, S., Chen, Y. Q. and May, S. (2014). Performance of Goodness-of-Fit Tests for the Cox Proportional Hazards Model with Time-Varying Covariates. *Lifetime Data Analysis* **20(3)** 355-368.

Hall, P. and Wang, J. L. (2006). Properties Of Principal Component Methods For Functional And Longitudinal Data Analysis. *Ann Statist* **34** 1493–1517.

Harrell, F. E., Lee, K. L. and Mark, D. B. (1996). Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine* **15** 361-387.

Heagerty, P. and Zheng, Y. (2005). Survival model predictive accuracy and ROC curves. *Biometrics,* **61** 92-105.

Huang, X. and Liu, L. (2007). A Joint Frailty Model for Survival and Gap Times Between Recurrent Events. *Biometrics* **63** 389-397.

Huang, X., Yan, F., Ning, J., Feng, Z., Choi, S. and Cortes, J. (2016). A Two-Stage Approach for Dynamic Prediction of Time-to-Event Distributions. *Statistics in Medicine* **35** 2167-2182.

Ibrahim, J. G., Chen, M.-H. and Sinha, D. (2004). Bayesian methods for joint modeling of longitudinal and survival data with applications to cancer vaccine trials. *Statistica Sinica* **14** 863-883.

James, G. M., Hastie, T. J. and A., S. C. (1999). Principal Component Models For Sparse Functional Data. *Biometrika* **87** 587-602.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association* **53** 457–481.

Leng, X. and Müller, H. (2006). Classification using functional data analysis for temporal gene expression data. *Bioinformatics* **22** 68-76.

Lin, D. Y., Wei, L. J. and Ying, Z. (1993). Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika* **80** 557-72.

Lin, J., Zhang, D. and Davidian, M. (2006). Smoothing spline-based score tests for proportional hazards models. *Biometrics* **62** 803-812.

Liu, L. and Huang, X. (2009). Joint analysis of correlated repeated measures and recurrent events processes in the presence of death, with application to a study on acquired immune deficiency syndrome. *Journal of the Royal Statistical Society* **58** 65-81.

Liu, X. and Yang, M. (2009). Identifying temporally differentially expressed genes through functional principal components analysis. *Biostatistics* **10** 667-79.

Pauler, D. and Finkelstein, D. (2002). Predicting time to prostate cancer recurrence based on joint models for non-linear longitudinal biomarkers and event time outcomes. *Statistics in Medicine* **21(24)** 3897-3911.

Quintas-Cardama, A., Choi, S., Kantarjian, H., Jabbour, E., Huang, X. and Cortes, J. (2014). Predicting Outcomes in Patients With Chronic Myeloid Leukemia at Any Time During Tyrosine Kinase Inhibitor Therapy. *Clinical Lymphoma Myeloma and Leukemia* **14** 327-334.

Rice, J. A. and Silverman, B. W. (1991). Estimating the mean and covariance structure nonparametrically when the data are curves. *J.royal Statist.soc* **53** 233-243.

Rizopoulos, D. (2010). JM: An R Package for the Joint Modelling of Longitudinal and Time-toEvent Data. *Journal of Statistical Software* **35 (9)** 1-33.

Rizopoulos and Dimitris (2011). Dynamic Predictions and Prospective Accuracy in Joint Models for Longitudinal and Time-to-Event Data. *Biometrics* **67** 819-829.

Rizopoulos, D. and Ghosh, P. (2011). A Bayesian semiparametric multivariate joint model for multiple longitudinal outcomes and a time-to-event. *Statistics in Medicine* **30** 1366C1380.

Rizopoulos, D. and Hatfield, L. A. (2013). Combining Dynamic Predictions from Joint Models for Longitudinal and Time-to-Event Data using Bayesian Model Averaging. *Journal of the American Statistical Association* **109** 1385-1397.

Silverman, B. W. (1996). Smoothed functional principal components analysis by choice of norm. *Annals of Statistics* **24** 1-24.

Slate, E. and Turnbull, B. (2000). Statistical models for longitudinal biomarkers of disease onset. *Statistics in Medicine* **19(4)** 617-637.

Song, X., Davidian, M. and Tsiatis, A. A. (2002). A semiparametric likelihood approach to joint modeling of longitudinal and time-to-event data. *Biometrics* **58** 742-753.

Staniswalis, J. G. and Lee, J. J. (1998). Nonparametric regression analysis of longitudinal data. *Journal of the American Statistical Association* **93** 1403–1418.

Tsiatis, A. A. and Davidian, M. (2001). A semiparametric estimator for the proportional hazards model with longitudinal covariates measured with error. *Biometrika* **88** 447-458.

Uno, H., Cai, T. and Wei, L. T. L. J. (2007). Evaluating Prediction Rules for t-Year Survivors with Censored Regression Models. *Journal of the American Statistical Association* **102** 527-537.

Wulfsohn, M. S. and Tsiatis, A. A. (1997). A joint model for survival and longitudinal data measured with error. *Biometrics* **53** 330-339.

Xu, J. and Zeger, S. L. (2001). Joint analysis of longitudinal data comprising repeated measures and times to events. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **50** 375-387.

Yao, F. and Lee, T. C. M. (2006). Penalized spline models for functional principal component analysis. *Journal of the Royal Statistical Society* **68** 3-25.

Yao, F., Müller, H.-G. and Wang, J. (2005). Functional Data Analysis for Sparse Longitudinal Data. *Journal of the American Statistical Association* **100** 577-590.

Yao, F., Müller, H.-G., Clifford, A. J., Dueker, S. R., Follett, J., Lin, Y., Buchholz, B. A. and Vogel, J. S. (2003). Shrinkage estimation for functional prin-

cipal component scores, with application to the population kinetics of plasma folate. *Biometrics* **59** 676-685.

Zheng, Y., Cai, T. and Feng, Z. (2006). Application of the time-dependent ROC curves for prognostic accuracy with multiple biomarkers. *Biometrics* **62** 279-287.

Research Center of Biostatistics
and Computational Pharmacy
China Pharmaceutical University
Nanjing 210009,P.R. China
E-mail: FYang@mdanderson.org

Department of Biostatistics
The University of Texas
MD Anderson Cancer Center
Houston, Texas 77030, USA
E-mail: xlhuang@mdanderson.org