Parameter estimation and uncertainty quantification in colon cancer-induced angiogenesis using a Fokker-Planck stochastic framework

Souvik Roy

Department of Mathematics,
University of Texas at Arlington

April 2, 2021

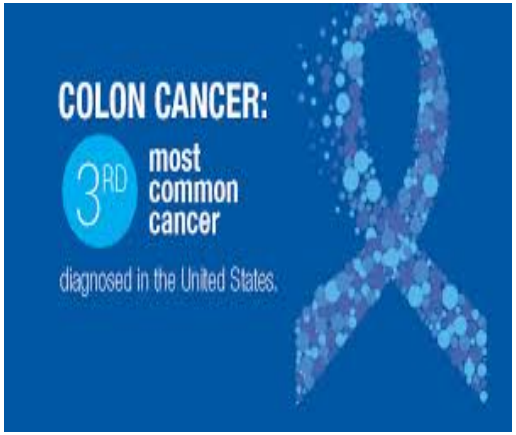
Acknowledgement

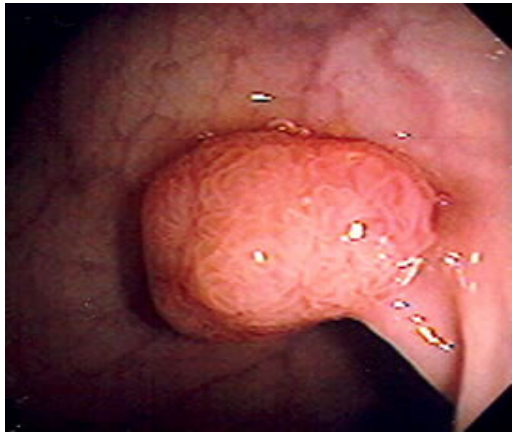
This is a joint work with Suvra Pal (Department of Mathematics and Zui Pan (College of Nursing and Health Innovation).

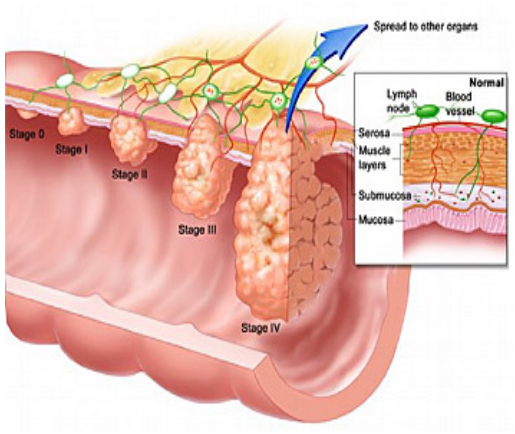
The work of S. Pal and S. Roy was funded by NCI-NIH (Award Number R21 CA242933). The work of Z. Pan was funded by NIH (Award Number R01 CA185055)

Reference: Suvra Pal, Zui Pan and Souvik Roy. A stochastic framework for parameter estimation and uncertainty quantification in colon cancer induced angiogenesis, *(Under review)*, 2021.

Colon cancer

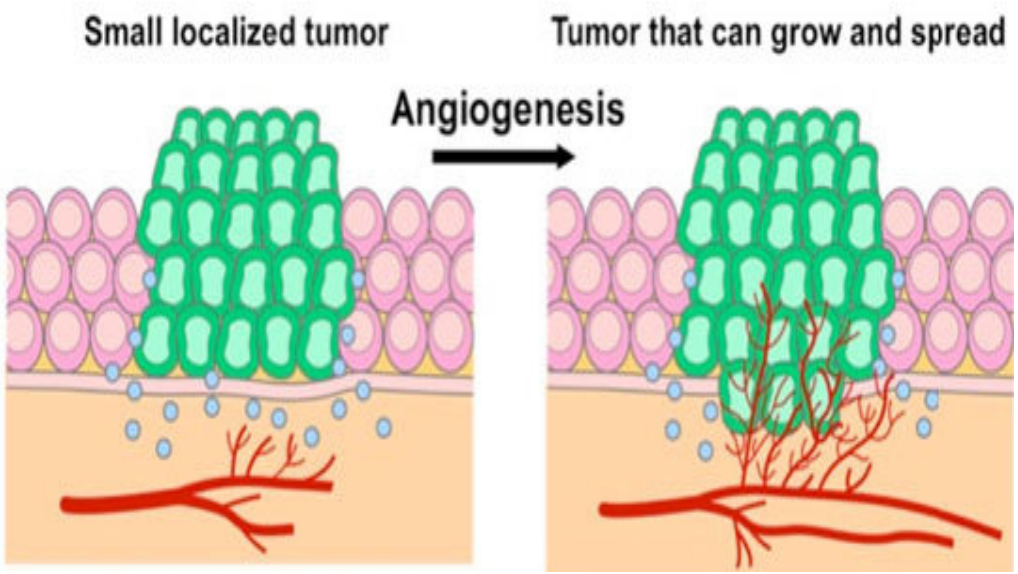






- Colon cancer is the 3^{rd} leading cause of cancer deaths in USA.
- Colon cancer begins in the large intestine (colon), which is the final part of the digestive tract.
- Though colon cancer usually affects older adults, it can happen to people of any age.
- It usually begins as small, noncancerous (benign) clumps of cells called polyps that form on the inside of the colon. Over time some of these polyps develop into colon cancers.
- Since polyps are usually small, and produce very few symptoms, colon cancer gets diagnosed at a very later stage.

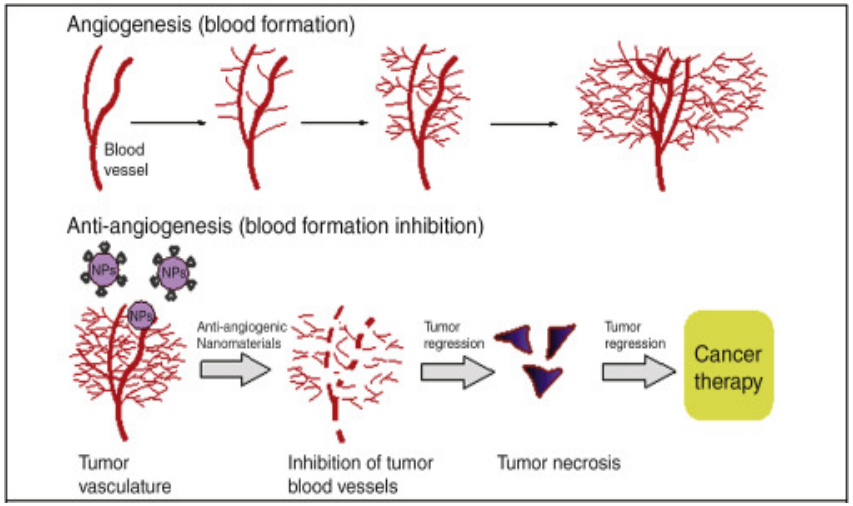
Angiogenesis I



- Angiogenesis is the process by which new blood vessels grow that allows the delivery of oxygen and nutrients to the body's tissues.
- It is an important process that is required for the growth and development as well as the healing of wounds.
- However, on the flipside, angiogenesis also helps in the growth of tumors.
- Growth factors are proteins that stimulate cell growth, differentiation, survival, inflammation, and tissue repair.

