bio-protocol

http://www.bio-protocol.org/e1644   Vol 5, Iss 21, Nov 05, 2015

Capturing the Driving Role of Tumor-host Crosstalk in a Dynamical Model of Tumor Growth

Sebastien Benzekry[1*], Afshin Beheshti[2], Philip Hahnfeldt[3] and Lynn Hlatky[3]

[1]INRIA Bordeaux Sud-Ouest MONC, Talence, France ; [2]Molecular Oncology Research Institute, Tufts Medical Center, Tufts Cancer Center, Tufts University School of Medicine, Boston, USA; [3]Center of Cancer Systems Biology, Tufts University School of Medicine, Boston, USA

[*]For correspondence: sebastien.benzekry@inria.fr

[Abstract] In 1999, Hahnfeldt *et al*. proposed a mathematical model for tumor growth as dictated by reciprocal communications between tumor and its associated vasculature, introducing the idea that a tumor is supported by a dynamic, rather than a static, carrying capacity. In this original paper, the carrying capacity was equated with the variable tumor vascular support resulting from the net effect of tumor-derived angiogenesis stimulators and inhibitors. This dynamic carrying capacity model was further abstracted and developed in our recent publication to depict the more general situation where there is an interaction between the tumor and its supportive host tissue; in that case, as a function of host aging (Benzekry *et al*., 2014). This allowed us to predict a range of host changes that may be occurring with age that impact tumor dynamics. More generally, the basic formalism described here can be (and has been), extended to the therapeutic context using additional optimization criteria (Hahnfeldt *et al*., 1999). The model depends on three parameters: One for the tumor cell proliferation kinetics, one for the stimulation of the stromal support, and one for its inhibition, as well as two initial conditions. We describe here the numerical method to estimate these parameters from longitudinal tumor volume measurements.

Software

1. Matlab (The Mathworks Inc) version 12 or later with optimization toolbox (can also be adapted for Scilab: http://www.scilab.org/index.php/en)

Procedure

This protocol describes how to use a dedicated mathematical model for analysis of tumor growth kinetics. Specifically, it deals with a method for determination of the coefficients of the model from longitudinal measurements of tumor growth. In the data example provided, the measures were obtained by using calipers to determine, every other day, largest (*L*) and smallest (*w*) diameters of subcutaneously implanted tumors (Benzekry *et al*., 2014) for additional details. The formula $V = w^2 L \pi / 6$ was then used to compute the tumor volume.

Researchers can use this protocol in order to compare two groups (or more) of tumor kinetics data and identify which of the coefficient(s) of the mathematical model significantly differ(s) (or not) among the groups. Longitudinal measurements of growth data are required.

A two-dimensional ordinary differential equation model for tumor growth as a function of the tumor volume, V, and its carrying capacity, K (formally, the maximum tumor volume that level of host environmental support could sustain), was used (t denotes time):

$$\begin{cases} \dfrac{dV}{dt} = aV \ln\left(\dfrac{K}{V}\right), & V(t_0) = V_0 \\[2ex] \dfrac{dK}{dt} = bV - dV^{\frac{2}{3}}K, & K(t_0) = K_0 \end{cases}$$

with V_0 = the volume of the tumor at $t = t_0$, K_0= the carrying capacity at $t = t_0$, a = proliferation of the tumor cells, b = carrying capacity stimulation, and d = carrying capacity inhibition. Numerical solutions of the model are fitted to experimental measurements of tumor volume at several time points, and the set of parameters (a, b, d) generating the best fit of the model are thereby determined. In other words, the values of (a, b, d) are determined as the ones minimizing the distance between the model simulation and the data (thus maximizing the likelihood of the data under the hypothesis that it has been generated by the model). The quantity V_0 is fixed to the first observed data point and K_0 is fixed to $2V_0$. For instance, for the first animal of group 1 of the data example, the resulting values of the parameters were: $V_0 = 836$ mm^3, $K_0 = 1672$ mm^3, $a = 0.224$ day^{-1}, $b = 0.710$ day^{-1} and $d = 0.0018$ mm^{-2} day^{-1}.

