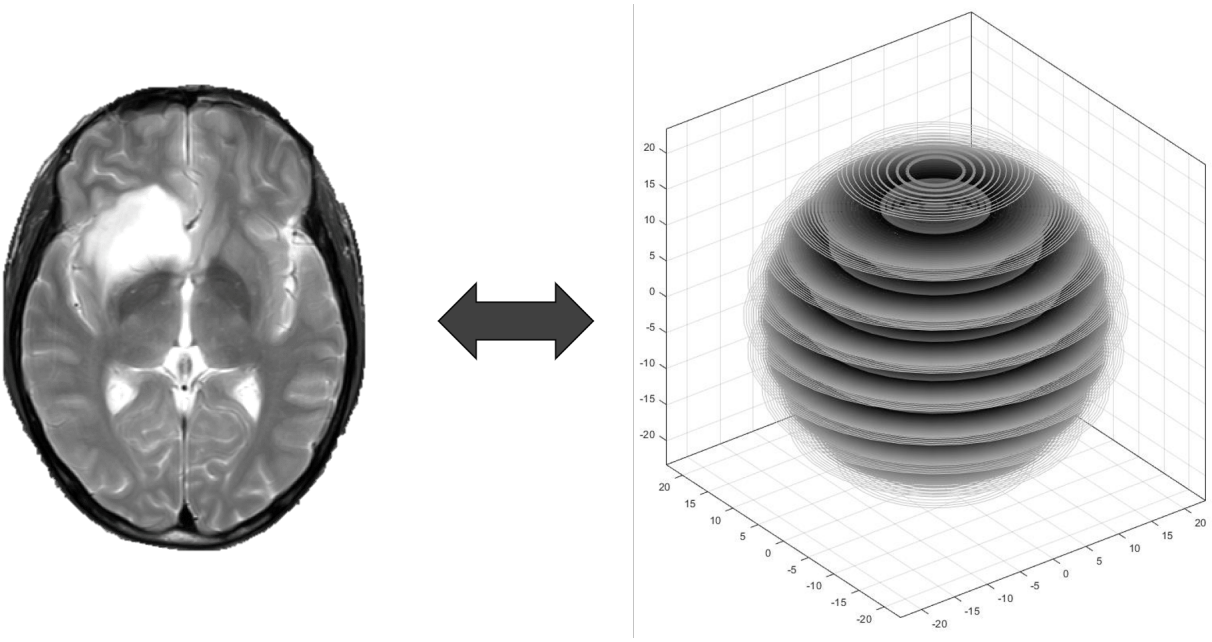
CHALMERS
UNIVERSITY OF TECHNOLOGY



Solving the Fisher Equation to Capture Tumour Behaviour for Patients with Low Grade Glioma

Master's thesis in Engineering Mathematics and Computational Science

JULIA LARSSON

Department of Mathematical Sciences
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2019

Master's thesis 2019

Solving the Fisher Equation to Capture Tumour Behaviour for Patients with Low Grade Glioma

JULIA LARSSON

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Mathematical Sciences
Division of Applied Mathematics and Statistics
Chalmers University of Technology
Gothenburg, Sweden 2019

Solving the Fisher Equation to Capture Tumour Behaviour for Patients with Low Grade Glioma
JULIA LARSSON

Supervisor: Philip Gerlee, Mathematical Sciences
Examiner: Philip Gerlee, Mathematical Sciences

Master's Thesis 2019
Department of Mathematical Sciences
Division of Applied Mathematics and Statistics
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Visualisation of how tumour properties estimated from Magnetic Resonance Images can be used to simulate tumour growth in MATLAB.

Typeset in LaTeX
Printed by Department of Mathematical Sciences
Gothenburg, Sweden 2019

Solving the Fisher Equation to Capture Tumour Behaviour for Patients with Low Grade Glioma

JULIA LARSSON
Department of Mathematical Sciences
Chalmers University of Technology

Abstract

Low Grade Glioma is a slow growing brain tumour, whose size is estimated using Magnetic Resonance Imaging and is treated with a combination of surgery, radiation and chemotherapy. However, cancer cells often remain after surgery leading to recurrence of the tumour and eventually death. To address this problem the Fisher equation has been considered, a partial differential equation describing how the cancer cell density changes over space and time. The Fisher equation can thus be fitted to patient data by comparing the tumour growth rate and the slope of the tumour interface, properties that is hypothesized to affect the survival time of the patients. The aim of this project is to simulate tumour growth to give additional information about Low Grade Glioma that can not be estimated directly from Magnetic Resonance images, but requires mathematical modelling. To do this, tumour properties have been estimated from data and statistically tested to investigate their potential effect on survival. Also, we have compared different methods of solving the Fisher equation and fitted the equation to data through parameter estimation techniques. The results show that two out of three investigated tumour properties have a significant effect on survival. The parameter estimation was successful and the different numerical methods for solving the Fisher equation yielded similar results for most cases. Additional information about the tumour was estimated from the Fisher equation, but the reliability of these results could be questioned. The main caveat is the simplicity of the Fisher equation and the small size of the patient data set. One solution could be to include the effect of surrounding tissue in the Fisher equation, but this requires accurate data containing many measurements for all patients. A second approach could therefore be to create a model using Non Linear Mixed Effect modelling, with the Fisher equation as the framework, in order to make the model more accurately capture the variation among patients.

Keywords: Low Grade Glioma, Fisher equation, survival analysis, particle swarm optimization.

Acknowledgements

I would like to thank Philip, for being my supervisor and inviting me to his group and research, as well as Alba Corell, Anja Smits and Asgeir Jakola from the Neuroscience department at Sahlgrenska for this collaboration and the data that they have provided me with. This thesis would not have been possible without these people. I would also like to thank the Bachelor students that wrote their Bachelor's Thesis with Philip in 2017. Their thesis taught me much about cancer modelling and gave me ideas about numerical solutions of the Fisher equation, amongst other things.

Lastly I would like to thank my family and friends, some close some far away, for always supporting me. Distance may part us, but I am always glad to have you.

Julia Larsson, Gothenburg, June 2019

Contents

List of Figures

List of Tables

1

Introduction

1.1 Background

Low Grade Gliomas are slow growing brain tumours that primarily affect adults with mean age at diagnosis of approximately 41 years [2, 3]. Glioma in general are classified according to a grade ranging from I to IV, where Low Grade Glioma corresponds to a Grade II glioma. Although patients suffering from Low Grade Glioma have relatively better chances of survival than patients with the more aggressive cancer type High Grade Glioma (Grade III/IV glioma) Low Grade Glioma eventually progress to High Grade Glioma, giving rise to a survival time of approximately seven years [3]. It is thus of great importance to detect the cancer in time, which can be done using Magnetic Resonance Imaging.

Magnetic Resonance Imaging (MRI) is the most commonly used tool for locating and estimating the size of the tumour [4], using either the acquisition sequence T1 or T2-FLAIR, where FLAIR stands for Fluid-Attenuated Inversion Recovery. These sequences enhances or suppress the imaging of different matter in the brain and thus give rise to images with different resolution [5, p. 95], where Low Grade Glioma typically has low signal intensity on T1 images and higher intensity on T2-FLAIR images [4]. An example of a Magnetic Resonance image (MR image) using T2-FLAIR can be seen in figure 1.1, where the tumour corresponds to the white region in the upper left part of the brain. After detecting the tumour, the doctors choose a treatment such that as many cancer cells are removed as possible while avoiding damaging surrounding healthy tissue.

The most common initial treatment of Low Grade Glioma is resection, removal of tumour by surgery, and is done on the basis of the MR image of the patient [6, 3, 4]. The problem however is that MR images can only detect cancer cells up to a certain density, meaning that the tumour is larger than the image shows [7]. Additional radiotherapy and chemotherapy are used to try to remove potential cancer leftovers from resection [6], but even after these safety measures cancer cells often remain, leading to recurrence and progression of the tumour.

The solution to this problem could be to use mathematical models that could simulate tumour growth, including the regions of the tumour that are invisible for the MR image. A popular model of choice is the Fisher equation [8, 9, 10, 1, 11], a Partial Differential Equation (PDE) that can be used to describe how the density of cancer cells changes in space and time according to how the cells migrate and

proliferate. The Fisher equation can be written as

$$\frac{\partial u(\mathbf{x},t)}{\partial t} = \nabla \cdot (D\nabla u(\mathbf{x},t)) + \rho u(\mathbf{x},t)\,(1 - u(\mathbf{x},t))$$

where $u(\mathbf{x},t)$ is the proportional cancer cell density over space $\mathbf{x}$ and time t, D is the diffusion coefficient governing the migration term and ρ is the growth rate, governing the proliferation term. The values of D and ρ can be estimated from certain tumour properties calculated from the MR images, properties that also are hypothesized to have an effect on survival. With the parameter values estimated from these tumour properties one could simulate tumour growth, as in figure 1.2, in order to learn more about the behaviour of Low Grade Glioma.

Estimating the parameters D and ρ can be done in different ways, either using parameter estimation techniques [9], or using an analytic expression of the parameters that is a result from an approximation of the Fisher equation [1, 12, 13]. The approximation however, as will be discussed further in chapter 2, requires assumptions that are rather unreasonable [14, pp. 439–447]. Although the analytic expression is commonly used because of its simplicity, one should question if the approximation really is sufficient to get a realistic representation of Low Grade Glioma.

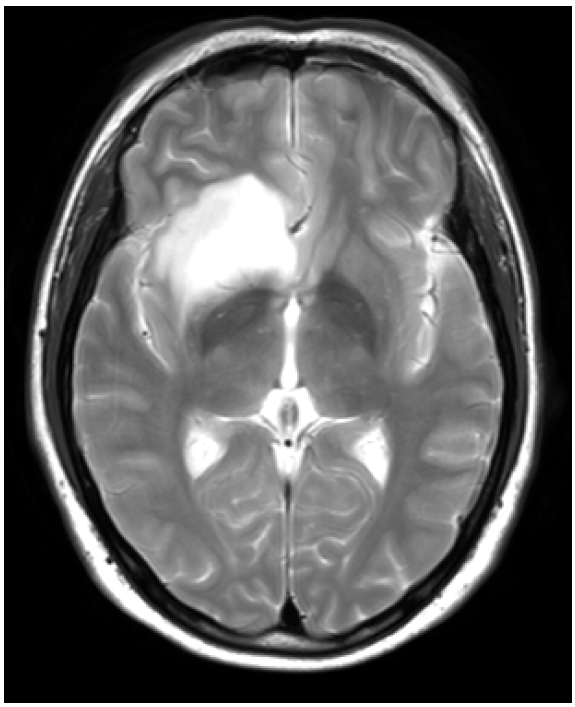


Figure 1.1: *Magnetic Resonance Image using acquisition sequence T2-FLAIR. The Low Grade Glioma corresponds to the white tract in the upper left of the brain.*

1.2 Aim

The aim of this project is to mathematically simulate tumour growth to give additional information about Low Grade Glioma, that can not be estimated directly from Magnetic Resonance images, in order to improve the treatment of Low Grade Glioma. To do this, the tumour properties of interest are extracted from patient data and statistically tested using survival analysis to investigate their potential effect on survival. Also, the different methods of solving the Fisher equation are compared and, using patient data, the values of diffusion coefficient D and growth rate ρ for each patient are determined using parameter estimation techniques.

The work behind this thesis was carried out in collaboration with the Neuroscience department at Sahlgrenska and combines two completely different research areas, mathematical modelling and medicine, in order to get the whole picture of how Low Grade Gliomas behave. The contribution of this thesis to the community of cancer modelling is to verify that the approximations of the Fisher equation are valid and to estimate the parameter values D and ρ using an optimisation algorithm, instead of using the analytic equations for the parameters.

Also, this thesis contributes to the doctors at Sahlgrenska by investigating the behaviour of Low Grade Glioma in depth, using the Fisher equation, and by giving an estimate of the volume of remaining cells after surgery. According to the neurosurgeon Asgeir Jakola, this estimate can be used to tune the additional radiotherapy more accurately, in order to remove as many cancer cells possible while minimising the damages done to surrounding healthy brain tissue. In addition, the size of the estimated remaining volume can be helpful when deciding which marginal should be used at surgery. If the estimate is large, it would be worth the risk to choose a marginal larger than the MR image would suggest to remove the remaining cells.

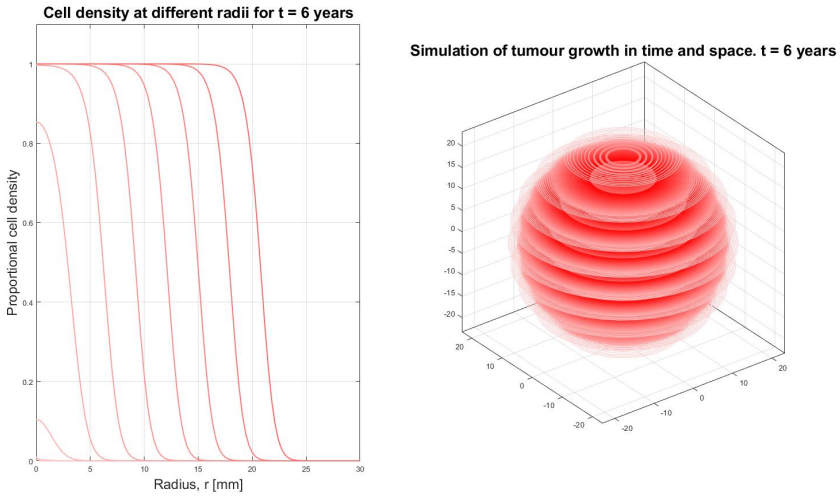


Figure 1.2: *Left: The solution to the Fisher equation in two dimensions. Right: The cancer cell distribution it would correspond to in three dimensions, assuming isotropy and spherical symmetry.*

2

Theory

In this chapter, the theoretical background of the Fisher equation and the properties of investigation, velocity, steepness and mismatch, is described. In addition, the theory behind the statistical tests will be introduced. Mainly studies and conclusions from previous work within the field of cancer modelling are presented here, while the application of the theory, amongst others, are presented in chapter 3.

2.1 The Fisher equation

In this section, the Fisher equation and the different simplifications, and expansions, of the equation is described. As shown in section 1.1, the Fisher equation in three spatial dimensions is defined as

$$\frac{\partial u(\mathbf{x},t)}{\partial t} = \nabla \cdot (D\nabla u(\mathbf{x},t)) + \rho u(\mathbf{x},t)\,(1 - u(\mathbf{x},t)) \tag{2.1}$$

where $u(\mathbf{x},t)$ is the proportional cancer cell density over space $\mathbf{x}$ $[mm]$ and time t $[years]$, D is the diffusion coefficient governing the migration term $[mm^2/years]$ and ρ is the growth rate, governing the proliferation term $[years^{-1}]$. The reason why $u(\mathbf{x},t)$ is a *proportional* cell density is because $u(\mathbf{x},t)$ is relative to a carrying capacity K which is the limiting amount of cancer cells [#cancer cells/mm^3] [1]. $u(\mathbf{x},t)$ is therefore a ratio, ranging from 0 to 1, describing the cancer cell density relative to the carrying capacity K. Since equation (2.1) is difficult and computationally heavy to solve, some simplifications will be considered.

One assumption is to consider a constant value of the diffusion coefficient in space, $D(\mathbf{x}) = D$. In reality, D is space dependent since it has been shown that gliomas tend to spread more rapidly through white matter than in grey matter tracts in the brain [15, p. 544]. Models have been created taking into account the non-constant value of D [10], but is neglected in this thesis for simplicity. Using this approximation, the Fisher equation can be written as

$$\frac{\partial u(\mathbf{x},t)}{\partial t} = D\left(\nabla^2 u(\mathbf{x},t)\right) + \rho u(\mathbf{x},t)\,(1 - u(\mathbf{x},t)) \tag{2.2}$$

where the nabla operator correspond to either Cartesian coordinates ($\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$)

or Spherical coordinates $\nabla = \left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \varphi}\right)$. If using Spherical coordinates, the spatial dimensions can be reduced from three to only one, if assuming isotropy and spherical symmetry. These assumptions means that the spread of cancer cells in space is uniform, which once again contradicts the biology behind the model but will be used for simplicity. In practice, isotropy and spherical symmetry leads to the proportional cancer cell density u being independent of the angles θ, φ, meaning that $\frac{\partial u}{\partial \theta} = \frac{\partial u}{\partial \varphi} = 0$. The Fisher equation, in Spherical coordinates, thus becomes

$$\frac{\partial u(r,t)}{\partial t} = D\left(\frac{\partial^2 u(r,t)}{\partial r^2} + \frac{2}{r}\frac{\partial u(r,t)}{\partial r}\right) + \rho u(r,t)\left(1 - u(r,t)\right) \tag{2.3}$$

where x has been changed to r, to clarify which spatial coordinates that are being used. Otherwise x and r share the exact same properties, thus if something is explained in terms of x, it is interchangeable with r. The reduction of dimensions in Spherical coordinates simplifies the computation of the Fisher equation, but another way to do it is to simply reduce the nabla operator in Cartesian coordinates, from three to one dimension. If once again assuming isotropy, the Fisher equation becomes

$$\frac{\partial u(x,t)}{\partial t} = D\left(\frac{\partial^2 u(x,t)}{\partial x^2}\right) + \rho u(x,t)\left(1 - u(x,t)\right) \tag{2.4}$$

but does not take into account the curvature of the surface of the tumour, that is assumed to be approximately spherical, which equation (2.3) does. With the Partial Differential Equations defined, the next step is to define the initial and boundary conditions.

