Systematic Modeling and Design Evaluation of Unperturbed Tumor Dynamics in Xenografts[S]

Zinnia P. Parra-Guillen,[1] Victor Mangas-Sanjuan,[1,2] Maria Garcia-Cremades,[3]
Iñaki F. Troconiz, Gary Mo, Celine Pitou, Philip W. Iversen, and Johan E. Wallin

Pharmacometrics and Systems Pharmacology Research Unit, Department of Pharmacy and Pharmaceutical Technology, School of Pharmacy and Nutrition, University of Navarra, Pamplona, Spain (Z.P.P.-G.,V.M.-S.,M.G.-C., I.F.T.); Navarra Institute for Health Research, Pamplona, Spain (Z.P.P.-G.,V.M.-S.,M.G.-C., I.F.T.); Global PK/PD & Pharmacometrics (G.M., C.P., J.E.W.) and Lilly Research Laboratories (P.W.I.), Eli Lilly and Company, Indianapolis, USA Solna, Sweden

ABSTRACT

Xenograft mice are largely used to evaluate the efficacy of oncological drugs during preclinical phases of drug discovery and development. Mathematical models provide a useful tool to quantitatively characterize tumor growth dynamics and also optimize upcoming experiments. To the best of our knowledge, this is the first report where unperturbed growth of a large set of tumor cell lines ($n = 28$) has been systematically analyzed using a previously proposed model of nonlinear mixed effects (NLME). Exponential growth was identified as the governing mechanism in the majority of the cell lines, with constant rate values ranging from 0.0204 to 0.203 day^{-1}. No common patterns could be observed across tumor types, highlighting the importance of combining information from different cell lines when evaluating drug activity. Overall, typical model parameters were precisely estimated using designs in which tumor size measurements were taken every 2 days. Moreover, reducing the number of measurements to twice per week, or even once per week for cell lines with low growth rates, showed little impact on parameter precision. However, a sample size of at least 50 mice is needed to accurately characterize parameter variability (i.e., relative S.E. values below 50%). This work illustrates the feasibility of systematically applying NLME models to characterize tumor growth in drug discovery and development, and constitutes a valuable source of data to optimize experimental designs by providing an a priori sampling window and minimizing the number of samples required.

Introduction

Drug discovery and development costs have been rising over the past years without translating into a larger number of drugs approved by the regulatory agencies (see the Pharmaceutical Research and Manufacturers of America 2016 profile; http://phrma-docs.phrma.org/sites/default/files/pdf/biopharmaceutical-industry-profile.pdf). This is especially relevant in the oncology area, where the attrition rates are among the highest (Hay et al., 2014). One approach to reducing these alarming numbers is to look for strategies to better predict clinical trial outcomes using early preclinical information (Zhang et al., 2006).

Xenograft and syngeneic models are widely used in drug development to evaluate the antitumor effects of oncological compounds and guide the selection and development of drug candidates (Sausville and Burger, 2006; Ocana et al., 2010). In these experiments, compounds exhibiting promising in vitro properties are tested on xenograft mouse experiments, typically across a small handful of different tumor cell lines. Success is then measured in terms of achieved tumor regression/cure in the treatment group compared with the control group.

Many experimental variables such as cell line aggressiveness, tumor size at dosing time, dose intensity, and dose frequency can significantly affect the evolution of tumor size over time, and consequently the outcome of these experiments in terms of observed shrinkage. Making an adequate choice of design can lead to more informative studies, with the additional benefit of reducing the number of animals, samples, or repetitions needed, and consequently reducing experimental costs, as recently illustrated by Lestini et al. (2016). However, pharmacokinetics, pharmacodynamics, and disease progression properties of the systems are rarely taken into account when designing xenograft experiments (Simeoni et al., 2013).

In this regard, semimechanistic pharmacokinetics/pharmacodynamic analysis could constitute a powerful tool to integrate information from multiple experiments, and obtain design-independent model parameters that would then

[1]Z.P.P.-G. and V.M.-S. contributed equally to this work.
[2]Current affiliation: Department of Pharmacy and Pharmaceutical Technology and Parasitology, University of Valencia, Valencia, Spain; and Molecular Recognition and Technological Development, Polytechnic University–University of Valencia, Valencia, Spain.
[3]Current affiliation: Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, California.
https://doi.org/10.1124/jpet.118.248286.
[S] This article has supplemental material available at jpet.aspetjournals.org.

ABBREVIATIONS: IAV, interanimal variability; ISV, interstudy variability; NLME, nonlinear mixed effects; RSE, relative standard error; TV, tumor volume; TV$_0$, initial tumor volume; TV$_{th}$, tumor volume threshold.

TABLE 1

Summary of the experimental data available for each of the tumor cell lines

The T_{max} value represents the latest sampling time available in the experiment; the T_{min} value represents the time at which the tumor could start to be measured.

Tumor	Cell	Number of Studies	Number of Mice	Number of Samples	Number of Days	
					T_{min}	T_{max}
Breast	MB-231	1	45	423	11	45
Colon	COLO205	5	85	668	6	120
Colon	HCT116	1	10	25	7	17
Glioblastoma	U-87 MG	3	100	382	7	63
Leukemia	MV411	2	33	388	7	73
Lung	A549	2	33	389	7	69
Lung	Calu-6	2	25	288	7	59
Lung	DMS53	1	6	120	7	98
Lung	H1650	2	21	277	7	62
Lung	H1975	3	61	361	7	41
Lung	H2122	8	92	763	7	50
Lung	H358	1	8	160	10	88
Lung	H441	17	287	3391	3	98
Lung	H460	1	13	109	8	47
Lung	HCC827	1	8	99	10	55
Lymphoma	JEKO-1	4	31	312	7	46
Lymphoma	OCI-LY-19	2	38	178	7	32
Lymphoma	WILL-2	1	24	80	7	23
Melanoma	A2058	2	20	177	7	35
Melanoma	A375	8	104	840	3	69
Melanoma	GAK	2	15	181	8	56
Melanoma	SK-MEL-30	2	24	229	0	45
Ovarian	A2780	1	8	68	8	37
Ovarian	SKVO-3	1	8	104	7	47
Pancreas	MIA PaCa-2	1	8	64	11	41
Renal	786-o	4	52	740	6	76
Renal	ACHN	3	50	604	7	60
Renal	Caki-1	1	13	116	13	51

enable the in silico exploration and optimization of potentially relevant scenarios. Nonlinear mixed effects (NLME) modeling represents a suitable methodology in this context to quantitatively describe not only the typical pharmacokinetics/pharmacodynamic profile but also to identify and quantify the different sources of variability (i.e., across animals and experiments), thus allowing for better data description and understanding.

To characterize tumor size data from xenograft mice experiments, a model that considers two different processes in the natural course of tumor progression (exponential and linear) (Simeoni et al., 2004) is commonly used. This model describes tumor growth dynamics in the absence or presence of anticancer treatment and enables the computation of a threshold concentration, such that if drug steady-state concentrations above that threshold are reached, the model will predict tumor shrinkage and eventually eradication. Indeed, Rocchetti et al. (2007) demonstrated a good correlation between the threshold concentration and the active dose used in clinic for several marketed drugs, thus suggesting its potential for scaling efficacy from animal to man, further supporting the use of this preclinical model. To the best of our knowledge, the NMLE methodology has been seldom used in pharmacokinetics/pharmacodynamic analyses of mice xenograft data, and when used most of the studies focused on a limited number of cell lines or tumor types only (Bueno et al., 2008; Parra-Guillen et al., 2013; Tate et al., 2014, 2016).

In this setting, the overall goal of our work was to illustrate how NLME modeling and optimal design theory can be systematically applied during the preclinical evaluation of new drug candidates to maximize the information that can be

extracted from xenograft experiments and enable its subsequent use in the design of future protocols. An accurate description of the tumor dynamics in the absence of drug is a key step before characterizing tumor growth inhibition drug effects. Therefore, unperturbed tumor growth data from a large set of tumor cell lines representing several cancer types was initially selected in this work with the following goals in mind: 1) to characterize tumor growth dynamics in the absence of active treatment; 2) to evaluate the impact of study design and number of animals on the precision of parameter estimates; and 3) to assess the potential benefit of optimizing the sampling time points.