Angiogenesis II

- Several growth factors, called tumor angiogenic factors (TAF), have been identified that regulate angiogenesis in colon cancer (e.g. vascular endothelial growth factor (VEGF)).



- Thus, in recent years, angiogenesis has become a potential biomarker and, subsequently, there arises a need to develop efficient anti-angiogenic drugs.

Some challenges in clinical cancer research

- It is essential to understand pharmaceutical concepts, such as physiological effects, tumor shrinkage, pharmacokinetics and pharmacodynamics (PK/PD), therapeutic window, maximum tolerable dose (MTD).
- A large number of clinical studies are often required to understand such complex mechanisms of cancers and associated treatments.
- Another major clinical challenge is to obtain an effective treatment strategy for each patient or at least identify a subset of patients who could benefit from a particular treatment.
- It is very expensive to test even one therapy during clinical trials, let alone testing combination therapies that are far more effective.
- In this context, mathematical and computational methods provide a fast and effective framework for treatment assessments.

Quantitative systems pharmacology

- One of the main computational tools used to discover, test, and predict dose exposure response is quantitative systems pharmacology (QSP) modeling.
- These models help us to understand biological processes, identify targets and biomarkers, predict treatment outcomes, and discover mechanisms of drug resistance.
- They are the main tools to test the impact of drug parameters and biological variances on drug efficacy and safety.
- QSP models are represented by a set of differential equations that describe the dynamical properties of the cancer mechanism and interaction with drugs.
- Several QSP models have been developed for cancer angiogenesis (e.g., Chaplain '00, Mohammadi et. al. '08, Capasso and Morale '09, Vilanova et. al. '17). However, almost all of these models are deterministic in nature, and, are unable to capture the inherent randomness in the cancer dynamics (e.g., due to proliferation).

Parameter estimation

- Traditionally, it has been thought that the precision of QSP's parameter estimation is not very important.
- Therefore, parameters are commonly estimated using the data that are often assembled from disparate sources rather than a single curated dataset.
- These may span multiple biological studies, in vitro, in vivo, and clinical assays, and may be qualitative. As a result, they cannot be easily validated or used for personalized treatments.
- Since each tumor has its own unique characteristics, they respond differently to different treatments.
- Therefore, we need to develop QSP models that take into account each individual's tumor characteristics.

A new framework

- In this work, we contribute to the field of cancer research by presenting a new stochastic QSP model for representing colon cancer-induced angiogenesis, using a Fokker-Planck (FP) setup.
- Furthermore, we develop a new FP optimization framework and sensitivity analysis approach to obtain and analyze the unknown parameters of the stochastic model from the patient data. This is an inverse problem.
- The results obtained suggests the type of drugs that can be used to control colon cancer-induce angiogenesis.

A QSP model for angiogenesis

- We start off with a QSP model formulated by the authors in Csercsik and Kovács '19.
- The QSP model describes the dynamics of the vasculature volume and the interplay between the tumor and vasculature volumes.
- The dynamics of TAF, which is produced by unsupported tumor and initiates the formation of new blood vessels from existing ones, is also included in the model
- The model variables are as follows
  1. $V(t)$- the total tumor cell volume (cm^3)
  2. $B(t)$- the vasculature volume in the tumor (cm^3)
  3. $T(t)$- the concentration of tumor angiogenic factors (TAF) in the tumor (mg/ml).

The ODE system

- The governing system of ODEs, representing the dynamics of the above defined variables are given as follows

$$\frac{dV}{d\tau} = c\gamma V,\ V(0) = V_0,$$
$$\frac{dB}{d\tau} = c_e c\gamma V + c_v TB,\ B(0) = B_0,$$
$$\frac{dT}{d\tau} = c_T(1-\gamma) - q_T T,\ T(0) = T_0.$$

- The unknown patient-specific parameters that need to be determined is the parameter vector $\boldsymbol{\theta} = (c, c_e, c_v, c_T, q_T, \gamma)$, defined as follows

1. c- growth rate of tumor (day^{-1})
2. c_e- rate of internalization of new vasculature from the environment
3. c_v- rate of formation of new blood vessels due to TAF (ml mg^{-1} day^{-1})
4. c_T- rate of production of TAF (mg ml^{-1} day^{-1})
5. q_T- rate of removal of TAF from tumor (day^{-1})
6. γ- ratio of well-supported tumor cells inside the tumor volume

The non-dimensionless ODE system

- For stabilization and scalability of the numerical algorithms, we non-dimensionalize the above ODE system using the following non-dimensionalized state and time variables, and parameters

$$\bar{V} = k_1 V,\ \bar{B} = k_2 B,\ \bar{T} = k_3 T,\ t = k_4 \tau,$$

$$\bar{c} = \frac{c}{k_4},\ \bar{c}_e = \frac{c_e k_2}{k_1},\ \bar{c}_v = \frac{c_v}{k_3 k_4},\ \bar{c}_T = \frac{c_T k_3}{k_4},\ \bar{q}_T = \frac{q_T}{k_4}.$$

- The transformed non-dimensionless ODE system is given as follows

$$\frac{d\bar{V}}{dt} = \bar{c}\gamma\bar{V},\ \bar{V}(0) = \bar{V}_0,$$

$$\frac{d\bar{B}}{dt} = \bar{c}_e\bar{c}\gamma\bar{V} + \bar{c}_v\bar{T}\bar{B},\ \bar{B}(0) = \bar{B}_0,$$

$$\frac{d\bar{T}}{dt} = \bar{c}_T(1-\gamma) - \bar{q}_T\bar{T},\ \bar{T}(0) = \bar{T}_0.$$

A stochastic differential dynamical model for angiogenesis I

- We extend the QSP model to the following Itó stochastic ODE

$$\begin{aligned}
\frac{d\bar{V}}{dt} &= \bar{c}\gamma\bar{V} + \sigma_1\, dW_1(t),\ \bar{V}(0) = \bar{V}_0,\\
\frac{d\bar{B}}{dt} &= \bar{c}_e\bar{c}\gamma\bar{V} + \bar{c}_v\bar{T}\bar{B} + \sigma_2\, dW_2(t),\ \bar{B}(0) = \bar{B}_0,\\
\frac{d\bar{T}}{dt} &= \bar{c}_T(1-\gamma) - \bar{q}_T\bar{T} + \sigma_3\, dW_3(t),\ \bar{T}(0) = \bar{T}_0.
\end{aligned}$$

  where $dW_i,\ i = 1, 2, 3$ are one-dimensional Wiener processes and $\sigma_i,\ i = 1, 2, 3$, are positive constants.

