# Modelling time-varying covariates effect on survival via functional data analysis: application to the MRC BO06 trial in osteosarcoma

Spreafico, M.; Ieva, F.; Fiocco, M.

**ORIGINAL PAPER**

# Modelling time-varying covariates effect on survival via functional data analysis: application to the MRC BO06 trial in osteosarcoma

**Marta Spreafico**[1,2,3] · **Francesca Ieva**[1,3,4] · **Marta Fiocco**[2,5,6]

## Abstract

Time-varying covariates are of great interest in clinical research since they represent dynamic patterns which reflect disease progression. In cancer studies biomarkers values change as functions of time and chemotherapy treatment is modified by delaying a course or reducing the dose intensity, according to patient's toxicity levels. In this work, a Functional covariate Cox Model (FunCM) to study the association between time-varying processes and a time-to-event outcome is proposed. FunCM first exploits functional data analysis techniques to represent time-varying processes in terms of functional data. Then, information related to the evolution of the functions over time is incorporated into functional regression models for survival data through functional principal component analysis. FunCM is compared to a standard time-varying covariate Cox model, commonly used despite its limiting assumptions that covariate values are constant in time and measured without errors. Data from MRC BO06/EORTC 80931 randomised controlled trial for treatment of osteosarcoma are analysed. Time-varying covariates related to alkaline phosphatase levels, white blood cell counts and chemotherapy dose during treatment are investigated. The proposed method allows to detect differences between patients with different biomarkers and treatment evolutions, and to include this information in the survival model. These aspects are seldom addressed in the literature and could provide new insights into the clinical research.

**Keywords** Functional data analysis · Time-varying covariates · Survival analysis · Osteosarcoma

Extended author information available on the last page of the article

# 1 Introduction

Osteosarcoma is a malignant bone tumour mainly affecting children and young adults. Although osteosarcoma is the most common primary malignant bone cancer, it is a rare disease and has an annual incidence of 3-4 patients per million (Smeland et al. 2019). Multidisciplinary management including neoadjuvant and adjuvant chemotherapy with aggressive surgical resection (Ritter and Bielack 2010) or intensified chemotherapy has improved clinical outcomes although the overall 5-year survival rate has remained unchanged in the last 40 years at 60–70% (Anninga et al. 2011). Therefore, it is extremely important to provide an effective tool to evaluate the prognosis for osteosarcoma and to guide the diagnosis.

Time-varying (or time-dependent) covariates are often of interest in clinical and epidemiological research: patients are followed during the study and subject-specific measurements are recorded at each visit. Well-known examples include biomarkers which change during follow-up or cumulative exposure to medications (Austin et al. 2020), such as chemotherapy. Depending on patients' treatment history or development of toxicity, biomarkers values change and chemotherapy treatment is modified by delaying a course or reducing the dose intensity. To study the association between time-varying responses with time-to-event outcome (e.g., death) is a challenging task which could offer new insights into the direction of personalised treatment.

In osteosarcoma treatment, patients usually undergo assessment of hematologic and serum biochemical parameters (Lewis et al. 2007), such as white blood cell (WBC) counts and alkaline phosphatase (ALP). The role of ALP as tumour marker for osteosarcoma has not been established, although several studies suggested that high ALP level is associated with poor overall or event-free survival and presence of metastasis (Ren et al. 2015; Hao et al. 2017). Chemotherapy is usually modelled by different allocated regimens, i.e., by Intention-To-Treat (ITT) analysis (Gupta 2011). ITT ignores anything that happens after randomization, such as protocol deviations or changes in drug intake over time, i.e., delays or dose reduction (Lancia et al. 2019). Lancia et al. (2019b) showed that there is mismatch between target and achieved dose of chemotherapy and the impact of dosis on patients' survival is still unclear. A novel method to study received chemotherapy dose and biomarkers as time-varying variables is proposed. This approach has never been applied to osteosarcoma treatment and provides new insight in understanding the effect of chemotherapy dosis intensity on sarcoma in childhood cancer. Moreover, as will be clear in the following, the application is inspiring from a statistical modelling perspective.

Models for time-to-event data which are able to deal with the dynamic nature of time-varying responses during follow-up are not well developed. One approach for using time-varying covariate data is the Time-Varying covariate Cox Model (TVCM) (Therneau and Grambsch 2000; Kalbfleisch and Prentice 2002), that is an extension of the Cox proportional hazard model (Cox 1972) accounting for covariates that can change value during follow-up. Since time-dependent observations are only available at the time of measurements, TVCM uses the last-observation-carried-forward

(LOCF) approach (Tsiatis and Davidian 2004), which leads to the pitfall of introducing bias due to the continuous nature of the process underlying the data, and fails to account for possible measurement errors (Arisido et al. 2019). Joint models address these issues by modelling simultaneously longitudinal and time-to-event data using shared random effects (Henderson et al. 2000; Tsiatis and Davidian 2004; Chi and Ibrahim 2006; Dantan et al. 2011; Rizopoulos 2012, 2016; Gould et al. 2015; Proust-Lima et al. 2016; Hickey et al. 2016, 2018). They are parametric models that allow for the inference on the association between the hazards characterizing the event outcome and the longitudinal processes. However, they require additional strong assumptions over TVCM that need to be carefully validated to avoid biased estimates (Arisido et al. 2019). Their benefits are hence strictly linked to the correct specification of longitudinal trajectories and baseline hazard function. In addition, inference computations could become prohibitive, especially for approaches developed in a Bayesian framework.

During the past two decades, Functional Data Analysis (FDA) has been increasingly used to analyse, model and predict dynamic processes (Ramsay and Silverman 2002, 2005; Müller 2005; Yao et al. 2005; Ferraty and Vieu 2006; Liu and Yang 2009; Ullah and Finch 2013; Ieva et al. 2013; Ieva and Paganoni 2016; Martino et al. 2019; Spreafico and Ieva 2021). The idea behind FDA and functional models is to express discrete observations arising from time series, i.e., longitudinal time-varying observations, in the form of functions (Ramsay and Silverman 2002, 2005). Functional representation incorporates trends and variations in the evolution of the process over time (Ullah and Finch 2013). Since functional data are infinite-dimensional covariates, some dimensionality reduction methods are needed to summarize and select a finite dimensional set of elements representing the most important features of each covariate. This information can then be included into time-to-event models. To model the relationship between survival outcomes and a set of finite and infinite dimensional predictors Functional Linear Cox Regression Models (FLCRM) have been recently proposed (Gellar et al. 2015; Lee et al. 2015; Qu et al. 2016; Kong et al. 2018; Li and Luo 2019). In case of an infinite dimensional process, Kong et al. (2018) characterized the joint effects of both functional and scalar predictors on time-to-event outcome employing Functional Principal Component Analysis (FPCA). FPCA is one of the most popular dimensionality reduction method in FDA and it is used to summarise each function to a finite set of covariates through FPC scores, while losing a minimum part of the information. An extended version of the FLCRM by Kong et al. (2018) to the case of multiple functional predictors—named Multivariate FLCRM (MFLCRM)—was introduced by Spreafico and Ieva (2021) to model recurrent events effect on long-term survival. However, since the main focus of the work was to develop a methodology for effectively modelling time-varying recurrent events in terms of the functional compensators underlying the processes of interest, the authors have neither compared MFLCRM with other survival models, nor considered its predictive performances over time. In case of multiple longitudinal processes, Li and Luo (2019) exploited the multivariate FPCA approach by Happ and Greven (2018) to extract the FPC scores from the multiple longitudinal trajectories in order to make personalized dynamic predictions. However, the authors did not focus on the smoothing and functional representation

aspects of the processes realized by the observed longitudinal data, on the clinical interpretation of the FPC scores and on their association with overall survival. Since it is often the changing patterns of the functional trajectories rather than the actual values that affects patients' survival, FDA provides a novel modelling and prediction approach, with a great potential for many applications in public health and biomedicine (Ullah and Finch 2013).

Motivated by a clinical question concerning the effect of biomarkers and dose variations during treatment on survival for osteosarcoma patients, an innovative FDA approach, named Functional covariate Cox Model (FunCM), is proposed and compared to a standard TVCM. In FunCM, FDA techniques are first exploited to represent time-varying processes and their derivatives over time in terms of functional data. Unlike joint models, FDA approach does not make assumptions on the distributions of longitudinal processes being computationally advantageous (Li and Luo 2019). Then, additional information contained into the evolution of the functions over time are included into MFLCRMs for overall survival through FPCA. A cross-validation method is implemented to compare MFLCRMs and standard TVCM in terms of their predictive performances at different time horizons. Three novelties of this work are listed here: (i) application of advanced statistical techniques to deal with time-varying covariates in the field of osteosarcoma treatment; (ii) reconstruction of the functional representations for biomarkers and chemotherapy dose values, and their rates of change, to retrieve information on the progression of processes over time; (iii) comparison between TVCM and FunCM in terms of both clinical interpretability and time-dependent predictive performances. This novel approach provides more information about the effect of individualized treatment adaption on survival for osteosarcoma patients.

The rest of this article is organized as follows. In Sect. 2 TVCM and FunCM to represent time-varying covariates by means of FDA and to include them into survival models are discussed. MRC BO06/EORTC 80931 Randomized Controlled Trial and longitudinal representations of time-varying covariates are described in Sect. 3. Results are presented in Sect. 4. Section 5 ends with a discussion of strengths and limitations of the current approach, identifying some developments for future research.

## 2 Statistical methods

### 2.1 Time-varying covariates and survival frameworks

A time-varying (or time-dependent) process is a covariate whose value can change over the duration of follow-up (e.g., time-varying biomarkers, current use of medication, and cumulative dose of drugs). In this study, the main interest is in analysing the association between patient's survival and variations during treatment of his/her multiple time-varying characteristics. The focus is hence on patients who had completed the entire chemotherapy treatment protocol in a pre-defined and clinically acceptable timing period.

Follow-up starts from date of randomization $T_0$ and is divided into a pre-defined 6-months *chemotherapy treatment period* $[T_0; T_0^*]$—also called *observation period*—considered for chemotherapy treatment completion, and a *post-treatment follow-up period* from $T_0^*$ onwards (see Fig. 1).

Under the TVCM framework, the Overall Survival (OS) is measured from randomization ($T_0$) to the date of death or last follow-up date, and the time-varying covariates can be defined over the entire follow-up period. Let $\mathcal{M}$ be a set of time-varying processes. Let $\mathbf{z}_i^{(m)} = \left\{ z_{il}^{(m)} = z_i^{(m)}(t_{il}), l = 1, \ldots, n_i^{(m)} \right\}$ be the vector of longitudinal values related time-varying process $m \in \mathcal{M}$ for each patient $i$, where $t_{il}$ is the time of the $l$-th measurement, $z_i^{(m)}(t_{il})$ is the value of the process at time $t_{il}$ and $n_i^{(m)}$ is the number of different measurements.