Materials and Methods

Experimental Data

Data were used from several experiments in which tumor volume (TV) measurements had been obtained from control mice. A total of 28 cell lines from 10 different tumor types were available for the analysis (see Table 1 for an overview of the experimental data). Briefly, tumor cells were subcutaneously inoculated into athymic nude mice, weighing a mean of 23 g (16–32 g). Tumor size was measured with a caliper at regular times, and TV was computed assuming the tumor has an ovoid form: $(l \times w^2)/2$ (Pierrillas et al., 2016), where l is the length and w the width of the tumor in millimeters ($l > w$). Mice were sacrificed when the measured TV exceeded a prespecified upper limit. All animal experiments were approved by the Eli Lilly and Company Institutional Animal Care and Use Committee.

Unperturbed Tumor Growth Modeling

The model first proposed by Simeoni et al. (2004) was used to characterize the unperturbed tumor growth dynamics in all cell lines

evaluated. Briefly, the model describes an initial exponential increase in TV followed by linear growth, assuming no spontaneous tumor cell death and homogenous tumor cell behavior:

$$\frac{d\text{TV}}{dt} = \frac{\lambda_0 \times \text{TV}}{\left\{1 + [(\lambda_0/\lambda_1) \times \text{TV}]^{\psi}\right\}^{1/\psi}} \quad (1)$$

where TV is the tumor volume at any time after cell inoculation; and λ_0 and λ_1 are the zero- and first-order rate constants characterizing the exponential and linear growth kinetics, respectively. The transition from exponential to linear growth occurs when the TV reaches the value of the tumor volume threshold (TV_{th}), which can be expressed as λ_1/λ_0. The shape factor (ψ) is fixed to the value of 20, ensuring a rapid transition between the two different growth rates once the TV reaches the TV_{th}. The initial condition of the system is given by the initial tumor volume (TV_0), defined as the TV immediately after tumor cell inoculation.

NLME modeling was used to analyze the data. The first-order conditional estimation method with an interaction algorithm implemented in NONMEM 7.3 (Icon Development Solutions, Ellicott City, MD) was selected. During the analysis, the typical population parameters TV_0, λ_0, and λ_1 were estimated. Interanimal (IAV) and interstudy variability (ISV) were explored using the three estimated typical parameters assuming a lognormal distribution of the individual parameters:

$$\theta_i = \theta_{\text{pop}} \times \exp(\eta_i) \quad (2)$$

where θ_i is the individual model parameter of mouse i; θ_{pop} is the population parameter (i.e., TV_0, λ_0, or λ_1); and η_i is the associated random effect of mouse i obtained from a normal distribution with mean 0 and estimated variance of the interindividual variability (ω^2).

The TV data were logarithmically transformed for the analysis, and residual variability (taking into account the discrepancy between individual predictions and observations) was described using an additive error model on the logarithmic scale:

$$\log(\text{TV}_{i,j}) = \log[f(\theta_i, t_j)] + \varepsilon_{i,j} \quad (3)$$

where $\text{TV}_{i,j}$ represents the TV observations for the ith animal at the jth measurement time; $f(\theta_i, t_j)$, corresponds to the predicted TV for mouse ith at time jth resulting from the vector of the ith individual parameters θ_i related as shown in eq. 1; and ε is the difference between the logarithm of the individual observed and predicted TVs. The set of ε values is independent between them and corresponds to a random variable mean 0 and estimated variance of the residual variability (σ^2).

Upper welfare limits for tumor size were not available for all of the different studies, neither were the reasons for dropouts. Therefore, no censoring or dropout model was considered in the analysis.

Model Selection and Evaluation. Different statistical models (i.e., IAV parameters, ISV parameters, and correlation between individual parameters) were explored and compared in terms of minimum value of the objective function value provided by NONMEM (approximately −2 log likelihood), precision of parameter estimates when obtained, and goodness of fit. A decrease in the objective function value of 3.84 between two nested models was considered significant at the 5% level.

Model performance was evaluated using the simulation-based diagnostics visual predictive checks generated with PsN version 4 (Lindbom et al., 2005) and the Xpose4 R package (Jonsson and Karlsson, 1999). For each cell line, 1000 studies with the same design characteristics as the original experiments were simulated. For each simulated scenario (tumor cell line) and measurement interval (or bin), the 2.5th, 50th, and 97.5th percentiles of the simulated TV values were calculated, and then the 95% confidence intervals of the aforementioned predicted percentiles were obtained and presented graphically together with the 2.5th, 50th, and 97.5th percentiles of the raw data.

Design Evaluation and Optimization

According to the Cramér-Rao inequality, the inverse of the Fisher information matrix is the lower bound of the variance/covariance matrix of any unbiased parameter estimate (Radhakrishna Rao, 1945; Cramér, 1946). Therefore, the S.E. value can be obtained from the square root of the diagonal elements of the inverse of the Fisher information matrix. The optimal design R package POPED (Foracchia et al., 2004) was used to compute parameter precision—i.e., relative standard errors (RSEs) calculated as the ratio between the S.E. values and the parameter estimates obtained for the different studied cell lines—given the selected model, the final set of parameter estimates, and an experimental design.

Parameter precision was evaluated under three distinct sampling schemas: 1) the standard schema of TV measurements taken every 2 days; 2) samples taken twice per week (i.e., two samples per week on days 1 and 4); or 3) once per week (i.e., one sample per week on day 1). For all of the evaluations, the original number of mice per study (ranging from 7 to 287) and the sampling window (see Table 1) were kept unchanged. Note that although the selected criteria commonly refer to the frequency of dose administration, in this context they are used to indicate the frequency at which TV measurements were taken. The RSEs obtained for the different model parameters across cell lines and evaluated sampling designs were analyzed.

In a second step, the sampling times were optimized assuming eight samples in all cases and allowing for only one tumor size sample per day. The study duration and the number of animals were not optimized, i.e., original designs were used. The adaptive random search methods implemented in POPED were used in the optimization step.

Results

Unperturbed Tumor Growth. The unperturbed tumor growth model proposed by Simeoni et al. (2004) provided a good description of the data for all tumor cell lines, as illustrated in Fig. 1 for the case of lung cancer cell lines and in Supplemental Fig. 1 for the rest of tumor cell lines. Parameter estimates can be found in Table 2. The estimates of the zero-order rate constant [reflecting the linear tumor growth (λ_1)] ranged from 2.58 to 464 mm^3/day. With respect to λ_0, taking into account the exponential growth, and TV_0, the corresponding ranges were much smaller, varying from 0.0204 to 0.203 day^{-1} and 16 to 148 mm^3, respectively. Only in the case of MB231 did the estimate for λ_0 have to be fixed to a high value of 0.5 day^{-1}, indicating that only linear growth was observed in the measurable range (TV_{th} of 5 mm^3).

IAV for the λ_0 and/or TV_0 parameters was identified in all tumor cell lines, with a low-to-moderate magnitude ranging from 8% to 70% and 12% to 61%, respectively. IAV in the λ_1 parameter could be estimated only in 10 out of 28 cell lines, with values ranging from 42% to 88% (Fig. 2A). This is due to the fact that for most of the cell lines little information was available over TV_{th} (marking the shift from exponential to linear growth), thus limiting the identification of IAV variability in λ_1. A correlation between individual parameters was explored during the modeling process and found to be significant only for the A549 and MV411 cell lines (Table 2).

