



# Patient specific simulation of tumor growth, response to the treatment and relapse of a lung metastasis: a clinical case

Thierry Colin, François Cornelis, Julien Jouganous, Jean Palussière, Olivier Saut

## ► To cite this version:

Thierry Colin, François Cornelis, Julien Jouganous, Jean Palussière, Olivier Saut. Patient specific simulation of tumor growth, response to the treatment and relapse of a lung metastasis: a clinical case. *Journal of Computational Surgery*, 2015, pp.18. 10.1186/s40244-014-0014-1 . hal-01102586

**HAL Id: hal-01102586**

**<https://inria.hal.science/hal-01102586>**

Submitted on 13 Jan 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Public Domain

## RESEARCH

# Patient specific simulation of tumor growth, response to the treatment and relapse of a lung metastasis: a clinical case.

Thierry Colin<sup>1\*</sup>, François Cornelis<sup>1,2</sup>, Julien Jouganous<sup>1</sup>, Jean Palussière<sup>3</sup> and Olivier Saut<sup>1</sup>

\*Correspondence:

colin@math.u-bordeaux1.fr

<sup>1</sup> Univ. Bordeaux, IMB, UMR 5251, F-33400, Talence, France  
CNRS, IMB, UMR 5251, F-33400, Talence, France  
Bordeaux INP, IMB, UMR 5251, F-33400, Talence, France  
INRIA Bordeaux-Sud-Ouest, F-33400, Talence, France  
Full list of author information is available at the end of the article

## Abstract

In this paper, a parametrization strategy based on reduced order methods is presented for tumor growth PDE models. This is applied to a new simple spatial model for lung metastasis including angiogenesis. The goal is to help clinicians monitoring tumors and eventually predicting its evolution or response to a particular kind of treatment. To illustrate the whole approach, a clinical case including the natural history of the lesion, the response to a chemotherapy and the relapse before a radiofrequency ablation is presented.

**Keywords:** Clinical data assimilation; Medical imaging; Partial Differential Equations; Tumor growth modeling

## Background

### Introduction

The metastatic disease to the lung is frequently encountered in patients with cancer whatever the primary location and it has been associated with poor prognosis. The incidence of such disease in patients who have died of an extrathoracic malignancy is reported to be of 20% to 54% [1]. Nevertheless, limited pulmonary metastatic disease can now be successfully treated not only for palliative reasons. By controlling the primary tumor and in the absence of widely disseminated disease in many organs, the resection or the ablation of pulmonary metastases may prolong survival, improve the quality of life and, in some cases, ensure cure [2].

During the last decade, a better management of the metastatic disease to the lung has been achieved by the evolution of imaging, medical oncology, and surgical techniques. There has been improvements in CT imaging quality and scan time [3], as well as advances in the field of nuclear medicine and MRI [4] which can give more precise information on the location and extent of the disease. In particular, there has been widespread use of PET/CT for evaluating patients with metastatic pulmonary disease, which can early detect metabolically active metastatic disease [5]. The targeted oncology treatments can achieve improved responsiveness and the resection of the lesion turns to be possible now with minimally invasive surgical techniques as well as percutaneous thermal ablations (see [6, 7, 8]).

In order to continue this trend and to improve the knowledge of the pulmonary metastatic disease, some authors advocated the development of new tools able to explore the first steps of metastatic implantation [9]. The majority of pulmonary metastases derive from cancer cells that enter the lungs through the pulmonary

arteries and disperse in alveolar capillaries. More rarely, metastatic disease is the consequence of lymphatic spreading or development directly in the bronchial tree (see [10]). Most of these cancer cells are able to adhere to the endothelium of the capillaries. However, the cell survival is ultimately determined by local and oncology factors. Although most of these malignant cells do not survive [2], those that survive penetrate the endothelium of the capillaries and install in the pulmonary parenchyma, where they grow. Therefore, a better understanding of this tumor growth could widely have direct clinical applications.

Cancer growth modeling aims at describing, understanding and predicting the evolution of tumors using numerical models. This science is constantly evolving, from the first cellular automata adapted to the microscopic scale to Ordinary Differential Equations describing the global dynamics of the tumor or Partial Differential Equations taking the spatial distribution of the cancer cells into account. If modeling can be used to help biologists understand the complex mechanisms of tumor growth, in this work, the motivation is to develop a tool for clinical oncologists to improve patient monitoring and eventually to predict the response to a treatment of some metastases to the lung.

As we deal with medical images, the spatial information is important and partial differential equations seem to be the most adapted method to model tumor growth in our context. A true difficulty is to include in a model the biological interactions, as complex as they are, responsible for the tumor evolution without neglecting that the model must be relevant for in vivo applications, and its parameters need to be recovered from clinical images.

Among the diversity of PDE models existing on this topic, we chose, as a starting point, the one developed in [11] for the promising results obtained. The diffusion term for nutrient is replaced by a micro model of angiogenesis (see [12]) which seems more relevant at the medical imaging scale. The goal is to simulate the evolution of a tumor from a given patient so we had to define a calibration method to personalize the model. This method must be a good trade-off between accuracy and computation time. For that purpose, the approach chosen is based on a reduced order method named Proper Orthogonal Decomposition (POD) (see [13] or [14]). The way we address this problem is similar to that of [11]: we first design a PDE model of tumor growth dealing with the spatial distribution of cancer cells with respect to time  $T(t, x, y)$ . This PDE model contains some parameters that are patient-specific. The problem is therefore to determine for which parameters the numerical solution fits the data. Once these parameters have been estimated (we call this step the calibration of the model), one can perform a prediction for a longer time scale.

The outline of this paper is the following: in the following paragraph, we present a clinical case of a single metastasis including natural growth, response to a treatment and relapse. The Methods section is devoted to the presentation of the PDE model of tumor growth that we use and the data assimilation technics. Then the results are discussed in a dedicated part.

### A typical clinical case

In this paper, we will focus on a clinical case of a patient with an isolated metastasis in the lung from a primary bladder tumor. The history of the patient is presented on Fig.1.

Figures 1a, 1b and 1c are CT-scans of the natural history of the lesion, *i.e.* without treatment. The patient has been seen for a radiofrequency ablation on 2008/12/10. At this point, the lesion was too large and the oncologist started a chemotherapy. Figures 1d and 1e are CT-scans during the treatment while Fig.1f shows the relapse a month after the end of the chemotherapy. A radiofrequency ablation was performed in July 2009. The questions that we want to answer in this paper are the following ones:

- 1 Is it possible with Fig.1a and Fig.1b to estimate Fig.1c?
- 2 Is it possible with Fig.1a, 1b, 1c and Fig.1d to estimate the efficacy of the treatment Fig.1e and the amount of the relapse Fig.1f?

## Methods

### A mathematical model for lung metastasis

The model we use in this work is derived from the one described in [11]. We here consider only one kind of cancer cells. The tumor microenvironment, and in particular the amount of nutrients available, is essential to explain its evolution. Consequently, instead of directly modeling the nutrient density, we use a very simplified angiogenesis model to take into account the process by which the tumor escapes the avascular stage. The biological hypotheses of the model are the following:

- One single type of nutrient is considered. Its concentration controls cellular proliferation and death.
- The amount of nutrients available is proportional to the quantity of blood vessels in the tissue. The nutrient diffusion is not taken into account as it is not relevant at the scale we consider.
- The only cellular motion we take into account is the passive transport due to volume variations caused by mitosis or cellular death.
- Cancer cells are continuously switching from the proliferative to the quiescent phenotype.