Under the FunCM framework, the *observation period* $[T_0; T_0^*]$ is used to reconstruct the functional representations of time-varying covariates. OS is then measured from the end of the *observation period* ($T_0^*$) to the date of death or last follow-up date. Only patients still alive at $T_0^*$ are included in the study cohort. To reconstruct the functional covariates, only measurements registered during the *observation period* (i.e., up to $T_0^*$) are considered, namely vector $\check{\mathbf{z}}_i^{(m)} = \left\{ z_{il}^{(m)} = z_i^{(m)}(t_{il}), l = 1, \ldots, v_i^{(m)} \right\} \subseteq \mathbf{z}_i^{(m)}$, where $v_i^{(m)}$ denotes the index of last measurement of type $m$ for patient $i$ in $[T_0; T_0^*]$, with $v_i^{(m)} \leq n_i^{(m)}$ and $t_{iv_i^{(m)}} \leq T_0^* < t_{iv_i^{(m)}+1}$.

In both cases, the observed time-to-death outcome for patient $i \in \{1, \ldots, N\}$ can be denoted as $(T_i, \delta_i^*)$, where $T_i = \min(T_i^*, C_i)$ is the observed event time (measured from $T_0$ or $T_0^*$ according to the framework), $T_i^*$ is the true event time, $C_i$ is the censoring time and $\delta_i^* = I(T_i^* \leq C_i)$ is the event indicator, with $I(\cdot)$ being the indicator function that takes the value 1 when $T_i^* \leq C_i$, and 0 otherwise.
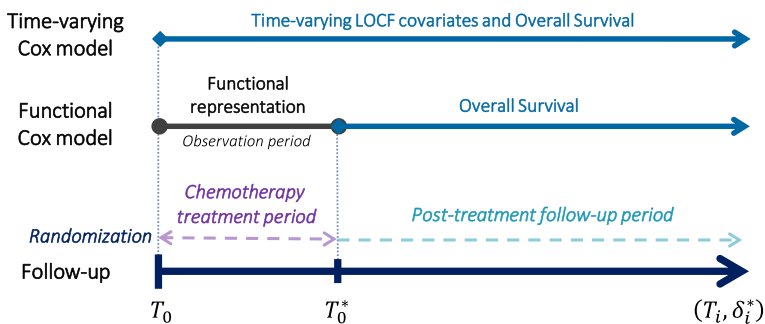


**Fig. 1** Follow-up periods. Time-varying (LOCF/functional) representation and Overall Survival (OS) for Time-Varying covariate Cox Model (TVCM) and Functional covariate Cox Model (FunCM). $T_0$ is the time of randomization. $T_0^* = T_0 + 180$ days is the end of the 6-months chemotherapy treatment period. LOCF = last-observation-carried-forward

## 2.2 Time-varying covariate Cox model

Starting from vector of longitudinal values $z_i^{(m)}$, a time-varying covariate $z_i^{(m)}(t)$ can be defined over the entire follow-up period, according to the LOCF approach (Tsiatis and Davidian 2004):

- when $z_i^{(m)}(t)$ is not observed at time $t \in \left[T_0; t_{in_i^{(m)}}\right]$, the most updated value is used: $z_{il}^{(m)} = z_i^{(m)}(t_{il})$ with $t_{il} \leq t < t_{il+1}$;

- from $t_{in_i^{(m)}}$ onwards, the last available measurement $z_i^{(m)}(t_{in_i^{(m)}})$ is considered.

The TVCM is an extension of the proportional hazard model by Cox (1972) accounting for covariates that can change value during follow-up (Therneau and Grambsch 2000; Kalbfleisch and Prentice 2002). Under TVCM, the proportional hazards model for patient $i$ has the form

$$h_i(t|\boldsymbol{\omega}_i, \boldsymbol{z}_i(t)) = h_0(t) \exp\left\{\boldsymbol{\gamma}^T \boldsymbol{\omega}_i + \boldsymbol{\alpha}^T \boldsymbol{z}_i(t)\right\} \tag{1}$$

where $h_0(t)$ is the baseline hazard function, $\boldsymbol{\omega}_i$ and $\boldsymbol{z}_i(t) = \left(z_i^{(1)}(t), \ldots, z_i^{(M)}(t)\right)$ are the vectors of baseline and time-varying covariates with regression parameters $\boldsymbol{\gamma}$ and $\boldsymbol{\alpha}$, respectively. Inference for coefficients $\boldsymbol{\theta} = (\boldsymbol{\gamma}, \boldsymbol{\alpha})$ is based on maximizing the partial likelihood (Kalbfleisch and Prentice 2002).

TVCM can also be stratified to allow for control by "stratification" of a predictor that does not satisfy the proportional hazard assumption (Kalbfleisch and Prentice 2002). Under stratified TVCM, the hazard function $h_{ig}(t|\boldsymbol{\omega}_i, \boldsymbol{z}_i(t))$ contains also a subscript $g$ that indicates the $g$-th stratum, as well as the baseline hazards $h_{0g}(t)$, where the strata are different categories of the stratification variable. Notice that the baseline hazard functions are different in each stratum.

## 2.3 Functional covariate Cox model

FunCM approach to represent time-varying covariates by means of FDA and to include them into survival models is now introduced. A summary of the proposed methodology is provided in Appendix A.1.

### 2.3.1 From longitudinal to functional representation

To model the continuous longitudinal vectors $\bar{z}_i^{(m)}$ defined over $[T_0; T_0^*]$ as functions $\tilde{x}_i^{(m)}(t)$, FDA techniques can be exploited, as discussed by Ramsay and Silverman (2002, 2005). The observed data $z_{il}^{(m)}$ are assumed as noisy measurements of the latent processes $\tilde{X}_i^{(m)}(t)$, where time $t \in [T_0; T_0^*]$ and $i$ is the patient's index.

For each process $m$, first the time-scale $t \in S_m \subseteq [T_0; T_0^*]$ is chosen. There are no restrictions on the choice of unit of measurement for $t$, though the specific choice can

simplify the computational process. According to the type of observed data (i.e., periodic or open-ended data) and the number of measurements $v_i^{(m)}$, the basis function system $\boldsymbol{\phi}_i^{(m)}(t)$ (e.g., polynomials, B-spline, Fourier, wavelets) is selected, with a number of basis less or equal to $v_i^{(m)}$. Functional data objects are usually expressed by a general functional form as linear combination of the basis functions $W_i^{(m)}(t) = \boldsymbol{\phi}_i^{(m)}(t)^T \boldsymbol{c}_i^{(m)}$, where $\boldsymbol{c}_i^{(m)}$ is the vector of coefficients for patient $i$. Other functional forms can be used to take into account the nature of the process itself (e.g., positive, increasing, decreasing). For example, for an increasing process, the functional data object can be defined using the monotone functional form $W_i^{(m)}(t) = \beta_{0i} + \beta_{1i} \int_{t_0}^{t} \exp[\boldsymbol{\phi}_i^{(m)}(u)^T \boldsymbol{c}_i^{(m)}] du$ (Ramsay and Silverman 2005). Once selected the type of basis functions and the functional form, data can be smoothed by regression analysis minimizing the (penalized) sum of squared errors, obtaining functions $\tilde{x}_i^{(m)}(t) = \hat{W}_i^{(m)}(t)$.

In the presence of constrain due to the specific application, data can be alternatively smoothed by regression analysis using the transformation $g(x) = \log \frac{x - L_m}{U_m - x}$, where $L_m$ and $U_m$ denote the lower and upper bounds respectively. For each patient $i$ the customized functional predictor $m$ is defined as:

$$\tilde{x}_i^{(m)}(t) = \frac{L_m + U_m \cdot \exp\left[\hat{W}_i^{(m)}(t)\right]}{1 + \exp\left[\hat{W}_i^{(m)}(t)\right]}. \tag{2}$$

Starting from the customized functional datum, the FDA approach also allows to reconstruct its derivative $d\tilde{x}_i^{(m)}(t)$ as function of the derivatives of the basis functions $d\boldsymbol{\phi}_i^{(m)}(t)$. The derivative of the functional process, indicated as $\tilde{x}_i^{(dm)}(t)$, represents the rate of change of process values over time. Both functional data $\tilde{x}_i^{(m)}(t)$ and derivatives $\tilde{x}_i^{(dm)}(t)$ can be incorporated as functional predictors into a functional Cox regression model for overall survival by taking into account that they are correlated.

### 2.3.2 Multivariate functional linear Cox regression model

MFLCRM extends the functional Cox regression model by Kong et al. (2018) to the case of multiple functional predictors (Spreafico and Ieva 2021). Let $\tilde{X}_i^{(1)}, \ldots, \tilde{X}_i^{(M)}$ be a set of $M$ functional predictors for individual $i$. MFLCRM includes the multiple functional predictors in the classical Cox model (Cox 1972) as:

$$h_i\left(t | \boldsymbol{\omega}_i, \tilde{x}_i^{(1)}, \ldots, \tilde{x}_i^{(M)}\right) = h_0(t) \exp\left\{\boldsymbol{\gamma}^T \boldsymbol{\omega}_i + \sum_{m=1}^{M} \int_{S_m} \tilde{x}_i^{(m)}(s) \alpha^{(m)}(s) ds\right\} \tag{3}$$

where $h_0(t)$ is the baseline hazard function, $\boldsymbol{\omega}_i$ is the vector of scalar (non functional) covariates with regression parameters $\boldsymbol{\gamma}$. The vector $\left(\tilde{x}_i^{(1)}, \ldots, \tilde{x}_i^{(M)}\right)$ is a realization of the $M$-variate functional data for individual $i$; $\alpha^{(m)}(s)$ are the functional regression

parameters respectively. Sets $S_m \subseteq [T_0; T_0^*]$ are compact sets in $\mathbb{R}$ and can be different (both in period length and time scale) among between different types $m$ of functional predictors.