This protocol is composed of several Matlab scripts and functions that are freely downloadable at the following address: https://github.com/benzekry/fit_tumor_growth. In the following, we detail each step that should be sequentially launched (by typing the script name in a Matlab command window) in order to perform the full task. The associated script is indicated in bold font. We perform these on an illustrative data set example composed of two tumor growth curve groups, each $n = 20$ mice.

1. Import the data from Excel (xlsx) file into Matlab. **importData.m**.
2. Manually explore the parameter space to determine an initial parameter guess that approximately fits the data. **play_model.m**.
3. (Optional) Determine bounds on the parameters. If the parameter space is not restricted for this model, one may face identifiability issues (several parameter sets yielding almost equal solutions). In our example, we chose bounds so that each parameter spans two orders of magnitude around the initial guess. This is specified in **fitGlobal.m** (within the function *lsqcurvefit*).
4. Determine the error variance model to be employed (uncertainty on the data). Options include proportional (common) or constant (unadvised for tumor growth measurements). Note that the former is equivalent to a fit (*i.e.,* sum of squared residuals minimization) of the log-transformed data by the log-transformed model. For caliper-derived measurements of tumor growth, a slightly more elaborate error model

might be employed (Ledzewicz and Schattler, 2008). It involves a power $\alpha < 1$ of the data as well as a minimal threshold and leads to weighted least-squares. In our example we used a proportional error model (specified in *lsqcurvefit* within **fitGlobal.m**).

5. Fit the model to the data using a built-in minimization function of Matlab. Options incude: lsqcurvefit, fmincon, and fminsearch. Personal experience suggests a better ability of fminsearch to escape local minima, although no bound constraint on the values of the parameters can be prescribed with this function. This can be circumvented by penalizing the model for parameter values outside the bounds. Alternatively, stochastic algorithms (such as Markov Chain Monte Carlo) or multi-start optimization might also be considered. We recommend fitting each animal growth curve separately, in particular when inter-animal variability is high, as opposed to fitting the average or median growth curve. The script that launches the fit and exports individual plots is **launch_fit.m**. Individual plots are exported into folders that have user-defined names ('Group1' and 'Group2' in our example).

6. Check the goodness-of-fit. This includes:

   a. Visual inspection of the simulated individual growth curves against the data (Figure 1).

   b. Distribution of the (weighted) residuals (all individuals pooled together). It should be gaussian, a hypothesis that can be assessed using the Kolmogorov-Smirnov test (as can be performed using the Matlab function kstest.m) (Massey, 1951). Two plots of the residuals are outputs when launching **launch_fit.m**.

   c. Computation of the coefficient of determination (R^2), defined by

   $$R^2 = 1 - \frac{\sum\left(y_i - f(t_i)\right)^2}{\sum(y_i - \bar{y})^2},$$ where y_i is the data point at time t_i, $f(t_i)$ is the solution of the model at this time and $\bar{y}$ is the average of the y_i's. The value of R^2 is printed at the screen by **launch_fit.m**. This number is a usual metric of goodness-of-fit that quantifies how better is the model at explaining the data in comparison to the average only (which would give $R^2 = 0$). The closer R^2 is to one, the better the fit.
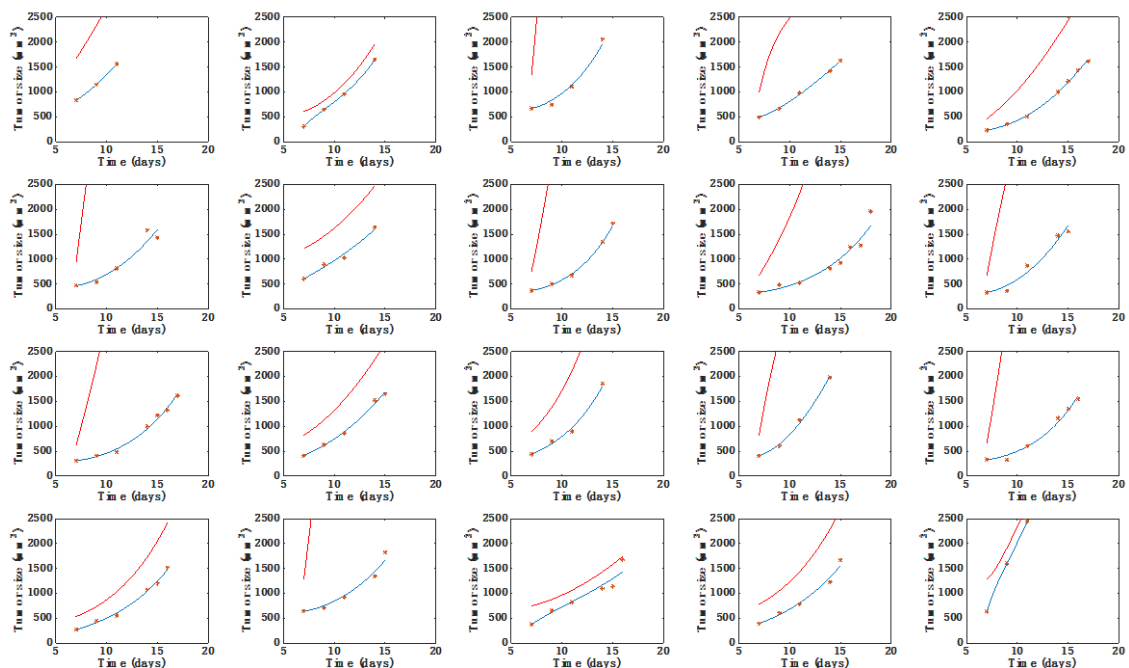
Figure 1. Individual fits of the tumor growth data by the model for each mouse from one of the two groups. The blue curve is the volume (V) curve while the red curve depicts the dynamics of the carrying capacity (K). Fits were performed by optimization of the coefficients *a, b* and *d*. Initial volume V_0 was set to the first volume measured. Initial carrying capacity was set to the double of this quantity. x-axis = time in days, y-axis = tumor size (in mm^3).

7. Investigate for statistical differences in the parameters of the model (here *a*, *b*, *d* and K_0) and across the groups under study. File manips_postFitAnalysis.m outputs two things:

   a. Boxplots of the parameter values, including statistical significance (*t*-test) (Figure 2). Statistical power depends on the number of mice employed. With n = 20 animals, the analysis performed here was able to detect a significant difference of 0.205 day^{-1} in parameter *a* between the two groups (0.295 ± 0.0758 in group 1 versus 0.09 ± 0.0132 in group 2, mean ± standard error, see Figure 2), *i.e.,* 69.4%, at the significance level of α = 0.05. To detect a similar difference of 0.2 day^{-1} with a probability of success (power) of 95% (*i.e.*, probability of false positive of 5%), only 10 animals are sufficient (and necessary).
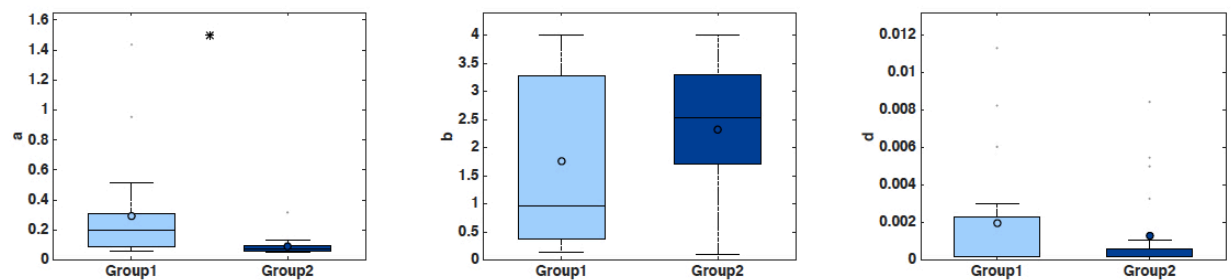
Figure 2. Boxplots of the distribution of three parameters of the model (a, b and d) for the two groups. Initial condition V_0 and carrying capacity K_0 were fixed as in Figure 1. x-axis = Group 1 and Group 2. y-axis = parameter values for *a*, *b* and *d*. * = $\alpha < 0.05$ (Student's *t*-test).