### Cell behavior

The tumor cell density is denoted by  $T$  and satisfies the following equation:

$$\frac{\partial T}{\partial t} + \nabla \cdot (\mathbf{v}T) = (\gamma_p - \gamma_d)T, \quad (1)$$

where  $\mathbf{v}$  is the transport velocity. The left part of this transport equation satisfies the mass conservation principle and the source term drives this evolution (proliferation and death) of the cells population. The coefficients for proliferation and death by hypoxia (lack of dioxygen),  $\gamma_p$  and  $\gamma_d$ , are detailed in (2) and (3). Above a given threshold of nutrient supply  $M_{th}$ , cancer cells tend to proliferate whereas below this threshold, they starve to death. The use of a hyperbolic tangent in both  $\gamma_p$  and  $\gamma_d$  expressions amounts to smoothing and regularizing the threshold functions, and  $K$  is a smoothing constant fixed. These functions are given by:

$$\gamma_p(M) = \gamma_0 \frac{1 + \tanh(K(M - M_{th}))}{2}, \quad (2)$$

$$\gamma_d(M) = \gamma_1 \frac{1 - \tanh(K(M - M_{th}))}{2}. \quad (3)$$

From the tumor cells density  $T$  and the field  $\gamma_p$  we can have an insight on the proliferative and quiescent cells location. Indeed the cells are proliferative where  $\gamma_p = \gamma_0$  and quiescent elsewhere, so proliferative cells are given by the expression  $P = \frac{\gamma_p}{\gamma_0}T$  whereas quiescent cells are given by  $Q = (1 - \frac{\gamma_p}{\gamma_0})T$ . In addition to cancer cells, we also take into account the healthy tissue whose density is denoted by  $S$ . We assume here that this healthy tissue is transported at the same velocity as the cancer cells but is globally neither proliferating nor dying. More precisely, cellular birth and death counterbalance each other at the timescales we consider so we have, for  $S$ , a transport equation without source term (4):

$$\frac{\partial S}{\partial t} + \nabla \cdot (\mathbf{v}S) = 0. \quad (4)$$

As we consider that the tissue is saturated, the sum of the two densities  $T$  and  $S$  equals to 1 which gives the value of  $S$  (5) (see [15]):

$$S = 1 - T. \quad (5)$$

Summing (1) and (4) we obtain an equation on  $\mathbf{v}$  (6):

$$\nabla \cdot \mathbf{v} = (\gamma_p - \gamma_d)T. \quad (6)$$

This is not sufficient to determine the velocity. However, to close the system of equations (see [16]), we consider that the velocity  $\mathbf{v}$  is obtained using a Darcy law in Eq.(7):  $\mathbf{v}$  is derived from a pressure or potential  $\pi$  in the tissue.

$$\mathbf{v} = -\nabla \pi. \quad (7)$$

This formulation amounts to saying that tumor cells are pushed out if they are proliferating or pulled in if they are dying. Here we could also use a Stokes equation to describe the velocity (see [15]) but it complicates the model without improving significantly its accuracy or biological relevance.

### *Angiogenesis*

At the end of the avascular stage of its development, the tumor reaches such a size that its direct environment is not able to supply enough nutrients to allow the tumor to keep on growing. At this point, cancer cells emit chemical signals such as Vascular Endothelial Growth Factor (VEGF) which result in the emergence of a tumoral neo-vasculature (see [17, 12, 18]). Research on this phenomenon is very active in particular to develop anti-angiogenic treatments. As most of the metastases visible on medical imaging are in the vascular stage, it was important to take the angiogenesis process into account. However, it is out of reach to model this extremely complex phenomenon in details. We therefore choose to use a very compact model in order to achieve a good balance between the complexity of the model and the feasibility of the data assimilation (see next section). It is described by the last two equations (8) and (9). The scalar variable  $\xi$  is the total amount of pro-angiogenic agent such as VEGF. It is secreted by quiescent cells given by the

expression  $\int_{\Omega}(1 - \frac{\gamma_p}{\gamma_0})Td\omega$ ,  $\Omega$  being the computational domain, and evacuated by the organism.

$$\frac{\partial \xi}{\partial t} = \alpha \int_{\Omega} \left(1 - \frac{\gamma_p}{\gamma_0}\right) T d\omega - \lambda \xi. \quad (8)$$

As we assume that the quantity of nutrient is proportional to the density of blood vessels in the tissue, we mixed these two notions in the unique variable  $M$  we shall call “vasculature”. It can be seen as the environment ability to meet the needs of the tumor in term of nutrients, and could be related with the carrying capacity introduced in the Hahnfeldt ODE model [19]. The vasculature  $M$  is damaged by proliferating tumor cells while angiogenesis is promoted by quiescent cells; it therefore obeys

$$\frac{\partial M}{\partial t} = -\eta TM + \beta \xi \left(1 - \frac{\gamma_p}{\gamma_0}\right) T. \quad (9)$$

The model takes into account some important mechanisms involved in tumor growth such as proliferation, death or angiogenesis. Moreover, it is simpler than the previous one [11] as we just have one sort of cancer cells and there is no transport or diffusion of the vasculature. This trade-off is made to keep the model biologically relevant yet simple enough to be parametrized.

#### *Taking therapeutical effects into account*

The model architecture makes easy to include different types of treatment. Chemotherapy effects can be simulated by adding a death term on Eq.(1) as  $-\delta T$ . In this case Eq.(1) becomes:

$$\frac{\partial T}{\partial t} + \nabla \cdot (vT) = (\gamma_p - \gamma_d)T - \delta T. \quad (10)$$

To fulfill the saturation assumption Eq.(6) is modified as follows:

$$\nabla \cdot \mathbf{v} = (\gamma_p - \gamma_d - \delta)T. \quad (11)$$

We can also take into account an anti-angiogenic drug. In this case, we define a therapeutical indicator through a function denoted by  $f$  as:

$$f(t) = \begin{cases} 1 & \text{without treatment,} \\ < 1 & \text{under treatment.} \end{cases}$$

If we consider that it inhibits the production of pro-angiogenic agent, Eq.(8) can be modified as follows:

$$\frac{\partial \xi}{\partial t} = \alpha f(t) \int_{\Omega} \left(1 - \frac{\gamma_p}{\gamma_0}\right) T d\omega - \lambda \xi.$$

If the anti-angiogenic drug also inhibits the growth factor membrane receptors on the endothelial cells, we can add a corresponding term to the vasculature equation Eq.(9):

$$\frac{\partial M}{\partial t} = -\eta TM + \beta \xi f(t) \left(1 - \frac{\gamma_p}{\gamma_0}\right) T.$$

We postpone the modeling of this effect of anti-angiogenic drugs to a future work.

