Parameter estimation and mathematical modeling for the quantitative description of drug resistance in gastrointestinal stromal tumor metastasis to the liver

Patricio Cumsille Atala

Group of Investigation in Tumor Angiogenesis (GIANT)
Department of Basic Sciences
University of Bío-Bío, Chillán, Chile
and
Centre for Biotechnology and Bioengineering (CeBiB) - University of Chile

1st CMM-Bath Workshop on Applied & Interdisciplinary Mathematics
March 2019

GIST liver metastases

- GIST: Gastro-Intestinal Stromal Tumors (sarcoma arising from the digestive tract).
- Incidence: 11-15 cases per million people per year.
- Metastases to the liver in 25% of the cases.
- Advanced cancer is not curable.
- There are palliative treatments, but their efficacy remains variable.
- Two different types of therapeutic resistances are observed in patients with this type of metastases.

Current treatment protocol

Treatment for metastatic GIST is tyrosine kinase inhibitor (TKI) therapy.

- First line: Imatinib, a cytotoxic drug which inhibits a specific receptor tyrosine kinase (KIT) triggering apoptosis of tumor cells.
- Second line: Sunitinib is recommended for patients who do not respond to imatinib. Sunitinib is a multi-targeted receptor tyrosine kinase inhibitor that yields to both cytotoxic and antiangiogenic effects.

(a) Imatinib controls metastatic lesions during a period more or less long: around 20-24 months in 85% of patients (see Fig. 1 left).

(b) The rest of the patients present a mutation of a proto-oncogenic KIT (see [1, 2]) that leads to an Imatinib insensitivity (see Fig. 1 right).

Data sets from two patients under a drug resistance scenario

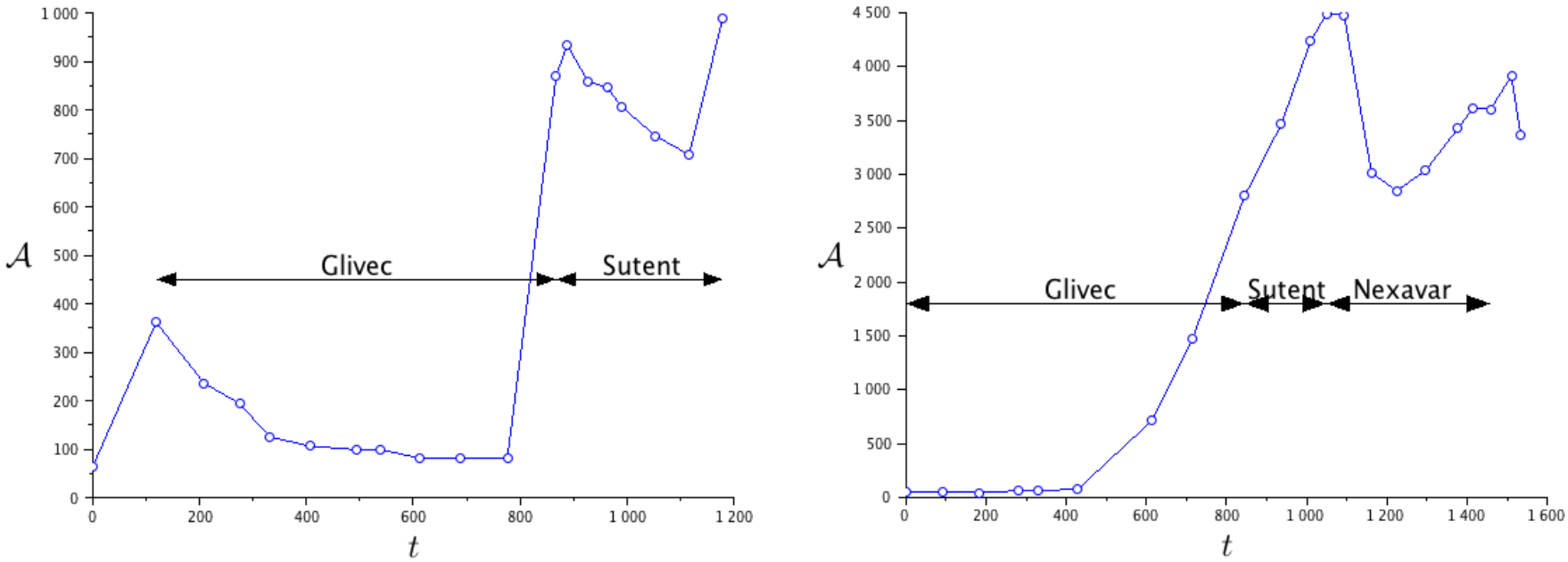

Figure 1: Profiles depict tumour area (in mm^2) versus time (in days). Left: Metastasis is controlled by Imatinib before a first relapse. Right: Treatment with Imatinib is ineffective, then this is switched to Sunitinib and subsequently to Sorafenib, which are relatively effectives until a relapse is observed.

Therapeutic Follow-up of patients

- CT-scans (usually every 2 months): Track the disease evolution and the response to the treatment.
- Clinicians challenge: optimizing cancer treatments and particularly the switch time from the first-line to the second-line treatment, in order to increase the overall survival time.
- RECIST criteria: the largest diameter of the lesion is the *only information* extracted from the CT-scan.

Our goal

General: Quantitative description of TKI therapy resistance in patients with GIST metastasis to the liver.
Novelty: This is the first attempt in the direction of predicting drug resistance.

Specific:

- Develop patient specific mathematical modeling based on medical images of GIST metastasis to the liver;
- Devise a practical identifiability approach (parameter estimation, model selection).

Brief overview of tumor growth models

We may classify the models by the features we want to capture/analyze:

1. Discrete systems (e.g. cellular automata): Growth at cellular scale.
2. ODE: Tissue scale, good approximation for tumor area/mass. Assumes spatial homogeneity.
3. PDE: Tissue scale, spatial heterogeneity. More complete description of spatial phenomena, however more difficult to analyze.

