Received: 28 February 2017 | Revised: 11 January 2018 | Accepted: 27 January 2018

DOI: 10.1002/sim.7640

RESEARCH ARTICLE

WILEY Statistics in Medicine

Multivariate joint frailty model for the analysis of nonlinear tumor kinetics and dynamic predictions of death

Agnieszka Król[1] | Christophe Tournigand [2] | Stefan Michiels[3] | Virginie Rondeau[1]

[1]INSERM U1219, Biostatistics team, University of Bordeaux, Bordeaux, France

[2]Hôpital Henri Mondor, Créteil, France

[3]Service de Biostatistique et d'Epidémiologie, Gustave Roussy, University Paris-Saclay, University Paris-Sud, CESP, INSERM U1018, Villejuif, France

Correspondence
Agnieszka Król, INSERM U1219, Biostatistics team, 146 rue Leo Saignat, Bordeaux Cedex 33076, France.
Email: krol@lunenfeld.ca.

The Response Evaluation Criteria in Solid Tumors are used as standard guidelines for the clinical evaluation of cancer treatments. The assessment is based on the anatomical tumor burden: change in size of target lesions and evolution of nontarget lesions (NTL). Despite unquestionable advantages of this standard tool, Response Evaluation Criteria in Solid Tumors are subject to some limitations such as categorization of continuous tumor size or negligence of its longitudinal trajectory. In particular, it is of interest to capture its nonlinear shape and model it simultaneously with recurrent progressions of NTL and overall survival. We propose a multivariate nonlinear mechanistic joint frailty model for longitudinal data, recurrent events, and a terminal event. In the model, the tumor size trajectory is described using an ordinary differential equation that accounts for the natural growth and treatment-induced decline. We perform a simulation study to validate the method and apply the model to a phase III clinical trial in colorectal cancer. In the results of the analysis, we determine on which component, tumor size, NTL, or death the treatment acts mostly and perform dynamic predictions of death. We compare the model with other models that consider parametric functions or splines for the tumor size trajectory in terms of goodness of fit and predictive accuracy.

KEYWORDS

joint modeling, longitudinal data, ordinary differential equation, survival analysis, tumor measurement

1 | INTRODUCTION

The increasing number of clinical trials for experimental cancer treatments has led to the creation of standard assessment tools for an endpoint definition. The most common criteria for tumor response are Response Evaluation Criteria in Solid Tumors (RECIST),[1] which are used in phase III clinical trials to determine the progression-free survival (PFS) endpoint for overall survival (OS). A progression is defined as at least 20% increase in the sum of the longest diameters (SLD) of target lesions over the smallest SLD recorder and/or detection of new lesions (NL) and/or the unequivocal progression of nontarget lesions (NTL). The PFS is widely used for drug approval despite its limitations. In particular, continuous tumor size (SLD) monitored in time is used as a categorical variable to define a patient's progression status. Statistical models for longitudinal data can be used to analyze the evolution of a biomarker according to the biological comportment of its trajectory. For instance, in advanced cancer, the initial decrease in tumor size and further increase may be observed due to a therapeutic failure.[2,3] During follow-up, measurements of small values of the biomarker can be subject to left censoring, due to detection limits.[4]

Joint modeling of longitudinal and survival data enables investigation of mutual associations between the processes, informative censoring for the biomarker and unbiased estimation of the treatment effect on survival compared with univariate analysis or a 2-stage approach.[5,6] Usually, a biomarker is analyzed with a linear mixed model; more flexible trajectory can be obtained using different forms of the parametric approach. For example, Proust-Lima et al[3] developed a 2-phase linear mixed-effects model for prostate-specific antigen. The longitudinal trajectory of a biomarker can also be approximated using splines, eg, B-splines as applied by Andrinopoulou et al[7] in the context of the joint modeling of patients with valve abnormalities. More flexible and sophisticated analyses are provided by approaches that assume mechanistic models for biomarker dynamics. These models could provide more accurate modeling of the biological process and account at the same time for the heterogeneity in the data and prognostic factors. Joint models for longitudinal and survival data using ordinary differential equations (ODEs) have been applied, eg, in the context of HIV[8] and prostate cancer.[9]

Mathematical modeling of tumor growth under a treatment is becoming an important tool in evaluating anticancer treatments in population analyses.[10] Ordinary differential equations provide a wide range of potential descriptions of changes in tumor size using a function of net growth, ie, difference between the growth and natural death, that can account for drug-induced decay or the emergence of resistance to treatment. For instance, Wang et al[11] proposed a tumor size model with exponential decay and linear growth for non–small cell lung cancer. Claret et al[2] developed a model for colorectal cancer that was expressed with ODE and assumed exponential growth and decay as well as time-dependent resistance to treatment. A more complex model, proposed by Ribba et al,[12] was formulated as a system of ODEs that considered the response of different types of tumor tissues to treatment. Finally, if an analytical solution of ODE is available, the data can be analyzed using a nonlinear mixed-effects model.

It is of great interest to model tumor size jointly with the detection of NL and NTL progressions in order to use the information that contributes to the progression, as defined by RECIST. Finally, to evaluate their effects on OS, a joint model with a survival part for death can be used. In our previous work, we proposed a joint model for the left-censored SLD analyzed within a framework of a linear mixed-effects model, recurrent events, and a terminal event and applied it to a clinical trial for colorectal cancer.[4] Here, we propose a multivariate joint frailty model for longitudinal data represented by a solution to ODE, recurrent events, and a terminal event. We perform a simulation study in which we compare this model to a joint model with a 2-phase linear mixed-effects model for the SLD and to a joint model applying B-splines for the SLD trajectory in terms of goodness of fit and predictive accuracy. All models are applied to a phase III randomized clinical trial for advanced colorectal cancer.[13]

This work is structured as follows. In Section 2, we present the model and the method of estimation as well as the methodology for the dynamic predictions and measures of predictive accuracy. Simulations for validation of the proposed method with comparison to alternatives are presented in Section 3. We apply the model to the colorectal dataset in Section 4 and conclude the paper in the discussion in Section 5.

2 | MULTIVARIATE JOINT FRAILTY MODEL

We define a multivariate joint frailty model for 3 types of data: longitudinal biomarker expressed by an ODE, recurrent events, and a terminal event associated with each other by random effects. The definition of the model in sections 2.1 to 2.2 is adapted to the framework of cancer data, but it can be easily generalized to other applications.

2.1 | Tumor-size evolution submodel

Let $y_i(t)$ denote the tumor size, ie, the SLD of target lesions, of individual i ($i = 1, \dots, N$) at time t. We describe the dynamics of $y_i(t)$ over time accounting for natural exponential net growth, $K_{G,i}$, and the anticancer treatment effect with the presence of resistance to drug developed with time, $K_{D,i}(t)$.[2] We adjust the model on time-independent prognostic factors $\boldsymbol{x}_i$ and represent the unobserved interindividual heterogeneity via random effects $\boldsymbol{b}_i$. Using an ODE model for a population, we express the tumor size dynamics for subject i as

$$\frac{dy_i(t)}{dt} = K_{G,i}y_i(t) - d_i(t)K_{D,i}(t)y_i(t), \qquad y_i(0) = y_{0,i}, \tag{1}$$

where $K_{G,i}$ is the tumor growth rate; $d_i(t)$ is the drug concentration (eg, dose) at time t (for all t, $d_i(t) > 0$); and $K_{D,i}$ is the drug-induced tumor decline rate that includes the effect of resistance to the drug. We propose to apply log-transformed parameters with normal random effects: $\log(y_{0,i}) = y_0 + b_{y_0,i}$, $\log(K_{G,i}) = K_{G,0} + b_{G,i} + \boldsymbol{x}_{G,i}^\top \boldsymbol{\beta}_G$, and $\log(K_{D,i}(t)) = K_{D,0} + b_{D,i} - t \times \exp(\lambda + b_{\lambda,i}) + \boldsymbol{x}_{D,i}^\top \boldsymbol{\beta}_D$, with λ being the rate of exponential tumor decay change with time due to development of resistance to drug. The random effects $\boldsymbol{b}_i^\top = (b_{y_0,i}, b_{G,i}, b_{D,i}, b_{\lambda,i})^\top$ follow a normal distribution with mean 0 and diagonal covariance matrix $\boldsymbol{B}_1$ with elements $\sigma^2_{b_{y_0}}$, $\sigma^2_{b_G}$, $\sigma^2_{b_D}$, and $\sigma^2_{b_\lambda}$. Two sets of prognostic factors are considered for each part of tumor dynamics, $\boldsymbol{x}_{G,i}$ for the tumor growth and $\boldsymbol{x}_{D,i}$ for tumor decline.

