Model-based prediction of progression-free survival for combination therapies in oncology

Received: 10 February 2023 | Revised: 12 May 2023 | Accepted: 22 May 2023

DOI: 10.1002/psp4.13003

ARTICLE

Model-based prediction of progression-free survival for combination therapies in oncology

Marcus Baaz[1,2] | **Tim Cardilin**[1] | **Mats Jirstrand**[1]

[1]Fraunhofer-Chalmers Research Centre for Industrial Mathematics, Gothenburg, Sweden

[2]Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, Gothenburg, Sweden

Correspondence
Marcus Baaz, Fraunhofer-Chalmers Research Centre for Industrial Mathematics, Gothenburg, Sweden.
Email: marcus.baaz@fcc.chalmers.se

Funding information
Merck KGaA

Abstract

Progression-free survival (PFS) is an important clinical metric for comparing and evaluating similar treatments for the same disease within oncology. After the completion of a clinical trial, a descriptive analysis of the patients' PFS is often performed post hoc using the Kaplan–Meier estimator. However, to perform predictions, more sophisticated quantitative methods are needed. Tumor growth inhibition models are commonly used to describe and predict the dynamics of preclinical and clinical tumor size data. Moreover, frameworks also exist for describing the probability of different types of events, such as tumor metastasis or patient dropout. Combining these two types of models into a so-called joint model enables model-based prediction of PFS. In this paper, we have constructed a joint model from clinical data comparing the efficacy of FOLFOX against FOLFOX + panitumumab in patients with metastatic colorectal cancer. The nonlinear mixed effects framework was used to quantify interindividual variability (IIV). The model describes tumor size and PFS data well, and showed good predictive capabilities using truncated as well as external data. A machine-learning guided analysis was performed to reduce unexplained IIV by incorporating patient covariates. The model-based approach illustrated in this paper could be useful to help design clinical trials or to determine new promising drug candidates for combination therapy trials.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

Combination therapies play a leading role in modern anticancer treatment and an important clinical metric for evaluating different treatments in oncology is progression-free survival (PFS). PFS is often estimated using the Kaplan–Meier estimator after the study is complete.

WHAT QUESTION DID THIS STUDY ADDRESS?

Can a joint tumor growth inhibition and event model be used to predict PFS before the completion of a clinical trial? How much data are required to perform adequate predictions and how well does it perform for different treatment arms, including combination therapies.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

The joint model we present was able to adequately describe the PFS for both treatment groups under analysis. The model's predictive capabilities were deemed to be good after two external validations were performed with it.

HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTICS?

The efficacy of a drug or drug combination could potentially be evaluated earlier in the drug development process.

INTRODUCTION

Combination therapies are one of the cornerstones of modern cancer treatment.[1] Their effectiveness has been attributed to their ability to combat patient variability in response, mitigate resistance, and induce synergistic effects.[2–4] The development of new and improved combination therapies is challenging due to high attrition rates and long development times.[5,6] Therefore, more research is needed to support early and accurate identification of promising new anticancer candidates and combinations.

In clinical cancer studies, overall survival (OS) is defined as the time from the start of treatment (or diagnosis) until death, and is the gold standard for evaluating treatments in oncology clinical trials.[7] Another important end point is progression-free survival (PFS), defined as the time until tumor progression (or death), based on the Response Evaluation Criteria in Solid Tumors (RECIST) criteria.[8] PFS can be used as a measure of clinical efficacy directly, or as a surrogate for OS for certain diseases.[9]

During clinical studies the disease status (e.g., tumor size and metastasis) is monitored and recorded, and used to construct a PFS curve, typically using the Kaplan–Meier estimator.[10] This allows, for example, comparisons of median PFS times between different treatments. However, predictions for alternative treatment schedules or patient populations require more sophisticated quantitative approaches.[11]

Mathematical modeling is commonly used to support the drug development process. Models are developed to capture tumor size dynamics, and to quantify the risks of detrimental events as a result of the disease.[12–15] Joint modeling can be utilized to connect time series and event data (e.g., tumor size, and metastasis or death).[16] This approach leverages data more efficiently than sequential approaches.[17,18]

By constructing a joint model for clinical data, using the nonlinear mixed-effects (NLME) framework,[19] and combining it with a model accounting for dropout, a model-based approach for determining PFS can be developed. With such a model, one can quantify between-subject variability, and estimate how different covariates affect the individual parameters. This should increase the accuracy when performing predictions on new patients. Moreover, the modeling approach also allows predictions of different treatment schedules or patients' future responses. Such models have previously been investigated and tested with several different monotherapies.[20,21]

In recent decades, large amounts of data have been gathered in oncology, where advances in, for example, single-cell technologies, and decreased sequencing costs have enabled the generation of large amounts of molecular data.[22,23] Machine learning (ML) is a popular approach for analyzing such high-dimensional data,[24] and has been used in oncology to for example, estimate the correlation between covariates and individual parameters, or PFS time, using methods such as RandomForest and LASSO.[25,26]

In this paper, we have three aims. First, to extend the model-based PFS prediction described by Yu et al.[21] to combination therapies. This is accomplished by joining a tumor growth inhibition (TGI) model to a time-to-event (TTE) model for non-target progression (NTP). An additional TTE model is also developed to account for patient dropout.[27] The models are calibrated to clinical data gathered from the PRIME study where FOLFOX was tested against FOLFOX + panitumumab.[28] Second, to test the PFS model's predictive capabilities by both predicting patients' future PFS as well as predicting the PFS for a different study. Last, we use ML to identify which covariates can help to explain inter-individual variability (IIV), potentially improving our understanding of how different subpopulations respond to treatment.