The conditions used when solving the equations in (2.3) and (2.4) is an initial exponential spread in space and no flux boundary at the origin and at the outer boundary. In equations, these conditions would correspond to

$$\begin{cases} u(x,0) = a_0 e^{-10x} \\ \frac{\partial u(x,t)}{\partial x} = 0, \quad x = 0 \text{ or } x = x_{max} \end{cases} \tag{2.5}$$

where a_0 ranges between 0 and 1. The initial condition is chosen in concordance with Philip Gerlees article from 2012 [8], while the boundary at the origin hinders the center of the tumour from moving from its initial position and the outer boundary prevents the tumour to spread indefinitely [11]. With these simplifications of the PDE and definitions of the initial condition and the boundary conditions, the Fisher equation can be solved numerically.

The numerical solution to equation (2.3) is a matrix containing values of $u(r,t)$ for each combination of r and t, where t varies over the columns and r over the rows in the matrix. The result is visualised in figure 2.1 in both two and three dimensions. In the rightmost plot $u(r,t)$ is plotted against both r and t, visualising how the values of $u(r,t)$ varies in the matrix of the numerical solution. In the leftmost plot, $u(r,t)$ is plotted against r at different time points. As can be seen, the proportional cancer cell density is initially small and grows larger over time. When maximum

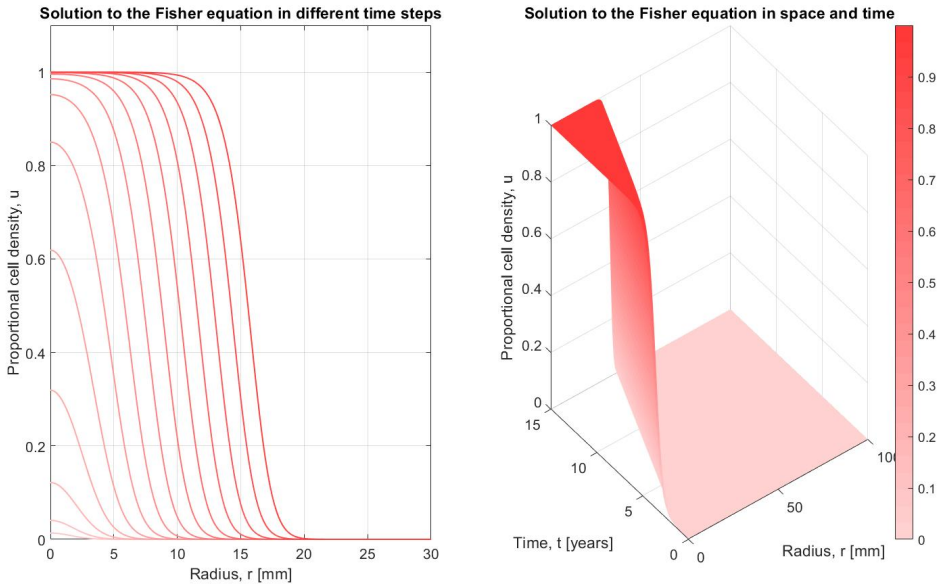


Figure 2.1: *Left: Solution to the Fisher equation in two dimensions, where each curve corresponds to the solution at a certain time point. Right: The solution to the Fisher equation plotted against both space and time.*

proportional cell density is reached, $u(r,t) = 1$, the distribution in space continues as a constant wave, whose shape is independent of t. To describe this constant travelling wave, if neglecting the initial phase, an analytic expression derived from the Fisher equation can be used.

Following the derivation done in the book *Mathematical Biology I: An Introduction*, written by J.D. Murray in 2002, the Fisher equation can be approximated as [14, p.446]

$$u(z) = \frac{1}{1+e^{\frac{\rho z}{c}}} + \frac{e^{\rho z/c}}{c^2(1+e^{\rho z/c})^2}\ln\left(\frac{4e^{\rho z/c}}{(1+e^{\rho z/c})^2}\right), \quad z = x - ct \tag{2.6}$$

where the z-coordinate travels with the solution and is dependent of both x and t, but also the travelling wave velocity c $[mm/year]$, expressed as $c = 2\sqrt{D\rho}$ which will be discussed further in section 2.2 below. This solution to the Fisher equation is done by using perturbation technique and power series, where each term in the series gives additional refinements of the approximation. However, as Murray also comments, the first term alone in equation (2.6) captures the behaviour of the Fisher equation quite accurately when comparing it to the numerically computed Fisher equation [14, p. 442]. Thus for simplicity, the second term in equation (2.6) will be neglected and the analytic equation of the Fisher equation becomes:

$$u(z) = \frac{1}{1+e^{\frac{\rho z}{c}}}, \quad z = x - ct \tag{2.7}$$

Although the analytic equation is claimed to capture the travelling wave it still

has drawbacks. Firstly, as already mentioned, it is only an approximation of the PDE. Secondly, it is derived from the Fisher equation in Cartesian coordinates, equation (2.4), and not equation (2.3) in Spherical coordinates, since no travelling wave solution exist for equation (2.3) due to the $1/r$ term in the migration term. Lastly, it is required for the analytic equation that the travelling wave has already been established, which theoretically requires that $t \to \infty$ and that no initial phase exists [14, pp. 439–447]. It is thus of interest to investigate if the approximation really can capture tumour behaviour and estimate its properties.

Variations of the Fisher equation and its boundaries exists, for example exponential growth (ρu) instead of logistic growth ($\rho u(1-u)$) [13], as well as a normally distributed initial cancer cell density rather than exponential [10]. In addition, expansions of the Fisher equation exists giving more complexity to the model. To give some examples, Amanda Swan added spatial dependency to the model to include the effect of cancer cells spreading more rapidly in white matter tracts [10], while Philip Gerlee incorporated the 'Go or grow' hypothesis, stating that cancer cells *either* migrate or proliferate [8]. These models are however implemented for modelling glioma of higher grade, but Magdalena Bogdańska in contrast focused on the model for Low Grade Glioma, modelling the phase where the Low Grade Glioma progress into High Grade Glioma [11].

As these examples show, there are many ways to vary or expand the Fisher equation, but in this thesis only the one dimensional Fisher equation in Spherical coordinates (2.3), along with the initial and boundary conditions in equation (2.5), and the analytic equation of the Fisher equation (2.7) are considered, called the PDE solution and Analytic Equation henceforward.

2.2 Properties of the Low Grade Glioma: Velocity, steepness and mismatch

Important tumour properties are velocity, steepness and mismatch, where the velocity illustrates how rapid the tumour grows while steepness and mismatch are two measures describing the spatial distribution of the cancer cells. From volume estimates of the tumours visualised on MR images, these properties can be calculated and used to help the doctors make more appropriate decisions regarding treatment, such as performing surgery earlier for a patient with high tumour velocity or use larger marginals at surgery for patients with more diffuse tumours.

The tumour velocity is the increase of volume over time and is used as an indicator of tumour grade and overall survival [16]. If assuming that the tumour is approximately elliptic or spherical, it has been shown that the diameter or radius of Low Grade Gliomas increase linearly over time with a constant diametric velocity of approximately $4mm/year$ [17, 13, 18, 19]. In addition, it has been shown that patients with a diametric velocity larger than $8mm/year$ have a higher risk of dying from their cancer [20].

The steepness is a measure of the spatial cell density distribution and is measured

as the slope of the density profile between radii estimated from two MR images, using acquisition sequence T1 and T2-FLAIR respectively, at the same time point [1, 12]. A small value of steepness would correspond to a diffuse tumour while a large value would correspond to a dense tumour and, in contrast to the velocity, steepness has been shown to *not* have an impact on survival for patients with glioblastoma multiforme, a more aggressive type of brain tumour. It could thus be hypothesized that it should not have an impact on Low Grade Glioma as well [21].

Lastly, the mismatch is used to measure how large the volume estimated from an MR image using T1 is, relative to the volume estimated from an image using T2-FLAIR [21]. Mismatch is a less commonly used measure than velocity and steepness and consequently, less is known about this measure. The doctor Asgeir Jakola at the Neuroscience department at Sahlgrenska University Hospital however believes that mismatch could be an interesting measure with potential and is thus considered in this thesis.

As a simplified explanation of how these properties are measured from MR images, see figure 2.2. The figure illustrates three MR images from one patient, two images taken with T2-FLAIR at two different time points, t_1 and t_2, and one image taken with T1 at one time point t_1. The velocity thus describes the increase of tumour radii on MR images taken with T2-FLAIR over time, the steepness how diffuse the tumour is on MR images using T1 in comparison with T2-FLAIR and the mismatch the difference in volume between the images taken with T1 and T2-FLAIR respectively.

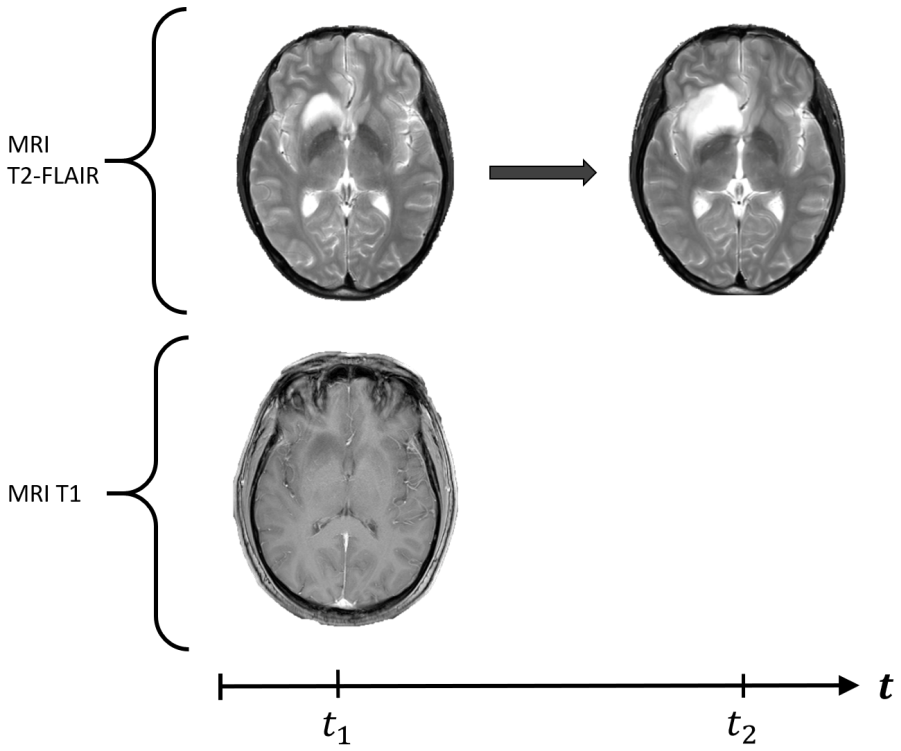


Figure 2.2: *MR images using either T1 or T2-FLAIR. The velocity is calculated using images taken with T2-FLAIR and steepness and mismatch are calculated using images taken with T1 and T2-FLAIR at the same time point.*

2.3 Velocity and steepness derived from the Fisher equation

Another central role for the properties velocity and steepness is that they are the key to estimate the values of the parameters D and ρ in the Fisher equation. Since the time span between the initial appearance of the tumour and diagnosis is unknown, it is impossible to fit the Fisher equation directly to the volume estimates. One must thus instead do this using the velocity and steepness, since they can both be estimated from data *and* the Fisher equation.

What velocity and steepness would correspond to in the Fisher equation is shown in figure 2.3, where c is the radial velocity $[mm/year]$ and s is the steepness $[mm^{-1}]$. The velocity is a measure of how fast the solution to the Fisher equation travels in time and the steepness corresponds to the value of the negative slope of the constant travelling wave. The properties c and s, as well as mismatch MM, can be estimated numerically which will be shown in section 3.3, but c and s can also be solved analytically through approximations of the Fisher equation [14, p. 442, 446]. The equations are

$$\begin{cases} c = 2\sqrt{\rho D} \\ s = \frac{\rho}{4c} = \frac{1}{8}\sqrt{\frac{\rho}{D}} \end{cases} \tag{2.8}$$

and have a time independent relation with D and ρ, where D is the diffusion coefficient $[mm^2/year]$ and ρ is the growth rate $[year^{-1}]$. Using the systems of equations in (2.8) one can thus easily calculate D and ρ and fit the Fisher equation to patient data [1]. The expression for the steepness s in (2.8) however is calculated from $u = 0.5$, where u is the proportional cancer cell density, when it is hypothesized that MR images using T1 and T2-FLAIR can detect cell densities of 0.8 and 0.16 of higher respectively [1], although these thresholds have been criticised [9]. In order to get a better correspondence to data, x is extracted from equation (2.7) and written as:

$$z = x - ct = \frac{c}{\rho}\ln\left(\frac{1}{u} - 1\right) \quad \Rightarrow \quad x = ct + \frac{c}{\rho}\ln\left(\frac{1}{u} - 1\right) \tag{2.9}$$

Using equation (2.9), the steepness of the secant can be calculated as

$$\begin{cases} c = 2\sqrt{\rho D} \\ s = \frac{-(0.16-0.80)}{x_{0.16}-x_{0.80}} = \frac{-(0.16-0.80)}{\frac{c}{\rho}\left(\ln\left(\frac{1}{0.16}-1\right)-\ln\left(\frac{1}{0.80}-1\right)\right)} \end{cases} \tag{2.10}$$

where, once again, the expressions are time independent since the term ct cancels out when taking the difference between $x_{0.80}$ and $x_{0.16}$, the radii at $u = 0.80$ and $u = 0.16$ respectively [21].

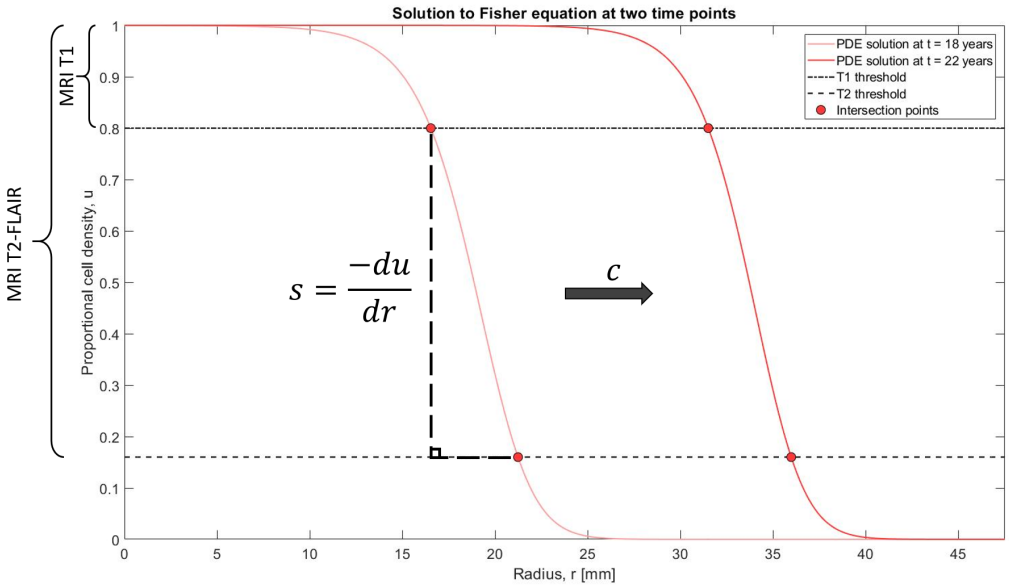


Figure 2.3: *Visualisation of how velocity c* $[mm/year]$ *and steepness s* $[mm^{-1}]$ *are estimated from the Fisher equation, as well as the two thresholds for the different MR images. The two curves are the same solution to the Fisher equation at two time points, where c is how fast the curve travels and s the steepness of the secant between* $u = 0.80$ *and* $u = 0.16$ *[1]. Images using T2-FLAIR can thus detect cell densities of 0.16 or higher and T1 can detect densities of 0.80 or higher.*

2.4 Survival analysis

In order to investigate if tumour properties have a significant effect on survival, statistical tests from survival analysis will be used where three central concepts for these tests are *survival time*, *event* and *censoring*.

The *survival time* is the time span after treatment [22, p. 58] and an *event* corresponds to the property of investigation in the survival analysis, which is often death of a patient. All other patients that leave the study for other reasons or have not experienced an event are called *censored* [23]. These definitions are used in the log-rank test, visualised using Kaplan-Meier plots, which is the model of choice for this purpose.