In practice, the tumor size measurements y_{ij} are observed at discrete time points t_{ik} ($k = 1, \ldots, n_i$) and transformed for the analysis, $f(y_{ij}) = y^*_{ij}$. Because of the fact that the measurements were taken with a possible error, we add a gaussian random error ε_{ij} with mean 0 and variance σ^2_ε. The model for the observations is[14]

$$y^*_{ik} = f(y_i(t_{ik})) + \varepsilon_{ij}, \tag{2}$$

where $f(\cdot)$ is a transformation function to satisfy the normality and homoscedascity of the error terms. Tumor size observations can be subject to left-censoring when it decreases below a level of detection s of the scan. Therefore, the vector of the independent variable $\boldsymbol{Y}^*_i$ consists of the observed and left-censored measurements, below s. In this case, the expression of the likelihood can be modified to account for the left censoring.

In the case of time-independent drug concentration, $d_i(t) = d_i$, for all $t \geq 0$; for instance, if the initial drug dose is considered, the model (1) has an analytical solution. Then, the tumor dynamics can be expressed as a nonlinear model:

$$y_i(t_{ik}) = \exp\left[y_0 + b_{y_0,i} + t_{ik} \times \mathrm{e}^{K_{G,0} + b_{G,i} + \boldsymbol{x}_{G,i}^\top \boldsymbol{\beta}_G} + d_i \mathrm{e}^{K_{D,0} + b_{D,i} - (\lambda + b_{\lambda,i}) + \boldsymbol{x}_{D,i}^\top \boldsymbol{\beta}_D} \times \left(\exp\left(\mathrm{e}^{\lambda + b_{\lambda,i}} t_{ik}\right) - 1\right)\right]. \tag{3}$$

Then, this expression for tumor size can be directly used in the mixed model (2). In the following, we will assume this nonlinear mixed-effects model for the analysis of tumor size dynamics and refer to the joint model as nonlinear mechanistic joint model or mechanistic model, for simplicity.

2.2 | Time-to-event data submodel

For subject i, we observe the time of the terminal event $T_i = \min(T^*_i, C_i)$, T^*_i is the true terminal event time and C_i the censoring time. We denote the observed recurrent times to nontarget (NT) progression $T_{ij} = \min(T^*_{ij}, C_i, T^*_i)$ with T^*_{ij} the jth time of the event, $j = 1, \ldots, r_i$. We define indicators $\delta_i = I_{\{T_i = T^*_i\}}$ and $\delta_{ij} = I_{\{T_{ij} = T^*_{ij}\}}$. Let $r_0(.)$ and $\lambda_0(.)$ denote baseline hazard functions related to the recurrent events and the terminal event, respectively. Let $\mathbf{x}_{R,i}$ and $\mathbf{x}_{T,i}$ be 2 vectors of covariates associated to the risks of the events. Let $\boldsymbol{\beta}_R$ and $\boldsymbol{\beta}_T$ be the constant effects of the covariates. Finally, a frailty v_i is a random effect that links the 2 risk functions and follows the normal distribution with mean 0 and variance σ^2_v; parameter α allows for more flexibility. The hazard functions are defined by a system of proportional hazards models:

$$\begin{cases} r_{ij}(t|\boldsymbol{u}_i) = r_0(t)\exp(v_i + \boldsymbol{x}_{R,i}^\top \boldsymbol{\beta}_R + h(y_i(t))^\top \boldsymbol{\eta}_R) & \text{(nontarget progression)} \\ \lambda_i(t|\boldsymbol{u}_i) = \lambda_0(t)\exp(\alpha v_i + \boldsymbol{x}_{T,i}^\top \boldsymbol{\beta}_T + g(y_i(t))^\top \boldsymbol{\eta}_T) & \text{(death)} \end{cases}, \tag{4}$$

where $\boldsymbol{u}_i = (\boldsymbol{b}_i, v_i)$ has the diagonal covariance matrix $\boldsymbol{B}$ with the diagonal elements from matrix $\boldsymbol{B}_1$ and σ^2_v; $h(y_i(t))$ and $g(y_i(t))$ are functions that link the hazards with the biomarker; $\boldsymbol{\eta}_T$ and $\boldsymbol{\eta}_R$ are the corresponding regression coefficients. Different link functions can be used to associate the processes.[9,15] In the following, we consider directly the random effects $\boldsymbol{b}_i$ (the whole vector or a subset), ie, individual deviations of the tumor size parameters, to represent the dependence.

2.3 | Inference

The conditional likelihood of the observation for subject i, $\Theta_i = \{\boldsymbol{Y}^*_i, \boldsymbol{T}^R_i, T_i, \boldsymbol{\delta}^R_i, \delta^T_i\}$ is based on the assumption that all the processes are independent given the random effects $\boldsymbol{u}_i$ and is given by

$$L(\Theta_i|\boldsymbol{u}_i;\boldsymbol{\xi}) = L(\boldsymbol{Y}_i^*|\boldsymbol{u}_i;\boldsymbol{\xi})L(\boldsymbol{T}_i^R,\boldsymbol{\delta}_i^R|\boldsymbol{u}_i;\boldsymbol{\xi})L(T_i,\delta_i^T|\boldsymbol{u}_i;\boldsymbol{\xi}), \tag{5}$$

where $\boldsymbol{\xi}$ is the vector of the parameters, $L(\boldsymbol{Y}_i^*|\boldsymbol{u}_i;\boldsymbol{\xi})$, $L(\boldsymbol{T}_i^R,\boldsymbol{\delta}_i^R|\boldsymbol{u}_i;\boldsymbol{\xi})$, and $L(T_i,\delta_i^T|\boldsymbol{u}_i;\boldsymbol{\xi})$ are the conditional likelihoods of $\boldsymbol{Y}_i$, $(\boldsymbol{T}_i^R,\boldsymbol{\delta}_i^R)$ $(\boldsymbol{T}_i^R=\{T_{ij}, j=1,\ldots,r_i\}, \boldsymbol{\delta}_i^R=\{\delta_{ij}, j=1,\ldots,r_i\})$ and (T_i,δ_i^T), respectively. The marginal likelihood $L(\boldsymbol{\Theta}_i;\boldsymbol{\xi})$ is obtained by integrating it over the random effects and is given by

$$\begin{aligned} L(\Theta_i;\boldsymbol{\xi}) =& \frac{1}{\left(\sqrt{2\pi\sigma_\varepsilon^2}\right)^{n_i}} \int_{\boldsymbol{u}_i} \{\lambda_i(T_i|\boldsymbol{u}_i)\}^{\delta_i} \prod_{j=1}^{r_i} \left\{r_{ij}(T_{ij}|\boldsymbol{u}_i)\right\}^{\delta_{ij}} \\ &\times \prod_{k=1}^{n_i} \left[\exp\left\{-\frac{(y_{ik}^* - f(y_i(t_{ik}))^2}{2\sigma_\varepsilon^2}\right\}\right] \prod_{j=1}^{r_i} \exp\left\{\int_{T_{i(j-1)}}^{T_{ij}} r_{ij}(t|\boldsymbol{u}_i)dt\right\} \\ &\times \exp\left\{\int_0^{T_i} \lambda_i(t|\boldsymbol{u}_i)dt\right\} \frac{|\boldsymbol{B}|^{-1/2}}{(2\pi)^{5/2}} \exp\left\{-\frac{\boldsymbol{u}_i^\top \boldsymbol{B}^{-1}\boldsymbol{u}_i}{2}\right\} d\boldsymbol{u}_i. \end{aligned} \tag{6}$$