By applying FPCA, each functional trajectory $\tilde{x}_i^{(m)}(s)$ can be approximated with a finite sum of $K_m$ orthonormal basis $\left\{ \xi_1^{(m)}, \ldots, \xi_{K_m}^{(m)} \right\}$:

$$\tilde{x}_i^{(m)}(s) \approx \mu^{(m)}(s) + \sum_{k=1}^{K_m} f_{ik}^{(m)} \xi_k^{(m)}(s) \tag{4}$$

where $\mu^{(m)}(s)$ is the functional mean and $f_{ik}^{(m)}$ is the FPC score of individual $i$ related to the $k$-th orthonormal base $\xi_k^{(m)}$. To select the truncation parameters $K_m$, representing the number of FPCs to be considered, Spreafico and Ieva (2021) chose the model with the highest Concordance index (Pencina and D'Agostino 2004), that is an overall measure of discrimination in survival analysis. In this work, the truncation parameters $K_m$ are selected in terms of predictive discrimination and calibration performances at different time horizons through the cross-validation procedure introduced in Sect. 2.3.3. From (4) the integrals in (3) can be approximated:

$$\int_{S_m} \left[ \tilde{x}_i^{(m)}(s) - \mu^{(m)}(s) \right] \alpha^{(m)}(s) ds \approx \int_{S_m} \sum_{k=1}^{K_m} f_{ik}^{(m)} \xi_k^{(m)}(s) \alpha^{(m)}(s) ds$$

$$= \sum_{k=1}^{K_m} f_{ik}^{(m)} \int_{S_m} \xi_k^{(m)}(s) \alpha^{(m)}(s) ds \tag{5}$$

$$= \sum_{k=1}^{K_m} f_{ik}^{(m)} \alpha_k^{(m)}$$

where $\alpha_k^{(m)}$ is the scalar representing the quantity $\int_{S_m} \xi_k^{(m)}(s) \alpha^{(m)}(s) ds$. Introducing approximation (5) in Eq. (3), the hazard function becomes:

$$h_i\left(t|\boldsymbol{\omega}_i, \tilde{x}_i^{(1)}, \ldots, \tilde{x}_i^{(M)}\right) = h_0(t) \exp\left\{ \boldsymbol{\gamma}^T \boldsymbol{\omega}_i + \sum_{m=1}^{M} \left[ \int_{S_m} \mu^{(m)}(s) \alpha^{(m)}(s) ds + \sum_{k=1}^{K_m} f_{ik}^{(m)} \alpha_k^{(m)} \right] \right\}$$

$$= h_0^*(t) \exp\left\{ \boldsymbol{\gamma}^T \boldsymbol{\omega}_i + \sum_{m=1}^{M} \sum_{k=1}^{K_m} f_{ik}^{(m)} \alpha_k^{(m)} \right\} \tag{6}$$

where $h_0^*(t) = h_0(t) \exp\left\{ \sum_{m=1}^{M} \int_{S_m} \mu^{(m)}(s) \alpha^{(m)}(s) ds \right\}$ is the baseline hazard function and $\alpha_k^{(m)} = \int_{S_m} \xi_k^{(m)}(s) \alpha^{(m)}(s) ds$ is the regression parameter related to the $k$-th FPC score related to process $m$. Therefore, defining the following quantities:

$$\boldsymbol{\theta} = \left[\boldsymbol{\gamma}^T, \left(\alpha_1^{(1)}, \ldots, \alpha_{K_1}^{(1)}\right), \ldots, \left(\alpha_1^{(M)}, \ldots, \alpha_{K_M}^{(M)}\right)\right]^T$$

$$\mathbf{w}_i = \left[\boldsymbol{\omega}_i^T, \left(f_{i,1}^{(1)}, \ldots, f_{i,K_1}^{(1)}\right), \ldots, \left(f_{i,1}^{(M)}, \ldots, f_{i,K_M}^{(M)}\right)\right]^T$$

and substituting them in Eq. (6), through FPCA the MFLCRM can be expressed as Cox model with hazard function

$$h_i(t) = h_0(t) \exp\left\{\boldsymbol{\theta}^T \mathbf{w}_i\right\}$$

where the vector of coefficients $\boldsymbol{\theta}$ can be estimated by maximising the partial likelihood function (Cox 1972). Notice that even in this case the hazard function in Eq. (6) could be stratified.

### 2.3.3 Selection of truncations parameters

The truncation parameters $K_m$ in Eq. (6) can be chosen in different ways: (i) the Proportion of Variance Explained (PVE) (Ramsay and Silverman 2005), (ii) Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC) or (iii) data-adaptive methods, such as cross-validation (Yao et al. 2005). In this analysis, a combination of these three methods is used. Let the sets of baseline and functional predictors be fixed. First, different combinations of increasing values of the truncation parameters $K_m$ for different time-varying processes $m$ are considered and the best models according to both AIC and BIC criteria are selected. Then, models according to five different thresholds for PVE ($K_m$ such that $PVE \geq 80, 85, 90, 95, 99\%$) are identified. Finally all the selected models are compared in terms of their predictive performances at different time horizons through cross-validation to identify the best one.

The predictive performance of the models is assessed in terms of discrimination and calibration. Discrimination is assessed through the time-dependent area under the curve (AUC), estimated through the nonparametric method by Li et al. (2018). Calibration is assessed by the weighted version of the Brier score under the assumption of independent censoring (Graf et al. 1999). Higher AUC and lower Brier score indicate better discrimination and calibration, respectively.

## 3 MRC BO06 randomized clinical trial data

### 3.1 Sample cohort selection and baseline characteristics

Clinical studies usually collect information about baseline characteristics and multiple time-varying processes to measure the disease progression. Data from the MRC BO06/EORTC 80931 Randomized Controlled Trial for patients with non-metastatic high-grade osteosarcoma recruited between 1993 and 2002 (Lewis et al. 2007) were analysed. Patients were randomized between conventional (*Reg-C*) and dose-intense (*Reg-DI*) regimens. Details concerning the trial protocol are provided in Appendix A.2.

The dataset included 497 eligible patients; 19 patients who did not start chemotherapy (13) or reported an abnormal dosage of one or both agents (6) were excluded. Motivated by the clinical research question concerning the effect of doses intensity on survival, only patients who completed all six cycles within 180 days (i. e., $T_0^*$ of the *observation period*) were included in the analyses. TVCM analysis was carried out on 377 patients (75.9% of the initial sample). Among them, one subject presented $T_i < T_0^*$ and was excluded from the FunCM cohort (376 patients—75.7% of the initial sample). The final cohorts for both TVCM and FunCM analyses are shown in Fig. 2.

Follow-up starts from date of randomization ($T_0$) and the *observation period* $[T_0; T_0^*]$ is given by the first 180 days after randomization (i.e., the 6-months chemotherapy treatment period). Patients' characteristics at baseline are provided in Table 1. Three age groups were defined according to Collins et al. (2013): *child* (male: 0–12 years; female: 0–11 years), *adolescent* (male: 13–17 years; female: 12–16 years) and *adult* (male: 18 or older; female: age 17 years or older). Median follow-up time, computed using the reverse Kaplan–Meier method by Schemper and Smith (1996), was 62.19 months (IQR = [38.93; 87.46]) and 245 patients (65%) were alive at the last follow-up visit.

## 3.2 Time-varying characteristics

Due to the skewed nature of the longitudinal trajectories of both ALP and WBC biomarkers, their logarithmic transformations shifted by one were considered. The
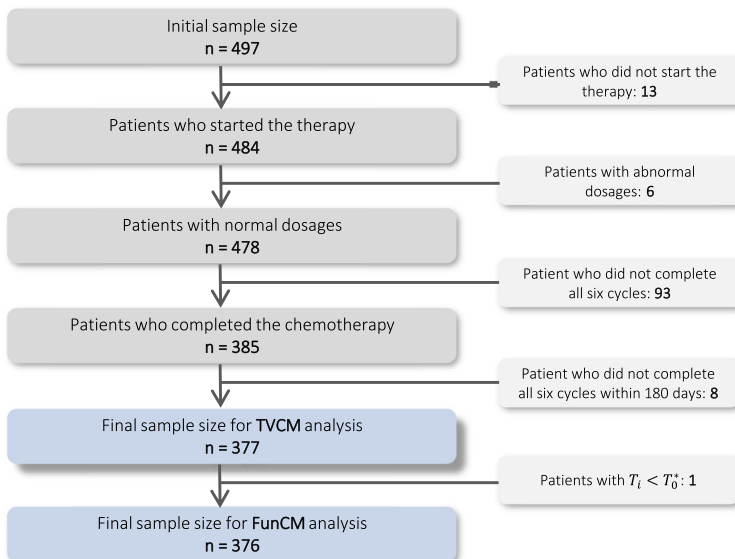


**Fig. 2** Flowchart of cohorts selection

**Table 1** Patients' characteristics at baseline

| Baseline characteristic | |
| --- | --- |
| Patients | 377 |
| Age [years] | |
| Median (IQR) | 15 (11; 18) |
| Minimum/maximum | 3/40 |
| *Child* | 109 (28.9%) |
| *Adolescent* | 154 (40.9%) |
| *Adult* | 114 (30.2%) |
| Gender | |
| *Female* | 148 (39.3%) |
| *Male* | 229 (60.7%) |
| Allocated treatment | |
| *Regimen-C* | 180 (47.7%) |
| *Regimen-DI* | 197 (52.3%) |
| White Blood Count [$\times 10^9$/L] | |
| Median (IQR) | 7.65 (6.30; 9.13) |
| Minimum/maximum | 3.60/16.20 |
| Alkaline Phospathase [IU/L] | |
| Median (IQR) | 311.5 (190.0; 551.5) |
| Minimum/maximum | 49.0/3680.0 |

vectors of longitudinal values of ALP and WBC measurements for patient $i$ are given as

$$z_i^{(ALP)} = \left\{ z_i^{(ALP)}(t_{il}), l = 1, \ldots, n_i^{(ALP)} \right\} \tag{7}$$

$$z_i^{(WBC)} = \left\{ z_i^{(WBC)}(t_{il}), l = 1, \ldots, n_i^{(WBC)} \right\} \tag{8}$$

where $t_{il}$ is the time of the $l$-th laboratory ALP or WBC test, $z_i^{(ALP)}(t_{il}) = \log(ALP_{il} + 1)$ and $z_i^{(WBC)}(t_{il}) = \log(WBC_{il} + 1)$ are the logarithmic values of ALP and WBC measurements at time $t_{il}$, $n_i^{(ALP)}$ and $n_i^{(WBC)}$ are the number of different ALP and WBC laboratory tests, respectively. Left and central panels of Fig. 3 show the longitudinal trajectories over time of $z_i^{(ALP)}$ and $z_i^{(WBC)}$ respectively. Each line represents the time-varying logarithmic biomarker values for a specific patient coloured by event status (black: *Censored*, red: *Dead*). Observed longitudinal data can be sparse and irregularly measured among patients and different biomarkers. ALP point-measurements $z_i^{(ALP)}(t_{il})$ observed among all patients over time ranged from a minimum of 2.708 to a maximum of 8.211 (corresponding to ALP values of 14 and 3680 IU/L, respectively). WBC point-measurements $z_i^{(WBC)}(t_{il})$ observed among all patients over time ranged from a minimum of 0.095 to a maximum of 4.771 (corresponding to WBC values of 0.1 and 117.0 $\times 10^9$/L, respectively). The presence in both biomarkers of extremely high/low levels compared to normal ranges
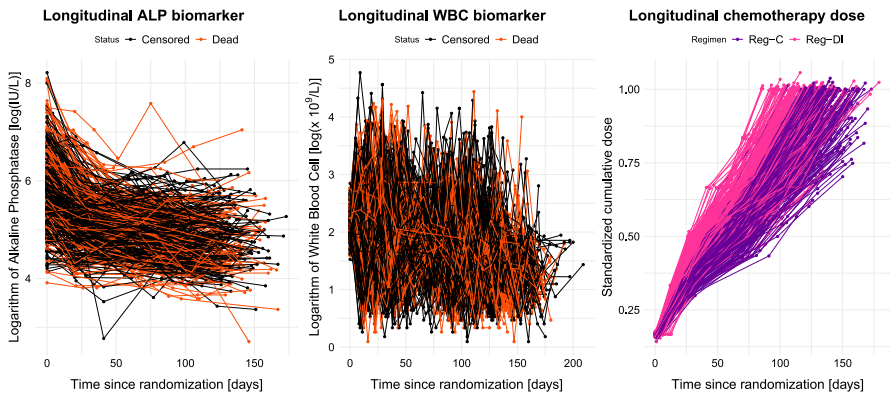
**Fig. 3** Time-varying covariates for each patient. Left panel: longitudinal logarithmic values of ALP biomarker over time coloured by event status (black: *Censored*, red: *Dead*). Central panel: longitudinal logarithmic values of WBC biomarker over time coloured by event status (black: *Censored*, red: *Dead*). Right panel: longitudinal values of standardized cumulative dose of chemotherapy coloured by allocated regimen (pink: *Reg-DI*, purple: *Reg-C*) (colour figure online)

is due to the presence of conditions usually experienced by patients in childhood cancer therapies, such as bone growth, tumour necrosis, inflammatory states, infections or toxicity (see Williamson et al. 2015).