Brief overview of inverse problems in biomedicine

- Classically, models used for clinical applications are based on ODE.
- They are parametrized using statistical methods.
- They may provide a prognosis of tumor volume, among others important aspects (see [3]).

Difficulties

- A non linear least squares problem in order to fit modeling to clinical data has to be solved.
- Good initial parameter values would be required in order to compute good parameter estimation, which is very difficult to obtain.
- No *a priori* knowledge on the system: neither the actual proportion of sensitive and resistant to treatments cells populations (at any time) nor their time of emergence is known.
- Modeling is complex since it contains nonlinear dynamic capable to reproduce different scenarios. In particular, there is no stability with respect to parameters ($\Longrightarrow$ no uniqueness of parameters).
- Data are sparse and only one partial combination of model variables is measured (tumor area).

Previous work

In *Spatial modelling of tumour drug resistance: the case of GIST liver metastases*, G. Lefebvre, F. Cornelis, **P. Cumsille,** T. Colin, C. Paignard and O. Saut, Mathematical Medicine and Biology, 34 (2017), pp. 151-176,
we propose a PDE model for GIST metastasis to the liver, its growth and therapy resistance. Here, assuming spatial homogeneity, we have obtained a general model and proposed five variants.

Mathematical modeling description

A strength of this work consists in the modeling of treatment: first-line (τ_1) and second-line (τ_2). Cytotoxic drugs do not impact similarly all the metastatic cancer cells since resistant phenotype can appear in the proliferative cell population.

Modeling considers a tumor is described by means of three proliferative tumor cells populations, which volumes are denoted by (P_1, P_2, P_3):

- P_1 is sensitive to τ_1 and τ_2;
- P_2 is sensitive only to τ_2; and
- P_3 is resistant to both τ_1 and τ_2.

Modeling also takes into account cell death and angiogenesis, which is a key factor in metastatic growth. The variable M represents vascularization and nutrient supply through angiogenesis.

Mathematical modeling

Equations for the volumes of proliferative tumor cells (1/4)

The proposed models are based on mass balance principle accounting for the different volume of proliferative tumor cells, as well as vascularization and nutrient supply through angiogenesis.

The volume of the proliferative tumor cells obeys a mass balance principle according to the following general equation:

$$P_i' = \left[\mu(M) - \delta(M) - \delta_i^{treat}(M)\right] P_i, \quad \text{for } i = 1, 2, 3. \tag{1}$$

$\mu(M)$ and $\delta(M)$ denote cellular growth and death rate, respectively, which depend on M as follows:

$$\mu(M) = \mu_{MAX} \frac{1 + \tanh(R(M - M_{hyp}))}{2}, \tag{2}$$

$$\delta(M) = \delta_{MAX} \frac{1 - \tanh(R(M - M_{hyp}))}{2} \tag{3}$$

Mathematical modeling

Equations for the volumes of proliferative tumor cells (2/4)

$M > M_{hyp}$ then $\mu(M) \approx \mu_{MAX}$ and $\delta(M) \approx 0 \Longrightarrow$ tumor cells undergo proliferation.

$\mu(M)$ and $\delta(M)$ are regularized versions of the Heaviside function, R is a numerical smoothing parameter (arbitrarily fixed to 5).

μ_{MAX} and δ_{MAX} are the maximum growth and death rates of the tumor cells, respectively.

Mathematical modeling

Equations for the volumes of proliferative tumor cells (3/4)

In (1), $\delta_i^{treat}(M)$ represents the death rate due to the treatments, which are related to the dose of drug delivered to the patient, among others factors.

Treatments act in different ways for the different proliferative tumor cell populations, P_i for $i = 1, 2, 3$. The functions $\delta_i^{treat}(M)$ are defined by:

$$\delta_1^{treat}(M) = [\delta_1 \chi_1(t) + \delta_2 \chi_2(t)] (\alpha + M), \quad (4)$$

$$\delta_2^{treat}(M) = \delta_2 \chi_2(t)(\alpha + M), \quad (5)$$

$$\delta_3^{treat}(M) = 0. \quad (6)$$

Mathematical modeling

Equations for the volumes of proliferative tumor cells (4/4)

In above equations, we have denoted by

$$\chi_1(t) = \mathbf{1}_{\{t<T_j\}}(t) \quad (\text{resp. } \chi_2(t) = \mathbf{1}_{\{t\geq T_j\}}(t)) \tag{7}$$

the characteristic function of treatment τ_1 (resp. τ_2), where T_j is the time at which physicians switch from τ_1 to τ_2 treatment for each patient $j = 1, 2$ considered in this work. Moreover, δ_k is the maximum death rate due to treatment τ_k for $k = 1, 2$.

Finally, the parameter α in Eqs (4)-(5) stands for a quantification of a basal vasculature, which is considered to be 0 or 1, depending on whether the model considers this basal level or not.

Mathematical modeling

Vascularization, nutrient supply and angiogenesis (1/3)

M describes vascularization and nutrient supply driven by tumor angiogenesis; see the review by Cumsille *et al.* [3] for a detailed overview on tumor growth. τ_2 effect has to be taken into account in these two related aspects.

Since the nutrients are supplied to the tumor by the vascularization, as a simple representation, only one variable is utilized to describe the nutrient concentration and the vascularization. Let us denote by M this variable, which is governed by a mass balance principle:

$$M' = \gamma \frac{\delta(M)}{\delta_{MAX}} \{(1 - \nu\chi_2(t))(P_1 + P_2) + \zeta P_3\}^{2/3} - \beta M P \cdot g\left(\frac{\mu(M)}{\mu_{MAX}}\right) \tag{8}$$

Mathematical modeling

Vascularization, nutrient supply and angiogenesis (2/3)

Since P_1, P_2 and P_3 produce TAFs, nutrient availability increases $\longrightarrow$ represented in Eq (8) by the term $\gamma \cdot \delta(M)/\delta_{MAX}$, where γ is the tumor angiogenic capacity.