The full marginal log-likelihood is $l(\boldsymbol{\xi}) = \sum_{i=1}^N \log L(\Theta_i;\boldsymbol{\xi})$, and it is estimated using maximum penalized likelihood estimation.[16] Penalization of the log-likelihood is introduced to provide smooth estimations of the baseline hazard functions. To allow for flexibility in the form of the baseline hazard functions, we propose to use m cubic M-splines with Q knots for their approximations.[17] The M-splines are a normalized version of B-splines and are well adapted to model hazard functions, as they are nonnegative and facilitate the integrals and derivatives in the calculation of the likelihood. Two parameters, κ_1 for the recurrent events and κ_2 for the terminal event, control the trade-off between the data fit and smoothness. The integrals of the likelihood are approximated using the pseudo-adaptive Gauss-Hermite quadrature.[18] In this rule, the quadrature points (9 points) are scaled and centered using the empirical Bayes estimates of random effects obtained from an appropriate nonlinear mixed-effects model for the longitudinal biomarker, and the integral over the frailty v_i is approximated using the nonadaptive rule with 20 points. The estimates of the parameters are obtained using the Marquardt algorithm,[19] and after convergence, the standard errors are calculated from the inverse Hessian matrix of the penalized log-likelihood.

2.4 | Dynamic predictions and predictive accuracy

Joint models are useful to calculate dynamic predictions of the terminal event in a finite time horizon. In the calculations, the information from model parameters estimates $\hat{\boldsymbol{\xi}}$ is complemented by a subject's history from repeated data (biomarker and recurrent events). We consider that values of all covariates included in the model, $\boldsymbol{x}_i$, are known. The estimated probability of the event T_i^* in an interval $[t, t+w]$, given that individual i was alive before time of prediction t and given the history and model's estimates, is given by

$$\hat{P}_i(t,t+w) = P(T_i^* \le t+w | T_i^* > t, \mathcal{R}_i(t), \mathcal{Y}_i(t), \boldsymbol{x}_i; \hat{\boldsymbol{\xi}}), \tag{7}$$

where $\mathcal{R}_i(t)$ and $\mathcal{Y}_i(t)$ are subject's i individual history of recurrent events and measurements of the biomarker, respectively, before time t. We propose to use marginal dynamic predictions, ie, conditional probabilities of the occurrence of the terminal event, integrated over the random effects distribution.[4] This approach is convenient when the interest is in predictions for new patients, for which the values of random effects are unknown. The uncertainty from the random effects is incorporated via the estimates of their variances.

Confidence intervals of the dynamic predictions can be obtained using the Monte Carlo (MC) method. The principle is to draw $\boldsymbol{\xi}_b$, $b=1,\ldots,B$ from the distribution $\mathcal{N}(\hat{\boldsymbol{\xi}}, \hat{Var}(\hat{\boldsymbol{\xi}})$ and for each $\boldsymbol{\xi}_b$ calculate $\hat{P}_i(t;t+w;\boldsymbol{\xi}_b)$. The estimate of the confidence intervals are the 2.5% and 97.5% percentiles of all B predictions.

The predictive accuracy of the model in the context of internal validation can be evaluated using appropriate statistical tools.[20,21] We propose to apply the data-based Brier score adapted for right-censored survival data using the inverse probability of the censoring weighted error method with tenfold cross-validation[22] and CVPOL_a, the estimator of the expected prognostic observed cross-entropy (EPOCE), a measure of prognostic information defined by the Kullback-Leibler distance between the true prognostic density and the density derived from the joint model.[23] For both measures, the closer their values are to zero, the better the prognostic model is. Definitions of EPOCE and Brier score are given in Appendix S2.

3 | SIMULATIONS

3.1 | Scenario and data generation

A simulation study of the proposed nonlinear mechanistic joint model was performed in order to evaluate the estimation method and to compare it with joint models with alternative modeling of the longitudinal data. Certainly, it is of interest to evaluate methods that include simplified models for tumor size trajectory, as, although misspecified in the context of biological knowledge, they can be useful in some cases due to computational facilities. They do not include biological parameters for the biomarker, but covariates for recurrent and terminal events as well as the frailty parameters can be estimated. The alternatives included a parametric approach in which the biomarker's trajectory was represented by 2 slopes of time and a semiparametric approach with B-splines for the trajectory (the models are presented in Appendix S1).

Five hundred datasets were generated using the nonlinear mechanistic joint model. In the model, the exposure to treatment was represented by a time-independent, subject-specific dose. In the proposed scenario, we assumed homogeneity of the resistance to the drug between patients, ie, no random effect for parameter λ to facilitate the calculations. The assumed model is given by

$$\begin{cases} y_{ik} = \exp\left[y_0 + b_{y_0,i} + t_{ik} \times \exp\left\{K_{G,0} + b_{G,i} + \beta_G x_{1,i}\right\} + d_i \times \exp\left\{K_{D,0} + b_{D,i} - \lambda \right.\right. \\ \qquad \left.\left. + \beta_D x_{2,i}\right\} \times \left(\exp\left(e^{\lambda} t_{ik}\right) - 1\right)\right] + \varepsilon_{ik} \\ r_{ij}(t|\boldsymbol{u}_i) = r_0(t) \quad \exp\left(v_i + \beta_R x_{1,i} + \eta_{R1} b_{y_0,i} + \eta_{R2} b_{G,i} + \eta_{R3} b_{D,i}\right) \\ \lambda_i(t|\boldsymbol{u}_i) = \lambda_0(t) \quad \exp\left(\alpha v_i + \beta_T x_{2,i} + \eta_{T1} b_{y_0,i} + \eta_{T2} b_{G,i} + \eta_{T3} b_{D,i}\right) \end{cases}.$$

The biomarker parameters were fixed so that the mean biomarker trajectory presented the initial drop and long-term regrowth, $y_0 = 1.0$ (the standard deviation of the associated random effect, $\sigma_{b_{y_0}} = 0.5$), $K_{G,0} = -1.0$ ($\sigma_{b_G} = 0.5$), $K_{D,0} = -1.0$ ($\sigma_{b_D} = 0.5$), and $\lambda = 0.8$. The data of N = 400 patients was generated with longitudinal measurements every 8 weeks with a maximum 15 measurements per subject. Dose d_i was generated from a normal distribution $\mathcal{N}(1, (0.2)^2)$. Two binary covariates $x_{1,i}$ and $x_{2,i}$ were generated from the Bernoulli distribution with $P = .5$. We set $\beta_G = 0.2$, $\beta_D = 0.1$, $\beta_R = 0.5$, and $\beta_T = 0.5$. The variance of the measurement error ε_{ik} was equal to 0.25.

The link between survival processes and the biomarker was explained via the random effects related to the biomarker initial value, growth, and decay parameters. We assumed a stronger relationship of the biomarker with the terminal event ($\eta_{Tl} = 0.4$, $l \in \{1, 2, 3\}$) than with the recurrent events ($\eta_{Rl} = 0.2$, $l \in \{1, 2, 3\}$). The variance of the frailty term was $(0.5)^2$ and $\alpha = 1.6$ so that the process of the recurrent events and the terminal event were positively associated.

For each subject, we generated times of the recurrent and terminal events. Firstly, we generated an exponential death time T_i^* assuming $\lambda_0(t) = 0.5$ and compared it to a fixed right-censoring time $C = 4.5$: if $T_i^* \leq C$, then T_i^* was considered for the terminal event and $\delta_i = 1$; otherwise, C was the time of the terminal event, and $\delta_i = 0$. Then, we generated exponential recurrent gap times R_{ij} using $r_0(t) = 0.8$. The corresponding observed calendar times were $T_{ij} = \min(T_i, C, \sum_{k=1}^{j} R_{ik})$. If $T_{ij} < T_i$, then the recurrent time was observed, and $\delta_{ij} = 1$. Generation of the recurrent times for individual i continued until $T_{ij} < T_i$ or if $j = 6$.

