

---

## Research Article

---

# 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

Philippe B. Pierrillas,<sup>1,3,5</sup> Michel Tod,<sup>1,2</sup> Magali Amiel,<sup>3</sup> Marylore Chenel,<sup>4</sup> and Emilie Henin<sup>1</sup>

Received 10 July 2015; accepted 15 December 2015; published online 12 January 2016

**Abstract.** The purpose of this study was to explore the interval censoring induced by caliper measurements on smaller tumors during tumor growth experiments in preclinical studies and to show its impact on parameter estimations. A new approach, the so-called interval-M3 method, is proposed to specifically handle this type of data. Thereby, the interval-M3 method was challenged with different methods (including classical methods for handling below quantification limit values) using Stochastic Simulation and Estimation process to take into account the censoring. In this way, 1000 datasets were simulated under the design of a typical of tumor growth study in xenografted mice, and then, each method was used for parameter estimation on the simulated datasets. Relative bias and relative root mean square error (relative RMSE) were consequently computed for comparison purpose. By not considering the censoring, parameter estimations appeared to be biased and particularly the cytotoxic effect parameter,  $k_2$ , which is the parameter of interest to characterize the efficacy of a compound in oncology. The best performance was noted with the interval-M3 method which properly takes into account the interval censoring induced by caliper measurement, giving overall unbiased estimations for all parameters and especially for the antitumor effect parameter (relative bias = 0.49%, and relative RMSE = 4.06%).

**KEY WORDS:** below quantification limit; interval censoring; interval-M3 method; simultaneous modeling continuous and categorical data; xenograft model.

## INTRODUCTION

Censored data is a practical and overriding question when analyzing data. There is no unique and simple answer on how to handle these data (1). If not considered properly, censoring may cause bias and misspecification during data analysis (2), so it is important to apprehend this problem and to propose strategies to get the best of experimental data.

---

**Electronic supplementary material** The online version of this article (doi:10.1208/s12248-015-9862-1) contains supplementary material, which is available to authorized users.

<sup>1</sup>EMR 3738, Ciblage Thérapeutique en Oncologie, Faculté de Médecine et de Maïeutique Lyon-Sud Charles Mérieux, Université Claude Bernard Lyon 1, Oullins, France.

<sup>2</sup>Pharmacie, Hôpital de la Croix Rousse, Hospices Civils de Lyon, Lyon, France.

<sup>3</sup>Centre de Pharmacocinétique et Métabolisme, Technologie Servier, Orléans, France.

<sup>4</sup>Division of Clinical Pharmacokinetics and Pharmacometrics, Institut de Recherches Internationales Servier, Suresnes, France.

<sup>5</sup>To whom correspondence should be addressed. (e-mail: philippe.pierrillas@gmail.com)

Traditionally for pharmacokinetic studies and more recently for pharmacodynamic studies, a major source of censored data is concentrations below the limit of quantification (LOQ) of the analytical method. In this context, several approaches were proposed to handle observed concentrations reported as below quantification limit (BQL) (3–5). These approaches have various degrees of complexity. The simpler method consists in discarding BQL values (the so-called M1 method) and applying the method of maximizing the conditional likelihood only for the observations above the LOQ. A more complex approach consists in the simultaneous modeling of above and below quantification limit observations, considering for the latter, the likelihood of the observation to be BQL given the predicted value (the so-called M3 method). Not considering BQL observations in PK studies can produce biased parameter estimations and biased interpretations. In tumor growth inhibition experiments, we were also facing a kind of censoring specifically induced by the measurement technique.

The xenografted animal model (generally on mice or rats) (6) is a widely used experimental model to evaluate anticancer effect of new compounds and possibly predict the efficacy in human thanks to translational approaches (7).

During tumor growth experiment on xenografted animals, the tumor size is measured by a technical operator using a caliper, and reported as the measures of two diameters,

often converted into tumor volume or tumor weight. One of the issues with this type of experimentation is the precision of the caliper measurement, reasonable for larger diameters (precision  $\pm 0.1$  mm) but questionable for smaller ones, whose measurement is often limited by operational difficulties (skin thickness, difficulties to delimit the tumor and to maintain correctly the tumor in the caliper ...), and only rounded values for tumor diameters (1, 2, 3, ... 5 mm) are reported. When converting into tumor volumes, rounded values for small tumor diameters result into discrete values and consequently to interval-censored tumor volumes.

The appreciation of the antitumor effect of a drug being essentially informed when the tumor size is at the lowest; the imprecision of small tumor sizes can largely affect the estimation of the drug effect.

A new approach, the so-called interval-M3 method, adapted to the censoring induced by the caliper measurement is being proposed and evaluated. This article reports the comparison of the performance of several methods (including methods discarding data, methods considering a single limit in the censoring process, and also methods considering interval censoring in its founding) to take into account the impact of interval censoring in tumor growth inhibition modeling exercise. We simulated replicated data of a tumor growth inhibition studies in xenografted mice and estimated model parameters under 12 different approaches for censored data in order to better estimate model parameters and thus to better quantify the antitumor effect.

## MATERIAL AND METHODS

### Issues of Interval Censoring in Tumor Sizes

In many tumor growth studies in xenografted animals, the dimensions of the tumors are measured with a caliper, whose precision is generally about 0.1 mm, and tumor volume, expressed in cubic millimeter, is often calculated, assuming the tumor has an ovoid form, following the approximation formula (demonstration in [Appendix](#)) (8,9):

$$\text{Tumor volume (mm}^3\text{)} = \frac{\text{length (mm)} \cdot \text{width}^2\text{(mm}^2\text{)}}{2}$$

However, due to the imprecision induced by skin thickness and palpability of the tumor, when the length and width are less than 5 mm, the measure of the caliper is then rounded with a precision of 0.5 mm, and reported tumor sizes seem to take discrete values, such as  $1 \times 1$ ,  $2 \times 2$ ,  $3 \times 3$ ,  $4 \times 4$ , and  $5 \times 5$  assuming the tumor has a spherical shape.

For example, if we consider a tumor of 2.2 mm long, and a second tumor of 2.8 mm long, the first one will be reported as a  $2 \times 2$  tumor, equivalent to a  $4 \text{ mm}^3$  tumor while the second one will be reported as a  $3 \times 3$  leading to a calculated tumor volume of  $13.5 \text{ mm}^3$ . Similarly, taking into account the censoring problem,  $2 \times 2$  tumors would include diameters between 1.5 and 2.5 mm and tumor volume ranging from  $\frac{1.5 \times 1.5^2}{2} = 1.69 \text{ mm}^3$  to  $\frac{2.5 \times 2.5^2}{2} = 7.8 \text{ mm}^3$ , reported as  $4 \text{ mm}^3$ .

