Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 10–20

IMAGE-DRIVEN PARAMETER ESTIMATION IN ABSORPTION-DIFFUSION MODELS OF CHROMOSCOPY

ISABEL N. FIGUEIREDO, PEDRO FIGUEIREDO AND NUNO ALMEIDA

Abstract: The administration of dyes and subsequent examination, with a colorimetry visual criterium, is a gastroenterology procedure for distinguishing, in endoscopic images, normal and aberrant colonic crypts. These are thought to be possible precursors of colon cancer. In this paper a combined image segmentation and parameter estimation model is proposed for *in vivo* colonic crypts' images, obtained with chromoscopic colonoscopy. The parameter estimation is an inverse problem. It is formulated as a partial differential equation constrained optimization problem, and involves an absorption-diffusion equation. A Lagrange multiplier formulation is employed and analyzed for resolving this inverse problem. Using only the segmentation of the medical endoscopic image, which separates normal and aberrant crypts, the mathematical model, proposed in this paper, performs a non invasive mathematical, and dimensionless, quantification of the dye absorption and diffusion coefficients, as well as, the dye absorbed, in normal and aberrant colonic crypts. This mathematical quantification can be important for clinicians, if it is able to provide a distinction between individuals with and without cancer. Numerical simulations, on a test image and on some medical endoscopic images, are presented for the validation and evaluation of the proposed mathematical model.

Keywords: inverse problem, PDE-constrained optimization, Lagrange multiplier, absorption-diffusion equation.

1. Introduction and motivation

Chromoscopic colonoscopy is a current medical technique, used in gastroenterology, to detect *in vivo* aberrant crypt foci (ACF) (see [1, 13, 25]). These are believed to be the precursors of colorectal cancer (see [3, 4, 5, 23]). From the endoscopic point of view, ACF are clusters of crypts (small pits in the colonic epithelium) which stain darker with some dyes (for instance methylene blue) than normal crypts. Colonoscopy allows the inspection of the entire colon and provides the possibility to perform a number of therapeutic operations during a single procedure. A tiny video camera at the tip of the endoscope generates images of the colonic mucosa, which are displayed on a monitor for real-time analysis. Chromoscopy consists in the instillation, by

Received June 11, 2010.
This work was partially supported by the research project UTAustin/MAT/0009/2008 of the UT Austin | Portugal Program (http://www.utaustinportugal.org/).

endoscopy, of a colored substance (a dye) in the colonic mucosa and subsequent visual analysis of the tissue stains in lesions, as ACF (their location and features, as for example, the dimension, number, pattern, shape, etc.). The *in vivo* endoscopic detection of ACF's external boundaries and the crypts' orifices in their interiors is extremely relevant for medical diagnosis and analysis. Up to now the methods applied for assessing the ACF are somewhat subjective and not standardized nor automated and rely on direct medical observation only.

Inspired by the definition of the *chromoscopic colonoscopy technique* we propose a mathematical (dimensionless) model for estimating the dye sink (or absorption) and diffusion parameters in normal and abnormal colonic crypts, as well as, the dye concentration. Remarkably, this is achieved solely with the endoscopic digital image as input datum, and assuming that the dye concentration and the pixel intensity (of the medical image) are related by the Beer-Lambert law (see [24]).

The overall model we suggest is a coupling of two problems. The first is the image segmentation of the endoscopic image and the second a parameter estimation problem. For solving the first problem we apply the technique described in [10], which is a Levenberg-Marquard Newton-type optimization method applied to the Chan and Vese segmentation model [6]. The second problem is an inversion model which is formulated as a PDE-constrained optimization model, *i.e.* an optimization problem, whose constraints are governed by a partial differential equation (PDE). This latter it is essentially a mass diffusion equation (see also [2, 19] where mass diffusion like-equations are used to represent the dye transport within tissues). More precisely, it is a time dependent PDE of parabolic type, whose unknown is the dye concentration. The coefficients of this parabolic equation, stand for the dye sink (or absorption) and diffusion parameters (considered as scalars) in the aberrant and normal colonic crypts. This separation is done throughout the segmentation curve (obtained in the first problem), which detects the ACF. The objective functional in the parameter estimation problem, is a fidelity term that matches the exact spatial dye concentration (expressed as a function of the medical image, by means of the Beer-Lambert law), with the dye concentration at final time, predicted by the parabolic equation.

To the best of our knowledge there are not in the literature papers reporting estimations for ACF absorption and diffusion parameters via *in vivo* endoscopic image segmentation. However, there is a related paper to ours (see [12]) for image-driven parameter estimation with application to brain tumor patient MR images.

The plan of this paper is as follows. Section 2 describes in detail the proposed model. In section 3 the characterization of the solutions of the parameter estimation problem is provided by means of the Lagrange multiplier theory, and an existence result is proved in theorem 3.1. The discretization of the PDE-constrained optimization model is explained in section 4, using finite elements and finite differences. Section 5 describes the algorithm employed. In section 6 we report on the numerical experiments. Finally some conclusions and future work are commented in the last section.

2. Description of the model

In order to describe, with detail, the proposed mathematical model, we assume the following approximations are valid:

- We suppose the dye solute, which is spread inside the lumen colon obeys to a mass diffusion equation equation, which relies on a mass conservation law (see [17, 18]). More exactly, we assume the dye diffuses according to a Fick's law (see [2, 19] for related papers, where Fick's law is used). In addition we also suppose that there is a sink going on, symbolizing the elimination or the absorption of the dye solute by the colon. In effect, it is known that certain dyes are absorbed by the colonic cells. For instance, the dye methylene blue is captured by cells which have absorption capabilities, such as the cells that exist in the colon and small intestine (see [15, 8, 20]).
- The diffusion is constant, with different values in normal and aberrant crypts. The sink term is a linear function of the dye concentration (this is also usual in many chemical reaction equations, see for instance [16]), and there are two sink (or absorption) scalar parameters: one in normal crypts and the other in aberrant crypts.
- Based on medical information, we know that the colorimetry criterium (only visual), used by the doctors, to assess the colonic crypts endoscopic images, reflect the dye concentration in normal and aberrant crypts. Therefore, for evaluating the performance of the model, we presume that the dye concentration, when the image is obtained, obeys to the Beer-Lambert law (see [24]). Briefly, this law states that the dye concentration is a function of the endoscopic image.

Let Ω be a bounded open subset of $\mathbb{R}^2$ (symbolizing a portion of the colon, as a plane object) $I_0 : \overline{\Omega} \longrightarrow \mathbb{R}$ be a given medical image and ω a subset of Ω, with boundary $\partial\omega$. Let $c = (c_1, c_2) \in \mathbb{R}$ be a constant vector symbolizing the two different regions in the image I_0 to be segmented (say ω and its complement $\omega^C = \Omega \setminus \overline{\omega}$) and finally let $\phi : \Omega \longrightarrow \mathbb{R}$, be the so-called level set function, defining implicitly the boundary $\partial\omega$. That is, the 0-level set $\{\phi = 0\}$ is $\partial\omega$, $\{\phi > 0\} = \omega$ and $\{\phi < 0\} = \omega^C$. We represent by $u(x,t)$ the dye concentration at the spatial point $x \in \Omega$ and at time $t \in [0,T]$. Here T is the given terminal time. Moreover we denote by $\sigma = (\sigma_1, \sigma_2)$ and $D = (D_1, D_2)$ two unknown scalars in $\mathbb{R}^2$. The Beer-Lambert law states that I_0 is a function of the real dye concentration at final time, u^*, that is,

$$I_0 = R_g I 10^{-\mu l u^*} \tag{1}$$

where I_0 is the transmitted light (represented by the given image), I is the incident light, R_g is the reflectance of the colon (unknown, see [11] for a related paper), μ is the molar absorptivity coefficient (unknown, it is a measure of the amount of light absorbed, per unit of dye concentration, within the colon), l is the thickness of the material (unknown). Resolving this equation with respect to u^* yields

$$u^* = -\frac{1}{b}\log_{10}\frac{I_0}{a} = \frac{1}{b}\log_{10}\frac{a}{I_0}, \quad \text{with} \quad a := R_g I \quad \text{and} \quad b := \mu l \tag{2}$$

where a and b are unknowns, that we suppose to be unknown scalars in the sequel.

In the combined mathematical model we propose in this paper, the segmentation is performed first, and then, in a second step, the inversion problem is carried out, for determining the sink (or absorption) and diffusion coefficients, in normal and aberrant crypts, as well as the scalars a and b. That is, we have to solve the following two problems:

- Problem 1 - Solve the segmentation problem

$$\min_{\phi,c} S(\phi, c). \tag{3}$$

- Problem 2 - Being ϕ the solution of (3) solve the inversion problem

$$\min_{u,\sigma,D,a,b} J(u,\sigma,D,a,b) := \frac{1}{2}\int_\Omega |\frac{1}{b}\log_{10}\frac{a}{I_0} - u(.,T)|^2$$

subject to

$$\left[\begin{array}{l} \dfrac{\partial u}{\partial t} + g(\sigma,\phi)u - \nabla\cdot\Big(g(D,\phi)\nabla u\Big) = 0 \quad \text{in} \quad \Omega\times(0,T] \\ u(x,0) = u_0(x) \quad \text{in} \quad \Omega\times\{t=0\} \\ \dfrac{\partial u}{\partial n}(x,t) = 0 \quad \text{in} \quad \partial\Omega\times[0,T]. \end{array}\right. \tag{4}$$

In problem 1, the function $S(\phi, c)$ is the objective functional we used in [10]. For convenience of the reader we recall that when S is the original Chan and Vese segmentation functional [6], we have

$$S(\phi,c) := \mu\int_\Omega |\nabla H_\epsilon(\phi)| + \lambda_1\int_\Omega |I_0 - c_1|^2 H_\epsilon(\phi) + \lambda_2\int_\Omega |I_0 - c_2|^2\big(1 - H_\epsilon(\phi)\big) \tag{5}$$

with μ, λ_1, λ_2 given positive scalar parameters, and H_ϵ a regularization of the Heaviside function H, *i.e.* (see [6])

$$H_\epsilon(z) := \frac{1}{\pi}\big(\arctan(\frac{z}{\epsilon}) + \frac{1}{2\epsilon})\big), \qquad H(z) := \begin{cases} 1, & z \geq 0 \\ 0, & z < 0 \end{cases} \tag{6}$$

with $\epsilon > 0$ a small given parameter.