- Using a compact notation, we can write the above equation as

$$\begin{aligned}
\frac{d\boldsymbol{X}}{dt} &= \boldsymbol{F}(\boldsymbol{X}, \boldsymbol{\theta}) + \boldsymbol{\sigma}\, d\boldsymbol{W}(t),\\
\boldsymbol{X}(0) &= \boldsymbol{X}_0,
\end{aligned}$$

A stochastic differential dynamical model for angiogenesis II

- Here

$$d\boldsymbol{W}(t) = \begin{pmatrix} dW_1(t) \\ dW_2(t) \\ dW_3(t) \end{pmatrix}$$

is a 3-dimensional Wiener process with stochastically independent components and

$$\boldsymbol{\sigma} = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix}$$

is the dispersion matrix.

- Aim: Given time series data for $\boldsymbol{X}$, find the unknown parameter set $\boldsymbol{\theta}$.

Optimization of stochastic processes

- An optimization problem can be formulated as follows: determine a parameter set $\boldsymbol{\theta}$ that optimizes a given objective given by a cost functional J.
- Since $\boldsymbol{X}$ is random and represents an outcome of a probability space, therefore a direct insertion of $\boldsymbol{X}$ into the objective J results into a random variable.
- For this reason, in stochastic optimal control, the following average of the cost functional is, usually, considered

$$J(\boldsymbol{X}(t), \boldsymbol{\theta}) = \mathbb{E}\left[\int_0^T L(t, \boldsymbol{X}(t), \boldsymbol{\theta})\, dt + \Psi[\boldsymbol{X}(T)]\right],$$

  where L and Ψ are continuous functions.
- To solve for $\boldsymbol{\theta}$, the method of dynamic programming can be applied in order to formulate the Hamilton-Jacobi-Bellman (HJB) equation for $\min_{\boldsymbol{\theta}} J$ with $\boldsymbol{\theta}$ the optimization variable.

The Fokker-Planck equation I

- On the other hand, one can note that the state of a random process is characterized by the shape of the corresponding probability density function (PDF).
- The evolution of the PDF, denoted with $f = f(x,t)$, associated to the stochastic process is modeled by the Fokker-Planck (FP) equation.

$$\partial_t f(x,t) + \nabla \cdot (\boldsymbol{F}(x,\boldsymbol{\theta})\, f(x,t)) = \frac{1}{2}\nabla \cdot (\boldsymbol{\sigma}^2 \nabla f(x,t)), \quad f(x,0) = f_0(x)$$

- Here $f(x,t)$ is the probability that $X(t) = x$, where $x \in \Omega \subset \mathbb{R}^3_+ = \{x \in \mathbb{R}^3 : x_i \geq 0,\ i = 1,2,3\}$.
- Notice that the FP equation can be written in flux form: $\partial_t f(x,t) = \nabla \cdot F$ where the flux F is given by

$$\mathcal{F}_j(x,t;f) = \frac{\sigma_j^2}{2}\partial_{x_j} f - \boldsymbol{F}_j(x,\boldsymbol{\theta}) f,\ j = 1,2,3.$$

- We chose non-reflecting boundary conditions as follows

$$F \cdot n = 0 \qquad \text{on } \partial\Omega \times (0,T).$$

The Fokker-Planck equation II

- We also have the conditions of positivity $f(x,t) \geq 0$ and conservativeness

$$\int_{\Omega} f(x,t)dx = \int_{\Omega} f_0(x)dx = 1.$$

- Our goal is to use the FP optimization framework to accurately and robustly obtain $\boldsymbol{\theta}$.

Results with the FP optimization framework I

The FP optimization has been successfully applied to Itō processes, subdiffusion processes, PDP processes arising in applications related to pedestrian motion, game theory, collective motion.

- M. Annunziato and A. Borzì, *A Fokker-Planck control framework for multidimensional stochastic processes*, Journal of Computational and Applied Mathematics, 237 (2013), 487–507.
- M. Annunziato and A. Borzì, *Optimal control of piecewise deterministic processes*, European Journal of Applied Mathematics, 25 (2014), 1–25.
- M. Annunziato, A. Borzì, M. Magdziarz, A. Weron, *A fractional Fokker-Planck control framework for subdiffusion processes*, Optimal Control, Applications and Methods, 2015.
- Souvik Roy, Mario Annunziato and Alfio Borzì. A Fokker–Planck feedback control-constrained approach for modelling crowd motion, *Journal of Computational and Theoretical Transport*, 45(6): 452–458, 2016.
- Souvik Roy, Alfio Borzì and Abderrahmane Habbal. Pedestrian motion constrained by FP-constrained Nash games, *Royal Society Open Science*, 4(9):170648, 2017.

Results with the FP optimization framework II

- Souvik Roy, Mario Annunziato, Alfio Borzì and Christian Klingenberg. A Fokker-Planck approach to control collective motion, *Computational Optimization and Applications*, 69(2):423–459, 2018.
- Souvik Roy. A sparsity-based Fokker-Planck optimal control framework for modeling traffic flows, *AIP Conference Proceedings*, 2302:110007, 2020.

To the best of our knowledge, this is the first work on FP optimization frameworks in colon cancer.

We formulate a FP optimization framework that takes into account inconsistencies and inaccuracies in data measurement, yet can provide an accurate estimate of the unknown parameter set $\boldsymbol{\theta}$.

Furthermore, we also carry out a sensitivity analysis of the optimal parameter set to determine the parameters that need to be controlled, thus, providing information of the type of drugs that can be used for treatment.