All the possible discrete values (experimental data) of tumor volume reported by the technical operator, according to length and width, and the corresponding "real" (uncensored) tumor volume intervals are summarized in [Table I](#).

This kind of interval censoring may result in biased parameter estimations and inflation of residual error if not correctly handled.

### Handling Interval Censoring in Tumor Size: Comparison Strategy

Stochastic Simulations and Estimations (SSE) methodology was used to compare and evaluate the different methods to handle interval-censored tumor volumes. SSE are composed of two steps: the first step consists in generating 1000 simulated replicates of datasets using a "true" reference model. The second step is an estimation step where both the reference and alternative models are fitted to the replicates. The alternative models are then compared to the reference model in terms of parameter values and precision, estimated in the 1000 replicates.

SSE process was performed with a Nonlinear Mixed Effects approach (10) using the *sse* script of PsN suite (Pearl-speaks to NONMEM) (11,12) and NONMEM software (version 7.3) (13).

In order to perform SSE, the following are to be defined: (i) the experimental study design; (ii) the reference simulation model; (iii) the alternative estimation methods to be compared.

### Definition of the Experimental Study Design

In each replicated study, four groups of 8 xenografted mice were simulated: a control group, and three groups receiving an anticancer drug dose of either 20, 45, or 100  $\mu\text{g}/\text{kg}$ . Tumors were implanted at time  $t=0$  h, and tumor size monitored every 24 h for 1000 h (42 days). Treatment was given orally for 7 days, from  $t=624$  h (26 days) to  $t=792$  h (33 days) every 24 h.

### Reference Model: Pharmacokinetic and Tumor Growth Inhibition Models

#### Model Structure

The pharmacokinetic model is a one-compartment disposition model with both first-order absorption and elimination processes.

$$\frac{dA_1}{dt} = -k_a \cdot A_1$$

$$\frac{dA_2}{dt} = k_a \cdot A_1 - k_{el} \cdot A_2$$

$$C_p = \frac{A_2}{V_c}$$

With  $k_a$  ( $\text{h}^{-1}$ ) the first-order rate of absorption,  $k_{el}$  ( $\text{h}^{-1}$ ) the first-order rate of elimination,  $V_c$  (L/kg) the apparent volume of distribution,  $C_p$  (ng/mL) the plasma concentration,  $A_1$  ( $\mu\text{g}/\text{kg}$ ) and  $A_2$  ( $\mu\text{g}/\text{kg}$ ) the amount of drug respectively in the depot compartment and in the central compartment. At each time of treatment administration, the corresponding dose was added to the amount  $A_1$  in the depot compartment.

**Table I.** Description of All Discrete Values of Tumor Volume and the Corresponding Range of Real Tumor Volume

Non-censored data	Interval-censored data	
Tumor size (mm <sup>3</sup> )	Length x width (mm x mm)	Experimental data (mm <sup>3</sup> )
0		
	0 x 0	0
0.0625	1 x 1	0.5
1.69	2 x 2	4
7.8	3 x 3	13.5
21.4	4 x 4	32
45.5	5 x 5	62.5
83.2		

The unperturbed tumor growth, without treatment, was described using an exponential tumor growth model ( $\lambda$  parameter,  $h^{-1}$ ), reasonable for smaller tumor sizes (14), and initial tumor size  $TG_0$  (mm<sup>3</sup>), as follows:

$$\frac{dTG}{dt} = \lambda \cdot TG$$

$$TG_{t=0} = TG_0$$

The cytotoxic effect of the drug ( $k_2$ , (ng/mL)<sup>-1</sup>.h<sup>-1</sup>) was modeled proportional to plasma concentration  $C_p$  and tumor size  $TG$ , as follows:

$$\frac{dTG}{dt} = \lambda \cdot TG - k_2 \cdot C_p \cdot TG$$

$$TG_{t=0} = TG_0$$

Typical profiles of the model for each dose are presented in Fig. 1.

#### Parameter Distributions in the Population

An exponential model was used for interindividual variability (IIV) on  $k_2$  and  $\lambda$  parameters, which implies an assumed log-normal distribution of the parameters between the 32 individuals in the population.

$$p_i = p_{pop} \cdot \exp(\eta_{p_i})$$

With  $p_{pop}$  the typical value of  $p$  parameter,  $p_i$  the value of  $p$  parameter for the  $i^{\text{th}}$  individual and  $\eta_{p_i}$  a random variable quantifying the variation from the typical (average)

individual. The  $\eta_{p_i}$  are assumed to be normally distributed, with a mean of zero and a variance of  $\omega_p^2$ .

#### Interval Censoring and Residual Error

When simulated tumor volumes were below 62.5 mm<sup>3</sup> (corresponding to a measured 5×5 tumor), the reported values were altered following the aforementioned interval-censoring limitation of caliper measurement for smaller tumors (as shown in Table I and Fig. 2).

Above this limit, the residual unexplained error linking the observed value to the corresponding predicted values was considered as a proportional error model, as follows:

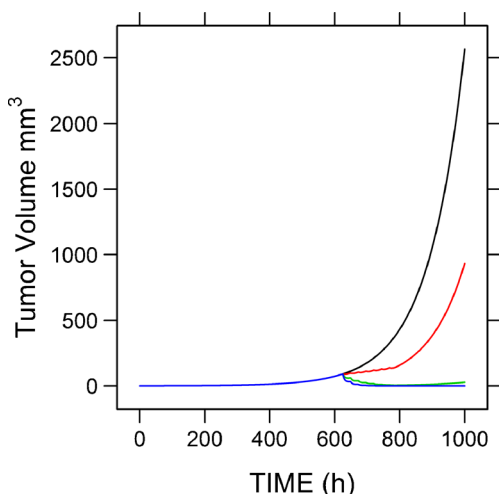
$$y_{ij} = f_{ij} \cdot (1 + \varepsilon_{ij})$$

With  $y_{ij}$ , the  $j^{\text{th}}$  observed tumor volume for the  $i^{\text{th}}$  individual and  $f_{ij}$  is the corresponding model predicted value,  $\varepsilon_{ij}$  a random variable with a normal distribution of mean 0 and standard deviation  $\sigma$ .

Parameter distributions (typical values and variabilities) are presented in Table II. Since PK samples are not drawn in most of tumor growth experiments, PK parameters were fixed to their typical values. The parameters to be estimated during the SSE exercise were the typical values of ( $TG_0$ ,  $\lambda$ ,  $k_2$ ), the variances of ( $\lambda$ ,  $k_2$ ) and the standard deviation of the proportional residual error,  $\sigma$ .