3.2 | Results

In the generated datasets, around 0.9 (range 0-6) recurrent events and 7 (range 1-15) repeated measurements of the biomarker per patient were observed. The results of the simulation study obtained from 500 replicates are presented in Table 1. For the proposed scenario, the estimation method provided satisfactory results with generally a small bias and coverage probability close to 95% of the expected level. We obtained precise estimations for the fixed regression coefficients and biomarker parameters. Only the standard error of the parameter related to tumor growth was less well estimated resulting in quite poor coverage probability (87.2%). Similarly, the variance of the random effect related to tumor growth was less well estimated than the other variances of the random effects. This can be explained by the relatively small number of quadrature points used in the simulations (9 points). However, the choice of the number of quadrature points is a trade-off between the parameter estimates and computational time. Finally, the link parameters were well estimated except for parameters related to the random effect of tumor growth.

Next, we compared the results of the mechanistic model with results of the same generated dataset estimated using the alternative models. For both alternative models, we found the form of the biomarker time function using the profile likelihood (a grid search aiming at maximizing the likelihood) for the first generated datasets and used the same form

TABLE 1 Results of the simulation study, 500 simulations (99.8% convergence rate), datasets of 400 individuals[a]

Parameter	Mean (SE)	ESE	Bias	CP	Parameter	Mean (SE)	ESE	Bias	CP
Fixed regression coefficients					Matrix B parameters				
$\beta_G = 0.2$	0.201 (0.06)	0.07	0.001	95.0%	$\sigma_{b_{y_0}} = 0.5$	0.499 (0.02)	0.02	0.001	94.6%
$\beta_D = 0.1$	0.096 (0.06)	0.06	0.004	96.2%	$\sigma_{b_G} = 0.5$	0.459 (0.09)	0.05	0.041	82.0%
$\beta_R = 0.5$	0.506 (0.09)	0.09	0.006	94.6%	$\sigma_{b_D} = 0.5$	0.490 (0.03)	0.03	0.010	94.0%
$\beta_T = 0.5$	0.544 (0.20)	0.16	0.044	94.2%	$\sigma_v = 0.5$	0.495 (0.05)	0.06	0.005	96.4%
Biomarker parameters					Link parameters with biomarker				
$K_{G,0} = -1.0$	-0.958 (0.09)	0.08	0.042	87.2%	$\eta_{R_1} = 0.2$	0.211 (0.10)	0.10	0.011	95.4%
$K_{D,0} = 1.0$	1.000 (0.05)	0.06	0.000	96.4%	$\eta_{R_2} = 0.2$	0.236 (0.27)	0.23	0.036	93.0%
$\lambda = 0.8$	0.775 (0.07)	0.06	0.025	91.8%	$\eta_{R_3} = 0.2$	0.208 (0.15)	0.14	0.008	94.4%
$y_0 = 1.0$	0.998 (0.02)	0.03	0.010	95.2%	$\eta_{T_1} = 0.4$	0.460 (0.20)	0.18	0.060	95.4%
					$\eta_{T_2} = 0.4$	0.538 (0.63)	0.47	0.138	96.6%
					$\eta_{T_3} = 0.4$	0.450 (0.38)	0.30	0.050	93.8%
Measurement error					Alpha parameter				
$\sigma_\epsilon = 0.5$	0.500 (0.01)	0.01	0.000	96.0%	$\alpha = 1.6$	1.789 (0.87)	0.56	0.189	87.4%

[a]The statistics are mean, mean empirical standard error (SE), mean estimated standard error (ESE), absolute bias, and coverage probability (CP).

of splines and parametric functions for the rest of the datasets. For splines, regarding rising complexity of the models with the increase in the degree and the number of knots (for each spline basis, we put a random effect), the convergence rate was poor. Among the converged models, the best one was the model obtained using quadratic B-splines with only boundary knots. Moreover, using the splines, we were not able to capture the tendency of the initial drop and long-term regrowth of the mean trajectory (Figure S1), but this can be related to the fact of using the simple form of splines without interior knots. However, given the composite structure of the models, it was not feasible to apply more complex splines. For the parametric model, we used the following functions of time: $f_1(t) = (1+t)^{-6.8}$ and $f_2(t) = t^{3.4}/(1+t)^{2.4}$. As the model parameters did not correspond to each other, we compared only the goodness of fit of the models using the approximated criterion likelihood cross-validation (LCV) proposed by Commenges et al[24] as well as the estimated population-averaged trajectory of the biomarker (Figure S1). As expected, we found the best value of the LCV (mean LCV for all the 500 estimated datasets) with the mechanistic model (mean LCV = 0.945) and the worst with the model using B-splines for time related to the biomarker (mean LCV = 1.539). For the model using 2 parametric functions for the biomarker time, we obtained the mean LCV = 1.273. The true generated mean trajectory was almost identical to the one obtained with the mechanistic model (Figure S1). Moreover, using the splines, we were not able to capture the tendency of the initial drop and long-term regrowth of the mean trajectory, but this can be related to the fact of using the simple form of splines without interior knots. However, given the complex structure of the models, it was not feasible to apply more accurate structure of splines.

We also evaluated the ability of the 2 misspecified models to estimate parameters related to the recurrent and terminal events. We found that both models estimated well the effect of the covariates related to the risk of recurrent events, but only the parametric model estimated also well the effect of the covariates influencing the risk of the terminal event, the variance of the frailty, and the power α. Finally, the approach with splines gave a poor level of model convergence with a convergence rate of 75% compared 100% and 99% with the mechanistic and parametric models, respectively. To sum up, in the proposed scenario of datasets generated using the mechanistic model, the parametric model with 2 functions of time performed better than the model with B-splines as an alternative model.

4 | APPLICATION TO A COLORECTAL CANCER DATASET

Colorectal cancer is one of the major forms of cancer with nearly 1.4 million new cases diagnosed in 2012. Developed countries are particularly affected by this public health problem, colorectal cancer being the second most common cancer leading to death in Europe.[25] Despite tremendous improvement in treatment, many patients still develop metastases. Therefore, there is a great need to develop effective chemotherapy to prolong OS and improve quality of life. In phase III clinical trials, the RECIST criteria constitute the major tool to define health status in terms of PFS. This study sought to apply directly the longitudinal data of tumor size and NT progressions to predict survival. We applied the mechanistic joint

frailty model to analyze the repeated data with the terminal event and used dynamic predictions to provide conclusions on the utility of this information in terms of survival. We compared the performance of the model with alternative models with simplified (parametric or splines) tumor size trajectory and the impact on OS. Finally, we analyzed the impact of prognostic factors on each component.

4.1 | GERCOR study

The randomized GERCOR study included 220 patients with metastatic colorectal cancer[13] to investigate 2 strategies: FOLFIRI followed by FOLFOX6 (arm A) and the reverse sequence (arm B). The patients were randomly assigned from 12/1997 to 9/1999 and the cutoff date for OS was August 30, 2002. Among the included patients, 177 (80%) died during the follow-up, median OS was 21.4 months (range, 17.0-25.9) in arm A and 20.6 months (range, 17.8-25.0) in arm B. Patients changed their treatment because of progression or unacceptable toxicity. If treatment needed to be changed a second time, patients received another treatment adjusted to their condition.

For the analysis, we included all the patients who had at least 1 assessment of the target lesions (212 patients). The data included information on the SLD of target lesions (1510 measurements in total, range 1-15); detection of NL; and NTL progressions. By considering the detection of NL and NTL progressions as one process of NT disease progression, we observed 217 recurrent events (1.02 per patient, range 0-7). The individual trajectories of the SLD are presented in Figure 1. In general, the trajectories were nonlinear, with an initial decrease in tumor size and further regrowth or stabilization in most patients.

The exposure to drug was represented by the individual initial daily dose, the sum of doses of drugs administered and was assumed to be time-independent in the model. The mean dose was 3289 mg/m2 (SD = 506), and for the joint modeling, its values were transformed, as is common in pharmacology, to avoid extreme values of the biomarker. We considered powers and linear scales of dose and using the profile likelihood method applied to a univariate nonlinear model with basic covariates (treatment, sex, age, and WHO performance status); we have chosen a model with Dose* = Dose/1000.