Angiogenesis is reduced by τ_2 which acts on population $P_1 + P_2$ which is sensitive to it $\longrightarrow$ represented for some models by the term $\nu\chi_2(t)$, ν stands for the anti-angiogenic effect of τ_2, whereas ζ is fixed to 1.

For other models, it is indirectly accounted for a relative increase of P_3 with respect to $P_1 + P_2$, ζ takes into account it, whereas ν is fixed to 1.

Mathematical modeling

Vascularization, nutrient supply and angiogenesis (3/3)

Finally, β is the rate of nutrients consumption and
$g\left(\mu(M)/\mu_{MAX}\right) = 1$ for some models, whereas
$g\left(\mu(M)/\mu_{MAX}\right) = \mu(M)/\mu_{MAX}$ for the others models.

The term $g\left(\mu(M)/\mu_{MAX}\right) = \mu(M)/\mu_{MAX}$ is meaningful since

$$\mu(M) \approx \mu_{MAX} \Longrightarrow \beta MP \cdot g\left(\mu(M)/\mu_{MAX}\right) \approx \beta MP$$

(High cell growth implies high nutrient consumption), whereas

$$\mu(M) \approx 0 \Longrightarrow \beta MP \cdot g\left(\mu(M)/\mu_{MAX}\right) \approx 0$$

(Low cell growth implies low nutrient consumption).

Finally, the exponent $2/3$ in Eq (8) accounts for the fact that nutrient availability must be proportional to the tumor cells' surface.

Mathematical modeling

Summary of the models and vector form

Table below summarizes all the models considered.

Table 1: **Models proposed.**

Model №	α	ν	ζ	$g\left(\mu(M)/\mu_{MAX}\right)$
1	0	Active	Fixed to 1	1
2	0	Fixed to 1	Active	1
3	1	Active	Fixed to 1	1
4	1	Active	Fixed to 1	$\mu(M)/\mu_{MAX}$
5	0	Active	Fixed to 1	$\mu(M)/\mu_{MAX}$

Finally, mathematical models previously described can be written under the vector form:

$$U' = F(t, U, \theta), \tag{9}$$

$$U(t_0) = U_0, \tag{10}$$

Mathematical modeling
Modeling capabilities

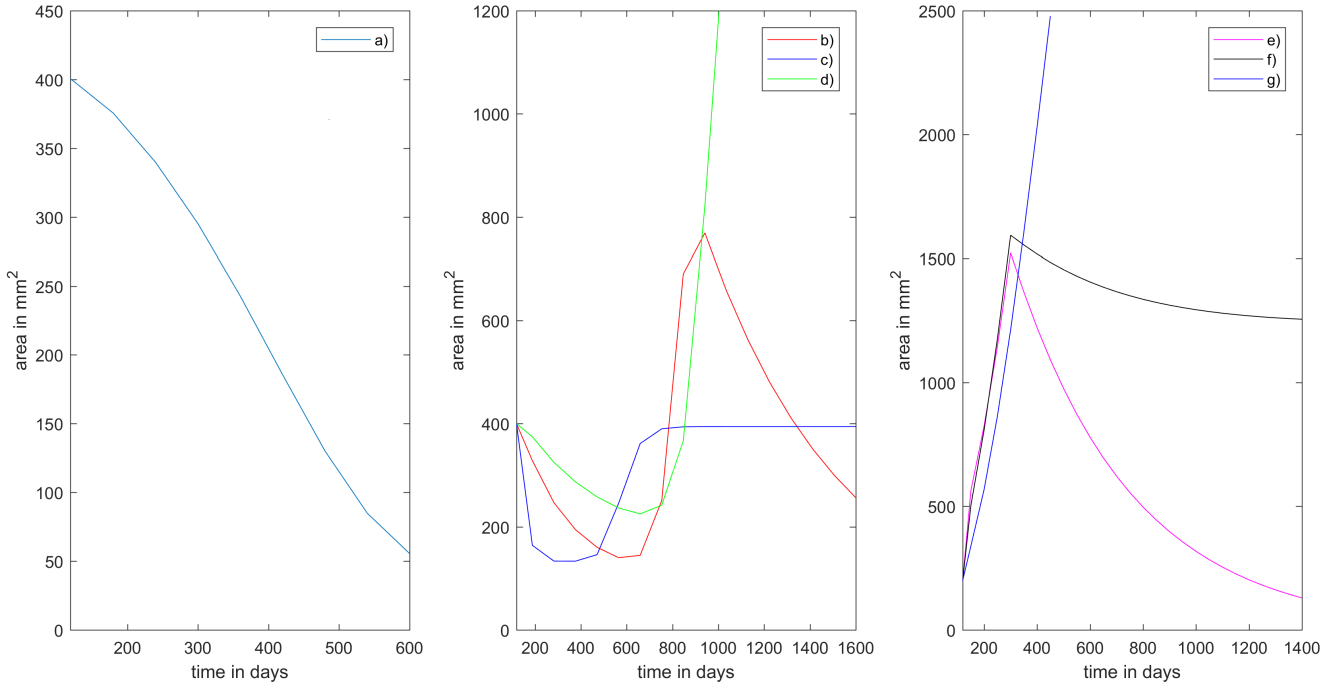

Figure 2: **Modeling capabilities.** Our modeling is able to reproduce the different scenarios reported by physicians. Left: a) τ_1 effective, no τ_2 applied. Center: τ_1 is applied then switched to τ_2, b) τ_1 partially effective, tumor regrows and τ_2 reduces tumor size; c) τ_1 is effective, τ_2 is effective in maintaining tumor size; d) τ_1 is effective, τ_2 is ineffective. Right: τ_1 becomes ineffective in reducing or maintaining tumor size and it is switched to τ_2, e) τ_2 is effective in reducing tumor size; f) τ_2 only stabilizes the tumor size; g) τ_2 is completely ineffective.