#### Alternative Error Models: Approaches to Handle Interval-Censored Tumor Size

Twelve different methods were evaluated to handle smaller tumor size values (Table III). In all cases, the model



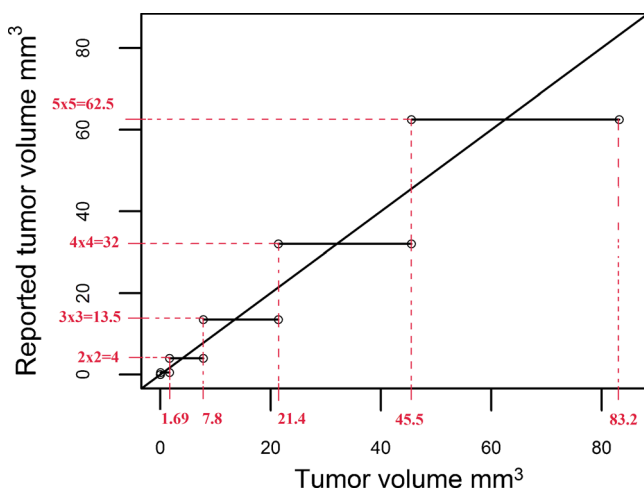
**Fig. 1.** Typical profiles for each dose group (control group in black, 10 µg/kg in red, 45 µg/kg in green, and 100 µg/kg in blue) of the tumor growth model

structure and the model for the interindividual variability of the parameters were the same as the reference model. Only the residual error model and method to handle interval-censored data were challenged. Observations greater than the limit of quantification were fitted using a proportional residual error model, unless otherwise noted. The twelve different methods were as follows:

– (a) Fitting of the whole dataset, not taking into account the censoring, using a combined (additive + proportional) residual error model.

– (b) and (c) ignoring tumor volumes inferior to LOQ (corresponding to the so-called M1 method (3)), defining LOQ to 0.5 and 4 mm<sup>3</sup> respectively (in order to consider several proportions of censored data), using a combined residual error model for observations above LOQ.

– (d) and (e) replacing values below LOQ by LOQ/2 and using for values below LOQ an additive error model with standard deviation equals to LOQ/4 (corresponding to the so-called M5 method (3)), defining LOQ to 0.5 and 4 mm<sup>3</sup> respectively.



**Fig. 2.** Observed versus reported tumor growth volumes: visualization of the interval censoring

– (f) and (g) taking into account the likelihood of the data to be below LOQ (corresponding to the so-called M3 method (3)), defining LOQ to 0.5 and 4 mm<sup>3</sup> respectively.

– (h) considering a quadratic residual error model (being more flexible than a combined residual error model) for observations corresponding to simulated values below 83.2 mm<sup>3</sup> (i.e., 5×5 tumors), and a proportional model for observations above this limit; the variance of the quadratic error can be written  $a + b*f + c*f^2$ , with  $a$  the additive term,  $b$  the first-order term and  $c$  the second-order term, and  $f$  the predicted tumor volume.

– (i) similar to h, with the first-order term  $b$  fixed to 0.

– (j) taking into account the interval censoring ( $1 \times 1 = 0.5$  mm<sup>3</sup>,  $2 \times 2 = 4$  mm<sup>3</sup>, ..., up to  $5 \times 5 = 83.2$  mm<sup>3</sup>) and defining a separate additive error model for each interval, setting the standard deviation to the quarter of the width of the corresponding interval (further referred to as multi-additive method) implying that 95% of the values belong to the interval. For example, observed values of  $3 \times 3 = 13.5$  mm<sup>3</sup>, corresponding to the censoring of the interval [7.8; 21.4 mm<sup>3</sup>], are associated to an additive residual error with a standard deviation fixed to  $\frac{21.4 - 7.8}{4} = 3.4$  mm<sup>3</sup>.

– (k) similar to j but defining a separate proportional error model for each interval (further referred to as multi-proportional method). For example, observed values of  $3 \times 3 = 13.5$  mm<sup>3</sup>, corresponding to the censoring of the interval [7.8; 21.4], are associated to a proportional residual error with a coefficient of variation fixed to  $\frac{3.4}{13.5} = 25.2$  %.

– (l) taking into account the likelihood of the data to be in each censored interval (further referred to as interval-M3 method). Therefore, for observed values below  $5 \times 5 = 83.2$  mm<sup>3</sup>, the probability of the observed tumor volume ( $TG_{obs}$ ) to be in each censored interval has to be defined:

• Below $1 \times 1$ : $P(0 < TG_{obs} < 0.0625)$	• $3 \times 3$ : $P(7.8 < TG_{obs} < 21.4)$
• $1 \times 1$ : $P(0.0625 < TG_{obs} < 1.69)$	• $4 \times 4$ : $P(21.4 < TG_{obs} < 45.5)$
• $2 \times 2$ : $P(1.69 < TG_{obs} < 7.8)$	• $5 \times 5$ : $P(45.5 < TG_{obs} < 83.2)$

Implementation of method (l) in NONMEM is detailed in an [Appendix](#).

## Performance Evaluation

Several criteria were considered to compare and evaluate the performances of the different methods. The ratio of parameter estimates ( $p_{pop}$ ,  $\omega_p^2$  and  $\sigma$ ) to the true value in the  $N=1000$  replicates of the pseudo-experimental study were depicted in box-plots for graphical comparison of the different methods.

Parameters estimated with the 12 methods were also compared in terms of relative bias, and relative Root Mean Squared Error (RMSE).

**Table II.** Parameters of the Simulated Model Used for the Stochastic Simulation and Estimation (SSE) Process

Parameters	Parameter definition	Values	Units
<b>PK model</b>			
$k_a$	First-order absorption rate	1	$h^{-1}$
$k_{el}$	First-order elimination rate	0.1386	$h^{-1}$
$V_c$	Apparent volume of distribution	5	$L \text{ kg}^{-1}$
<b>Tumor growth model</b>			
$TG_0$	Initial tumor size	0.35	$mm^3$
$\lambda$	Tumor growth rate	0.0089	$h^{-1}$
$\omega_\lambda$	Interindividual variability on $\lambda$	15	%
$k_2$	Cytotoxic effect	0.00001	$mL \text{ ng}^{-1} h^{-1}$
$\omega_{k_2}$	Interindividual variability on $k_2$	15	%
$\sigma_{prop}$	Proportional residual error	20	%

$$\text{Relative bias} = 100\% \cdot \frac{1}{N} \sum_i \frac{est_i - true_i}{true_i}$$

$$\text{Relative RMSE} = 100\% \cdot \sqrt{\frac{1}{N} \cdot \sum_i \frac{(est_i - true_i)^2}{true_i}}$$

The rate of successful minimization runs was also used for the comparison of all methods. Finally, the magnitudes of the residual error were compared to judge the impact of the censoring handling of low values on the prediction of larger values of tumor volumes.