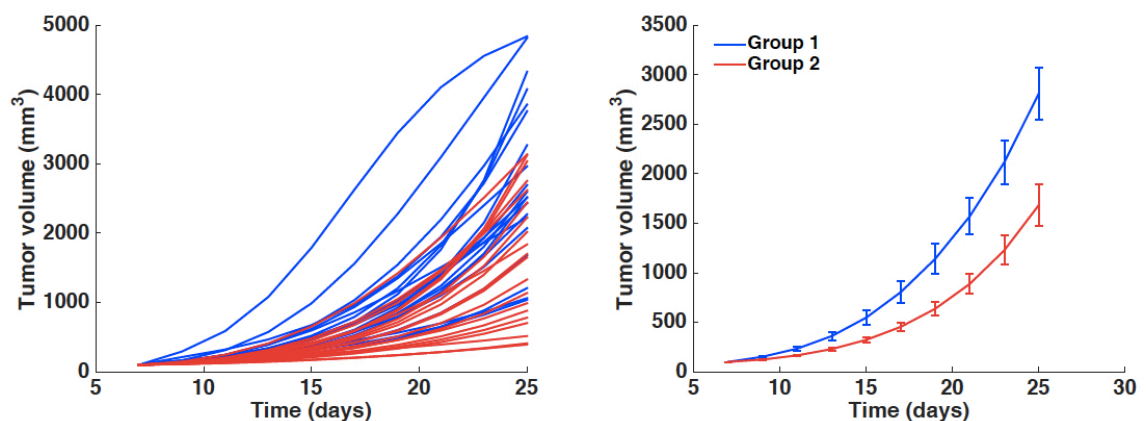b. Simulated growth curves: Either all the animals or averaged other the groups (Figure 3).

Figure 3. Simulated growth curves. Left panel: All individual simulations corresponding to each animal-specific fitted parameter set. Right panel: Average growth curves (mean ± standard error). x-axis = time in days, y-axis = tumor volume in mm³. Blue curves = Group 1. Red curves = Group 2.

Acknowledgments

References

1. Beheshti, A., Benzekry, S., McDonald, J. T., Ma, L., Peluso, M., Hahnfeldt, P. and Hlatky, L. (2015). Host age is a systemic regulator of gene expression impacting cancer progression. *Cancer Res* 75(6): 1134-1143.

2. Benzekry, S., Lamont, C., Beheshti, A., Tracz, A., Ebos, J. M., Hlatky, L. and Hahnfeldt, P. (2014). Classical mathematical models for description and prediction of experimental tumor growth. *PLoS Comput Biol* 10(8): e1003800.

3. Hahnfeldt, P., Panigrahy, D., Folkman, J. and Hlatky, L. (1999). Tumor development under angiogenic signaling: a dynamical theory of tumor growth, treatment response, and postvascular dormancy. *Cancer Res* 59(19): 4770-4775.

4. Ledzewicz, U. and Schattler, H. (2008). Optimal and suboptimal protocols for a class of mathematical models of tumor anti-angiogenesis. *J Theor Biol* 252(2): 295-312.

5. Massey, F. J., Jr. (1951). The Kolmogorov-Smirnov test for goodness of fit. *J Am Statist Assoc* 46(253): 68-78.