The log rank test is a statistical test that checks if there is a significant difference in survival time between two or more patient groups, based on a certain property of the patients [22, p. 67–71]. For example, Johan Pallud et.al. divided a patient data set in two groups where the first group had a tumour diametric velocity below $8mm/year$ and the second group above $8mm/year$ and showed that the patient group with higher velocity had worse survival compared with the other group [20]. In general, the hypotheses for the log rank test is

$$\begin{cases} H_0 : \text{There is no difference in survival between the groups} \\ H_1 : \text{There is a difference in survival between the groups} \end{cases} \tag{2.11}$$

and the log–rank test is a large-sample chi-square test with statistic

$$\text{Log-rank statistic} = \frac{(O_i - E_i)^2}{\text{Var}(O_i - E_i)}, \quad i = 1, 2 \tag{2.12}$$

where O_i is the observed number of events for group i, E_i is the calculated expected number of events for group i and $\text{Var}(O_i - E_i)$ is the variance of the difference between O_i and E_i. The log-rank test statistic under the null hypothesis H_0 is approximately chi-square with one degree of freedom and the p-value is thus calculated using the chi-square distribution [22, p. 67–71]. The reason why log-rank test is commonly used in survival analysis is because the test can use information from both patients having an event *and* censored patients, since it is assumed that the censored patients have the same likelihood of survival as for those patients where an event have occurred [23]. The log rank test can be expanded to be applied to a data set with more groups than two, but is neglected in this thesis.

To illustrate the survival time in the different groups, a Kaplan-Meier plot is commonly used. The plot illustrates the surviving fraction of patients in the two groups over time. When an event occurs, the curve takes a step down illustrating that the frequency of surviving patients has decreased. When a patient leaves because of censoring, it is marked with a + sign [23].

3

Methods

In this chapter the application of the theory will be presented, as well as the methods used in this specific project. The sections below describes the idea of how to answer each question of the aim: Extract tumour properties from data, investigating tumour properties using statistics, compare different methods of solving the Fisher equation and estimate D and ρ using parameter estimation techniques. For each method, arguments are presented answering the question of why the method of choice is sufficient for this purpose.

To clarify what has been done by the thesis author and not, the PhD student Alba Corell from the Neuroscience department at Sahlgrenska University Hospital is main responsible for the tumour volume estimates from MR images and the clinical data. The rest of the work is done by the author, with inspiration from the Bachelor thesis from 2017 with Philip Gerlee as supervisor [24]. Lastly, the survival analysis was done using the program language R (version 3.5.2) together with the packages `readxl`, `survival` and `survminer` [25, 26, 27], while the rest was made using MATLAB (R2017b).

3.1 Patient data and properties measured from data

In this section, the extraction of tumour volumes from MR images and the methods for estimating velocity, steepness and mismatch are presented. Here, as well as section 3.3 below, volumes and radii estimated using MR images with either acquisition sequence T1 or T2-FLAIR will be frequently mentioned. Since MR images using T1 and T2-FLAIR are assumed to detect cell densities of 0.80 and 0.16 respectively, the subscripts for the volumes and radii will be 0.80 and 0.16, rather than T1 and T2-FLAIR.

The data consists of 28 patients and was produced mainly by the PhD student Alba Corell, with help of the neurosurgeon Asgeir Jakola. In addition, this thesis is a part of a larger study with granted ethical permission and inclusion of patients is based on informed consent. The data set consists of patient age at diagnosis, information regarding if the patient is alive or not at check-up at 2019-01-01 and tumour volumes estimated from Magnetic Resonance images using acquisition sequence T1 and T2-FLAIR respectively.

When doing an MRI scan, the images received are 3D images segmented into 2D images using a constant segmentation depth d between the images. From these images, a tumour area is estimated by a doctor for each segment and the volume is calculated as the sum of the volumes of each segment, calculated as the estimated area multiplied with the constant depth d. For an illustration, see figure 3.1. Each patient have two or more volume estimates from MR images using T2-FLAIR at different time points and only one volume estimate from an MR image using T1, taken at the same time point as one of the images using T2-FLAIR. Using these estimated volumes, velocity, steepness and mismatch can estimated.

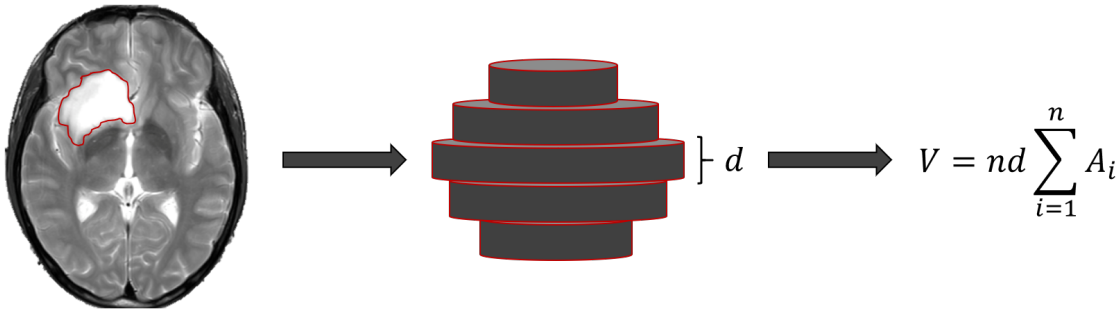


Figure 3.1: *Illustration of how the volume is estimated from MR images. From an MRI scan, segmented images are retrieved with constant segmentation depth d between the images. From every image, an area is estimated by a doctor and the volume is calculated as the sum of the volumes estimated from each segmented disc.*

The tumour velocity is estimated by measuring the tumour radii from MR images using T2-FLAIR and calculated using linear regression

$$r^i = r_0^i + c^i t \quad i = 1, 2, ... \tag{3.1}$$

where the radius r increases linearly with the constant velocity c over time t for all patients i [16]. The velocity c is then estimated from equation (3.1) using least squares.

The steepness is measured as the negative slope of the density profile between the radii at the visibility thresholds for MR images using T1 and T2-FLAIR respectively, and is calculated as

$$s = \frac{-(0.16 - 0.80)}{r_{0.16} - r_{0.80}} \tag{3.2}$$

where s is the steepness $[mm^{-1}]$, $r_{0.16}$ is the radius estimated from an MR image taken with T2-FLAIR and $r_{0.80}$ the radius estimated from an MR image taken with T1 $[mm]$. Lastly, the mismatch is a measure of how large volume that is measured from an MR image using T1, relative the volume estimated from an MR image using T2-FLAIR. The mismatch MM is calculated as

$$\text{MM} = \frac{V_{0.16} - V_{0.80}}{V_{0.16}} = 1 - \frac{V_{0.80}}{V_{0.16}} \in [0, 1] \tag{3.3}$$

where $V_{0.80}$ and $V_{0.16}$ are the volumes measured from MR images using T1 and T2-FLAIR respectively $[cm^3]$. Other equations for the mismatch exist where they instead use the ratio $V_{0.80}/V_{0.16}$ [21]. However since $V_{0.16} > V_{0.80}$ [4], the values of the mismatch becomes bounded between 0 and 1 if using equation (3.3), which is easier to interpret. If mismatch is zero, the difference in estimated volume is zero and shows a perfect match between the images, while a mismatch of 1 means that $V_{0.80} = 0$ and shows total mismatch.

3.2 Survival analysis

In this section, the implementation of the theory described in section 2.4 is presented. The *survival time* is defined as the time span after surgery that the patient has been alive, used in concordance with Alba Corells article from 2018 [2], which is either until a patient dies or until check-up 2019-01-01. If a patient dies from her tumour, it is called an *event* and all other patients that have either died of other reasons or are alive at check-up are called *censored*. With these concepts defined the log-rank test can be applied to the data set, using a p-value of significance of 0.05.

Three properties are investigated using the log-rank test: Mismatch, invasiveness and velocity. In each case, the patients are divided into two groups based on their values of the properties of investigation. The division in groups for the velocity is done in concordance with an article from 2006, written by Johan Pallud et.al. [20], and the mismatch and invasiveness based on their median value.

For the mismatch, the patients were divided in small mismatch ($MM < 0.5$) and large mismatch ($MM > 0.5$), while for the velocity the patients were divided into groups based on their radial velocity, $c < 4mm/year$ and $c > 4mm/year$ respectively. The invasiveness differs slightly from the steepness, but is also a measure of the cancer cell density profile and is used in concordance with previous studies [21]. The invasiveness is defined as ρ/D and is calculated according to the equation for the secant, also presented in section 2.2 in equation (2.10):

$$s = \frac{-(0.16 - 0.80)}{2\left(\ln\left(\frac{1}{0.16} - 1\right) - \ln\left(\frac{1}{0.80} - 1\right)\right)}\sqrt{\frac{\rho}{D}} = A\sqrt{\frac{\rho}{D}} \tag{3.4}$$

The invasiveness is thus calculated as $\rho/D = (s/A)^2$ and the patients were once again divided into two groups, $\rho/D < 10$ and $\rho/D > 10$ mm^{-2} respectively.

3.3 Velocity, steepness and mismatch estimated numerically from the Fisher equation

From Magnetic Resonance images the tumour velocity, steepness and mismatch can be estimated and, if using the velocity and steepness, the parameters D and ρ in the Fisher equation can be estimated. As mentioned in section 2.2, if using the analytic estimates of the velocity c and steepness s one can get a direct correlation between the properties and the parameters D and ρ. However, for the PDE solution of the Fisher equation, no such direct correlation between properties and parameters exists. In addition, c and s are time dependent for the PDE solution of the Fisher equation and it is thus of interest to investigate how c and s changes over time and how they depend on parameters D and ρ.

What velocity and steepness would correspond to in the Fisher equation was shown in figure 2.3, where c is the velocity of the constant travelling wave and s the negative slope of the secant. To numerically calculate the velocity and steepness from the PDE solution of the Fisher equation the following system of equation is used:

$$\begin{cases} c(t_k) = \frac{u_{half}(r,t_{k+1}) - u_{half}(r,t_k)}{t_{k+1} - t_k} \\ s(t_k) = \frac{-(0.16 - 0.80)}{r_{0.16}(t_k) - r_{0.80}(t_k)} \end{cases} \tag{3.5}$$

Here u_{half} is the maximum proportional cell density divided in half, $u_{half} = 0.5$ if the constant travelling wave has been reached and $u_{half} < 0.5$ else. In addition, t_k and t_{k+1} is the time at time point k and $k+1$ respectively and $r_{0.16}$ and $r_{0.80}$ are the radii measured at $u = 0.16$ and $u = 0.80$ respectively, the hypothesized thresholds for T2-FLAIR and T1. These methods of calculating the velocity and steepness numerically are inspired by the method used in the Bachelor thesis from 2017, with Philip Gerlee as supervisor [24], and gives estimates of the properties that can be plotted over time.

The velocity and steepness profiles over time, using the equations in (3.5), are shown in figure 3.2. As can be seen, the velocity can be measured from $t = 0$ while the steepness of the secant can only be calculated from the time when $u = 0.8$ has been reached and forward. During the initial phase the values of c and s varies but becomes approximately constant after the travelling wave has been reached.

As mentioned, c and s are time dependent in equation (3.5) while c and s estimated from data have a constant value. In order to avoid the time dependency it is assumed that the constant travelling wave has already been reached, where c and s are approximately constant, when the measurements from the patients have been taken. This assumption is strengthened by the reported observations that Low Grade Gliomas grows linearly with a radial velocity of approximately $2mm/year$ [17].

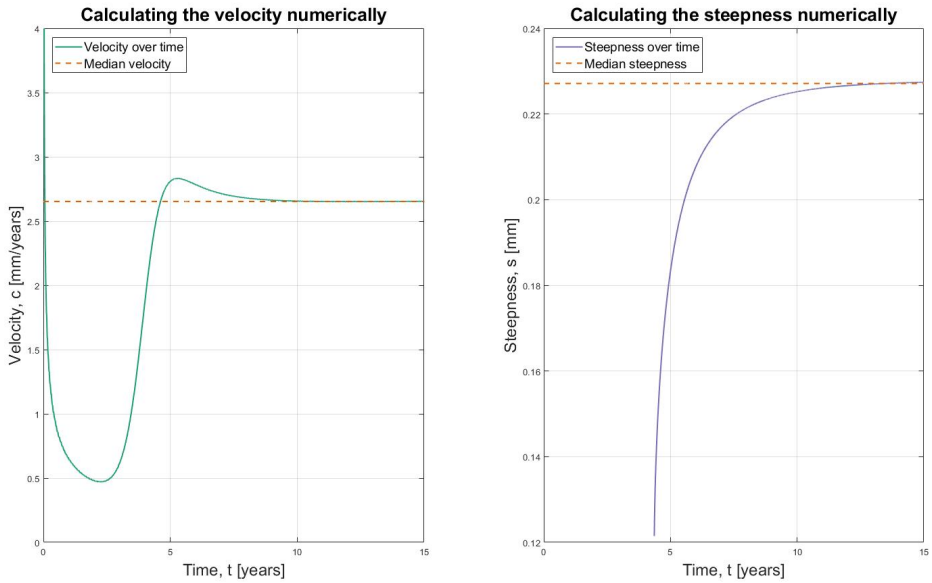


Figure 3.2: *Velocity and steepness profile for the Fisher equation over time, calculated according to the equation* (3.5)*. To get the velocity and steepness of the constant wave, the median of the last 20% of the two properties are calculated.*

To calculate the constant velocity and steepness from the numerical estimates of c and s, the median of the last 20% of the measured velocities and steepnesses at each time point is calculated. In this way, only the velocities and steepnesses measured during the constant travelling wave are used and the median of the properties are taken to reduce the effect of noise. So far only how the velocity and steepness are calculated from the Fisher equation have been described, but a numerical method for calculating the mismatch has been developed as well.

To calculate the mismatch numerically, a method very different to the estimation of c and s is used. As mentioned in section 3.1, the equation for the mismatch is defined as

$$\mathrm{MM} = \frac{V_{0.16} - V_{0.80}}{V_{0.16}} = \frac{4\pi}{3}\left(\frac{r_{0.16}^3 - r_{0.80}^3}{r_{0.16}^3}\right) \tag{3.6}$$

if assuming the tumour to be spherical. One must thus find the radii where the proportional cancer cell density $u(r,t)$ equals 0.16 and 0.80 in the matrix containing the numerical solution. This is done by, for a fix t, searching in the matrix for the column index where $u(r,t) = 0.16$ and $u(r,t) = 0.80$ are fulfilled respectively, since $u(r,t)$ varies with respect to r over the columns. After finding the column indices, they are inserted in the vector of radii that was used to get the numerical solution to $u(r,t)$ and gives back the radii $r_{0.80}$ and $r_{0.16}$. These radii are lastly inserted in equation (3.6) and the mismatch is calculated.

An contour plot, illustrating the values contained in the matrix, is visualised in figure 3.3 in order to illustrate the process. For each radius and time point, there

exists a proportional cancer cell density and to extract one property, for example the radius, the remaining two needs to be known and fix.

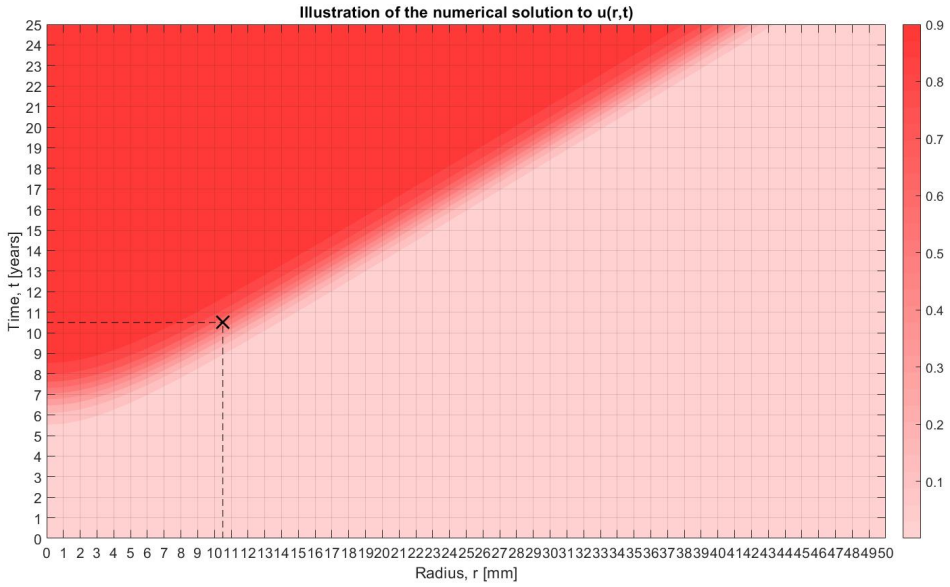


Figure 3.3: *Illustration of the matrix of the numerical solution and how the corresponding values of r and t for a certain cell density can be determined from the matrix.*

3.4 Comparison of solvers of the Fisher equation

Since the Analytic Equation is commonly used while PDE solution of the Fisher equation theoretically better corresponds to reality, it would be of interest to compare the different ways of solving the Fisher equation, both concerning estimating different properties and the solution to the Fisher equation itself. For this comparison, the parameter values used are $D \in [0.3285, 0.973]$ $mm^2/year$ for the diffusion coefficient and $\rho \in [0.73, 2.92]$ $year^{-1}$ for the growth rate. The choice is based on parameter values from literature that give a velocity around 2 $mm/year$ and all give reasonable results within the intervals $t \in [0, 30]$ $years$, $r \in [0, 100]$ mm with an initial density of 0.1 [11, 18].