Categorical Visual Predictive Check methodology (VPC) was also used for one example dataset for method (a) (the combined residual error model method) and method (l) (the interval-M3 method). Five hundred simulated replicates of the dataset were generated using the corresponding estimated set of parameters. Evolving over time, the observed proportion of data belonging to each category (i.e., intervals of tumor volume) was compared to the 95% confidence interval of the corresponding proportions computed from simulated data. Simulations from estimated models were also compared to the “true” model to better highlight the impact of biased parameters as well as the calculation of the threshold concentration (also known as  $C_T$  which is a secondary

parameter derived from the model (8)) representing concentration at steady state which eradicates the tumor.

The performance of the new method was also evaluated on two other tumor growth inhibition models, as proposed in published works (15,16): in this way, the methods were compared using different tumor growth models, different set of parameters, and different study designs.

### Application to a Real Dataset

Finally, the method was tested with data issued from a real experiment on a new compound currently under development. Data were extracted from an experiment involving three groups (each composed of 4 mice): a control group and two dose groups in which the drug was given orally once a day, 5 days a week for 2 weeks, and 47% of the observations were concerned by the interval censoring.

## RESULTS

### Simulated Data Description

The proportions of tumor volumes being in each interval after the 1000 simulations are summarized in Table IV. The median frequency in each censored interval was about 10% each, leading to a total interval censoring, i.e., values below  $83.2 \text{ mm}^3$ , of about 60%.

**Table III.** Description of All Estimation Methods

Method label	Description	Data	Estimation method
a	Combined error model	No censoring accounted	FOCE-I
b	Combined error model	Ignore data $< LOQ = 0.5 \text{ mm}^3$	FOCE-I
c	Combined error model	Ignore data $< LOQ = 4 \text{ mm}^3$	FOCE-I
d	M5 method	BLQ censoring with $LOQ = 0.5 \text{ mm}^3$	FOCE-I
e	M5 method	BLQ censoring with $LOQ = 4 \text{ mm}^3$	FOCE-I
f	M3 method	BLQ censoring with $LOQ = 0.5 \text{ mm}^3$	Laplacian
g	M3 method	BLQ censoring with $LOQ = 4 \text{ mm}^3$	Laplacian
h	Quadratic error model	Interval censoring	FOCE-I
i	Quadratic error 2 model	Interval censoring	FOCE-I
j	Multi-additive error model	Interval censoring	FOCE-I
k	Multi-proportional error model	Interval censoring	FOCE-I
l	Interval-M3 method	Interval censoring	Laplacian

### Performance of Each Method

The box plots of all methods for each estimated parameter are shown in Fig. 3. Table V shows the results about relative bias and relative RMSE.

The typical value for the tumor growth rate  $\lambda$  did not seem to be affected by the censoring of lower values and was well estimated with all methods. Both relative bias and relative RMSE were about the same value for all methods. Indeed, the information necessary to estimate this parameter was well and sufficiently supported by data from the control group and larger tumors, and as expected censoring had low impact on this parameter. The least biased method was the interval-M3 method (l) with a relative bias of  $-0.53\%$ .

Typical values of the cytotoxic effect parameter  $k_2$  and the initial tumor volume  $TG_0$  appeared to be more impacted by the censoring of the data, being generally underpredicted and overpredicted, respectively. Indeed, the most informative data were expected to be in the range of the interval-censored area for both parameters. The multi-additive and multi-proportional methods (j and k) gave the worst estimations for  $TG_0$  in terms of relative bias (respectively  $-23.3$  and  $22.7\%$ ) and relative RMSE (respectively  $26.1$  and  $27.1\%$ ) while interval-M3 method (l) gave the best results with a relative bias of  $-0.27\%$  and a relative RMSE of  $11.8\%$ . Regarding the typical value of  $k_2$ , M5-methods (d and e) were the most biased methods (with a relative bias respectively of  $-14.1$  and  $-16.5\%$ ), and the best estimations were given by the interval-M3 method (l) with the smallest relative bias of  $0.49\%$  and the most precise estimation with a relative RMSE equal to  $4.06\%$ .

Regarding interindividual variability parameters, similarly, the variance in  $\lambda$  parameter,  $\omega_\lambda^2$ , was not much impacted by the interval censoring, and all methods gave approximately the same range of estimates, but the estimations were generally more biased and less precise than typical value estimations. The imprecision of IIV variance parameters was certainly due to the small number of animals in the simulated study design.

The interindividual variance on  $k_2$ ,  $\omega_{k_2}^2$ , was more sensitive to the interval censoring of the data. Estimations of  $\omega_{k_2}^2$  were more biased with the M5-methods (d and e) (relative bias of  $69.0$  and  $222\%$ ), and the largest relative RMSE were found with M5-methods (d, e), M3-methods (f, g) and quadratic methods (h, i) (from  $48.4$  to  $260\%$ ). For  $\omega_{k_2}^2$  parameter, the less biased method was the combined error method (a) with a relative bias equal to  $0.99\%$ .

We noticed that method (c) gives the worst results regarding bias and precision which can be probably due to the omission of  $20\%$  of the data (LOQ being fixed to  $4 \text{ mm}^3$ ).

In summary, the interval-M3 method (l) was found to be the most adequate method, performing well in terms of bias

and precision for all the parameters, resulting in the lowest relative RMSE and relative bias for all typical parameters.

### Visual Predictive Check

Categorical VPC for the combined error model method (a) and interval-M3 method (l) are presented in Fig. 4. The interval-M3 method (l) better described the evolution of the proportion of tumor volumes to be in each interval than method (a) especially in the range  $[0; 6.8] \text{ mm}^3$ . It is to say, in general, classical methods implementing a combined error model would be poorly informed by tumor size observations belonging to the interval  $[0; 4\sigma_{\text{add}}] \text{ mm}^3$  (where  $\sigma_{\text{add}}$  is the standard deviation of the additive component of the combined residual error model).