In problem 2, the PDE equation is a second-order parabolic equation describing the time evolution of the dye concentration u, in the region Ω (see for instance [9] and [21], respectively, for theoretical and numerical/discretization issues of parabolic equations). The scalar function g is defined by

$$g(s,\phi) := s_1 H_\epsilon(\phi) + s_2\big(1 - H_\epsilon(\phi)\big), \quad \forall s = (s_1, s_2) \in \mathbb{R}^2. \tag{7}$$

The second order term $\nabla\cdot\Big(g(D,\phi)\nabla u\Big)$ describes diffusion and the term $g(\sigma,\phi)u$ represents the sink term, or depletion. The parameters $\sigma = (\sigma_1, \sigma_2)$ and $D = (D_1, D_2)$, involved in the definition of g, are, respectively, the dye sink (or absorption) and diffusion coefficients, for subscript 1 in aberrant crypts, and for subscript 2 in normal crypts. The separation between normal and abnormal crypts, is done by means of ϕ, through the regularized Heaviside function $H_\epsilon(\phi)$, which is included in the definition of g.

We assume a given dye concentration $u(.,0) := u_0$ at initial time $t = 0$ (the time at which the dye is delivered). In the numerical experiments we take u_0 a positive constant, because according to medical information, a dye concentration is instilled during colonoscopy. Moreover, supposing homogeneous Neumann boundary conditions for the dye concentration, appears to be in good agreement with the physical problem. We also observe that, without loss of generality, the given final time T (which is, from now on, the time at which the endoscopic image is obtained), in the dye concentration equation, can be considered equal to 1. In effect, if we take another final time $T \neq 1$, then a change in the time variable ($t \in [0,1] \to t_1 = tT \in [0,T]$) results in multiplying the absorption and diffusion coefficients, $\sigma = (\sigma_1, \sigma_2)$ and $D = (D_1, D_2)$, by the constant T.

In this problem 2, the objective functional J is the fidelity term, which adjusts the spatial dye concentration u^* (expressed as a function of I_0, by means of the Beer-Lambert law), with the dye concentration at final time, $u(.,T)$, predicted by the model. Therefore, it appears reasonable, for evaluating the performance of the model, to take as fitting term the difference between u^* and $u(.,T)$, measured in the $L^2(\Omega)$ Sobolev space. We remark that here I_0 symbolizes the intensity of the given input image.

Using the PDE equation in (4) to express the dye concentration u as a function of (σ, D), *i.e.*, $u = f(\sigma, D)$, then the PDE-constrained optimization problem (4) reduces to

$$\min_{\sigma,D,a,b} \tilde{J}(\sigma, D, a, b) = J\big(f(\sigma, D,), \sigma, D, a, b\big). \tag{8}$$

This last model is referred to as the reduced formulation of (4) and $\tilde{J}$ as the reduced objective functional.

The procedure we employ in this paper, for the numerical treatment of problem 3 is a Levenberg-Marquard Newton-type optimization method as described in [10]. For solving the parameter estimation problem (4), we use an optimization algorithm applied to the reduced model (8). More precisely, it is a gradient descent method. In this latter approach the computation of the gradient of $\tilde{J}$ relies on the Lagrange multiplier theory. In effect the Lagrangian functional couples the functional J to the time-dependent dye concentration equation through a Lagrange multiplier. It then provides a mechanism for deriving optimality conditions for the PDE-constrained optimization problem (4), and for evaluating the gradient of the reduced objective functional $\tilde{J}$.

3. Characterization of solutions

In this section we start by describing the Lagrangian formulation applied to problem (4), and the corresponding first-order optimality conditions, in a weak variational setting (we remark that this variational framework is also necessary for discretization by finite elements in section 4). We also detail the gradient of $\tilde{J}$. At the end of the section, in theorem 3.1, we prove an existence result for problem (4) and the existence of a Lagrange multiplier, when, in the parabolic equation, the Neumann boundary condition is replaced by a homogeneous Dirichlet condition.

We consider the Sobolev space $H^1(\Omega) = \{u : \ u \in L^2(\Omega), \ u' \in L^2(\Omega)\}$, with u' the time derivative of u in the weak sense, and the dual space $\left[H^1(\Omega)\right]'$ to $H^1(\Omega)$. Moreover, we introduce the following functional spaces involving time

$$L^2\big(0,T;H^1(\Omega)\big) \quad \text{and} \quad L^2\big(0,T;\left[H^1(\Omega)\right]'\big), \tag{9}$$

where

$$L^2\big(0,T;H^1(\Omega)\big) := \Big\{u : [0,T] \to H^1(\Omega) : \ \|u\|_{L^2\left(0,T;H^1(\Omega)\right)} = \Big(\int_0^T \|u(.,t)\|^2_{H^1(\Omega)} dt\Big)^{\frac{1}{2}}\Big\} \tag{10}$$

and $L^2\big(0,T;\left[H^1(\Omega)\right]'\big)$ is defined in a similar way (see [9] for a detailed definition of Sobolev spaces and functional spaces involving time). Then we set

$$\begin{aligned} U &:= \{u \in L^2\big(0,T;H^1(\Omega)\big), \ u' \in L^2\big(0,T;\left[H^1(\Omega)\right]'\big)\} \\ P &:= \{p \in L^2\big(0,T;H^1(\Omega)\big), \ p' \in L^2\big(0,T;\left[H^1(\Omega)\right]'\big)\} \end{aligned} \tag{11}$$

and form the Lagrangian functional L associated to problem (4), by adding to the objective functional J a duality pairing of the dye concentration equation with a Lagrange multiplier p (also known as the adjoint variable). The definition of L is

$$\begin{aligned} L(u,\sigma,D,a,b,p) \ &:= \ \tfrac{1}{2}\int_\Omega \Big|\frac{1}{b}\log_{10}\frac{a}{I_0} - u(.,T)\Big|^2 \\ &+ \ \int_0^T\int_\Omega \Big(-\frac{\partial p}{\partial t}u + g(\sigma,\phi)u\,p + g(D,\phi)\nabla u\cdot\nabla p\Big) \\ &+ \ \int_\Omega u(.,T)\,p(.,T) - u_0\,p(.,0). \end{aligned} \tag{12}$$

for any $u \in U$, $p \in P$ and $(\sigma, D, a, b) \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2$.

Under appropriate conditions (see theorem 3.1 at the send of this section) a minimizing solution (u, σ, D, a, b) of (4) is a stationary point of the Lagrangian functional L. The first-order necessary condition for optimality is that the first variation of the Lagrangian functional vanishes at stationary points. Thus, by taking variations of L with respect to the variables u, σ, D, a, b, p, we obtain the weak form of the first-order optimality conditions:

- Variation of L with respect to p yields the weak dye concentration equation (also called the state equation)

$$\begin{aligned} 0 &= \partial_p L(u,\sigma,D,a,b,p)(\hat{p}) \\ &= \int_0^T \int_\Omega \Big(\frac{\partial u}{\partial t}\,\hat{p} + g(\sigma,\phi)\,u\,\hat{p} + g(D,\phi)\nabla u\cdot\nabla\hat{p}\Big), \qquad \forall \hat{p}\in P, \end{aligned} \tag{13}$$

whose strong from is

$$\frac{\partial u}{\partial t} + g(\sigma,\phi)u - \nabla\cdot\Big(g(D,\phi)\nabla u\Big) = 0 \quad \text{in} \quad \Omega\times(0,T]. \tag{14}$$

This dye concentration equation evolves forward in time, with initial condition $u(x,0) = u_0(x)$ in Ω, and Neumann boundary condition $\frac{\partial u}{\partial n}(x,t) = 0$, in $\partial\Omega\times[0,T]$ (cf. (4)).