The properties of interest are the velocity c and steepness s, either solved according to the analytic equations in (2.10) or numerically using the equations in (3.5) mentioned above. The properties c and s are calculated for each combination of parameters D and ρ and the difference in estimated value for c and s between the solvers is calculated as

$$\begin{cases} d_c = (c_{PDE} - c_{Analytic})/c_{Analytic} \\ d_s = (s_{PDE} - s_{Analytic})/s_{Analytic} \end{cases} \tag{3.7}$$

where d_c and d_s is the difference in velocity c and steepness s respectively, and the subscripts *Analytic* and PDE tells if the velocity and steepness have been estimated using the analytic or numerical approach respectively.

If instead considering the comparison between solvers of the Fisher equation, the PDE and the Analytic Equation respectively, the sum of squares is used as an error measure. The sum of squares is calculated as

$$\mathrm{SS}(t) = \frac{1}{N}\sum_{j=1}^{N}\left(u(r_j,t) - u(z_j)\right)^2 \tag{3.8}$$

where $\mathrm{SS}(t)$ is the sum of squares over time, N is the total number of spatial steps, $u(r_j,t)$ is the PDE solution to the Fisher equation, and $u(z_j)$ is the solution to the Fisher equation using the Analytic Equation. A problem however is that since the Analytic Equation does not have an initial phase it reaches the constant wave phase more rapidly. To compensate for this, the Analytic Equation is time adjusted such that it grows on the same time scale as the PDE solution. As a result, $\mathrm{SS}(t)$ is only calculated when the PDE has reached the constant travelling wave. To investigate if the sum of squares differs for different combinations of parameter values, $\mathrm{SS}(t)$ is averaged over the time interval were the travelling wave has been reached and then plotted against D and ρ.

3.5 Parameter estimation

The goal with the parameter estimation is to determine the values of the parameters D and ρ in the Fisher equation from the velocity and steepness measured from data, in order to simulate tumour growth for specific patients. This can be done by either calculating the parameters analytically or estimating them from the PDE using parameter estimation techniques. If considering the analytic calculations no estimation is needed, since D and ρ can be solved using the systems of equations:

$$\begin{cases} c = 2\sqrt{\rho D} \\ s = \frac{1}{8}\sqrt{\frac{\rho}{D}} \end{cases} \quad \Leftrightarrow \quad \begin{cases} D = 0.0625\frac{c}{s} \\ \rho = 4cs \end{cases} \tag{3.9}$$

However, as mentioned in section 2.2, to get the steepness of the secant which would correspond better to data one should instead use

$$\begin{cases} c = 2\sqrt{\rho D} \\ s = \frac{-(0.16-0.80)}{\frac{c}{\rho}\left(\ln\left(\frac{1}{0.16}-1\right)-\ln\left(\frac{1}{0.80}-1\right)\right)} \end{cases} \quad \Leftrightarrow \quad \begin{cases} D \approx 0.053\frac{c}{s} \\ \rho \approx 4.757cs \end{cases} \tag{3.10}$$

where one can see that they differ slightly in value. To estimate the parameters from the PDE solution of the Fisher equation however, an optimization algorithm is required.

Estimating the parameters D and ρ from the PDE is a more difficult task, since there is no direct connection between velocity c and steepness s and the parameters. This means that there exists no equation that couples the properties and the parameters together, which excludes all optimization algorithms that is based on derivatives, such as gradient descent and Newton [28]. The solution is therefore to use a derivative free optimization algorithm.

The optimization algorithm of choice is the Particle Swarm Optimization algorithm (PSO), a stochastic optimization algorithm inspired by swarming animals such as birds and fish [29, p. 117]. This algorithm does not require that the objective function is dependent of D and ρ and can thus be defined as

$$f\left(c_{numeric}, s_{numeric}\right) = \frac{\left(c_{numeric} - c_{data}\right)^2}{c_{data}} + \frac{\left(s_{numeric} - s_{data}\right)^2}{s_{data}} \tag{3.11}$$

inspired by the objective function used in the Bachelor thesis from 2017 [24]. Here, f is the objective function that should be minimized, $c_{numeric}$ and $s_{numeric}$ are the velocity and steepness calculated numerically, using equation (3.5), and c_{data} and s_{data} are the velocity and steepness calculated from data. The goal is thus to find the combination of parameter values D and ρ that gives a value of $c_{numeric}$ and $s_{numeric}$ as close as possible to the values measured from data, yielding an objective function that tends to zero. Although mismatch also can be included in the objective function, the numerical value of the mismatch is heavily dependent of the step size of time t and radius r respectively. To avoid this dependency, mismatch is excluded from the objective function but will be calculated numerically using the result from the PSO, whose value $MM_{numeric}$ will be compared with MM_{data}.

The algorithm is based on the PSO described in Mattias Wahdes book *Biologically inspired optimization methods: an introduction* with slight alterations, especially regarding the evaluation of the solutions [29, p. 123]. A flowchart illustrating the algorithm is shown in figure 3.4 and a more thorough description of the algorithm can be found in Appendix A. With the parameters D and ρ estimated using the PSO algorithm, it would be of interest to investigate how and if the results from the parameter estimation differs from the parameter values calculated analytically, using the system of equations in (3.10).

Three plots created for each patient are meant to answer this question. Firstly, a figure showing the PDE solution of the Fisher equation is presented, where the parameters are either calculated analytically or estimated by the PSO algorithm. Secondly, a plot illustrating the PDE solution of the Fisher equation, using parameter values from the PSO, and the Analytic Equation, using analytically calculated parameter values, is shown. For this case the time is adjusted for the Analytic Equation such that the solutions grow on the same time scale, just as when calculating the sum of squares in section 3.4.

Lastly, a plot visualizing the radii measured from the MR images using T2-FLAIR, together with the simulated radii using different solvers, is shown. The solvers in question are, once again, the PDE solution of the Fisher equation together with the parameter values from the PSO, the Analytic Equation with analytically calculated

parameters, but also a simple linear regression. For the linear regression, the linear radial growth is calculated using equation (3.1) and the time is adjusted according to the equation

$$0 = r_0^i + c^i t_0 \Rightarrow t_0 = \frac{-r_0^i}{c^i} \tag{3.12}$$

to simulate growth from $r = 0$ rather than $r = r_0$ for each patient i. The simulated radii thus becomes $c^i t$ and are measured at the time points $t + t_0$.

For the two first cases however, the simulated radii are found by searching in the numerical solution of $u(r, t)$ in a manner similar to the numerical calculation of the mismatch described in section 3.3. What is different is that only the radius r needs to be estimated for the mismatch, while in this case both r and t need to be extracted from the matrix at separate occasions. Firstly, the numerical radius closest to the radius of the first measurement from data, r_1, is found from the vector of radii and its index in the vector is extracted. Since the number of columns in the matrix is the same as the number of elements in the vector of radii, inserting the index in the matrix gives all values of $u(r_1, t)$ for varying t. From here, the time for the first measurement t_1 is extracted by finding the row index where $u(r_1, t) = 0.16$ and inserting the index in the vector of time points to get t_1. Since t_1 is known and the time *difference* between the measurements is known from data, the time for all measurements can be calculated. With the time known, the corresponding radius at each measurement is extracted as above where the roles are exchanged, that is *find* r_m *such that* $u(r, t_m) = 0.16$ *for all measurements* $m = 2, 3, ...$

For each of the three cases mentioned above, the simulated radii are compared with the radii measured from data and the sum of squares is calculated as a measure of the error, according to the equation

$$\mathrm{SS} = \frac{1}{n} \sum_{k=1}^{n} \left(r_{data}(t_k) - r_{simulated}(t_k) \right)^2 \tag{3.13}$$

where n is the number of data points, $r_{data}(t_k)$ is the radius calculated from patient data at time t_k and $r_{simulated}(t_k)$ is the radius estimated from simulated data, using one of the solvers described above.

This method of extracting radii and time points from the solution of the Fisher equation will also be used to give an estimate of the age of the tumour and the size of the invisible radius and volume respectively. The age of the tumour will basically be calculated as the time at diagnosis, t_1, estimated using the different solvers. In addition, the invisible radius will be calculated as $(r_{0.01} - r_{0.16})$, i.e. the radii at $u = 0.16$ and $u = 0.01$ respectively, at the time for the measurement before operation. Similarly the invisible volume will be calculated using the same radii but as $(\frac{4\pi}{3}(r_{0.01}^3 - r_{0.16}^3)$, where the choice of $u = 0.01$ is arbitrary in both cases.

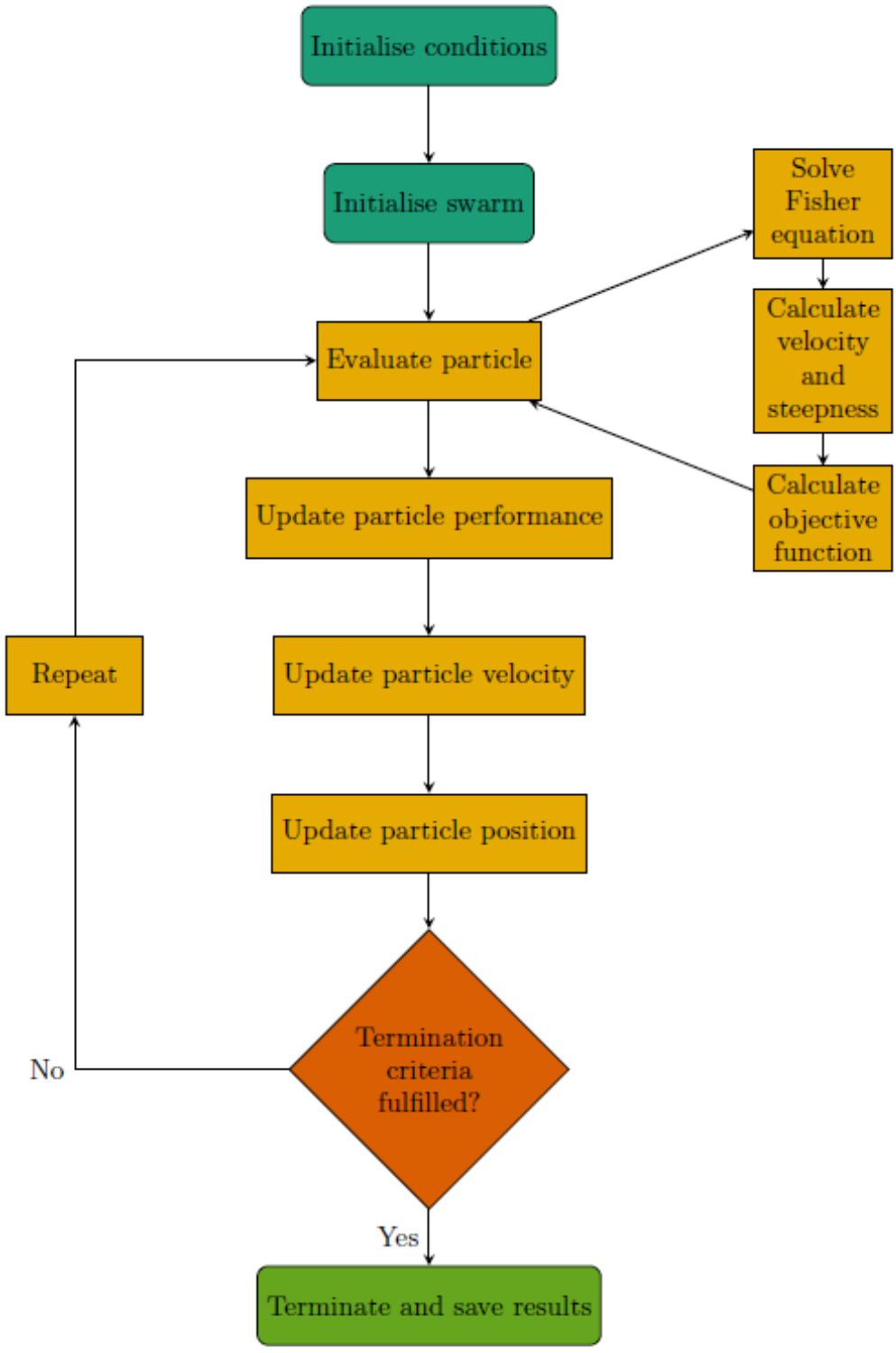


Figure 3.4: *Flowchart illustrating in general how the Particle Swarm Optimisation algorithm works. The colors represents different phases of the code: Initialisation phase in dark green, the iteration process in yellow/orange and the termination step in light green.*

4

Results

In this chapter, the results for each question at issue are presented. In section 4.1 the extraction of properties from patient data and the result from the log-rank tests are presented, while in 4.2 and 4.3 the comparison of solvers and the parameter estimation is introduced. Mainly tables and figures will be illustrated in the sections below, while the interpretation of the results is presented in chapter 5.

4.1 Statistics

4.1.1 Summary of patient data

The patient data set consists of 28 patients in total, all diagnosed with Low Grade Glioma during 2005-2015. Of the 28 patients, 9 have died from cancer, 3 have died of other causes and 16 are still alive after follow-up 2019-01-01. A summary of patient properties is shown in table 4.1 and a plot illustrating the radial growth for all patients is shown in figure 4.1. The problem however with the whole data set is that it shows unreasonable properties, such as an initial radius of zero and a velocity, steepness and mismatch below zero. Fortunately, there exists a reasonable explanation for all cases, which will be discussed in chapter 5.

Due to the unreasonable values, subsets of the data must be defined to only use reasonable data, while keeping as many patients as possible from the already small

Table 4.1: *Summary of tumour properties for all 28 patients. As can be seen, the values for certain properties are rather unreasonable, indicating that not all patients can be used in further studies of the data set.*

Property	Min	Max	Mean	Median
Age at diagnosis $[years]$	21	69	49.32	49.50
Time to operation $[years]$	0.15	8.05	1.52	0.48
Tumour volume at diagnosis $[cm^3]$	0.00	138.97	36.60	19.34
Tumour radius at diagnosis $[mm]$	0.00	32.13	17.79	16.65
Radial velocity $[mm/year]$	-5.51	83.34	5.48	0.77
Steepness $[mm^{-1}]$	-2.21	2.37	0.13	0.13
Mismatch	-0.71	0.87	0.17	0.17

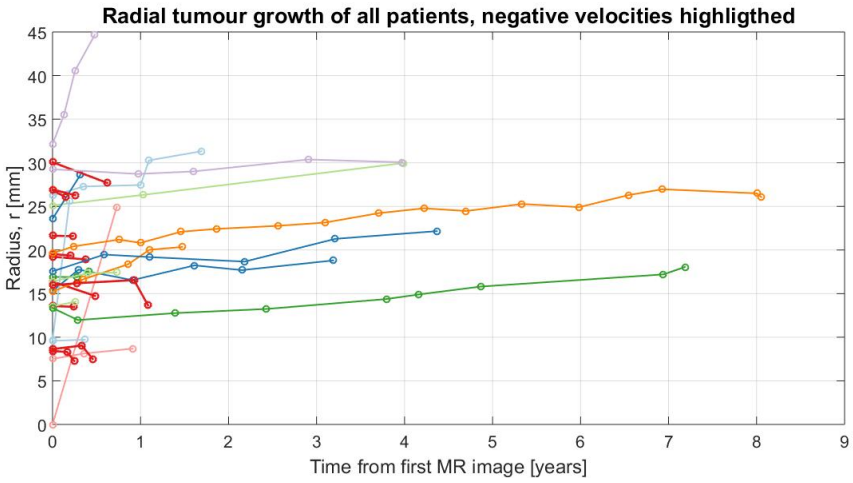


Figure 4.1: *Radial growth over time for all patients. The data points in red highlights the patients that seem to have negative tumour growth.*

data set. The subsets of the data in consideration are

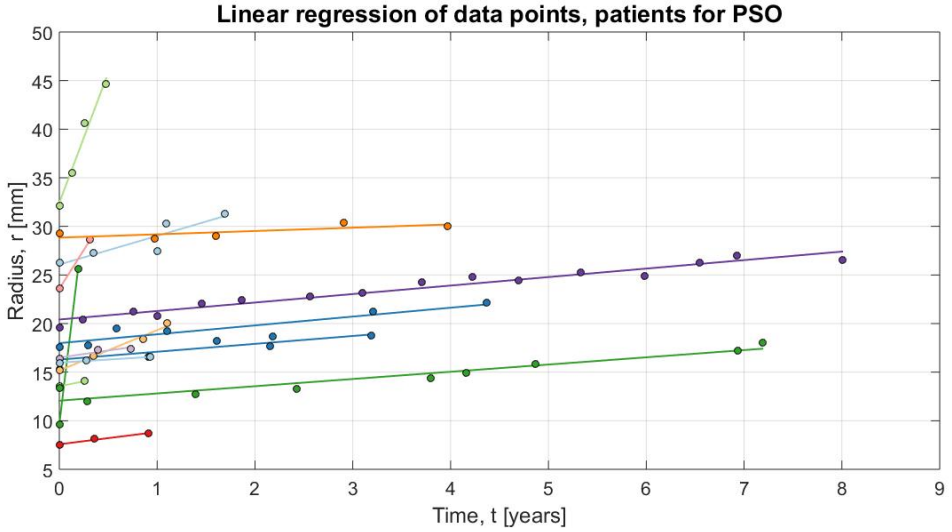
1. All patients (28 patients)
2. Filtered data, positive steepness (S-data, 20 patients)
3. Filtered data, positive velocity (C-data, 17 patients)
4. Filtered data, positive velocity and steepness (SC-data, 14 patients)

where the first subset is presented above, S-data and C-data are used in the survival analysis and the last subset, SC-data, is the data used in the parameter estimation. In addition, the patient with zero initial radius is removed from all subsets. A summary of the patient properties for the SC-data is presented in table 4.2 and two plots illustrating linear growth and time adjusted growth according to equation (3.1) and (3.12) respectively are shown in figure 4.2a and 4.2b. For additional plots, see Appendix B and for the complete list of tumour properties for all patients, see Appendix C.