A FP optimization problem for parameter estimation I

- We will estimate the patient specific unknown parameter vector $\boldsymbol{\theta}$, given the values of $f(x,t)$ at specific time instants $t_1, \cdots, t_N$ as $f_i^*(x),\ i = 1, \cdots, N$, i.e, the data is given as PDFs that helps account for inconsistencies and inaccuracies in data measurements.
- We solve the following constrained optimization problem

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta} \in U_{ad}}{\operatorname{argmin}}\, J(f, \boldsymbol{\theta}) := \frac{\alpha}{2} \int_Q \left(f(x,t) - f^*(x,t)\right)^2 \, dx + \frac{\beta}{2} \|\boldsymbol{\theta}\|_{l^2}^2,$$

  subject to the FP system, where U_{ad} is an admissible set for the range of values of $\boldsymbol{\theta}$.
- The function $f^*(x,t)$ is the data function formed by interpolating the patient data $f_i^*(x)$ for all times t.
- The first term in J is a least-squares data fitting term whereas the second term is a l^2 regularization term that is based on the fact that the biological parameter values are bounded.
- We denote the FP equation above as $\mathcal{E}(f_0, \boldsymbol{\theta}) = 0$

A FP optimization problem for parameter estimation II

- Notice that the map $\boldsymbol{\theta} \to f = f(\boldsymbol{\theta})$ is twice differentiable. We introduce the reduced cost functional $\hat{J}$ as follows

$$\hat{J}(\boldsymbol{\theta}) = J(f(\boldsymbol{\theta}), \boldsymbol{\theta})).$$

- Thus, we can formulate an unconstrained minimization problem as

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta} \in U_{ad}}{\operatorname{argmin}} \hat{J}(f(\boldsymbol{\theta}), \boldsymbol{\theta}).$$

Theory of the FP optimization problem I

Proposition (Solution of FP equation)

Let $f_0 \in H^1(\Omega)$, $f_0 \geq 0$, and $\boldsymbol{\theta} \in U_{ad}$. Then, there exists an unique non-negative solution of $\mathcal{E}(f_0, \boldsymbol{\theta}) = 0$ given by $f \in L^2(0, T; H^1(\Omega)) \cap C([0, T]; L^2(\Omega))$.

Proposition (Conservation of total probability)

The FP system is conservative.

Proposition ($L^2(\Omega)$ stability)

The solution f of the FP system satisfies the following stability estimate

$$\|f(t)\|_{L^2(\Omega)} \leq \|f_0\|_{L^2(\Omega)} \exp\left(\|\boldsymbol{\sigma}^{-1}\|_2^2 N^2 t\right),$$

where $N = \sup_{\Omega \times U_{ad}} |F(x, \boldsymbol{\theta})|$.

Theory of the FP optimization problem II

Proposition (Other stability results)

Let $f_0 \in H^1(\Omega)$, $f_0 \geq 0$, and $\boldsymbol{\theta} \in U_{ad}$. Then, if f is a solution to $\mathcal{E}(f_0, \boldsymbol{\theta}) = 0$, the following inequalities hold

$$\|f\|_{L^\infty(0,T;L^2(\Omega))} \leq c_1 \|f_0\|_{L^2(\Omega)},$$

$$\|\partial_t f\|_{L^2(0,T;H^{-1}(\Omega))} \leq (c_2 + c_3 N)\|f_0\|_{L^2(\Omega)},$$

where c_1, c_2, c_3 are positive constants and N is defined above. Further, if $\|\boldsymbol{\sigma}^{-1}\|_2^2 > \frac{1}{N}$, then the following inequality holds

$$\|f\|_{L^2(0,T;H^1(\Omega))} \leq c_4 \|f_0\|_{L^2(\Omega)},$$

where c_4 is a positive constant.

Theorem (Existence of a minimizer)

Let $f_0 \in H^1(\Omega)$. Then, there exists a pair $(f^*, \boldsymbol{\theta}^*) \in C([0,T]; H^1(\Omega)) \times U_{ad}$ such that f^* is a solution to $\mathcal{E}(f_0, u^*) = 0$ and $\boldsymbol{\theta}^*$ minimizes J in U_{ad}.

First-order necessary optimality condition I

Proposition (Differentiability of the reduced functional)

The reduced functional $\hat{J}(u)$ is differentiable and its derivative is given by

$$d\hat{J}(\boldsymbol{\theta}) \cdot \boldsymbol{\psi} = \left\langle \beta\boldsymbol{\theta} - \nabla_{\boldsymbol{\theta}} \boldsymbol{F} \cdot \nabla p, \boldsymbol{\psi} \right\rangle_{l^2}, \qquad \forall \boldsymbol{\psi} \in \mathbb{R}^6,$$

where p is the solution to the adjoint equation

$$-\partial_t p(x,t) - f(x,t)(\boldsymbol{F}(x,\boldsymbol{\theta}) \cdot \nabla p(x,t)) - \frac{1}{2}\nabla \cdot (\boldsymbol{\sigma}^2 \nabla p(x,t)) = -\alpha \sum_{i=1}^{N} (f(x,t_i) - f_i^*(x)),$$

$$\frac{\partial p}{\partial n} = 0, \qquad \text{on } \partial\Omega \times (0,T),$$

with $p(x,T) = 0$.

First-order necessary optimality condition II

- The optimality conditions can now be written as

$$\partial_t f(x,t) + \nabla \cdot (\boldsymbol{F}(x,\boldsymbol{\theta})\, f(x,t)) = \frac{1}{2}\nabla \cdot (\boldsymbol{\sigma}^2 \nabla f(x,t)), \quad \text{in } \Omega \times (0,T),$$
$$f(x,0) = f_0(x) \ \text{ in } \Omega, \quad \mathcal{F} \cdot \hat{n} = 0 \ \text{ on } \partial\Omega \times (0,T).$$

(FOR)

$$-\partial_t p(x,t) - f(x,t)(\boldsymbol{F}(x,\boldsymbol{\theta}) \cdot \nabla p(x,t)) - \frac{1}{2}\nabla \cdot (\boldsymbol{\sigma}^2 \nabla p(x,t))$$
$$= -\alpha(f(x,t) - f^*(x,t)), \quad \text{in } \Omega \times (0,T),$$
$$p(x,T) = 0 \ \text{ in } \Omega, \quad \frac{\partial p}{\partial n} = 0 \ \text{ on } \partial\Omega \times (0,T).$$

(ADJ)

$$\left\langle \beta\boldsymbol{\theta} - \nabla_{\boldsymbol{\theta}} \boldsymbol{F} \cdot \nabla p, \boldsymbol{\psi} - \boldsymbol{\theta} \right\rangle_{l^2} \geq 0, \qquad \forall \boldsymbol{\psi} \in \mathbb{R}^6. \tag{OPT}$$