The time-varying standardized cumulative dose of chemotherapy is now introduced. Let $j \in \{1, \ldots, 6\}$ be the cycle index and $t_{ij}$ the time of the $j$-th cycle for the $i$-th patient. The standardized cumulative dose of chemotherapy (DOX+ CDDP) for the $i$-th patient at time $t_{ij}$ is defined as:

$$z_i^{(\delta)}(t_{ij}) = \frac{\text{Cumulative dose of } DOX + CDDP \text{ until cycle } j [mg/m^2]}{\text{Total target dose at the end of six cycles } [mg/m^2]}$$

$$= \frac{1}{175 \, [mg/m^2] \cdot 6} \cdot \sum_{k=1}^{j} \frac{DOX_{ik} + CDDP_{ik}}{surface \, area_{ik}} \left[\frac{mg}{m^2}\right]. \tag{9}$$

This can be interpreted as the regulated Received Dose Intensity (rRDI) introduced by Lancia et al. (2019) evaluated over real time and not over cumulative time on treatment. For each patient $i$, the vector of longitudinal values of standardized cumulative dose of chemotherapy over time is defined as $z_i^{(\delta)} = \{z_i^{(\delta)}(t_{ij}), j = 1, \ldots, 6\}$. The right panel of Fig. 3 shows the longitudinal trajectories $z_i^{(\delta)}$ over time. Each line represents the individual time-varying standardized cumulative chemotherapy dose coloured by allocated regimen (pink: *Reg-DI*, purple: *Reg-C*). Patients - also within the same regimen - reported different values of standardized cumulative dose during time, depending on the delays and dose reductions required during chemotherapy due to toxicity. In particular, the lines form a tight bundle in the early phase of the treatment, but later they open up in a hand-fan shape because treatment adjustments are generally more frequent towards the end of the protocol. Median value of total standardized cumulative dose $z_i^{(\delta)}(t_{i6})$ was 0.998

(IQR = [0.901; 1.000]), with minimum and maximum final values equal to 0.613 and 1.056, respectively. Median value of time from randomization to last cycle $t_{i6}$ was 127 days (IQR = [114; 179]), with minimum and maximum periods of 85 and 179 days, respectively.

## 4 Results

Since the role of received chemotherapy dose and biomarkers on patient's survival is still unclear for osteosarcoma (Ren et al. 2015; Hao et al. 2017; Lancia et al. 2019b), a new time-varying/functional perspective may help in understanding their relationship, providing new insights for childhood cancer. In this regard, the methodologies proposed in Sect. 2 were applied to the MRC BO06 osteosarcoma trial. Statistical analyses were performed in the R-software environment (Core Team 2020).

### 4.1 Time-varying covariate Cox model

To study the effect of time-varying biomarkers and doses on survival, a TVCM was fitted on the cohort of 377 patients (see Fig. 2). In particular, the hazard function in Eq. (1) was adjusted for gender at randomization ($\omega_i$) and stratified by age group $g \in \{child, adolescent, adult\}$, as follows:

$$
h_{ig}(t|\omega_i, z_i(t)) = h_{0g}(t) \exp\Bigg\{ \gamma_1 \cdot gender_i + \alpha_1 \cdot z_i^{(ALP)}(t)
$$
$$
+ \alpha_2 \cdot z_i^{(WBC)}(t) + \alpha_3 \cdot 100 z_i^{(\delta)}(t) \Bigg\}
\tag{10}
$$

where $h_{0g}(t)$ is the baseline hazard function for the $g$-th age stratum, $z_i^{(ALP)}(t)$, $z_i^{(WBC)}(t)$ and $z_i^{(\delta)}(t)$ are the time-varying covariates of ALP and WBC biomarkers and standardized cumulative dose (multiplied by 100 due to its different values scale), obtained applying LOCF method to longitudinal vectors $z_i^{(ALP)}$, $z_i^{(WBC)}$ and $z_i^{(\delta)}$ respectively. In Table 2 hazard ratios along with their 95% confidence interval are shown. Gender at randomization and time-varying WBC were associated to survival, whereas time-varying ALP biomarker and chemotherapy dose showed no effects on survival. Being a male was associated to a 1.5-times faster experience of the event.

**Table 2** Estimated hazard ratios (HR) along with 95% confidence intervals (CI) from the stratified time-varying covariate Cox model (TVCM) in Eq. (10)

| Covariates | HR | 95% CI |
|---|---|---|
| Gender (male) | 1.539 | [1.046; 2.263] |
| $z_i^{(ALP)}(t)$ | 0.991 | [0.711; 1.383] |
| $z_i^{(WBC)}(t)$ | 0.647 | [0.477; 0.877] |
| $z_i^{(\delta)}(t) \cdot 100$ | 1.005 | [0.984; 1.027] |

The higher the value of WBC at time $t$, the higher the risk of death. This model ignored the continuous nature of the processes underlying the data.

## 4.2 Functional covariate Cox model

### 4.2.1 Functional representation of time-varying biomarkers and chemotherapy dose

To convert the longitudinal values of ALP and WBC biomarkers registered during the *observation period*, $\bar{z}_i^{(ALP)}$ and $\bar{z}_i^{(WBC)}$, into the functions $\tilde{x}_i^{(ALP)}(t)$ and $\tilde{x}_i^{(WBC)}(t)$, measurements by cycles were used. This implies that all time-varying values were on the same temporal domain, i.e., $t \in S_{ALP} = S_{WBC} = [1, 6]$ cycles. For both ALP and WBC biomarkers ($m = \{ALP, WBC\}$), B-spline basis functions $\phi_i^{(m)}(t)$ (ALP: 2 or 3 basis of order 2 or 3; WBC: 6 or 7 basis of order 5, according to each patient $i$) and a general functional form were used. Clinical bounds $[L_m; U_m]$ (ALP: $[0;9]$; WBC: $[0;5]$) were employed in order to include the extremely high/low levels experienced by patients during treatment. Lower bounds equal to 0 were chosen to ensure the non-negativity of the functional values. A data driven approach was used to select the upper bounds defined as $U_m = \left\lceil \max_{i,l} z_i^{(m)}(t_{il}) \right\rceil$. For each patient $i$ the following functional ALP and WBC predictors were provided:

$$\tilde{x}_i^{(ALP)}(t) = \frac{9 \cdot \exp\left[\phi_i^{(ALP)}(t)^T \widehat{\mathbf{c}}_i^{(ALP)}\right]}{1 + \exp\left[\phi_i^{(ALP)}(t)^T \widehat{\mathbf{c}}_i^{(ALP)}\right]}, \tag{11}$$

$$\tilde{x}_i^{(WBC)}(t) = \frac{5 \cdot \exp\left[\phi_i^{(WBC)}(t)^T \widehat{\mathbf{c}}_i^{(WBC)}\right]}{1 + \exp\left[\phi_i^{(WBC)}(t)^T \widehat{\mathbf{c}}_i^{(WBC)}\right]} \tag{12}$$

where $\widehat{\mathbf{c}}_i^{(m)}$ ($m = \{ALP, WBC\}$) are the vectors of coefficients estimated by regression analysis using the transformation $g(x) = \log \frac{x - L_m}{U_m - x}$. Starting from the customized functional data in Eqs. (11) and (12), the derivatives $\tilde{x}_i^{(dm)}(t)$ ($m = \{ALP, WBC\}$), which represents the rate of change in the biomarkers values over time, were reconstructed. A graphical representation of functional biomarkers curves and their derivatives are shown in Figs. 4 and 5, respectively (left panels: ALP biomarker; central panels: WBC biomarker). Each line represents the functional predictor for patient $i$ coloured according to the death-event status.

To convert the longitudinal values of standardized cumulative chemotherapy dose $z_i^{(\delta)}$ into the functional form $\tilde{x}_i^{(\delta)}(t)$, measurements in days were considered since different duration in treatment is a key-point in the chemotherapy protocol. Based on clinical motivations, the interval $S_\delta = [0, 180]$ days was selected, since all the patients completed the therapy within 180 days from randomization. B-spline basis functions $\phi_i^{(\delta)}(t)$ (5 basis of order 5), a monotone functional form and clinical bounds
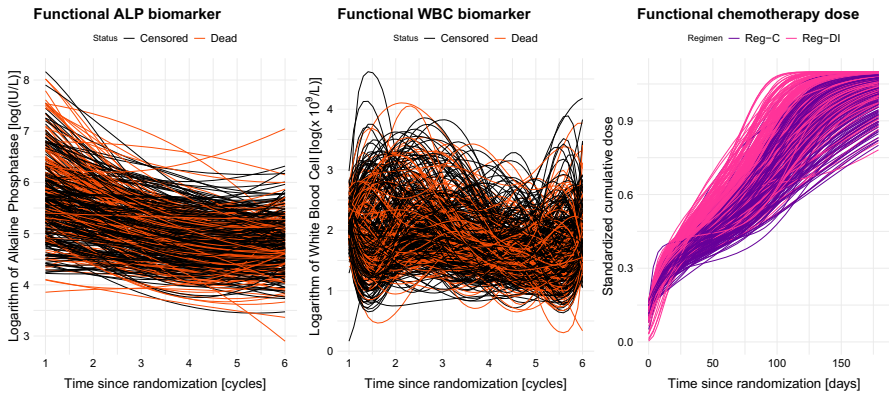
**Fig. 4** Left panel: functional representations of ALP biomarker over cycles coloured by status (black: *Censored*, red: *Dead*). Central panel: functional representations of WBC biomarker over cycles coloured by status (black: *Censored*, red: *Dead*). Right panel: functional representations of standardized cumulative dose of chemotherapy over time coloured by allocated regimen (pink: *Reg-DI*, purple: *Reg-C*). Each line is the graphical representation of the functional predictor of a patient (colour figure online)
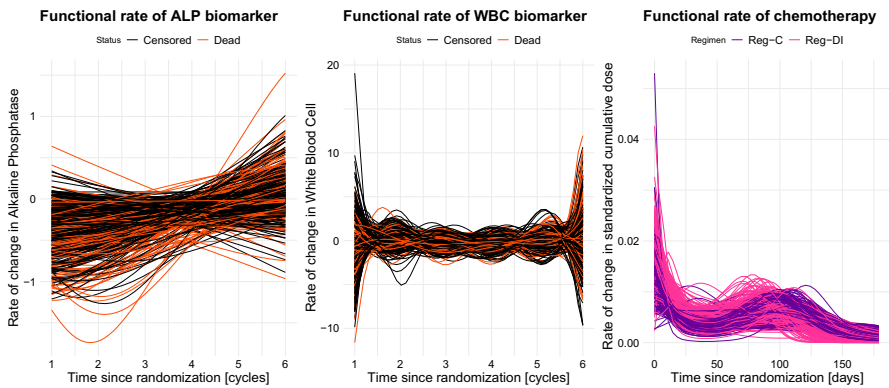