METHODS

Data

The PRIME study was a randomized phase III study that investigated FOLFOX + panitumumab ($N = 325$) against FOLFOX alone ($N = 331$), in patients with wild-type rat

sarcoma virus (RAS) mutated metastatic colorectal cancer (mCRC).[28] FOLFOX is a combination chemotherapy regimen consisting of oxaliplatin, folinic acid, and fluorouracil and is a first-line standard treatment for this type of cancer.[29] FOLFOX was administered every second week by the following regimen: oxaliplatin 85 mg/m^2 intravenous (i.v.) infusion (day 1) and leucovorin 200 mg/m^2 i.v. (day 1 + 2) followed by fluorouracil (5-FU) 400 mg/m^2 i.v. bolus (day 1 + 2) followed by 5-FU 600 mg/m^2 i.v. (day 1 + 2). Panitumumab is a humanized monoclonal antibody that inhibits the epidermal growth factor receptor (EGFR),[30] and was administered once every second week (6 mg/kg). Data from this study were collected via Project Data Sphere.[31] We extracted clinical parameters (e.g., laboratory tests, age, and RAS status) alongside PFS data, including tumor size time series and event data for NTP, death, and uninformative dropout.

According to the RECIST (version 1.1) guidelines, tumor lesions are classified as either target or non-target lesions. At most, five target lesions (usually the largest) are selected. Any remaining lesions are classified as non-target lesions. Target lesions are quantitatively measured during the trial, whereas non-target lesions are monitored only qualitatively.[32]

The PFS is determined by either target progression (TP) or NTP. TP occurs if the sum of the largest target diameters (SLDs) has increased by 20% (and 5 mm) compared to nadir.[30] NTP occurs if a new lesion appears, non-target lesions are deemed unequivocally progressing, or the patient dies. Rarely, TP and NTP occur simultaneously, in which case the patient is classified as TP. Patients without TP or NTP at their final checkup are censored (dropout). The most common reasons for dropout cited in the literature are adverse events and personal issues.[33] We denote the time these events occur by T_{TP}, T_{NTP}, and T_{DO}, respectively.

Patients were monitored once every 2 months, by two different radiologists, and to avoid the need to model the inherent variability of this we primarily used data from the first radiologist. Data from the second radiologist were only used when measurements from the first were not available. We exclude patients with a mutation in RAS exon 2 or that only had one SLD measurement. The remaining data comprise 127 and 121 patients in the FOLFOX and FOLFOX + panitumumab arms, respectively. Panitumumab alone was not tested in this study. However, it was tested in the ASPECCT study, in which two patient populations with mCRC baseline wild-type RAS with emergent or non-emergent RAS mutations were treated.[34] We extracted the PFS curves corresponding to the two arms, and constructed a single PFS curve as a weighted average with respect to the number of patients in each arm. An overview of patient characteristics for both studies can be found in Appendix S1.

Tumor growth inhibition model

To model the SLD time series, we use a TGI model shown in Figure 1.[20] If a patient is given FOLFOX, the tumor dynamics are described by the following equations:

$$\frac{\mathrm{dSLD}}{\mathrm{d}t} = k_{\mathrm{g}}\,\mathrm{SLD} - k_{\mathrm{s,FOLFOX}}\,\mathrm{SLD},$$

$$\mathrm{SLD}(0) = \mathrm{SLD}_0,$$

where k_{g} is the inherent tumor growth rate, SLD_0 the initial SLD, and $k_{\mathrm{s,FOLFOX}}$ the drug-induced killing rate of FOLFOX. We model FOLFOX as an exponentially decaying chemotherapy by:

$$k_{s,\mathrm{FOLFOX}} = a_{\mathrm{FOLFOX}}\,D_{\mathrm{FOLFOX}}\,e^{-\gamma_{\mathrm{FOLFOX}}\,t},$$

where a_{FOLFOX} and γ_{FOLFOX} are parameters corresponding to the potency and development of resistance, respectively. As we do not have pharmacokinetic (PK) data for each patient, we drive the pharmacodynamic (PD) model using the total dose per cycle (14 days), D_i, assuming the same treatment schedule for all individuals, as specified in the study protocol. In the case of FOLFOX, we combine the three individual drugs' total doses.

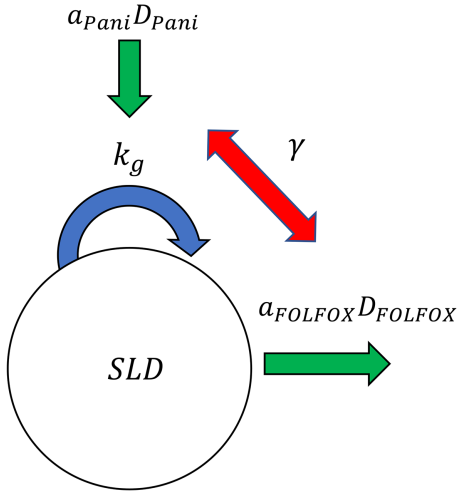

FIGURE 1 A schematic representation of the TGI model for two drugs. SLD denotes the sum of the longest diameters of the target lesions, and k_g the net tumor growth rate constant before the start of treatment. Parameters a_i and D_i are the potency and the total dose per cycle of drug i, respectively. SLD, sum of the largest target diameter; TGI, tumor growth inhibition.

Because panitumumab is an EGFR inhibitor, we model it as reducing the tumor growth rate.[35] The tumor dynamics for patients receiving FOLFOX and panitumumab simultaneously are given by:

$$\frac{\mathrm{dSLD}}{dt} = \left(k_g - k_{s,\mathrm{Pani}}\right) \mathrm{SLD} - k_{s,\mathrm{FOLFOX}}\, \mathrm{SLD},$$

$$\mathrm{SLD}(0) = \mathrm{SLD}_0.$$

where

$$k_{s,\mathrm{Pani}} = a_{\mathrm{Pani}}\, D_{\mathrm{Pani}}\, e^{-\gamma_{\mathrm{Pani}}\, t}.$$

We also tried using a $I_{\max}$ function for panitumumab but because we only have one dose level and use the average dose this led to similar but slightly worse results. Moreover, the potency function we use can be seen as a linear approximation of an $I_{\max}$ function valid for small (non-saturating) exposures.

Without data on panitumumab, given as a monotherapy, we must assume that the interaction effects between the drugs are so small that they can effectively be ignored. It is reported in the literature that the addition of FOLFOX to panitumumab resulted in no additional efficacy compared to monotherapy.[36] We evaluate this using the model to predict the PFS for panitumumab and compare it to the ASPECCT study.