- Variation of L with respect to u gives the weak adjoint dye concentration equation

$$\begin{aligned} 0 &= \partial_u L(u,\sigma,D,a,b,p)(\hat{u}) \\ &= -\int_\Omega \big(\frac{1}{b}\log_{10}\frac{a}{I_0} - u(.,T)\big)\cdot\hat{u}(.,T) \\ &+ \alpha\int_\Omega (u(.,T)-d_1)\,\hat{u}(.,T)\,H_\epsilon(\phi) + \beta\int_\Omega (u(.,T)-d_2)\,\hat{u}(.,T)\,\big(1-H_\epsilon(\phi)\big) \\ &+ \int_0^T\int_\Omega \Big(-\frac{\partial p}{\partial t}\,\hat{u} + g(\sigma,\phi)\,p\,\hat{u} + g(D,\phi)\nabla p\cdot\nabla\hat{u}\Big) \\ &+ \int_\Omega p(.,T)\,\hat{u}(.,T), \qquad \forall\hat{u}\in U, \end{aligned} \tag{15}$$

yielding the strong form

$$-\frac{\partial p}{\partial t} + g(\sigma,\phi)p - \nabla\cdot\Big(g(D,\phi)\nabla u\Big) = 0 \quad \text{in} \quad \Omega\times(0,T]. \tag{16}$$

The adjoint dye concentration equation evolves backward in time, with the weak terminal condition

$$\int_\Omega p(.,T)\,\hat{u}(.,T) = \int_\Omega \big(\frac{1}{b}\log_{10}\frac{a}{I_0} - u(.,T)\big)\cdot\hat{u}(.,T), \qquad \forall\hat{u}\in U, \tag{17}$$

which results directly from (15), and Neumann boundary condition $\frac{\partial p}{\partial n}(x,t) = 0$, in $\partial\Omega\times[0,T]$.

- Variation of L with respect to σ_1, σ_2, D_1, D_2 leads to the weak σ_1, σ_2, D_1, D_2 -gradient equations, respectively

$$\begin{aligned}
0 &= \partial_{\sigma_1} L(u,\sigma,D,a,b,p)(\hat{\sigma}_1) &&= \int_0^T \int_\Omega u\, p\, H_\epsilon(\phi)\, \hat{\sigma}_1, && \forall \hat{\sigma}_1 \in \mathbb{R},\\
0 &= \partial_{\sigma_2} L(u,\sigma,D,a,b,p)(\hat{\sigma}_2) &&= \int_0^T \int_\Omega u\, p\, \big(1 - H_\epsilon(\phi)\big)\, \hat{\sigma}_2, && \forall \hat{\sigma}_2 \in \mathbb{R},\\
0 &= \partial_{D_1} L(u,\sigma,D,a,b,p)(\hat{D}_1) &&= \int_0^T \int_\Omega \nabla u \cdot \nabla p\, H_\epsilon(\phi)\, \hat{D}_1, && \forall \hat{D}_1 \in \mathbb{R},\\
0 &= \partial_{D_2} L(u,\sigma,D,a,b,p)(\hat{D}_2) &&= \int_0^T \int_\Omega \nabla u \cdot \nabla p\, \big(1 - H_\epsilon(\phi)\big)\, \hat{D}_2, && \forall \hat{D}_2 \in \mathbb{R}.
\end{aligned} \tag{18}$$

- Variation of L with respect to a, b leads to the weak a, b -gradient equations, respectively

$$\begin{aligned}
0 &= \partial_a L(u,\sigma,D,a,b,p)(\hat{a}) &&= \int_\Omega \Big(\frac{1}{b}\log_{10}\frac{a}{I_0} - u(.,T)\Big)\frac{1}{ab\log 10}\,\hat{a}, && \forall \hat{a} \in \mathbb{R},\\
0 &= \partial_b L(u,\sigma,D,a,b,p)(\hat{b}) &&= \int_\Omega \Big(\frac{1}{b}\log_{10}\frac{a}{I_0} - u(.,T)\Big)\frac{-1}{b^2}\,\log_{10}\frac{a}{I_0}\,\hat{b}, && \forall \hat{b} \in \mathbb{R}.
\end{aligned} \tag{19}$$

Now, recalling the well known results from Lagrangian multiplier theory (see [14]), the gradient of the reduced objective functional $\tilde{J}$ in (8), can be evaluated using a combined procedure involving the computation of the state and adjoint variables, *i.e.* u and p. More precisely, the first-order necessary optimality condition for $\tilde{J}$ is

$$\begin{aligned}
0 \;=\; & \nabla \tilde{J}(\sigma,D,a,b)(\hat{\sigma},\hat{D},\hat{a},\hat{b}) = \nabla_{\sigma,D,a,b} J\big(u,\sigma,D,a,b\big)(\hat{\sigma},\hat{D},\hat{a},\hat{b})\\
& + \int_0^T \int_\Omega \Big(- \frac{\partial p}{\partial t} u + g(\hat{\sigma},\phi) u\, p + g(\hat{D},\phi) \nabla u \cdot \nabla p \Big),\\
& \forall (\hat{\sigma},\hat{D},\hat{a},\hat{b}) \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2,
\end{aligned} \tag{20}$$

with u and p the solutions of the state and adjoint equations (13) and (15), respectively, and clearly, $\hat{\sigma} = (\hat{\sigma}_1, \hat{\sigma}_2)$, $\hat{D} = (\hat{D}_1, \hat{D}_2)$. Thus, from (20), we infer (compare with (18))

$$\begin{array}{rcl} 0 & = & \partial_{\sigma_1}\tilde{J}(\sigma,D,a,b)(\hat{\sigma}_1) = \displaystyle\int_0^T\int_\Omega u\,p\,H_\epsilon(\phi)\,\hat{\sigma}_1, \qquad \forall\hat{\sigma}_1\in\mathbb{R},\\ 0 & = & \partial_{\sigma_2}\tilde{J}(\sigma,D,a,b)(\hat{\sigma}_2) = \displaystyle\int_0^T\int_\Omega u\,p\,\big(1-H_\epsilon(\phi)\big)\,\hat{\sigma}_2, \qquad \forall\hat{\sigma}_2\in\mathbb{R},\\ 0 & = & \partial_{D_1}\tilde{J}(\sigma,D,a,b)(\hat{D}_1) = \displaystyle\int_0^T\int_\Omega \nabla u\cdot\nabla p\,H_\epsilon(\phi)\,\hat{D}_1, \qquad \forall\hat{D}_1\in\mathbb{R},\\ 0 & = & \partial_{D_2}\tilde{J}(\sigma,D,a,b)(\hat{D}_2) = \displaystyle\int_0^T\int_\Omega \nabla u\cdot\nabla p\,\big(1-H_\epsilon(\phi)\big)\,\hat{D}_2, \qquad \forall\hat{D}_2\in\mathbb{R},\\ 0 & = & \partial_a\tilde{J}(\sigma,D,a,b)(\hat{a}) = \displaystyle\int_\Omega\Big(\frac{1}{b}\log_{10}\frac{a}{I_0}-u(.,T)\Big)\,\frac{1}{ab\log 10}\,\hat{a}, \qquad \forall\hat{a}\in\mathbb{R},\\ 0 & = & \partial_b\tilde{J}(\sigma,D,a,b)(\hat{b}) = \displaystyle\int_\Omega\Big(\frac{1}{b}\log_{10}\frac{a}{I_0}-u(.,T)\Big)\,\frac{-1}{b^2}\,\log_{10}\frac{a}{I_0}\,\hat{b}, \qquad \forall\hat{b}\in\mathbb{R}. \end{array} \tag{21}$$

We prove now that a minimum of (4) is indeed a stationary point of the Lagrangian functional L, for the particular case when the Neumann boundary condidition is replaced by a homogeneous Dirichlet condition, in the parabolic equation. For this purpose, let us first denote by $a(t,v,z)$ the bilinear form associated to the parabolic equation, that is, for any $t\in[0,T]$

$$a(t,v,z) := \int_\Omega\Big(g(\sigma,\phi)\,v\,z+g(D,\phi)\nabla v\cdot\nabla z\Big), \quad \forall(v,z)\in\big[H_0^2(\Omega)\big]^2. \tag{22}$$

Theorem 3.1. *Consider problem (4), with the Neumann boundary condition replaced by a homogeneous Dirichlet condition (i.e., $u(x,t)=0$ in $\partial\Omega\times[0,T]$). Assume in addition that the bilinear form $a(.,.,.):[0,T]\times[H_0^1(\Omega)]^2\to\mathbb{R}$ satisfies the coercivity condition, for any $v\in H_0^1(\Omega)$*

$$a(t,v,v)\geq\alpha\|v\|^2_{H_0^1(\Omega)}-\beta\|v\|^2_{L^2(\Omega)} \tag{23}$$

with α and β positive constants independent of $t\in[0,T]$. Moreover, we suppose the initial condition for the parabolic equation verifies $u_0\in L^2(\Omega)$. Then :

- *There exists a unique solution $u\in L^2\big(0,T;H_0^1(\Omega)\big)\bigcap H^1\big(0,T;H^{-1}(\Omega)\big)$ of the parabolic equation defined in (4). Consequently $u\in C^0\big(0,T;L^2(\Omega)\big)$.*
- *The PDE constrained problem (4) has at least a solution $(u^*,\sigma^*,D^*,a^*,b^*)$, and a corresponding Lagrange multiplier p^*, such that $(u^*,\sigma^*,D^*,a^*,b^*,p^*)$ is a stationary point of the Lagrangian functional L.*

Note: The space $C^0\big(0,T;L^2(\Omega)\big)$ comprises all continuous functions $u:[0,T]\to L^2(\Omega)$ and the space $H^1\big(0,T;H^{-1}(\Omega)\big):=\Big\{u:\ u,\ u'\in L^2\big(0,T;H^{-1}(\Omega)\big)\Big\}$, with $H^{-1}(\Omega)$ the dual space to $H_0^1(\Omega)$.

Proof: We give here a sketch of the proof. The arguments rely on existence and regularity results, on energy estimates for the parabolic equation (see for instance [9] or [22]) and on Lagrange multiplier theory for variational problems (see [14]).