There were no relevant relationships between the magnitude of the different variability parameters and its typical estimate or the experimental size (Fig. 2B). Inclusion of ISV did not show significance for the case of lung and melanoma
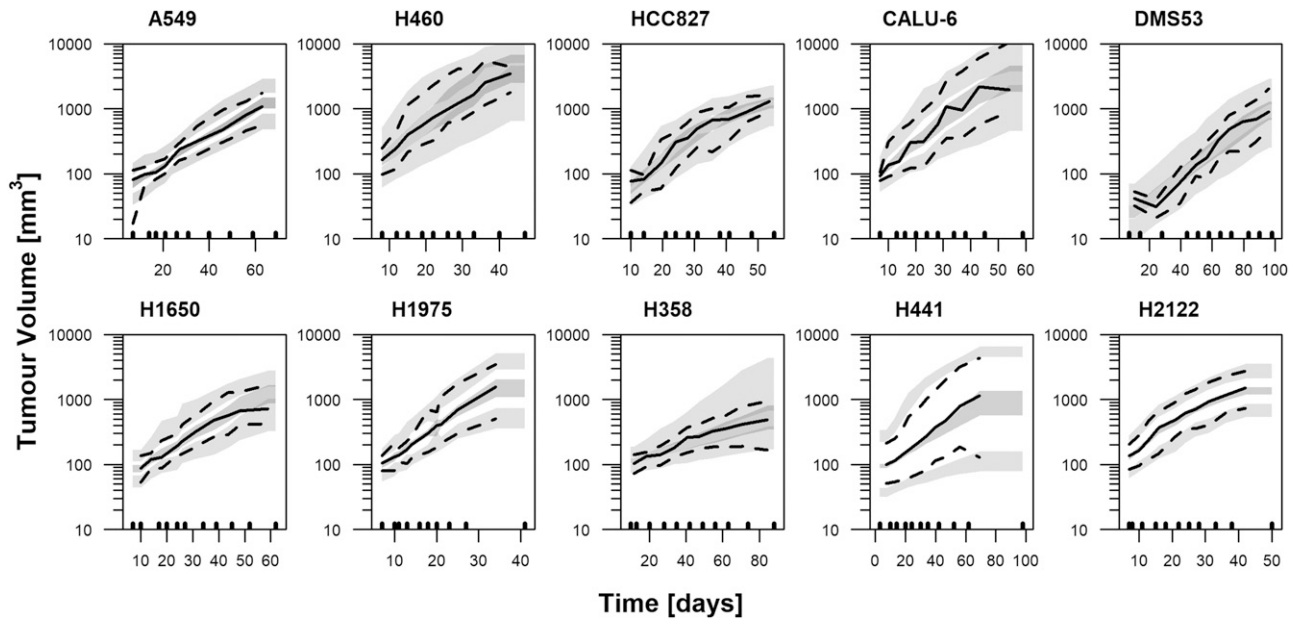
Fig. 1. Visual predictive checks of the unperturbed tumor growth model corresponding to lung tumor cell lines. Lines represent the 2.5th (dashed), 50th (solid), and 97.5th (dashed) percentiles of raw data. Gray areas correspond to the 95% confidence interval of the 2.5th, 50th, and 97.5th percentiles computed from 1000 simulated studies. Black lower marks indicate the used binning intervals for plotting.

cell lines, for which a large number of studies were available ($P > 0.05$).

Design Evaluation and Optimization. POPED was used to evaluate the adequacy of the original experimental designs to precisely estimate all model parameters. In general, good precision (RSE below 10% in average) was obtained for the typical parameters when samples were measured every 2 days (Fig. 3A; Supplemental Table). The largest

TABLE 2
Model parameter estimates corresponding to all the tumor cell lines analyzed
The model parameters are defined in *Materials and Methods*.

Tumor Line	Cell Line	λ_0	λ_1	TV_0	TV_{th}	IAV λ_0	IAV λ_1	IAV TV_0	RUV
		day^{-1}	mm^3/day	mm^3	mm^3	CV%	CV%	CV%	$Log\ mm^3$
Lung	A549	0.051	58.5	48.2	1150	19.8[a]		29.5[a]	0.205
Lung	H460	0.104	237	71.3	2280	25.2		25.9	0.148
Lung	HCC827	0.100	35.1	21.8	351	16.9			0.194
Lung	Calu-6	0.0875	126	49.9	1440	27.2	53.2	12.6	0.147
Lung	DMS53	0.0438	27.6	16.0	630	11.0			0.316
Lung	H1650	0.0575	48.0	50.6	835	20.1		12.8	0.201
Lung	H1975	0.104	188	46.3	1810	14.8			0.202
Lung	H358	0.0204	115	88.6	5640	30.0			0.129
Lung	H441	0.0342	75.0	69.6	2190	68.6		40.1	0.223
Lung	H2122	0.103	42.6	60.2	414	20.4	45.1	23.0	0.151
Renal	ACHN	0.0249	12.8	123	514	68.7	42.5	19.7	0.162
Renal	CAKI-1	0.0265	64.2	102	2420			14.3	0.239
Renal	786-o	0.0489	29.5	82.4	603	22.4	67.7	26.4	0.176
Melanoma	A2058	0.129	55.9	64.7	433	8.20	39.6		0.154
Melanoma	A375	0.0948	101	69.3	1070	39.8	67.2	36.3	0.193
Melanoma	GAK	0.0277	19.0	88.8	686	45.5		17.6	0.191
Melanoma	SK-MEL-30	0.0640	84.8	148	1330	45.9		23.7	0.200
Lymphoma	JEKO-1	0.0825	217	62.9	2630	10.5		19.3	0.284
Lymphoma	OCI-LY-19	0.138	464	30.0	3360	9.70			0.320
Lymphoma	WILL-2	0.203	352	25.3	1730			28.0	0.263
Colon	COLO205	0.106	31.9	65.6	301	25.1	88.1	61.5	0.211
Colon	HCT116	0.121	176	67.5	1460	19.3			0.409
Glioblastoma	U-87-MG	0.0547	51.9	80.4	949	38.9		29.0	0.400
Leukemia	MV411	0.0695	66.5	44.3	957	25.4[b]	72.6[b]	34.7	0.194
Ovarian	A2780	0.103	254	59.0	3660	23.8			0.216
Ovarian	SKVO-3	0.0798	55.4	61.4	694	10.0			0.203
Breast	MB231	0.500 FIX	2.58	81.9	5.00		41.8	23.6	0.122
Pancreas	MIA PaCa-2	0.120	21.6	43.1	180		58.2	15.4	0.114

RUV, residual unexplained variability.
[a]Correlation between IAV λ_0 and IAV TV_0 of −90%.
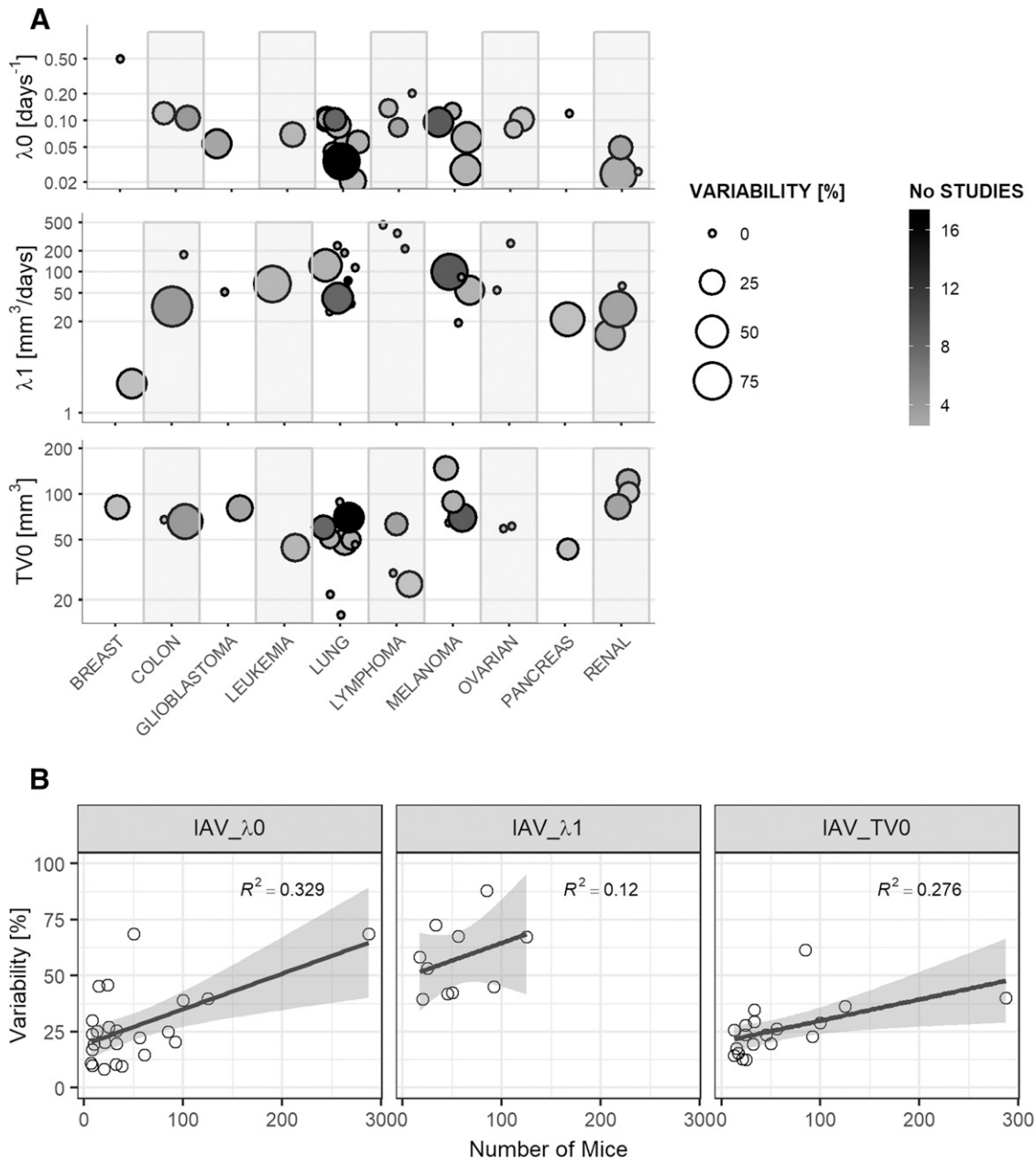[b]Correlation between IAV λ_0 and IAV λ_1 of 100%.