#### Data assimilation technique

The idea is to use the model that has been described in the previous section to obtain a forecast of the evolution of the lung tumor presented in the Background. This model contains numerous parameters that have to be recovered using the two images in order to perform the prediction. Some appear explicitly in the equations, such as the scalars  $\alpha$ ,  $\beta$ ,  $\gamma_0$ ,  $\gamma_1$ ,  $\eta$ ,  $\lambda$  and  $M_{th}$ . Others are implicit and imposed as initial conditions, such as the scalar field  $M(t = 0)$  and the scalar  $\xi(t = 0)$ , and in the case where the patient is under treatment at least one additional parameter has to be determined. The goal of this section is to build a calibration method fast and accurate enough to recover an adequate set of parameters describing the tumor evolution.

#### Simplifying assumptions

As mentioned before, we have 8 scalars and one spatial field to identify. These quantities cannot be estimated by *in vitro* or *in vivo* experiments. Furthermore, the parameterization problem is ill-posed. We can fix  $\xi(t = 0)$  arbitrarily without loss of generality since the variations of  $\xi(t = 0)$  can be taken into account with parameters  $\alpha$ ,  $\beta$  and  $\lambda$ . For convenience, concerning the order of magnitude in our numerical code, we take  $\xi(t = 0) = 0.1$ . For  $M$ , the situation is more complex and we have no universal solution to propose. We choose to take  $M(t = 0) = 0.8 \times T + S$ . This means that the available quantity of nutrients is lower inside the lesion than outside and that initially these quantities are constant. Note that this property is not satisfied for  $t > 0$  because of equation (9). This is one of the strengths of the model: it is possible to obtain an heterogeneous distribution of nutrients within the tumor during the evolution that accounts for complex evolution and that makes the difference with scalar models dealing only with the volume of the tumor. Again the ratio 0.8 in the initialization of  $M$  is arbitrary and has to be related to  $M_{th}$  and  $K$ .

#### Sensitivity analysis

In order to restrict the parameter space, we perform a sensitivity analysis on the model. It aims at determining if some parameters have more influence on the results than others. In this case, it means that the variation ranges of the less meaningful parameters must be modified or that we can simply fix them to a nominal value. For each parameter  $(p_i)_{i=1\dots 7} = (\alpha, \beta, \gamma_0, \gamma_1, \lambda, \eta, M_{th})$ , we choose a range of variation  $[a_i; b_i]$  given in Fig.1. Denoting by  $T(t, x, y, p_1, \dots, p_7)$  the solution given by the PDE system described previously, we compute numerically  $\frac{\partial}{\partial p_i} \int_{\Omega} T(t_f, x, y, p_1, \dots, p_7) dx dy$  for each  $i = 1$  to 7,  $t_f$  being the final time. We evaluate the mean value  $m_i$  and the standard deviation  $S_i$  of these quantities when

the values of the parameters  $(p_i)_{i=1\dots 7}$  are uniformly distributed on  $[a_i; b_i]$ . We use Morris method (see [20]) to perform this computation. The corresponding values of  $(m_i, S_i)$  are plotted in Fig.2.

The influence of the parameter can be established by the distance to the origin of the corresponding point  $(m_i, S_i)$ . In our case, we can see that the most influent parameter is  $\gamma_0$  which is not surprising as it rules the exponential growth of the tumor. Conversely, the model is almost not sensitive to variations of  $\alpha$  and  $\beta$  in the ranges we consider so a part of the inverse problem can be simplified by fixing them thus reducing the degrees of freedom.

#### *Formulation of the inverse problem*

Given a sequence of medical images or snapshots of the tumor, we aim at finding a suitable set of parameters able to reproduce the observed behavior. In other words, simulations with this set must fit as well as possible the clinical data. For this, our approach is to use a scalar objective function, which basically quantifies the distance between the observable data and the model simulation, and try to minimize it. There are different ways to measure this distance and as a criterion we chose a combination of the comparison of the mass and the  $L^2$  distance as follows. The cost function is denoted by  $f_{obj}$ :

$$\begin{aligned} f_{obj}(p_1, \dots, p_7) = & w_1 \times \sum_{i=2}^{n_s} \left[ \sqrt{\int_{\Omega} (T_{data}(t_i) - T_{model}(t_i, p_1, \dots, p_7))^2 d\omega} \right] \\ & + w_2 \times \sum_{i=2}^{n_s} \left[ \left| \int_{\Omega} T_{data}(t_i, x, y) dx dy - \int_{\Omega} T_{model}(t_i, x, y, p_1, \dots, p_7) dx dy \right| \right], \end{aligned} \quad (12)$$

where  $(p_1, \dots, p_k)$  is the parameter set used in the model simulation,  $w_1$  and  $w_2$  are weights to balance the influence of each term, and  $n_s$  is the number of snapshots we have. The first term in the objective function accounts for the shape while the second one only accounts for its volume. Of course, as initial data for our simulation, we use the “exact” value given by the data at time  $t = t_1$ , therefore the sum starts at  $i = 2$ . As a consequence, we need at least two images of the metastasis at different times to personalize the model. The tumor cell density  $T_{data}(t_i)$  is extracted from the  $i^{th}$  snapshot (corresponding to time  $t_i$ ). The quantity  $T_{model}(t_i, p_1, \dots, p_k)$  is the tumor cell density given by the simulation at time  $t_i$  and for the parameters  $(p_1, \dots, p_k)$ . To compute the  $L^2$  distance, a data registration is necessary. We simply translate the images in order that their centers of mass match with the center of mass of the first image  $T_0$ . The masses are computed by integrating the density on the domain  $\Omega$ .

From a computational point of view, whatever the minimization method is, one has to evaluate many times the objective function. It implies to simulate the model for lots of parameter sets which could be quite expensive. For instance, if we use a gradient algorithm to estimate seven parameters, at each iteration, the model is simulated eight times.

To make the calibration faster, we have developed a strategy based on a reduced order method called Proper Orthogonal Decomposition (POD) (see [21]).



*Building a reduced order model to speed up computations*

POD resolution method for infinite dynamical systems consists in approaching partial differential equation systems with ordinary differential equations by decoupling the time and space variables (see [13] and [14]). The initial infinite dimension problem is replaced by a finite dimension (the smallest as possible) problem.

Let us first describe the POD used on the tumor cell density variable  $T$ . As we want to decouple space and time variables, we use the following representation for  $T$  (or any variable of interest):

$$T(X, t) = \sum_{i=1}^d a_i^T(t) \Phi_i^T(X) + \epsilon(X, t),$$

where  $a_i^T$  are scalar functions depending on time and  $\Phi_i^T$  are spatial functions called modes and represent the geometry of the variable  $T$ . The dimension of the reduced problem is denoted by  $d$ . As it is an approximation, an error  $\epsilon(t, X)$  is committed. The goal of POD is to provide us the best basis of spatial functions  $\Phi_i^T$  to minimize the error.