The first claim is a direct consequence of existence results for linear evolution equation, more precisely second-order parabolic equations (see [9], [22]). In particular, we note that the coercivity condition (23) is always verified if $g(D,\phi)$ is positive, and this latter happens, if D_1 and D_2 are both positive (we remark that the regularized Heaviside function H_ϵ ranges between 0 and 1).

Regarding the second claim, one can show, using standard weak limit arguments, the existence of a solution $(u^*,\sigma^*,D^*,a^*,b^*)$ to problem (4), where u^* is the solution of the parabolic equation, depending on (σ^*,D^*), that is $u^* := u(\sigma^*,D^*)$. Then, for demonstrating the existence of a corresponding Lagrange multiplier p^* (or in other words, to showing $(u^*,\sigma^*,D^*,a^*,b^*,p^*)$ is a stationary point of the Lagrangian functional L), it is necessary to verify that the map $(\sigma,D,a,b) \to J(u,\sigma,D,a,b)$ is Fréchet differentiable at $(u^*,\sigma^*,D^*,a^*,b^*,p^*)$ and that (σ^*,D^*,a^*,b^*) is regular (see theorem 1.6 in [14]). The Fréchet differentiability of J at (a^*,b^*) does not cause problems and at (u^*,σ^*,D^*) is a consequence of the Fréchet differentiability of the mapping $(\sigma,D) \to u$ at (σ^*,D^*). In effect, let us denote by

$$\begin{aligned} &u^{*\prime}(\sigma_1) := u'(\sigma^*,D^*)(\sigma_1), \quad && u^{*\prime}(\sigma_2) := u'(\sigma^*,D^*)(\sigma_2),\\ &u^{*\prime}(D_1) := u'(\sigma^*,D^*)(D_1), && u^{*\prime}(D_2) := u'(\sigma^*,D^*)(D_2), \end{aligned} \tag{24}$$

the partial derivatives of u at (σ^*,D^*) in the direction of σ_1, σ_2, D_1 and D_2, respectively. Then, they satisfy

$$\left[\begin{array}{l} \dfrac{\partial u^{*\prime}(\sigma_1)}{\partial t} + g(\sigma^*,\phi)u^{*\prime}(\sigma_1) - \nabla\cdot\Big(g(D^*,\phi)\nabla u^{*\prime}(\sigma_1)\Big) = -\sigma_1 H_\epsilon(\phi)u^* \quad \text{in} \quad \Omega\times(0,T] \\ u^{*\prime}(\sigma_1)(x,0) = 0 \quad \text{in} \quad \Omega\times\{t=0\} \\ u^{*\prime}(\sigma_1)(x,t) = 0 \quad \text{in} \quad \partial\Omega\times[0,T]. \end{array}\right. \tag{25}$$

$$\left[\begin{array}{l} \dfrac{\partial u^{*\prime}(\sigma_2)}{\partial t} + g(\sigma^*,\phi)u^{*\prime}(\sigma_2) - \nabla\cdot\Big(g(D^*,\phi)\nabla u^{*\prime}(\sigma_2)\Big) = -\sigma_2\big(1-H_\epsilon(\phi)\big)u^* \quad \text{in} \quad \Omega\times(0,T] \\ u^{*\prime}(\sigma_2)(x,0) = 0 \quad \text{in} \quad \Omega\times\{t=0\} \\ u^{*\prime}(\sigma_2)(x,t) = 0 \quad \text{in} \quad \partial\Omega\times[0,T]. \end{array}\right. \tag{26}$$

$$\left[\begin{array}{l} \dfrac{\partial u^{*\prime}(D_1)}{\partial t} + g(\sigma^*, \phi)u^{*\prime}(D_1) - \nabla \cdot \Big(g(D^*, \phi)\nabla u^{*\prime}(D_1)\Big) = \\ \qquad\qquad\qquad\qquad\qquad\qquad \nabla \cdot \Big(D_1 H_\epsilon(\phi)\nabla u^*\Big) \quad \text{in} \quad \Omega \times (0, T] \\ u^{*\prime}(D_1)(x, 0) = 0 \quad \text{in} \quad \Omega \times \{t = 0\} \\ u^{*\prime}(D_1)(x, t) = 0 \quad \text{in} \quad \partial\Omega \times [0, T]. \end{array}\right. \tag{27}$$

$$\left[\begin{array}{l} \dfrac{\partial u^{*\prime}(D_2)}{\partial t} + g(\sigma^*, \phi)u^{*\prime}(D_2) - \nabla \cdot \Big(g(D^*, \phi)\nabla u^{*\prime}(D_2)\Big) = \\ \qquad\qquad\qquad\qquad \nabla \cdot \Big(D_2(1 - H_\epsilon(\phi))\nabla u^*\Big) \quad \text{in} \quad \Omega \times (0, T] \\ u^{*\prime}(D_2)(x, 0) = 0 \quad \text{in} \quad \Omega \times \{t = 0\} \\ u^{*\prime}(D_2)(x, t) = 0 \quad \text{in} \quad \partial\Omega \times [0, T]. \end{array}\right. \tag{28}$$

Because of the existence and uniqueness results for second-order parabolic equations, these four parabolic equations thus guarantee the Fréchet differentiability of J at $(u^*, \sigma^*, D^*, a^*, b^*)$.

Finally, the point $(\sigma^*, D^*, a^*, b^*)$ is regular if, for this point, the four mappings associated to the parabolic equations (25), (26), (27), (28) are surjective onto the space $L^2\big(0, T; H^{-1}(\Omega)\big)$ (see Definition 1.5 in [14]). This is true, and is again a consequence of existence and uniqueness results for second-order parabolic equations (see for instance, theorem 11.3 and lemma 11.4 in [22]). ∎

We observe that the results of this theorem are still true if we replace the Dirichlet boundary condition by a mixed Dirichlet-Neumann boundary condition. In effect, Poincaré inequality plays a key role for proving the existence of solution to the parabolic equation, and it is still valid for mixed Dirichlet-Neumann boundary condition. However, for pure Neumann boundary conditions, as defined in (4), the Poincaré inequality does not hold, so it might not be possible to claim the results of this theorem apply directly to the inversion problem (4). Nevertheless, we decided to formulate (4) with only Neumann conditions, because we think it is the most appropriate framework for the *medical chromoscopy technique* we are modelling through equation (4). Also in the numerical results (see section 6) we always took homogeneous Neumann conditions.

We finally state another property, regarding the energy decay of the dye concentration.

Proposition 3.1. *If $g(\sigma, \phi)$ and $g(D, \phi)$ are positive, then the energy $E(t) = \int_\Omega u(.,t)^2$ of the parabolic equation, in (4), decays in time.*

Proof: Multiplying the parabolic equation in (4) by u and integrating in Ω we have

$$\int_\Omega \Big(\frac{du}{dt}u + g(\sigma,\phi)\,u^2 + g(D,\phi)|\nabla u|^2\Big) = 0, \tag{29}$$

which implies

$$\frac{dE}{dt}(t) \le 0, \quad \forall t \in (0,T]. \tag{30}$$

Also integrating from 0 to t, we get

$$E(t) \le E(0) \Longleftrightarrow \int_\Omega u(.,t)^2 \le \int_\Omega u_0^2 \quad \forall t \in (0,T], \tag{31}$$

which concludes the statement. ■

4. Discretization

In order to actually compute the finite-dimensional approximation to the solution of (4), or equivalently (8), we need to define the discretization of the state (13) and adjoint (15) equations, as well as, the reduced objective functional $\tilde{J}$ and the first-order optimality conditions (21). To do so we employ finite elements, for discretizing the space variable x, and finite differences for the time variable t. This yields a nonlinear system of algebraic equations for the unknown (spatio-temporal) dye concentration u, (spatio-temporal) adjoint dye concentration p, and unknown scalar parameters σ_1, σ_2, D_1, D_2.

The first thing that needs to be explained is the finite-element matrix formulation corresponding to the several spatial integrals appearing in (21) and also in the weak formulations of the forward and adjoint dye concentration equations. The next formula (32) lists these correspondences, for the fixed ϕ (defining the segmentation) obtained in problem 1 :

$$\begin{array}{llllll} z'Mv & \approx & \displaystyle\int_\Omega z\,v & z'Kv & \approx & \displaystyle\int_\Omega \nabla z\cdot\nabla v \\ z'M(\phi)v & \approx & \displaystyle\int_\Omega H_\epsilon(\phi)\,z\,v & z'K(\phi)v & \approx & \displaystyle\int_\Omega H_\epsilon(\phi)\,\nabla z\cdot\nabla v \\ z'M(\sigma,\phi)v & \approx & \displaystyle\int_\Omega g(\sigma,\phi)\,z\,v & z'K(D,\phi)v & \approx & \displaystyle\int_\Omega g(D,\phi)\,\nabla z\cdot\nabla v \end{array} \tag{32}$$

where M, K , $M(\phi)$, $K(\phi)$, $M(\sigma,\phi)$, $K(\sigma,\phi)$ are finite element matrices, z and v are arbitrary functions in $H^1(\Omega)$, and we also represent by the same letters z and v the corresponding finite element vectors (the prime in z' means the transpose of vector z). Thus, by applying the finite element method, and using these matrix correspondences we arrive at the following semidiscrete

Galerkin formulation (discrete in space and continuous on time), for the forward-in-time dye concentration equation (compare with (13))

$$\left[\begin{array}{l} M\dfrac{\partial u}{\partial t}(t) + M(\sigma,\phi)u(t) + K(D,\phi)u(t) = 0 \\ u(0) = u_0 \end{array}\right. \tag{33}$$

and for the backward-in-time adjoint dye concentration equation (compare with (15))

$$\left[\begin{array}{l} -M\dfrac{\partial p}{\partial t}(t) + M(\sigma,\phi)p(t) + K(D,\phi)p(t) = 0 \\ p(T) = \frac{1}{b}\log_{10}\frac{a}{I_0} - u(T). \end{array}\right. \tag{34}$$

Here, we denote by $u(t)$ and $p(t)$ the finite element approximations of $u(.,t)$ and $p(.,t)$, respectively. This means that now $u(t)$ and $p(t)$ are vectors of unknowns at the finite element nodes. We represent by $\frac{\partial u}{\partial t}$, $\frac{\partial p}{\partial t}$ the derivatives of $u(t)$, $p(t)$ with respect to time. The definition of $p(T)$ comes directly from (17), and $\log_{10}\frac{a}{I_0} = \log_{10} a - \log_{10} I_0$, where $\log_{10} I_0 := \big(\log_{10} i_0\big)$ with i_0 the generic element of the given image I_0 (the dimensions of vectors I_0 and $u(T)$ must be equal, which can require an interpolation of the given image I_0).