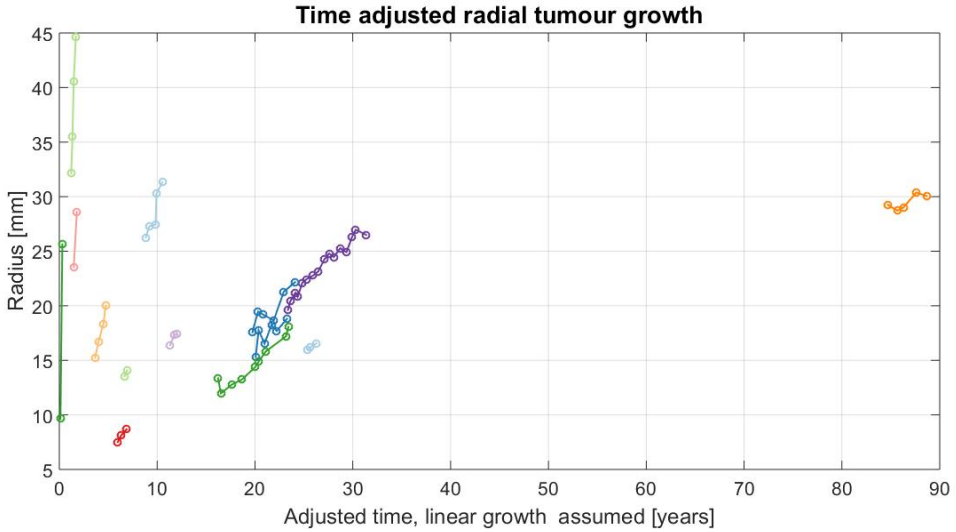
Table 4.2: *Summary of tumour properties for the patients in subset SC-data, showing more reasonable values than for the full patient data set.*

Property	Min	Max	Mean	Median
Age at diagnosis $[years]$	21	69	52.43	56
Time to operation $[years]$	0.19	8.05	2.54	1.47
Tumour volume at diagnosis $[cm^3]$	1.78	138.97	37.13	17.71
Tumour radius at diagnosis $[mm]$	7.52	32.13	20.70	16.17
Radial velocity $[mm/year]$	0.34	83.34	10.16	1.36
Steepness $[mm]$	0.047	2.37	0.63	0.28
Mismatch	0.037	0.87	0.38	0.34

Figure 4.2a illustrates the radial growth for the patients included in the SC-data and how the growth would be estimated assuming linear growth. For most patients the linear regression fits very well and the median R^2 gets a value of 0.91, after excluding the patients with only two measurements since they automatically will get a perfect R^2 of 1. As can be noted, the range of the velocities is very large, which is even more visible in figure 4.2b, where the time is adjusted according equation (3.12). Most of the measurements from the patients are located to the left in the figure except for one patient, Patient 16, that will be described in more detail in section 4.3.1.3.



(a) *Data points and the corresponding linear growth with constant tumour growth velocity c [mm/year].*



(b) *Time adjusted radial growth, adjusted according to equation* (3.12) *and used to get an estimate of the age of the tumour.*

Figure 4.2: *Radial growth for the patients included in the subset SC-data, modified to show certain features of the data.*

4.1.2 Survival analysis

The results of the survival analysis of mismatch, invasiveness and velocity are shown in figure 4.3, 4.4 and 4.5 respectively. The plus signs correspond to censored patients and the discrete steps corresponds to events occurring in each group. The p-value is presented in the lower left corner and calculated using the log-rank test. For the analysis of the mismatch and steepness, S-data containing patients with positive steepness was used while C-data with positive velocities was used for the analysis of velocity. The p-value for mismatch, invasiveness and velocity are 0.13, 0.024 and 0.0022 respectively, implying that the effect of mismatch is not significant, while the effect of invasiveness and velocity are significant.

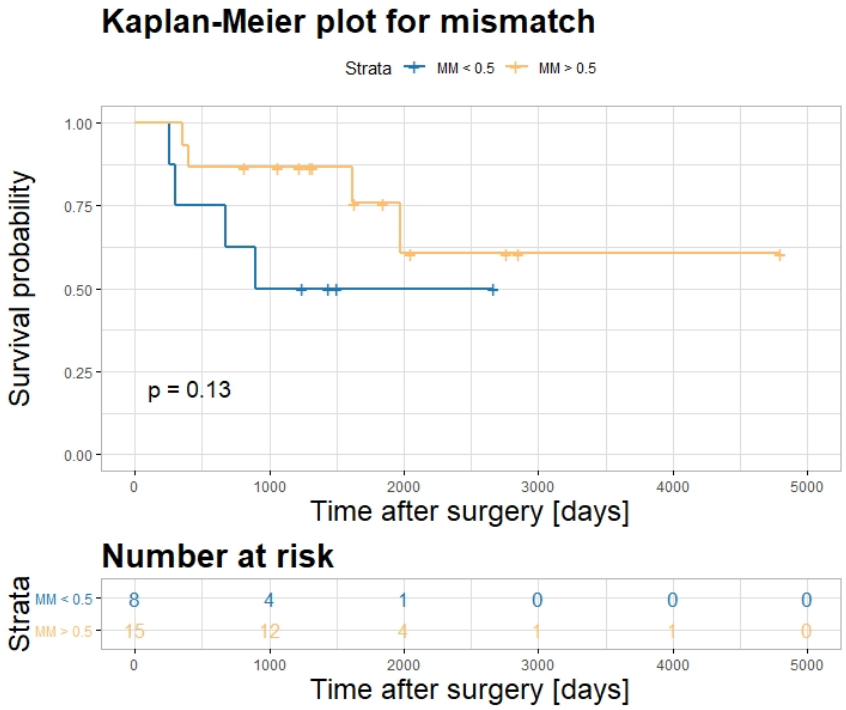


Figure 4.3: *Kaplan-Meier plot visualising the survival of patients with different values of mismatch. The subset of use is the S-data and the p-value indicates that mismatch has no significant effect on survival.*

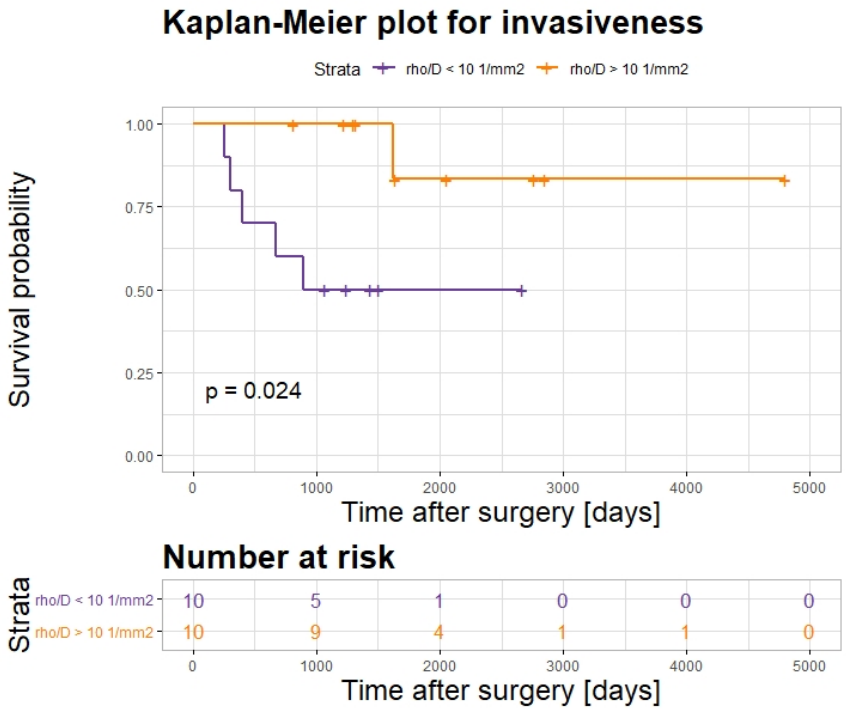


Figure 4.4: *Kaplan-Meier plot visualising the survival of patients with different values of invasiveness. The subset of use is the S-data and the p-value indicates that the invasiveness has a significant effect on survival.*

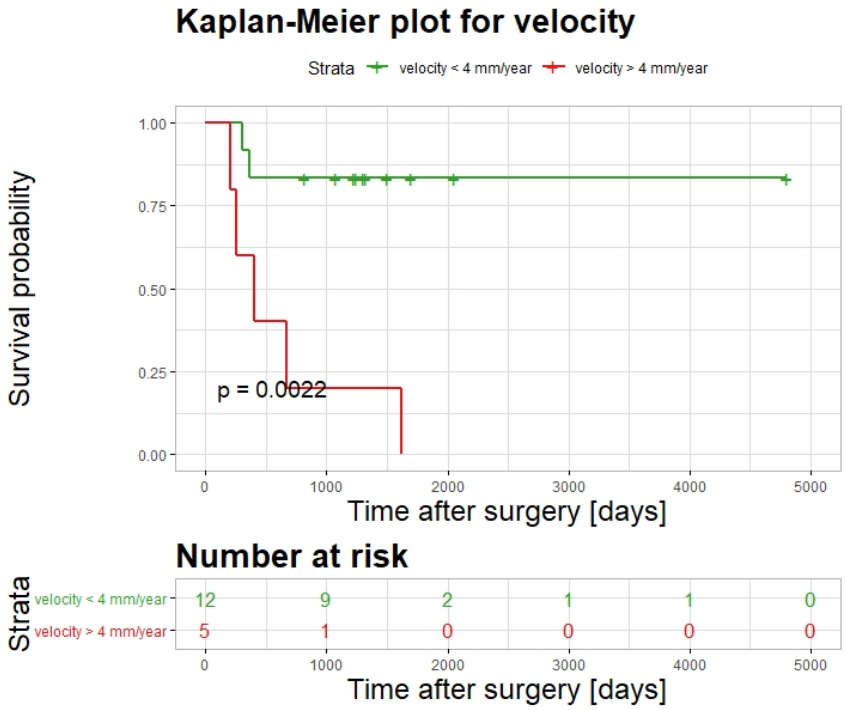


Figure 4.5: *Kaplan-Meier plot visualising the survival of patients with different values of velocity. The subset of use is C-data and the p-value indicates that it has a significant effect on survival.*

4.2 Comparison of methods for solving velocity, steepness and the Fisher equation and their dependence on parameter values

In this section, the results from the different comparisons are presented and visualised as surface and contour plots, using the settings described in section 3.4. For the velocity c and steepness s, their dependence on parameter values are presented as well as the differences d_c and d_s, calculated according to the equations in (3.7), between the analytic and numerical estimates of c and s. For the comparison between solvers of the Fisher equation, the solution to the equation using either (2.3) or (2.7) are compared for one pair of parameter values D and ρ, and the sum of squares calculated according to equation (3.8) is plotted for all combinations of parameter values.

The results from the velocity and steepness comparisons are shown in figure 4.6 and 4.7, which illustrates that the velocity increases for large values of D and ρ and steepness increases for small values of D and large values of ρ. The contour plots visualising d_c and d_s implies that the difference is very small between the two different methods of calculating velocity and steepness, with the exception for small values of ρ. The inverted colours in figure 4.7 relative figure 4.6 is because the velocity calculated numerically is slightly larger than the one calculated analytically, while the numerical steepness is slightly smaller than the analytic steepness, yielding a positive and negative scale respectively.

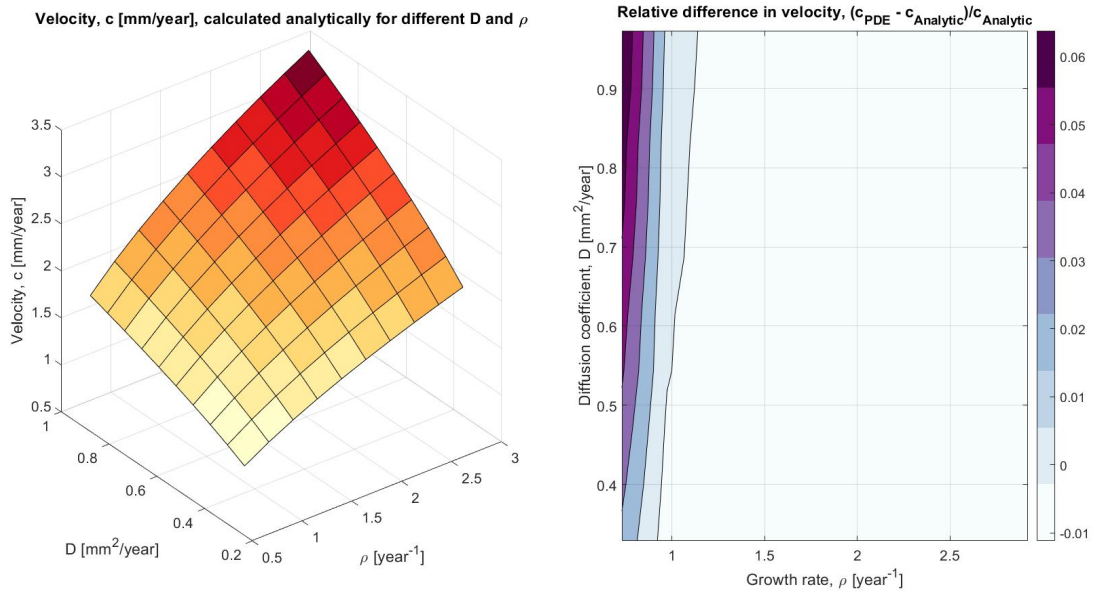


Figure 4.6: *Left: Velocity, calculated numerically, plotted against D and ρ. Right: The difference, relative the analytic solution, between the solvers. The largest difference is shown for small values of ρ and large values of D.*

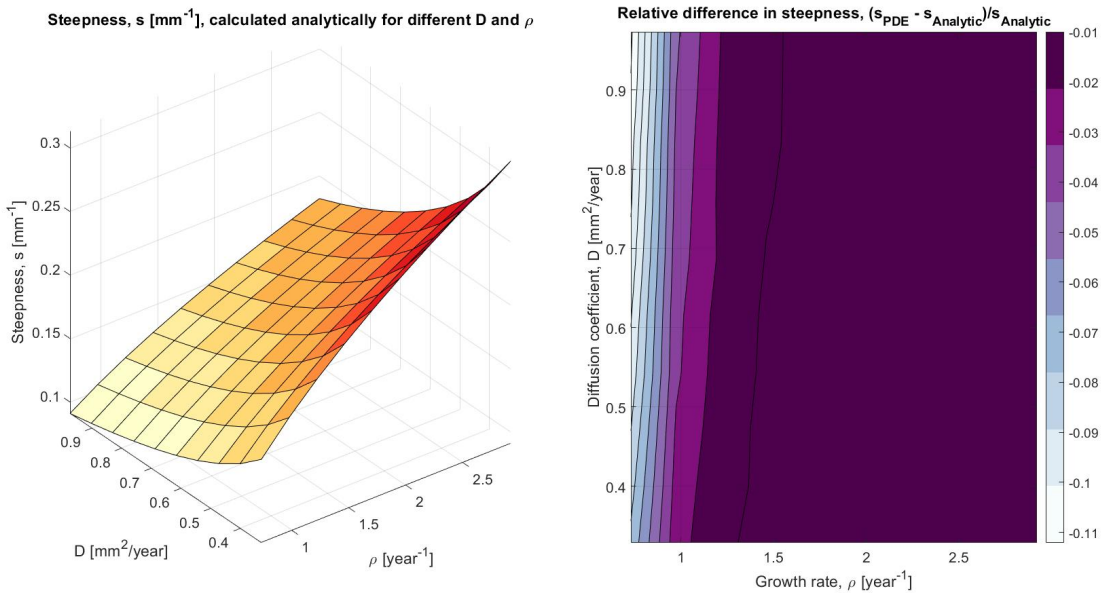


Figure 4.7: *Left: Steepness, calculated numerically, plotted against D and ρ. Right: The difference, relative the analytic solution, between the solvers. The largest difference is shown for small values of ρ and large values of D, although the colors are inverted in comparison with figure 4.6.*

The results from the comparison between solvers of the Fisher equation are shown in figure 4.8 and 4.9 respectively, where figure 4.8 illustrates that the largest difference is in the curvature of the travelling wave. The edges are sharper for the Analytic Equation in comparison with the PDE solution, giving rise to two similar curves, except for this feature. If looking at figure 4.9 it shows that the sum of squares is small but varies with D and ρ, indicating that the parameters have a different effect on the two solutions of the Fisher equation.