Parameter estimation problem

In the proposed models, summarized in Table 1, the quantity $P(t) = (P_1 + P_2 + P_3)(t)$ represents the total number of tumor cells at time $t \longrightarrow$ it is proportional to the tumor area $\mathscr{A}(t)$. Therefore, only this partial combination of model variables is observed, i.e. it is the *model's observation function*. In consequence, to identify the model's parameters one has to find the vector $\theta^j \in \mathbb{R}^8_+$ such that the sum of squares

$$S(\theta^j) = \sum_{i=1}^{N^j} \left(\mathscr{A}_i^j - P(t_i^j, \theta^j) \right)^2 \tag{11}$$

is minimized with respect to the data $\{\mathscr{A}_i^j\}_{i=1}^{N^j}$ observed in time points $\{t_i^j\}_{i=1}^{N^j}$, for each model's observation function $P(t_i^j, \theta^j)$, for each data set $j = 1, 2$ (patient 1 and 2), under constant variance data assumption.

Main difficulty: ill-posedness

Unfortunately, inverse problems are frequently ill-posed, issue that has been observed in systems biology; see *Inverse problems in systems biology*, H. Engl, C. Flamm, P. Kugler, J. Lu, S. Muller & P. Schuster, Inverse Problems, Vol. 25, No. 12 (2009).

When comparing the three curves in Fig. 2(right) during the first 400 days only, we could infer that the corresponding sets of parameters should be almost the same implying model stability, which is not the case.

One can find different sets of parameters that produce the same initial tumor behavior but that lead to a drastically different long-term evolution.

Regularization technique

To overcome this difficulty, we introduce a regularization term, i.e. instead of minimizing $S(\theta^j)$, we have minimized

$$\min_{\theta^j \in \mathbb{R}^8_+} S_\alpha(\theta^j) \left(\equiv S(\theta^j) + \alpha \|\theta^j\|_2^2 \right), \tag{12}$$

where $\alpha > 0$ is the regularization parameter. We have numerically solved the optimization problem in Eq (12) for several values of α small enough obtaining the best results for $\alpha = 1e-3$.

An algorithm to perform the minimization

Proposed solution

We define a 'box' for the 'feasible' parameters θ given by an upper and lower bound for each particular parameter

$$\Theta_{ad} = [\gamma_{0,\min}, \gamma_{0,\max}] \times [\gamma_{1,\min}, \gamma_{1,\max}] \times \ldots \times [\nu_{\min}, \nu_{\max}].$$

We pick n random samples distributed in this set and solve the direct problem on each of the n obtained random parameters $\tilde{\theta}_i^j$, we choose $\theta_0^j = \text{argmin}_i S(\tilde{\theta}_i^j)$.

This method allows to deal with the problem and, in fact, it is a basic idea for several **global** solvers.

The Algorithm

1. Load the data $tdata$, $Area(tdata)$, initial conditions and previously saved $\hat{\theta}^j$ (optimal parameters from last run).
2. Set n (number of random samples) and Θ_{ad} (feasible parameters)
3. Sample the possible initial parameters $\tilde{\theta}_i^j \in \Theta_{ad}$, $i = 1, \ldots, n$, $j = 1, 2$.
4. Define $\theta_0^j := \text{argmin}_i S(\tilde{\theta}_i^j)$,
5. If $S(\theta_0^j) < S(\hat{\theta}^j)$ solve the problem

$$\min_{\theta^j \in \mathbb{R}_+^8} S_\alpha(\theta^j),$$

by using *fminsearch* from Matlab® with initial guess θ_0^j. Else, break.