Fig. 2. (A) Graphical representation of the parameter estimates for the different tumor cells lines categorized by the number of studies available and the associated variability magnitude. (B) Relation between the estimate of the different IAVs and the number of mice included in the analysis of each cell line.

imprecision was detected in the linear growth rate (λ_1) for those tumors with a large tumor threshold (TV_{th}). On the other hand, larger RSEs were observed for the variability parameters (Fig. 3A). Precision of the IAV parameters was highly dependent on the number of experimental mice (Fig. 4), regardless of the sampling schema evaluated. For a standard experimental size of 8–10 mice, RSEs between 50% and 60% would be observed. This clear trend was not significant for typical parameters and was independent of the value of IAV precision.

Similarly, the influence of reducing the number of sampling time measurements on parameter precision was evaluated. The twice per week sampling schema (i.e., two samples per week) showed little impact on parameter precision, with a mean absolute loss <5% and a relative loss <25% for all parameters, except for the residual error estimate (i.e., the residual unexplained variability) (Fig. 3B, Supplemental Table 3). Further reducing to one sample per week worsened the precision of the typical parameter estimates, showing mean absolute and relative losses >10%, but also more heterogeneous results (Fig. 3B; Supplemental Table 3) highly dependent on the rate of tumor growth. This loss was less pronounced for the variability parameters, especially for λ_0.

Optimizing the sampling times assuming a design of eight samples per study provided results similar to assuming a fixed twice per week schema (Fig. 3B). A wide range of sampling times, depending on the individual cell growth kinetics, was observed (Supplemental Fig. 2). However, in most of the cases
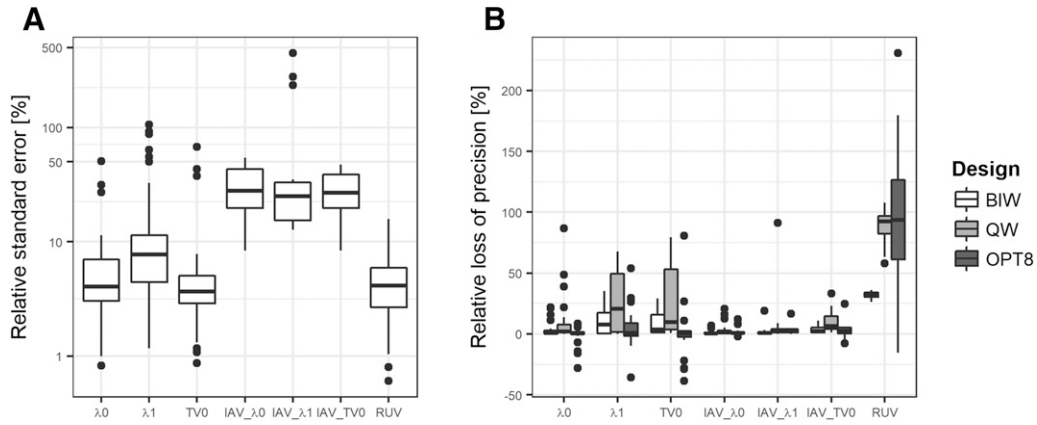
Fig. 3. (A) Box plot of the relative S.E. (RSE) values for the structural and variability model parameters (parameter definition provided in *Materials and Methods*) of the different cell lines assuming samples taken every 2 days (Q2D). (B) Box plot representing the predicted precision when evaluating the sampling designs twice per week (BIW), once per week (QW), or the optimized sampling schema (OPT8) with respect to the Q2D reference design. The box represents the interquartile range (IQR) and the whiskers expand up to 1.5 times the IQR range. Dots represent outliers; RUV denotes residual unexplained variability.

clusters of sampling points around the earliest sampling time point, the tumor threshold (TV_{th}, when relevant), and the latest possible collection sampling time were observed, as illustrated for three cell lines in Fig. 5.

Discussion

Although there is an on-going discussion regarding the capability of xenograft mouse models to generate meaningful data for human extrapolations, they are frequently used to evaluate the efficacy of anticancer drugs early in discovery and development. To overcome some of the limitations of these approaches, mathematical models have been proposed as a tool to derive meaningful parameters independent of the experimental settings, which can be later used to optimize upcoming experiments or guide dose rationale in humans (Simeoni et al., 2013).

In this work, tumor growth dynamics of several tumor cell lines representing different tumor types were modeled in the
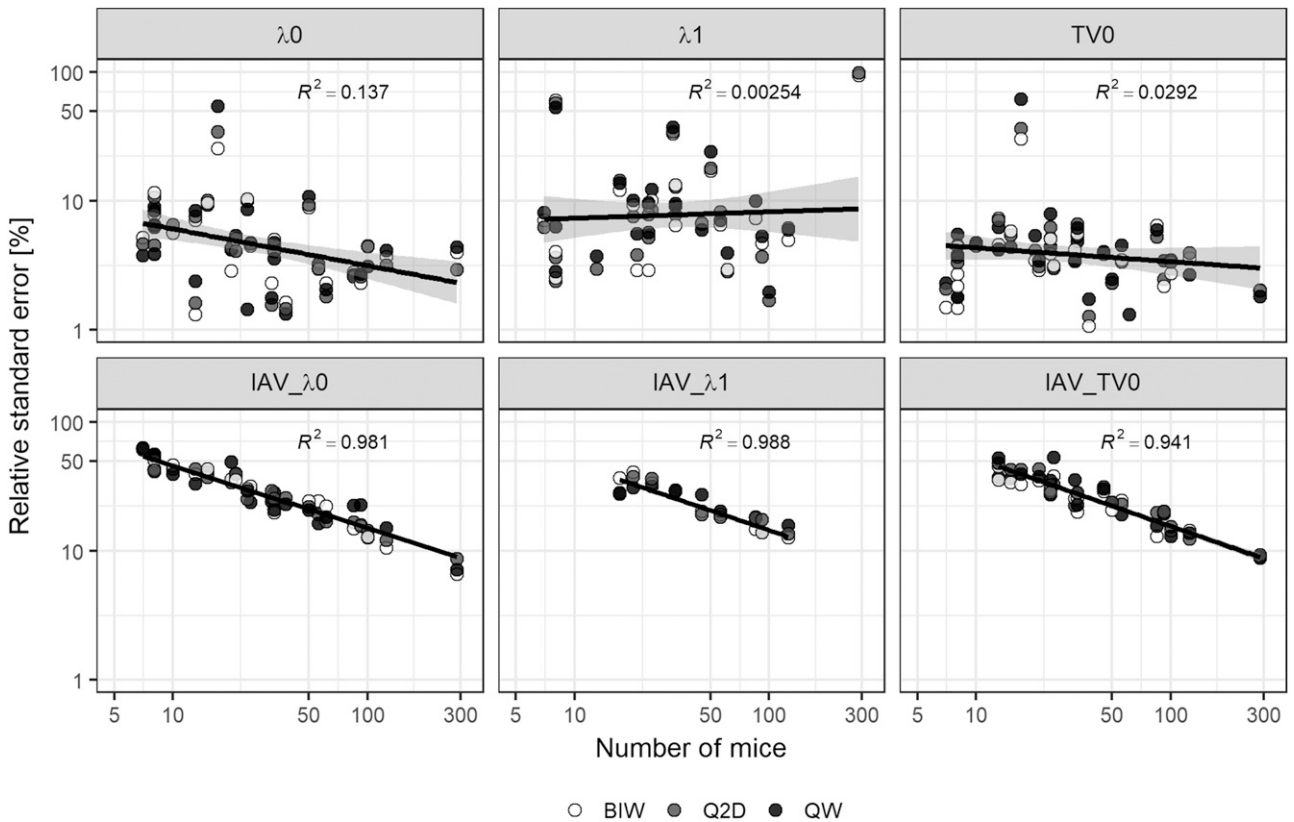