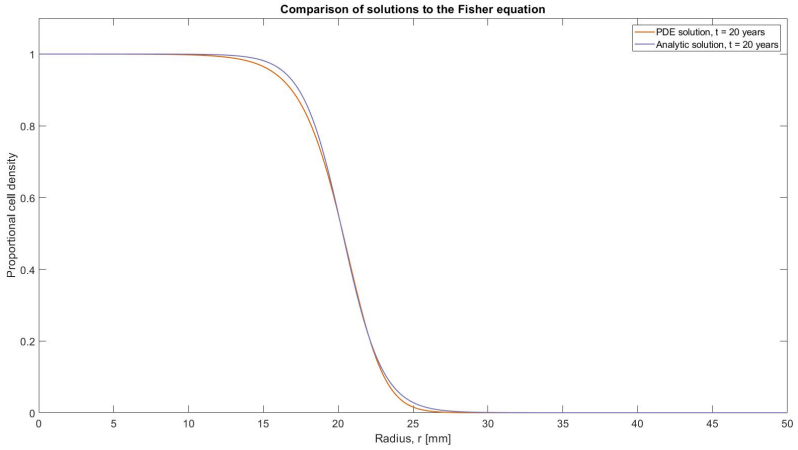


Figure 4.8: *The solution to the Fisher equation using either the PDE solution and Analytic Equation. The largest different is in the curvature of the constant travelling wave.*

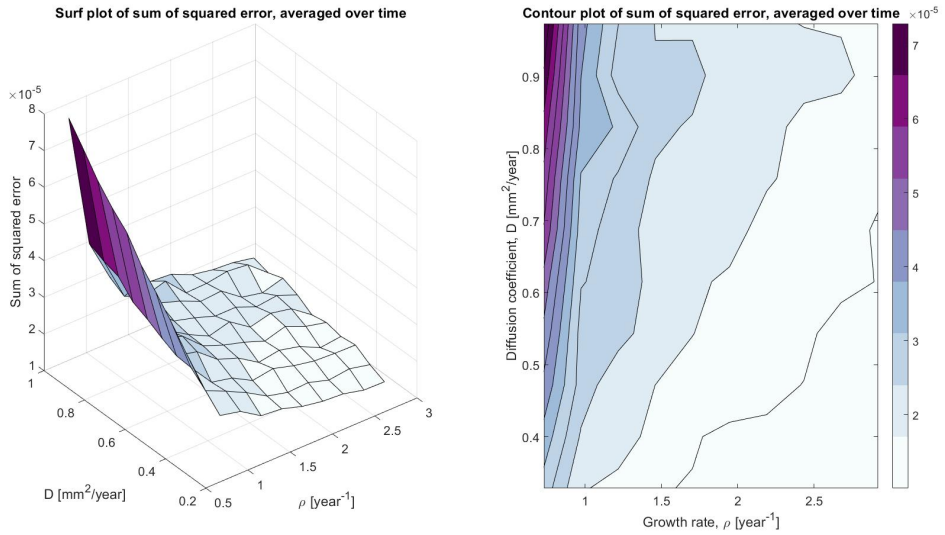


Figure 4.9: *Averaged sum of squares plotted against D and ρ. The largest difference is found for small values of ρ.*

4.3 Parameter estimation

In this section, the results from the parameter estimation using Particle Swarm Optimization (PSO) will be presented, together with a comparison between the numerical and analytic methods of solving the Fisher equation and calculating the parameter values. If the results are similar it would indicate that the analytic equations are good approximations, simplifying both the parameter estimation and computation of the Fisher equation. The result from three patients will be presented in particular and the results for all patients will be presented in general. For the details of the results for the remaining 11 patients, see Appendix D.

Three different combinations of how the Fisher equation is solved and how the parameter values D and ρ have been estimated will be consequently investigated and compared in the sections below. The two first variants are the Fisher equation solved using Spherical coordinates, but with parameters either estimated from the PSO algorithm or calculated according to the analytic equations in (3.10). The third variant is the Fisher equation solved using the Analytic Equation, equation (2.7), together with parameter values calculated according to the analytic equations. To clarify which variant that is currently described in the text below, they will be called PDE_{pso}, PDE_{acp} and AE_{acp} respectively, where AE stands for Analytic Equation of the Fisher equation, while the subscript acp stands for analytically calculated parameters. The subscripts will in addition be used to emphasize which method the parameters were calculated with, for example D_{pso} or D_{acp}.

4.3.1 Three examples from the PSO result

The patients of choice are Patient 1, 15 and 16, and are selected to show the different results can that be retrieved from the parameter estimation. All patients yield great estimations of the velocity and steepness respectively, but behaves differently relative AE_{acp}. Three plots will be presented for each patient. The first one compares PDE_{pso} with PDE_{acp}, while the second plot compares PDE_{pso} with AE_{acp}. The last figure shows how different solvers would simulate patient data, as described in section 3.5, and suggests at which time the measurements were taken relative the initial formation of the tumour. The solvers in question for the last comparison are PDE_{pso}, AE_{acp}, but also a simple linear regression, the same as when calculating the velocity and shifted in time according to equation (3.12), and will be abbreviated LR henceforward. The sum of squares for each case is calculated according to equation (3.13) and is presented above each plot.

4.3.1.1 Patient 1

The summary of the results from Patient 1 can be found in table 4.3, showing great accuracy regarding velocity and steepness, less accurate estimation of the mismatch and slightly different parameter values using the different methods. Since the analytically calculated parameters D_{acp} and ρ_{acp} are estimated using the values of c and s from data, the analytical estimates of c and s would simply be the velocity and steepness from data. Of that reason, c and s for the analytic case are not printed in table 4.3 or in future similar tables.

Table 4.3: *Summary of the velocity c, steepness s, mismatch MM and the parameter values D and ρ for Patient 1 (67 years), calculated either from Data, using Analytic calculations or estimated using PSO.*

	c $[mm/year]$	s $[mm^{-1}]$	MM	D $[mm^2/year]$	ρ $[year^{-1}]$
Data	2.95	0.071	0.636	-	-
Analytic	-	-	-	2.171	1.002
PSO	2.95	0.071	0.683	1.705	1.019

The plots for Patient 1 are shown in figure 4.10 and 4.11. For this patient PDE_{pso}, PDE_a and AE_{acp} differs relatively much from each other, which can be seen in figure 4.10, where the parameters D_{acp} and ρ_{acp} gives a different shape of the curves.

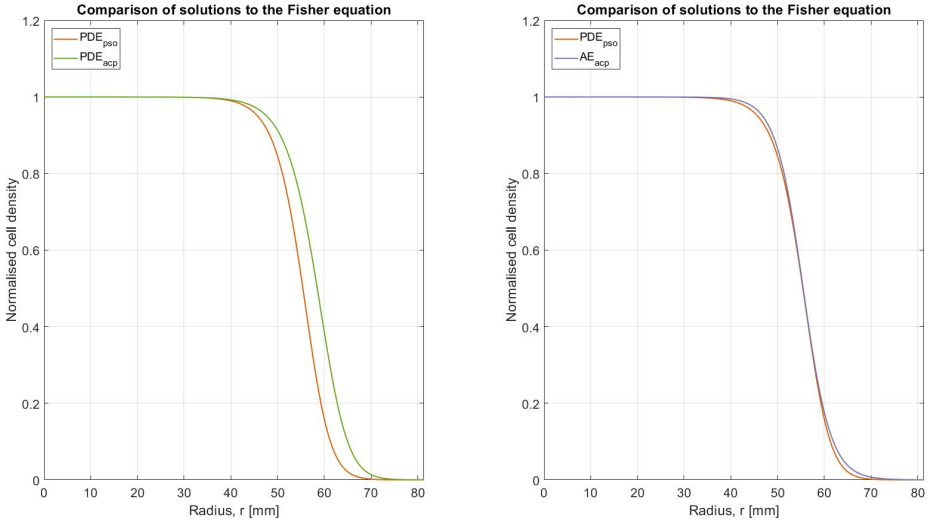


Figure 4.10: *Comparison of solutions of the Fisher equation. Left:* PDE_{pso} *and* PDE_{acp}*. Right:* PDE_{pso} *and* AE_{acp}*.*

For the comparison between patient and simulated data, all solvers perform similarly but PDE_{pso} gives the highest value of the sum of squares. If comparing the time axes, both AE_{acp} and LR suggests that the tumour is not older than maximum 10 years, while PDE_{pso} suggests that the tumour in reality is approximately 20 years old at surgery.

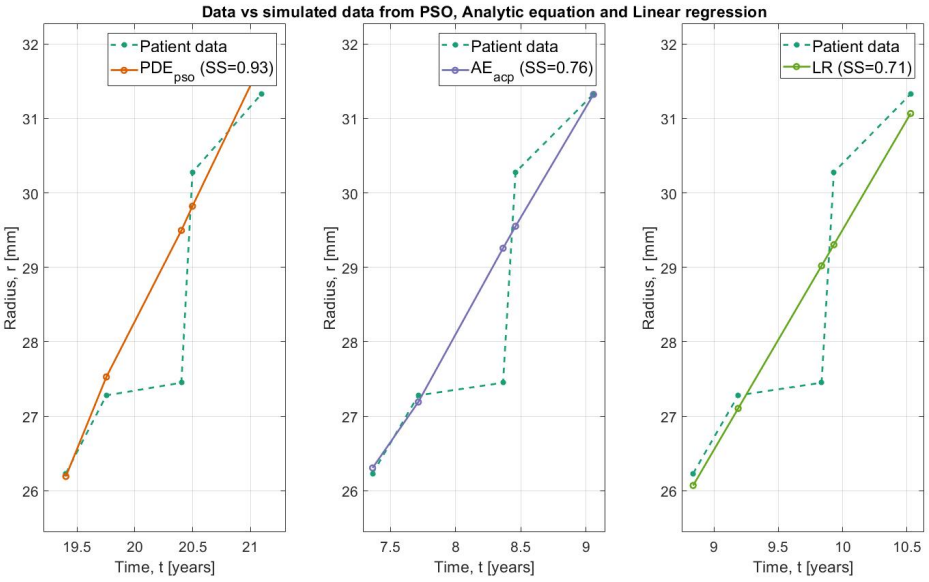


Figure 4.11: *How different solvers would simulate patient data. Left:* PDE_{pso}*. Middle:* AE_{acp}*. Right:* LR*. The sum of squares is noted in the upper right corner for each case and the main difference is the time axis.*

4.3.1.2 Patient 15

The summary of the results from Patient 15 can be found in table 4.4, showing great accuracy regarding velocity, steepness and mismatch and larger similarities between the different methods of estimating D and ρ.

Table 4.4: *Summary of the velocity c, steepness s and the parameter values D and ρ for Patient 15 (25 years), calculated either from Data, using Analytic calculations or estimated using PSO.*

	$c\ [mm/year]$	$s\ [mm^{-1}]$	MM	$D\ [mm^2/year]$	$\rho\ [year^{-1}]$
Data	4.156	0.552	0.161	-	-
Analytic	-	-	-	0.395	10.923
PSO	4.156	0.552	0.160	0.388	11.063

The plots for Patient 15 are shown in figure 4.12 and 4.13 and show much more similarities with AE_{acp} than Patient 1. In the leftmost plot in figure 4.12 both curves are plotted, but the plots representing PDE_{pso} and PDE_{acp} respectively lies on top of each other. A very reasonable result if comparing the values of the parameters in table 4.4, calculated analytically or using the PSO. For the comparison between PDE_{pso} and AE_{acp}, the difference between the curves is smaller, if comparing with figure 4.10.

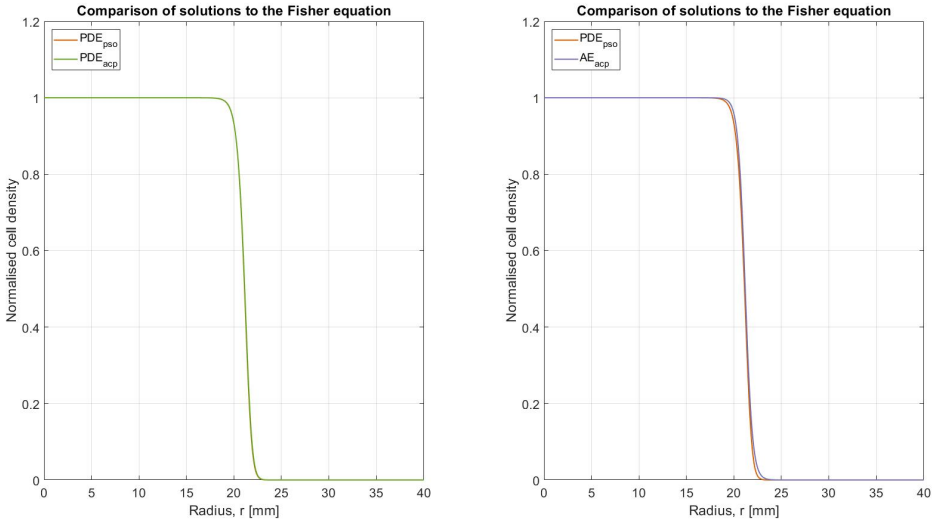


Figure 4.12: *Comparison of solutions of the Fisher equation. Left: PDE_{pso} and PDE_{acp}. Right: PDE_{pso} and AE_{acp}.*

Lastly for figure 4.13, the PDE_{pso} performs better than AE_{acp}, but once again LR yields the smallest value of the sum of squares. Here the time scale is similar for all solvers, in contrast to Patient 1.

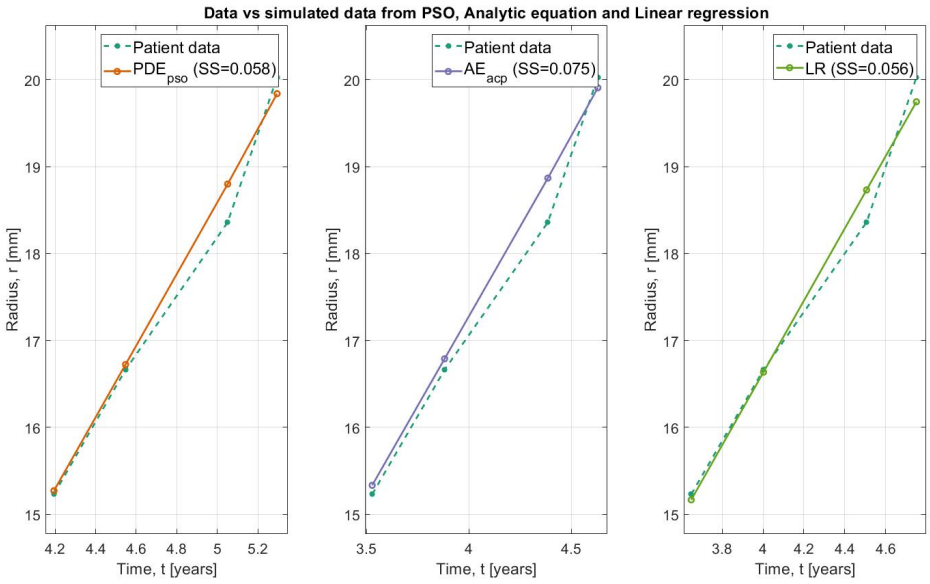


Figure 4.13: *How different solvers would simulate patient data. Left: PDE_{pso}. Middle: AE_{acp}. Right: LR. The sum of squares is noted in the upper right corner for each case.*

4.3.1.3 Patient 16

The summary of the results from Patient 16 can be found in table 4.5, showing great accuracy regarding velocity, steepness and mismatch but large difference between the different methods of estimating D.

Table 4.5: *Summary of the velocity c, steepness s, mismatch MM and the parameter values D and ρ for Patient 16 (69 years), calculated either from Data, using Analytic calculations or estimated using PSO.*

	c $[mm/year]$	s $[mm^{-1}]$	MM	D $[mm^2/year]$	ρ $[year^{-1}]$
Data	0.341	1.70	0.037	-	-
Analytic	-	-	-	0.0105	2.754
PSO	0.341	1.70	0.037	0.0044	2.77

The plots for Patient 16 are shown in figure 4.14 and 4.15. If looking at the leftmost plot in figure 4.14, the solutions are very different due to the large difference in D. However, the curve visualising the PDE_{pso} in the rightmost plot is barely visible, meaning the curves PDE_{pso} and AE_{acp} are very similar despite the difference in value of the diffusion coefficient.

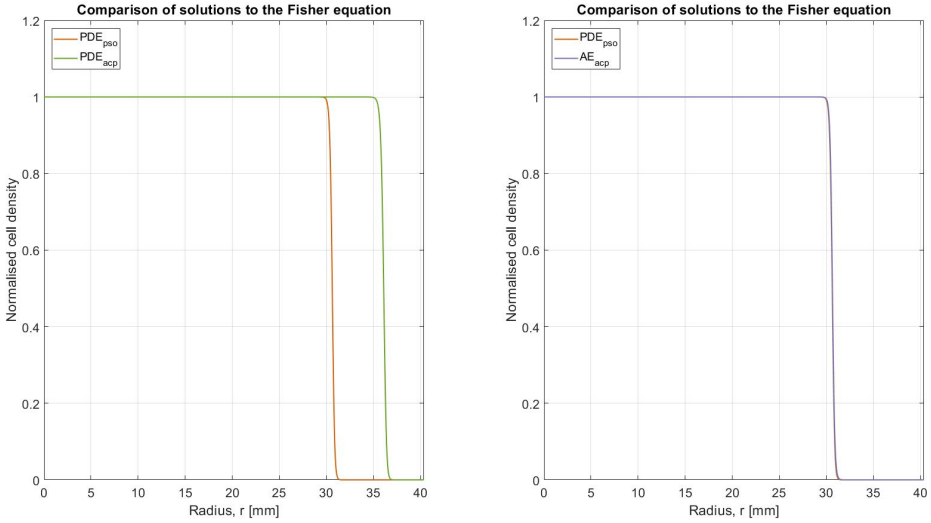


Figure 4.14: *Comparison of solutions of the Fisher equation. Left:* PDE_{pso} *and* PDE_{acp}*. Right:* PDE_{pso} *and* AE_{acp}*.*

In figure 4.15, the way of estimating the time for the PDE_{pso} and AE_{acp} is a little different. For this case, and also Patient 28 in Appendix D, the first measurement is an outlier, leading to an overestimation of the radius if only the first measurement, $r(t_1)$, was used when estimating the time. To correct for this, the mean value of the two first measurements, $(r(t_1) + r(t_2))/2$, was used instead.

