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

Agnieszka Król, Christophe Tournigand, Stefan Michiels, Virginie Rondeau

S1 Alternative joint models

S1.1 Trivariate joint model with parametric longitudinal trajectory (two slopes of time).

The alternative trivariate joint model for the mechanistic joint model is a model in which the longitudinal trajectory is expressed with a linear mixed-effects model that includes random intercept and two random slopes. The biomarker $y_i^*(t_{ik})$ is expressed by:

$$y_i^*(t_{ik}) = \boldsymbol{x}_{0,i}^\top \boldsymbol{\beta}_0 + b_{0,i} + (\boldsymbol{x}_{1,i}^\top \boldsymbol{\beta}_1 + b_{1,i}) f_1(t_{ik}) + (\boldsymbol{x}_{2,i}^\top \boldsymbol{\beta}_2 + b_{2,i}) f_2(t_{ik}) + \epsilon_i(t_{ik}).$$

The individual trajectory of the biomarker $y_i^*(\cdot)$ is described by its intercept (fixed part represented by the vector of covariates $\boldsymbol{x}_{0i}$ with the first element equal to 1 for the intercepts, and random part, b_{10}), first slope $f_1(\cdot)$ that interacts with the fixed part, the covariates $\boldsymbol{x}_{0i}$, and random part, b_{1i}, and the second slope $f_2(\cdot)$, that interacts, again with the covariates $\boldsymbol{x}_{0i}$ and the random effect b_{1i}.

The random effects included in the longitudinal part are normally distributed with the covariance matrix:

$$\boldsymbol{B}_1 = \begin{pmatrix} \sigma_0^2 & \sigma_{01} & \sigma_{02} \\ \sigma_{01} & \sigma_1^2 & \sigma_{12} \\ \sigma_{02} & \sigma_{12} & \sigma_2^2 \end{pmatrix}.$$

For the sake of the trivariate model's identifiability, we assume that the random effects b_{li} ($l \in \{0, 1, 2\}$) are independent from the frailty term v_i.

The functions for the slopes can be found using the profile likelihood method. In our work they were chosen among

$$f_1(t) = \begin{cases} \exp(-\gamma t) \\ (1+t)^{\theta} \\ -\log(t+\delta) \end{cases}$$

and $f_2(t) = t^{1+\alpha}/(1+t)^{\alpha}$. These functions allow different forms of trajectories, including a commonly observed in advanced cancers, shape of initial drop and long-term regrowth under a treatment [7], [4].

In the profile likelihood method, we compared the likelihoods of models with different combinations of these functions and related parameters $\gamma, \theta, \delta, \alpha$. We applied the univariate non-linear mixed-model and for in the application part the covariates at that stage were: sex, age, treatment and WHO performance status. We started with the parameters values from the interval $[-5, 5]$ every 0.2 and if we found that the maximum value of the likelihood was inside the interval, we appropriately decreased the interval. For the simulation study, we performed this method only for the first generated data set and reused the chosen functions for all replications in order to ease the computational burden.

Model for simulation study

To compare the mechanistic and parametric joint models, for the datasets generated as described in Section 3.1 of the main manuscript, we performed estimations using the following model:

$$\begin{cases} y_{ik} = & \beta_0 + b_{0,i} + (\beta_{10} + \beta_{11}x_{1,i} + b_{1,i})f_1(t_{ik}) + (\beta_{20} + \beta_{21}x_{2,i} + b_{2,i})f_2(t_{ik}) + \epsilon_{ik} \\ r_{ij}(t|\boldsymbol{u}_i) = & r_0(t)\, \exp\left(v_i + \beta_3 x_{1,i} + \eta_{r1} b_{0,i} + \eta_{r2} b_{1,i} + \eta_{r3} b_{2,i}\right) \\ \lambda_i(t|\boldsymbol{u}_i) = & \lambda_0(t)\, \exp\left(\alpha v_i + \beta_4 x_{2,i} + \eta_{t1} b_{0,i} + \eta_{t2} b_{1,i} + \eta_{t3} b_{2,i}\right), \end{cases}$$

where functions $f_1(t) = (1+t)^{-6.8}$ and $f_2(t) = t^{3.4}/(1+t)^{2.4}$ were chosen using the profile likelihood method for the first generated dataset using the mixed-effect model. The coefficients β and the elements of matrix $\boldsymbol{B}_1$ in the longitudinal part were initialized with values 0.5. Given, that the datasets were generated using independent random effects b, we assumed for the parametric model that the non-diagonal elements of matrix $\boldsymbol{B}_1$ were 0. The rest of the parameters were initialized as in the mechanistic model, i.e. the values for the parameters related to recurrent events and a terminal event (spline basis, prognostic factors coefficients, variance of the frailty term, parameter α) were taken from the corresponding bivariate joint frailty models for recurrent events and a terminal event. The values of the coefficients of the link functions were set to the real values.