Discretization of the FP equation and of its adjoint I

- We implement and analyze an time splitting Douglas-Gunn (DG3) scheme combined with the positive and conservative second-order Chang-Cooper (CC) space-discretization scheme.
- The DG3-CC scheme for the FPK equation can be written as follows

$$
\begin{aligned}
\frac{f^{m^*}_{i,j,k} - f^{m}_{i,j,k}}{\delta t} &= \frac{1}{2h}(\mathcal{F}^{m^*}_{i+\frac{1}{2}} - \mathcal{F}^{m^*}_{i-\frac{1}{2}}) \\
&+ \frac{1}{2h}(\mathcal{F}^{m}_{i+\frac{1}{2}} - \mathcal{F}^{m}_{i-\frac{1}{2}}) + \frac{1}{h}(\mathcal{F}^{m}_{j+\frac{1}{2}} - \mathcal{F}^{m}_{j-\frac{1}{2}}) + \frac{1}{h}(\mathcal{F}^{m}_{k+\frac{1}{2}} - \mathcal{F}^{m}_{k-\frac{1}{2}}), \\
\frac{f^{m^{**}}_{i,j,k} - f^{m}_{i,j,k}}{\delta t} &= \frac{1}{2h}(\mathcal{F}^{m^*}_{i+\frac{1}{2}} - \mathcal{F}^{m^*}_{i-\frac{1}{2}}) + \frac{1}{2h}(\mathcal{F}^{m^{**}}_{j+\frac{1}{2}} - \mathcal{F}^{m^{**}}_{j-\frac{1}{2}}) \\
&+ \frac{1}{2h}(\mathcal{F}^{m}_{i+\frac{1}{2}} - \mathcal{F}^{m}_{i-\frac{1}{2}}) + \frac{1}{2h}(\mathcal{F}^{m}_{j+\frac{1}{2}} - \mathcal{F}^{m}_{j-\frac{1}{2}}) + \frac{1}{h}(\mathcal{F}^{m}_{k+\frac{1}{2}} - \mathcal{F}^{m}_{k-\frac{1}{2}}), \\
\frac{f^{m+1}_{i,j,k} - f^{m}_{i,j,k}}{\delta t} &= \frac{1}{2h}(\mathcal{F}^{m^*}_{i+\frac{1}{2}} - \mathcal{F}^{m^*}_{i-\frac{1}{2}}) + \frac{1}{2h}(\mathcal{F}^{m^{**}}_{j+\frac{1}{2}} - \mathcal{F}^{m^{**}}_{j-\frac{1}{2}}) + \frac{1}{2h}(\mathcal{F}^{m+1}_{k+\frac{1}{2}} - \mathcal{F}^{m+1}_{k-\frac{1}{2}}) \\
&+ \frac{1}{2h}(\mathcal{F}^{m}_{i+\frac{1}{2}} - \mathcal{F}^{m}_{i-\frac{1}{2}}) + \frac{1}{2h}(\mathcal{F}^{m}_{j+\frac{1}{2}} - \mathcal{F}^{m}_{j-\frac{1}{2}}) + \frac{1}{2h}(\mathcal{F}^{m}_{k+\frac{1}{2}} - \mathcal{F}^{m}_{k-\frac{1}{2}}),
\end{aligned}
$$

Discretization of the FP equation and of its adjoint II

- At each step, the DG3-CC scheme is implicit in 1 direction and explicit in the other 2 directions.
- The discrete scheme for the adjoint FP equation is obtained using a combination DG3 scheme for the time discretization, one sided finite difference discretization for the advection term, and central difference for the diffusion term.

The Chang-Cooper scheme

- The FP equation can be written in flux form, $\partial_t f = \nabla \cdot F$, where

$$\nabla\cdot\mathcal{F} = \frac{1}{h}\left[(\mathcal{F}^m_{i+\frac{1}{2},j,k} - \mathcal{F}^m_{i-\frac{1}{2},j,k}) + (\mathcal{F}^m_{i,j+\frac{1}{2},k} - \mathcal{F}^m_{i,j-\frac{1}{2},k}) + (\mathcal{F}^m_{i,j,k+\frac{1}{2}} - \mathcal{F}^m_{i,j,k-\frac{1}{2}})\right].$$

- Here $\mathcal{F}^m_{i+\frac{1}{2},j,k}$, $\mathcal{F}^m_{i,j+\frac{1}{2},k}$, $\mathcal{F}^m_{i,j,k+\frac{1}{2}}$ represent the numerical flux in the i, j, k directions, respectively, at the point (x_{1i}, x_{2j}, x_{3k}).
- The numerical flux $\mathcal{F}^m_{i+\frac{1}{2},j,k}$ in the i^{th} direction is given as follows

$$\mathcal{F}^m_{i+\frac{1}{2},j,k} = \left[(1-\delta_i)B_{i+\frac{1}{2},j,k,m} + \frac{\sigma_i^2}{2h}\right] f^m_{i+1,j} - \left[\frac{\sigma_i^2}{2h} - \delta_i B_{i+\frac{1}{2},j,k,m}\right] f^m_{i,j},$$

where

$$B_{i+\frac{1}{2},j,m} = -\boldsymbol{F}_1(x_{1i+\frac{1}{2}}, x_{2j}, x_{3k}, \boldsymbol{\theta}),$$

and

$$\delta_i = \frac{1}{w^m_{i+\frac{1}{2},j}} - \frac{1}{\exp(w^m_{i+\frac{1}{2},j,k}) - 1}, \qquad w^m_{i+\frac{1}{2},j,k} = 2hB_{i+\frac{1}{2},j,k}/\sigma_i^2.$$

- A similar formulae also holds true for the fluxes in the other directions.

Numerical analysis of the DG3-CC scheme I

We assume that $\boldsymbol{F}$ is Lipschitz continuous with Lipshitz constant Γ.

Theorem (Positivity)

The DG3-CC scheme is positive under the Courant-Friedrichs-Lewy (CFL)-like condition

$$\delta t < \min\left\{\frac{2}{\Gamma}, \frac{2h^2}{V}\right\},$$

where

$$V = \frac{h\underline{B}}{e^{2h\underline{B}/\overline{C}} - 1} + \frac{h\overline{B}}{1 - e^{-2h\overline{B}/\overline{C}}},$$

$$\underline{B} = \min_{x,t}\{\boldsymbol{F}(x,t)\},\ \overline{B} = \max_{x,t}\{\boldsymbol{F}(x,t)\},\ \overline{C} = \max_{i}\sigma_i^2$$

Numerical analysis of the DG3-CC scheme II

Define the discrete L^1 norm as

$$\|f\|_1 = h^3 \sum_{i,j,k=0}^{N_x} |f_{i,j,k}|,$$

Lemma (Conservativeness)

The DG3-CC scheme is conservative, i.e., $\|f^m\|_1 = \|f^0\|_1$.