We tried estimating a resistance term for each drug, but this resulted in very poor precision on the estimates. Therefore, we let $\gamma_{\mathrm{FOLFOX}} = \gamma_{\mathrm{Pani}} = \gamma$, and thus the model assumes that a patient given the combination therapy becomes resistant to both treatments simultaneously, which is not necessarily the case in reality.[37]

We use the NLME framework to account for the IIV in the data and let k_g, SLD_0, γ, and a_{FOLFOX} be log-normally distributed in the population. Hence, the specific value of the parameters for individual i is described by, $\theta_i = \theta_m e^{\eta_i}$, where $\theta_m = (k_g, \mathrm{SLD}_0, \gamma, \text{and } a_{\mathrm{FOLFOX}})$ denotes the median values and $\eta_i \sim \mathrm{MN}(0, \Omega)$ the random effects for individual i. The Ω is the covariance matrix and no correlation between random effects was assumed. We also tried to estimate a random effect for a_{Pani}, but this led to very high correlation with a_{FOLFOX}, most likely due to the lack of monotherapy data for panitumumab. Finally, the observational model for SLD contained an additive and a proportional term.

Time-to-event models

The NTP and dropout events are accounted for using TTE models. The survival function, $S(t)$, is defined as the probability of no event before time t (i.e., $S(t) = \Pr(T > t)$), where T is the time the event occurs. Survival is often expressed using a hazard function, defined as the event rate at time t conditional on survival until time t or later.[38]

We utilize a joint-modeling approach where the rate of change of SLD is linked with the hazard function for NTP events, denoted by h_1. That is:

$$h_1(t) = \max\left(\alpha \frac{\mathrm{dSLD}(t)}{dt} + \beta,\ 0\right),$$

where β is the baseline hazard of the event, and α is a scaling factor linking the hazard to SLD rate of change. Thus, better target lesions' response decreases the risk of NTP.

Dropout is not linked with patient response and is modeled separately using a parametric survival model. We tested several different common models, such as the exponential, Gompertz, and Weibull, and found that a Weibull model best fit the data. The hazard function associated with the Weibull survival model is given by:

$$h_2(t) = \frac{k}{\lambda} \cdot \left(\frac{t}{\lambda}\right)^{k-1},$$

where k is the shape parameter and λ the scale parameter or characteristic time.

Progression-free survival prediction

To predict PFS, we generate virtual patients (sample size and times based on the actual study) from the estimated population distributions. Time series, including measurement noise, are then created and T_{TP} is calculated for each virtual patient. Individual survival curves for NTP are created and from these T_{NTP} is sampled, and T_{DO} for each patient is sampled from the parametric model. These three times are then be compared and the one that occurs first is chosen as the PFS time and event for each patient. If dropout occurs first, that individual is censored at that time. This process is then repeated 1000 times to obtain a good prediction as well as a 95% confidence interval for the prediction. The algorithm for the prediction is summarized below.

1. Generate virtual patients and simulate (noisy) time series of SLD from the estimated population distributions.
2. Estimate T_{TP} from the time series.
3. Construct individual survival curves and sample T_{NTP}.
4. Sample T_{DO} using the parametric model.
5. Pick $\min\left(T_{\mathrm{Tp}}, T_{\mathrm{NTP}}, \text{and } T_{\mathrm{DO}}\right)$, record which event triggered PFS, and repeat 1000 times.

We perform two validation analyses using truncated and external data, respectively. First, we predict the PFS for panitumumab monotherapy based on the ASPECCT study, using the algorithm above, and assuming similar patient characteristics to the PRIME study. Second, we recalibrate the PFS model using truncated data from the PRIME study (at 3, 7, or 27 months) and make a forward prediction of the survival curve, median PFS, and hazard ratio for the remainder of the study. As data are increasingly truncated, the precision of the parameter estimates becomes worse. To take this uncertainty into account, we assume normally distributed fixed effects with the point estimates as the means, and the standard errors as standard deviations.

Covariate analysis

The above analysis assumes the patient characteristics of simulated studies as the PRIME study. We now quantify the correlation between the recorded baseline covariates and Empirical Bayes Estimates (EBEs) of each patient, to reduce unknown IIV and allow for more accurate predictions of new patients. We thus obtain new random effects on the following form:

$$\theta_i = \theta_m \, e^{\eta_i + Z_i \, w},$$

where Z_i are the measured covariate values for patient i and w is the estimated weight of each covariate effect.

To identify covariates to include, we use three regression methods: linear regressions, LASSO, and ridge regression along with two decision tree methods: Recursive Partitioning and RandomForest.[39] The hyperparameters for the penalized regression methods are tuned by performing cross-validation with 10 folds, and choosing the hyperparameters resulting in the smallest mean square error. Covariates that are significant by the regression methods, or top three in importance by the decision tree methods, are chosen as candidates to be included in the final model.

The parameter estimation is then rerun with all candidate covariates added, and the fixed effects are fixed at their previously estimated values. The covariate whose weight has the largest RSE for each parameter is iteratively removed until all weights are estimated with sufficient precision (RSE < 50%). We test the following 28 baseline covariates: albumin, alkaline phosphatase, creatine, lactate dehydrogenase, hemoglobin, platelets, white blood cells, carcinoembryonic antigen, age, weight, height, ethnicity, sex, Eastern Cooperative Oncology Group status, number of metastases, histology, number of target lesions at baseline, baseline SLD, type of diagnosis (colon or rectal cancer), metastasis to the liver, time since diagnosis, prior surgery, mutation/wild type/failure in KRAS exon 3 or 4, a mutation in NRAS exon 2, 3, or 4, and a mutation in BRAF exon 15. We compare our results to those obtained using the automated covariate builder in Monolix (covSAMBACOSSAC) on the full set of covariates.

Computational methods

The parameters of the TGI model and NTP TTE model were estimated simultaneously using the joint modeling functionality of Monolix[40] and the Stochastic Approximation Expectation Maximization algorithm. The joint model was exported to R studio[41] to generate visual predictive checks (VPCs) of the predicted population SLD curves. The parametric dropout model was also calibrated in Monolix and all three models were exported to Mathematica,[42] where the PFS model was created. The covariate analysis was performed in R[41] using the packages glmnet, randomForest, and rpart.