**Fig. 5** Left panel: functional representations of the rate of change of ALP biomarker over cycles coloured by status (black: *Censored*, red: *Dead*). Central panel: functional representations of the rate of change of WBC biomarker over cycles coloured by status (black: *Censored*, red: *Dead*). Right panel: functional representations of the rate of change of standardized cumulative dose of chemotherapy over time coloured by allocated regimen (pink: *Reg-DI*, purple: *Reg-C*). Each line is the graphical representation of the functional predictor of a patient (colour figure online)

$L_\delta = 0$ and $U_\delta = 1.1$ were used. For each patient $i$ a functional predictor of standardized cumulative dose of chemotherapy was obtained:

$$\tilde{x}_i^{(\delta)}(t) = \frac{1.1 \cdot \exp\left(\widehat{\beta}_{0i} + \widehat{\beta}_{1i} \int_0^t \exp\left[\boldsymbol{\phi}_i^{(\delta)}(u)^T \widehat{\mathbf{c}}_i^{(\delta)}\right] du\right)}{1 + \exp\left(\widehat{\beta}_{0i} + \widehat{\beta}_{1i} \int_0^t \exp\left[\boldsymbol{\phi}_i^{(\delta)}(u)^T \widehat{\mathbf{c}}_i^{(\delta)}\right] du\right)} \tag{13}$$

where $\widehat{\mathbf{c}}_i^{(\delta)}$ is the vector of coefficients estimated by penalized regression analysis using the transformation $g(x) = \log \frac{x - L_\delta}{U_\delta - x}$. Finally, starting from the customized

functional data in Eq. (13), the derivatives $\tilde{x}_i^{(d\delta)}(t)$, which represents the rate of change of chemotherapy dose over time, were reconstructed. A graphical representation of functional standardized cumulative dose curves $\tilde{x}_i^{(\delta)}(t)$ and their derivatives $\tilde{x}_i^{(d\delta)}(t)$ are shown in right panels of Figs. 4 and 5, respectively. Each line represents the functional predictor for patient $i$ coloured according to the allocated regimen. Functional standardised cumulative dose curves $\tilde{x}_i^{(\delta)}(t)$ (right panel in Fig. 4) also provide information on treatment adjustments. Dose reductions are represented by final standardised cumulative dose smaller than 1. For patients with a similar final dose, the slope displays information on the duration of treatment: the lower the slope, the longer the duration of treatment, reflecting delays compared to protocol (see Appendix A.2).

Figures 4 and 5 show that, taking into account the continuous nature of the processes underlying the data, a customized functional representation of the time-varying covariates and their derivatives highlights trends and variations in the shape of the processes over time.

### 4.2.2 Functional principal component analysis for time-varying biomarkers and chemotherapy

The functional trajectories provided in Eqs. (11), (12) and (13) and their derivatives were summarised into a finite set of covariates by applying Functional Principal Component Analyses (FPCAs). Only results of FPCA on functional predictors $\tilde{x}_i^{(ALP)}(t)$ and $\tilde{x}_i^{(\delta)}(t)$ are presented. In both cases, two principal components were enough to account for at least 95% of the observed variability.
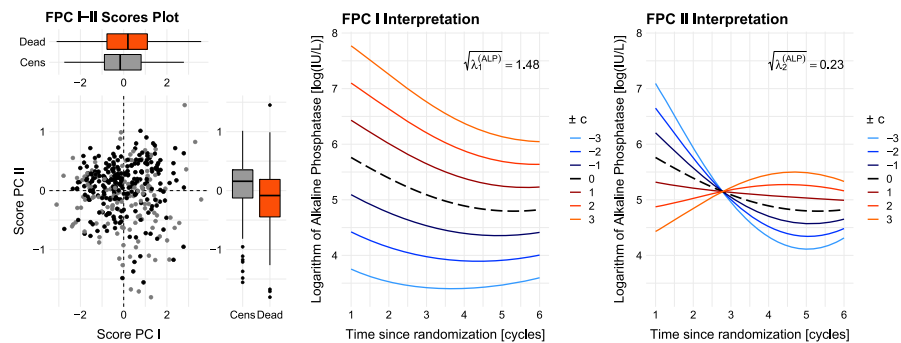


**Fig. 6** FPCA for functional Alkaline Phosphatase $\tilde{x}_i^{(ALP)}(t)$ Left panel: Functional PC scores plot $\left(f_{i1}^{(ALP)}, f_{i2}^{(ALP)}\right)$ with boxplots (black: *Censored*, red: *Dead*) Central panel: Interpretation of first FPC $\xi_1^{(ALP)}$—average standardized cumulative dose $\mu^{(ALP)}(t) \pm c\sqrt{\lambda_1^{(ALP)}} \cdot \xi_1^{(ALP)}$, with $\sqrt{\lambda_1^{(ALP)}} = 1.48$ and $\pm c = \pm 1, \pm 2, \pm 3$ Right panel: Interpretation of second FPC $\xi_2^{(ALP)}$—average standardized cumulative dose $\mu^{(ALP)}(t) \pm c\sqrt{\lambda_2^{(ALP)}} \cdot \xi_2^{(ALP)}$, with $\sqrt{\lambda_2^{(ALP)}} = 0.23$ and $\pm c = \pm 1, \pm 2, \pm 3$ (colour figure online)

Results of FPCA on functional ALP predictors $\tilde{x}_i^{(ALP)}(t)$ are provided in Fig. 6. Left panel reports the FPC scores plot $\left(f_{i1}^{(ALP)}, f_{i2}^{(ALP)}\right)$ with relative boxplots, which show the distributions of the estimated FPC score values among censored and dead patients. Each point represents a patient coloured by status (black: *Censored*, red: *Dead*). Central and right panels displays how to interpret the first two Principal Components $\xi_k^{(ALP)}$, showing the average ALP curve $\mu^{(ALP)}(t) \pm c\sqrt{\lambda_k^{(ALP)}} \cdot \xi_k^{(ALP)}$ where $\lambda_k^{(ALP)}$ is the is eigenvalue related to the $k$-th component and $c$ are constants chosen in order to let the scores values lie within one, two or three ($\pm c = \pm 1, \pm 2, \pm 3$) standard deviations (i.e., square roots of $\lambda_k^{(ALP)}$). The first component $\xi_1^{(ALP)}$ explained 83.8% of the variability and a positive (negative) score reflected higher (lower) values of ALP trajectories during treatment compared to the mean (left panel). The second component $\xi_2^{(ALP)}$ explained 13.1% of the variability and positive scores reflected "flat" curves, whereas negative score reflected curves with highly negative slopes in the first cycles (right panel). The lower the score, the higher the ALP levels during the first two cycles of the treatment. FPC scores thus summarize the different patterns of the functional biomarker trajectories between patients during treatment, being a more informative representation than the baseline value or the last available measure used through LOCF.

Results of FPCA on functional standardized cumulative dose $\tilde{x}_i^{(\delta)}(t)$ are shown in Fig. 7. Left panel reports the FPC scores plot $\left(f_{i1}^{(\delta)}, f_{i2}^{(\delta)}\right)$ with relative boxplots, which show the distributions of the estimated FPC score values among the two regimens. Each point corresponds to a patient. Different colours represent the two regimens. Central and right panels displays how to interpret the first two Principal
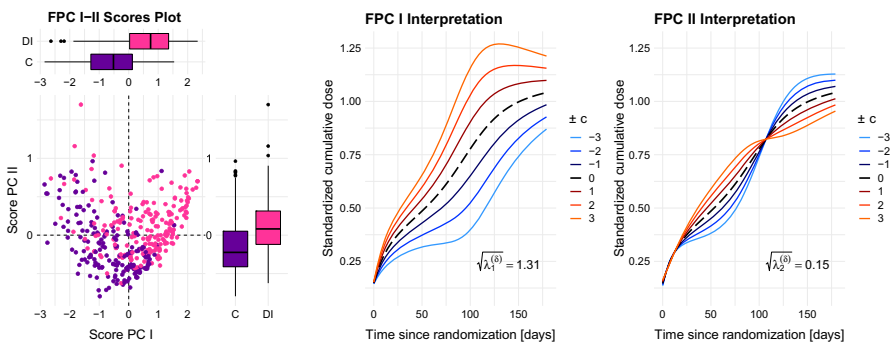


**Fig. 7** FPCA for functional standardized cumulative dose $\tilde{x}_i^{(\delta)}(t)$ Left panel: Functional PC scores plot $\left(f_{i1}^{(\delta)}, f_{i2}^{(\delta)}\right)$ with boxplots (pink: *Reg-DI*, purple: *Reg-C*) Central panel: Interpretation of first FPC $\xi_1^{(\delta)}$— average standardized cumulative dose $\mu^{(\delta)}(t) \pm c\sqrt{\lambda_1^{(\delta)}} \cdot \xi_1^{(\delta)}$, with $\sqrt{\lambda_1^{(\delta)}} = 1.31$ and $\pm c = \pm 1, \pm 2, \pm 3$ Right panel: Interpretation of second FPC $\xi_2^{(\delta)}$—average standardized cumulative dose $\mu^{(\delta)}(t) + \pm c\sqrt{\lambda_2^{(\delta)}} \cdot \xi_2^{(\delta)}$, with $\sqrt{\lambda_1^{(\delta)}} = 0.15$ and $\pm c = \pm 1, \pm 2, \pm 3$ (colour figure online)

Components $\xi_k^{(\delta)}$, showing the average curve $\mu^{(\delta)}(t) \pm c\sqrt{\lambda_k^{(\delta)}} \cdot \xi_k^{(\delta)}$ where $\lambda_k^{(\delta)}$ is the is eigenvalue related to the $k$-th component and $c$ are constants chosen in order to let the scores values lie within one, two or three ($\pm c = \pm 1, \pm 2, \pm 3$) standard deviations (i.e., square roots of $\lambda_k^{(\delta)}$). The first component $\xi_1^{(\delta)}$ explained 86.9% of the variability and reflects information on treatment administration and adjustments with respect to protocol. Positive scores (i.e., curves above the average $\mu^{(\delta)}(t)$ in the left panel) indicate patients without dose-reduction (i.e., their final standardized cumulative dose is greater or equal to 1) and with possible delays in treatment: the lower the positive score, the higher the time needed to end the treatment. Negative scores (i.e., curves below the average $\mu^{(\delta)}(t)$) represent patients with both time-delays and dose-reduction: the lower the negative score, the higher the total dose-reduction. The second component $\xi_2^{(\delta)}$ explained 9.8% of the variability and a positive score indicated a faster growth in the chemotherapy assumption in the first period compared to the second one, with respect to the mean (right panel). Every two patients reported different values of FPC scores, reflecting delays or dose reductions during chemotherapy. This representation illustrates different treatment dynamics, also among patients allocated to the same regimen. Summarizing differences in both trends and variations related to the shape of chemotherapy doses consumption processes over time, the use of FPC scores is more informative than an IIT analysis by different allocated regimens or a LOCF approach that considers only the last available value.