Theorem (Stability)

The solution $f^m_{i,j,k}$ *obtained using the DG3-CC scheme for the FP equation with a source* $g(x,t)$*, under the CFL-like condition given above, satisfies the following* L^1 *stability result*

$$\|f^m\|_1 \leq \|f^0\|_1 + \delta t \sum_{n=0}^{m} \max(\|g^n\|_1, \|g^{n-1/2}\|_1), \qquad m = 0, \ldots N_t - 1,$$

where $\|\cdot\|_1$ *is the discrete* L^1 *norm.*

Numerical analysis of the DG3-CC scheme III

Lemma (Truncation error)

The truncation error of the DG3-CC scheme is of order $\mathcal{O}(\delta t^2 + h^2)$ under the aforementioned CFL-like condition.

Theorem (Convergence)

The DG3-CC scheme is convergent with an error of order $\mathcal{O}(\delta t^2 + h^2)$ under the aforementioned CFL condition in the discrete L^1 norm.

A projected NCG scheme

Algorithm (Projected NCG Scheme)

1. *Input: initial approx.* $\boldsymbol{\theta}_0$. *Evaluate* $d_0 = -\nabla_{\boldsymbol{\theta}} \hat{J}(\boldsymbol{\theta}_0)$, *index* $k = 0$, *maximum* $k = k_{max}$, *tolerance* = tol.
2. *While* $(k < k_{max})$ *do*
3. *Set* $\boldsymbol{\theta}_{k+1} = P_U\left[\boldsymbol{\theta}_k + \alpha_k\, d_k\right]$, *where* α_k *is obtained using a line-search algorithm.*
4. *Compute* $g_{k+1} = \nabla_{\boldsymbol{\theta}} \hat{J}(\boldsymbol{\theta}_{k+1})$.
5. *Compute* β_k^{HG} *using Hager-Zhang formula.*
6. *Set* $d_{k+1} = -g_{k+1} + \beta_k^{HG} d_k$.
7. *If* $\|\boldsymbol{\theta}_{k+1} - \boldsymbol{\theta}_k\|_{l^2} < tol$, *terminate.*
8. *Set* $k = k + 1$.
9. *End while.*

Sensitivity analysis using LHS-PRCC I

- Specify a PDF for each uncertain parameter in the model. We use the PDF corresponding to the Weibull distribution, since the parameter values are positive.
- Divide each pdf into M non-overlapping and equiprobable intervals.
- For each uncertain parameter, randomly sample every equiprobable interval exactly once so that the entire range of each parameter is explored.
- Using these values, create the Latin hypercube sampling matrix of size $M \times 6$.
- Use each row (consisting of 6 values) of the LHS matrix to generate the output.
- Rank all 6 columns of the LHS matrix and call the resulting ranked matrix as $X_R = [X_{1R}, X_{2R}, \cdots, X_{6R}]$, where each X_{iR}, $i = 1, \cdots, k$, is a vector of dimension $(M \times 1)$ representing the rank transform for the i-th parameter. In a similar way, also rank the output vector of dimension $(M \times 1)$, and call the resulting ranked vector of output values as Y_R.

Sensitivity analysis using LHS-PRCC II

- For the i-th uncertain parameter ($i = 1, 2, \cdots, 6$), run two multiple linear regression (MLR) models. The first one is the MLR of X_{iR} on all $\{X_{jR} : j = 1, 2, \cdots, 6 \text{ and } j \neq i\}$ and the second one is the MLR of Y_R on all $\{X_{jR} : j = 1, 2, \cdots, 6 \text{ and } j \neq i\}$.
- Calculate the residuals from both MLR models. The Partial rank correlation coefficient (PRCC) value for the i-th parameter is the Pearson's correlation coefficient between these two sets of residuals.
- For each uncertain parameter, perform test of significance to assess if the corresponding PRCC value is significantly different from zero.
- Identify parameters with large PRCC values (> 0.5 or < -0.5) and corresponding small p-values (say, < 0.05 at 5% significance level).
- Finally, rank the sensitive parameters based on the magnitude of their PRCC values.

Numerical setup

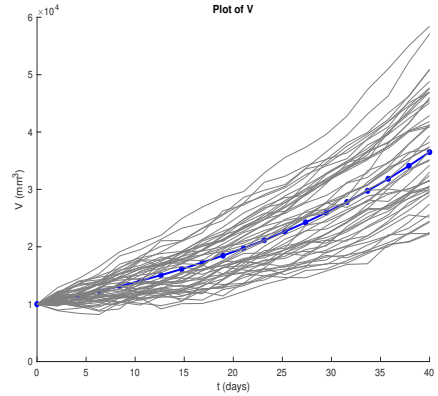
- We choose our domain $\Omega = (0, 6)^3$ and discretize it using $N_{x_i} = 51$ points for $i = 1, 2, 3$.
- The final time t is chosen to be 4 and the maximum number of time steps N_t is chosen to be 50.
- The patient data is represented by the target PDFs $f_i^*(x),\ i = 1, \cdots, N$ with $N = 10, 20$, where f_i^* are described by a normal distribution about the measured mean value $\mathbb{E}[f_i^*]$ and variance 0.05.
- We perform a 4D interpolation to obtain the data function $f^*(x, t)$ at all discrete times $t_k,\ k = 1, \cdots, N_t$.
- The values of $\sigma_1, \sigma_2, \sigma_3$ are chosen to be 0.2 to account for the measurement errors.
- The regularization parameters are chosen to be $\alpha = 1,\ \beta = 0.02$.

Datasets

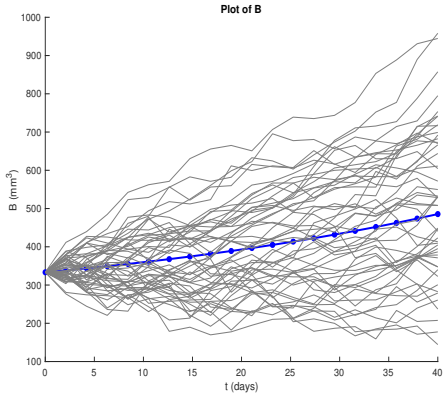
We use 2 data sets:

- Synthetic data measurements by solving the original ODE in the time interval $t = [0, 4]$ with $N = 10, 20$, and with the non-dimensional parameter set $\boldsymbol{\theta} = (1.3400, 0.0350, 0.1200, 1.1400, 0.2473, 0.5000)$.
- Real data based from Sápi et. al. '15, Csercsik et. al. '19 where mice specimens were transplanted subcutaneously with C38 colon adenocarcinoma, and small animal MRI was used to measure the tumor volume $\bar{V}$ in days 3, 5, 7, 9, 11, 15, 17, 19, 21, 23 and obtain the corresponding data for $\bar{B}$ and $\bar{T}$.