These functions  $\Phi_i^T$  are extracted from a database of admissible behaviors of  $T$ . To generate this database, we sample the parameter space, run the direct model for each parameter set thus obtained and keep several snapshots  $(S_k^T)_k$  of the variable  $T$ . If the sample is correctly chosen, we have a representative set of geometrical configurations for the tumor cell density. Then we look for the functions  $\Phi_i^T$  in the  $d$ -dimensional vectorial space generated by the snapshots from the database (we denote by  $E_h = \text{Span}((S_k^T)_k)$  this approximation space). These functions are taken as an orthonormal basis verifying:

$$\forall h \in \{1, \dots, d\}, \|T(X, t) - \sum_{i=1}^h a_i^T(t) \Phi_i^T(X)\| = \min_{(F_i)_{i \in E_h}} \|T(X, t) - \sum_{i=1}^h a_i^T(t) F_i(X)\|.$$

This property is very interesting as it allows us to use a small number of modes without losing too much accuracy.

The basis solution to this minimization problem is extracted from the autocorrelation matrix of the snapshots  $M^T$  ( $M_{i,j}^T = (S_i^T | S_j^T)$ ,  $(\cdot | \cdot)$  being a scalar product on  $\text{Span}((S_k^T)_k)$ ) by the formula (13):

$$\forall i \in \{1, \dots, d\}, \Phi_i^T(X) = \frac{1}{\sqrt{\lambda_i^T}} \sum_{k=1}^{n_s} v_i^T[k] S_k^T, \quad (13)$$

where  $\lambda_i^T$  denotes the  $i^{th}$  eigenvalue of  $M^T$  and  $v_i^T[k]$  the  $k^{th}$  component of the  $i^{th}$  eigenvector of  $M^T$ .

Once we have obtained these POD modes, we compute the projection of the PDE system on the approximation space  $E_h$  using the eigenmodes  $\Phi_i^T(X)$ . We use the POD approach on both the tumor cell density  $T$  and the pressure field  $\pi$  which are the two fields driven by PDEs in our system. Thus we obtain the approximations  $T(X, t) = \sum_{i=1}^{n_T} a_i^T(t) \Phi_i^T(X)$  and  $\pi(X, t) = \sum_{i=1}^{n_\pi} a_i^\pi(t) \Phi_i^\pi(X)$ ,  $n_T$  and  $n_\pi$  being the numbers of modes we use to represent  $T$  and  $\pi$ .

*Reduced order model on  $T$ :* We inject the POD expression of  $T$  in Eq.(1), where we denote  $\Gamma = \gamma_p - \gamma_d$ :

$$\sum_i \frac{\partial a_i^T}{\partial t} \Phi_i^T + \sum_i a_i^T \nabla \cdot (v \Phi_i^T) = \sum_i a_i^T \Gamma \Phi_i^T.$$

Then this equation is projected on the orthogonal basis  $(\Phi_k^T)_k$ , which gives:

$$\forall j \in \{1, \dots, n_T\}, \frac{\partial a_j^T}{\partial t} + \sum_i a_i^T (\nabla \cdot (v \Phi_i^T) | \Phi_j^T) = \sum_i a_i^T (\Gamma \Phi_i^T | \Phi_j^T).$$

After simplification, we have a system of ODE on the POD coefficients  $(a_i^T)_i$  which is solved with an implicit Euler scheme:

$$\forall j \in \{1, \dots, n_T\}, \frac{\partial a_j^T}{\partial t} = \sum_i a_i^T (\Gamma \Phi_i^T - \nabla \cdot (v \Phi_i^T) | \Phi_j^T). \quad (14)$$

That can be rewritten as:

$$\frac{dA^T}{dt}(t) = N^T(t)A^T(t),$$

where  $N_{i,j}^T(t) = (\Gamma(t)\Phi_j^T - \nabla \cdot (v(t)\Phi_j^T) | \Phi_i^T)$ .

Therefore, we avoid the time interpolation phase that was essential in [11] to approximate the time derivative. Consequently the hypothesis that two consecutive medical data are close in time is not necessary in our case and the accuracy of the scheme is improved.

*Reduced order model on  $\pi$ :* From Eq.(7) and Eq.(6) we obtain:

$$-\Delta\pi = \Gamma T.$$

Then we use the POD expression for  $\pi$ ,

$$-\sum_j a_j^\pi \Delta \Phi_j^\pi = \Gamma T.$$

Finally we project this last equation on the orthogonal basis  $(\Phi_i^\pi)_i$ :

$$\forall i \in \{1, \dots, n_\pi\}, -\sum_j a_j^\pi (\Delta \Phi_j^\pi | \Phi_i^\pi) = (\Gamma T | \Phi_i^\pi). \quad (15)$$

Here again, the system above can be written in the compact form:

$$N^\pi A^\pi(t) = -B^\pi(t),$$

where  $N_{i,j}^\pi = (\Delta \Phi_i^\pi | \Phi_j^\pi)$  and does not depend on the time and  $B_j^\pi(t) = (\Gamma(t)T(t) | \Phi_i^\pi)$ .

*Initialization:* For the simulations, we also need to set initial values for the coefficients  $(a_i^T)_i$  and  $(a_i^\pi)_i$ . These values are obtained by projecting the initial configurations of the fields associated on the POD basis:  $\forall i \in \{1, \dots, n_T\}, a_i^T(t=0) = (T(t=0)|\Phi_i^T)$  and  $\forall j \in \{1, \dots, n_{\pi}\}, a_j^\pi(t=0) = (\pi(t=0)|\Phi_j^\pi)$ .

#### *Resolution of the inverse problem*

The aim of this section is to sum up the algorithm we use in order to solve the inverse problem associated to the cost function Eq.(12), that is find  $(p_1, \dots, p_7) = \text{argmin}(f_{obj}(p_1, \dots, p_7))$ . We proceed in two steps.

*Step 1:* We perform a Monte Carlo method in order to determine a first approximation of the parameters and to avoid a local minimizer. This consists in comparing simulations performed with parameters sets randomly chosen in an empirical parameters space. We keep the parameters set corresponding to the lower value of the cost function  $f_{obj}$  (see Eq.12).

*Step 2:* Once we have this first approximation, we start a gradient descent method using the result of the first step as initialization. In this second step, the value of parameters  $\alpha$  and  $\beta$  are fixed to those obtained in step 1, since the sensitivity analysis performed in paragraph 4.2 shows a low variability of the results with respect to their values.

Note that the use of the POD method implies that we deal with the evaluation of solutions of ODEs and not of PDEs during this process and this fact saves a lot of CPU time.

## **Results and discussion**

### **Tumor growth**

The method previously described was applied to our clinical case (Fig.1). The first scan (on 2008/06/07) represents the initial condition for the tumor cell density  $T$ . The other one (on 2008/09/22) is used in the objective function from Eq.(12) as  $T_{data}$ . The process to extract information from this type of medical images is quite simple. First, the tumor is delineated by a radiologist. The scan measures the permeability to X-rays which we assume is proportional to the density of cells. So we consider that the tumor cell density  $T$  is proportional to the normalized intensity of the corresponding voxel on the image.