RESULTS

Tumor growth inhibition and time-to-event models

The TGI model was able to describe the SLD data adequately. Estimates of parameters are shown in Table 1. All parameters were estimated with acceptable precision (RSE < 25%). Examples of individual fits are shown in Figure 2, and VPCs can be found in Figure S1. Model predictions versus the observations, along with residual plots, can also be found in Figures S2 and S3.

The TTE models were also able to describe the data sufficiently well. We found, as previously mentioned, that a Weibull model fits the dropout data best and Kaplan–Meier VPC plots for the NTP model can be found in Figure S4. Parameters were estimated with acceptable precision for both TTE models and are presented in Table 1.

Progression-free survival model

After validating each model, they were combined to form the PFS model. Predictions with this model were performed and the results are shown in Figure 3. All simulations involved in the predictions were performed 1000 times. The figure shows that the model was able to describe the PFS well for both treatment groups in the PRIME study. Furthermore, the proportions of events in the TP, NTP, and dropout categories for virtual patients were in good agreement with what was observed in the PRIME study. For the FOLFOX treatment arm, 47%,

38%, and 15% of patients' PFS were set by TP, NTP, and dropout, respectively. For the virtual patients, the same numbers were 44%, 35%, and 21%, respectively, with a standard deviation of 4%. The numbers for the patients given FOLFOX + panitumumab arm were 42%, 44%, and 14%, and for the virtual patients, the numbers were 48%, 32%, and 19%, with a standard deviation of 5%.

TABLE 1 Estimated parameters for the TGI and TTE models.

Parameter	Units	Estimate (RSE%)	BSV (RSE%)
TGI			
SLD_0	mm	123 (5)	0.76 (5)
k_g	1/month	0.074 (7)	0.53 (11)
a_{FOLFOX}	m^2/(mg month)	2.0×10^{-4} (4)	0.41 (7.66)
a_{Pani}	kg/(mg month)	7.3×10^{-3} (15)	
γ	1/month	0.18 (10)	1.06 (7)
TTE			
α	1/mm	0.0041 (22)	
β	1/month	0.035 (19)	
k	–	1.50 (11)	
λ	month	36 (14)	
Observational error			
s_1 [a]	Mm	2.27 (12)	
s_2 [b]	–	0.15 (5)	

Abbreviations: BSV, between-subject variability; SLD, sum of the largest target diameter; TGI, tumor growth inhibition; TTE, time-to-event.

[a]Standard deviation of additive error.

[b]Standard deviation of proportional error.

The predicted median PFS time for patients with wild-type RAS mutated mCRC given panitumumab as monotherapy was 4.7 (4.2–7.0) months and a prediction for the entire PFS curve can be seen in Figure 3. Furthermore, the model was recalibrated with truncated data and three examples of how well the model could predict the rest of the data are shown in Figure 4. The data were cut after ~27, 7, and 3 months, respectively. Here, the random effect on a_{FOLFOX} was not included because including it led to high correlation between random effects. Observed hazard ratios and median PFS for the entire PRIME study and the data we modeled, along with the predictions, are shown in Table 2.

Covariate analysis

The ML algorithms identified 20 candidate covariates that could influence the random effects. The ones estimated with sufficient precision are shown in Table 3. For example, the most important covariate for explaining the variability in SLD_0 was the number of lesions at baseline.