Let us now proceed and subdivide the time interval $[0,T]$ into $N-1$ subintervals

$$[0,T] = \bigcup_{i=1}^{N-1}[t_i, t_{i+1}], \qquad 0 = t_1 < T_2 < \ldots < t_i < \ldots < t_N = T$$

We assume, for simplicity, the time step size $\triangle t = t_i - t_{i-1} = \frac{T}{N-1}$ is constant over the time interval. In addition, we approximate the time derivatives $\frac{\partial u}{\partial t}(t_{i+1})$ and $\frac{\partial p}{\partial t}(t_{i+1})$ by the forward time difference scheme

$$\begin{array}{l} \dfrac{\partial u}{\partial t}(t_{i+1}) \approx \dfrac{u(t_{i+1}) - u(t_i)}{\triangle t} = \dfrac{u_{i+1} - u_i}{\triangle t} \\ \dfrac{\partial p}{\partial t}(t_{i+1}) \approx \dfrac{p(t_{i+1}) - p(t_i)}{\triangle t} = \dfrac{p_{i+1} - p_i}{\triangle t} \end{array} \tag{35}$$

where the notations are self-explanatory. Consequently, we fully discretize the state and adjoint equations (33) and (34), by the following system of equations

$$\left[\begin{array}{l} M\dfrac{u_{i+1} - u_i}{\triangle t} + M(\sigma,\phi)u_{i+1} + K(D,\phi)u_{i+1} = 0, \quad \forall i = 1,\ldots,N-1 \\ u_1 = u_0 \end{array}\right. \tag{36}$$

and

$$\left[\begin{array}{l} -M\dfrac{p_{i+1} - p_i}{\triangle t} + M(\sigma,\phi)p_i + K(D,\phi)p_i = 0, \quad \forall i = N-1,\ldots,1 \\ p_N = \frac{1}{b}\log_{10}\frac{a}{I_0} - u_N. \end{array}\right. \tag{37}$$

These two schemes correspond to implicit (or backward) Euler time discretization methods. The systems (36) and (37) can equivalently be rewritten as

$$\left[\begin{array}{l} \Big(\frac{M}{\triangle t} + M(\sigma,\phi) + K(D,\phi)\Big)u_{i+1} = \frac{M}{\triangle t}u_i, \quad \forall i = 1,\dots,N-1 \\ u_1 = u_0 \end{array}\right. \tag{38}$$

and

$$\left[\begin{array}{l} \Big(\frac{M}{\triangle t} + M(\sigma,\phi) + K(D,\phi)\Big)p_i = \frac{M}{\triangle t}p_{i+1}, \quad \forall i = N-1,\dots,1 \\ p_N = \frac{1}{b}\log_{10}\frac{a}{I_0} - u_N. \end{array}\right. \tag{39}$$

Now the discrete reduced objective functional is (compare with J in (4))

$$\tilde{J}(\sigma, D, a, b) = \frac{1}{2}(\frac{1}{b}\log_{10}\frac{a}{I_0} - u_N)' M (\frac{1}{b}\log_{10}\frac{a}{I_0} - u_N), \tag{40}$$

where the upper-script prime means transpose vector. Then the discretization of the first-order optimality conditions, corresponding to (21), with respect to the optimization variables σ_1, σ_2, D_1, D_2 is now

$$\begin{array}{lllll} 0 &=& \partial_{\sigma_1}\tilde{J}(\sigma,D,a,b) &=& \frac{\triangle t}{2}\Big(\sum_{i=1}^{N-1} u_i^T M(\phi)p_i + \sum_{i=2}^{N} u_i^T M(\phi)p_i\Big) \\ 0 &=& \partial_{\sigma_2}\tilde{J}(\sigma,D,a,b) &=& \frac{\triangle t}{2}\Big(\sum_{i=1}^{N-1} u_i^T \big(M - M(\phi)\big)p_i + \sum_{i=2}^{N} u_i^T \big(M - M(\phi)\big)p_i\Big) \\ 0 &=& \partial_{D_1}\tilde{J}(\sigma,D,a,b) &=& \frac{\triangle t}{2}\Big(\sum_{i=1}^{N-1} u_i^T K(\phi)p_i + \sum_{i=2}^{N} u_i^T K(\phi)p_i\Big) \\ 0 &=& \partial_{D_2}\tilde{J}(\sigma,D,a,b) &=& \frac{\triangle t}{2}\Big(\sum_{i=1}^{N-1} u_i^T \big(K - K(\phi)\big)p_i + \sum_{i=2}^{N} u_i^T \big(K - K(\phi)\big)p_i\Big), \\ 0 &=& \partial_a\tilde{J}(\sigma,D,a,b) &=& (\frac{1}{b}\log_{10}\frac{a}{I_0} - u_N)' M \frac{1}{ab\log 10} v \\ 0 &=& \partial_b\tilde{J}(\sigma,D,a,b) &=& -(\frac{1}{b}\log_{10}\frac{a}{I_0} - u_N)' M \frac{1}{b^2}\log_{10}\frac{a}{I_0}, \end{array} \tag{41}$$

where v represents a vector with all components equal to 1.

We finish this section with a property of numerically stability for the schemes (36)-(37).

Proposition 4.1. *If $g(\sigma,\phi)$ and $g(D,\phi)$ are positive, then the numerical schemes* (36) *and* (37) *are unconditionally stable.*

Proof: The reasoning is similar to both problems, thus we only give the arguments for one of them, (36). Multiplying the equation of (36) by u_{i+1} we get, for any i

$$(u_{i+1} - u_i, u_{i+1})_{L^2(\Omega)} + \triangle t \int_\Omega g(\sigma,\phi)\, u_{i+1}^2 + \triangle t \int_\Omega g(D,\phi)\, |\nabla u_{i+1}|^2 = 0. \tag{42}$$

But for any v and z in $L^2(\Omega)$ we have the inequality

$$\int_\Omega (v-z)\,v \geq \frac{1}{2}\big(\|v\|^2_{L^2(\Omega)} - \|z\|^2_{L^2(\Omega)}\big), \tag{43}$$

due to Cauchy-Schwartz inequality and Young inequality. Therefore, the first term in the left hand side of (42) can be re-written according to this last inequality. So for any i

$$\|u_{i+1}\|^2_{L^2(\Omega)} + \triangle t \int_\Omega \Big(g(\sigma,\phi)\,u^2_{i+1} + g(D,\phi)\,|\nabla u_{i+1}|^2\Big) \leq \|u_i\|^2_{L^2(\Omega)}. \tag{44}$$

Thus summing from $i=1$ to $k-1$, for any $k \geq 2$, we get

$$\|u_k\|^2_{L^2(\Omega)} + \triangle t \sum_{i=1}^{k-1} \int_\Omega \Big(g(\sigma,\phi)\,u^2_{i+1} + g(D,\phi)\,|\nabla u_{i+1}|^2\Big) \leq \|u_0\|^2_{L^2(\Omega)}. \tag{45}$$

■

5. Algorithm

The optimization method we apply to problem (4) is a gradient descent method with the following steps:

Step 1 Set $k=0$ and initialize with $\sigma^k = (\sigma_1^k, \sigma_2^k)$, $D^k = (D_1^k, D_2^k)$, a^k, b^k.

Step 2 Solve the state and adjoint equations (38)-(39) and get, respectively, the dye concentration and adjoint dye concentration vectors, u^{k+1} and p^{k+1}.