4.2 | Data analysis

For the GERCOR data, we applied the nonlinear mechanistic joint model (model 1) as well as the alternative trivariate joint models: with 2 parametric functions (model 2) and B-splines (model 3) to describe the trajectory of the biomarker. For all the approaches, we used the same strategy (described below) to identify the final models.

Using the profile likelihood, we found a transformation for the biomarker to satisfy the normality assumption (Box-Cox transformation $\text{SLD}^* = (\text{SLD}^{0.2} - 1)/0.2$), the form of the parametric functions for model 2 ($f_1(t) = \exp(-3t)$, $f_2(t) = t^{1.1}/(t+1)^{0.1}$) and the number of knots and type of B-splines for model 3 (quadratic B-splines with no interior knots). For

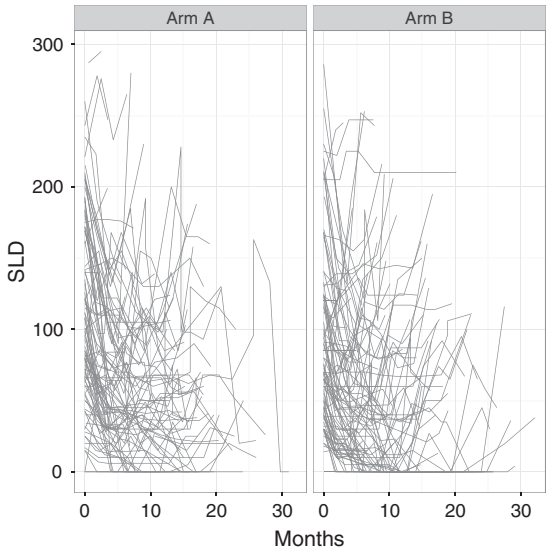

FIGURE 1 Subject-specific longitudinal trajectories of the sum of the longest diameters (SLD) in the treatment arms

the splines model, similarly to the case of the simulation study, we were limited in choosing the form of splines due to poor convergence of the mixed-effects models with splines of higher degree and bigger number of knots at the stage of the preliminary analysis.

For all the parts of the models, we used global backward elimination to select covariates using the appropriate univariate models. The covariates were retained in each process individually if the related P value was less than .1. We chose among the following variables measured at baseline: treatment (arm B/arm A), sex (female/male), age (<60/60-69/≥70), WHO performance status (PS, 0/1/2), BMI (<25/≥25), primary site (colon/rectum), surgery (no surgery/curative/palliative surgery), previous adjuvant chemotherapy (yes/no), previous adjuvant radiotherapy (yes/no), metastases (metachronous/synchronous), liver metastatic site (yes/no), lung metastatic site (yes/no), and number of metastatic sites (>1/1). In the selection process, we imposed the covariates treatment, sex, age, and WHO PS to be kept in the models. The exception was the mechanistic submodel for the biomarker, in which only treatment was included in the tumor decline part (K_D). Other covariates were not found to be significant for any kinetics parameter. Also, the covariate age was not included in the biomarker part in model 3 due to a convergence problem present in the joint analysis with this variable.

Values of some of the parameters were initialized: Parameters associated to splines of the baseline hazard functions were given values obtained from the appropriate bivariate joint frailty model for recurrent and terminal events, and parameter values of the biomarker were initialized with values from the mixed-effects models. Finally, we applied left-censored SLD since all the target lesions of certain patients decreased below the detection limit (5 mm). In the GERCOR dataset, 13.6% of the observed measurements were left-censored.

To approximate the baseline hazards, we used cubic M-splines with 7 knots and values of smoothing parameters found using cross-validation for the univariate models. Finally, for the numerical integration, we used the pseudo-adaptive Gauss-Hermite quadrature with 9 points for the part related to the biomarker (posteriori estimates of the random effects and their variances were found using the appropriate mixed-effects model with non-adaptive quadrature with 20 points) and non-adaptive quadrature with 20 points for the part related to the frailty term.

4.3 | Results

4.3.1 | Results of the mechanistic joint model

The results of the nonlinear mechanistic joint model (model 1) are presented in Tables 2 and 3. The treatment effect was not found to be significant either for the risk of NT progressions or for the risk of death. For the nontarget progressions, we found only a prognostic effect of the previous chemotherapy. Patients who had previously received a chemotherapy had an increased risk of progressions related to nontarget lesions (NL or nontarget lesions progressions). For the terminal event, a significant effect of WHO performance status at baseline confirmed that more severe disease was associated with an increased risk of death. Moreover, the presence of metachronous (consecutive) metastases was associated with a higher risk of death compared with synchronous metastases.

TABLE 2 Results of mechanistic joint frailty model for GERCOR data: estimation of covariates and tumor size parameters

Covariate	NT Progression HR (95% CI)	Death HR (95% CI)	Parameter	Tumor Size[a] Est. (95% CI)
Treatment (B/A)	1.16 (0.83; 1.62)	1.56 (0.92; 2.63)	Treatment(B/A)[b]	0.86 (0.49;1.23)***
Age (60-70/<60)	1.02 (0.70; 1.47)	1.44 (0.82; 2.55)	$K_{G,0}$	−2.05 (−2.60; −1.50)***
Age (≥70/<60)	1.14 (0.73; 1.76)	1.36 (0.68; 2.73)	$K_{D,0}$	0.01 (−0.44;0.42)
Sex (female/male)	0.85(0.60; 1.21)	1.14 (0.68;1.92)	λ	1.28 (0.67;1.89)***
WHO PS (1/0)	1.19 (0.84; 1.69)	2.05 (1.09; 3.83)*	y_0	4.46 (4.36;4.56)***
WHO PS (2/0)	1.12 (0.58; 2.15)	5.76 (2.02; 16.41)***		
Prev. chemotherapy	1.70 (1.18; 2.45)**	…		
Metachron. metast.	…	2.53 (1.41; 4.53)**		

Abbreviations: CI, confidence interval; HR, hazard ratio; SE, standard error.

[a]SLD transformed using Box-Cox transformation (0.2).

[b]Variable related to the tumor decline $K_{D,0}$.

*P value ≤ .05.

**P value ≤ .01. ***P value ≤ .001.

TABLE 3 Results of mechanistic joint frailty model for GERCOR data: estimations related to the random effects and link functions

Association Parameters	Est. (SE)		Est. (SE)		Est. (SE)
$\sigma_{b_{y_0}}$	0.63 (0.05)***	α	1.78 (0.82)*	η_{T_4}	1.69 (0.04)***
σ_{b_G}	1.55 (0.18)***	σ_ε	1.69 (0.04)***	η_{R_1}	0.19 (0.10)
σ_{b_D}	1.43 (0.16)***	η_{T_1}	0.35 (0.20)	η_{R_2}	−0.19 (0.11)
σ_{b_λ}	2.34 (0.34)***	η_{T_2}	−0.58 (0.24)*	η_{R_3}	0.21 (0.06)***
σ_v	0.44 (0.06)***	η_{T_3}	0.58 (0.16)***	η_{R_4}	0.15 (0.32)

Abbreviation: SE, standard error.
*P value ≤ .05
**P value ≤ .01.
***P value ≤ .001.

Figure S2 illustrates the baseline hazard functions obtained with the model. For death, we observed an increasing function, but for the recurrent events, the average risk was decreasing after around 1 year. Indeed, most of nontarget disease progressions occurred before in the first year of treatment.

The parameter of the biomarker kinetics, $K_{D,0}$, was not significantly different from 0. However, we found a significant effect of treatment arm on the decline in tumor size, so we concluded that the drug induced greater tumor decline in patients who received FOLFOX6 and then FOLFIRI (treatment A) than those who received treatment B (0.86, SE = 0.19). The interpretation of the treatment effect is not straightforward. In particular, after 1 year of treatment, the mean SLD was estimated at 40.2 mm in arm B compared with 11.9 mm in arm A (after retransforming SLD*). To ease the interpretation, the tumor size trajectories of average individuals from arms A and B are presented in Figure S3.