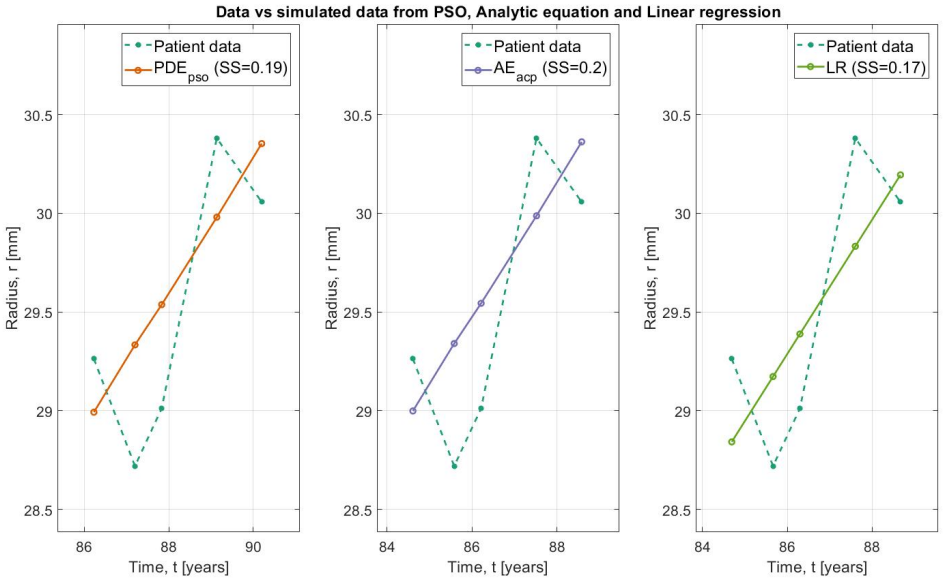


Figure 4.15: *How different solvers would simulate patient data. Left:* PDE_{pso}*. Middle:* AE_{acp}*. Right: LR. The sum of squares is noted in the upper right corner for each case and the main difference is the time axis.*

All solvers give similar sum of squares, where once again the LR is the best. However, if checking the time axis, the estimated time since the creation of the tumour is highly unreasonable. These results will be further discussed in chapter 5, but the reader is reminded that this patient is 69 years old at diagnosis whereas the tumour is estimated to have grown for approximately 85 years before diagnosis.

4.3.2 Summary of result from the parameter estimation for all patients

For this section, values of parameters and properties are summarised and differences between solvers and data are visualised. Starting off with the summary, table 4.6 shows the minimum, maximum, mean and median value of different properties and parameter values. They are divided into sections where the first section is values taken from data, the second section shows the values of D and ρ calculated either analytically or using the PSO algorithm and the third section shows properties calculated from the results from the PSO, hence the sometimes unreasonable values. The invisible radius and volume here are calculated according to section 3.5, thus as $(r_{0.01} - r_{0.16})$ and $(\frac{4\pi}{3}(r_{0.01}^3 - r_{0.16}^3))$ respectively. For the complete table of results from the PSO, see Appendix C.

Table 4.6: *Summary of properties measured from the patient data used in the parameter estimation, estimates of the parameters D and ρ and properties measured using the result from the PSO algorithm.*

Property	Min	Max	Mean	Median
Age at diagnosis $[years]$	21	69	52.429	56
Velocity $[mm/year]$	0.341	83.341	10.157	1.361
Steepness $[mm]$	0.047	2.369	0.629	0.279
$D_{Analytic}$ $[mm^2/year]$	0.0105	86.256	8.903	0.352
D_{PSO} $[mm^2/year]$	$2.337 \cdot 10^{-7}$	67.320	7.167	0.347
$\rho_{Analytic}$ $[year^{-1}]$	0.341	20.131	6.070	4.380
ρ_{PSO} $[year^{-1}]$	0.350	20.148	6.053	4.422
Objective function	$4.395 \cdot 10^{-10}$	0.0321	0.0023	$5.111 \cdot 10^{-9}$
Time before diagnosis $[years]$	1	87	21.429	18
Invisible radius $[mm]$	0.294	13.258	3.717	1.726
Invisible volume $[cm^3]$	0.381	421.824	58.259	8.276

The parameter combinations for all patients, using the parameters estimated from the PSO, are shown in figure 4.16, where the raw values are visualised in the left-most plot and the logarithmic values to the right. The lines visible in figure 4.16 corresponds to how D depends on ρ for constant values of c and s. If considering the analytic calculation of c and s, the logarithmic form of the equations becomes

$$\begin{cases} c = 2\sqrt{D\rho} \Rightarrow \ln c = \ln 2 + \frac{1}{2}\ln D + \frac{1}{2}\ln \rho \\ s = A\sqrt{\frac{\rho}{D}} \Rightarrow \ln s = \ln A + \frac{1}{2}\ln \rho - \frac{1}{2}\ln D \end{cases} \tag{4.1}$$

where $A = \frac{-(0.16-0.80)}{2\left(\ln\left(\frac{1}{0.16}-1\right)-\ln\left(\frac{1}{0.80}-1\right)\right)}$. Solving for $\ln D$, the equations becomes

$$\begin{cases} \ln D = 2\ln c - 2\ln 2 - \ln \rho \\ \ln D = -2\ln s + 2\ln A + \ln \rho \end{cases} \tag{4.2}$$

which is done in figure 4.16 for c equal to 5 and 0.5 $mm/year$ and steepness equal to 0.5 and 0.05 mm^{-1} respectively. In this way it is easier to interpret which parameter combinations correspond to which values of c and s.

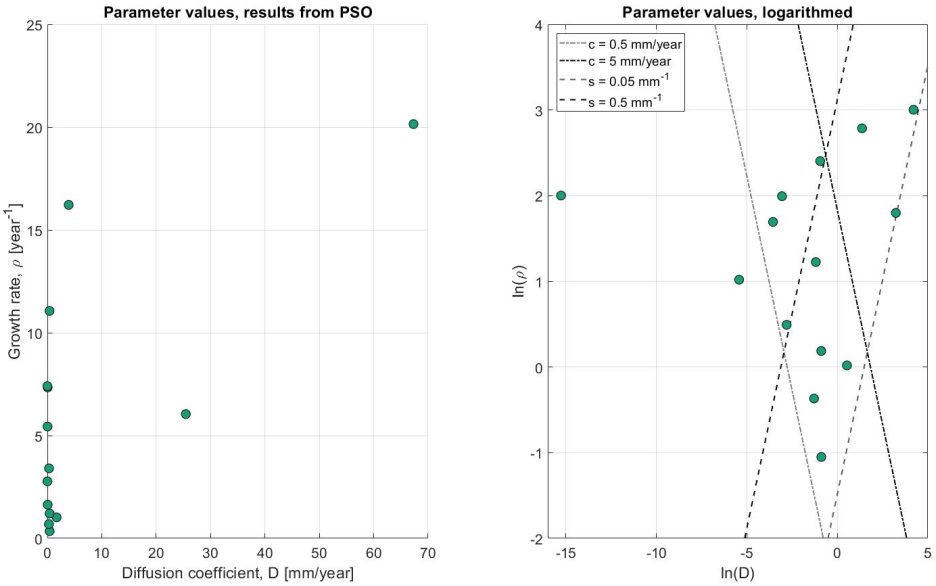


Figure 4.16: *Left: Parameter values for each patient plotted against D and ρ. Right: The logarithm of the parameter values. The lines visualises how* $\ln D$ *varies with respect to* $\ln \rho$ *for c equal to 5 and 0.5 $mm/year$ and steepness equal to* 0.5 *and* 0.05 mm^{-1} *respectively.*

Figure 4.17 shows the difference in parameter values between results from PSO and analytically calculated parameters relative the analytically calculated parameters, i.e. $(D_{pso} - D_{acp})/D_{acp}$ and $(\rho_{pso} - \rho_{acp})/\rho_{acp}$ respectively. As can be seen, the difference in parameter values is larger for the diffusion coefficient D and in addition, D_{pso} is always underestimated in relation to D_{acp} while ρ_{pso} is almost always overestimated in relation to ρ_{acp}.

If instead considering the difference in properties values, that is the velocity, steepness and mismatch, the result is visualised in figure 4.18. Here, the values of the properties estimated from data are compared with the values estimated from the PSO using the same equation for the difference as in figure 4.17.

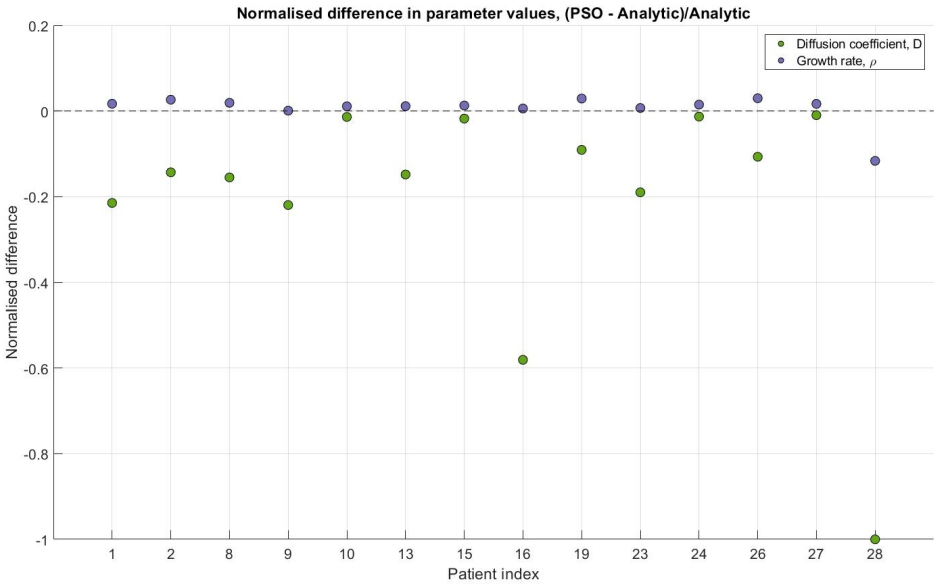


Figure 4.17: *Difference between the PSO result and the analytically calculated parameter values, in relation to the analytic parameter values. They are relatively similar in most cases, with some exceptions.*

In figure 4.18, the spread in values for the mismatch is larger than for the velocity and steepness, while the velocity and steepness are in general quite accurately estimated. Especially the velocity, if comparing the axes.

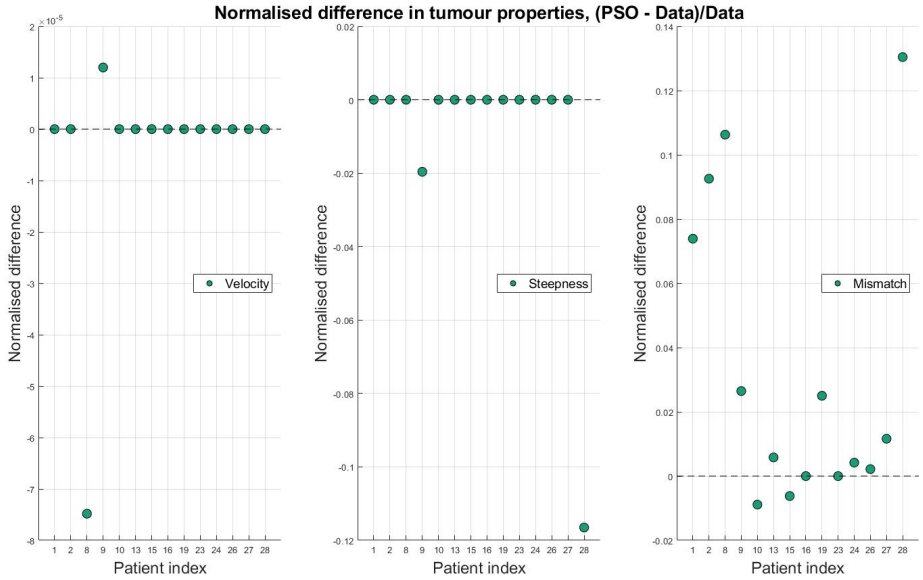


Figure 4.18: *Difference of the tumour properties measured from data and the properties estimated from the PSO algorithm, relative the values calculated from data. The difference is smallest for the velocity and largest for the mismatch.*

Lastly, the invisible tumour radius and volume are plotted against the velocity and steepness respectively and are visualised in figure 4.19. No trend can be seen for the velocity, but it seems that both the invisible radius and volume decreases for increasing steepness.

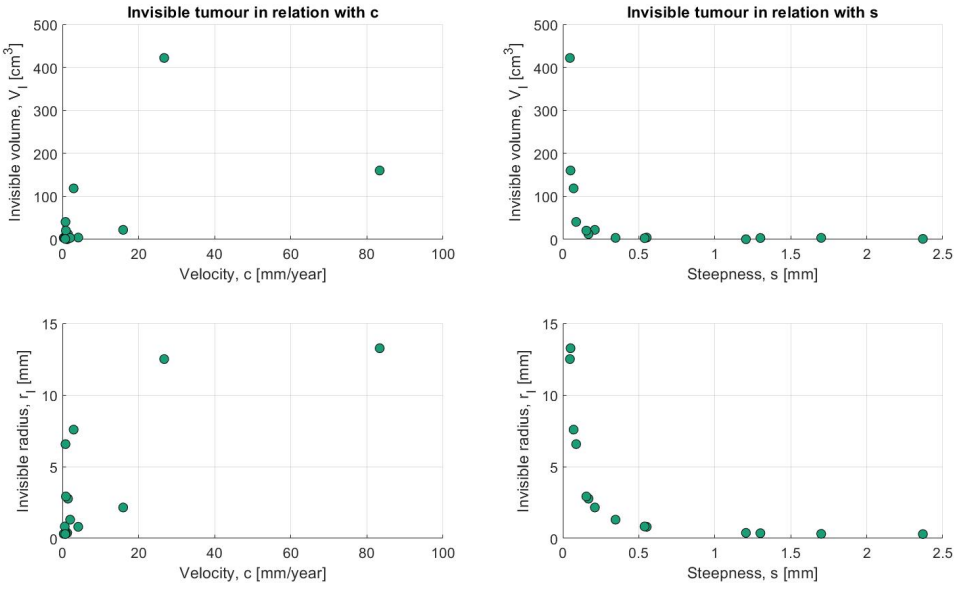


Figure 4.19: *How the invisible volume and radius depend on velocity c and steepness s.*

5

Discussion

In this chapter, the main results of the project will be summarised and discussed. The aim was to simulate tumour growth to give additional information about Low Grade Glioma that can not be estimated directly from Magnetic Resonance images, but requires mathematical modelling. To fulfill the aim, the tumour properties were extracted from patient data and statistically tested to investigate their potential effect on survival. Also, the different methods of solving the Fisher equation were compared and, using patient data, the values of diffusion coefficient D and growth rate ρ for each patient were determined using parameter estimation techniques. The result for each goal is evaluated in the sections below together with a discussion concerning if the aim has been reached.

5.1 Statistics

In this section, the results from section 4.1 are discussed. The questions to answer here are *What are the values of the tumour properties for the data set* and *Does the different properties have an effect on survival*?

If considering table 4.1 and figure 4.1, some values were rather unreasonable. For one patient the initial tumour radius was zero and for some patients the velocity, steepness and mismatch were negative. These observations should not be possible, since the patients were untreated during the measurements and MR images taken with T2-FLAIR should detect cancer cell densities at higher rate in comparison with T1. Fortunately these observations are reasonable, since they are due to an inconsistency and error when estimating tumour volumes from images with T2-FLAIR and T1 respectively.