To further simplify the problem, we assume, as a first approximation, that tumor cells do not die from hypoxia so we fix  $\gamma_1 = 0$ . Let us remind that we also set the initial conditions for  $M$  and  $\xi$ . We still have six parameters to estimate. A parameter space is empirically defined and we generate a database of simulations on this space. In this case the database contains 374 simulations and for each simulation we save 5 snapshots for both  $T$  and  $\pi$ . From this collection of snapshots, we extract the POD modes and the calibration process is launched with the reduced model. For the simulation of the POD model we only use 36 modes on  $T$  and 10 on  $\pi$  which is here a good balance between accuracy and computation speed. We obtain the set of parameters given in Table 2.

These parameters are used to make a simulation on the direct model and try to predict the evolution of the tumor. A third CT scan on 2008/12/10 is available to compare our prediction and the real evolution.

The result of this simulation is compared with the observed growth of the metastasis on Fig.3. The prediction from the model is satisfactory. To have a quantitative criterion of evaluation of the prediction, we can consider the mass of the tumor which is drawn on Fig.4.

We can see that the horizontal error of prediction of the third scan is less than 10 days. Moreover, the relative error is below 5% of the total duration which is lower than the 10% segmentation error made by the clinician. If we look at the curves, we observe a small difference between the reduced model simulation and the direct model simulation. This is due to the POD hypothesis and the truncation of the POD basis. To quantify the accuracy in shape of the simulation, we calculated spatial indicators as the DICE and the volume concordance given by:  $DICE = 100 \times \frac{2 * |T_{model} \cap T_{data}|}{|T_{model}| + |T_{data}|}$  and  $VC = 100 \times (1 - \frac{|T_{model} - T_{data}|}{|T_{data}|})$ . Moreover, the temporal prediction error is another significant indicator. Indeed, clinicians often need to know when the tumor will reach a critical size. It could be, for instance, the size over which the tumor cannot be treated using radio-frequencies. We denote by  $t_i$  the time of the  $i^{th}$  exam and  $t'_i$  the time when the simulated tumor reaches the size of the real tumor at the  $i^{th}$  exam. It is relevant to look at the delay between the simulation and the real case  $t_i - t'_i$  and the normalized delay  $100 \times \frac{t_i - t'_i}{t_i - t_0}$ ,  $i = 1, 2$ . The values at the calibration time point (2008/09/22) and the final scan (2008/12/10) are given on Table 3.

### Chemotherapy

Given the fast growth of this metastasis, clinicians decided to treat the patient with a chemotherapy. The treatment starts just after the last scan on 2008/12/10 and ends on 2009/06/29. To monitor the efficiency of the treatment, three control scans were planned: two during the treatment (on 2009/03/21 and 2009/05/27) and a last one, one month after the end of the therapy (on 2009/07/27). This last one shows a relapse as the tumor starts growing again from the end of the treatment. It is interesting for us to calibrate the model on this decreasing phase to try to predict the response to the treatment and possibly the relapse. **As the model has already been parametrized on the growth phase, we keep the same parameters from this calibration** and use the cytotoxic drug modeling given in Eq.(10) and Eq.(11). This second inverse problem is easier as the only parameter to estimate is  $\delta$ .

This time again, we need two images. We use the last one before treatment (on 2008/12/10) as the initial condition and the first control (on 2009/03/21) to parametrize. We find  $\delta = 1.05$  and then the direct model is simulated up to after the date of the last scan (2009/07/27).

The evolution of the tumor mass during the treatment provides a good insight on the therapeutical efficacy. It is given on Fig.5 and we can see that here again the model is predictive for this case and provides a good estimation of the response of the patient to chemotherapy.

The tumor shape during treatment obtained by simulation is compared to the clinical data on Fig.6 and the shape indicators are given on Table 4.

The tumor shape is quite well reproduced by the model for the first control scan which we use to find the treatment parameter  $\delta$ . However, on the last control scan made during the cure (on 2009/05/27), the tumor is quite hard to delineate due to various physiological phenomena that are not taken into account in the model (such as oedema or fibrosis). It results from the cytotoxic effects of the treatment on the tumor cells. Finally, the model thus calibrated gives a good prediction of the relapse after the end of the chemotherapy. For this last time point, shapes comparison provides a good result.

#### Another clinical test case

The whole calibration method described previously was used on another case of tumor growth. Here again, we use two scans at different time points to calibrate the model and a third image to quantify the accuracy of the prediction. The mass comparison is given on Fig.7.

In this case, the growth is slower than in the previous case and the model is able to reproduce such a kind of dynamics. Indeed, the time error in prediction is about 5.6 days which, at the time scale we work and considering the tumor registration uncertainties, is a satisfactory result. The shape of the simulated tumor is also satisfying according to the spatial distribution of tumor cell densities on Fig.8 and the indicators on Table 5.

## Conclusion

To try to answer the questions raised by our clinical case a new approach was developed. First, a model was written that takes important mechanisms of tumor growth like hypoxia or angiogenesis into account. Yet the model is simple enough to be calibrated for a specific patient. We have described a new calibration method based on Proper Orthogonal Decomposition to fit the patient data and perform a prediction of the tumor evolution. Prediction which is confirmed qualitatively and quantitatively by the medical imaging. As the model contains various kinds of treatments we could, on the same clinical case, establish a precise quantitative prediction of the response to the treatment at the end of the protocol. As shown in Fig.3, the evolution is really quick between 2008/09/22 and 2008/12/10 and is difficult to predict by the clinicians only using the scans. In that kind of circumstances, a mathematical approach like ours could have helped the oncologist in his diagnosis. On this particular case, the interest of the simulation is clear.

For the time being we do not use any data on the real vasculature. With the constant progress of medical imaging one can imagine that such data will soon be available. This advancement will validate our modeling of the vasculature  $M$ .

We mainly used here two dimensional data extracted from scans (that naturally give a 3D view of the tumor). The whole method was coded for three dimensional data, from the direct model simulation to the POD projection and the calibration algorithm. Currently, it is computationally too expensive to build a database in 3D so we cannot use the whole 3D process on the clinical case. However, by improving and making the database generation faster, it could be interesting to use a 3D

method as the reduced model simulation calculation time is almost the same as in 2D. Moreover it would allow us to free ourselves from the choice of a particular slice.

#### Competing interests

The authors declare that they have no competing interests.

#### Acknowledgements

This study has been carried out within the frame of the LABEX TRAIL, ANR-10-LABX-0057 with financial support from the French State, managed by the French National Research Agency (ANR) in the frame of the "Investments for the future" Programme IdEx (ANR-10-IDEX-03-02).

Experiments presented in this paper were carried out using the PLAFRIM experimental testbed, being developed under the Inria PlAFRIM development action with support from LABRI and IMB and other entities: Conseil Régional d'Aquitaine, FeDER, Université de Bordeaux and CNRS (see <https://plafrim.bordeaux.inria.fr/>).

#### Author details

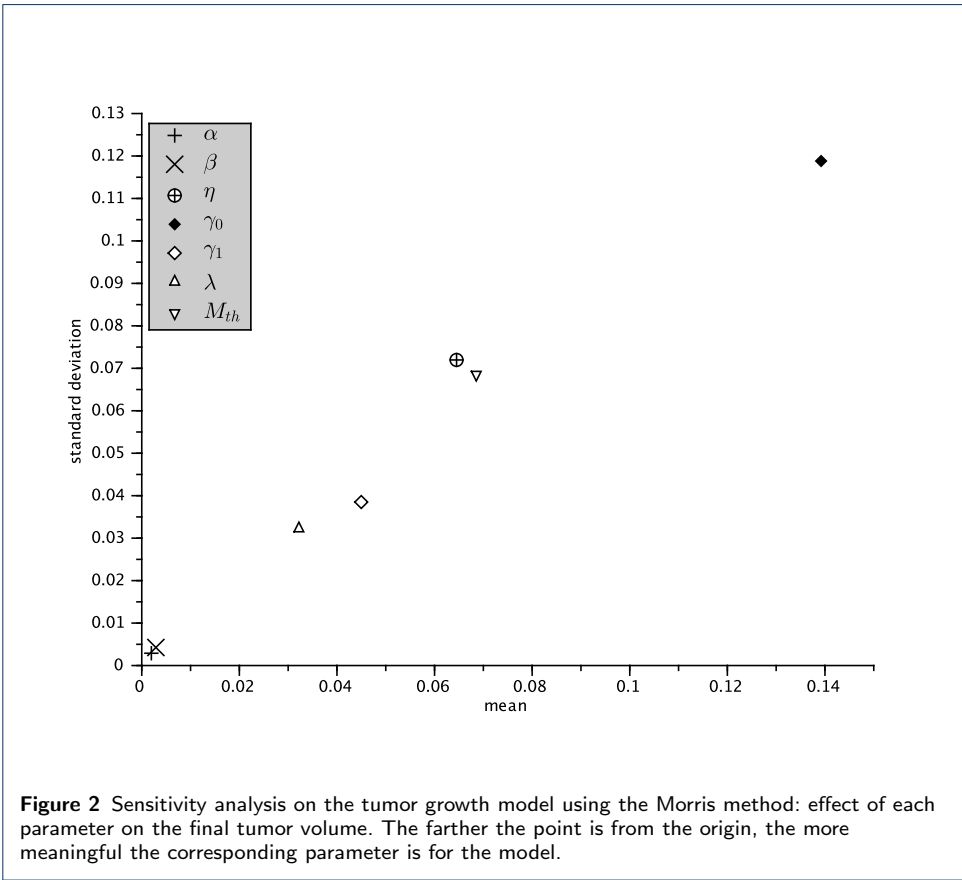
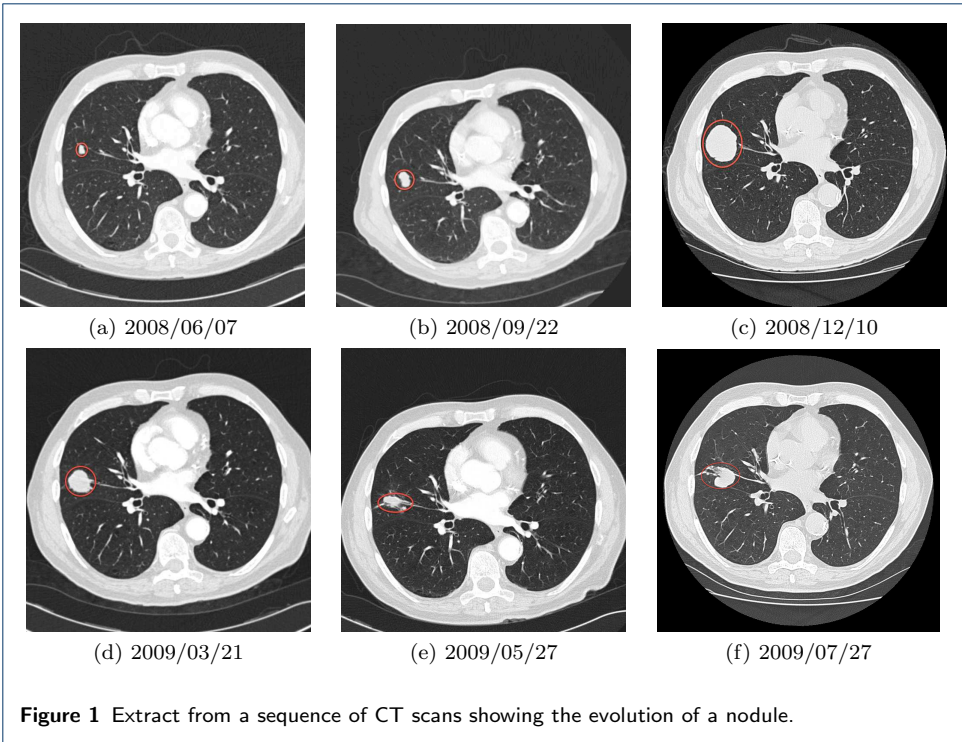
<sup>1</sup> Univ. Bordeaux, IMB, UMR 5251, F-33400, Talence, France CNRS, IMB, UMR 5251, F-33400, Talence, France Bordeaux INP, IMB, UMR 5251, F-33400, Talence, France INRIA Bordeaux-Sud-Ouest, F-33400, Talence, France. <sup>2</sup> Hôpital Pellegrin, CHU Bordeaux, Place Amélie Raba-Léon, 33000 Bordeaux, France. <sup>3</sup> Institut Bergonié, 229 cours de l'Argonne, 33000 Bordeaux, France.

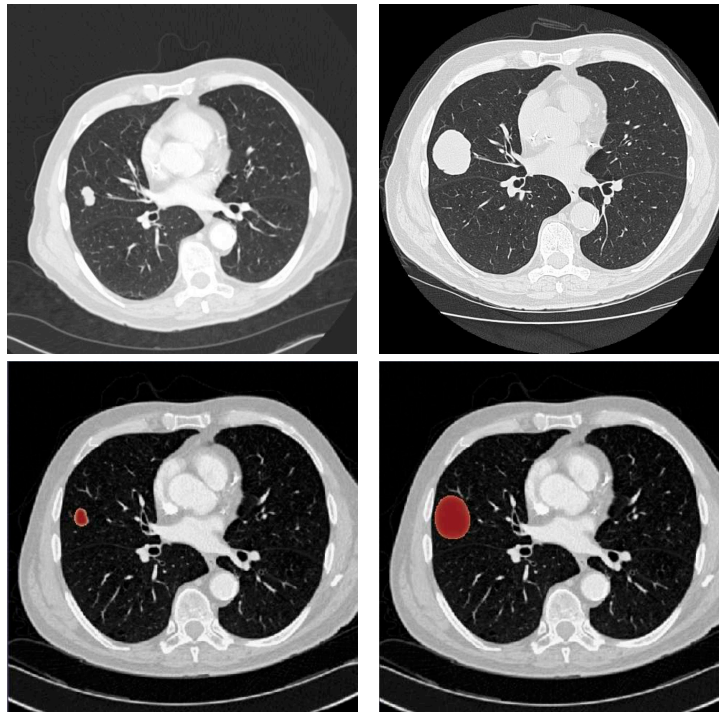
#### References

- Hu, H., Wan, M., Xu, C., Zhan, P., Zou, J., Zhang, Q., *et al.*: Histological subtypes of solitary pulmonary nodules of adenocarcinoma and their clinical relevance. *J Thorac Dis.* **5**(6), 841–6 (2013)
- Zisis, C., Tsakiridis, K., Kougioumtzi, I., Zarogoulidis, P., Darwiche, K., Machairiotis, N., Zaric, B., Katsikogiannis, N., Kesisis, G., Stylianaki, A., *et al.*: The management of the advanced colorectal cancer: management of the pulmonary metastases. *Journal of thoracic disease* **5**(Suppl 4), 383 (2013)
- Prokop, M.: Lung cancer screening: The radiologist's perspective. *Seminars in respiratory and critical care medicine* **35**(1), 91–8 (2014)
- Ohno, Y.: New applications of magnetic resonance imaging for thoracic oncology. *Seminars in respiratory and critical care medicine* **35**(1), 27–40 (2014)
- Lin, Y., Lin, W., Kao, C., Yen, K., Chen, S., Yeh, J.: Prognostic value of preoperative metabolic tumor volumes on pet-ct in predicting disease-free survival of patients with stage i non-small cell lung cancer. *Anticancer Res* **32**(11), 5087–91 (2012)
- Simon, C., Dupuy, D., DiPetrillo, T., Safran, H., Grieco, C., Ng, T., *et al.*: Pulmonary radiofrequency ablation: long-term safety and efficacy in 153 patients. *Radiology* **243**(17392258), 268–75 (2007)
- von Meyenfeldt, E., Gooiker, G., van Gijn, W., Post, P., van de Velde, C., Tollenaar, R., *et al.*: The relationship between volume or surgeon specialty and outcome in the surgical treatment of lung cancer: a systematic review and meta-analysis. *J Thorac Oncol* **7**(7), 1170–8 (2012)
- von Meyenfeldt, E., Wouters, M., Fat, N., Prevoo, W., Burgers, S., van Sandick, J., *et al.*: Local treatment of pulmonary metastases: from open resection to minimally invasive approach? less morbidity, comparable local control. *Surgical endoscopy* **26**(8), 2312–21 (2012)
- Wood, S., Pernemalm, M., Crosbie, P., AD, W.: The role of the tumor-microenvironment in lung cancer-metastasis and its relationship to potential therapeutic targets. *Cancer Treat Rev* (2013)
- Quiros, R., Scott, W.: Surgical treatment of metastatic disease to the lung. *Semin Oncol* **35**(2), 134–46 (2008)
- Colin, T., Iollo, A., Lombardi, D., Saut, O.: System identification in tumor growth modeling using semi-empirical eigenfunctions. *Mathematical Models and Methods in Applied Sciences* **22**(06) (2012)
- Carmeliet, P., Jain, R.K.: Angiogenesis in cancer and other diseases. *Nature* **407**(6801), 249–257 (2000)
- Kunisch, K., Volkwein, S.: Galerkin proper orthogonal decomposition methods for parabolic problems. *Numer. Math.* **90**, 117–148 (2001)
- Kunisch, K., Volkwein, S.: Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. *SIAM Journal on Numerical Analysis* **40**, 492–515 (2002)
- Bresch, D., Colin, T., Grenier, E., Ribba, B., Saut, O.: Computational modeling of solid tumor growth: The avascular stage. *SIAM Journal on Scientific Computing* **32**(4), 2321–2344 (2010)
- Ambrosi, D., Preziosi, L.: On the closure of mass balance models for tumor growth. *Mathematical Models and Methods in Applied Sciences* **12**(05), 737–754 (2002)
- Billy, F., Ribba, B., Saut, O., Morre-Trouilhet, H., Colin, T., Bresch, D., Boissel, J.-P., Grenier, E., Flandrois, J.-P.: A pharmacologically-based multi-scale mathematical model of angiogenesis, and its use in analyzing the efficacy of a new anti-cancer treatment strategy. *Journal of Theoretical Biology* **260**(4), 545–562 (2009)
- Mantzaris, N.V., Webb, S., Othmer, H.G.: Mathematical modeling of tumor-induced angiogenesis. *Journal of Mathematical Biology* **49**(2) (2004)
- Hahnfeldt, P., Panigrahy, D., Folkman, J., Hlatky, L.: Tumor development under angiogenic signaling: a dynamical theory of tumor growth, treatment, response and postvascular dormancy. *Cancer Research* **59** (1999)
- Morris, M.D.: Factorial sampling plans for preliminary computational experiments. *Technometrics* **33**(2), 161–174 (1991)
- Sirovich, L.: Low dimensional description of complicated phenomena. *Contemp. Math.* **99**, 277–305 (1989)

#### Figures

#### Tables





**Figure 3** Comparison between the prediction and the observation. Top row shows the real medical image, below the simulation. The dates are on the left 2008/09/22 (calibration) and on the right 2008/12/10 (prediction).

**Table 1** Parameter space used for the sensitivity analysis.

parameter	variation range
$\alpha$	[5; 25]
$\beta$	[5; 25]
$\eta$	[0.1; 1]
$\gamma_0$	[0.2; 1.4]
$\gamma_1$	[0; 0.3]
$\lambda$	[0.2; 1.2]
$M_{th}$	[0.7; 1.1]

**Table 2** Parameter set obtained by the inverse problem.

$\alpha$	angiogenic agent synthesis factor	8.109
$\beta$	angiogenic growth factor	7.241
$\eta$	vasculature destruction factor	0.673
$\gamma_0$	proliferation rate	1.116
$\gamma_1$	death by hypoxia rate	0
$\lambda$	angiogenic agent destruction rate	0.865
$M_{th}$	hypoxia threshold	1.045

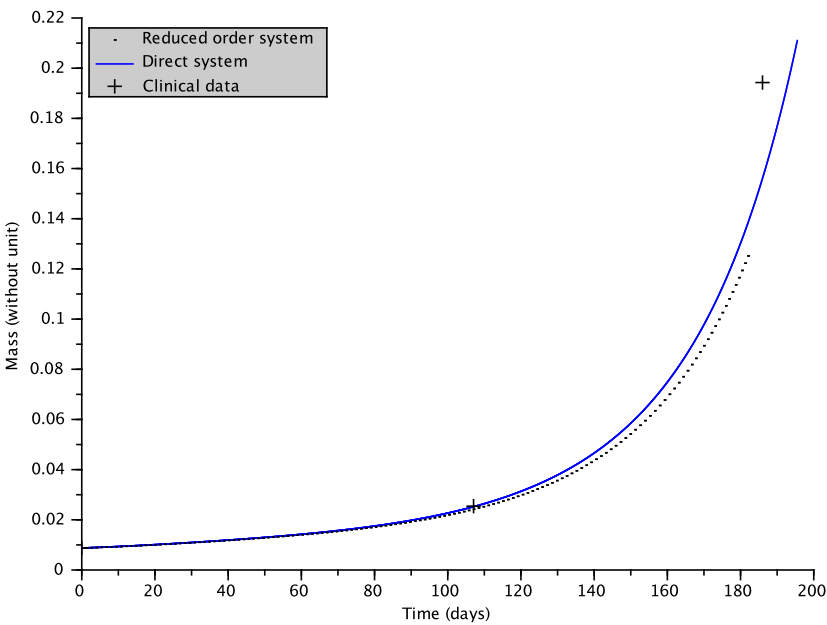
**Table 3** Scalar indicators for the tumor growth of the first clinical case: DICE, Volume Concordance and delays.