S1.2 Trivariate joint model with longitudinal trajectory approximated with B-splines.

The second alternative trivariate joint model is a model with a longitudinal trajectory approximated with B-splines. This approach allows for more flexibility for the biomarker profiles than the parametric approach. It is particularly useful when the longitudinal patterns seem to be non-linear. On the other hand, the interpretation of the results is more difficult but graphical methods have been proposed [1].

The model for the longitudinal biomarker is given by:

$$y_i^*(t_{ik}) = \boldsymbol{x}_{0,i}^\top \boldsymbol{\beta}_0 + b_{0,i} + \sum_{l=1}^{L} (\beta_{1l} + b_{1l,i}) B_l(t_{ik}) + \epsilon_i(t_{ik}),$$

where $B_l(\cdot)$ is the B-spline basis function. In this approach we do not include interactions of the functions of time with covariates as this would not be interpretable. The form of the splines, i.e. number of knots and degree of splines, is found using the profile likelihood method with the univariate model.

Model for simulation study

Using the datasets generated as described in Section 3.1 of the main manuscript, the estimations were obtained using the following model:

$$\begin{cases} y_{ik} = & \beta_0 + b_{0,i} + \sum_{l=1}^{2} (\beta_{1l} + b_{1l,i}) B_l(t_{ik}) + \epsilon_{ik} \\ r_{ij}(t|\boldsymbol{u}_i) = & r_0(t) \exp\left(v_i + \beta_3 x_{1,i} + \eta_{r1} b_{0,i} + \eta_{r2} b_{1,i} + \eta_{r3} b_{2,i}\right) \\ \lambda_i(t|\boldsymbol{u}_i) = & \lambda_0(t) \exp\left(\alpha v_i + \beta_4 x_{2,i} + \eta_{t1} b_{0,i} + \eta_{t2} b_{1,i} + \eta_{t3} b_{2,i}\right), \end{cases}$$

a model assuming quadratic B-splines for the tumor size trajectory with no interior knots. Therefore, we used three random effects related to the biomarker, one for the intercept and two for the splines. The initialization of the parameters was performed as with the parametric model.

S2 Predictive Accuracy Measures

S2.1 Brier score

A quadratic prediction error, often called the Brier score, evaluates calibration and discrimination of a prognostic model. The Brier Score represents the difference between the predictions and the observed patient's status at a given time point. For an estimated dynamic prediction $\hat{P}(t, t+w; \boldsymbol{\xi})$ and T^* time of death, the expected Brier score is defined by:

$$BS_{t+w} = \mathbb{E}(I_{\{T^* \leq t+w\}} - \hat{P}(t, t+w; \boldsymbol{\xi}))^2$$

and in a population of N_t individuals still at risk at time t it is estimated by $\hat{BS}_{t+w} = \sum_{i=1}^{N_t}(I_{\{T_i^* \leq t+w\}} - \hat{P}_i(t, t+w; \boldsymbol{\xi}))^2/N_t$. The Brier score can take values between 0 and 1 and it is said to be a proper measure as it is minimized when the predicted probabilities are equal to the true probabilities in the population.

The data-based Brier score uses the Inverse Probability of Censoring Weighting (IPCW) method based on weighting the observations according to their probability of being observed [8]. It can be written as:

$$\hat{BS}_{t+w} = \frac{1}{N_t}\sum_{i=1}^{N_t}\left(I_{\{T_i^* \leq t+w\}} - \hat{P}_i(t, t+w, \boldsymbol{\xi}))\right)^2 \hat{w}_i(t+w),$$

where the weights are:

$$\hat{w}_i(t+w) = \frac{I_{\{T_i^* \leq t+w\}}\delta_i}{\hat{G}(T_i^*)/\hat{G}(t)} + \frac{I_{\{T_i^* > t+w\}}}{\hat{G}(t+w)/\hat{G}(t)},$$

with $\hat{G}(t)$ is the Kaplan-Meier estimator of $G(t) = \mathbb{P}(C > t)$, C the censoring time. The incorporation of the estimator $\hat{G}(t)$ takes into account the condition that the patients are still at risk at the time of prediction t [5].

The Brier score is a popular measure of predictive accuracy in survival analysis but it is not straightforward to interpret. In order to illustrate the estimated Brier score, a graphical representation of the results at different time points $t + w$ can be performed.

Often, in the presentation of the results of the Brier score the uncertainty in the parameters and hence in predictions is not considered. However, confidence intervals may add value to conclusions on the models predictive accuracy and make it more relevant. For the evaluation of predictive accuracy for right-censored survival data, the Monte Carlo method or theoretical confidence regions proposed by Blanche et al., 2015 [2] can be applied. The theoretical confidence bands are constructed, as in the case of binary events, using the property of the asymptotic normality of the estimator of the data-based Brier score. The variance of the estimated Brier score is consistently estimated by the empirical variance of the influence function of the estimator [2]. However, in case of an internal validation, adjustments for the cross-validation should be taken into account. A solution could be to propose approximated cross-validation using the penalized approach that enables also the construction of confidence regions, as in the case of of prognostic cross-entropy (EPOCE) described below [3].

S2.2 Prognostic cross-entropy (EPOCE)

The expected cross-entropy is a measure of predictive accuracy that instead of using a binary status of the event occurrence in a given time window, it is based on the whole

distribution of the event time. As the Brier score, it evaluates discrimination and calibration at the same time. In the case of right-censored observations, a modified criterion, the expected observed cross-entropy (EPOCE) has been proposed. Both criteria as well as their estimators derived from information theory to describe the risk of an estimator were developed by Commenges et al., 2012 [3]. The EPOCE represents the loss function using the expected Kullback-Leibler divergence between the true prognostic density function $f^*_{T^*|\mathcal{F}(t),X,T^*>t}$ and the prognostic density derived from the model $g^{\hat{\boldsymbol{\xi}}}_{T^*|\mathcal{F}(t),X,T^*>t}$, with $\hat{\boldsymbol{\xi}}$ the ML estimator of the model parameters $\boldsymbol{\xi}$. Both densities are conditioned on the repeated biomarker values or recurrent events history $\mathcal{F}(t)$, fixed covariates X and that a particular subject is at risk of the survival event at prediction time t, $\{T^* > t\}$. The EPOCE is defined as the expectation under the true density function of the loss function conditioned on the information from the censoring. Two estimators of EPOCE were proposed: the mean prognostic observed loss (MPOL) in case of external validation and the cross-validated prognosis observed loss (CVPOL) as well as it approximation CVPOL_a for internal validation. The CVPOL_a is an approximate formula for the leave-one-out cross-validation estimator of the EPOCE. It is defined as follows:

$$\begin{aligned}\text{CVPOL}_a(g^{\hat{\boldsymbol{\xi}}}_{T^*|\mathcal{F}(t),X,T^*>t}, t) &= \text{MPOL}(g^{\hat{\boldsymbol{\xi}}}_{T^*|\mathcal{F}(t),X,T^*>t}, t) + Ntrace(\boldsymbol{H}^{-1}\boldsymbol{K}_t) \\ &= -\frac{1}{N_t}\sum_{i=1}^{N_t} F_i(\hat{\boldsymbol{\xi}}, t) + Ntrace(\boldsymbol{H}^{-1}\boldsymbol{K}_t),\end{aligned}$$

where $\boldsymbol{H}$ is the Hessian matrix of the (penalized) log-likelihood, $\boldsymbol{K}_t = \frac{1}{N_t(N-1)}\sum_{i=1}^{N} I_{\{T_i^*\geq t\}}\hat{\boldsymbol{v}}_i(t)\hat{\boldsymbol{d}}_i^\top$ with $\hat{\boldsymbol{v}}_i(t) = \frac{\partial F_i(\boldsymbol{\xi},t)}{\partial \boldsymbol{\xi}}|_{\hat{\boldsymbol{\xi}}}$ and $\hat{\boldsymbol{d}}_i = \frac{\partial l_i(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}}|_{\hat{\boldsymbol{\xi}}}$ and $F_i(\hat{\boldsymbol{\xi}}, t) = I_{\{T_i^*>t\}}\log(g^{\hat{\boldsymbol{\xi}}}_{T^*|\mathcal{F}(t),X,T^*>t})$, the individual contribution to the conditional log-likelihood.

The EPOCE estimator is a useful tool for comparison of different models in terms of their predictive accuracy as a kind of confidence intervals, i.e. 95% tracking intervals of the difference between them can be computed in order to enable a better evaluation. It is also natural to apply this method as the models are fitted using the likelihood and the EPOCE estimator is based on the log of the density [6]. It does not use a time window as the Brier score but the whole distribution of the event time until a chosen prediction time. Finally, using the CVPOL_a the evaluation is corrected for over-optimism without performing directly the procedure of the cross-validation which may be time consuming in the case of the joint modeling. On the other hand there is no specific interval of values of CVPOLa, its single value is difficult to interpret. However, the decreasing risk in time can be evaluated in comparison with different models.

S3 Results of simulation study

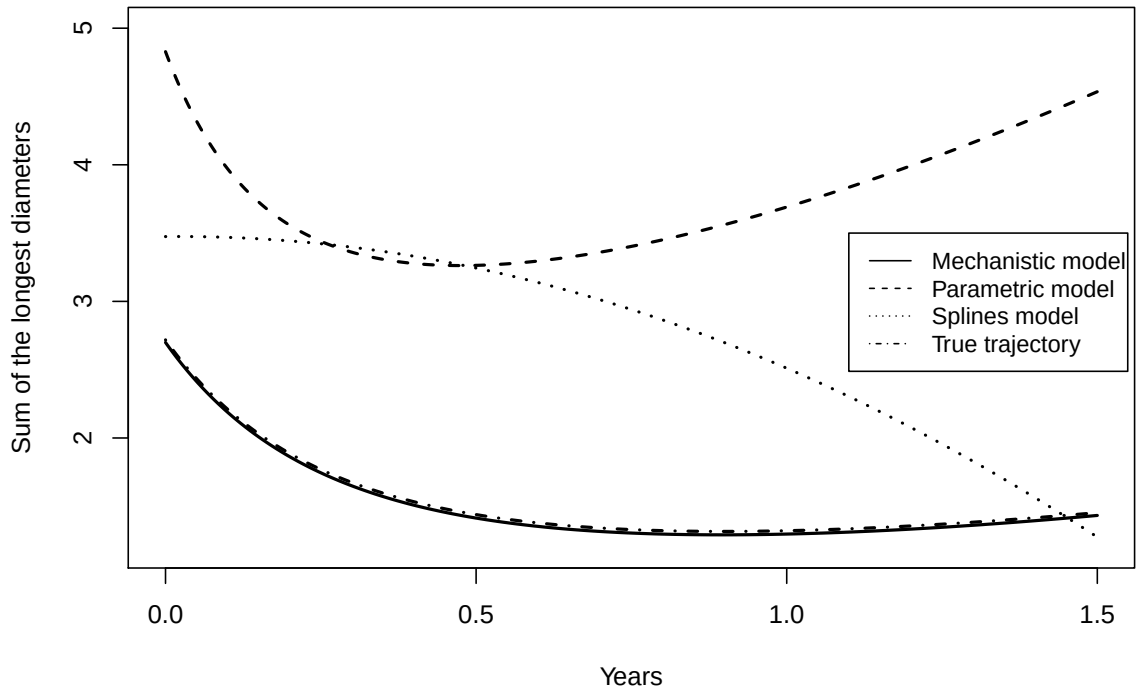

Figure S1: Population-averaged tumor size (SLD) trajectories for the three types of trivariate joint models: mechanistic, parametric (two parametric functions of time) and splines (quadratic B-splines for biomarker time) and the true generated mean trajectory.

S4 Application to the GERCOR study: Results of alternative joint models and predictive accuracy

Table S1: Results of joint model for GERCOR data with two parametric functions for time in the longitudinal submodel ($f_1(t) = \exp(-3t)$, $f_2(t) = t^{1.1}/(t+1)^{0.1}$ (estimates of covariances of random effects are not shown for the sake of clarity).

	NT Progression	Death	Tumor Size[1]
Covariate	HR (95% CI)	HR (95% CI)	Est. (SE)
Treatment(B/A)	1.22 (0.85 - 1.74)	1.13 (0.77 - 1.66)	-2.87 (0.41)***
Age (60-70/<60)	0.97 (0.65 - 1.45)	0.94 (0.62 - 1.45)	0.11 (0.41)
Age (≥70/<60)	1.01 (0.63 - 1.62)	0.88 (0.53 - 1.46)	-0.03 (0.17)
Sex (Female/Male)	0.89 (0.62 - 1.28)	1.01 (0.68 - 1.48)	-0.23 (0.13)
WHO PS (1/0)	1.27 (0.86 - 1.88)	2.52 (1.58 - 4.01)***	2.90 (0.20)***
WHO PS (2/0)	1.40 (0.72 - 2.74)	8.21 (3.78 - 17.84)***	3.58 (0.34)***
Prev. chemotherapy (Yes/No)	1.62 (1.09 - 2.40)*	-	-0.76 (0.17)***
Metachron. metast. (Yes/No)	-	1.82 (1.19 - 2.78)**	
Intercept	-	-	3.87 (0.32)***
$f_1(t)$	-	-	0.71 (0.51)
Treatment (B/A) $*f_1(t)$	-	-	2.87 (0.56)***
WHO PS (1/0) $*f_1(t)$	-	-	-2.33 (0.37)***
WHO PS (2/0) $*f_1(t)$	-	-	-2.31 (0.63)***
Metast. site (Liver/Other) $*f_1(t)$	-	-	1.65 (0.32)***
$f_2(t)$	-	-	-2.58 (0.30)***
Treatment (B/A) $*f_2(t)$	-	-	1.85 (0.41)***
Primary site (Colon/Rectum) $*f_2(t)$	-	-	-3.29 (0.32)***
Prev. radiotherapy (Yes/No) $*f_2(t)$	-	-	4.69 (0.37)***

Association Parameters	Est. (SE)		Est. (SE)
σ_{b_1}	3.84 (0.19)***	η_{T_1}	-0.29 (0.11)**
σ_{b_2}	3.12 (0.40)***	η_{T_2}	-0.60 (0.18)**
σ_{b_3}	5.83 (0.19)***	η_{T_3}	0.47 (0.07)***
σ_v	0.87 (0.06)***	η_{R_1}	-0.13 (0.08)
α	0.32 (0.38)	η_{R_2}	-0.27 (0.13)
σ_ϵ	2.02 (0.10)***	η_{R_3}	0.13 (0.03)

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

* p-value ≤ 0.05, ** p-value ≤ 0.01, *** p-value ≤ 0.001

HR - hazard ratio, CI - confidence interval, SE - standard error

Table S2: Results of trivariate joint frailty model for GERCOR data using quadratic B-splines for time in the longitudinal submodel (estimates of covariances of random effects are not shown for the sake of clarity).

	NT Progression	Death	Tumor Size[1]
Covariate	HR (95% CI)	HR (95% CI)	Est. (SE)
Treatment(B/A)	1.17 (0.84 - 1.64)	1.42 (0.85 - 2.37)	−0.33 (0.14)*
Age (60-70/<60)	1.00 (0.69 - 1.46)	1.26 (0.70 - 2.24)	-
Age (≥70/<60)	1.12 (0.72 - 1.73)	1.22 (0.65 - 2.32)	-
Sex (Female/Male)	0.87 (0.61 - 1.22)	1.06 (0.64 - 1.74)	−0.17 (0.15)
WHO PS (1/0)	1.28 (0.90 - 1.83)	2.74 (1.49 - 5.03)***	1.02 (0.15)***
WHO PS (2/0)	1.30 (0.67 - 2.53)	8.79 (3.27 - 23.61)***	1.83 (0.25)***
Prev. chemotherapy (Yes/No)	1.76 (1.22 - 2.54)*	-	−0.67 (0.19)***
Metachron. metast. (Yes/No)	-	2.39 (1.28 - 4.47)**	
Intercept	-	-	5.05 (0.27)***
Metast. site (Liver/Other)	-	-	1.42 (0.20)***
Metast. site (Lung/Other)	-	-	0.66 (0.17)***
bs(time,2)1	-	-	−7.46 (0.47)***
bs(time,2)2	-	-	0.53 (0.64)

Association Parameters	Est. (SE)		Est. (SE)
σ_{b_1}	2.19 (0.14)***	η_{T_1}	−3.38 (2.23)
σ_{b_2}	4.31 (0.15)***	η_{T_2}	1.71 (1.01)
σ_{b_3}	2.58 (0.02)***	η_{T_3}	1.79 (1.27)
σ_v	0.55 (0.11)***	η_{R_1}	−0.31 (1.22)
α	2.00 (0.47)***	η_{R_2}	0.21 (0.55)
σ_ϵ	2.37 (0.20)***	η_{R_3}	0.06 (0.71)

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

* p-value ≤ 0.05, ** p-value ≤ 0.01, *** p-value ≤ 0.001

HR - hazard ratio, CI - confidence interval, SE - standard error

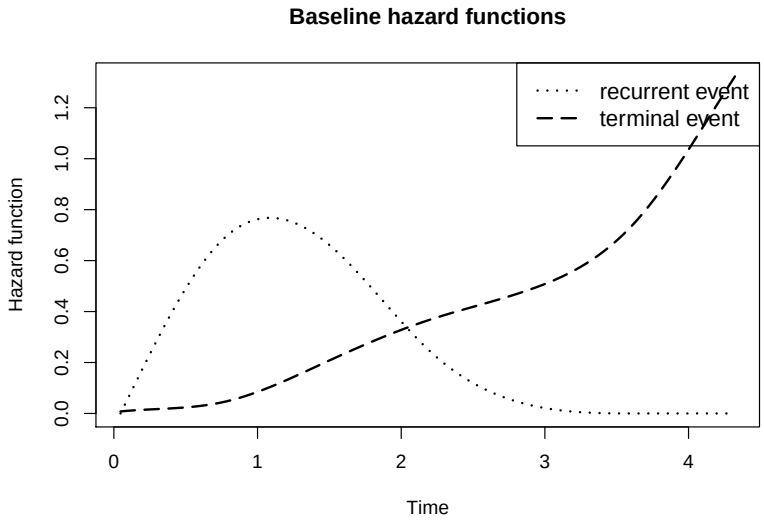

Figure S2: Estimated baseline hazard functions for recurrent and terminal events using cubic M-splines with 7 knots for the GERCOR study.

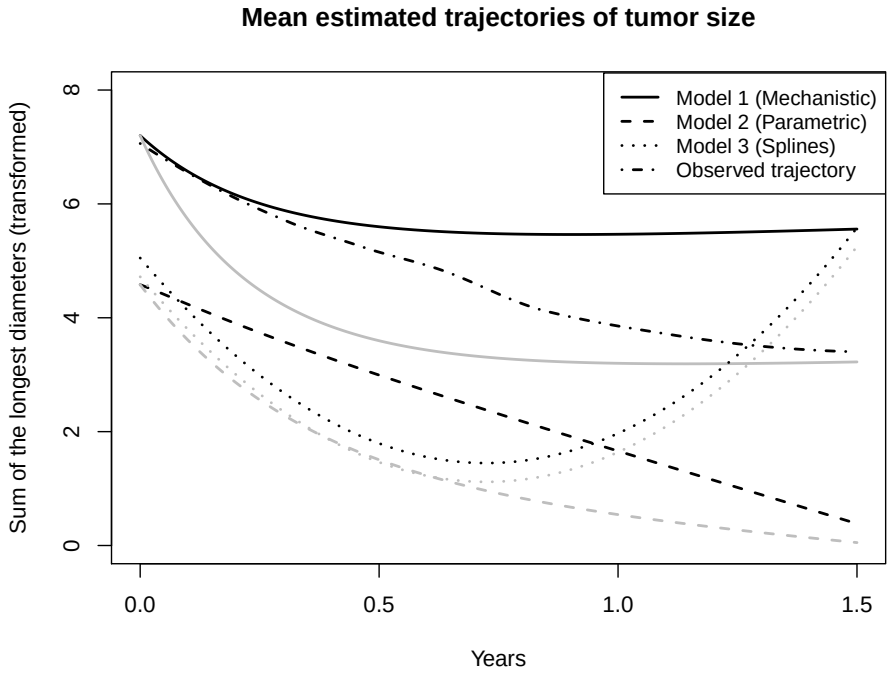

Figure S3: Population-averaged tumor size (SLD) trajectories for trivariate joint models: mechanistic (Model 1), parametric (Model 2) and splines (Model 3) and the observed mean trajectory (obtained with the `loess` function of the `R` software). Black lines correspond to patients treated with drug A and grey lines to patients treated with drug B.