We found a significant presence of the effect of the drug resistance ($\hat{\lambda} = 1.28$, SE = 0.31). The estimates of the random effects included in the model indicated the presence of considerable heterogeneity among the patients. The obtained values resulted in capturing the tendency of the initial decline and the long-term regrowth at the population level (Figure S3) even though the observed decline was not pronounced.

The analyzed processes were significantly associated with each other via the random effects (Table 3). Therefore, the kinetics of SLD significantly influenced the risk of developing nontarget progression and the risk of death. We found that the higher deviation from the average drug-induced tumor decline, the higher the risk of nontarget progression ($\hat{\eta}_{R_3} = 0.21$, SE = 0.06) and the risk of death ($\hat{\eta}_{T_3} = 0.58$, SE = 0.16). Individuals that have a tumor that does not reduce in a standard way have more chances to have a progressive disease or to die. We also observed a strong and positive relationship between the risk of NT progressions and death ($\hat{\sigma}_v^2 = 0.44$, SE = 0.06, the flexibility parameter $\hat{\alpha} = 1.78$, SE = 0.82). NT progressions and tumor size were associated significantly via the random effect related to tumor size decline b_D and the terminal event and the tumor size via the random effects of the tumor growth, b_G; tumor decline, b_D; and drug resistance, b_λ.

4.3.2 | Dynamic predictions

Using the estimated model, we calculated the probability of death for 4 typical patients with similar characteristics (treatment arm B, age 60-70, male, WHO PS 1, no previous chemotherapy, synchronous metastases) in 2 scenarios of their disease history. In the first scenario, 2 patients were different in terms of the history of tumor size measurements (patient 1 had an initial decrease in SLD values, but after 0.7 year, the tumor size increased; patient 2 had a continuous decrease in tumor size), and no recurrent event was observed for neither of them. In the second scenario, we assumed the same trajectories of the SLD (initial decline and further regrowth), but for patient 3, we observed an NT progression at 0.7 year but not for patient 4. In both scenarios, the time of prediction was fixed at $t = 1$ year, and the predictions were calculated for the horizons from 0.1 to 1.5 year and changed every 0.1 year.

The results of the predictions are presented in Figure 2. For the first scenario, as expected, the predicted probabilities of death were higher for patient 1, with reincreasing values of SLD, than for patient 2, with decreasing SLD. The probability of death in the interval [1-2.6] was .34 for patient 1 and .22 for patient 2. For the second scenario, again, the patient with the worst prognosis (presence of NT progression), patient 3, had much higher predicted probabilities of death than patient 4. In this scenario, the difference between the curves was greater than in the scenario in which dissimilar SLD trajectories were assumed. However, the confidence intervals obtained with the MC method using 100 samples in both scenarios did not suggest that the prediction curves of the patients in a given scenario were significantly different from each other.

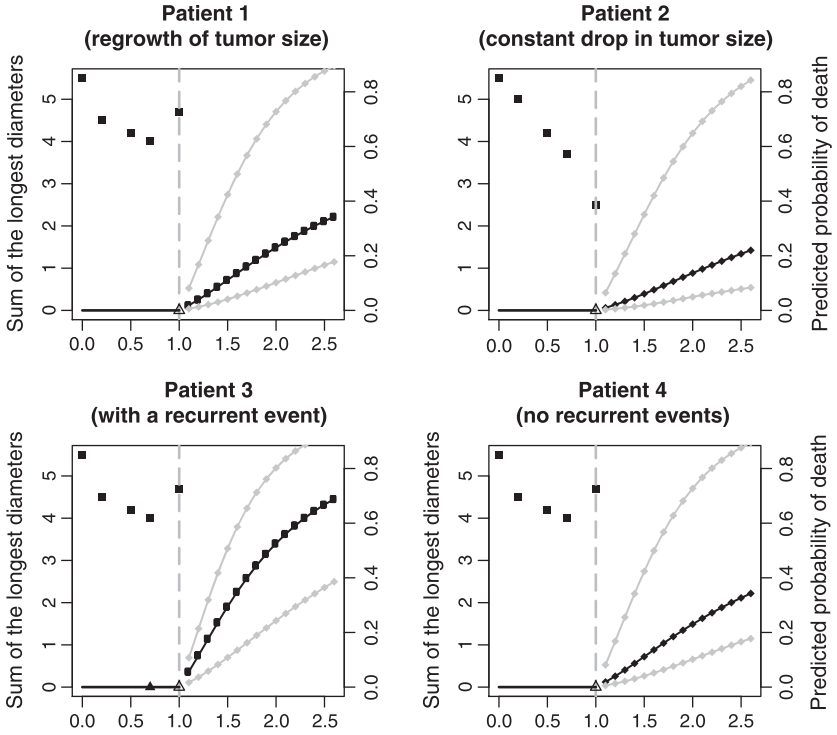

FIGURE 2 Predicted probabilities of death with time of prediction $t = 1$ and moving window from 0.1 to 1.5 years for typical patients. Top graphs represent predictions for patients different from each other in terms of tumor size measurements (black squares), and bottom graphs for patients with different history of recurrent events (black triangles). The vertical line represents the time of predictions, and the gray lines are the confidence intervals

4.3.3 | Comparison with the alternative models

Results of the models (models 2 and 3) are presented in Tables S1 and S2. We obtained similar results for the recurrent and terminal event parts as we did for the mechanistic model. We found a prognostic effect of treatment on tumor size in model 3 (−0.33, SE = 0.14); tumor size was greater in treatment A than that in treatment B. In model 2, there was a higher decrease of SLD* in treatment B in the short-term evolution (2.87, SE = 0.56) and a lower long-term decrease in treatment B than in treatment A (1.85, SE = 0.41). As for model 1, the effect of the interaction between treatment arm and time function on tumor size can be easily interpreted using average plots of SLD^* for both treatments (Figure S3). For model 2, we found a significant association between the biomarker and death and between death and recurrent events via the random effects but no significant link between recurrent events and tumor size. For model 3, we found a significant link only between recurrent events and death. Moreover, we found a prognostic effect of the treatment arm on tumor size at baseline for both models and a significant interaction with functions of time for model 2. Figure S3 shows the mean tumor-size trajectories. We found the tendency of tumor regrowth using the spline model and the mechanistic model, and the slope of the tumor size increase was steeper than that in the mechanistic model. Regarding the observed mean trajectory of the SLD, the mechanistic model resulted in the best approximation of the mean curve.

In terms of goodness of fit, the mechanistic model was the best model with the approximated LCV = 1.93 compared with 2.31 for the parametric model and 2.17 for the spline model. We compared the predictive accuracy of the models using the EPOCE and the Brier Score. The results of the EPOCE calculated for time points from 0.3 to 2.5 years are presented in Figure 3. Using the estimator $CVPOL_a$ (left graph) and the differences in $CVPOL_a$ with 95% tracking intervals (right graph), we found that in the first period (0.3-1 year), the spline model (model 3) provided the best predictive accuracy and thereafter; it was the mechanistic model (model 1) that had the best predictive accuracy. The parametric model had visibly the worst predictive accuracy throughout the time points. Regarding the tracking intervals of the differences in predictive accuracy, we found that only until 0.6 year, the spline model performed significantly better than the mechanistic model. Moreover, until 1.2 year for the spline model and for several time points throughout the considered time interval for the mechanistic model, these models were much better than the parametric model in terms of predictive accuracy.