Step 3 Update the control variables $\sigma^{k+1} = (\sigma_1^{k+1}, \sigma_2^{k+1})$, $D^{k+1} = (D_1^{k+1}, D_2^{k+1})$, a^{k+1}, b^{k+1} by using a gradient descent algorithm, *i.e.* (compare with (41))

$$\begin{aligned}
\sigma_1^{k+1} &= \sigma_1^k - \alpha_{\sigma_1}\frac{\triangle t}{2}\Big(\sum_{i=1}^{N-1} u_i^{k+1} M(\phi) p_i^{k+1} + \sum_{i=2}^{N} u_i^{k+1} M(\phi) p_i^{k+1}\Big)\\
\sigma_2^{k+1} &= \sigma_2^k - \alpha_{\sigma_2}\frac{\triangle t}{2}\Big(\sum_{i=1}^{N-1} u_i^{k+1}\big(M - M(\phi)\big) p_i^{k+1} + \sum_{i=2}^{N} u_i^{k+1}\big(M - M(\phi)\big) p_i^{k+1}\Big)\\
D_1^{k+1} &= D_1^k - \alpha_{D_1}\frac{\triangle t}{2}\Big(\sum_{i=1}^{N-1} u_i^{k+1} K(\phi) p_i^{k+1} + \sum_{i=2}^{N} u_i^{k+1} K(\phi) p_i^{k+1}\Big)\\
D_2^{k+1} &= D_2^k - \alpha_{D_2}\frac{\triangle t}{2}\Big(\sum_{i=1}^{N-1} u_i^{k+1}\big(K - K(\phi)\big) p_i^{k+1} + \sum_{i=2}^{N} u_i^{k+1}\big(K - K(\phi)\big) p_i^{k+1}\Big)\\
a^{k+1} &= a^k - \alpha_a\Big(\big(\frac{1}{b^k}\log_{10}\frac{a^k}{I_0} - u_N\big)' M \frac{1}{a^k b^k \log 10}\Big)\\
b^{k+1} &= b^k - \alpha_b\Big(-\big(\frac{1}{b^k}\log_{10}\frac{a^k}{I_0} - u_N\big)' M \frac{1}{{b^k}^2}\log_{10}\frac{a^k}{I_0}\Big)
\end{aligned} \tag{46}$$

where the positive scalars α_{σ_1}, α_{σ_2}, α_{D_1}, α_{D_2}, α_a, α_b are step lengths, not necessarily equal.

Step 4 Repeat steps 1-3, with $k = 0$ replaced by $k + 1$.

Step 5 Stop the algorithm with the following stopping criterium

$$\begin{aligned} &\frac{optc(\sigma_1^k)}{optc(\sigma_1^0)} < tol, \quad \frac{optc(\sigma_2^k)}{optc(\sigma_2^0)} < tol, \quad \frac{optc(D_1^k)}{optc(D_1^0)} < tol, \quad \frac{optc(D_2^k)}{optc(D_2^0)} < tol, \\ &\frac{optc(a^k)}{optc(a^0)} < tol, \quad \frac{optc(b^k)}{optc(b^0)} < tol, \end{aligned} \tag{47}$$

where *optc* stands for optimality condition, the upper-scripts ^{k} and ^{0} represent the current k and 0 iterate, respectively, and *tol* is a prescribed termination tolerance on the optimization variables (for instance $tol = 5 \times 10^{-3}$).

6. Numerical results

We describe now the results obtained with the combined model (4). We perform first a test-case and afterwards numerical simulations on three medical ACF endoscopic images.

For all the experiments, we take dimensionless values, and the following have been used, unless otherwise mentioned. The time interval is $[0, 1]$, thus $T = 1$, and the final time step size $\triangle t = 1/10 = 0.1$. The spatial domain $\Omega = [0, 1]^2$ is discretized with squared finite element meshes, with N^2 finite elements (where $N = 50, 100, 150, 200, 300$ or 400) and bilinear shape functions. The initial dye concentration u_0 is 0.007, for the test problem, and 0.002, for the medical images. The stopping criterium for the algorithm was the termination tolerance on the optimization variables set to 5×10^{-3} (see(47)). For the parameter ϵ, in the regularized Heaviside function (6), we took $\epsilon = 0.002$. The implementation is done in MATLAB® [26]; we use Comsol Multiphysics® [7] for extracting the finite element matrices (see (32)).

6.1. Test problem. The intention is to investigate if and how close the original sink (or absorption) and diffusion coefficients can be retrieved with the combined model (4). Therefore, it appears reasonable, for evaluating the performance of the model, to take for u^* in (2), the generated dye concentration, via the forward problem, for a given choice of 4 parameters, and with a predefined segmentation curve (the choice is a circumference, see Figure 1). Consequently, the objective functional we consider, for the test problem, is

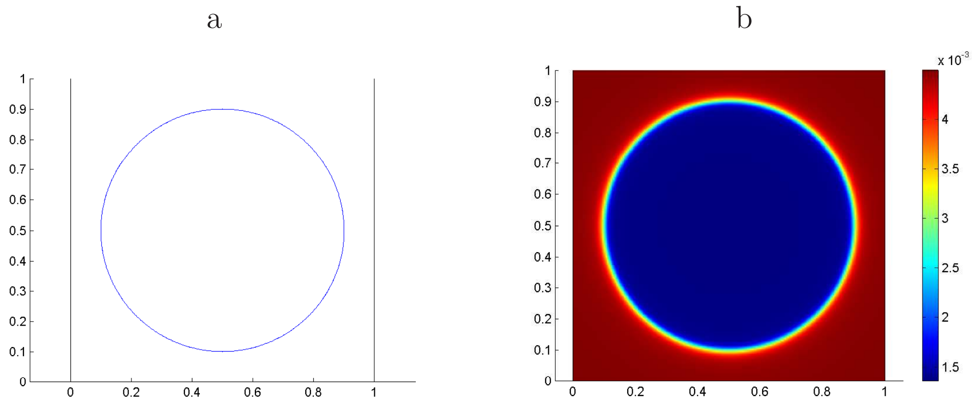

FIGURE 1. Test problem with 200^2 finite elements. a- The circumference, symbolizing the segmentation curve. b- The generated dye concentration, at final time, with $(\bar{\sigma_1}, \bar{\sigma_2}, \bar{D_1}, \bar{D_2}) = (2.0, 0.5, 0.00012, 0.00016)$

of the form

$$J(u, \sigma, D) := \frac{1}{2}\int_{\Omega} |u_{(\bar{\sigma},\bar{D})}(., T) - u(., T)|^2 \tag{48}$$

where $u_{(\bar{\sigma},\bar{D})}(., T)$ is the solution of the forward problem, at final time T, and the given parameters are $(\bar{\sigma_1}, \bar{\sigma_2}, \bar{D_1}, \bar{D_2}) = (2.0, 0.5, 0.00012, 0.00016)$.

The results are displayed in Table 1, for 4 different finite element meshes, and two different initial guesses $(\sigma_1^0, \sigma_2^0, D_1^0, D_2^0)$: $(2.5, 0.9, 0.00014, 0.00012)$ is guess1, and $(1.5, 0.3, 0.0001, 0.00014)$ is guess 2. We remark that we used two different step lengths in the gradient descent algorithm : $\alpha_{D_1} = \alpha_{D_2} = 10^4$ for the diffusion parameters and $\alpha_{\sigma_1} = \alpha_{\sigma_2} = 10^8$ for the absorption parameters (see (46)).

Table 1 - Optimization results for the test problem

Guess	N	σ_1	σ_2	D_1	D_2	Iterations
guess 1	50	2.00046	0.500013	0.000120433	0.0001607	39
	100	2.00153	0.49986	0.000119443	0.000159388	76
	150	2.00145	0.500035	0.000121371	0.00016247	42
	200	2.00127	0.499878	0.000119449	0.00015929	45
guess 2	50	1.99736	0.500053	0.000118844	0.000157691	34
	100	1.99669	0.500405	0.000121677	0.000161489	35
	150	2.0023	0.499737	0.00011856	0.00015805	28
	200	1.99708	0.499994	0.000117927	0.000156082	37
Exact solution		2 .0	0.5	0.00012	0.00016	

The stopping criterium for the algorithm was the termination tolerance on the optimization variables set to 10^{-3} (see (47)).

6.2. Medical endoscopic images. For performing the segmentation of the three ACF endoscopic images, *i.e.* for resolving problem 1 in (3), we used the technique described in [10], for an RGB input image. However, we observe that many other types of vector-valued images, for segmenting the ACF, could have been used, with successful segmentation results (see [10] for a list and a detailed explanation). For the inversion problem we consider the grayscale image corresponding to I_0, and we use it in equations (1)-(2).

The Tables 2, 3 and 4 display the results obtained for the endoscopic images, named M1, M3 and M2. The Figures 2, 3 and 4 show the the segmentation, dye concentration and adjoint state at final time T, for M1, M2 and M3, respectively. The pictures correspond to the values indicated in Tables 2, 3 and 4.