Fig. 4. Relative S.E. values of the typical (upper panels) or IAV (lower panels) model parameters vs. the number of mice included in the analysis of the different cell lines. BIW, tumor size samples taken twice per week; Q2D, tumor size samples taken every two days; QW, tumor size samples taken every week.
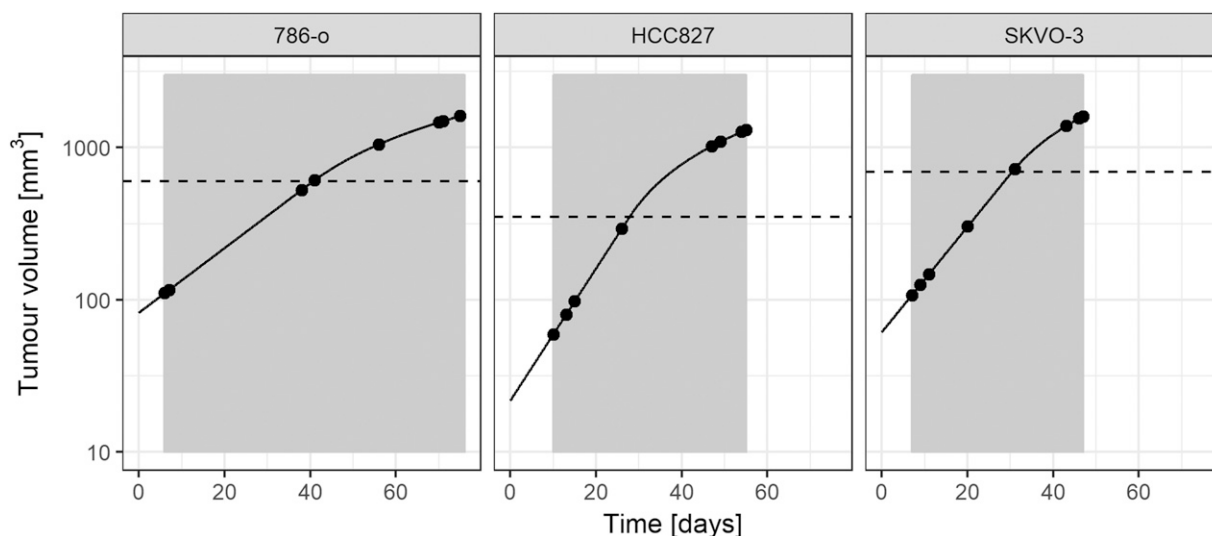
Fig. 5. TV profiles for three selected cell lines. Lines represent model predictions and dots represent the optimized sampling times.

absence of drug using the model proposed by Simeoni et al. (2004). This model was selected as a common structure to systematically analyze our data given its flexibility and proven capability to characterize tumor growth data from different xenograft mouse experiments, as supported by its extended literature use. This work adds to previous efforts characterizing tumor growth dynamics with a common model (Wong et al., 2012; Delgado-SanMartin et al., 2015), and represents the first report where NLME modeling and optimal design techniques have been undertaken to quantitatively evaluate a large set of tumor cell lines.

The model proposed by Simeoni et al. (2004) allowed for an adequate description of the unperturbed tumor growth under the different scenarios analyzed. The obtained exponential growth rates (λ_0) were in good agreement with published estimates available for some cell lines (Simeoni et al., 2004; Rocchetti et al., 2005, 2007; Haddish-Berhane et al., 2013; Terranova et al., 2013; Tate et al., 2014, 2016). It should be highlighted that in some cases data did not fully support the two mechanisms of tumor growth. In those situations, model structure could have been simplified to consider just linear or exponential growth resulting in little impact on the rest of the parameter estimates, as was done for the case of MB211. The maximum tumor burden allowed or dropout information was not available, and therefore was not included in the analysis. This could have potentially induced some parameter estimation bias, especially for λ_1 (Martin et al., 2016; Pierrillas et al., 2016). However, given the large tumor size values observed in our experiments, the large number of mice available for some studies, and the lack of obvious trends or misspecifications observed for the larger tumor size predictions (Fig. 1; Supplemental Fig. 1) we believe that considering the upper welfare limit in the analysis would have had little impact.

Given the availability of several independent experiments for the same cell line, inclusion of ISV in addition to IAV was explored. ISV did not provide further explanation of parameter variability, supporting study reproducibility. Therefore, this work provides valuable information in relation to anticipating the dynamics of tumor growth in xenograft mice and supporting experimental design.

The dynamics of tumor growth in the absence of drug treatment varied widely across the 28 different cell lines. The lower growth rate values were observed for the renal tumor type. However, no additional common patterns could be detected across cell lines for the same type of tumor. Indeed, a large range of tumor growth rate values was observed for lung cancer cell lines. This variability observed across cell lines suggests that information from multiple cell lines obtained during preclinical studies should be integrated to improve extrapolation of results to clinical setting, rather than relying on single cell line experiments. Similar results have been recently illustrated by García-Cremades et al. (2016) for ovarian and pancreatic cancer. A natural extension of this work would be to collect genetic, mutation, and tumor histology information of different cell lines and perform a bioinformatic analysis to identify (if any) common patterns that could account for similarities and differences in growth rates across the different cell types. This information could be potentially used to anticipate growth dynamics of new cell types based on their genetic signature.

An essential aspect when interpreting modeling results is the reliability of the obtained parameter estimates, i.e., large parameter imprecision translates into high uncertainty in model predictions and therefore limited extrapolation capability. In this regard, the experimental design plays an essential role in parameter estimates (al-Banna et al., 1990). Measuring tumor size twice per week allowed for a precise estimation of typical model parameters in all evaluated cell lines, with small precision loss compared with measuring every 2 days. However, further reducing the sampling schema to once per week showed a less consistent impact on precision loss with lower impact on parameter precision in slow growing tumors compared with fast growing tumors. The results from this evaluation design exercise showed that the underlying properties of the cell line can be used to guide experimental design, in terms of sampling intensity and/or size of the experiment.

Among the unperturbed growth model parameters, special attention should be paid to λ_0, since this parameter is potentially used to predict the target human exposure of a

candidate drug (Rocchetti et al., 2007). Except for MIA PaCa-2 cell line, where linear rather than exponential growth was the predominant mechanism, the RSE of λ_0 was always below 20%, regardless of the experimental setting (e.g., number of mice and parameter value) and sampling schema. These observations illustrate the adequacy of the NLME approach to precisely characterize and rank the typical exponential growth rate of the different cell lines. Another important aspect to consider, due to its impact on the confidence interval around the predicted tumor size profiles, is the precision of the estimated variability. The accuracy of the IAV parameters was largely driven by the number of mice rather than the sampling schema or the IAV magnitude, with approximately 50 animals required for RSE <50%. The NLME model represents a suitable approach to achieve such large sample size, since it enables the integration of data from multiple experiments and even control and treatment groups. In addition, the parameters reported in this study could be used as prior information to further reduce the size of control groups.

One step further in the application of optimal design theory is the optimization of the selected sampling times, i.e., identifying those times that, given a model structure and some design constrains, provide the maximum possible information for the estimation of the different model parameters. Optimal design strategies applied to xenograft data have been recently proposed by Lestini et al. (2016). Their results showed how optimal sampling schemas can improve the information obtained from experimental data and highlight the importance of early and late sampling for adequate parameter characterization. Comparable results were obtained in our analysis of control groups, identifying clusters of points at early sampling times (informed about TV_0), late sampling times (informed about the linear growth), or TV_{th} (exponential and linear growths) (Fig. 5; Supplemental Fig. 2). Similarities or differences between optimal times across cell types are therefore explained by the similarities or differences in their growth behavior. Nonetheless, evaluation of experimental designs that can be systematically applied under different conditions might be of more relevance than optimizing the sampling times for each of the cell lines, especially given the low cost and time consumption for obtaining tumor size measurements, and the convenience of standard designs in terms of experimental organization.