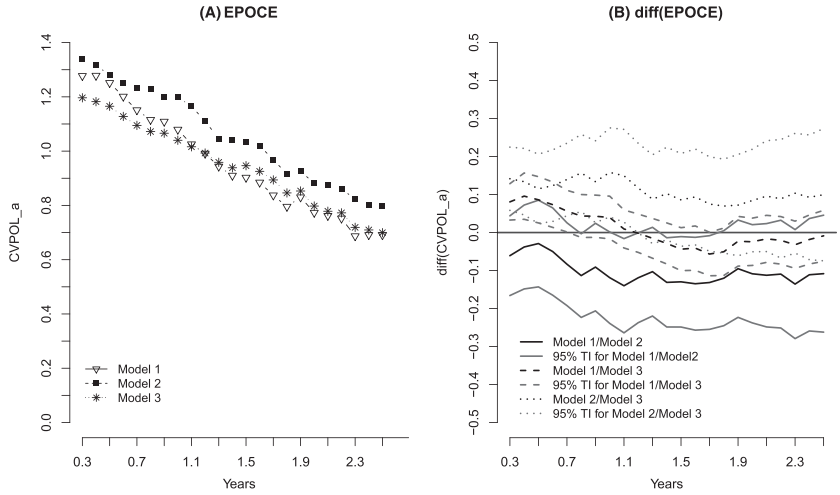

FIGURE 3 A, Estimator $extCVPOL_a$ of the expected prognostic observed cross-entropy (EPOCE) in the time window 0.3 to 2.5 years for the models applied to the GERCOR study. B, Differences in the $extCVPOL_a$ with the 95% tracking intervals between the analyzed models. Model 1, the mechanistic model; model 2, the parametric model; and model 3, the spline model

Finally, we calculated 10 times the cross-validated error of prediction using the GERCOR study for each model (internal validation). The time of prediction was set to 1 year, and we varied the window of prediction from 0.1 to 1.5 year. The results are presented in Figure S4. This time, we observed a similar Brier score for the mechanistic and parametric models for all the horizons. The spline model had similar predictive accuracy only when the window was relatively small, until $w = 0.5$. In the longer time horizons, the spline model had worse Brier score than the mechanistic and parametric models. All models had the Brier score smaller than 0.15 if the prediction window was less than half year. For larger windows, the predictive accuracy of all model was less good, but the Brier score did not exceed 0.20 for any window for the mechanistic and parametric model.

To summarize, the proposed mechanistic joint model performed better in terms of goodness of fit and had generally better or similar predictive accuracy compared with the standard joint models that used parametric functions or splines to describe the tumor size trajectory.

5 | DISCUSSION

In this work, we proposed a multivariate mechanistic joint frailty model for the analysis of a nonlinear longitudinal biomarker (possibly left censored), recurrent events, and a terminal event. The advantage of this approach compared with a standard trivariate model proposed in our previous work[4] is that the application of an ODE on the biomarker dynamics provides a better understanding of the underlying biological process. In addition, one can directly include the information on drug dose and, finally, improve the predictive accuracy for the terminal event. Indeed, for our simulation study and in the application to the GERCOR trial, we observed better performance in terms of goodness of fit and predictive accuracy when using the mechanistic model than when estimation was based on joint models with a standard mixed-effects models for the biomarker with a specific form of its trajectory (using splines or 2 functions of time).

The presented approach could contribute to research in oncology importantly. For clinical trials of an experimental treatment, the model would be a complement to the standard criteria (RECIST) that use only summarized and categorical information on tumor size evolution and nontarget disease progression. We showed that the evolution of tumor size and recurrent events of NT progressions were significantly associated with death. Finally, we provided a statistical tool to calculate dynamic predictions of death with confidence intervals, using the history of a patient's tumor size evolution, NT progression, and baseline characteristics. We applied 2 measures of predictive accuracy, the approximated CVPOL and Brier score, and found that the proposed model was generally better or similar to the alternative models, although the Brier score was quite elevated if the prediction window was large potentially due to the small sample size.

We made some important assumptions. Firstly, we assumed that the change in tumor size can be described by linear natural growth and exponential decay with the effect of resistance to the drug. This approach is similar to those proposed by Claret et al[2] and Wang et al[11] and applied to external studies on colorectal cancer[26] and other types of cancer.[27] Different

models can be applied (see review in Bender et al[28] and Ribba et al[10]), eg, a model formulated as a system of ODEs including different types of tumor tissues[12] or a model using a logistic term for tumor growth,[29] and it will be of interest to compare them in terms of performance and predictive accuracy. Secondly, in our approach, we did not consider how the nontarget progression or changes of treatment lines influenced the kinetics of tumor size of target lesions. The effect of nontarget disease process on the follow-up of SLD could be investigated by adding a further change in slope in the SLD trajectory once a nontarget progression is detected. This subject will be elaborated in future research.

We used the same mechanistic model for both lines of treatment considered as therapeutic strategies since the separation of lines in the ODE model would considerably complexify the model. In this work, we assumed a time-independent drug dose. However, estimation including a drug dose that changes in time would require a numerical solution of the ODE. Moreover, available datasets from randomized clinical trials often do not include pharmacokinetic information. However, such an approach would be interesting to apply in future research. We would like also to use our methods for a meta-analysis in order to verify the performance of the mechanistic model applied to multiple trials. Finally, it should be noted that applying such a complex model, one might encounter convergence or identifiability issues if there is a small number of observations per individual. Explicit recommendations on the required amount of information for the joint model require further investigation.

In our model, we used random effects as the structure of dependence. Compared with other commonly used link functions, current biomarker level, and its derivative, the association via random effects has the advantage of being time-independent, and thus, facilitates computations. For the analyzed data, the complexity of the model with the current tumor size as the link function in our approach resulted in convergence problems, with the stopping criterion not being achieved regardless different initialization of the parameters or the number of random effects for the biomarker. Recently, a more sophisticated dependence has been proposed by Mauff et al,[30] a cumulative effect of the biomarker weighted by temporal proximity but incorporating such a complex link function could increase substantially the computational burden of our method.

We proposed predictions for the terminal event, but it can be also of interest to estimate the probability of the occurrence of the next recurrent event given the history of the biomarker, the history of nontarget disease, and given that an individual was alive at the time of prediction. This method would enable us to provide a complete information on the predictive abilities of SLD and nontarget disease process on both risks of death and progressions.

For the estimation of the mechanistic joint model, we used the approach of penalized likelihood maximized with the Marquardt algorithm. It would be of interest to perform a simulation study to compare the algorithm with other algorithms for mechanistic joint models, for instance, methods that use the Laplace approximation, implemented in NONMEM[31] software or the SAEM algorithm implemented in both Monolix software[32] and NONMEM. However, in the case of more complex model with no analytical solution, numerical schemes are required, which would demand implementation of built-in algorithms for ODEs numerical solutions in the programs for the estimation. On the other hand, our proposed method provides estimations of baseline hazard functions, the penalized approach facilitates assumptions of smooth hazards common in biology. Moreover, the standard errors are obtained directly from the Hessian matrix. In the case of SAEM, they are obtained by stochastic approximation, and in NONMEM, they are estimated from the data using asymptotic statistical theory. Finally, dynamic predictions or evaluation of predictive accuracy are not available with Monolix and NONMEM software. Recently, Desmée et al[33] proposed dynamic predictions using Hamiltonian MC algorithm implemented in Stan software[34] to find the posterior distribution of the random effects when using Monolix software for the model.

Our approach is, to our knowledge, the first to use the dynamic model for tumor growth jointly with the evolution of the nontarget disease and death. Methods already proposed for using longitudinal information from tumor size measurements to predict survival were usually using a 2-stage approach.[2,11,35] However, such an approach can lead to biased population parameters and thus to misleading results as shown in a simulation study.[36] The application of joint modeling overcomes the problem of informative censoring by death by analyzing the processes simultaneously. Therefore, we believe that the proposed statistical tool, together with the methods for evaluating predictive accuracy, will be useful for future assessments of treatment effects and as a dynamic tool for assessing the prognosis.

6 | SOFTWARE

All the models were estimated using the extensions of the R package frailtypack.[37] The mechanistic joint frailty model that uses the analytical solution (3) to the ODE is available in the package on CRAN since version 2.12.0.

ACKNOWLEDGMENTS

The authors thank the reviewers for their insightful comments and suggestions. The authors thank GERCOR for providing the data, in particular Prof. Aimery de Gramont. The authors gratefully acknowledge EHESP and OpenHealth Institute for funding the research of A. Król. Computer time was provided by the MCIA (Mésocentre de Calcul Intensif Aquitain) computing facilities at the Université de Bordeaux and the Université de Pau et des Pays de l'Adour.