6. Define $\hat{\theta}^j$ as the output of the solver and save.

Fit to Patient 1 data

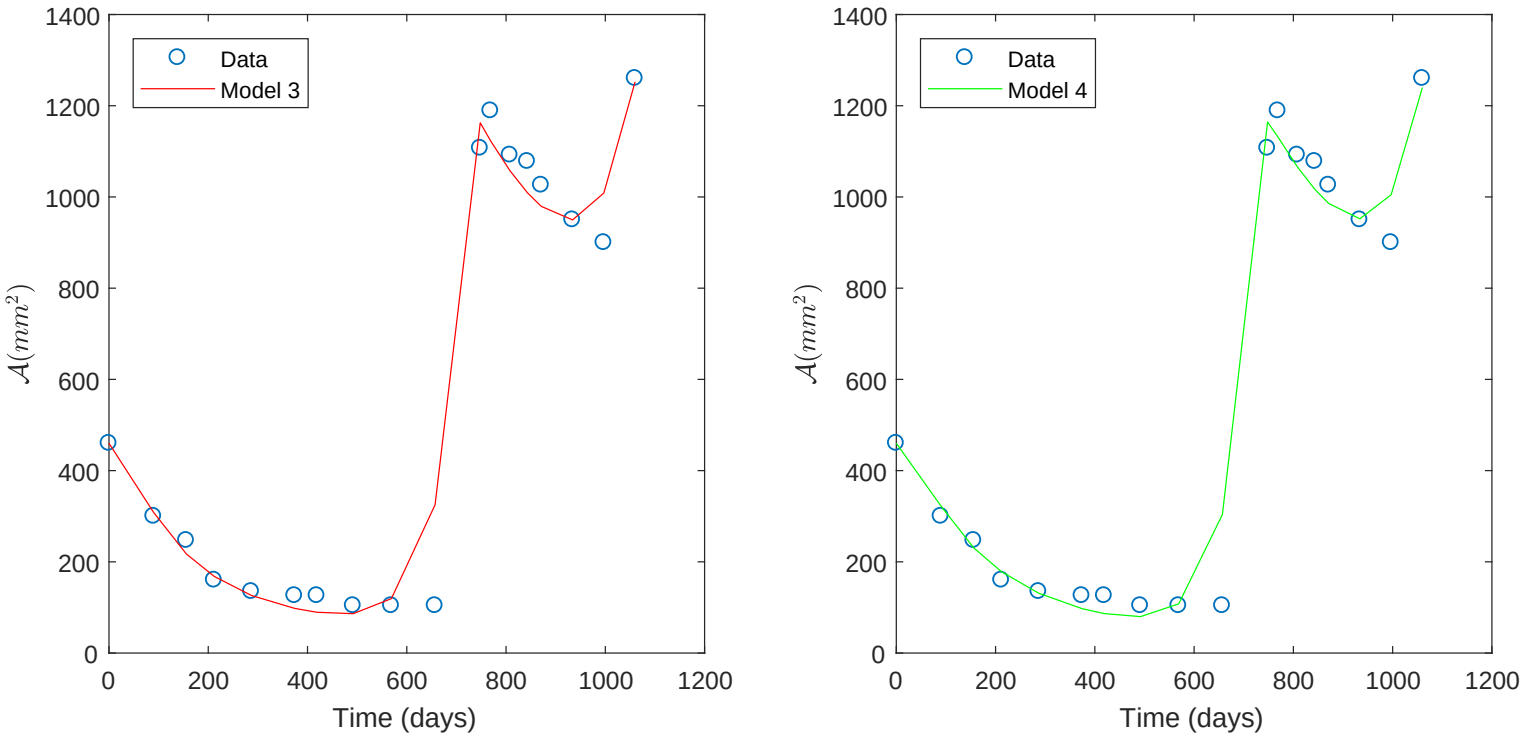

Figure 3: **Fit to Patient 1 data**

Fit to Patient 1 data

Validating statistical assumption (1/2)

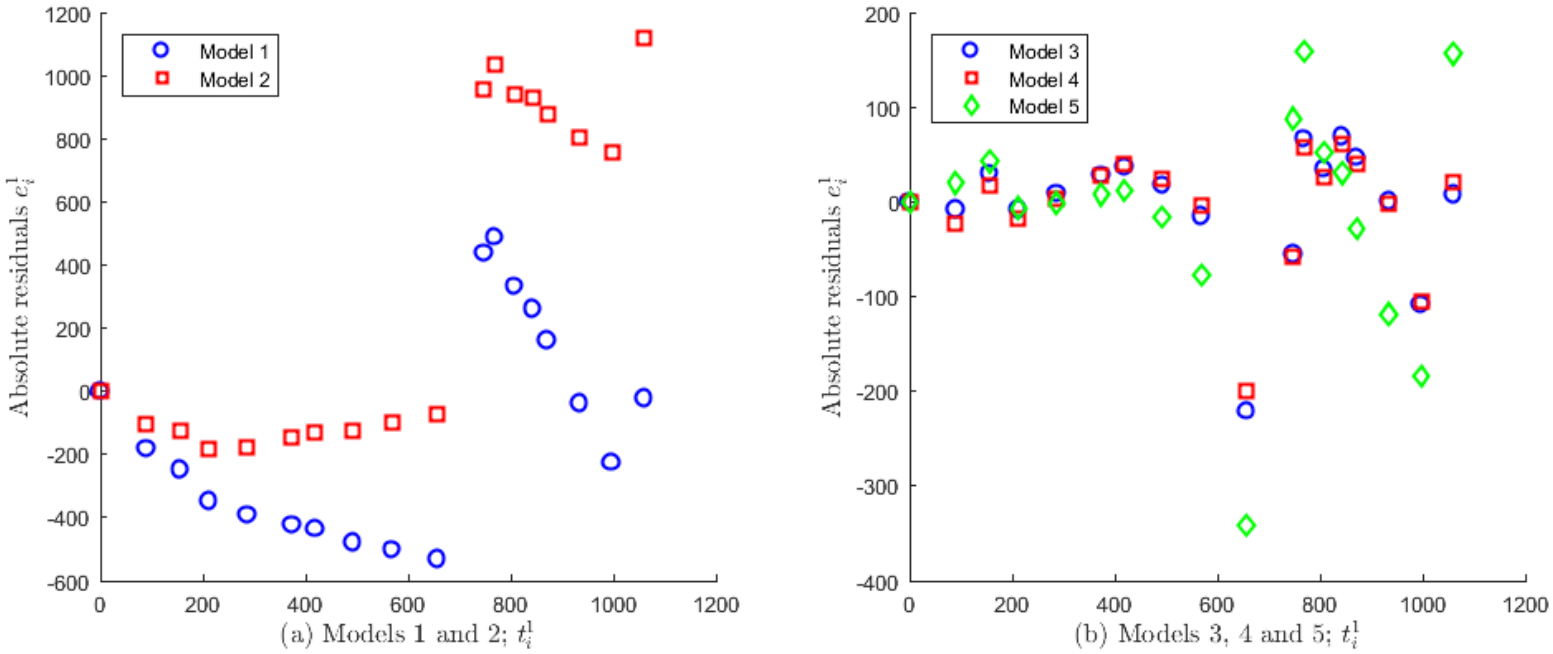

Figure 4: **Absolute residuals vs. time for Patient 1 data under constant variance data assumption**. A random pattern is observed for Models 3, 4 and 5, whereas a certain tendency is observed for models 1 and 2.