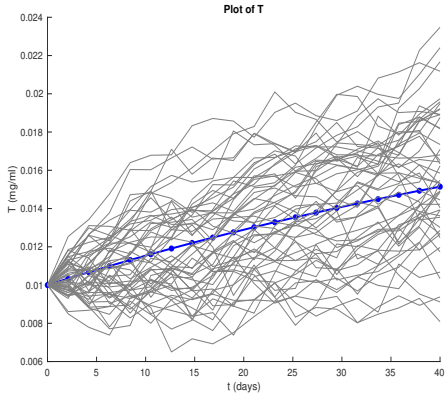
Test Case: Synthetic data ($N = 10$)



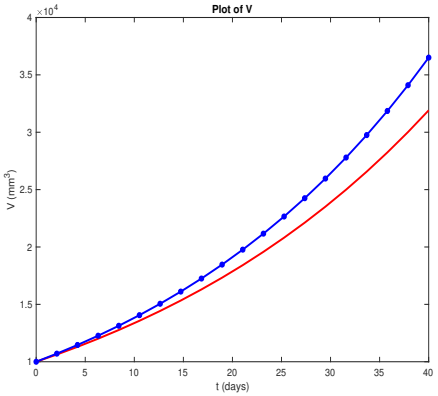
(a) V - MC plots



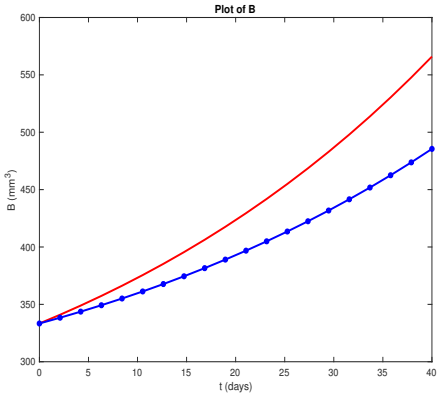
(b) B - MC plots



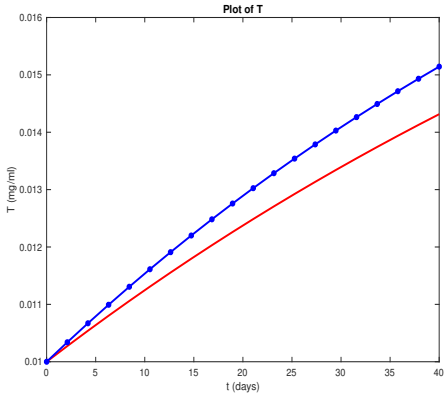
(c) T- MC plots



(d) V

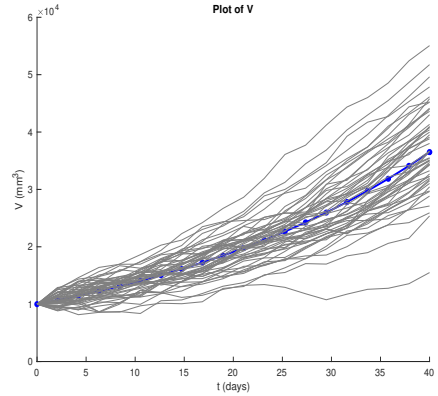


(e) B

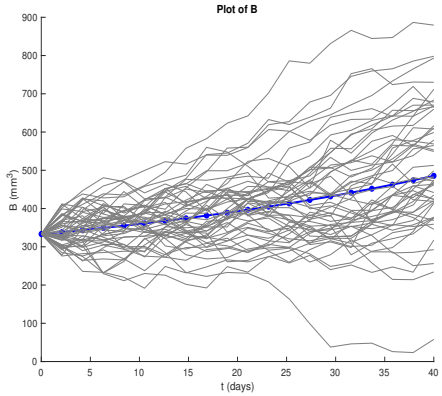


(f) T

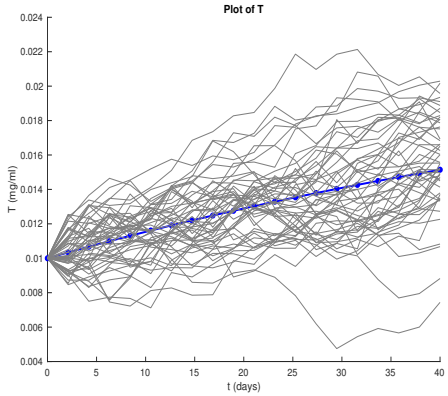
Test Case: Synthetic data ($N = 20$)