In addition, the comparison of tumor growth typical profiles for method (a) and method (l) with the true profile is shown in a supplemental figure (Supplemental Figure 1). The profile simulated with the interval M3 method appears to be the closest to the true profile and especially for high dose for which we observe the lowest tumor volumes. The  $C_T$  defined by Simeoni *et al.* (8,17) was also impacted. For the “true model”  $C_T = 890 \text{ ng ml}^{-1}$ , with the set of parameters estimated with method (a)  $C_T = 974 \text{ ng ml}^{-1}$  and with method (l)  $C_T = 886 \text{ ng ml}^{-1}$ .

### Magnitude of the Residual Error for Larger Tumors

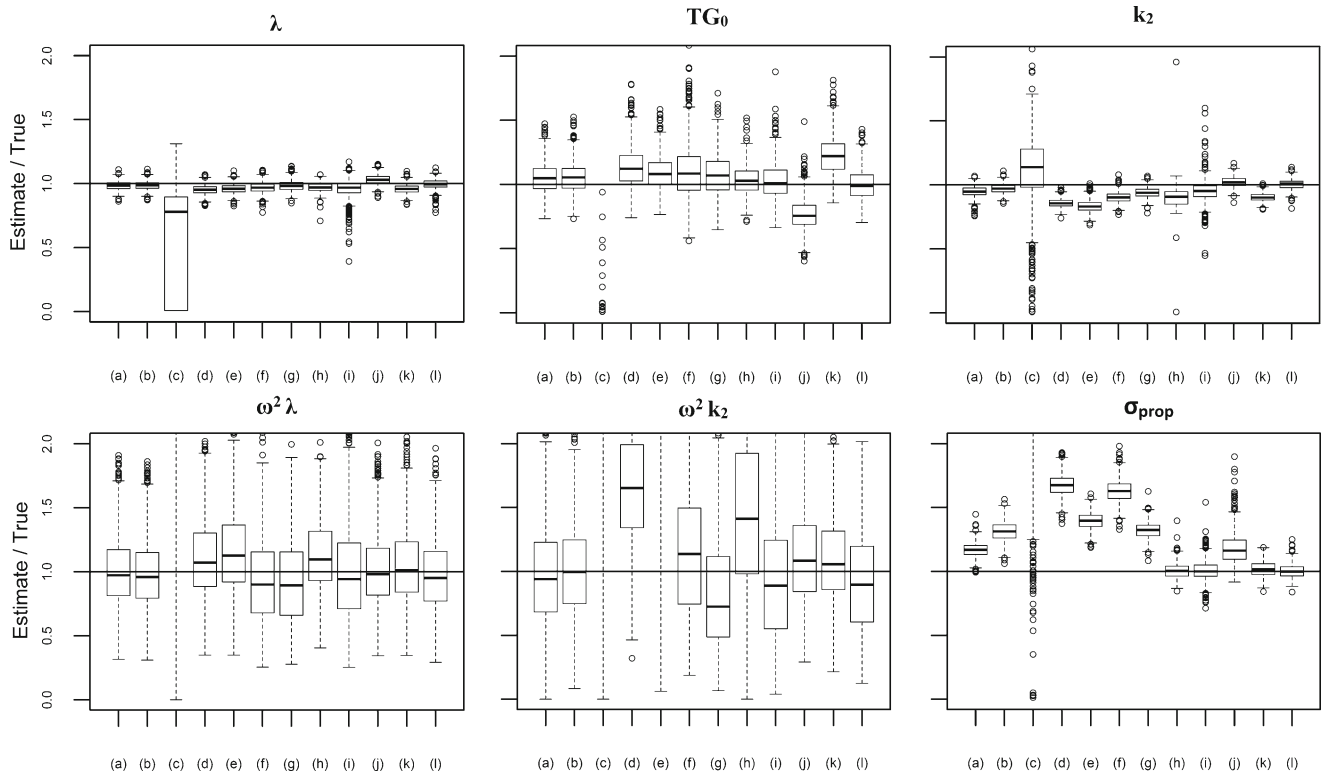
The magnitudes of the proportional residual error,  $\sigma_{\text{Prop}}$ , for larger observations obtained from each method are presented in Table V. This parameter represented the value of the residual error for non-censored tumor sizes and as a result the impact of the interval censoring of smaller tumors on the predictions of larger tumor volumes. Predictions for larger tumors were more altered with methods considering the censoring as a single limit (range from  $23$  to  $49\%$ ). Methods considering the interval censoring due to caliper measurement limitation (h, i, j, k, l) gave a residual error value close to  $20\%$  (range from  $20$  to  $23\%$ ),  $20\%$  being the value of the residual error in the reference model. This suggests that understanding and taking into account the source of error when measuring improve the quality of prediction of both smaller and larger tumor volumes.

### Rate of Minimization Successful Runs

The rates of minimization successful runs for each method are summarized in a supplemental material (supplemental table 1). We observed the lowest rate for the methods using Laplacian estimation methods with the lowest value for method (l) ( $42\%$ ). This implies that the complexification of the estimation method has an impact on the rate of successful

**Table IV.** Distribution of the Proportion of Simulated Tumor Volumes in Each Censored Interval: Median, First and Third Quartiles of the Proportion Among the 1000 Replicates