Table 2 - Optimization results for medical image M1

tol	N	σ_1	σ_2	D_1	D_2	a	b	Iterations
guess0		3	4	2e-5	4e-5	1500	30000	
tol1	100	5.36143	5.51566	3.73587e-5	5.18381e-5	1983.77	29927.9	71
	150	5.35848	5.51564	3.98597e-5	5.24625e-5	1983.75	29927.8	71
	200	5.35519	5.51654	4.02067e-5	5.34485e-5	1983.83	29927.8	71
tol2	100	5.3807	5.52688	3.74771e-5	5.19782e-5	1989.9	29927	113
	150	5.37916	5.52581	3.99988e-5	5.26239e-5	1989.92	29927	113
	200	5.37718	5.52596	4.03656e-5	5.36265e-5	1990.09	29926.9	114
guess1		2.5	3.5	3.5e-5	3e-5	1500	30000	
tol1	150	5.36514	5.50455	6.63366e-5	5.71782e-5	1959.02	29933.4	59
tol2	150	5.38577	5.5307	6.63696e-5	5.72826e-5	1968.68	29931.9	101
guess2		4	3.5	1.5e-5	4.5e-5	1500	30000	
tol1	150	5.36962	5.51957	1.8845e-5	4.76797e-5	1972.06	29930.4	75
guess3		3	4	2e-5	4e-5	1400	29000	
tol1	150	5.32389	5.48248	3.98238e-5	5.24378e-5	1910.88	28926	68
tol2	150	5.34352	5.4919	3.99612e-5	5.25972e-5	1917	28925.9	107

The medical images M1 and M3 have 200×200 pixels and M3 560×560 has pixels. We used finite element discretizations with 100^2, 150^2, 200^2, 300^2 and 400^2 elements. Moreover, for the medical image M1, we indicate in Table 2, the values for the termination tolerance set to two values : $tol1 = 5 \times 10^{-3}$ and $tol2 = 5 \times 10^{-4}$. For M1 and M3 only the tolerance $tol1$ was used. We remark that for these medical images, the initial dye concentration was set $u(.,0) = 0.002$ and we used three different step lengths in the gradient descent algorithm : $\alpha_{D_1} = \alpha_{D_2} = 10$ for the diffusion parameters, $\alpha_{\sigma_1} = \alpha_{\sigma_2} = 10^8$ for the sink (or absorption) parameters and $\alpha_a = \alpha_b = 10^{14}$ (see (46)).

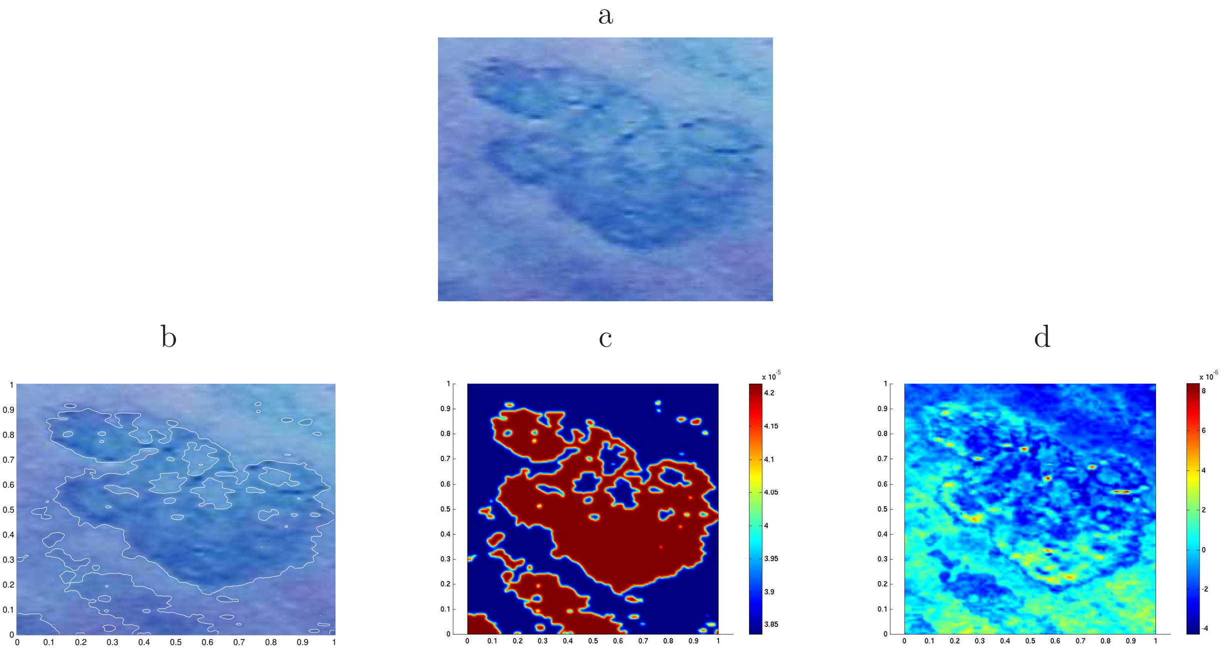

FIGURE 2. Graphics corresponding to the results displayed in Table 2 for $N = 200$, guess0 and tol1. a: Medical image M1 with 200^2 pixels. b: Segmentation. c and d: The optimal dye concentration and optimal adjoint dye concentration at final time.

Table 3 - Optimization results for medical image M2

Image	N	σ_1	σ_2	D_1	D_2	a	b	Iterations
guess0		3	4	2e-5	4e-5	1500	30000	
M2	200	5.38206	5.65368	4.4406e-5	5.33728e-5	1994	29933.3	106
tol1	300	5.38246	5.65369	4.5894e-5	5.2829e-5	1993.74	29933.4	106
	400	5.3826	5.65493	4.67463e-5	5.25886e-5	1993.8	29933.4	107
guess1		2.5	3.5	3.5e-5	3e-5	1500	30000	
tol1	200	5.45484	5.64016	9.03615e-5	8.37244e-5	1839.84	29962.5	86
guess2		4	3.5	1.5e-5	4.5e-5	1500	30000	
tol1	200	5.37042	5.69566	2.26725e-5	4.9905e-5	2014.66	29933.1	199
guess3		3	4	2e-5	4e-5	1400	29000	
tol1	200	5.34598	5.6242	4.41564e-5	5.32138e-5	1921.95	28932.5	103

We emphasize that the inverse problem is a mathematical model, with dimensionless values for the domain Ω, the final time T, the six inversion parameters σ_1, σ_2, D_1, D_2, a, b, as well as, the initial dye concentration $u(.,0)$. We briefly justify now the initial values chosen for σ_1, σ_2, D_1, D_2, a

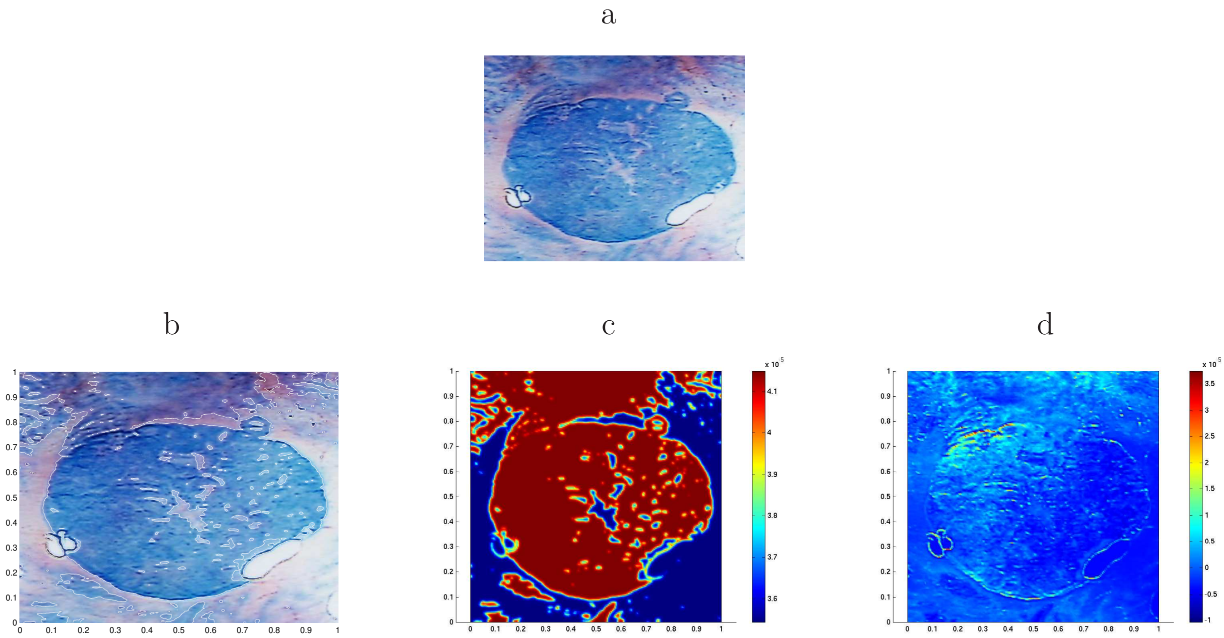

FIGURE 3. Graphics corresponding to the results displayed in Table 3 for $N = 200$, guess0 and tol1. a: Medical image M2 with 560^2 pixels. b: Segmentation. c and d: The optimal dye concentration and optimal adjoint dye concentration at final time.

and b, in order to initialize the algorithm. For the sink (or absorption) and diffusion parameters, we decided to initialize, with values, which are of the same order as those used in the test problem. Regarding the initial choice for a and b, it relies in the following observations. As we know that the energy of the parabolic equation decreases in time (see proposition 3.1), and we choose $u(.,0) = 0.002$, than at final time the dye concentration would be very small, say $u^* \approx 10^{-5}$. On the other hand, due to the Beer-Lambert law (1), we have $b = \frac{1}{u^*} \log_{10} \frac{a}{I_0}$ and $a = I_0 10^{bu^*}$. Thus if $\log_{10} \frac{a}{I_0}$ ranges between 1 and 9 (since the reflectance unit is percent or a factor between 0 and 1, and we assume $I > I_0$), then $b \approx \frac{1}{u^*}$ and consequently $b \approx 10^5$, and $a \approx 10 I_0$.