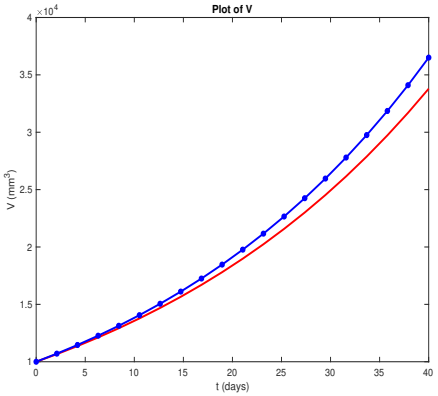
(a) V - MC plots



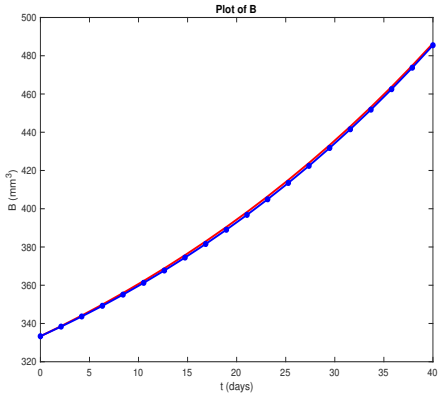
(b) B - MC plots



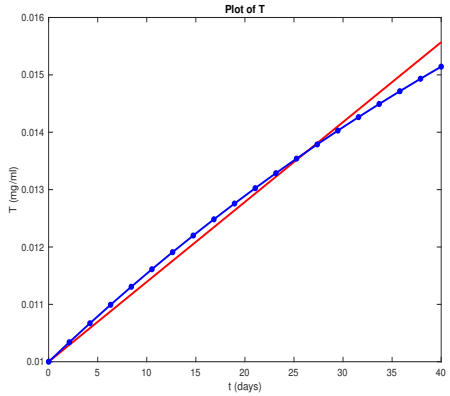
(c) T- MC plots



(d) V



(e) B



(f) T

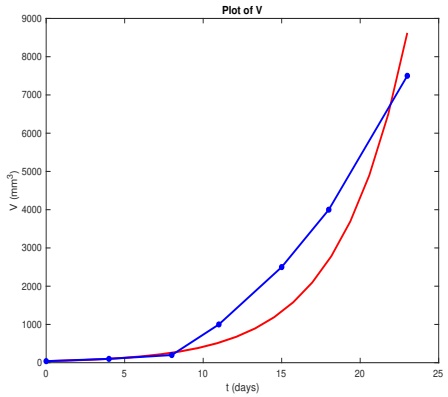
Sensitivity analysis

The obtained optimal parameter set $\boldsymbol{\theta}^*$ for $N = 10$ is $(0.9349, 0.042, 0.1478, 0.8505, 0.18, 0.6726)$ and with $N = 20$ is $(1.0359, 0.0383, 0.1940, 1.0877, 0.2000, 0.5775)$

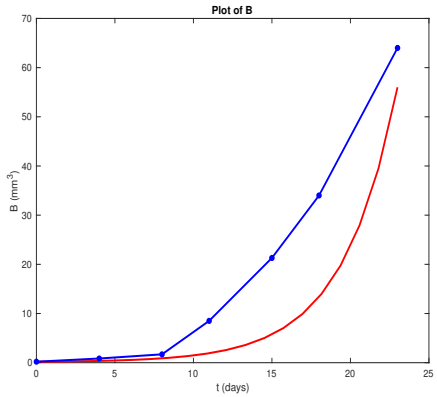
Parameter	$N = 10$		$N = 20$	
	$p-$value	PRCC value	$p-$value	PRCC value
c	2.652e-30	0.870	1.842e-29	0.864
c_e	0.629	-0.050	0.913	0.011
c_v	0.302	0.107	0.622	-0.051
c_T	0.840	0.021	0.368	0.093
q_T	0.696	0.041	0.125	0.158
γ	1.301e-44	0.938	2.933e-48	0.949

Table: $p-$values and PRCC values for the optimal parameter set $\boldsymbol{\theta}^*$

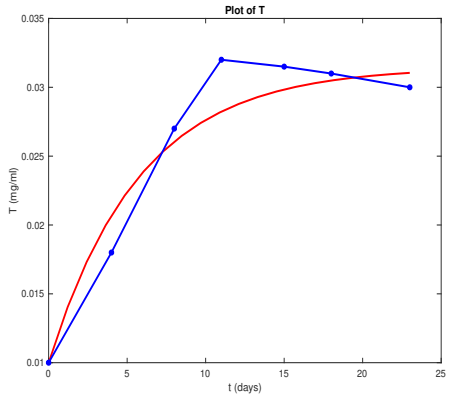
Test Case: Real data



(a) V



(b) B



(c) T

Optimal value	**Parameter**	$p-$value	PRCC value
1.5	c	7.166e-33	0.886
0.067	c_e	0.557	-0.061
0.2	c_v	0.287	0.110
1.22	c_T	0.179	0.139
0.39	q_T	0.311	-0.105
0.6	γ	3.040e-66	0.979

Treatment inference

- We observe that the parameters c and γ are the most sensitive ones with respect to the tumor volume V at the final time.
- This indicates that we need to consider a combination of anti-angiogenic drugs that can control the colon cancer growth rate c and ratio of well-supported tumor cells γ.
- Choices of such combination therapies include a mix of angiogenic inhibitors Bevacizumab, Ramucirumab, Regorafenib, Ziv-aflibercept and the chemotherapy drug FOLFIRI.
- A challenging question is to obtain the optimal dosage for such combination therapies such that the toxicity levels and drug costs are minimized.

Conclusions

- We have presented a new stochastic framework for parameter estimation and uncertainty quantification in colon cancer-induced angiogenesis.
- We characterize the stochastic process using the PDF, whose evolution is governed by the FP equation.
- The coefficients in the FP equation represent the unknown patient specific parameters that we estimate using the patient data, by formulating a PDE-constrained optimization problem.
- We presented robust discretization schemes to solve for the parameter estimation problem.
- Numerical results with synthetic data and real data using experimental mice demonstrates that the unknown parameters can be estimated real-time with high accuracy.
- In an ongoing work, we are formulating an optimal control problem for obtaining and assessment of different optimal combination therapies in colon cancer.

References

1. H. M. Byrne and M. A. J. Chaplain. Mathematical models for tumour angiogenesis: Numerical simulations and nonlinear wave solutions, *Bulletin of Mathematical Biology*, 57:461–486, 1995.
2. V. Capasso and D. Morale. Stochastic modelling of tumour-induced angiogenesis, *Journal of Mathematical Biology*, 58(1-2):219–233, 2009.
3. M. A. J. Chaplain. Mathematical modelling of angiogenesis, *Journal of Neurooncology*, 50:37–51, 2000.
4. D. Csercsik and L. Kovács. Dynamic modeling of the angiogenic switch and its inhibition by Bevacizumab, *Complexity*, 9079104, 2019.
5. M. A. J. Chaplain, S. R. McDougall, and A. R. A. Anderson. Mathematical modeling of tumor-induced angiogenesis, *Annual Review of Biomedical Engineering*, 8:233–257, 2006.
6. J. Folkman. The role of angiogenesis in tumor growth, *Seminars in Cancer Biology*, 3(2):65-71, 1992.
7. J. Folkman. Role of angiogenesis in tumor growth and metastasis, *Seminars in Oncology*, 29(6):15–18, 2002.
8. J. Sápi, L. Kovács, D. A. Drexler, P. Kocsis, D. Gajári and Z. Sápi. Tumor volume estimation and quasi- continuous administration for most effective bevacizumab therapy, *PLoS ONE*, 10(11):1-20, 2015.
9. G. Vilanova, I. Colominas and Hector Gomez. A mathematical model of tumour angiogenesis: growth, regression and regrowth, *Journal of the Royal Society Interface*, 14(126):20160918, 2017.

THANK YOU FOR YOUR ATTENTION