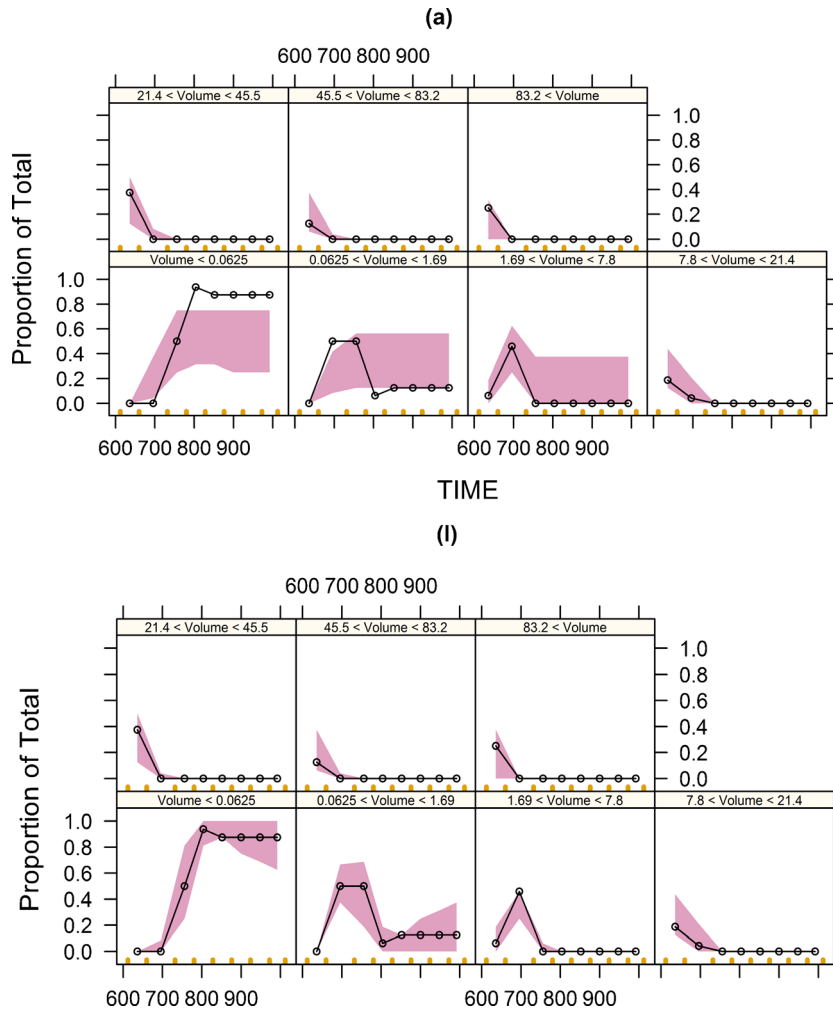
	[0; 0.0625]	[0.0625; 1.69]	[1.69; 7.8]	[7.8; 21.4]	[21.4; 45.5]	[45.5; 83.2]
1st quartile	11.2%	7.9%	7.9%	7.7%	8.6%	8.5%
Median	12.7%	9.6%	9.6%	9.2%	10.1%	9.9%
3rd quartile	14.2%	11.4%	11.0%	10.7%	11.9%	11.2%



**Fig. 3.** Box plots of ratio of parameter estimates divided by the true parameter for each tested method.  $\lambda$  is the exponential growth parameter,  $TG_0$  is the initial tumor size,  $k_2$  is the cytotoxic effect parameter, and  $\omega_\lambda^2$  and  $\omega_{k_2}^2$  the corresponding inter individual variance parameters;  $\sigma_{Prop}$  is the magnitude of the proportional component of the residual error model

**Table V.** Relative Bias and Relative Root Mean Squared Error (RMSE) for Each Methods of the Estimation of the Parameters and Proportional Component of the Residual Error for Tumor Volumes Superior to  $83.2 \text{ mm}^3$  for Each Method

Methods	Parameters					
	$\lambda$	$TG_0$	$k_2$	IIV $\lambda$	IIV $k_2$	$\sigma_{prop}$
Relative bias (%)						Estimate
a	-1.54	5.18	-5.12	0.58	0.99	0.23
b	-1.37	5.51	-2.85	-1.47	2.51	0.26
c	616.3	$>10^3$	$>10^3$	$>10^3$	$>10^3$	0.49
d	-4.78	13.1	-14.1	10.7	69.0	0.33
e	-3.92	8.70	-16.5	15.9	222.6	0.28
f	-3.12	10.0	-9.72	-5.61	16.0	0.33
g	-1.57	7.59	-6.03	-6.36	-16.2	0.26
h	-2.88	3.52	-11.2	15.1	$>10^3$	0.20
i	-4.12	2.81	-4.87	11.4	8.86	0.20
j	3.16	-23.3	2.25	1.20	12.6	0.23
k	-4.02	22.7	-9.65	5.19	10.3	0.20
l	-0.53	-0.27	0.49	-1.34	-7.06	0.20
Relative RMSE (%)						
a	3.70	12.4	6.54	27.3	52.9	
b	3.62	13.0	4.71	27.0	39.2	
c	$>10^3$	$>10^3$	$>10^3$	$>10^3$	$>10^3$	
d	6.03	20.5	14.5	32.8	84.4	
e	5.32	15.4	17.1	36.9	259.9	
f	5.26	22.2	10.5	35.9	56.3	
g	4.34	17.0	7.17	36.2	48.4	
h	4.83	12.4	18.9	52.7	$>10^3$	
i	8.25	14.3	10.1	110.1	125.9	
j	4.85	26.1	4.41	28.2	40.8	
k	5.41	27.1	10.1	30.2	37.9	
l	4.28	11.8	4.06	32.1	40.0	



**Fig. 4.** Categorical Visual Predictive Check for method (a) and method (I) for a given dataset. The panels show simulation based 95% confidence intervals (area) around the observations (solid line) for the fraction of censored observations

minimization runs. Decreasing the proportion of interval-censored values results in an increase of successful convergence rate (not shown).

The additional comparisons performed using other tumor growth parameters from published works (including the Simeoni model (15) with  $\lambda_0$ ,  $\lambda_1$ ,  $\omega^2\lambda_0$ ,  $\omega^2\lambda_1$ ,  $TG_0$ ,  $k_1$ ,  $k_2$ ,  $\omega^2k_2$ , and  $\omega^2k_1$  and including an example with a K-PD model (16)) led to similar results (as shown in [supplemental material](#)), reinforcing the benefit of the use of the interval-M3 method when interval censoring occurs in the data.

### Application to a Real Dataset

The real dataset has been modeled using a bi-phasic tumor growth model (9) and the cytotoxic effect has been described using a sigmoidal function. It appeared that the use of the interval-M3 method (I) improved the fit of the data with a significant decrease of the objective function (889 for method (a), 839 for method (g) and 783 for method (I)). In addition, the precision of the estimated parameters (as shown in [supplemental material](#)) were the best when using the interval-M3 method which is probably due to a better

consideration of the information provided by the data. A comparison of the  $C_T$  parameter calculated with methods (a), (g), and (I) was also done, and a fourfold difference was found between the different methods (1393 vs. 501 vs. 330  $\text{ng}\cdot\text{ml}^{-1}$  respectively for method (a), (g), and (I)).

### DISCUSSION

Data censoring, if not handled properly, may alter the quality of parameter estimation and model development. In this work, stochastic simulations and estimations were performed to investigate the impact of interval censoring in smaller tumor measurements on the estimation of model parameters and evaluate different approaches to handle these data. Clearly, not considering the underlying process of measurement censor can lead to biased parameters and perhaps incorrect parameter interpretations. A novel approach, the so-called interval-M3 method, is proposed and has been evaluated by stochastic simulations and estimations, taking into account the interval censoring of smaller tumors, due to caliper measurement limitations in xenografted animals for tumor growth inhibition experiments.