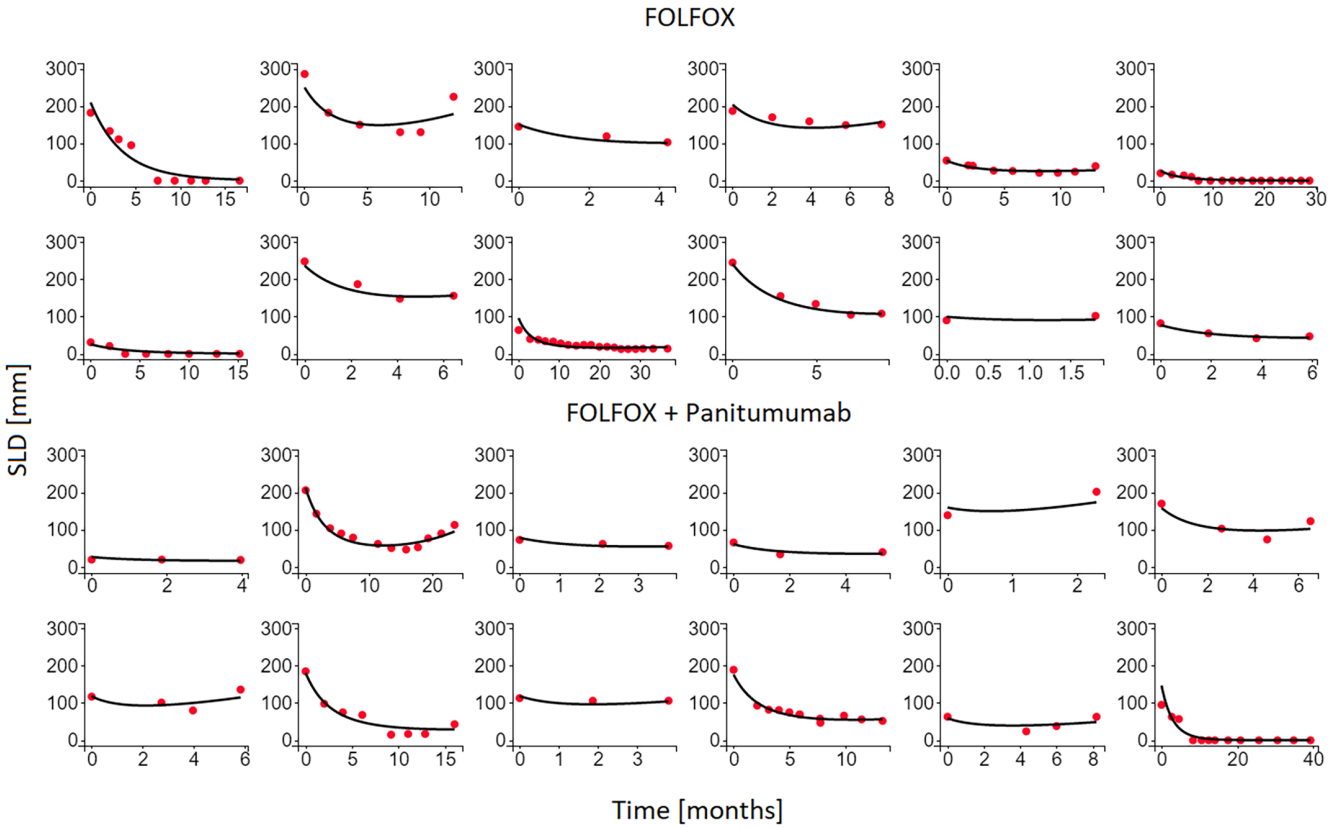

FIGURE 2 Individual model predictions (black curves) plotted together with the experimental data (red dots). The upper two rows show examples for the FOLFOX treatments arm, and the lower two for the FOLFOX + panitumumab arm. SLD, sum of the largest target diameter.

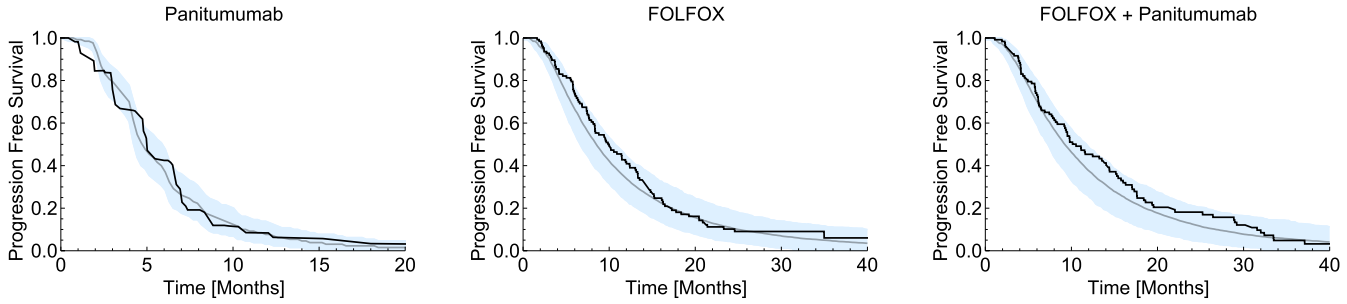

FIGURE 3 Prediction of PFS for panitumumab, FOLFOX, and FOLFOX + panitumumab treatment arms. Black lines are the observed PFS, gray lines the median prediction, and the blue areas a 95% confidence interval of the prediction. PFS, progression-free survival.

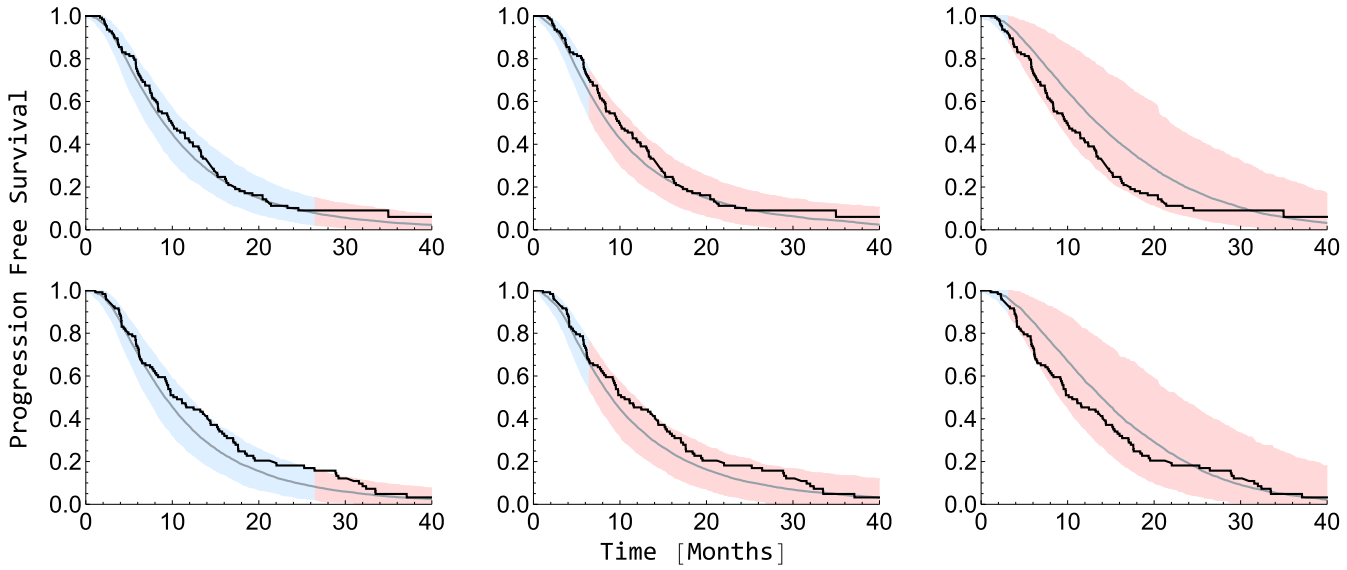

FIGURE 4 Prediction of PFS for FOLFOX (upper row) and FOLFOX + panitumumab (lower row) treatment arms with truncated data. Data truncation is indicated in red. The black line indicates observed PFS, the gray line the median prediction, and the colored areas are 95% confidence intervals for the prediction. PFS, progression-free survival.

We estimate the total reduction of variability as $1 - \frac{\mathrm{tr}(\Omega_C)}{\mathrm{tr}(\Omega)}$, where Ω_C is the covariance matrix estimated after including covariates. The automated covariate builder found many similar covariates, see Appendix S1, and it reduced the variability by 33%, which compares well with our method which reduced it by 30%.