### 4.2.3 Multivariate functional linear Cox regression model

To study the effect of risk factors on survival, several MFLCRMs based on different sets of baseline and functional predictors (see Table 3) were estimated. Since functional trajectories and their relative derivatives are correlated, in each MFLCRM only one type was considered. Each model was adjusted for gender and stratified by

**Table 3** Selected truncation parameters $K_m$ and integrated AUC (iAUC) for different sets of baseline and functional predictors

| Model | Baseline | Truncation parameters $K_m$ | | | | | | iAUC |
|---|---|---|---|---|---|---|---|---|
| | $\omega_i$ | ALP | dALP | WBC | dWBC | $\delta$ | $d\delta$ | |
| 1 | $(gender_i)$ | 2 | – | 7 | – | 1 | – | 0.650 |
| 2 | $(gender_i)$ | 2 | – | 7 | – | – | 1 | 0.635 |
| 3 | $(gender_i, wbc_i)$ | 2 | – | – | 4 | 2 | – | 0.666 |
| 4 | $(gender_i, wbc_i)$ | 2 | – | – | 4 | – | 3 | 0.664 |
| 5 | $(gender_i, alp_i, wbc_i)$ | – | 2 | – | 4 | 2 | – | 0.650 |
| 6 | $(gender_i, alp_i, wbc_i)$ | – | 2 | – | 4 | – | 3 | 0.647 |
| 7 | $(gender_i, alp_i)$ | – | 1 | 7 | – | 1 | – | 0.645 |
| 8 | $(gender_i, alp_i)$ | – | 1 | 7 | – | – | 1 | 0.641 |
| TVCM | | | | | | | | 0.592 |

iAUC for stratified time-varying covariate Cox model (TVCM) in Eq. (10)
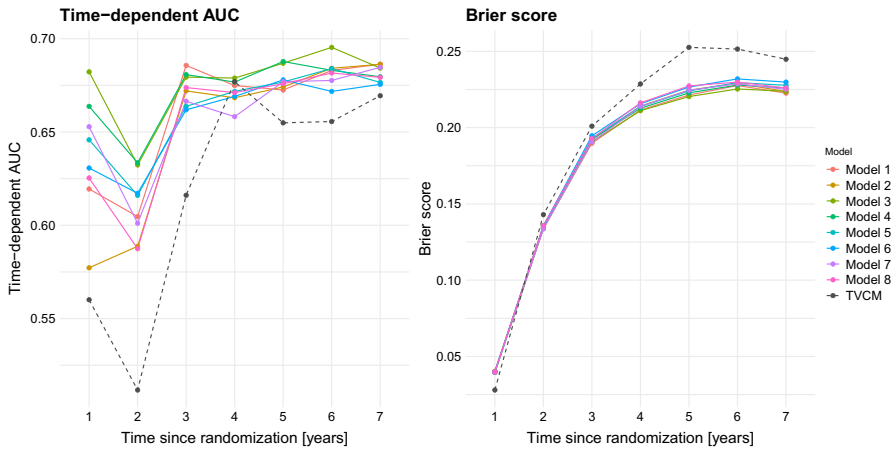
**Fig. 8** Left panel: time-dependent AUC over different time horizons (from 1 to 7 years after randomization) for Models 1–8 of Table 3 (solid coloured lines) and TVCM in Eq. (10) (dashed black line). Right panel: Brier score over different time horizons (colour figure online)

age group at randomization $g \in \{child, adolescent, adult\}$. When functional rate of changes of ALP or WBC biomarkers were included in the models, the values of logarithmic ALP or WBC levels at randomization were also considered as adjusting baseline covariates. Cross-validation with five folds was employed to select the truncation parameters $K_m$ for each set of covariates (see Table 3). Time-dependent AUCs and Brier scores were estimated with R packages tdROC (function tdROC) by Li and Wu (2016) and ipred (function sbrier) by Peters and Hothorn (2019), respectively. Figure 8 shows the cross-validated mean values of time-dependent AUC and Brier score over different time horizons for all estimated models (solid lines) and for TVCM in Eq. (10) (dashed black lines). All functional models outperformed TVCM and showed similar Brier score measures over time, therefore time-dependent AUC was used to select the final model. Weighted averages of the several time-dependent AUCs over time, estimated through the integrated AUCs (iAUC) by Heagerty and Zheng (2005), are reported in Table 3. According to the highest iAUC, the best MFLCRM was Model 3, defined as follows:

$$
h_{ig}\left(t|\boldsymbol{\omega}_i, \tilde{x}_i^{(ALP)}(t), \tilde{x}_i^{(dWBC)}(t), \tilde{x}_i^{(\delta)}(t)\right)
$$

$$
= h_{0g}(t) \exp\left\{\gamma_1 gender_i + \gamma_2\, wbc_i + \sum_{k=1}^{2} f_{ik}^{(ALP)}\alpha_k^{(ALP)}\right.
$$

$$
\left. + \sum_{k=1}^{4} f_{ik}^{(dWBC)}\alpha_k^{(dWBC)} + \sum_{k=1}^{2} f_{ik}^{(\delta)}\alpha_k^{(\delta)}\right\} \tag{14}
$$

where $h_{0g}(t)$ is the baseline hazard function for the $g$-th age stratum, $\boldsymbol{\omega}_i = (gender_i, wbc_i)$ is the vector of baseline covariates; $\tilde{x}_i^{(ALP)}(t)$, $\tilde{x}_i^{(dWBC)}(t)$ and $\tilde{x}_i^{(\delta)}(t)$ are the functional predictors of ALP biomarker, rate of change of WBC and

standardized cumulative dose, respectively, with relative FPC scores $f_{ik}^{(m)}$ ($k = 1, \ldots, K_m; m \in \{ALP, dWBC, \delta\}; K_{ALP} = 2; K_{dWBC} = 4; K_\delta = 2$).

To estimate the effect of the selected functional predictors on survival, MFLCRM (14) was fitted on the FunCM cohort of 376 patients (see Fig. 2). In Table 4 hazard ratios along with their 95% confidence interval are shown. Level of WBC at randomization and the FPC scores related to alkaline phosphatase $f_{i1}^{(ALP)}, f_{i2}^{(ALP)}$ were associate to survival. The higher the value of WBC at randomization the higher the risk of death, whereas no effects were observed due to the rate of change in WBC during the protocol observation period. Patients with high ALP trajectories had poor survival, especially in case of curves with highly negative slopes during the first cycles of chemotherapy protocol. FPC scores related to functional chemotherapy dose showed no effects on survival. Estimated survival probabilities are shown in Fig. 9. High values of baseline WBC corresponded to poor survival (top-left panel). The score $f_{i1}^{(\delta)}$ related to the first PC of functional chemotherapy indicated that there was no improvement on survival due to dose-intense profiles (top-right panel). The effect of functional ALP biomarker suggested that patients with high ALP trajectories over time (i.e., high value of $f_{i1}^{(ALP)}$—bottom-left panel), especially during the first cycles of the chemotherapy protocol (i.e., low value of $f_{i2}^{(ALP)}$ - bottom-right panel), had poor survival.

## 5 Discussion

To study the association between time-varying processes and time-to-event data is a challenging problem in clinical research and the development of models and methods able to deal with dynamic time-varying covariates is of statistical interest and of clinical relevance. Research into functional modelling have received considerable attention in recent years. In this work, a novel approach based on FDA techniques to investigate the dynamics of time-varying processes over time and to include

**Table 4** Estimated hazard ratios (HR) along with 95% confidence intervals (CI) from the multivariate functional linear Cox regression model

| Covariates | HR | 95% CI |
|---|---|---|
| gender (male) | 1.431 | [0.964; 2.123] |
| wbc | 3.169 | [1.525; 6.585] |
| $f_1^{(ALP)}$ | 1.210 | [1.018; 1.437] |
| $f_2^{(ALP)}$ | 0.554 | [0.399; 0.768] |
| $f_1^{(\delta)}$ | 0.869 | [0.719; 1.051] |
| $f_2^{(\delta)}$ | 0.885 | [0.547; 1.432] |
| $f_1^{(dWBC)}$ | 0.990 | [0.889; 1.102] |
| $f_2^{(dWBC)}$ | 0.916 | [0.789; 1.062] |
| $f_3^{(dWBC)}$ | 1.161 | [0.892; 1.512] |
| $f_4^{(dWBC)}$ | 1.219 | [0.898; 1.655] |
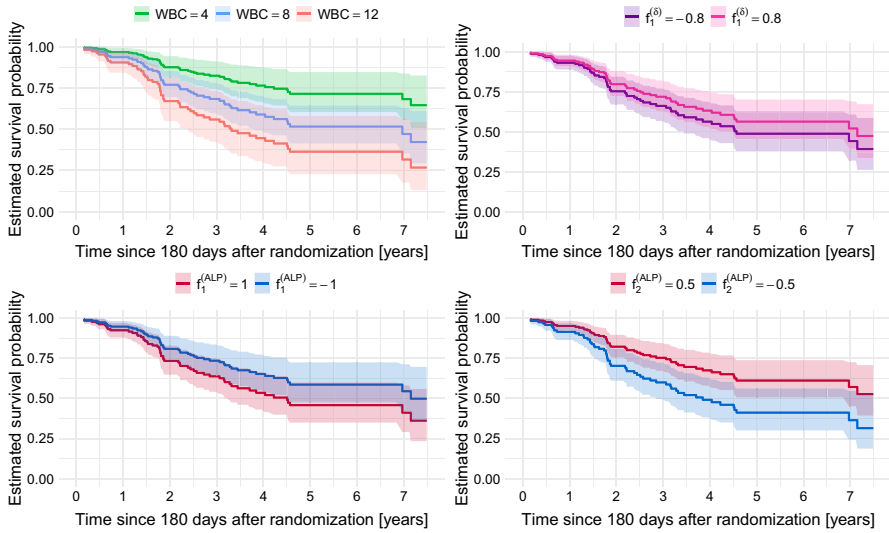
**Fig. 9** Estimated survival probability based on the multivariate functional linear Cox regression model (14). Time $t_0 = 0$ corresponds to $T_0^*$ in Fig. 1. Top-left panel: patients with different values of WBC [$\times 10^9$/L] at randomization (green: $WBC = 4$; blue: $WBC = 8$; red: $WBC = 12$). Top-right panel: patients with different values of the first PC score for functional chemotherapy (purple: $f_1^{(\delta)} = -0.8$; pink: $f_1^{(\delta)} = 0.8$). Bottom-left panel: patients with different values of the first PC score for functional ALP biomarker (red: $f_1^{(ALP)} = 1$; blue: $f_1^{(ALP)} = -1$). Bottom-right panel: patients with different values of the second PC score for functional ALP biomarker (red: $f_2^{(ALP)} = 0.5$; blue: $f_2^{(ALP)} = -0.5$). When not specified, the other risk factors are fixed to the most frequent class for categorical covariates, i.e., *adolescent males*, and to the median value for continuous ones, i.e., $WBC = 7.65 \times 10^9/L$ at randomization, $f_1^{(\delta)} = 0.08$, $f_2^{(\delta)} = -0.03$, $f_1^{(ALP)} = -0.10$, $f_2^{(ALP)} = 0.07$, $f_1^{(dWBC)} = -0.12$, $f_2^{(dWBC)} = -0.02$, $f_3^{(dWBC)} = -0.07$ and $f_4^{(dWBC)} = -0.08$ (colour figure online)

additional information that may be related to the survival into the time-to-event model was presented. Data from the MRC BO06/EORTC 80931 randomized clinical trial for osteosarcoma treatment were analysed. Biomarkers and chemotherapy dose were incorporated as time-varying covariates into time-to-event models using both a TVCM and a FunCM approach. The standard TVCM with LOCF approach ignored the continuous nature of the processes underlying the data. To overcome this issue, FunCM exploited FDA techniques to represent time-varying characteristics in terms of functions, enriching the information available for modelling survival with relevant time-varying features related to the evolution of the processes over time. These features were included into MFLCRMs by FPCA to study the effects of functional risk factors on patients' overall survival.