ORCID

Agnieszka Król http://orcid.org/0000-0002-9405-3722

REFERENCES

1. Eisenhauer E, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45(2):228-247.
2. Claret L, Girard P, Hoff P, et al. Model-based prediction of phase III overall survival in colorectal cancer on the basis of phase II tumor dynamics. *J Clin Oncol*. 2009;27(25):4103-4108.
3. Proust-Lima C, Taylor J, Williams S, et al. Determinants of change in prostate-specific antigen over time and its association with recurrence after external beam radiation therapy for prostate cancer in five large cohorts. *Int J Radiat Oncol Biol Phys*. 2008;72(3):782-791.
4. Król A, Ferrer L, Pignon JP, et al. Joint model for left-censored longitudinal data, recurrent events and terminal event: predictive abilities of tumor burden for cancer evolution with application to the FFCD 2000–05 trial. *Biometrics*. 2016;72(3):907-916.
5. Wulfsohn M, Tsiatis A. A joint model for survival and longitudinal data measured with error. *Biometrics*. 1997;53(1):330-339.
6. Ibrahim J, Chu H, Chen L. Basic concepts and methods for joint models of longitudinal and survival data. *J Clin Oncol*. 2010;28(16):2796-2801.
7. Andrinopoulou E, Rizopoulos D, Takkenberg J, Lesaffre E. Joint modeling of two longitudinal outcomes and competing risk data. *Stat Med*. 2014;33(18):3167-3178.
8. Guedj J, Thiébaut R, Commenges D. Joint modeling of the clinical progression and of the biomarkers' dynamics using a mechanistic model. *Biometrics*. 2011;67(1):59-66.
9. Desmée S, Mentré F, Veyrat-Follet C, Sébastien B, Guedj J. Using the SAEM algorithm for mechanistic joint models characterizing the relationship between nonlinear PSA kinetics and survival in prostate cancer patients. *Biometrics*. 2017;73(1):305-312.
10. Ribba B, Holford N, Magni P, et al. A review of mixed-effects models of tumor growth and effects of anticancer drug treatment used in population analysis. *CPT Pharmacometrics Syst Pharmacol*. 2014;3(5):1-10.
11. Wang Y, Sung C, Dartois C, et al. Elucidation of relationship between tumor size and survival in non-small-cell lung cancer patients can aid early decision making in clinical drug development. *Clin Pharmacol Ther*. 2009;86(2):167-174.
12. Ribba B, Kaloshi G, Peyre M, et al. A tumor growth inhibition model for low-grade glioma treated with chemotherapy or radiotherapy. *Clin Cancer Res*. 2012;18(18):5071-5080.
13. Tournigand C, André T, Achille E, et al. FOLFIRI followed by FOLFOX6 or the reverse sequence in advanced colorectal cancer: A randomized GERCOR study. *J Clin Oncol*. 2004;22(2):229-237.
14. Guedj J, Thiébaut R, Commenges D. Maximum likelihood estimation in dynamical models of HIV. *Biometrics*. 2007;63(4):1198-1206.
15. Rizopoulos D. *Joint Models for Longitudinal and Time-to-Event Data: With Applications in R*: Boca Raton, FL: CRC Press; 2012.
16. Joly P, Commenges D, Lettenneur L. A penalized likelihood approach for arbitrarily censored and truncated data: application to age-specific incidence of dementia. *Biometrics*. 1998;54(1):185-194.
17. Ramsay J. Monotone regression splines in action. *Stat Sci*. 1988;3(4):425-461.
18. Rizopoulos D. Fast fitting of joint models for longitudinal and event time data using a pseudo-adaptive Gaussian quadrature rule. *Comput Stat Data Anal*. 2012;56(3):491-501.
19. Marquardt D. An algorithm for least-squares estimation of nonlinear parameters. *SIAM J Appl Math*. 1963;11(2):431-441.
20. Proust-Lima C, Séne M, Taylor J, Jacqmin-Gadda H. Joint latent class models for longitudinal and time-to-event data: a review. *Stat Methods Med Res*. 2014;23(1):74-90.
21. Blanche P, Proust-Lima C, Loubère L, Berr C, Dartigues JF, Jacqmin-Gadda H. Quantifying and comparing dynamic predictive accuracy of joint models for longitudinal marker and time-to-event in presence of censoring and competing risks. *Biometrics*. 2015;71(1):102-113.
22. Gerds T, Schumacher M. Consistent estimation of the expected brier score in general survival models with right-censored event times. *Biom J*. 2006;48(6):1029-1040.
23. Commenges D, Liquet B, Proust-Lima C. Choice of prognostic estimators in joint models by estimating differences of expected conditional Kullback-Leibler risks. *Biometrics*. 2012;68(2):380-387.
24. Commenges D, Joly P, Gégout-Petit A, Liquet B. Choice between semi-parametric estimators of Markov and non-Markov multi-state models from coarsened observations. *Scand Stat Theory Appl*. 2007;34(1):33-52.
25. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer*. 2013;49(6):1374-1403.

26. Sharma M, Gray E, Goldberg R, Sargent D, Karrison T. Resampling the N9741 trial to compare tumor dynamic versus conventional end points in randomized phase ii trials. *J Clin Oncol*. 2015;33(1):36-41.
27. Zecchin C, Gueorguieva I, Enas N, Friberg L. Models for change in tumour size, appearance of new lesions and survival probability in patients with advanced epithelial ovarian cancer. *Br J Clin Pharmacol*. 2016;82(3):717-727.
28. Bender B, Schindler E, Friberg L. Population pharmacokinetic pharmacodynamic modelling in oncology: a tool for predicting clinical response. *Br J Clin Pharmacol*. 2015;79(1):56-71.
29. Tham L, Wang L, Soo R, et al. A pharmacodynamic model for the time course of tumor shrinkage by Gemcitabine + Carboplatin in non-small cell lung cancer patients. *Clin Cancer Res*. 2008;14(13):4213-4218.
30. Mauff K, Steyerberg E, Nijpels G, Heijden A, Rizopoulos D. Extension of the association structure in joint models to include weighted cumulative effects. *Stat Med*. 2017;36(23):3746-3759.
31. Beal SL, Scheiner LB. *NONMEM Users Guides*. San Francisco, CA: University of California; 1992.
32. Mbogning C, Bleakley K, Lavielle M, Mbogning C, Bleakley K, Lavielle M. Joint modeling of longitudinal and repeated time-to-event data using nonlinear mixed-effects models and the SAEM algorithm. *J Stat Comput Simul*. 2015;85(8):1512-1528.
33. Desmée S, Mentré F, Veyrat-Follet C, Sébastien B, Guedj J. Nonlinear joint models for individual dynamic prediction of risk of death using Hamiltonian Monte Carlo: application to metastatic prostate cancer. *BMC Med Res Methodol*. 2017;17(1):105.
34. Stan Development Team. Stan modeling language user's guide and reference manual. Version 2.8.0; 2015.
35. Claret L, Gupta M, Han K, et al. Evaluation of tumor-size response metrics to predict overall survival in Western and Chinese patients with first-line metastatic colorectal cancer. *J Clin Oncol*. 2013;31(17):2110-2114.
36. Bonate P, Suttle A. Modeling tumor growth kinetics after treatment with pazopanib or placebo in patients with renal cell carcinoma. *Cancer Chemother Pharmacol*. 2013;72(1):231-240.
37. Król A, Mauguen A, Mazroui Y, Lauent A, Michiels S, Rondeau V. Tutorial in joint modeling and prediction: a statistical software for correlated longitudinal outcomes, recurrent events and a terminal event. *J Stat Softw*. 2017;81(3):1-52.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Król A, Tournigand C, Michiels S, Rondeau V. Multivariate joint frailty model for the analysis of nonlinear tumor kinetics and dynamic predictions of death. *Statistics in Medicine*. 2018;1–14. https://doi.org/10.1002/sim.7640