Comparison of different methods showed there was a large difference in the quality of estimations (regarding relative bias and relative RMSE) depending on the handling method for tumor size censoring and particularly for the drug cytotoxic effect parameter  $k_2$ , which is the parameter of interest when evaluating drug efficacy. The use of the interval-M3 method (method 1), combining the modeling of continuous and categorical data and taking into account the interval censoring process allowed an unbiased and precise estimation of all model parameters. Moreover, this method can be easily implemented in NONMEM (version 6 or later) using the variable  $F\_FLAG$  to simultaneously fit continuous ( $F\_FLAG=0$ ) and categorical data ( $F\_FLAG=1$ ).

The results presented in this article were issued from an example where 60% of observations were interval-censored ( $<83.2 \text{ mm}^3$ ). Similar analyses were performed, varying the proportion of interval-censored data (data not shown): the larger the proportion of censored observations, the larger the bias in parameter estimation. In any case, the interval-M3 method (1) appeared to be the best method with less biased parameter estimations. However, it is important to note that bias of parameters decreases with the proportion of censored data; when few tumor measurements are interval-censored, handling those values with the use of the interval-M3 method does not have a significant impact on model building.

The most commonly used method (the combined error method (a)) showed some limitations regarding tumor volume predictions (as shown on categorical VPC, Fig. 4), especially for smaller tumor sizes corresponding to the interval  $[0; 4\sigma_{\text{add}}]$ . Consequently, this method gives imprecise predictions in this interval. The structural model used in this work was simpler than those usually required, explaining the small value of the additive component. Probably, the use of our interval-M3 method will be even more beneficial than classical methods as the value of  $\sigma_{\text{add}}$  increases.

The modeling exercise performed on real data showed the impact on a real situation and the difference in parameter estimates that can occur and especially for  $C_T$  parameter determination which is one of the key parameter in a xenograft experiment.

A possible limitation in the use of the interval-M3 method (method 1) could be the higher rate of non-successful minimization runs (in a Supplemental Table 1). In fact, we noticed a higher non-successful termination rate for methods requiring LAPLACIAN option. This fact may be explained by different elements. First, the Laplacian method runs more complex computations than FOCE-I since Laplacian is a second-order approximation method for the integral of the likelihood while FOCE-I is a first-order approximation (18). It has been also shown that methods maximizing the likelihood for the censored data have a lower rate of successful termination than methods which only apply extended least squares and do not properly consider the censoring (1,4). Moreover, these methods need a longer computation time; a twofold factor was found compared to methods using FOCE-I method (data not shown). Nevertheless, in our simulation-estimation exercise, initial values were not optimized for each dataset, and no retry was done although

LAPLACIAN methods are known to be more sensitive to initial values. In total, given the non-negligible gain in terms of bias and precision in parameter estimations, the interval-M3 method seems valuable.

The interval-M3 method is an extension of the M3-method used generally for BQL values. In this approach, there is no natural difference between BQL and rounded values since they are handled in the same way: BQL data corresponding to the lowest interval and the borders of the interval-censoring method being adjustable as needed.

In addition, another interesting area to dig in would be to explore the performances of our interval-M3 method with more complex models such as models dealing with combination therapies which could exhibit more efficacy (and thus more interval-censored data) due to synergistic effect.

## CONCLUSION

In conclusion, the interval-M3 method appeared to perform the best (the most precise and the least biased) to take into account interval censoring due to caliper measurement limitations with smaller tumors. This advantage should be important when evaluating antitumor drug effect in tumor growth inhibition experiments in xenografted animals.

## ACKNOWLEDGMENTS

Emilie Héning was funded by *Fondation Synergie Lyon Cancer* and *La Ligue Nationale contre le Cancer*. This work is a part of a PhD project (Philippe Pierrillas) granted by Servier Laboratories. We also want to thank François Bouzom for all the interesting discussions and his wise advice.

## APPENDIX 1

The volume of an ovoid form can be written as:

$$Volume = \frac{4}{3} \cdot \pi \cdot \frac{p}{2} \cdot \left(\frac{q}{2}\right)^2$$

With p and q the diameters of the ovoid.  
Then after development, we have:

$$Volume = \frac{4}{3} \cdot \pi \cdot \frac{p}{2} \cdot \frac{q^2}{4} = \frac{\pi}{3} \cdot \frac{p \cdot q^2}{2}$$

Approximating  $\frac{\pi}{3} \cong 1.047^{-1}$ , the volume of an ovoid can be written as:

$$Volume = \frac{p \cdot q^2}{2}$$

This approximated formula is often applied to convert tumor diameters into tumor volumes.

## APPENDIX 2

NONMEM control stream for the interval-M3 method

**NONMEM control stream for the interval-M3 method**

\$PROBLEM SSE Tumor Growth

\$DATA

\$INPUT

\$\$SUBROUTINE ADVAN13 TOL=6

\$MODEL NCOMP=3

COMP=DEP\_PK ; Depot compartment 1  
 COMP=CENT\_PK ; Central compartment 2  
 COMP=TUMOR\_GR ; Tumor growth 3

\$PK

F1=1  
 KA=1  
 K10=0.1386  
 V2=5  
 S2=V2

LAMB=THETA(1)\*EXP(ETA(1))  
 TG0=THETA(2)\*EXP(ETA(2))  
 K2=THETA(3)\*EXP(ETA(3))

A\_0(1)=0 ; Initialization of depot compartment  
 A\_0(2)=0 ; Initialization of central compartment  
 A\_0(3)=TG0 ; Initialization of tumor growth compartment

\$DES

;-----PK Model-----;

CP=A(2)/S2

DADT(1)=-KA\*A(1) ; Depot compartment First order Absorption  
 DADT(2)=KA\*A(1) - K10\*A(2) ; Central compartment First order Elimination

;-----Tumor Growth Model-----;

DADT(3)=LAMB\*A(3) - K2\*CP\*A(3) ; Exponential Tumor Growth with linear antitumor effect

\$ERROR

;-----Residual Error model-----;