Differences in results for TVCM and MFLCRM were due to the different nature of the information incorporated in the two models. TVCM considered as constants the last biomarkers/dose levels over different time points (expressed in days). In practice, among the measurements recorded during the *observation period*, only the last value had any real impact on overall survival, as only one patient presented with a time-to-event of less than 180 days. This discarded both information about the

continuous nature of the processes and the history of the actual levels measured. MFLCRM included information related to different levels variations and timing during the entire *observation period*, and functional biomarkers were defined over cycles. Thanks to the introduction of relevant dynamic features related to the continuous functional nature of the processes, MFLCRM resulted more informative than TVCM, outperforming it both in terms of calibration and discrimination over time. MFLCRM results suggested that osteosarcoma patients with high ALP trajectories during treatment, especially during the first cycles of the chemotherapy protocol, have poor overall survival. Dose-intense profiles were not associated with survival, even if functional chemotherapy representations were able to capture individual realisations of the intended treatment, detecting differences between patients randomised to the same regimen. This suggested that considering only the assumed dose as treatment proxy is not enough. Chemotherapy presents some particular aspects, such as latent accumulation of toxicity, which must be taken into account (Lancia et al. 2019).

The proposed FunCM focused on the representation and the reconstruction of the functional trajectories related to the time-varying processes of interest. Such data are usually considered in a very simplistic way in cancer prediction models, where they act as fixed baseline or as time-dependent LOCF covariates. In this way the amount of information they may provide is not considered, as it is often the changing patterns of the functional trajectories rather than the baseline/last value that affects patients' survival. The strength and innovation of FunCM was the ability to capture the individual realisations of the process over time through a customized functional reconstruction. The developed techniques allowed (i) to account for the continuous time-varying nature of the processes underlying the data and their properties, such as nonlinearity, positivity, constraints, monotonicity, (ii) to move from sparse and irregular longitudinal data to functions defined over a common continuous domain, overcoming the issues of values missingness and different temporal grids, and (iii) to reconstruct and provide derivatives information in a tailored way. The use of derivatives is important both in extending the range of simple graphical exploratory methods and in the development of more detailed methodology (Ramsay and Silverman 2005). In fact, interesting patterns are often much more apparent in derivatives than in the original curves. Furthermore, through a proper dimensionality reduction technique, this methodology allowed to extract additional information contained in the functions. This result is an effective exploratory and modelling technique to highlight trends and variations in the evolution of the processes over time.

In contrast to a TVCM approach, the use of FunCM requires that patients survived for a period at least equal to the length of the *observation period* used to compute the functional predictors. This might imply a loss of information in situations with high rate of mortality during the *observation period* (that is not the case under study as only one of the cohort patients who had completed the chemotherapy treatment protocol died during the first 6-months after randomization—see Fig. 2). In those cases, a joint modelling approach can be used to overcome both LOCF and selection bias issues, since its allows the simultaneous modelling of longitudinal and time-to-event outcomes. However, joint models are computational expensive in case of

multiple longitudinal outcomes and require assumptions on the distributions of the processes that need to be carefully validated to avoid biased estimates.

This work opens doors to many further developments, both in the field of statistical methods and in cancer research. The dimensionality reduction via FPCA is just one way to work with these data in order to use them within inferential contexts. In fact, the reconstruction via FDA allows to properly use the functional data to address relevant clinical research questions, according to the needs of the analysis and the outcomes of interest. From a clinical point of view, it will be necessary to simultaneously consider chemotherapy treatment modifications and the occurrence of adverse events. This aspect need to be taken into account into the representation of the dynamic evolution of these processes. To model them simultaneously is not a trivial task.

The complexity of the chemotherapy treatment asks for the developments of new methodologies. This study shows that working in this direction is a difficult but profitable approach, which could lead to new improvements for subject-specific survival predictions and personalised treatment.
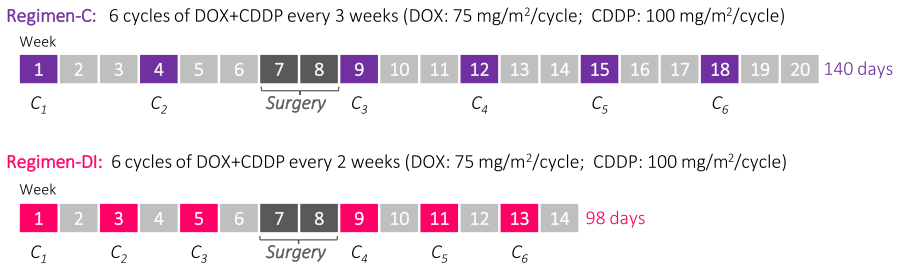
## Appendix

### A.1. Functional covariate Cox model: summary of the procedure

The developed Functional covariate Cox Model (FunCM) may be summarized in four steps. Steps 1 and 2 are devoted to reconstruct the functional trajectories. Steps 3 and 4 provide a suitable framework for including such time-varying covariates in a time-to-event model.

1. *Data preprocessing and time-varying characteristics* Cohort of patients is selected. For each process $m$ and patient $i$, longitudinal time-varying characteristics $\bar{z}_i^{(m)}$ during the *observation period* $[T_0; T_0^*]$ (see Fig. 1) under study are identified.
2. *Reconstruct functional predictors* For each process $m$ and patient $i$, the functional data $\tilde{x}_i^{(m)}(t)$ and their derivatives $\tilde{x}_i^{(dm)}(t)$ are reconstructed, applying the FDA techniques introduced in Sect. 2.3.1.
3. *Summarize functional trajectories through FPCA* Functional Principal Component Analysis (FPCA) is applied in order to perform a dimensionality reduction and summarise information emerging from the functional predictors to a finite set of covariates, i.e., the FPC scores $f_{ik}^{(m)}$ ($k = 1, \ldots, K_m$) in Eq. (6).
4. *Multivariate functional linear Cox regression model (MFLCRM)* K-fold cross validation is applied to select the best set of covariates and the number of principal components $K_m$ to consider for each process $m$, according to time-dependent AUC and Brier score as mentioned in Sect. 2.3.3. Finally, the MFLCRM given by Eq. (6) is fitted on the whole dataset in order to quantify the association between time-varying processes and patients' overall survival.

Regimen-C: 6 cycles of DOX+CDDP every 3 weeks (DOX: 75 mg/m$^2$/cycle; CDDP: 100 mg/m$^2$/cycle)



Regimen-DI: 6 cycles of DOX+CDDP every 2 weeks (DOX: 75 mg/m$^2$/cycle; CDDP: 100 mg/m$^2$/cycle)



**Fig. 10** Patients are randomized at baseline to one of the two regimens, with the same anticipated cumulative dose but different duration

## A.2. MRC BO06/EORTC 80931 randomized controlled trial protocol

Data from the MRC BO06/EORTC 80931 Randomized Controlled Trial for patients with non-metastatic high-grade osteosarcoma recruited between 1993 and 2002 were analysed (Lewis et al. 2007). The trial randomised patients between conventional treatment with doxorubicin (DOX) and cisplatin (CDDP) given every 3 weeks (*Reg-C*) versus a dose-intense regimen of the same two drugs given every 2 weeks (*Reg-DI*), supported by granulocyte colony-stimulating factor. Chemotherapy was administered for six cycles (a cycle is a period of either 2 or 3 weeks depending on the allocated regimen), before and after surgical removal of the primary osteosarcoma. In both arms, DOX (75 mg/m$^2$) plus CDDP (100 mg/m$^2$) were given over six cycles. Surgery to remove the primary tumour was scheduled at week 6 after starting treatment in both arms, that is, after 2 cycles (2 × [DOX+CDDP]) in regimen-C and after 3 cycles (3 × [DOX+CDDP]) in regimen-DI. Postoperative chemotherapy was intended to resume 2 weeks after surgery in both arms. Figure 10 shows the trial design. Planned total cumulative dose was 1,050 mg/m$^2$ in both regimens. Planned treatment time from beginning first cycle was 140 and 98 days for *Reg-C* (6 cycles · 3 weeks/cycle · 7 days/week + 14 days of surgery period) and *Reg-DI* (6 cycles · 2 weeks/cycle · 7 days/week + 14 days of surgery period), respectively. Laboratory tests were usually performed before each cycle of chemotherapy (in some cases also during and after the cycle) in order to monitor patient's health status and the development of toxicities or adverse events. Delays or chemotherapy dose reductions during treatment were possible in case of toxicity. Additional details can be found in the primary analysis of the trial by Lewis et al. (2007).

**Data availability** Data are not publicly available due to privacy restrictions. Access to the full dataset of MRC BO06 trial can be requested to MRC Clinical Trials Unit at UCL, Institute of Clinical Trials and Methodology, UCL, London.