Fit to Patient 1 data

Validating statistical assumption (2/2)

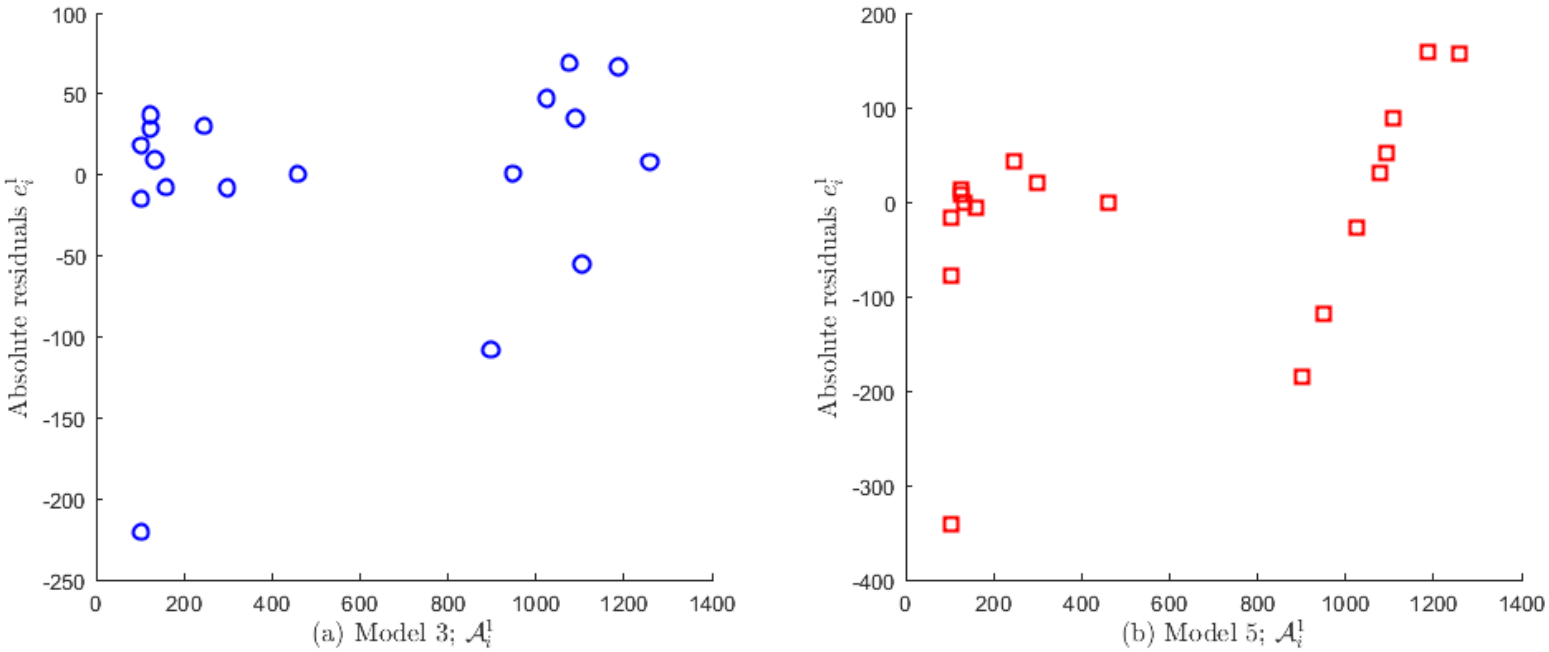

Figure 5: **Absolute residuals vs. observations for Patient 1 data under constant variance data assumption.** A random pattern is observed for Models 3 and 4 supporting the assumption of constant variance generated data, whereas a certain tendency is observed for Model 5.

Goodness of fit criteria for Patient 1 data

Comparison of the models proposed

We compared the five models by using the estimated variance $(\tilde{\sigma})^2$, RMSE $\hat{\sigma}$ and the coefficient of determination R^2 for data from Patient $j = 1$.

Table 2: **Statistics for the goodness of fit to Patient 1 data under constant variance data assumption.**

Model №	$\tilde{\sigma}^2$	RMSE	R^2
3	$4.4474e+3$	89.4726	0.9785
4	$3.7755e+3$	82.4368	0.9818

From Table 2, it can be observed that the best fit is achieved by Model 4 followed by Model 3. In consequence, once validated the statistical assumption on measuring errors, we can assure that Model 4 fits best to Patient 1 data.