IPRED=A(3)

```

W=THETA(4)*IPRED           ; Weighting of the proportional residual error model

IF (DV.GT.62.5) THEN
  F_FLAG=0                 ; values above 62.5 mm3 treated as continuous data
  Y=IPRED+W*EPS(1)
ENDIF

DUM0=(0.0625-IPRED)/W
PROB0=PHI(DUM0)           ; Definition of the probability tumor volume to be < 0.0625 mm3

DUM1=(1.69-IPRED)/W
PROB1=PHI(DUM1)           ; Definition of the probability tumor volume to be < 1.69 mm3

DUM2=(7.8-IPRED)/W
PROB2=PHI(DUM2)           ; Definition of the probability tumor volume to be < 7.8 mm3

DUM3=(21.4-IPRED)/W
PROB3=PHI(DUM3)           ; Definition of the probability tumor volume to be < 21.4 mm3

DUM4=(45.5-IPRED)/W
PROB4=PHI(DUM4)           ; Definition of the probability tumor volume to be < 45.5 mm3

DUM5=(83.2-IPRED)/W
PROB5=PHI(DUM5)           ; Definition of the probability tumor volume to be < 83.2 mm3

IF (DV.EQ.0.25) THEN
  F_FLAG=1
  Y= PROB0                 ; Definition of the probability tumor volume to be < 0.0625 mm3
ENDIF

IF (DV.EQ.0.5) THEN
  F_FLAG=1
  Y= PROB1-PROB0           ; Definition of the probability tumor volume is in the range [0.0625 ; 1.69] mm3
ENDIF

IF (DV.EQ.4) THEN
  F_FLAG=1
  Y= PROB2-PROB1           ; Definition of the probability tumor volume is in the range [1.69 ; 7.8] mm3
ENDIF

IF (DV.EQ.13.5) THEN
  F_FLAG=1
  Y= PROB3-PROB2           ; Definition of the probability tumor volume is in the range [7.8 ; 21.4] mm3
ENDIF

IF (DV.EQ.32) THEN
  F_FLAG=1
  Y= PROB4-PROB3           ; Definition of the probability tumor volume is in the range [21.4 ; 45.5] mm3
ENDIF

IF (DV.EQ.62.5) THEN

```

```

F_FLAG=1
Y= PROB5-PROB4      ; Definition of the probability tumor volume is in the range [45.5 ; 83.2] mm3
ENDIF

IRES=DV-IPRED       ; Definition of individual residuals
IWRES=IRES/W        ; Definition of individual weighted residuals

$THETA              ; Initial estimates for typical values

$SIGMA
  1 FIX             ; EPS 1

$OMEGA              ; Initial estimates for random effects

$ESTIMATION METHOD=1 INTER NOABORT MAXEVAL=9999 LAPLACIAN NUMERICAL
SLOW                ;; Estimation with LAPLACIAN options

$STABLE            ;; Output Tables

```

## REFERENCES

- Bergstrand M, Karlsson MO. Handling data below the limit of quantification in mixed effect models. *AAPS J.* 2009;11(2):371–80. doi:10.1208/s12248-009-9112-5.
- Duval V, Karlsson MO. Impact of omission or replacement of data below the limit of quantification on parameter estimates in a two-compartment model. *Pharm Res.* 2002;19(12):1835–40.
- Beal SL. Ways to fit a PK model with some data below the quantification limit. *J Pharmacokinet Pharmacodyn.* 2001;28(5):481–504.
- Ahn JE, Karlsson MO, Dunne A, Ludden TM. Likelihood based approaches to handling data below the quantification limit using NONMEM VI. *J Pharmacokinet Pharmacodyn.* 2008;35(4):401–21. doi:10.1007/s10928-008-9094-4.
- Nguyen TH, Comets E, Mentre F. Extension of NPDE for evaluation of nonlinear mixed effect models in presence of data below the quantification limit with applications to HIV dynamic model. *J Pharmacokinet Pharmacodyn.* 2012;39(5):499–518. doi:10.1007/s10928-012-9264-2.
- Sausville EA, Burger AM. Contributions of human tumor xenografts to anticancer drug development. *Cancer Res.* 2006;66(7):3351–4. doi:10.1158/0008-5472.CAN-05-3627.
- Rocchetti M, Simeoni M, Pesenti E, De Nicolao G, Poggessi I. Predicting the active doses in humans from animal studies: a novel approach in oncology. *Eur J Cancer.* 2007;43(12):1862–8. doi:10.1016/j.ejca.2007.05.011.
- Simeoni M, Magni P, Cammia C, De Nicolao G, Croci V, Pesenti E, *et al.* Predictive pharmacokinetic-pharmacodynamic modeling of tumor growth kinetics in xenograft models after administration of anticancer agents. *Cancer Res.* 2004;64(3):1094–101.
- Koch G, Walz A, Lahu G, Schropp J. Modeling of tumor growth and anticancer effects of combination therapy. *J Pharmacokinet Pharmacodyn.* 2009;36(2):179–97. doi:10.1007/s10928-009-9117-9.
- Sheiner LB, Grasela TH. An introduction to mixed effect modeling: concepts, definitions, and justification. *J Pharmacokinet Biopharm.* 1991;19(3):S11–S24.
- Lindbom L, Pihlgren P, Jonsson EN. PsN-Toolkit—a collection of computer intensive statistical methods for non-linear mixed effect modeling using NONMEM. *Comput Methods Prog Biomed.* 2005;79(3):241–57. doi:10.1016/j.cmpb.2005.04.005.
- Lindbom L, Ribbing J, Jonsson EN. Perl-speaks-NONMEM (PsN)—a Perl module for NONMEM related programming. *Comput Methods Prog Biomed.* 2004;75(2):85–94. doi:10.1016/j.cmpb.2003.11.003.
- Beal SL, Sheiner LB, Boeckmann A, Bauer R. NONMEM user's guide (1989–2009). Ellicott City: Icon Development Solutions; 2009.
- Herman AB, Savage VM, West GB. A quantitative theory of solid tumor growth, metabolic rate and vascularization. *PLoS One.* 2011;6(9), e22973. doi:10.1371/journal.pone.0022973.
- Simeoni M, Poggessi I, Germani M, De Nicolao G, Rocchetti M. Population modeling of tumor growth inhibition in vivo: application to anticancer drug development. Uppsala: Population Approach Group in Europe; 2004.
- Sostelly A, Payen L, Guitton J, Di Pietro A, Falson P, Honorat M, *et al.* Quantitative evaluation of the combination between cytotoxic drug and efflux transporter inhibitors based on a tumour growth inhibition model. *Fundam Clin Pharmacol.* 2014;28(2):161–9. doi:10.1111/fcp.12005.
- Magni P, Simeoni M, Poggessi I, Rocchetti M, De Nicolao G. A mathematical model to study the effects of drugs administration on tumor growth dynamics. *Math Biosci.* 2006;200(2):127–51. doi:10.1016/j.mbs.2005.12.028.
- Wang Y. Derivation of various NONMEM estimation methods. *J Pharmacokinet Pharmacodyn.* 2007;34(5):575–93. doi:10.1007/s10928-007-9060-6.