**Code availability** R code is provided here: https://github.com/mspreafico/BO06-FunCM.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

## References

Anninga JK, Gelderblom H, Fiocco M, Kroep JR, Taminiau AH, Hogendoorn PC, Egeler RM (2011) Chemotherapeutic adjuvant treatment for osteosarcoma: where do we stand? Eur J Cancer 47 (16):2431–2445. https://doi.org/10.1016/j.ejca.2011.05.030

Arisido MW, Antolini L, Bernasconi D, Valsecchi MG, Rebora P (2019) Joint model robustness compared with the time-varying covariate Cox model to evaluate the association between a longitudinal marker and a time-to-event endpoint. BMC Med Res Methodol 19:222. https://doi.org/10.1186/s12874-019-0873-y

Austin PC, Latouche A, Fine JP (2020) A review of the use of time-varying covariates in the Fine-Gray subdistribution hazard competing risk regression model. Stat Med 39(2):103–113. https://doi.org/10.1002/sim.8399

Chi Y, Ibrahim JG (2006) Joint models for multivariate longitudinal and multivariate survival data. Biometrics 62(2):432–445. https://doi.org/10.1111/j.1541-0420.2005.00448.x

Collins M, Wilhelm M, Conyers R et al (2013) Benefits and adverse events in younger versus older patients receiving neoadjuvant chemotherapy for osteosarcoma: findings from a meta-analysis. J Clin Oncol 31(18):2303–2312. https://doi.org/10.1200/JCO.2012.43.8598

Cox DR (1972) Regression models and life-tables (with discussion). J R Stat Soc B 34(2):187–220

Dantan E, Joly P, Dartigues JF, Jacqmin-Gadda H (2011) Joint model with latent state for longitudinal and multistate data. Biostat 12(4):723–736. https://doi.org/10.1093/biostatistics/kxr003

Ferraty F, Vieu P (2006) Nonparametric Functional Data Analysis: Theory and Practice. Springer Series in Statistics, Springer, New York

Gellar JE, Colantuoni E, Needham DM, Crainiceanu CM (2015) Cox regression models with functional covariates for survival data. Stat Model 15(3):256–278. https://doi.org/10.1177/1471082X14565526

Gould LA, Boye ME, Crowther MJ, Ibrahim JG, Quartey G, Micallef S, Bois FY (2015) Joint modeling of survival and longitudinal non-survival data: current methods and issues. report of the DIA Bayesian joint modeling working group. Stat Med 34(14):2181–2195. https://doi.org/10.1002/sim.6141

Graf E, Schmoor C, Sauerbrei W, Schumacher M (1999) Assessment and comparison of prognostic classification schemes for survival data. Stat Med 18(17–18):2529–2545. https://doi.org/10.1002/(sici)1097-0258(19990915/30)18:17/18<2529::aid-sim274>3.0.co;2-5

Gupta SK (2011) Intention-to-treat concept: a review. Perspect Clin Res 2(3):109–112. https://doi.org/10.4103/2229-3485.83221

Hao H, Chen L, Huang D, Ge J, Qiu Y, Hao L (2017) Meta-analysis of alkaline phosphatase and prognosis for osteosarcoma. Eur J Cancer Care 26(5):e12536. https://doi.org/10.1111/ecc.12536

Happ C, Greven S (2018) Multivariate functional principal component analysis for data observed on different (dimensional) domains. J Am Stat Assoc 113(522):649–659. https://doi.org/10.1080/01621459.2016.1273115

Heagerty PJ, Zheng Y (2005) Survival model predictive accuracy and ROC curves. Biometrics 61(1):92–105. https://doi.org/10.1111/j.0006-341X.2005.030814.x

Henderson R, Diggle P, Dobson A (2000) Joint modelling of longitudinal measurements and event time data. Biostat 1(4):465–480. https://doi.org/10.1093/biostatistics/1.4.465

Hickey GL, Philipson P, Jorgensen A, Kolamunnage-Dona R (2016) Joint modelling of time-to-event and multivariate longitudinal outcomes: recent developments and issues. BMC Med Res Methodol 16 (117). https://doi.org/10.1186/s12874-016-0212-5

Hickey GL, Philipson P, Jorgensen A, Kolamunnage-Dona R (2018) A comparison of joint models for longitudinal and competing risks data, with application to an epilepsy drug randomized controlled trial. J R Stat Soc A 181(4):1105–1123. https://doi.org/10.1111/rssa.12348

Ieva F, Paganoni AM (2016) Risk prediction for myocardial infarction via generalized functional regression models. Stat Methods Med Res 25(4):1648–1660. https://doi.org/10.1177/0962280213495988

Ieva F, Paganoni AM, Pigoli D, Vitelli V (2013) Multivariate functional clustering for the morphological analysis of electrocardiograph curves. J R Stat Soc C 62(3):401–418. https://doi.org/10.1111/j.1467-9876.2012.01062.x

Kalbfleisch J, Prentice R (2002) The Statistical Analysis of Failure Time Data, Second Edition. Wiley

Kong D, Ibrahim JG, Lee E, Zhu H (2018) FLCRM: functional linear Cox regression model. Biometrics 74 (1):109–117. https://doi.org/10.1111/biom.12748

Lancia C, Anninga J, Sydes MR, Spitoni C, Whelan J, Hogendoorn PCW, Gelderblom H, Fiocco M (2019) A novel method to address the association between received dose intensity and survival outcome: benefits of approaching treatment intensification at a more individualised level in a trial of the European Osteosarcoma Intergroup. Cancer Chemother Pharmacol 83(5):951–962. https://doi.org/10.1007/s00280-019-03797-3

Lancia C, Anninga J, Sydes MR, Spitoni C, Whelan J, Hogendoorn PCW, Gelderblom H, Fiocco M (2019b) Method to measure the mismatch between target and achieved received dose intensity of chemotherapy in cancer trials: a retrospective analysis of the MRC BO06 trial in osteosarcoma. BMJ Open 9(5).https://doi.org/10.1136/bmjopen-2018-022980

Lancia C, Spitoni C, Anninga J, Whelan J, Sydes MR, Jovic G, Fiocco M (2019) Marginal structural models with dose-delay joint-exposure for assessing variations to chemotherapy intensity. Stat Methods Med Res 28(9):2787–2801. https://doi.org/10.1177/0962280218780619

Lee E, Zhu H, Kong D, Wang Y, Giovanello KS, Ibrahim JG (2015) BFLCRM: a Bayesian functional linear Cox regression model for predicting time to conversion to Alzheimer's disease. Ann Appl Stat 9(4):2153–2178. https://doi.org/10.1214/15-AOAS879

Lewis I, Nooij M, Whelan J et al (2007) Improvement in histologic response but not survival in osteosarcoma patients treated with intensified chemotherapy: a randomized phase III trial of the European osteosarcoma intergroup. J Natl Cancer Inst 99(2):112–128. https://doi.org/10.1093/jnci/djk015

Li K, Luo S (2019) Dynamic prediction of Alzheimer's disease progression using features of multiple longitudinal outcomes and time-to-event data. Stat Med 38(24):4804–4818. https://doi.org/10.1002/sim.8334

Li L, Wu C (2016) tdROC: Nonparametric estimation of time-dependent ROC curve from right censored survival data. https://CRAN.R-project.org/package=tdROC, R package version 1.0

Li L, Greene T, Hu B (2018) A simple method to estimate the time-dependent receiver operating characteristic curve and the area under the curve with right censored data. Stat Methods Med Res 27 (8):2264–2278. https://doi.org/10.1177/0962280216680239

Liu X, Yang MCK (2009) Identifying temporally differentially expressed genes through functional principal components analysis. Biostatistics 10(4):667–679. https://doi.org/10.1093/biostatistics/kxp022

Martino A, Ghiglietti A, Ieva F, Paganoni AM (2019) A k-means procedure based on a Mahalanobis type distance for clustering multivariate functional data. Stat Methods Appl 28(2):301–322. https://doi.org/10.1007/s10260-018-00446-6

Müller H (2005) Functional modelling and classification of longitudinal data. Scand J Stat 32:223–240. https://doi.org/10.1111/j.1467-9469.2005.00429.x

Pencina MJ, D'Agostino RB (2004) Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. Statist Med 23:2109–2123. https://doi.org/10.1002/sim.1802

Peters A, Hothorn T (2019) ipred: Improved predictors. https://CRAN.R-project.org/package=ipred, R package version 0.9-9

Proust-Lima C, Dartigues J, Jacqmin-Gadda H (2016) Joint modeling of repeated multivariate cognitive measures and competing risks of dementia and death: a latent process and latent class approach. Stat Med 35(3):382–398. https://doi.org/10.1002/sim.6731

Qu S, Wang JL, Wang X (2016) Optimal estimation for the functional Cox model. Ann Stat 44(4):1708–1738. https://doi.org/10.1214/16-AOS1441

R Core Team (2020) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Ramsay JO, Silverman BW (2002) Applied Functional Data Analysis: Methods and Case Studies. Springer Series in Statistics, Springer, New York

Ramsay JO, Silverman BW (2005) Functional Data Analysis. Springer Series in Statistics, Springer, New York

Ren H, Sun L, Li H, Ye Z (2015) Prognostic significance of serum alkaline phosphatase level in osteosarcoma: a meta-analysis of published data. BioMed Res Int 2015(Article ID 160835). https://doi.org/10.1155/2015/160835

Ritter J, Bielack SS (2010) Osteosarcoma. Ann Oncol 21(suppl 7):vii320–vii325. https://doi.org/10.1093/annonc/mdq276

Rizopoulos D (2012) Joint Models for Longitudinal and Time-to-Event Data with Applications in R. Chapman & Hall/CRC

Rizopoulos D (2016) The R Package JMbayes for fitting joint models for longitudinal and time-to-event data using MCMC. J Stat Softw 72(7):1–46. https://doi.org/10.18637/jss.v072.i07

Schemper M, Smith TL (1996) A note on quantifying follow-up in studies of failure time. Control Clin Trials 17(4):343–346. https://doi.org/10.1016/0197-2456(96)00075-x

Smeland S, Bielack SS, Whelan J et al (2019) Survival and prognosis with osteosarcoma: outcomes in more than 2000 patients in the EURAMOS-1 (European and American Osteosarcoma Study) cohort. Eur J Cancer 109:36–50. https://doi.org/10.1016/j.ejca.2018.11.027

Spreafico M, Ieva F (2021) Functional modeling of recurrent events on time-to-event processes. Biom J 63(5):948–967. https://doi.org/10.1002/bimj.202000374

Therneau TM, Grambsch PM (2000) Modeling Survival Data: Extending the Cox Model. Springer, New York

Tsiatis AA, Davidian M (2004) Joint modelling of longitudinal and time-to-event data: an overview. Stat Sin 14(3):809–834

Ullah S, Finch CF (2013) Applications of functional data analysis: a systematic review. BMC Med Res Methodol 13(43). https://doi.org/10.1186/1471-2288-13-43

Williamson MA, Snyder LM, Wallach JB (2015) Wallach's interpretation of diagnostic tests: pathways to arriving at a clinical diagnosis, 10th edn. Wolters Kluwer Health, Philadelphia

Yao F, Müller H, Wang J (2005) Functional data analysis for sparse longitudinal data. J Am Stat Assoc 100(470):577–590. https://doi.org/10.1198/016214504000001745

## Authors and Affiliations

**Marta Spreafico**[1,2,3] ⬦ · **Francesca Ieva**[1,3,4] · **Marta Fiocco**[2,5,6]

✉ Marta Spreafico
marta.spreafico@polimi.it

Francesca Ieva
francesca.ieva@polimi.it

Marta Fiocco
m.fiocco@math.leidenuniv.nl

[1]    MOX - Laboratory for Modeling and Scientific Computing, Department of Mathematics, 20133 Politecnico di Milano, Milan, Italy

[2]    Mathematical Institute, Leiden University, Leiden, The Netherlands

[3]    CHRP - National Center for Healthcare Research and Pharmacoepidemiology, University of Milano-Bicocca, 20126 Milan, Italy

[4]    CHDS - Centre for Health Data Science, Human Technopole, 20157 Milan, Italy

[5]    Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands

[6]    Trial and Data Center, Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands