Using the SAEM algorithm for mechanistic joint models characterizing the relationship between nonlinear PSA kinetics and survival in prostate cancer patients

Solène Desmée[1,2,*], France Mentré[1,2], Christine Veyrat-Follet[3], Bernard Sébastien[4] and Jérémie Guedj[1,2]

[1]INSERM, IAME, UMR 1137, F-75018 Paris, France

[2]Université Paris Diderot, IAME, UMR 1137, Sorbonne Paris Cité, F-75018 Paris, France

[3]Drug Disposition, Disposition Safety and Animal Research Department, Sanofi, Alfortville, France

[4]Biostatistics and Programming, Sanofi, Chilly-Mazarin, France

*_email:_ solene.desmee@inserm.fr

SUMMARY: Joint modelling is increasingly popular for investigating the relationship between longitudinal and time-to-event data. However numerical complexity often restricts this approach to linear models for the longitudinal part. Here we use a novel development of the Stochastic-Approximation Expectation Maximization algorithm that allows joint models defined by nonlinear mixed-effect models. In the context of chemotherapy in metastatic prostate cancer, we show that a variety of patterns for the Prostate Specific Antigen (PSA) kinetics can be captured by using a mechanistic model defined by nonlinear ordinary differential equations. The use of a mechanistic model predicts that biological quantities that cannot be observed, such as treatment-sensitive and treatment-resistant cells, may have a larger impact than PSA value on survival. This suggests that mechanistic joint models could constitute a relevant approach to evaluate the efficacy of treatment and to improve the prediction of survival in patients.

KEY WORDS: Joint model; Metatastic prostate cancer; Nonlinear mixed effect model; Prostate specific antigen ; SAEM algorithm; Survival.

1. Introduction

The field of joint modeling, which aims to characterize the relationship between longitudinal biomarkers and a time-to-event, has received a lot of attention from biostatisticians in the last decade (Tsiatis and Davidian, 2004; Rizopoulos et al., 2009; Wu et al., 2011). Moreover large efforts have been made to reach out beyond the academic community and packages in R or codes and macro in SAS are now available (Guo and Carlin, 2004; Rizopoulos, 2010; Garcia-Hernandez and Rizopoulos, 2015). However a major limitation of these tools is that they essentially rely on linear models for the longitudinal part (Tsiatis and Davidian, 2004; Rizopoulos, 2012; Asar et al., 2015). Although these models are often sufficient to describe sparse longitudinal data, such as those obtained in epidemiological studies, they may not be appropriate to describe the kinetics of frequently sampled markers such as those obtained in clinical trials for treatment evaluation. In this context, nonlinear mechanistic models, for instance based on ordinary differential equations (ODEs), may be necessary to characterize the dynamic changes in the longitudinal markers but the difficulty to calculate the likelihood in nonlinear mixed effect models (NLMEM) complicates considerably the use of these models. In the last years, several stochastic methods have been developed to propose efficient algorithms for inference in NLMEM (Kuhn and Lavielle, 2005; Plan et al., 2012). One of them is the Stochastic Approximation Expectation-Maximization (SAEM) algorithm (Delyon et al., 1999). As in other EM algorithms, the algorithm is an iterative process where each iteration is divided into a step where the complete likelihood conditional on observations is calculated (E-step), and a step where the complete likelihood is maximized (M-step). In addition, in the SAEM algorithm, the E-step is divided into two parts: a simulation of individual parameters using a Markov Chain Monte Carlo (MCMC) algorithm (S-step), and then a calculation of the expected likelihood using a stochastic approximation (A-step). Recently the SAEM algorithm implemented in Monolix (www.lixoft.eu) has been expanded

to the context of joint models and was shown by simulation to provide precise estimates when the longitudinal marker was defined by a NLMEM (Mbogning et al., 2015; Desmée et al., 2015).

Here we applied this approach to characterize, on real data, the relationship between the kinetics of the Prostate-Specific Antigen (PSA) and survival in metastatic Castration-Resistant Prostate Cancer (mCRPC) using data from a phase 3 clinical trial (Tannock et al., 2013). In this context of patients with advanced disease, the incidence of death is high and the PSA kinetics is closely monitored after treatment initiation to rapidly detect a breakthrough in PSA and propose rescue strategies. Thus, unlike studies where long-term PSA levels can be described by linear models (Proust-Lima et al., 2008) here, the rapid changes in PSA levels require dynamical nonlinear models. Similarly to what is done to characterize the effect of an anti-viral treatment on a pathogen (Perelson and Guedj, 2015), the PSA kinetics can be schematically viewed as a dynamic interaction between the chemotherapy and the process of production and elimination of cancer cells (Seruga, Ocana, and Tannock, 2011). This interaction can naturally be modeled using a system of nonlinear ODEs where both parameters and processes have a biological interpretation. Here, we aimed to show how a mechanistic joint approach can be used to characterize the relationship between the nonlinear kinetics of a biomarker and the time-to-death in the context of treatment evaluation.

The outline is as follows: in Section 2, we introduce the clinical data and the methods used to characterize the relationship between PSA kinetics and survival. In Section 3, we show how to construct and evaluate a joint model in the context of a mCRPC clinical trial. Lastly, in section 4, we conclude by pointing out the advantages and the limits of this approach.

2. Material and Methods

2.1 *Description of the Data*

We analyzed the data of 598 men with metastatic Castration-Resistant Prostate Cancer (mCRPC) treated with docetaxel and prednisone, the first-line reference chemotherapy, which constituted the control arm of a phase 3 clinical trial (Tannock et al., 2013). In the protocol, PSA had to be measured within 8 days before treatment initiation, every 21 days during treatment and then every 84 days after treatment. Baseline available covariates were body size, age, race and time elapsed since hormone therapy. Two patients had no PSA measurement and were not included in the analysis.

For the sake of internal validation, the dataset was randomly split into a training and a validation dataset of 400 and 196 patients, respectively. In the training dataset, 6,627 PSA measurements were collected, among which 1,385 (20.9%) were pre-treatment and 3,934 (59.4%) were on treatment. The median [minimum ; maximum] duration between the first PSA measurement and the treatment initiation was 104 days [2 ; 1,195]. The median number of measurements per patient was 15 [3 ; 57] with a median number of pre-, on- and post-treatment measurements of 4 [1 ; 14], 9 [0 ; 48] and 2 [0 ; 26], respectively. The limit of quantification (LoQ) was 0.1 ng.ml^{-1} and 165 observations (2.5%) from 21 patients (5.3%) were below LoQ. In the validation dataset, the median numbers of PSA measurements per patient was also equal to 15 and the total number of PSA measurements was 3,185. In the training dataset 286 patients deceased (71.5%), leading to a median survival [Kaplan-Meier 95% confidence interval] of 656 days [598 ; 741]. In the validation dataset, 145 patients deceased (74.0%), leading to a median survival of 598 days [547 ; 732]. Unlike it is stated otherwise, all results given below are obtained using the training dataset only.

2.2 *Mechanistic PSA Kinetic Model*

In this model (Figure 1), PSA is produced by two types of cells, namely treatment-sensitive cells (S) and treatment-resistant cells (R) (Seruga et al., 2011). In absence of treatment, sensitive and resistant cells proliferate with rates α_S and α_R, respectively, and are eliminated with a similar rate d. The total number of sensitive and resistant cells is limited by a saturation term, noted N_{max}. Mutations from S to R and from R to S occur with an identical rate, noted g. PSA is secreted by both treatment-sensitive and -resistant cells with a rate p and is cleared from the blood with a rate δ.

The associated system of ordinary differential equations (ODEs) is:

$$\begin{cases} \frac{dS}{dt} = \alpha_S(1 - \frac{S+R}{N_{max}})S + g(R - S) - dS \\ \frac{dR}{dt} = \alpha_R(1 - \frac{S+R}{N_{max}})R + g(S - R) - dR \\ \frac{dPSA}{dt} = pS + pR - \delta PSA \end{cases} \quad (1)$$

Let PSA_b, S_b and R_b be the initial values of PSA, sensitive and resistant cell counts, respectively, i.e., their values at first PSA measurement. In order to determine their values, we used the fact that by definition the treatment resistant cells are less fit to grow than the treatment-sensitive cells ($RF = \alpha_R/\alpha_S < 1$). Therefore treatment-sensitive cells should be largely predominant at baseline ($S_b \gg R_b$) and thus quasi-steady state approximations can be made to derive baseline conditions:

$$S_b = \frac{\delta}{p}\ PSA_b$$

Then one can obtain after few calculations on equations (1):

$$R_b = \frac{g}{d - RF \times (g + d)} \times \frac{\delta}{p}\ PSA_b$$

.

Time $t = 0$ indicates the beginning of treatment that has a constant and non-null effectiveness against the treatment-sensitive cells ($\varepsilon > 0$) while it has no efficacy against resistant cells. Two mechanisms of actions for docetaxel were considered (Herbst and Khuri, 2003;

Petrylak, 2005): it can either inhibit angiogenesis (i.e., decreases the cancer cells proliferation from α_S before treatment initiation to $\alpha_S \times (1 - \varepsilon)$ afterwards) or increase cell apoptosis (i.e., stimulates the cancer cells elimination from d before treatment initiation to $d \times (1 + \varepsilon)$ afterwards).

[Figure 1 about here.]

The following reparameterizations were done to improve the model identifiability. First since before treatment initiation PSA increased, the proliferation rate of sensitive cells is necessarily larger than the apoptosis rate ($\alpha_S > d$). Second after treatment initiation, no patient achieved a sustained low level of PSA, which means that the proliferation rate of resistant cells is also larger than the apoptosis rate ($\alpha_R > d$). Thus $\alpha_S > \alpha_R > d$ and we parameterized d as $d = RE \times RF \times \alpha_S$ with $0 < RE < 1$. For the sake of parameter identifiability, we fixed δ to 0.23 day^{-1}, corresponding to a PSA half-life in blood of about 3 days (Ruffion, Rebillard, and Grima, 2005) while p and g were determined by a sensitivity analysis (see Web Appendix A).

Finally, the mathematical model for PSA kinetics was defined by the vector parameter: $\boldsymbol{\mu} = (\alpha_S,\ RF,\ RE,\ \varepsilon,\ PSA_b,\ N_{max})$.

2.3 *Statistical Model for PSA Measurements*

Nonlinear mixed-effect models (NLMEM) were used to analyze all the longitudinal PSA measurements (before, during and after treatment).

Let N be the number of patients and $\boldsymbol{y}_i = (y_{i1},\ \ldots, y_{in_i})$ the vector of observations in patient i, where y_{ij} is the observed Naperian logarithm of PSA+1 for the patient i, $i = 1, \ldots, N$, at time t_{ij}, $j = 1, \ldots, n_i$. A constant error model is assumed on the logarithm of PSA+1:

$$y_{ij} = \log(PSA(t_{ij}, \boldsymbol{\psi}_i) + 1) + e_{ij} \tag{2}$$

where $\boldsymbol{\psi}_i$ is the vector of the individual parameters, $PSA(t_{ij}, \boldsymbol{\psi}_i)$ is given by the system of ODEs (1) and e_{ij} is the residual Gaussian error of mean 0 and variance σ^2. $\boldsymbol{\psi}_i$ is decomposed as a vector of fixed effects $\boldsymbol{\mu}$ representing median parameters of the population and random effects $\boldsymbol{\eta}_i$ specific for each individual. It is assumed that $\boldsymbol{\eta}_i \sim \mathcal{N}(0, \boldsymbol{\Omega}^2)$ with $\boldsymbol{\Omega} = diag(\omega_{\alpha_S}, \omega_{RF}, \omega_{RE}, \omega_{\varepsilon}, \omega_{PSA_b}, \omega_{N_{max}})$. We assumed log-normal distribution for α_S, PSA_b and N_{max} and logit-normal distribution for RF and RE. Regarding ε, we used a logit-normal distribution when an effect in blocking cell proliferation was evaluated and a log-normal distribution when an effect of treatment in enhancing cell death was evaluated (see section 2.2).

2.4 *Characterization of the Relationship between PSA Kinetics and Survival*

Let X_i and C_i denote the time-to-death and the censoring time, respectively, for patient i. $T_i = \min(X_i, C_i)$ and $\delta_i = \mathbf{1}_{\{X_i \leqslant C_i\}}$ are observed. The time-to-death was modeled using a parametric risk-proportional model assuming a Weibull function for the baseline hazard function: for $t \geqslant 0$,

$$h_i(t|\mathbf{PSA}(t, \boldsymbol{\psi}_i)) = h_0(t) \exp(\boldsymbol{\gamma}^T \boldsymbol{w}_i + \boldsymbol{\beta}^T f(t, \boldsymbol{\psi}_i)) \tag{3}$$

where $\mathbf{PSA}(t, \boldsymbol{\psi}_i) = \{PSA(s, \boldsymbol{\psi}_i); 0 \leqslant s < t\}$ denotes the history of the true unobserved longitudinal process up to t, h_0 is the Weibull baseline hazard function $h_0(t) = \frac{k}{\lambda}(\frac{t}{\lambda})^{k-1}$, $\boldsymbol{\gamma}$ is the vector of coefficients associated with the vector of baseline covariates $\boldsymbol{w}_i$ and $\boldsymbol{\beta}$ is the vector of coefficients associated with the ODE model outputs $f(t, \boldsymbol{\psi}_i)$ (Eq. 1).

The following models for the relationship between survival and PSA kinetics were considered:

- No link: $f(t, \boldsymbol{\psi}_i) = 0$,
- Initial PSA: $f(t, \boldsymbol{\psi}_i) = \log(PSA(0, \boldsymbol{\psi}_i) + 1)$,
- Current PSA: $f(t, \boldsymbol{\psi}_i) = \log(PSA(t, \boldsymbol{\psi}_i) + 1)$,

- Current PSA slope: $f(t, \boldsymbol{\psi}_i) = \frac{d \log(PSA(t,\boldsymbol{\psi}_i)+1)}{dt}$,
- Area under PSA: $f(t, \boldsymbol{\psi}_i) = \int_0^t \log(PSA(u, \boldsymbol{\psi}_i) + 1)du$,
- Current S: $f(t, \boldsymbol{\psi}_i) = \log(S(t, \boldsymbol{\psi}_i))$,
- Current R: $f(t, \boldsymbol{\psi}_i) = \log(R(t, \boldsymbol{\psi}_i))$,
- Current S and R: $f(t, \boldsymbol{\psi}_i) = (\log(S(t, \boldsymbol{\psi}_i)), \log(R(t, \boldsymbol{\psi}_i)))$. In this model $\boldsymbol{\beta}$ has two components, $\boldsymbol{\beta} = (\beta, \beta')$.

Of note the last three models exploit the mechanistic nature of the model as they relate to quantities that are not observed.

2.5 *Maximum Likelihood Estimation*

The log-likelihood for subject i is given by:

$$LL_i(\boldsymbol{\theta}) = \log \int p(\boldsymbol{y}_i|\boldsymbol{\eta}_i; \boldsymbol{\theta}) p(T_i, \delta_i|\boldsymbol{\eta}_i; \boldsymbol{\theta}) p(\boldsymbol{\eta}_i; \boldsymbol{\theta}) d\boldsymbol{\eta}_i \tag{4}$$

where $\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\Omega}, \sigma, \lambda, k, \gamma, \boldsymbol{\beta})$ is the vector of parameters to estimate, $p(\boldsymbol{y}_i|\boldsymbol{\eta}_i; \boldsymbol{\theta})$ is the probability density function of the longitudinal observations conditionally on the random effects $\boldsymbol{\eta}_i$, $p(T_i, \delta_i|\boldsymbol{\eta}_i; \boldsymbol{\theta}) = h_i(T_i|\boldsymbol{PSA}(T_i, \boldsymbol{\psi}_i); \boldsymbol{\theta})^{\delta_i} S_i(T_i|\boldsymbol{PSA}(T_i, \boldsymbol{\psi}_i); \boldsymbol{\theta})$ is the likelihood of the survival part, with $S_i(t|\boldsymbol{PSA}(t, \boldsymbol{\psi}_i); \boldsymbol{\theta}) = \exp(-\int_0^t h_i(s|\boldsymbol{PSA}(s, \boldsymbol{\psi}_i); \boldsymbol{\theta})ds)$ and $p(\boldsymbol{\eta}_i; \boldsymbol{\theta})$ is the probability density function of the random effects.

The likelihood was maximized using the SAEM algorithm (Delyon et al., 1999; Kuhn and Lavielle, 2005) implemented in Monolix version 4.3.2. The likelihood was estimated by Importance Sampling with a Monte-Carlo size of 200,000 to ensure efficient precision. SAEM was run using one chain after checking that using 3 chains gave similar results (Vigan et al., 2014). All calculations were performed in a i7 64bits 3.33 GHz.

2.6 *Model Selection and Evaluation*

We first selected the model for the longitudinal data using only longitudinal data. For each couple of parameters (p, g) explored in the sensitivity analysis (Web Appendix A), the two

possible mechanisms of action of treatment were tested and only the model giving the lowest Bayesian Information Criterion (BIC) was retained.

Next the selection of the joint model was based on the BIC (Park and Qiu, 2014). Of note baseline covariates were tested in the joint model only if they were significant in univariate analysis using a Weibull survival model and a likelihood ratio test with a significance level of 20%. Then, the model evaluation relied on the analysis of the model individual predictions and residuals in the training dataset, as well as the analysis of the survival predictions in the validation dataset. The residuals for the longitudinal part were assessed using the Individual Weighted Residuals, noted IWRES, defined by $IWRES_{ij} = \frac{y_{ij} - \log(PSA(t_{ij}, \hat{\boldsymbol{\psi}}_i))}{\hat{\sigma}}$ where $\hat{\sigma}$ is the estimated standard deviation of the residual error and $\hat{\boldsymbol{\psi}}_i$ is the vector of the estimated individual parameters, i.e., the Empirical Bayes Estimates (EBEs) defined as the mode of the conditional distribution $p(\boldsymbol{\psi}_i|\boldsymbol{y}_i; \hat{\boldsymbol{\theta}})$ with $\hat{\boldsymbol{\theta}}$ the estimation of the population parameters $\boldsymbol{\theta}$. IWRES were plotted versus time and versus predicted PSA values.

Because it is difficult to interpret the individual predicted hazard functions, the evaluation of the survival part of the model essentially relied on the residuals. Cox-Snell and Martingale residuals, noted r_i^{CS} and r_i^M, respectively, are defined by $r_i^{CS} = \int_0^{T_i} h_i(s|\boldsymbol{PSA}(T_i, \hat{\boldsymbol{\psi}}_i))ds$ and $r_i^M = \delta_i - r_i^{CS}$, $i = 1, ..., n$. The Kaplan-Meier curve for the r^{CS} was compared to the survival curve of the unit exponential distribution $\exp(-t)$, while the Martingale residuals were displayed versus the predicted values inferred from the PSA kinetics and introduced in the final joint model (Rizopoulos, 2012; Sène, Bellera, and Proust-Lima, 2014). In addition to the residuals and in order to evaluate the overall prediction for the survival, the mean survival curve, defined by $S(t) = \frac{1}{N}\sum_{i=1}^{N} S_i(t|\boldsymbol{PSA}(t, \hat{\boldsymbol{\psi}}_i); \hat{\boldsymbol{\theta}})$ was calculated and compared to the Kaplan-Meier curve.

In order to evaluate the ability of the model to predict the survival in a different dataset, the mean survival curve was also calculated in the validation dataset. For that purpose,

population parameters were fixed to the values found in the training dataset (e.g., $\hat{\boldsymbol{\theta}}$) and individual parameters $\boldsymbol{\psi}_i$ were estimated from the EBEs. Of note the time-to-death in the validation dataset was not used to estimate the mean survival curve of this dataset.

Lastly, the parametric assumption for the baseline hazard function was relaxed and spline functions for the baseline hazard h_0, namely piecewise constant, linear and restricted cubic, were tested. More details can be found in the Web Appendix B.

3. Results

3.1 *Model Selection*

A model assuming an effect of docetaxel on cell angiogenesis systematically provided a better fit to the longitudinal data than a model assuming an effect on cell apoptosis. Therefore only the results of the model with an effect on cell angiogenesis will be discussed below. The values for the PSA production rate and for the mutation rate, noted p and g, respectively, were fixed to 20 ng.day^{-1} and 10^{-7} day^{-1} after the sensitivity analysis study (see Web Appendix A). None of the baseline covariates had a p-value lower than 0.2 in univariate analysis and therefore none of them were tested in the joint model ($\boldsymbol{\gamma} = 0$).

[Table 1 about here.]

Parameter estimates obtained with the 8 candidate joint models are summarized in Table 1. PSA kinetic parameters were largely insensitive to the choice of the survival part of the model. In all cases they were precisely estimated with relative standard error smaller than 8% for both fixed effects and variance components. In particular the treatment effect in blocking angiogenesis, ε, was consistently estimated to about 43% ($p < 10^{-15}$ by likelihood ratio test) and the fitness of resistant cells was close to that of sensitive cells ($RF = 99.98\%$). The model using the current PSA value outperformed all models relying on PSA in terms of BIC. Surprisingly the models based on the current PSA slope and on the cumulative PSA

(Area under PSA), did not lead to a large improvement in the BIC compared to the No Link model, i.e., the parametric Weibull model with no effect of PSA on survival.

One of the advantages of mechanistic models is that unobserved quantities can also be evaluated. Interestingly a model relying on the current number of resistant cells, R, led to a BIC smaller to what was obtained with the current PSA model. Further in the model relying on both the current number of treatment-sensitive and -resistant cells (Current S and R), the shape parameter, k, was not significantly different from 1, suggesting that the inclusion of R and S could capture most of the time-dependent change in the hazard function. This model led to the lowest BIC and therefore was retained as the final joint model for evaluation since no further combination was found to improve the BIC. Using this model and a Monte Carlo size of 200,000, the CPU times for parameter estimation and likelihood estimation were 8.72×10^3 seconds and 2.13×10^5 seconds, respectively.

Further, we studied the effect of relaxing the assumption of parametric baseline hazard function. However there was no reduction in BIC using piecewise constant, linear or restricted cubic splines on the final model (see Web Appendix B).

3.2 *Model Evaluation*

[Figure 2 about here.]

As shown in Figure 2, various kinetic patterns of PSA before, during and after treatment could be well captured by the model. The analysis of the IWRES did not suggest any model misspecification (Figure 3) in the fitting accross time or PSA values.

[Figure 3 about here.]

For the survival part, the Kaplan Meier curve of the Cox-Snell residuals (Figure 3) was close to the theoretical survival curve (unit exponential distribution). However there was a slight overestimation of the survival for small values of the residuals, indicating a possible

underestimation of the risk of death in the patients with a small risk of death. Martingale residuals (Figure 3) were satisfactory according to the number of sensitive cells but tended to underestimate the survival in patients with large number of resistant cells. Concerning the survival prediction, the mean survival curve was close to the Kaplan-Meier curve (Figure 4). In order to evaluate the capability of the model to capture the fact that PSA value at treatment initiation is associated with survival, we also compared the mean survival curves and the Kaplan-Meier curves according to the initial PSA value, PSA_0. Interestingly the model still well captured the survival function in the patients with $PSA_0 \leqslant 40$ or $40 < PSA_0 \leqslant 140$. However we note that the model tended to underestimate the survival in patients with $PSA_0 \geqslant 140$, i.e., patients with very high baseline PSA values, but the mean survival curve remained included in the Kaplan-Meier 95% confidence interval.

[Figure 4 about here.]

Finally we applied the joint model on the 196 mCRPC patients from the validation dataset. For that purpose, we fixed the population parameters to the values found in the training dataset and we calculated the EBEs for each patient (see Section 2.6). In other words, the individual trajectories for both PSA and the hazard function in a patient were calculated using only the observed PSA measurements of that patient and the information on the vital status in this dataset was not used. As can be seen in Figure 5, the Kaplan-Meier curve in the training and in the validation dataset did not exactly overlay and there was a 58 days difference in the median survival between the two datasets. In spite of this discrepancy, the mean survival curve well fitted the Kaplan-Meier curve of the validation dataset, showing that the inclusion of the individual PSA kinetics was sufficient to correctly predict the survival in this dataset.

[Figure 5 about here.]

All data and codes needed to reproduce the results, evaluate the residuals and generate the figures are available as supplementary materials with a README file.

4. Discussion

A mechanistic joint model was proposed to study the relationship between survival and longitudinal data described by a system of nonlinear ODEs. Unlike what is done in virology and bacteriology, the field of mechanistic model for treatment evaluation in oncology is still in its infancy. Here, following the efforts of other groups that aimed to develop cancer models that can be used for clinical purposes (Ribba et al., 2014), we showed that mechanistic joint models could be used for prediction of time-to-event in clinical trials. Our model allowed to capture a variety of patterns in PSA kinetics observed before, during and after chemotherapy. Further the model predicted that two quantities, namely docetaxel-sensitive and resistant cancer cells, may have a large impact on survival. Of note these two quantities were not observed, and the only observed quantity was the PSA value, which was assumed to be proportional to the total number of cancer cells. The kinetics of docetaxel-resistant cells had a larger impact on survival than the kinetics of treatment-sensitive cells. Since the kinetics of resistant cells drives the increase in PSA levels on the long-run, our prediction is consistent and expands the observation that the final tumor growth rate is highly predictive of the time-to-death (Stein et al., 2008). The relevance of the model was further reinforced by the fact that it could be used to predict the survival curve of the 196 patients that had not been used for model building. By including only the information on PSA and ignoring the information on survival in these patients, the model prediction well matched with the Kaplan-Meier curve observed in these patients.

One of the main advantages of mechanistic joint model is that parameters have a biological interpretation. Therefore putative scenarios can be performed to anticipate the impact of changes in experimental settings. Here, for instance, one could evaluate by simulation the

effect of a drug that could affect the resistant cells, either by blocking the proliferation of resistant cells or by having a complementary mechanism of action. Further, this approach could be useful to guide and design cancer clinical trials by using model predictions (such as here S and R) as surrogate markers and thus possibly reduce the follow-up time or the number of patients needed to assess treatment efficacy or to compare different treatment strategies. However, this will require to pursue the development of these models and to integrate data that are rarely collected or analyzed longitudinally such as the count of circulating tumor cells (CTC) (Wilbaux et al., 2015), the size of the tumors and/or drug pharmacokinetics. ODE models are particularly suitable for that purpose since correlations between the biomarkers can be naturally taken into account through their mechanistic interactions. Obviously these models will need to be validated internally but also externally, which was not done here, to assess their predictive ability in a different context.

However the complexity of these models comes with a cost. In terms of feasibility these models require frequent sampling longitudinal measurements to characterize the kinetics. This limits their use in the context of epidemiological studies where usually large cohorts of patients are followed with sparse measurements. Although SAEM algorithm, unlike other methods such as adaptive Gaussian quadratures (Guedj et al., 2011; Prague et al., 2013), can deal with a large number of random effects (Lavielle et al., 2011), it remains time-consuming. Further, in our case and in spite of this computation time, the calculation of the likelihood was still associated with a non-negligible standard errors, making the use of these models not reliable to identify small effects. Lastly, it remains to be demonstrated that such mechanistic joint models can be used in practice for dynamic predictions, i.e., to improve the prediction of an individual patient followed prospectively (Rizopoulos, 2011).

In conclusion, we used a novel feature of SAEM algorithm of Monolix that allows joint modelling with a longitudinal process described by ODEs to develop a mechanistic joint

model for PSA kinetics and time-to-death. This model sheds a new light on the relationship between PSA and death in mCRPC patients and this approach opens the way for the use of more complex and physiological models to improve treatment evaluation and prediction.

Supplementary Materials

Web Appendices referenced in Sections 2 and 3, as well as training dataset, MLXTRAN code, R codes and a README file are available with this paper at the *Biometrics* website on Wiley Online Library.

Acknowledgements

The authors would like to thank the Drug Disposition Department, Sanofi, Paris, which supported Solène Desmée by a research grant during this work. They also thank Guilhem Darche for the data management and Hervé Le Nagard for the use of the computer cluster services hosted on the "Centre de biomodélisation UMR1137".

References

Asar, Ö., Ritchie, J., Kalra, P. A., and Diggle, P. J. (2015). Joint modelling of repeated measurement and time-to-event data: an introductory tutorial. *International Journal of Epidemiology* **44,** 334–344.

Delyon, B., Lavielle, M., and Moulines, E. (1999). Convergence of a stochastic approximation version of the em algorithm. *Annals of Statistics* pages 94–128.

Desmée, S., Mentré, F., Veyrat-Follet, C., and Guedj, J. (2015). Nonlinear Mixed-Effect Models for Prostate-Specific Antigen Kinetics and Link with Survival in the Context of Metastatic Prostate Cancer: a Comparison by Simulation of Two-Stage and Joint Approaches. *The AAPS Journal* **17,** 691–699.

Garcia-Hernandez, A. and Rizopoulos, D. (2015). %JM Macro v2.01 Reference Manual. *Available at: www.jm-macro.com* .

Guedj, J., Thiébaut, R., and Commenges, D. (2011). Joint modeling of the clinical progression and of the biomarkers' dynamics using a mechanistic model. *Biometrics* **67,** 59–66.

Guo, X. and Carlin, B. P. (2004). Separate and Joint Modeling of Longitudinal and Event Time Data Using Standard Computer Packages. *The American Statistician* **58,** 16–24.

Herbst, R. S. and Khuri, F. R. (2003). Mode of action of docetaxel a basis for combination with novel anticancer agents. *Cancer Treatment Reviews* **29,** 407–415.

Kuhn, E. and Lavielle, M. (2005). Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics & Data Analysis* **49,** 1020–1038.

Lavielle, M., Samson, A., Karina Fermin, A., and Mentré, F. (2011). Maximum Likelihood Estimation of Long-Term HIV Dynamic Models and Antiviral Response. *Biometrics* **67,** 250–259.

Mbogning, C., Bleakley, K., and Lavielle, M. (2015). Joint modelling of longitudinal and repeated time-to-event data using nonlinear mixed-effects models and the stochastic approximation expectation maximization algorithm. *Journal of Statistical Computation and Simulation* **85,** 1512–1528.

Park, K. Y. and Qiu, P. (2014). Model selection and diagnostics for joint modeling of survival and longitudinal data with crossing hazard rate functions. *Statistics in Medicine* **33,** 4532–4546.

Perelson, A. S. and Guedj, J. (2015). Modelling hepatitis c therapy - predicting effects of treatment. *Nature Reviews Gastroenterology & Hepatology* **12,** 437–445.

Petrylak, D. (2005). Therapeutic options in androgen-independent prostate cancer: building on docetaxel. *BJU International* **96,** 41–46.

Plan, E. L., Maloney, A., Mentré, F., Karlsson, M. O., and Bertrand, J. (2012). Performance comparison of various maximum likelihood nonlinear mixed-effects estimation methods

for dose–response models. *The AAPS journal* **14,** 420–432.

Prague, M., Commenges, D., Guedj, J., Drylewicz, J., and Thiébaut, R. (2013). Nimrod: A program for inference via a normal approximation of the posterior in models with random effects based on ordinary differential equations. *Computer Methods and Programs in Biomedicine* **111,** 447–458.

Proust-Lima, C., Taylor, J. M. G., Williams, S. G., Ankerst, D. P., Liu, N., Kestin, L. L., Bae, K., and Sandler, H. M. (2008). 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. *International Journal of Radiation Oncology, Biology, Physics* **72,** 782–791.

Ribba, B., Holford, N. H., Magni, P., Trocóniz, I., Gueorguieva, I., Girard, P., Sarr, C., Elishmereni, M., Kloft, C., and Friberg, L. E. (2014). A review of mixed-effects models of tumor growth and effects of anticancer drug treatment used in population analysis. *CPT: Pharmacometrics & Systems Pharmacology* **3,** e113.

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

Rizopoulos, D. (2011). Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics* **67,** 819–829.

Rizopoulos, D. (2012). *Joint Models for Longitudinal and Time-to-Event Data: With Applications in R.* CRC Press.

Rizopoulos, D., Verbeke, G., and Lesaffre, E. (2009). Fully exponential Laplace approximations for the joint modelling of survival and longitudinal data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **71,** 637–654.

Ruffion, A., Rebillard, X., and Grima, F. (2005). Psa doubling time and method of calculation. *Progres en Urologie: Journal de l'Association Francaise d'Urologie et de*

la Societe Francaise d'Urologie **15,** 1035–1041.

Sène, M., Bellera, C. A., and Proust-Lima, C. (2014). Shared random-effect models for the joint analysis of longitudinal and time-to-event data: application to the prediction of prostate cancer recurrence. *Journal de la Société Francaise de Statistique* **155,** 134–155.

Seruga, B., Ocana, A., and Tannock, I. F. (2011). Drug resistance in metastatic castration-resistant prostate cancer. *Nature Reviews. Clinical Oncology* **8,** 12–23.

Stein, W. D., Figg, W. D., Dahut, W., Stein, A. D., Hoshen, M. B., Price, D., Bates, S. E., and Fojo, T. (2008). Tumor Growth Rates Derived from Data for Patients in a Clinical Trial Correlate Strongly with Patient Survival: A Novel Strategy for Evaluation of Clinical Trial Data. *The Oncologist* **13,** 1046–1054.

Tannock, I. F., Fizazi, K., Ivanov, S., Karlsson, C. T., Fléchon, A., Skoneczna, I., Orlandi, F., Gravis, G., Matveev, V., Bavbek, S., et al. (2013). Aflibercept versus placebo in combination with docetaxel and prednisone for treatment of men with metastatic castration-resistant prostate cancer (venice): a phase 3, double-blind randomised trial. *The lancet Oncology* **14,** 760–768.

Tsiatis, A. A. and Davidian, M. (2004). Joint modeling of longitudinal and time-to-event data: an overview. *Statistica Sinica* **14,** 809–834.

Vigan, M., Stirnemann, J., and Mentré, F. (2014). Evaluation of estimation methods and power of tests of discrete covariates in repeated time-to-event parametric models: application to gaucher patients treated by imiglucerase. *The AAPS Journal* **16,** 415–423.

Wilbaux, M., Tod, M., De Bono, J., Lorente, D., Mateo, J., Freyer, G., You, B., and Hénin, E. (2015). A joint model for the kinetics of ctc count and psa concentration during treatment in metastatic castration-resistant prostate cancer. *CPT: Pharmacometrics & Systems Pharmacology* **4,** 277–285.

Wu, L., Liu, W., Yi, G. Y., and Huang, Y. (2011). Analysis of longitudinal and survival

data: joint modeling, inference methods, and issues. *Journal of Probability and Statistics* **2012,** Article ID 640153.

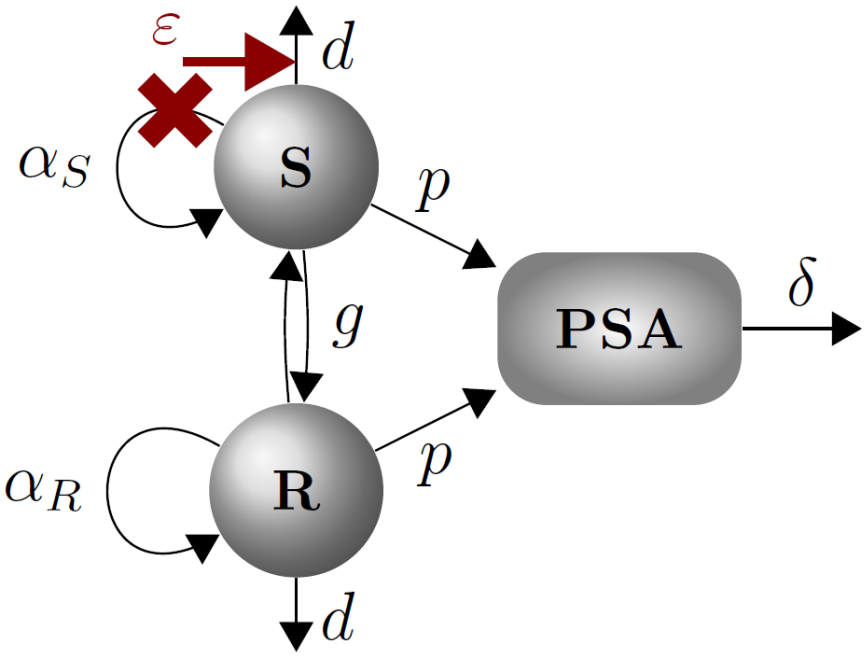

Figure 1. Schema of the secretion of PSA by sensitive (S) and resistant (R) cells. PSA is expressed in ng.mL^{-1} and S and R in mL^{-1}. In absence of treatment, α_S and α_R are the rates of S and R proliferation (day^{-1}), respectively, d the rate of S and R elimination (day^{-1}), g is the mutation rate (day^{-1}), p the rate of PSA secretion by S and R (ng.day^{-1}) and δ the rate of PSA elimination (day^{-1}). Treatment can inhibit S cells proliferation (big cross) or stimulate S cells elimination (big arrow) with the constant effectiveness ε. This figure appears in color in the electronic version of this article.

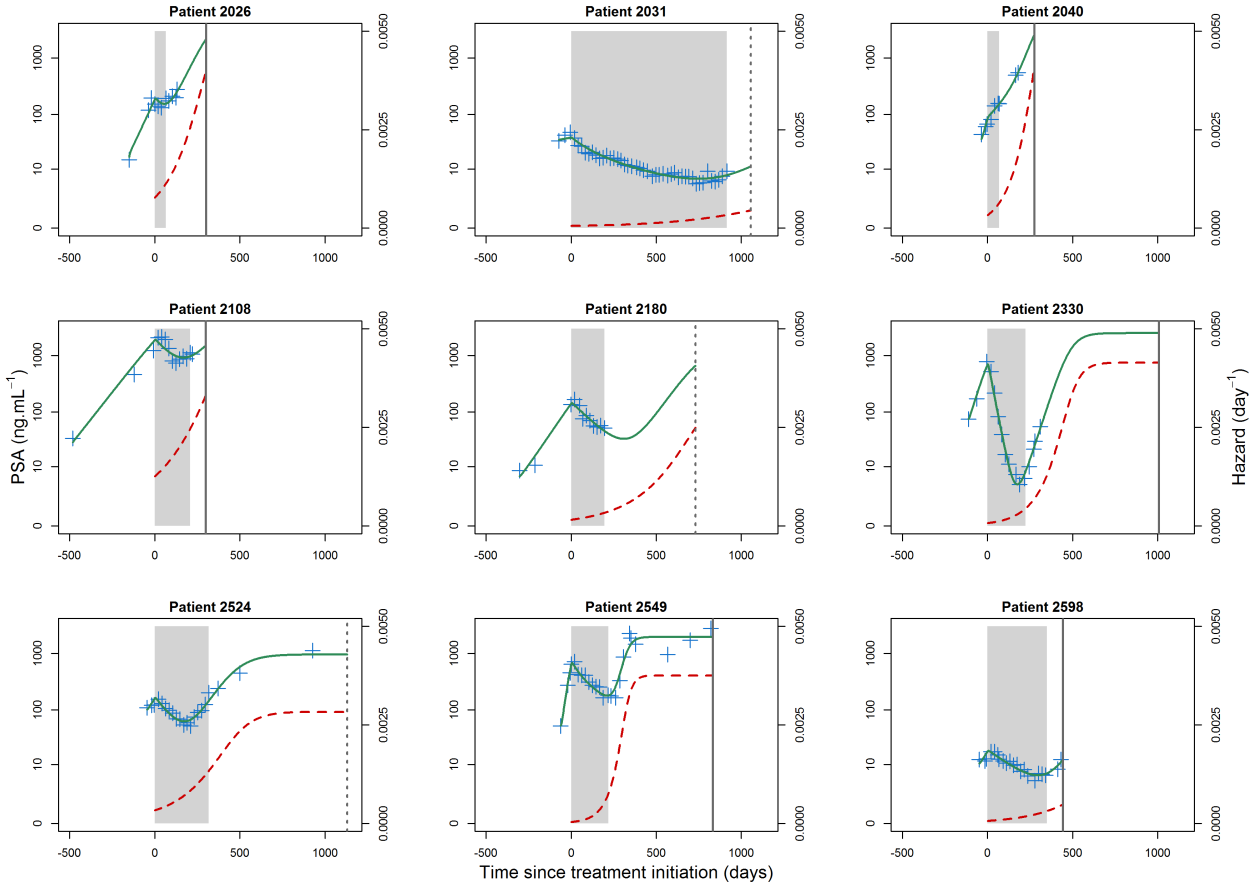

Figure 2. Individual fits of PSA kinetics and hazard function in patients with various PSA profiles. Crosses denote PSA measurements and solid lines are the model predictions using the final joint model. The vertical lines indicate the vital status (solid=death, dotted=censor) at the last recorded time and dashed curves are the individual hazard functions predicted by the final joint model. The grey area indicates the treatment period. This figure appears in color in the electronic version of this article.

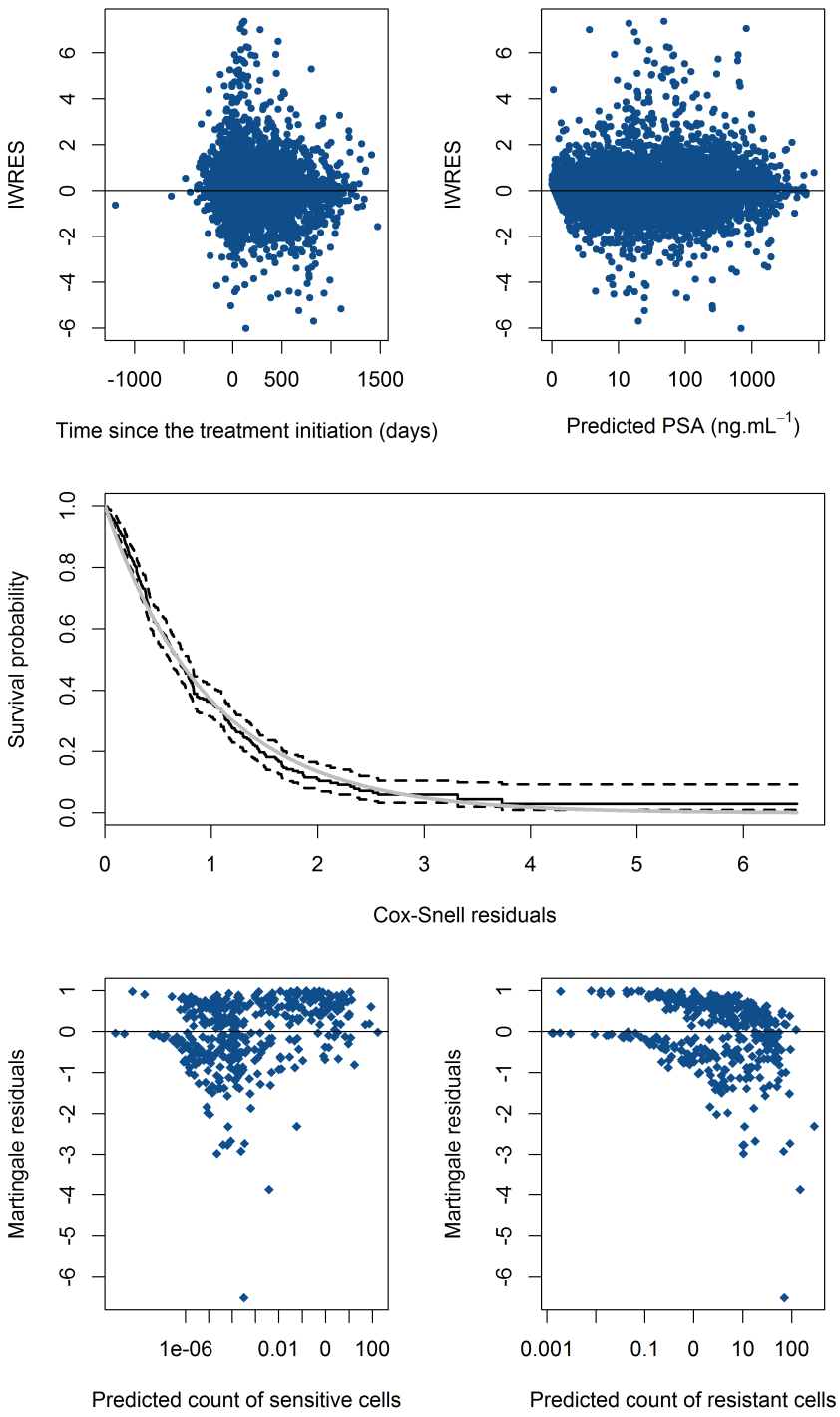

Figure 3. Residuals for the longitudinal and survival parts for the 400 patients included in the training dataset: *Top*: Individual weighted residuals (IWRES) of the longitudinal PSA model versus time *(left)* and versus predicted PSA values *(right)*. *Middle*: Kaplan-Meier estimates of the Cox-Snell residuals (solid black line) and its confidence interval (dashed black lines) and the survival function of the unit exponential distribution (solid grey line). *Bottom*: Martingale residuals versus predicted count of sensitive cells *(left)* and versus predicted count of resistant cells *(right)*. This figure appears in color in the electronic version of this article.

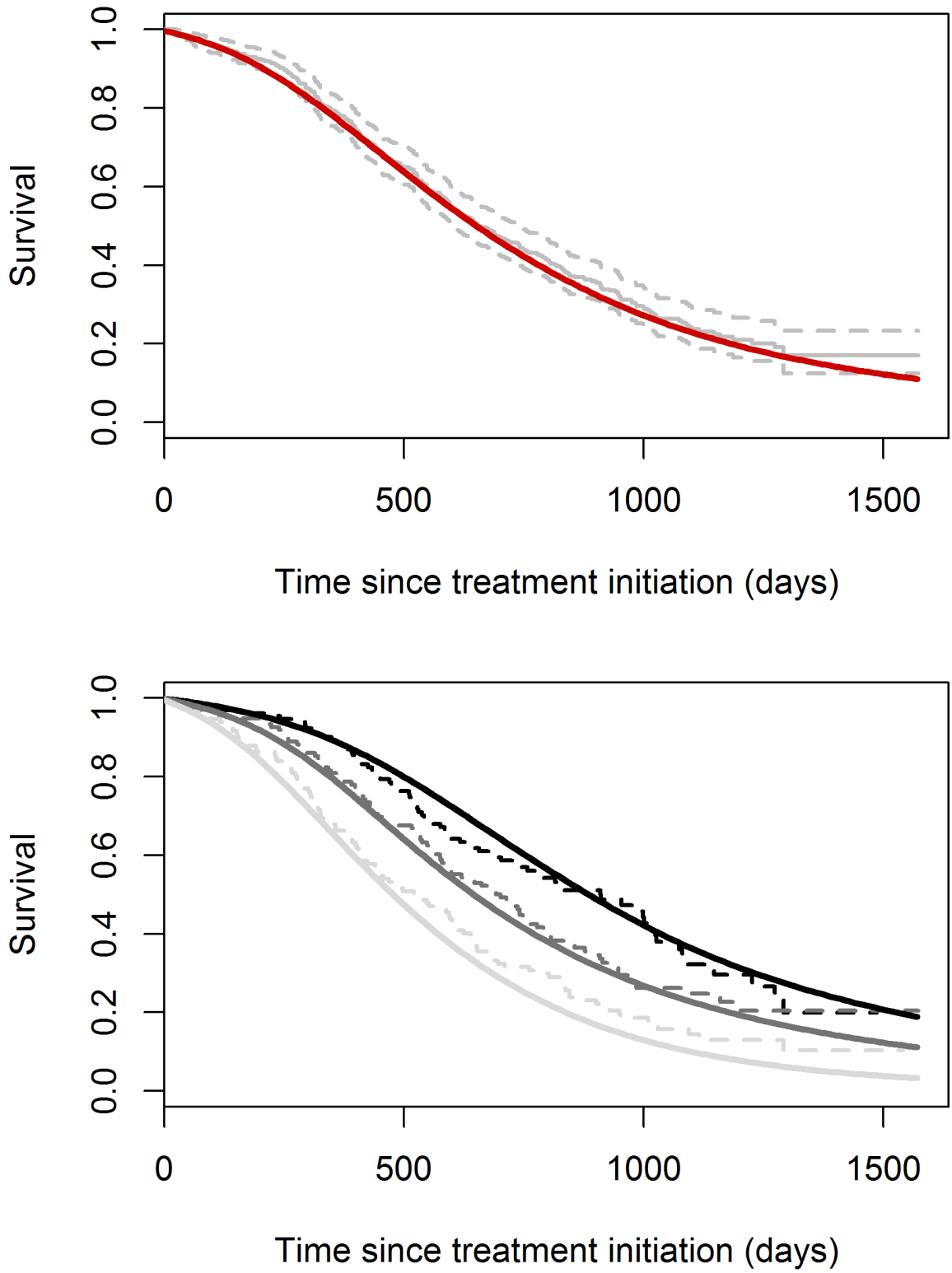

Figure 4. For the 400 patients included in the training dataset: *Top*: Estimated Kaplan-Meier survival curve (grey solid line) and its confidence interval (grey dashed line) and mean survival curve (solid line). *Bottom*: Estimated Kaplan-Meier survival curves (dotted lines) and mean survival curves (solid lines) stratified by observed PSA values at treatment initiation (black: $PSA_0 \leqslant 40$, dark grey: $40 < PSA_0 \leqslant 140$ and light grey: $PSA_0 \geqslant 140$). This figure appears in color in the electronic version of this article.

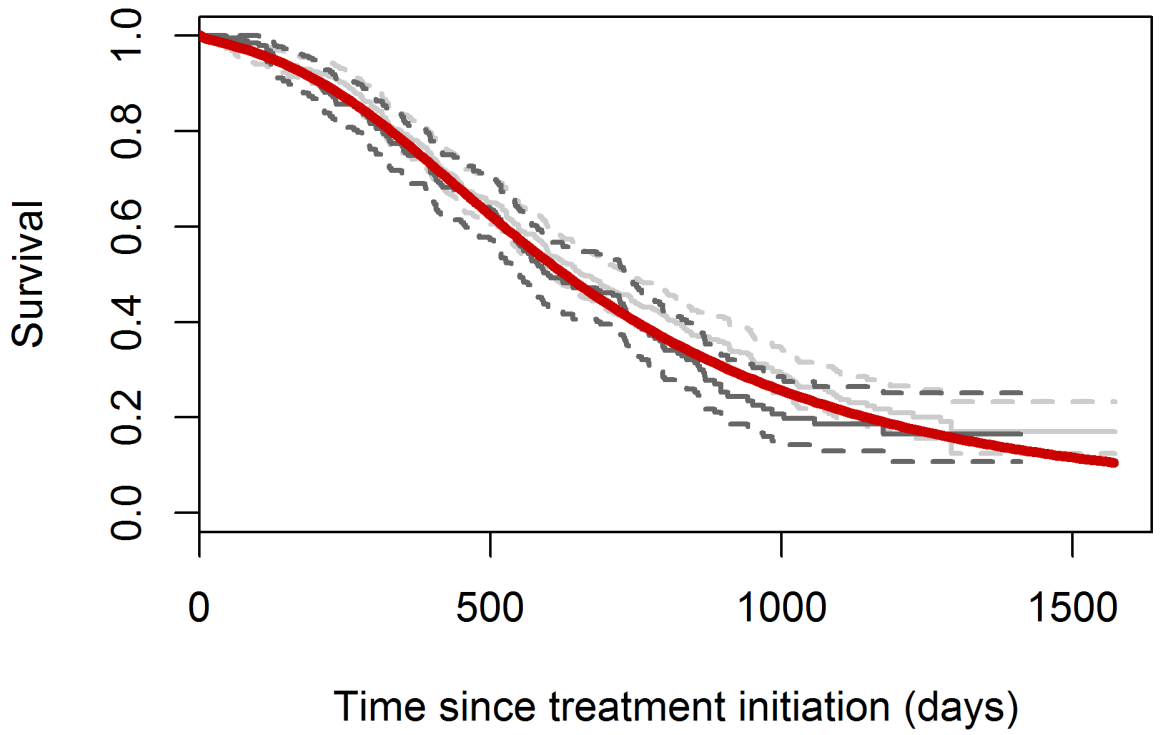

Figure 5. Estimated Kaplan-Meier survival curves (thin solid lines) and their confidence intervals (dashed lines) for the 196 patients included in the validation dataset (dark grey) and for the 400 patients included in the training dataset (light grey) and mean survival curve for the 196 patients included in the validation dataset (thick solid line). This figure appears in color in the electronic version of this article.

Table 1
Goodness-of-fit statistics (standard error) and parameters estimates (relative standard error (%)) of PSA kinetics and survival in the 400 patients of the training dataset for the 8 joint models.

Models	**No link**		**Initial PSA**		**Current PSA**		**PSA slope**		**Area under PSA**		**Current S**		**Current R**		**Current S and R**	
-2LL	14508	(7)	14486	(7)	14350	(7)	14485	(7)	14479	(7)	14471	(7)	14334	(7)	14326	(8)
BIC	14598	(7)	14582	(7)	14446	(7)	14581	(7)	14575	(7)	14567	(7)	14430	(7)	14421	(8)
α_S (day^{-1})	0.066	(3)	0.060	(3)	0.078	(3)	0.078	(3)	0.061	(3)	0.062	(3)	0.068	(3)	0.067	(3)
RF	0.9997	(0)	0.9996	(0)	0.9998	(0)	0.9998	(0)	0.9997	(0)	0.9997	(0)	0.9998	(0)	0.9998	(0)
RE	0.81	(1)	0.79	(1)	0.84	(1)	0.84	(0)	0.79	(1)	0.81	(1)	0.81	(1)	0.82	(1)
ε	0.42	(4)	0.46	(4)	0.35	(4)	0.35	(5)	0.47	(4)	0.45	(4)	0.43	(3)	0.43	(3)
PSA_b (ng.mL^{-1})	22.2	(8)	22.2	(8)	22.0	(8)	22.5	(8)	22.2	(8)	22.3	(8)	21.7	(8)	21.9	(8)
N_{max} (mL^{-1})	56	(4)	57	(4)	81	(4)	77	(4)	57	(4)	71	(4)	65	(4)	120	(4)
λ (day)	885	(4)	1615	(8)	4259	(15)	920	(4)	1435	(7)	675	(5)	877	(6)	906	(7)
k	1.52	(5)	1.53	(3)	1.28	(2)	1.48	(2)	1.19	(2)	1.73	(5)	1.01	(0)	1	-
β	-		0.21	(12)	0.40	(7)	17	(17)	0.00023	(8)	0.067	(15)	0.38	(6)	0.00032	(21)
β'	-		-		-		-		-		-		-		0.39	(7)
ω_{α_S}	0.48	(5)	0.53	(5)	0.49	(5)	0.45	(6)	0.52	(5)	0.49	(4)	0.50	(5)	0.48	(5)
ω_{RF}	2.68	(4)	2.72	(4)	2.55	(4)	2.51	(4)	2.66	(4)	2.59	(4)	2.6	(4)	2.55	(4)
ω_{RE}	0.63	(4)	0.64	(4)	0.58	(5)	0.60	(5)	0.67	(4)	0.66	(4)	0.64	(5)	0.59	(5)
ω_ε	1.14	(6)	1.21	(6)	1.10	(6)	1.04	(7)	1.15	(6)	1.21	(6)	0.90	(6)	0.95	(6)
ω_{PSA_b}	1.63	(4)	1.63	(4)	1.66	(4)	1.62	(4)	1.62	(4)	1.63	(4)	1.64	(4)	1.64	(4)
$\omega_{N_{max}}$	2.21	(6)	2.22	(6)	2.30	(6)	2.14	(6)	2.17	(6)	2.22	(6)	2.27	(6)	2.57	(7)
σ	0.32	(1)	0.32	(1)	0.32	(1)	0.32	(1)	0.32	(1)	0.32	(1)	0.32	(1)	0.32	(1)