Table 4 - Optimization results for medical image M3

Image	N	σ_1	σ_2	D_1	D_2	a	b	Iterations
guess0		3	4	2e-5	4e-5	1500	30000	
M3 tol1	100	4.67692	4.80533	4.20916e-5	5.43009e-5	1734.64	29952.4	49
	150	4.67702	4.80352	4.49877e-5	5.48639e-5	1734.44	29952.3	48
	200	4.67582	4.80274	4.48888e-5	5.55619e-5	1735.06	29952.1	47

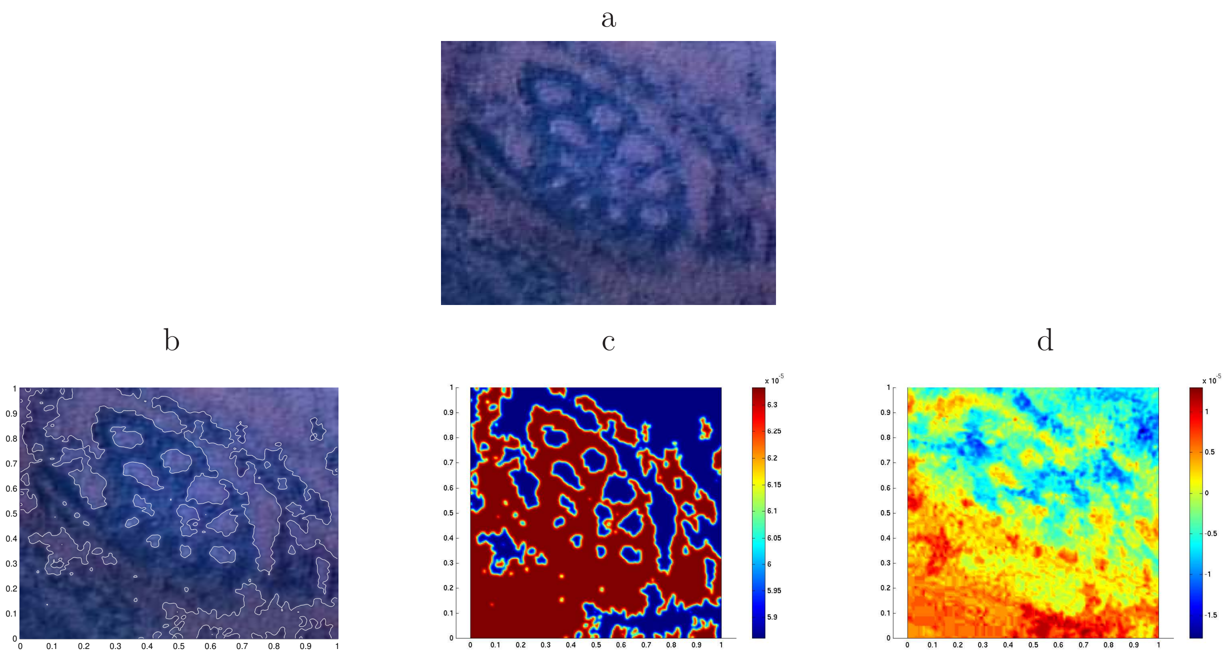

FIGURE 4. Graphics corresponding to the results displayed in Table 4, for $N = 200$, guess0 and tol1. a: Medical image M3 with 200^2 pixels. b: Segmentation. c and d: The optimal dye concentration and optimal adjoint dye concentration at final time.

7. Conclusions and future work

In this paper a novel coupled segmentation and parameter estimation model has been proposed for assessing and extracting physical information about colonic crypts *in vivo* endoscopic images. The overall mathematical model permits to obtain dimensionless parameter values, namely the dye sink (or absorption) and diffusion parameters, in normal and aberrant colonic crypts. Mathematical analysis of the inversion problem in a Lagrangian framework has been carried out.

The results obtained with a test image reveal an excellent performance of the model. Moreover, the results for three medical endoscopic ACF images appear to be very promising.

In order to completely validate this model we intend, in the future, to do tests in a data-base of medical images, taken either from one individual, or from different groups of individuals. The aim is to establish a correlation between the values of the dye sink (or absorption) and diffusion parameters in two different groups of individuals: with and without colon cancer.

We believe the model presented in this paper is a valuable and innovative contribution towards the understanding of ACF, and it will lead to a better clarification regarding the possible cancer potential of ACF. From the medical point of view, the methodology involves only *in vivo* techniques, which is an extremely important feature of this procedure.

References

[1] D. G. Adler, C. J. Gostout, D. Sorbi, et al. Endoscopic identification and quantification of the aberrant crypt in the human colon. *Gastrointestinal Endoscopy*, 56:657–662, 2002.

[2] A. N. Bashkatov, E. A. Genina, V. V. Tuchin, E. E. Chikinab, A. B. Knyazevb, and O. V. Mareev. Methylene blue laser therapy for the treatment of chronic maxillary sinusitis. In *Therapeutic Laser Applications and Laser-Tissue Interactions II, Proc. of SPIE-OSA Biomedical Optics*, volume 5863, pages 58630Z–1, 2005.

[3] R. P. Bird. Observation and quantification of aberrant crypts in the murine mucosa treated with a colon carcinogen: preliminary findings. *Cancer Letters*, 37:147–151, 1987.

[4] R. P. Bird. Role of aberrant crypt foci in understanding the pathogenesis of colon cancer. *Cancer Letters*, 93:55–71, 1995.

[5] R. P. Bird and C. K. Good. The significance of aberrant crypt foci in understanding the pathogenesis of colon cancer. *Toxicology Letters*, 112-113:395–402, 2000.

[6] T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Trans. Image Processing*, 10(2):266–277, 2001.

[7] Comsol Multiphysics®. *http://www.comsol.com/*.

[8] M. Dinis-Ribeiro and L. Moreira-Dias. There is no clinical evidence of consequences after methylene blue chromoendoscopy. *Gastrointestinal Endoscopy, Letters to the Editor*, 67(7):1209, 2008.

[9] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[10] I. N. Figueiredo, P. N. Figueiredo, G. Stadler, O. Ghattas, and A. Araújo. Variational image segmentation for endoscopic human colonic aberrant crypt foci. *IEEE Transactions on Medical Imaging*, 2010.

[11] D. Hidovic-Rowe and E. Claridge. Modelling and validation of spectral reflectance for the colon. *Phys. Med. Biol.*, 50:1071–1093, 2005.

[12] C. Hogea, C. Davatzikos, and G. Biros. An image-driven parameter estimation problem for a reaction-diffusion glioma growth model with mass effects. *J. Math. Biol.*, 56(6):793–825, 2008.

[13] D. Hurlstone et al. Rectal aberrant crypt foci identified using high-magnification-chromoscopic colonoscopy: biomarkers for flat and depressed neoplasia. *American Journal of Gastroenterology*, pages 1283–1289, 2005.

[14] K. Ito and K. Kunisch. *Lagrange multiplier approach to variational problems and applications*, volume 15 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.

[15] M. Kida, K. Kobayashi, and K. Saigenji. Routine chromoendoscopy for gastrointestinal diseases: indications revised. *Endoscopy*, 35:590–596, 2003.

[16] R. A. Ludolph, W. R. Vleth, and H. L. Frisch. Effects of linear reaction and sorption processes on the penetrant time lag. *The Journal of Physical Chemistry*, 83(21):2793–2796, 1979.

[17] R. M. M. Mattheij, S. W. Rienstra, and J. H. M. ten Thije Boonkkamp. *Partial differential equations*. SIAM Monographs on Mathematical Modeling and Computation. Philadelphia, PA, 2005.

[18] J. D. Murray. *Mathematical biology. I*, volume 17 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, third edition, 2002. An introduction.

[19] R. J. Naftalin, P. S. Zammit, and K. C. Pedley. Concentration polarization of fluorescent dyes in rat descending colonic crypts: evidence of crypt fluid absorption. *Journal of Physiology*, 487:479–495, 1995.

[20] J. R. Olliver, C. P. Wild, P. Sahay, S. Dexter, and L. J. Hardie. Chromoendoscopy with methylene blue and associated DNA damage in barretts oesophagus. *The Lancet*, 362(21):373–374, 2003.

[21] A. Quarteroni, R. Sacco, and F. Saleri. *Numerical mathematics*, volume 37 of *Texts in Applied Mathematics*. Springer-Verlag, Berlin, second edition, 2007.

[22] M. Renardy and R. C. Rogers. *An introduction to partial differential equations*, volume 13 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1993.

[23] L. Roncucci, A. Medline, and W. R. Bruce. Classification of aberrant crypt foci and microadenomas in human colon. *Cancer Epidemiology, Biomarkers & Prevention*, 1:57–60, 1991.

[24] D. F. Swinehart. The Beer-Lambert law. *J. Chem. Educ.*, 39(7):333–335, 1962.

[25] T. Takayama, S. Katsuki, Y. Takahashi, et al. Aberrant crypt foci of the colon as precursors of adenoma and cancer. *The New England Journal of Medicine*, 339:1277–1284, 1998.

[26] The Mathworks, Inc. *http://www.matlab.com.*

Isabel N. Figueiredo
CMUC, Department of Mathematics, University of Coimbra, 3001-454 Coimbra, Portugal.
E-mail address: isabelf@mat.uc.pt

Pedro Figueiredo
Faculty of Medicine, University of Coimbra, 3004-504 Coimbra, and Department of Gastroenterology, University Hospital of Coimbra, 3000-075 Coimbra, Portugal.
E-mail address: pedro.n.figueiredo@clix.pt

Nuno Almeida
Department of Gastroenterology, University Hospital of Coimbra, 3000-075 Coimbra, Portugal.
E-mail address: nuno.p.almeida@clix.pt