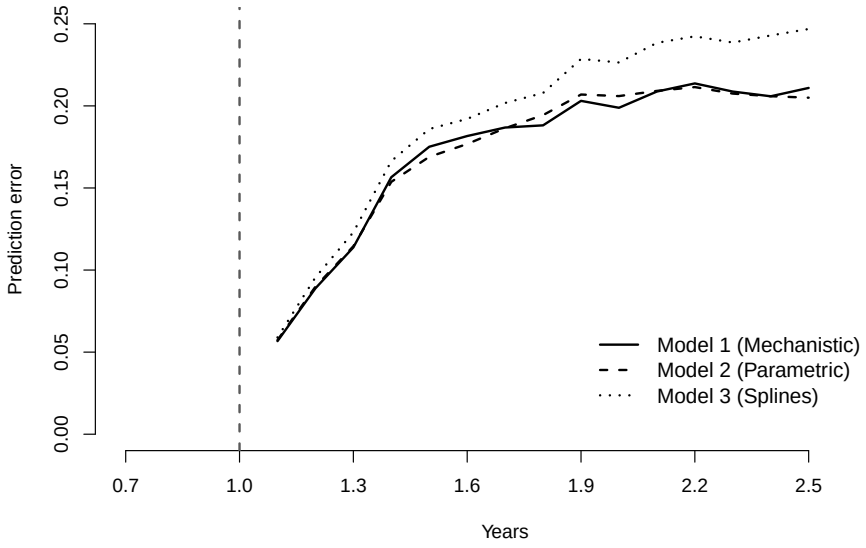

Figure S4: Error of prediction using 10-fold cross-validation with time of prediction t = 1 year and varying window w from 0.1 to 1.5 for the models applied to the GERCOR study.

References

[1] E Andrinopoulou, D Rizopoulos, J Takkenberg, and E Lesaffre. Joint modeling of two longitudinal outcomes and competing risk data. *Stat Med*, 33(18):3167–3178, 2014.

[2] P Blanche, C Proust-Lima, L Loubère, C Berr, JF Dartigues, and H Jacqmin-Gadda. Quantifying and comparing dynamic predictive accuracy of joint models for longitudinal marker and time-to-event in presence of censoring and competing risks. *Biometrics*, 71(1):102–113, 2015.

[3] D Commenges, B Liquet, and C Proust-Lima. Choice of prognostic estimators in joint models by estimating differences of expected conditional Kullback-Leibler risks. *Biometrics*, 68(2):380–387, 2012.

[4] C Ferté, S Koscielny, L Albiges, L Rocher, J Soria, R Iacovelli, Y Loriot, K Fizazi, and B Escudier. Tumor growth rate provides useful information to evaluate sorafenib and everolimus treatment in metastatic renal cell carcinoma patients: An integrated analysis of the TARGET and RECORD phase 3 trial data. *Eur Urol*, 65(4):713–720, 2013.

[5] A Mauguen, B Rachet, S Mathoulin-Pélissier, G MacGrogan, A Laurent, and V Rondeau. Dynamic prediction of risk of death using history of cancer recurrences in joint frailty models. *Stat Med*, 32(30):5366–5380, 2013.

[6] C Proust-Lima, M Séne, J Taylor, and H Jacqmin-Gadda. Joint latent class models for longitudinal and time-to-event data: A review. *Stat Methods Med Res*, 23(1):74–90, 2014.

[7] C Proust-Lima, J Taylor, S Williams, D Ankerst, N Liu, L Kestin, K Bae, and H Sandler. 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*, 72(3):782–791, 2008.

[8] R Schoop, M Schumacher, and E Graf. Measures of prediction error for survival data with longitudinal covariates. *Biom J*, 53(2):275–93, 2011.