In conclusion, a common structure based on the model proposed by Simeoni et al. (2004) characterized well the typical and dispersion tendencies of the longitudinal TV data obtained from a wide range of tumor cell lines providing a robust and general platform: 1) to quantify and compare tumor growth, and possibly 2) to design future treatment experiments based on information from treated groups, in vitro data regarding drug potency, and drug pharmacokinetics. Experimental designs measuring tumor size twice weekly, or even once a week for slow growing tumors, are sufficient for accurate characterization of the tumor growth dynamics. However, analyses of pooled experiments to increase the number of mice are required for adequate quantification of variability. The approach presented here can be easily extrapolated also to drug studies, where modeling and simulation together with design optimization is anticipated to have a more relevant role given the greater experimental study alternatives, e.g., dose level, dose intensity, or group size.

References

al-Banna MK, Kelman AW, and Whiting B (1990) Experimental design and efficient parameter estimation in population pharmacokinetics. *J Pharmacokinet Biopharm* **18**:347–360.

Bueno L, de Alwis DP, Pitou C, Yingling J, Lahn M, Glatt S, and Trocóniz IF (2008) Semi-mechanistic modelling of the tumour growth inhibitory effects of LY2157299, a new type I receptor TGF-β kinase antagonist, in mice. *Eur J Cancer* **44**:142–150.

Cramér H (1946) *Methods of Mathematical Statistics*, Princeton University Press, Princeton, NJ.

Delgado-SanMartin JA, Hare JI, de Moura AP, and Yates JW (2015) Oxygen-driven tumour growth model: a pathology-relevant mathematical approach. *PLOS Comput Biol* **11**:e1004550.

Foracchia M, Hooker A, Vicini P, and Ruggeri A (2004) POPED, a software for optimal experiment design in population kinetics. *Comput Methods Programs Biomed* **74**:29–46.

Garcia-Cremades M, Pitou C, Iversen PW, and Troconiz IF (2016) A comparison of different model-based approaches to scale preclinical to clinical tumour growth inhibition in gemcitabine-treated pancreatic cancer, in Abstracts of the Annual Meeting of the Population Approach Group in Europe, Abstract 5704.

Haddish-Berhane N, Shah DK, Ma D, Leal M, Gerber HP, Sapra P, Barton HA, and Betts AM (2013) On translation of antibody drug conjugates efficacy from mouse experimental tumors to the clinic: a PK/PD approach. *J Pharmacokinet Pharmacodyn* **40**:557–571.

Hay M, Thomas DW, Craighead JL, Economides C, and Rosenthal J (2014) Clinical development success rates for investigational drugs. *Nat Biotechnol* **32**:40–51.

Jonsson EN and Karlsson MO (1999) Xpose—an S-PLUS based population pharmacokinetic/pharmacodynamic model building aid for NONMEM. *Comput Methods Programs Biomed* **58**:51–64.

Lestini G, Mentré F, and Magni P (2016) Optimal design for informative protocols in xenograft tumor growth inhibition experiments in mice. *AAPS J* **18**:1233–1243.

Lindbom L, Pihlgren P, and Jonsson EN (2005) PsN-Toolkit—a collection of computer intensive statistical methods for non-linear mixed effect modeling using NONMEM [published correction appears in *Comput Methods Programs Biomed* (2005) 80: 277]. *Comput Methods Programs Biomed* **79**:241–257.

Martin EC, Aarons L, and Yates JWT (2016) Accounting for dropout in xenografted tumour efficacy studies: integrated endpoint analysis, reduced bias and better use of animals. *Cancer Chemother Pharmacol* **78**:131–141

Ocana A, Pandiella A, Siu LL, and Tannock IF (2010) Preclinical development of molecular-targeted agents for cancer. *Nat Rev Clin Oncol* **8**:200–209.

Parra-Guillen ZP, Berraondo P, Ribba B, and Trocóniz IF (2013) Modeling tumor response after combined administration of different immune-stimulatory agents. *J Pharmacol Exp Ther* **346**:432–442.

Pierrillas PB, Tod M, Amiel M, Chenel M, and Henin E (2016) Improvement of parameter estimations in tumor growth inhibition models on xenografted animals: a novel method to handle the interval censoring caused by measurement of smaller tumors. *AAPS J* **18**:404–415.

Radhakrishna Rao C (1945) Information and the accuracy attainable in the estimation of statistical parameters. *Bull Calcutta Math Soc* **37**:81–91.

Rocchetti M, Poggesi I, Germani M, Fiorentini F, Pellizzoni C, Zugnoni P, Pesenti E, Simeoni M, and De Nicolao G (2005) A pharmacokinetic-pharmacodynamic model for predicting tumour growth inhibition in mice: a useful tool in oncology drug development. *Basic Clin Pharmacol Toxicol* **96**:265–268.

Rocchetti M, Simeoni M, Pesenti E, De Nicolao G, and Poggesi I (2007) Predicting the active doses in humans from animal studies: a novel approach in oncology. *Eur J Cancer* **43**:1862–1868.

Sausville EA and Burger AM (2006) Contributions of human tumor xenografts to anticancer drug development. *Cancer Res* **66**:3351–3354 [discussion 3354].

Simeoni M, De Nicolao G, Magni P, Rocchetti M, and Poggesi I (2013) Modeling of human tumor xenografts and dose rationale in oncology. *Drug Discov Today Technol* **10**:e365–e372.

Simeoni M, Magni P, Cammia C, De Nicolao G, Croci V, Pesenti E, Germani M, Poggesi I, and Rocchetti M (2004) Predictive pharmacokinetic-pharmacodynamic modeling of tumor growth kinetics in xenograft models after administration of anticancer agents. *Cancer Res* **64**:1094–1101.

Tate SC, Burke TF, Hartman D, Kulanthaivel P, Beckmann RP, and Cronier DM (2016) Optimising the combination dosing strategy of abemaciclib and vemurafenib in BRAF-mutated melanoma xenograft tumours. *Br J Cancer* **114**:669–679.

Tate SC, Cai S, Ajamie RT, Burke T, Beckmann RP, Chan EM, De Dios A, Wishart GN, Gelbert LM, and Cronier DM (2014) Semi-mechanistic pharmacokinetic/pharmacodynamic modeling of the antitumor activity of LY2835219, a new

cyclin-dependent kinase 4/6 inhibitor, in mice bearing human tumor xenografts. *Clin Cancer Res* **20**:3763–3774.

Terranova N, Germani M, Del Bene F, and Magni P (2013) A predictive pharmacokinetic-pharmacodynamic model of tumor growth kinetics in xenograft mice after administration of anticancer agents given in combination. *Cancer Chemother Pharmacol* **72**:471–482.

Wong H, Choo EF, Alicke B, Ding X, La H, McNamara E, Theil FP, Tibbitts J, Friedman LS, Hop CE, et al. (2012) Antitumor activity of targeted and cytotoxic agents in murine subcutaneous tumor models correlates with clinical response. *Clin Cancer Res* **18**:3846–3855

Zhang L, Sinha V, Forgue ST, Callies S, Ni L, Peck R, and Allerheiligen SRB (2006) Model-based drug development: the road to quantitative pharmacology. *J Pharmacokinet Pharmacodyn* **33**:369–393.

Address correspondence to: Johan E. Wallin, Eli Lilly and Company, Gustav III Boulevard 42, P.O Box 721, SE 169 27 Solna, Sweden. E-mail: wallinjo@lilly.com

Title: Systematic modelling and design evaluation of unperturbed tumour dynamics in xenografts

Authors and affiliations:

Zinnia P Parra-Guillen, Victor Mangas-Sanjuan, Maria Garcia-Cremades, Iñaki F Troconiz, Gary Mo, Celine Pitou, Philip W Iversen, Johan E Wallin.

ZPP-G,VM-S,MG-C and IFT: Pharmacometrics & Systems Pharmacology Research Unit, Department of Pharmacy and Pharmaceutical Technology, School of Pharmacy and Nutrition, University of Navarra, Pamplona, Spain.