DISCUSSION

Tumor growth inhibition and time-to-event models

As no individual PK data were available, we let the total dose per cycle drive the PD model. Similar approaches can be found in the literature.[20,21] Moreover, a PK model describing the daily dynamics of the drugs would most likely not significantly change the results, because the size of the tumors is only measured approximately once every second month.

The TGI model we use is similar to models previously used to describe SLD time series data.[20,21,43] Our model was able to describe data well on both an individual level (Figure 2) and a population level (Figure S1). All parameters were estimated with acceptable precision and the residuals were approximately normally distributed. Moreover, we calculate a doubling time of ~9.6 months, which compares well to 7.5[44] and 11.5[20] months, as reported in the literature for mCRC. γ has previously been estimated to be between 0.13 and 1.21/months for different cancers, and our estimate of 0.33 compares well with these results.[21,43,45]

No unexplained trends could be seen in the diagnostic plots for either treatment arm. The parameters of the two TTE models were estimated with acceptable precision, and the VPC Kaplan–Meier plots capture data well for both monotherapy and combination therapy. We tried estimating random effects on α and β, but it led to clear patterns in the EBEs and high shrinkage (95%).

Progression-free survival

The combined PFS model was able to describe the recorded PFS times for the population well, as can be seen in Figure 3. The proportion of PFS events between the patients of the PRIME study and the virtual patients was in good agreement. As we did not have data for panitumumab as a monotherapy, we had to assume that the interaction effects between the two treatments were so small they could effectively be ignored. Both to evaluate this assumption and challenge the model's predictive capabilities, we predicted the PFS time for panitumumab given as a monotherapy and compared it with the results from the ASPECCT study. They found a median PFS time of ~5.5 months for 164 patients with wild-type RAS, which compares well with our prediction of 4.8 (4.3–5.7) months. Figure 3 also shows the observed PFS curve is well-predicted by our model. Thus, our results seem to be consistent with the hypothesis of negligible interaction effects.

Being able to predict the median PFS for panitumumab, using only data for FOLFOX and FOLFOX + panitumumab, implies that it should also be possible to predict the result of the combination having only monotherapy data for both drugs (assuming no interaction effects). Thus, this type of PFS model can potentially be beneficial when determining which drugs to test in combination.

To further test the predictive capabilities of the model we re-estimated it with truncated data and made a forward prediction of the PFS curve, median PFS, and hazard ratio. As can be seen in Figure 4, the confidence intervals become broader with less data, because the uncertainty in the parameter estimates is included in the PFS predictions. The median predictions when the data was truncated at 27 and 7 months can still be seen to be good for both treatment arms. The large confidence interval of the predictions from the model calibrated with data truncated at 3 months is mainly due to the inability to properly estimate the resistance parameter, which is to be expected given that only short time series were used.

As can be seen in Table 2, the hazard ratio predictions all cover the observed value, but none of them show a significant difference between the treatments. This is most likely because the difference between the two treatments, in the subset of data we model, is not that large, and a significant hazard ratio was not found in the raw data either. An interesting and clinically relevant next step could be to test this approach with data where the difference is larger, and then to estimate how early one could predict a significant difference in hazard ratio. PFS-curve predictions indicate that the model could be calibrated with data from

TABLE 2 Observed and predicted hazard ratios and median PFS.

	Hazard ratio (95% CI)	Median[a] (95% CI)
	Observed values	
Full PRIME study	0.8 (0.67–0.95)	10 (9.3–11.4) vs. 8.6 (7.5–9.5)
Modeled data	0.89 (0.67–1.17)	10.5 (8.9–14) vs. 10 (8.2–11.9)
Data length (months)[b]	**Predicted values**	
All	0.92 (0.76–1.11)	9.1 (7.2–11.8) vs. 8.5 (6.9–10.9)
27	0.98 (0.83–1.17)	9.1 (7.3–11.8) vs. 9 (7.1–11.3)
7	0.95 (0.76–1.20)	8.9 (7.1–11.6) vs. 8.6 (6.9–11.2)
3	0.99 (0.79–1.31)	14 (8.9–24) vs. 13.5 (8.7.1–23)

Abbreviations: CI, confidence interval; PFS, progression-free survival.

[a]FOLFOX + Pani versus FOLFOX.

[b]Time when the data were truncated.

TABLE 3 Weights of covariates.

	SLD_0 (RSE%)	k_g(RSE%)	γ (RSE%)	a_{FOLFOX} (RSE%)
x_{Age}	–	1×10^{-2} (29)	–	4×10^{-3} (37)
x_{AP}	5×10^{-5} (27)			
x_{HG}	-9×10^{-3} (5)		-5×10^{-3} (29)	
x_{LD}	-2×10^{-4} (23)			
x_{LB}	0.26 (5)		0.17 (25)	
x_{PL}	–	-2×10^{-3} (37)		
x_{WBC}	–			-3×10^{-2} (40)
$x_{NRAS2,F}$	–		−2 (36)	
BSV	0.4	0.44	1.0	0.36

Abbreviations: AP, alkaline phosphatase; BSV, between-subject variability; HG, hemoglobin; LB, lesions at baseline; LD, lactate dehydrogenase; PL, platelets; WBC, white blood cell.

a, for example, phase II study, and used to guide the design of the next phases of the drug development process.

Nagase et al.[46] recently published an article detailing an analogous method for describing and predicting PFS based on a parametric multistate ordinary differential equation model. Instead of modeling the tumor dynamics directly, they model the number of patients remaining, with disease progression, dead, and censored, with the corresponding hazard rates as transition rates from one state to another. They highlight that, for example, a mixed effect TGI model could be incorporated into their approach to describe individual data. Thus, the research we present in this paper could perhaps facilitate this.

Alongside PFS, patient quality of life is an important consideration when testing new treatments. Winter et al.[27] have developed a joint modeling approach for linking patient dropout to health-related quality of life data. Based on our data, we have assumed that dropout follows a Weibull distribution. It could be interesting to investigate if a more descriptive model, akin to the one Winter et al. presents, could be incorporated into our framework.