	2008/09/22	2008/12/10
DICE	90.96%	87.21%
Volume Concordance	82.54%	77.76%
Delay (days)	0	-6.7
Normalized Delay	0%	-3.6%

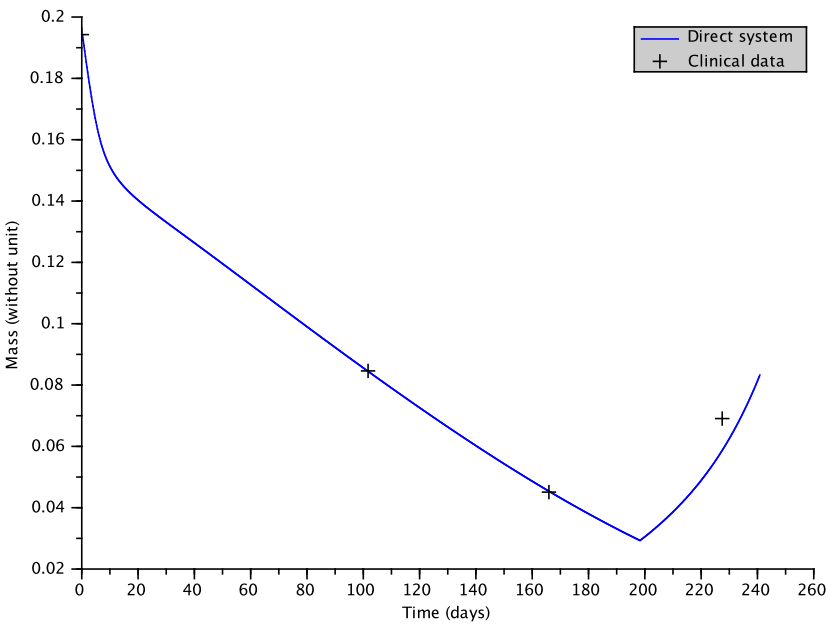
**Table 4** Scalar indicators for the tumor under chemotherapy and rebound of the first clinical case: DICE, Volume Concordance and delays.

	2009/03/21	2009/05/27	2009/07/27
DICE	92.26%	87.44%	84.79
Volume Concordance	84.41%	74.56%	69.9%
Delay (days)	0.5	-0.8	-6.4
Normalized Delay	0.5%	-0.5%	-2.8%

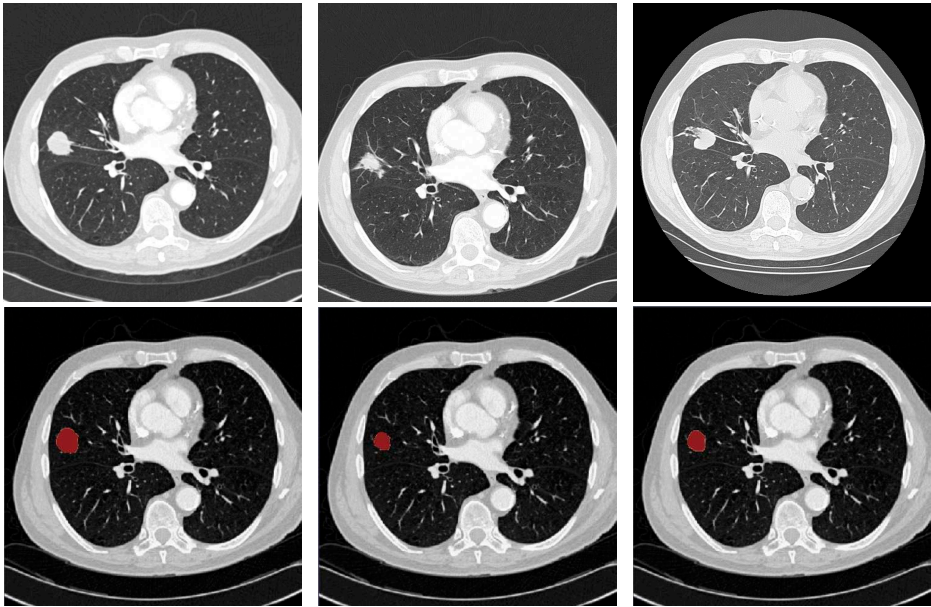




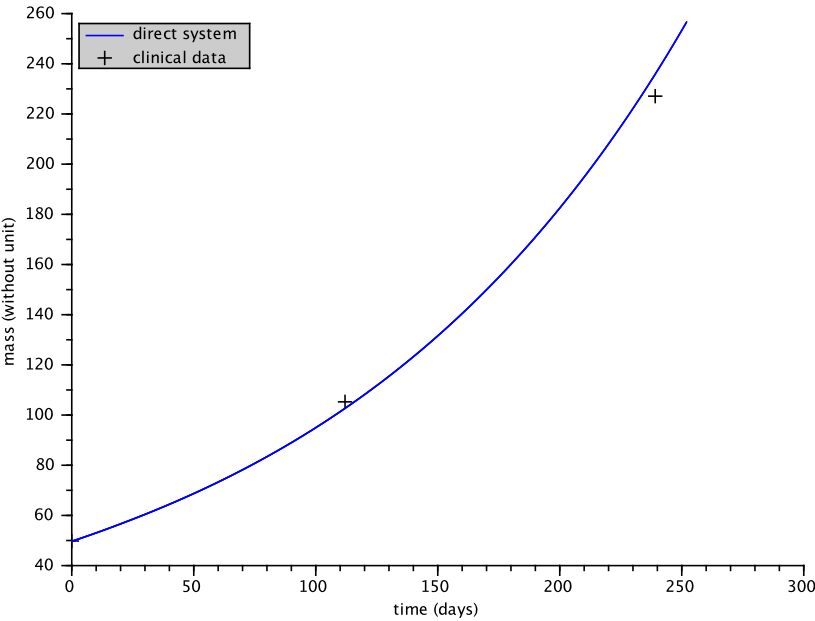
**Figure 4** Mass evolution comparison between the data and the simulations for the tumor growth.



**Figure 5** Mass evolution comparison between the data and the simulations under treatment.



**Figure 6** Comparison between the prediction and the observation for the tumor decrease due to treatment and the relapse post treatment: above the real medical image, below the simulation. The dates are on the left 2009/03/21 (calibration) in the middle 2009/05/27 (prediction for the response to treatment) and on the right 2009/07/27 (prediction for the relapse).



**Figure 7** Second test case: mass evolution comparison between the data and the simulations.



**Figure 8** Comparison between the prediction and the observation for the second test case: above the real medical image, below the simulation. The dates are on the left 2010/03/11 (calibration) and on the right 2010/07/16 (prediction).

**Table 5** Scalar indicators for the tumor growth of the second test case: DICE, Volume Concordance and delays.

	2010/03/11	2010/07/16
DICE	85.41%	88.69%
Volume Concordance	70.59%	76.45%
Delay (days)	0	5.6
Normalized Delay	0%	2.3%