ZPP-G,VM-S,MG-C and IFT: IdiSNA, Navarra Institute for Health Research, Pamplona, Spain

GM, CP and JEW: Global PK/PD & Pharmacometrics, Eli Lilly and Company.

PWI: Lilly Research Laboratories, Eli Lilly and Company.

VM-S: Current affiliation: Department of Pharmacy and Pharmaceutical Technology and Parasitology, University of Valencia, Valencia, Spain; Molecular Recognition and Technological Development, Polytechnic University-University of Valencia, Valencia, Spain

MG-C: Current affiliation: Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, California, USA

ZPP-G and VM-S share first authorship

Supplementary Figure 1
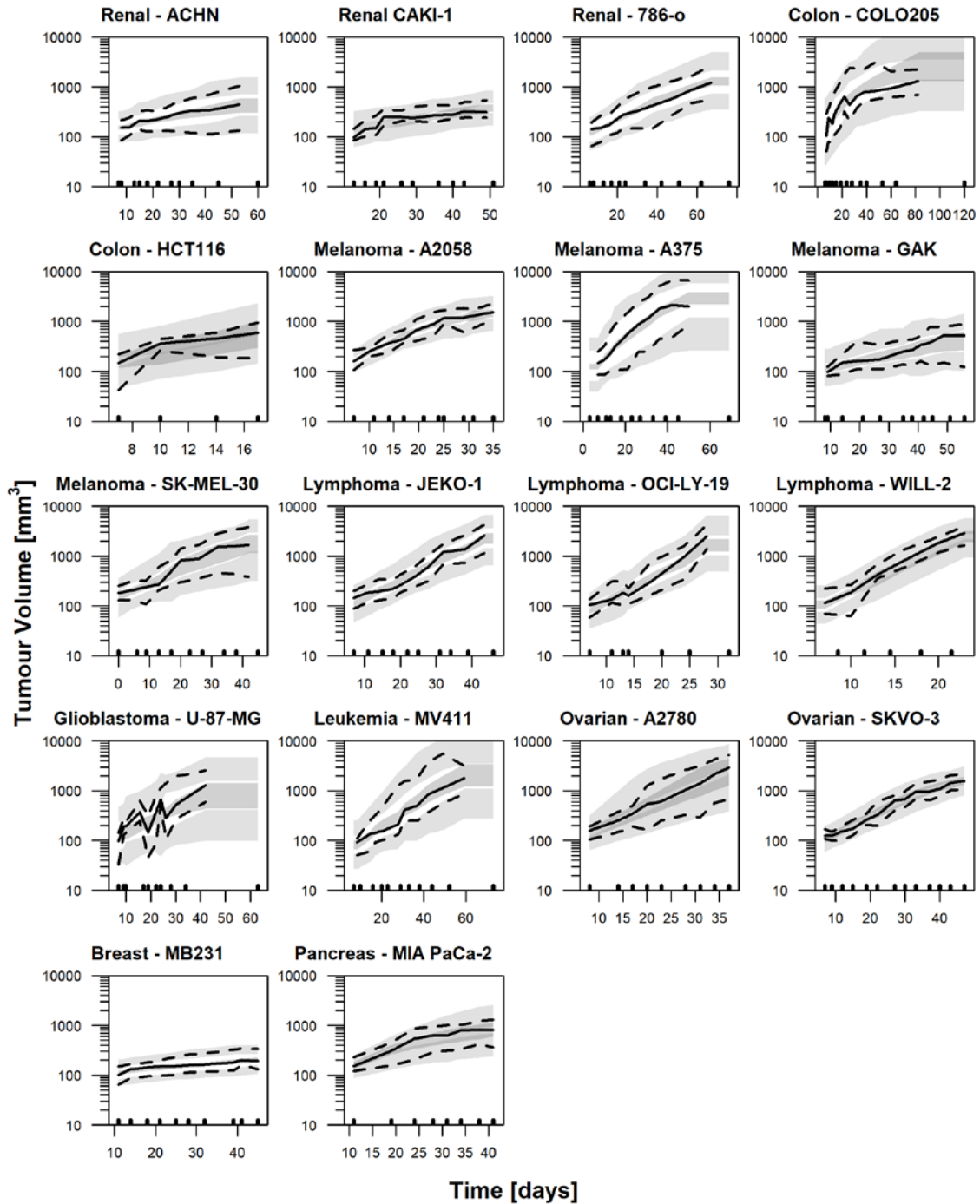
Supplementary Figure 1: Visual predictive checks of the unperturbed tumour growth model corresponding to all tumour cell lines but lung. Lines represent the 2.5th (dashed), 50th (solid) and 97.5th (dashed) percentiles of raw data. Grey areas correspond to the 95% confidence interval of the 2.5th, 50th and 97.5th percentiles computed from 1000 simulated studies. Yellow lower marks indicate the used binning intervals for plotting.

Supplementary Figure 2
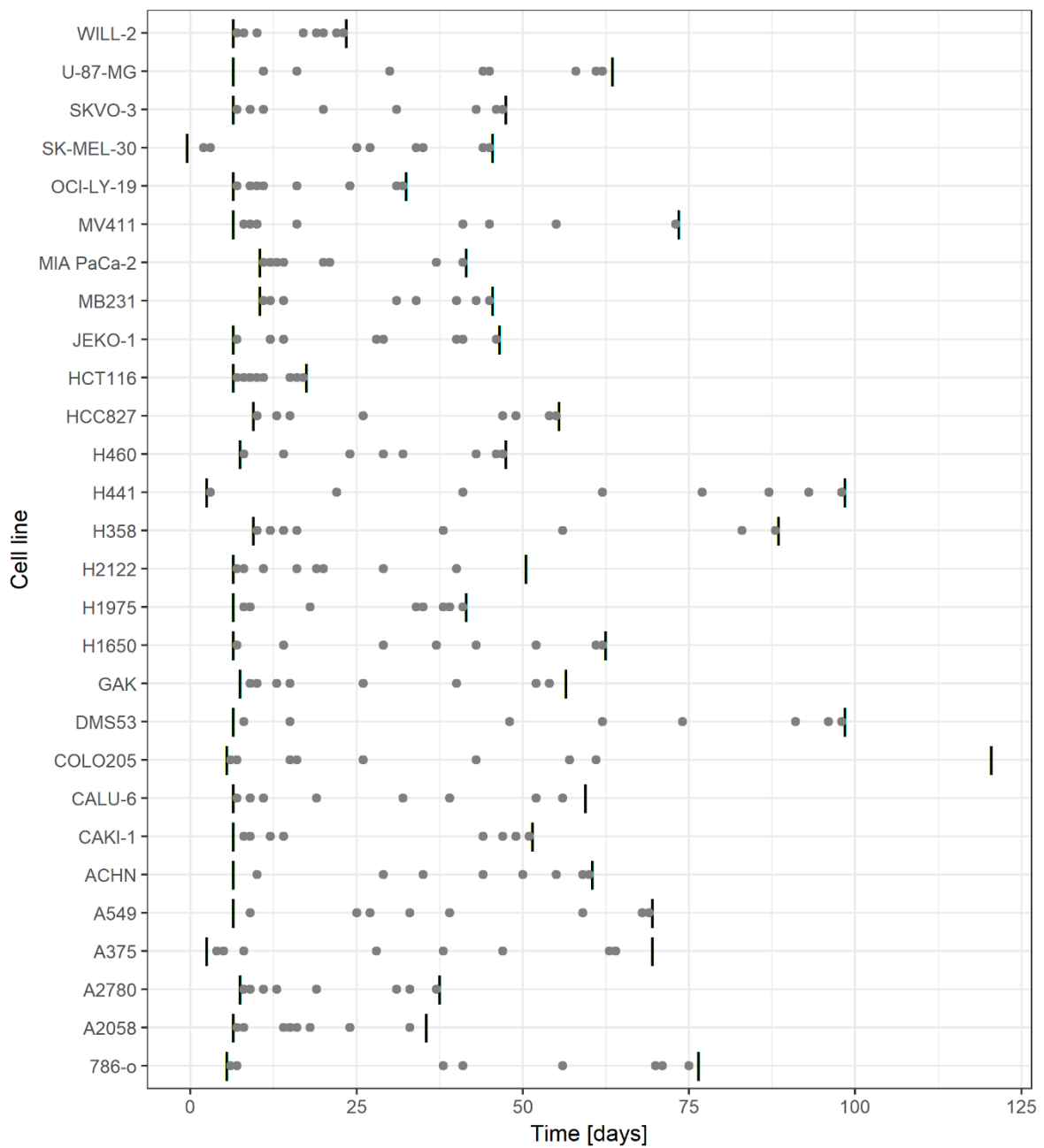
Supplementary Figure 2: Optimal sampling time points per cell lines. Points represent optimised sampling times for each cell line assuming 8 samples taken; vertical lines delimit the allowed sampling window based on the original design.