Covariate analysis

To better understand why patients respond differently to treatments, we estimated the correlation between the patients' EBEs and their baseline covariate. We tested five ML algorithms to find potentially important covariates, and then used Monolix to estimate the weight of each covariate estimated with sufficient precision.

Searching among covariates indicated by the ML algorithms gave similar results to when we used the automated covariate builder in Monolix. Both approaches explained ~30% of the total IIV. The largest contribution came from SLD_0, which does not affect TP according to our model. Thus, the clinical relevance of these results is not clear, and the next step could be to investigate more covariates (e.g., more genomic data), because 70% of IIV remains to be explained.

Our analysis shows the possibilities of combining the NLME framework with ML approaches to improve the quantification of intersubject variability. This could allow us to identify genomic profiles corresponding to the probability of response, leading to more accurate predictions when designing new studies, and facilitating personalized treatment on a subgroup or individual level.

CONCLUSIONS

We successfully calibrated a joint model using clinical SLD and TTE data for the treatments FOLFOX and FOLFOX + panitumumab. Predictions with the model were performed on truncated data sets and led to good predictions of the entire PFS curve (40 months) using only 7 months of data. Moreover, predictions were also made for an external data set and the model was shown to have good predictive capabilities here as well. Thus, this modeling approach can potentially provide early insights into the efficacy of new drug combinations and be used to support decision making. We also performed an ML-guided analysis to identify patient covariates related to treatment response. Several covariates were identified, and their importance was estimated with good precision.

AUTHOR CONTRIBUTIONS

All authors wrote the manuscript. M.B. and T.C. designed the research. M.B. performed the research. M.B. analyzed the data.

ACKNOWLEDGMENTS

The authors would like to thank Anup Zutshi for his valuable comments helping to improve the manuscript.

FUNDING INFORMATION

M.B. was supported by an educational research grant from Merck KGaA, Darmstadt, Germany (CrossRef Funder ID: 10.13039/100009945).

CONFLICT OF INTEREST STATEMENT

The authors declared no competing interests for this work.

DATA AVAILABILITY STATEMENT

Some example code is provided in Appendix S1 and all code is available on reasonable request to the corresponding author. Data are available through ProjectDataSphere.

ORCID

Marcus Baaz https://orcid.org/0000-0002-5098-017X
Mats Jirstrand https://orcid.org/0000-0002-6612-8037

REFERENCES

1. Mokhtari RB, Homayouni TS, Baluch N, et al. Combination therapy in combating cancer. *Oncotarget*. 2017;8:38022-38043. doi:10.18632/oncotarget.16723
2. Palmer AC, Sorger PK. Combination cancer therapy can confer benefit via patient-to-patient variability without drug additivity or synergy. *Cell*. 2017;171:1678-1691.e13. doi:10.1016/j.cell.2017.11.009
3. Fitzgerald JB, Schoeberl B, Nielsen UB, Sorger PK. Systems biology and combination therapy in the quest for clinical efficacy. *Nat Chem Biol*. 2006;2:458-466. doi:10.1038/nchembio817
4. Vakil V, Trappe W. Drug combinations: mathematical modeling and networking methods. *Pharmaceutics*. 2019;11:208. doi:10.3390/pharmaceutics11050208

5. Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates? *Nat Rev Drug Discov*. 2004;3:711-715. doi:10.1038/nrd1470
6. Arrowsmith J, Miller P. Phase II and phase III attrition rates 2011–2012. *Nat Rev Drug Discov*. 2013;12:569. doi:10.1038/nrd4090
7. Zhuang SH, Xiu L, Elsayed YA. Overall survival: a gold standard in search of a surrogate: the value of progression-free survival and time to progression as end points of drug efficacy. *Cancer J*. 2009;15:395-400. doi:10.1097/PPO.0b013e3181be231d
8. Gutman SI, Piper M, Grant MD, Basch E, Oliansky DM, Aronson N. Background. Agency for Healthcare Research and Quality (US); 2013.
9. Gyawali B, Eisenhauer E, Tregear M, Booth CM. Progression-free survival: it is time for a new name. *Lancet Oncol*. 2022;23:328-330. doi:10.1016/S1470-2045(22)00015-8
10. Goel MK, Khanna P, Kishore J. Understanding survival analysis: Kaplan-Meier estimate. *Int J Ayurveda Res*. 2010;1:274-278. doi:10.4103/0974-7788.76794
11. Etikan I. The Kaplan Meier estimate in survival analysis. *Biom Biostat Int J*. 2017;5:5. doi:10.15406/bbij.2017.05.00128
12. Simeoni M, Magni P, Cammia C, et al. Predictive pharmacokinetic-pharmacodynamic modeling of tumor growth kinetics in xenograft models after administration of anticancer agents. *Cancer Res*. 2004;64:1094-1101. doi:10.1158/0008-5472.CAN-03-2524
13. Gabrielsson J, Gibbons FD, Peletier LA. Mixture dynamics: combination therapy in oncology. *Eur J Pharm Sci*. 2016;88:132-146. doi:10.1016/j.ejps.2016.02.020
14. Cardilin T, Almquist J, Jirstrand M, et al. Tumor static concentration curves in combination therapy. *AAPS J*. 2017;19:456-467. doi:10.1208/s12248-016-9991-1
15. Le-Rademacher J, Wang X. Time-to-event data: an overview and analysis considerations. *J Thorac Oncol*. 2021;16:1067-1074. doi:10.1016/j.jtho.2021.04.004
16. Zhudenkov K, Gavrilov S, Sofronova A, et al. A workflow for the joint modeling of longitudinal and event data in the development of therapeutics: tools, statistical methods, and diagnostics. *CPT Pharmacomet Syst Pharmacol*. 2022;11:425-437. doi:10.1002/psp4.12763
17. Ibrahim JG, Chu H, Chen LM. Basic concepts and methods for joint models of longitudinal and survival data. *J Clin Oncol*. 2010;28:2796-2801. doi:10.1200/JCO.2009.25.0654
18. Bruno R, Mercier F, Claret L. Evaluation of tumor size response metrics to predict survival in oncology clinical trials. *Clin Pharmacol Ther*. 2014;95:386-393. doi:10.1038/clpt.2014.4
19. Leander J, Almquist J, Johnning A, Larsson J, Jirstrand M. Nonlinear mixed effects modeling of deterministic and stochastic dynamical systems in Wolfram Mathematica. *IFAC-Pap*. 2021;54:409-414. doi:10.1016/j.ifacol.2021.08.394
20. Claret L, Girard P, Hoff PM, 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;27:4103-4108. doi:10.1200/JCO.2008.21.0807
21. Yu J, Wang N, Kågedal M. A new method to model and predict progression free survival based on tumor growth dynamics. *CPT Pharmacomet Syst Pharmacol*. 2020;9:177-184. doi:10.1002/psp4.12499
22. Marx V. The big challenges of big data. *Nature*. 2013;498:255-260. doi:10.1038/498255a
23. Jiang P, Sinha S, Aldape K, Hannenhalli S, Sahinalp C, Ruppin E. Big data in basic and translational cancer research. *Nat Rev Cancer*. 2022;22:1-15. doi:10.1038/s41568-022-00502-0
24. Pugliese R, Regondi S, Marini R. Machine learning-based approach: global trends, research directions, and regulatory standpoints. *Data Sci Manag*. 2021;4:19-29. doi:10.1016/j.dsm.2021.12.002
25. Capobianco E. High-dimensional role of AI and machine learning in cancer research. *Br J Cancer*. 2022;126:523-532. doi:10.1038/s41416-021-01689-z
26. Terranova N, French J, Dai H, et al. Pharmacometric modeling and machine learning analyses of prognostic and predictive factors in the JAVELIN gastric 100 phase III trial of avelumab. *CPT Pharmacomet Syst Pharmacol*. 2022;11:333-347. doi:10.1002/psp4.12754
27. Winter A, Cuer B, Conroy T, et al. Flexible modeling of longitudinal health-related quality of life data accounting for informative dropout in a cancer clinical trial. *Qual Life Res*. 2022;32:669-679. doi:10.1007/s11136-022-03252-6
28. Douillard J-Y, Oliner KS, Siena S, et al. Panitumumab–FOLFOX4 treatment and RAS mutations in colorectal cancer. *N Engl J Med*. 2013;369:1023-1034. doi:10.1056/NEJMoa1305275
29. Ikoma N, Raghav K, Chang G. An update on randomized clinical trials in metastatic colorectal carcinoma. *Surg Oncol Clin N Am*. 2017;26:667-687. doi:10.1016/j.soc.2017.05.007
30. Chua YJ, Cunningham D. Panitumumab. *Drugs Today Barc Spain*. 1998;2006(42):711-719. doi:10.1358/dot.2006.42.11.1032061
31. Project Data Sphere 2022. https://www.projectdatasphere.org/.
32. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45:228-247. doi:10.1016/j.ejca.2008.10.026
33. Kim J, Kim MG, Lim K-M. Participation in and withdrawal from cancer clinical trials: a survey of clinical research coordinators. *Asia-Pac J Oncol Nurs*. 2022;9:197-201. doi:10.1016/j.apjon.2021.12.015
34. Kim TW, Peeters M, Thomas A, et al. Impact of emergent circulating tumor DNA RAS mutation in panitumumab-treated chemoresistant metastatic colorectal cancer. *Clin Cancer Res*. 2018;24(22):5602-5609. doi:10.1158/1078-0432.ccr-17-3377
35. Vera-Yunca D, Parra-Guillen ZP, Girard P, Trocóniz IF, Terranova N. Relevance of primary lesion location, tumour heterogeneity and genetic mutation demonstrated through tumour growth inhibition and overall survival modelling in metastatic colorectal cancer. *Br J Clin Pharmacol*. 2022;88:166-177. doi:10.1111/bcp.14937
36. Amgen. Highlights of Prescribing Information VECTIBIX 2006.
37. Loria R, Vici P, Di Lisa FS, Soddu S, Maugeri-Saccà M, Bon G. Cross-resistance among sequential cancer therapeutics: an emerging issue. *Front Oncol*. 2022;12.
38. Moore DF. *Applied Survival Analysis Using R*. Springer International Publishing; 2016. doi:10.1007/978-3-319-31245-3
39. Hastie T, Tibshriani R, Friedman J. *The Elements of Statistical Learning*. 2nd ed. Springer; 2017.
40. Monolix 2021R2 Lixoft SAS, a Simulations Plus company.
41. R Core Team. R: A Language and Environment for Statistical Computing. 2022.

42. Wolfram Research, Inc. Mathematica, Version 13.0. Wolfram Research, Inc.; 2022.
43. Krishnan SM, Laarif SS, Bender BC, Quartino AL, Friberg LE. Tumor growth inhibition modeling of individual lesion dynamics and interorgan variability in HER2-negative breast cancer patients treated with docetaxel. *CPT Pharmacomet Syst Pharmacol*. 2021;10:511-521. doi:10.1002/psp4.12629
44. Burke JR, Brown P, Quyn A, Lambie H, Tolan D, Sagar P. Tumour growth rate of carcinoma of the colon and rectum: retrospective cohort study. *BJS Open*. 2020;4:1200-1207. doi:10.1002/bjs5.50355
45. Fostvedt LK, Nickens DJ, Tan W, Parivar K. Tumor growth inhibition modeling to support the starting dose for dacomitinib. *CPT Pharmacomet Syst Pharmacol*. 2022;11:1256-1267. doi:10.1002/psp4.12841
46. Nagase M, Doshi S, Dutta S, Lin C-W. Estimation of time to progression and post progression survival using joint modeling of summary level OS and PFS data with an ordinary differential equation model. *J Pharmacokinet Pharmacodyn*. 2022;49:455-469. doi:10.1007/s10928-022-09816-w

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Baaz M, Cardilin T, Jirstrand M. Model-based prediction of progression-free survival for combination therapies in oncology. *CPT Pharmacometrics Syst Pharmacol*. 2023;00:1-11. doi:10.1002/psp4.13003