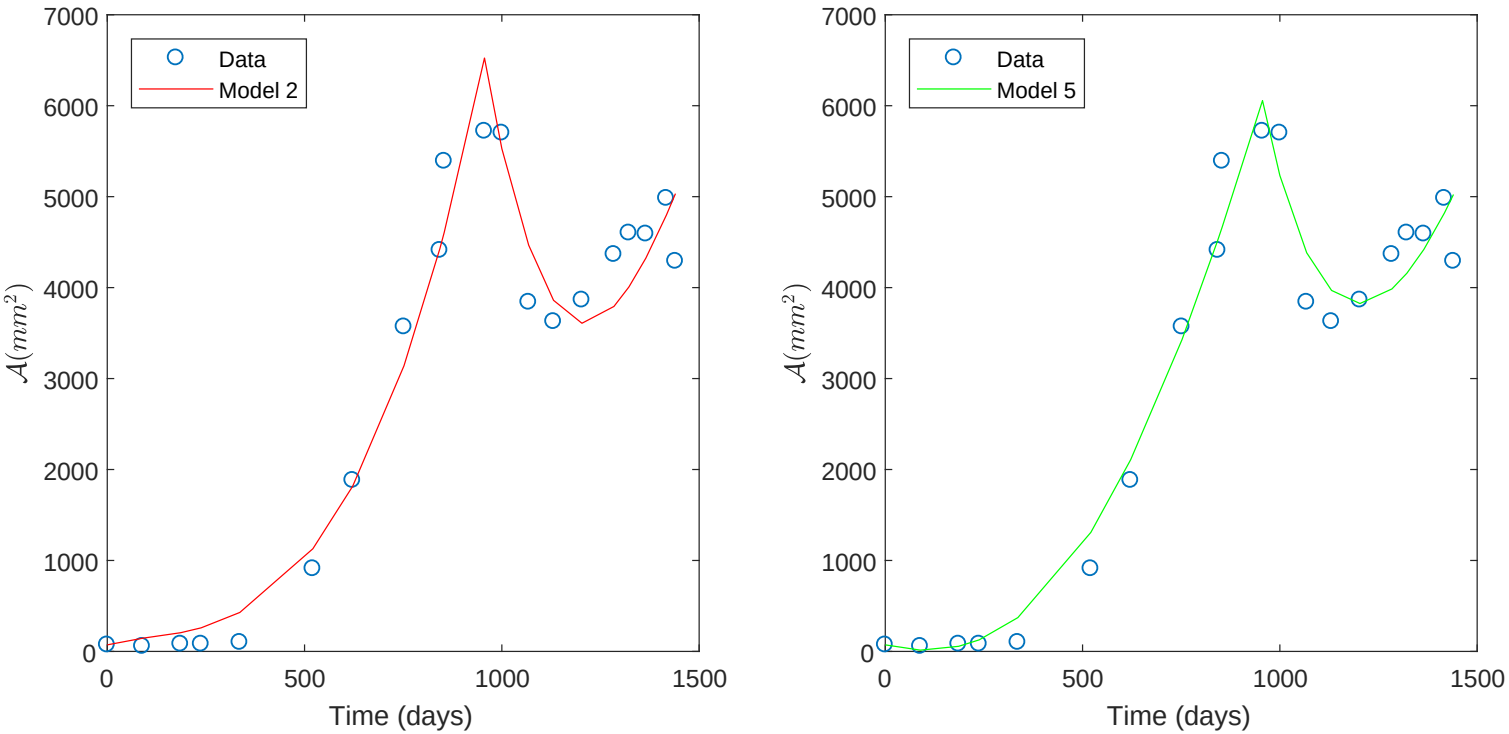

Figure 6: **Fit to Patient 2 data**

Fit to Patient 2 data

Validating statistical assumption (1/2)

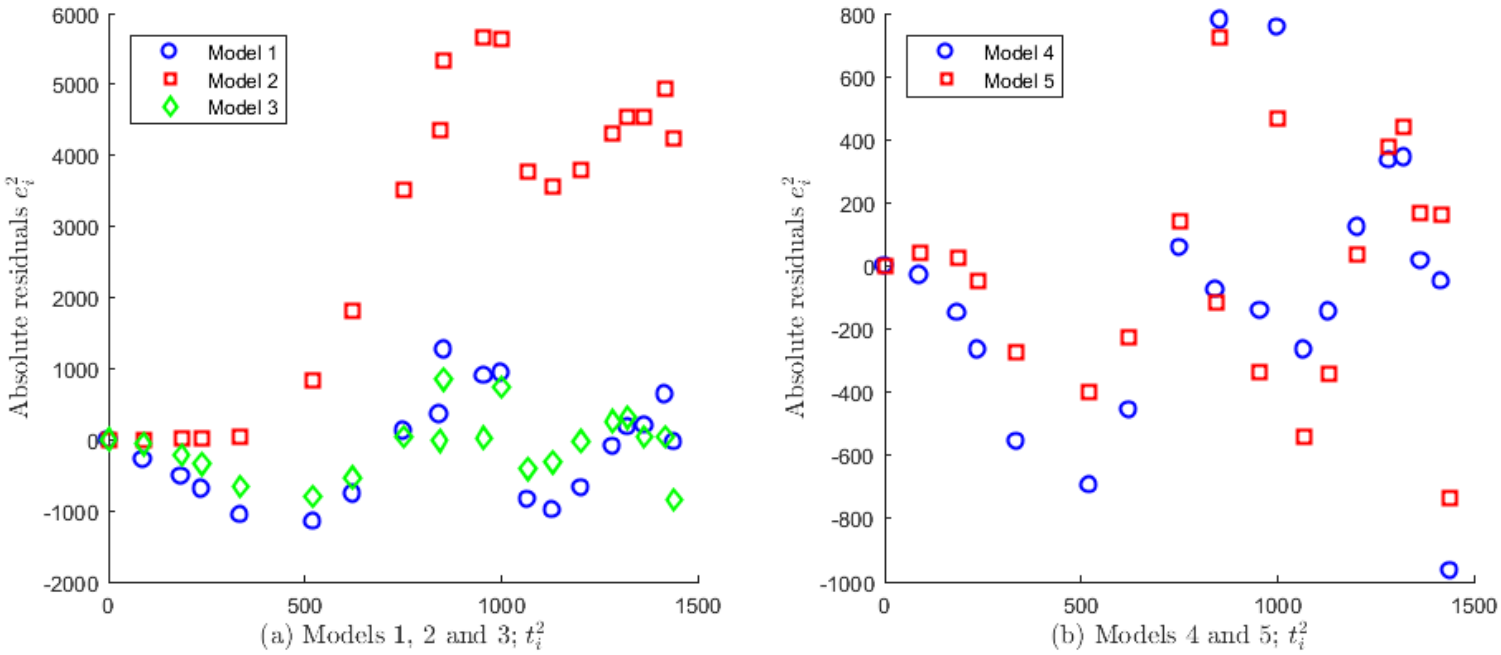

Figure 7: **Absolute residuals vs. time for Patient 2 data under constant variance data assumption.** A certain tendency is observed for Models 1, 2 and 3, whereas a random pattern is observed for Models 4 and 5.