Supplementary Table 1: Relative standard errors reported by POPED for the every two days sampling schema

Tumour Line	Cell Line	Parameter						
		λ_0	λ_1	TV_0	$IAV\ \lambda_0$	$IAV\ \lambda_1$	$IAV\ W_0$	RUV
Lung	A549	3.44	5.64	5.09	24.88	n.a	25.13	2.25
	H460	6.94	2.73	7.36	39.69	n.a	42.19	4.50
	HCC827	6.20	2.03	3.68	50.28	n.a	n.a	5.21
	DMS53	4.20	5.74	1.33	53.63	n.a	n.a	3.94
	H1650	4.39	3.10	3.00	31.30	n.a	35.30	2.97
	H1975	1.91	3.02	0.86	18.19	n.a	n.a	2.20
	H358	10.46	*	1.31	50.12	n.a	n.a	4.00
	H441	3.67	87.25	2.29	8.35	n.a	8.40	0.61
	H2122	2.28	4.52	2.69	16.72	14.95	18.46	1.65
Renal	ACHN	8.83	19.63	2.84	20.12	233.98	21.07	2.00
	CAKI-1	1.64	*	4.18	n.a	n.a	40.07	4.18
	786-o	3.01	8.25	3.54	19.51	19.11	19.70	1.64
Melanoma	A2058	3.18	8.70	3.90	35.31	32.57	n.a	4.38
	A375	3.44	5.48	3.19	12.73	12.71	12.98	1.14
	GAK	11.27	*	4.67	36.97	n.a	39.33	3.81
	SK-MEL-30	8.95	3.38	4.87	29.00	n.a	29.85	3.08
Lymphoma	JEKO-1	1.90	28.02	3.53	26.07	n.a	26.92	2.87
	OCI-LY-19	1.62	*	1.17	23.14	n.a	n.a	3.18
	WILL-2	0.83	7.65	6.10	n.a	n.a	29.32	5.10
Colon	COLO205	2.95	8.22	6.33	18.36	15.35	16.14	1.03
	HCT116	6.77	*	4.72	45.57	n.a	n.a	10.00
Leukemia	MV411	4.38	11.38	5.95	24.89	24.85	25.24	2.21
Ovarian	A2780	8.37	50.34	2.57	50.09	n.a	n.a	6.45
	SKVO-3	3.86	3.65	2.65	50.99	n.a	n.a	5.59
Breast	MB231	n.a	6.30	3.84	n.a	23.03	25.45	2.63
Pancreas	MIA PaCa-2	27.07	13.20	37.69	n.a	34.64	40.29	4.58

*Parameter fixed. n.a.: not applicable

Supplementary Table 2: Relative standard errors reported by POPED for the twice per week sampling schema

Tumour Line	Cell Line	Parameter						
		λ_0	λ_1	TV_0	IAV λ_0	IAV λ_1	IAV W_0	RUV
Lung	A549	3.44	5.64	5.09	24.88	n.a	25.13	2.25
	H460	6.94	2.73	7.36	39.69	n.a	42.19	4.50
	HCC827	6.20	2.03	3.68	50.28	n.a	n.a	5.21
	DMS53	4.20	5.74	1.33	53.63	n.a	n.a	3.94
	H1650	4.39	3.10	3.00	31.30	n.a	35.30	2.97
	H1975	1.91	3.02	0.86	18.19	n.a	n.a	2.20
	H358	10.46	*	1.31	50.12	n.a	n.a	4.00
	H441	3.67	87.25	2.29	8.35	n.a	8.40	0.61
	H2122	2.28	4.52	2.69	16.72	14.95	18.46	1.65
Renal	ACHN	8.83	19.63	2.84	20.12	233.98	21.07	2.00
	CAKI-1	1.64	*	4.18	n.a	n.a	40.07	4.18
	786-o	3.01	8.25	3.54	19.51	19.11	19.70	1.64
Melanoma	A2058	3.18	8.70	3.90	35.31	32.57	n.a	4.38
	A375	3.44	5.48	3.19	12.73	12.71	12.98	1.14
	GAK	11.27	*	4.67	36.97	n.a	39.33	3.81
	SK-MEL-30	8.95	3.38	4.87	29.00	n.a	29.85	3.08
Lymphoma	JEKO-1	1.90	28.02	3.53	26.07	n.a	26.92	2.87
	OCI-LY-19	1.62		1.17	23.14	n.a	n.a	3.18
	WILL-2	0.83	7.65	6.10	n.a	n.a	29.32	5.10
Colon	COLO205	2.95	8.22	6.33	18.36	15.35	16.14	1.03
	HCT116	6.77	*	4.72	45.57	n.a	n.a	10.00
Leukemia	MV411	4.38	11.38	5.95	24.89	24.85	25.24	2.21
Ovarian	A2780	8.37	50.34	2.57	50.09	n.a	n.a	6.45
	SKVO-3	3.86	3.65	2.65	50.99	n.a	n.a	5.59
Breast	MB231	n.a	6.30	3.84	n.a	23.03	25.45	2.63
Pancreas	MIA PaCa-2	27.07	13.20	37.69	n.a	34.64	40.29	4.58

*Parameter fixed. n.a.: not applicable

Supplementary Table 3: Relative standard errors reported by POPED for the every week sampling schema

Tumour Line	Cell Line	Parameter						
		λ_0	λ_1	TV_0	IAV λ_0	IAV λ_1	IAV W_0	RUV
Lung	A549	3.48	7.77	5.19	25.35	n.a	26.11	4.35
	H460	7.05	4.32	7.82	40.41	n.a	46.90	8.76
	HCC827	6.62	3.37	5.74	50.97	n.a	n.a	9.45
	DMS53	4.29	9.62	2.34	54.06	n.a	n.a	7.41
	H1650	4.47	5.01	3.36	31.96	n.a	43.42	5.82
	H1975	1.96	3.93	1.36	18.36	n.a	n.a	4.05
	H358	10.56	*	2.21	50.32	n.a	n.a	7.22
	H441	3.67	105.39	2.30	8.36	n.a	8.50	1.16
	H2122	2.50	4.57	3.00	18.82	15.27	21.57	3.24
Renal	ACHN	8.91	26.71	2.99	20.30	447.51	22.99	4.06
	CAKI-1	2.44	*	4.53	n.a	n.a	41.66	7.41
	786-o	3.10	8.34	3.67	20.55	19.52	21.01	3.34
Melanoma	A2058	3.88	8.92	4.94	42.55	34.09	n.a	9.11
	A375	3.46	5.50	3.26	12.88	12.82	13.51	2.24
	GAK	11.39	*	4.96	37.60	n.a	44.17	7.45
	SK-MEL-30	9.00	4.88	5.02	29.21	n.a	31.55	5.89
Lymphoma	JEKO-1	1.99	32.62	3.76	27.45	n.a	29.88	5.57
	OCI-LY-19	1.67		1.78	23.46	n.a		5.74
	WILL-2	1.15	8.94	6.50		n.a	30.01	8.33
Colon	COLO205	3.20	8.23	6.48	21.35	15.37	16.62	1.98
	HCT116	7.00	*	5.81	46.22	n.a	n.a	15.81
Leukemia	MV411	4.42	11.46	6.08	25.38	25.20	26.34	4.35
Ovarian	A2780	8.49	63.65	4.10	50.22	n.a	n.a	11.18
	SKVO-3	4.40	5.64	4.21	53.34	n.a	n.a	10.21
Breast	MB231	n.a	6.70	4.27	n.a	25.04	30.51	5.20
Pancreas	MIA PaCa-2	50.60	13.45	67.63	n.a	34.99	47.07	8.57

*Parameter fixed. n.a.: not applicable