Fit to Patient 2 data

Validating statistical assumption (2/2)

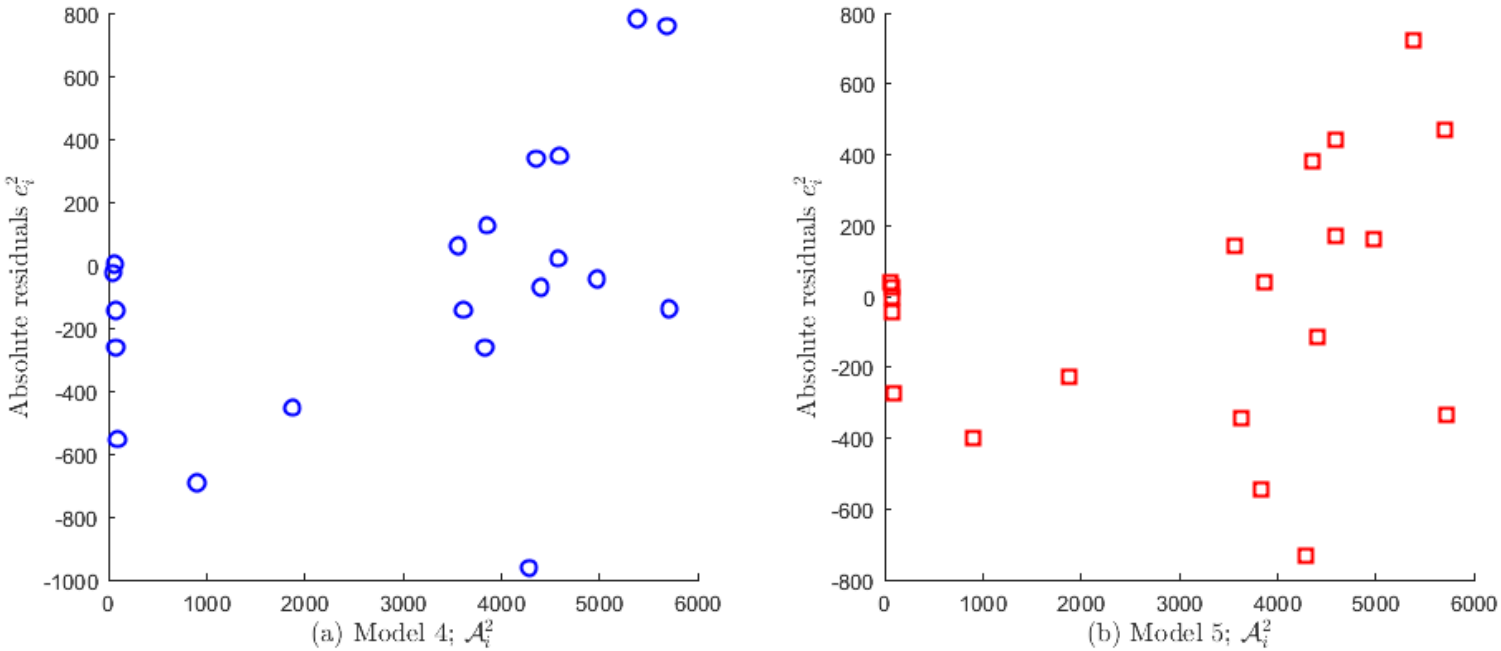

Figure 8: **Absolute residuals vs. observations for Patient 2 data under constant variance data assumption.** A random pattern is observed for Models 4 and 5 supporting the assumption of constant variance generated data.

Goodness of fit criteria for Patient 2 data

Comparison of the models proposed

We compared the five models by using the estimated variance $(\tilde{\sigma})^2$, RMSE $\hat{\sigma}$ and the coefficient of determination R^2 for data from Patient $j = 2$.

Table 3: **Statistics for the goodness of fit for Patient 2 data under constant variance data assumption.**

Model №	$\tilde{\sigma}^2$	RMSE	R^2
4	1.7843e+05	5.4533e+02	0.95811
5	1.2611e+05	4.5846e+02	0.97039

Data in Table 3 shows that the best fit is achieved for Model 5 followed by Model 4. On the other hand, as before, the accuracy of parameter estimation for Models 4 and 5 is considered qualitatively assessed through validation of the statistical assumption satisfied for the measuring errors.

Summary

So, what have we done?

1. Propose a reduction from a PDE model into a general ODE model, which direct problem represents different realistic situations for the disease.
2. Propose an inverse problem in order to estimate parameters for the model (therefore, obtaining a prediction of the disease), starting only from CT scans.
3. Propose regularization techniques in order to deal with noisy data and possible numerical problems in 'the bad case'.
4. Propose an algorithm to solve the inverse problem, taking into account our poor knowledge of good initial parameters.
5. Obtain, finally, suitable solutions for the inverse problem for the existing data.

Future Work

1. What about theoretical properties of proposed modeling? Which is the minimal number of observations N to have, for example, an unique set of parameters $\hat{\theta}$ which globally minimizes $S(\hat{\theta})$? (if it exists).
2. How to pick α for the regularization process? The size of α may change a lot the results. There are techniques to calibrate it, for instance Morozov principle.
3. To improve modeling with richer data in order to predict drug resistance (functional imaging, TEP or MRI).
4. Is it computationally possible to work in the context of spatial heterogeneity? (PDE models).
5. Is it possible to adapt the model for other diseases?

Some references I


Jean-Yves Blay,
A decade of tyrosine kinase inhibitor therapy: Historical and current perspectives on targeted therapy for GIST.
Cancer Treatment Reviews, 37(5):373-384, 2011.


JY Blay, A Le Cesne, PA Cassier, and IL Ray-Coquard,
Gastrointestinal stromal tumors (GIST): a rare entity, a tumor model for personalized therapy, and yet ten different molecular subtypes.
Discov Med, 13(72):357-67, May 2012.


P. Cumsille, A. Coronel, C. Conca, C. Quiñinao, and C. Escudero,
Proposal of a hybrid approach for tumor progression and tumor-induced angiogenesis,
Theoretical Biology and Medical Modelling, 12 (2015), p. 13.

Acknowledgments

Centre for Biotechnology and Bioengineering
PIA-CONICYT grant PFBasal-01

Thank you for your attention