Firstly, the patient with zero initial volume continuously took MR images before she was diagnosed with Low Grade Glioma for another disease she had, explaining that observation. Secondly, the negative velocities are due to the combination of few measurements and a change in segmentation depth when estimating the tumour volume. As mentioned in section 3.1, the tumour volumes are estimated from MR images by segmenting 3D images to 2D images with a constant segmentation depth d and the total volume is calculated as the sum of the segment volumes. According to neurosurgeon Asgeir Jakola, the segmentation depth used when estimating the volume is most often 5mm, but for some patients the MR image taken before surgery

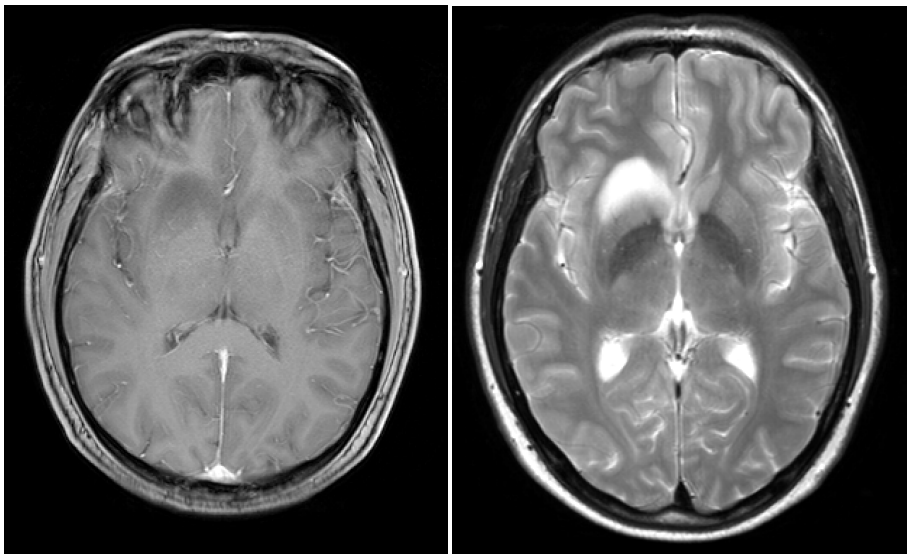


Figure 5.1: *MR image of a Low Grade Glioma using acquisition sequence T1 and T2-FLAIR respectively. The tumour is located at the exact same place in both images but in the T1 image it is difficult to determine what is tumour and not, even for a doctor.*

was made using a depth of 1mm, to get a more accurate estimate of the tumour volume. Thus, if a patient only has two measurements where the first tumour volume was estimated using $d = 5mm$ and the second $d = 1mm$, it looks like the volumes *decreases in size* and consequently yields a negative velocity.

Lastly, the negative steepness and mismatch is due to the difficulty of defining what is tumour and what is surrounding tissue in MR images taken with acquisition sequence T1. One only needs to look at figure 5.1 to understand that estimating the volume is not an easy task, even for a doctor. Due to this difficulty, healthy surrounding tissue has been included in the volume estimate, yielding a volume estimated from MR images using T1 that is larger than the estimated volume from T2-FLAIR, giving rise to negative steepness and mismatch.

To make the data set consistent, all measurements with $d = 1mm$ and all patients with larger T1 volume than T2-FLAIR volume were removed, as well as the patient with zero initial radius, leading to the subsets of data mentioned in section 4.1.1. After filtering, the data was more reasonable but showed large variety for all measured properties, see table 4.2. For this reason, the median is a better measure of the overall values for the data set, since it is more robust to outliers than the mean value. The median age is higher than the previously reported mean age of 41 [3] and the median radial velocity is slightly smaller than the reported mean velocity of $2mm/year$, but could be due to the small data set of 14 patients.

The small data set has in turn affected the results from the survival analysis as well, summarised in section 4.1.2. As can be seen in figure 4.3, the mismatch is insignificant while the invasiveness and velocity are significant, which is illustrated in figure 4.4 and 4.5 respectively. In this context however, significance is a strong word. Due to the small data set the statistical power of the test is very low, which means

that the probability that the significant effects truly are significant is very low. In addition to the log-rank test, the original idea was to also use the Cox Proportional Hazards model to verify the findings in the log-rank test and investigate potential combination effects of the properties [23]. The problem however was that the data set was so small that the test yielded contradictory p-values and the test was thus ignored.

Although the statistical power is small, the log-rank test can still give an indication of which factors that seem to have an effect on survival and not. The test for the velocity indicates that the velocity has an effect on survival, which is in concordance with the conclusion made by Johan Pallud et.al. [20]. The invasiveness is indicated to also have an effect on survival and that patients with diffuse tumours ($\rho/D < 10mm^{-2}$) has a higher risk of dying from their cancer. This is a reasonable result since diffuse tumours tend to infiltrate surrounding healthy brain tissue and becomes harder to detect on an MR image. However, these findings are *not* in concordance with the paper written by Aymeric Amelot et.al. in 2017, where the conclusion instead is that the invasiveness does not have an effect on survival [21]. It should be noted however that in his study, glioblastoma multiforme was of investigation and not Low Grade Glioma. The invasiveness could thus have an effect on survival for patients with Low Grade Glioma, which is especially interesting since the steepness can be calculated after only one measurement. Then the doctors can take faster decisions regarding treatment in contrast to the velocity, since a time series of volumes is needed to accurately measure the velocity. It would thus be of interest to investigate the effect of steepness, as well as mismatch although it was not significant, using a larger data set.

In conclusion, the difficulties in retrieving a large data set with non contradictory data has affected the result. The summary of the data shows large variety between the patients and median values that are close but not in correspondence with previously measured values. Although the power of the log-rank tests is small, the velocity should be truly significant based on previous results, while it would be interesting to test the significance of invasiveness and mismatch on a larger data set.

5.2 Comparison of methods for solving velocity, steepness and the Fisher equation and their dependence on parameter values

Let us now focus on the deterministic modelling, where the results from section 4.2 are discussed. The question of investigation is *does the different methods of solving the Fisher equation, as well as the velocity and steepness, differs*?

For the comparison of the different methods of calculating velocity c and s, illustrated in figure 4.6 and 4.7, the result shows that the different methods give reasonable and very similar results. The dependence of the parameters D and ρ is as excepted for both c and s, where the highest velocity c is found for large D and

ρ while the largest steepness s is found for large D and small ρ. Considering the biology behind the model, it is very reasonable that a tumour with rapidly dividing and moving cancer cells grows faster in size. It is also reasonable that the cancer cell distribution in space is small if the cancer cells does not migrate, thus giving a high steepness.

In addition, the contour plots in figure 4.6 and 4.7 show that the difference between the two ways of solving c and s is small. The largest difference is found for small ρ in both cases with a maximum difference of 0.06 for the velocity and -0.11 for the steepness. This indicates that the dependency of the parameters D and ρ differs between the methods for small ρ but is in all other cases quite similar.

Two things to take into consideration regarding the conclusions above however are the ranges of choice and the method of numerically calculating the properties. Firstly, as mentioned in section 3.4, the choice of the ranges for D and ρ were based on previously reported parameter values, but also because all combinations reached the constant travelling wave within the intervals $t \in [0, 30]$ $years$ and $r \in [0, 100]$ mm. As shown in section 4.3 where the results from the parameter estimation are presented, the variation in parameter values is larger than the range used in this test. Therefore, it is a possibility that the methods of calculating c and s differ for parameter values that are not included in the chosen range. Also, the solutions to the Fisher equation for small ρ takes longer time to reach the constant travelling wave. If the phase where the solution has reached the constant travelling wave is shorter, there is a risk that the velocity and steepness for the constant phase is not as accurately calculated as for those solutions where the phase is long. A restriction to only include the constant wave phase has been implemented in the functions, but may not have been as strict as it should have been.

A similar error could have occurred for the comparison between the solutions of the Fisher equation, either using the PDE solution or the Analytic Equation. As can be seen in figure 4.9, the sum of squares for all combinations of D and ρ is the largest once again for small values of ρ. It could be due to the same reasons mentioned above, or basically because there exists a larger difference for these combinations of parameter values. Although the difference is at largest for small ρ, overall the difference between the solvers is very small implying that the curves are similar. The difference is in the curvature of the solutions to the Fisher equation, as can be seen in figure 4.8. For the PDE solution of the Fisher equation, the edges of the travelling wave is smoother than for the Analytic Equation. Although the curves for only one set of parameter values are illustrated in figure 4.8, other parameter combinations have been tested as well yielding the same behaviour, i.e. the edges of the PDE solution are always behind the Analytic Equation. If looking very closely at 4.8, one can see that the difference in curvature goes within the interval $u \in [0.16, 0.80]$ and could be the reason of why the difference when estimating the steepness is larger than when estimating the velocity.

To conclude this section, the different methods of calculating the velocity, steepness and solving the Fisher equation are similar. There exists a small difference between them that seems to be largest for small values of the growth rate ρ and varying values of diffusion coefficient D, but could potentially be due to numerical errors. From a

mathematical point of view these small differences could be interesting, but if the goal is to use the Fisher equation as a tool in medicine these small differences would be negligible. For that reason, these findings indicate that the Analytic Equation of the Fisher equation, together with the analytical estimates of velocity and steepness, suffices if it would be used in medicine.

5.3 Parameter estimation

Lastly, the result from the parameter estimation from section 4.3 is discussed, where the question in consideration is *can the parameters D and ρ be determined from patient data using parameter estimation*?

As all three examples shows, see table 4.3, 4.4 and 4.5 respectively as well as figure 4.18, the numerically calculated velocities and steepnesses are very accurately estimated for the optimal values of D and ρ. In figure 4.18 where the values of the properties estimated from data are compared with the values estimated from the PSO, it can be seen that for only some exceptions the values differ for the velocity and steepness. The exceptions are Patient 8, 9 and 28 respectively which are explained by their abnormally large values of velocity and steepness. For patient 8 and 9 the tumour velocities are approximately 26 and 83 $mm/year$, where the median velocity is 1.36 $mm/year$, and Patient 28 has the largest recorded value of steepness of 2.4, see Appendix C for additional information. These findings indicate that the parameter estimation performs worse for extreme values of velocity and steepness.

For the mismatch illustrated in figure 4.18 the spread is much larger, which is reasonable since the algorithm has no information regarding the mismatch. Although it was not used in the optimization process, the difference relative data is not that large. The largest differences in mismatch is found for patient 1, 2, 8 and 28, where the reason for the first three could be that the steepness for these patients is very small. For Patient 28, the steepness is too large for the PDE to estimate, giving rise to a result that does not capture the data as well as for the other patients.

These results indicate that the Particle Swarm Optimization algorithm was a good choice of algorithm, although it has its drawbacks. Firstly, the algorithm terminates after a finite set of iterations even if the termination criteria has not been reached. In this case however, the total number of iterations were only reached for 3 of 14 patients, since the remaining 11 patients reached a value of the objective function below 10^{-8}, which was used as a termination criterion. For the 3 patients where 10^{-8} was not reached, Patient 1, 9 and 28 respectively, one had major numerical issues and another had a steepness so large it could not be estimated from the Fisher equation, see Appendix C and D for additional information. Why Patient 1 gets a relatively large value of $1.19 \cdot 10^{-6}$ is because the total number of iterations were smaller when the parameters for this patient were estimated. The number of iterations was changed for the rest of the patients and the optimisation process was not repeated for Patient 1 using the new criteria. This inconsistency yielded a larger objective function value for Patient 1, but when this error was noted it was too late to do anything about it.

A further drawback of the PSO algorithm is that one gets no information regarding the robustness or sensitivity of the solution found. Also, since it is a stochastic algorithm, there is no guarantee that it exactly finds the same optimum every time. A test was made were the PSO solves the same minimization problem ten times using the same settings to investigate if the same solution is found every time, which it does up to at least 3 decimals accuracy. For a more detailed description, see Appendix A. The result of this test indicates that the parameter estimation is not affected by this stochastic feature of the algorithm, and since the PSO is shown to be an accurate algorithm [29, p. 145], it is a good choice of algorithm at least among the stochastic optimization algorithms. But, if there exists an algorithm that can perform as well as the PSO and give more additional information regarding the performance, it would be of interest to use that algorithm for this purpose.

One thing that could have been done differently for the PSO is the inclusion of the constraints. As mentioned in Appendix A, where the Particle Swarm Optimization algorithm is described in more detail, the particles that minimises the objective function steers the swarm while the particles violating the constraints automatically gets an unreasonably high value of the objective function. An option, described in Mattias Wahdes book, could be to instead use penalty methods to restrict the particles [29, p. 143], yielding a similar result as the one implemented here but using a technique with a more thorough mathematical background.

Since the PSO algorithm seems to have accurately and reliably estimated the parameters D and ρ, even though improvements could be made, it is to note that the results differ from the analytically calculated parameter values. If considering figure 4.17 the difference in parameter values can differ considerably for some patients. Why figure 4.17 contradicts the conclusions made in section 5.2 could be due to the range of parameter values, since the range here is much larger than the one used during the comparisons. In addition, the same methods of calculating the velocity and steepness are used in the parameter estimation as in the comparisons, meaning that the potential error for small values of ρ can have affected these results as well.

In summary, it seems that the parameter values of diffusion coefficient D and growth rate ρ can be estimated from patient data. The Particle Swarm Optimization algorithm gives reliable results with large accuracy, despite the improvements of the algorithm that could be done, and a set of parameters has been estimated for all 14 patients. PDE_{pso} and AE_{acp} can either be different for different parameter values, similar for similar parameter values, or similar *for different parameter values*, as described in section 4.3.1 for the three patients of choice. For most patients the second case is the most common, thus PDE_{pso} and AE_{acp} are similar for similar parameter values. These findings implies once again that the analytic equations, both regarding the solution to the Fisher equation *and* the estimation of the parameter values, are accurate approximations. The exceptions are, as previously noted, when the values of the velocity and especially steepness are rather extreme.

5.4 Has the aim been fulfilled?

So far only the results of the goals have been discussed, but have the goals fulfilled the overall aim to give additional information of the tumour to improve the treatment of Low Grade Glioma? Unfortunately, it does not seem like it.

As for the survival analysis discussed in section 5.1, the only thing that is certain is that the velocity has a significant effect on overall survival of the patients, since it has been shown to be significant in previous papers [20]. The invasiveness could potentially have an effect on survival and would be interesting to test again on a larger data set. However, the method of calculating steepness is rather unreliable since it is dependent of the visibility thresholds. The thresholds used throughout this thesis are 0.80 for the Magnetic Resonance image with acquisition sequence T1 and 0.16 for the MR image with T2-FLAIR, suggested by Kristin Swanson in 2008 [1]. However, an article from 2013 by Chloé Gerin concludes that it is uncertain if these constant thresholds even exist [7]. This not only affects the steepness but all measures that are influenced by these thresholds, such as the numerically calculated mismatch and the invisible volume.

If recalling the results mentioned in section 4.3.2, the invisible volume was the difference in volume calculated using the radii measured at $u = 0.16$ and $u = 0.01$ respectively for PDE_{pso}. The invisible volume was meant to be used as a tool for the doctors at Sahlgrenska to better estimate the amount of remaining cancer cells after surgery, but with the arguments regarding the visibility thresholds above these retrieved numbers are probably not that reliable. It could still give an indication of how large the invisible volume is, which might could help the doctors when taking a decision regarding the marginals of choice when performing surgery or tuning the radiotherapy. Also, since a visible trend is seen between invisible volume and steepness in figure 4.19, although maybe not accurately calculated, one can after only one measurement determine if the invisible volume possibly is large or small.

Secondly, one should really question the estimated time made from the results of the PSO algorithm. Although it is unknown which of the methods to estimate the age of the tumour, PDE_{pso}, AE_{acp} or linear regression LR, is the most accurate, the PDE_{pso} for some patients gives unreasonably high age of the tumours. One example is Patient 2, see Appendix D, where the estimated age of the tumour is approximately 15-20 years for AE_{acp}, 20-25 years for LR and 50-55 years for PDE_{pso}. A questionable time considering that the patient age of diagnosis was 63, thus the result from the PSO implies that this patient got the tumour 13 years of age.

This result puts into question if the Fisher equation really is a good model after all. As all models, the Fisher equation lacks a lot of details concerning Low Grade Glioma. At the same time that is exactly what models are made for, to only consider the central concept of a process in order to return an accurate approximation of reality. The problem here therefore could be that an important factor regarding tumour growth is not included in this variant of the Fisher equation. Such a central concept could the anisotropy, the non-homogeneous spatial distribution of the tumour. In this model isotropy is assumed, resulting in a PDE with spherical

symmetry and a diffusion coefficient independent of space. The MR images shown throughout this thesis however, for example figure 5.1, are clearly not spherical or independent of its surroundings. There already exists at least one model taking into account anisotropy, but this is used to simulate a glioma of higher grade [10]. It could thus be of interest to create a similar model to instead simulate growth of Low Grade Glioma, to investigate if it improves the result.

Although the Fisher equation could use some improvements, one can not exclude the effect of deficient data that can make all models fail. The most obvious case is Patient 16, the 69 year old patient with an approximately 85 year old tumour, see figure 4.15. A possible explanation could be that the diffusion coefficient D and growth rate ρ are not constant over time, yielding a higher growth velocity at the initial phase of the tumour. How to investigate these changes in values over time is uncertain, since for a patient the tumour has already grown for some time when the patient is diagnosed.

Another explanation could also simply be that errors have occurred when estimating tumour volumes from MR images, which would not be the first case in this thesis. It is very difficult and time consuming to estimate tumour volumes, especially when using acquisition sequence T1. As mentioned in section 4.1.1, data from 14 patients was discarded in order to get a subset of the data with reasonable values and even that subset was not flawless. This approach of estimating specific values of D and ρ for each patient requires accurate and many volume estimates *for all patients*, which certainly is not the case. As already mentioned, 14 patients were discarded due to contradictory data and if that data could have been used more efficiently, or if data from other patients could give information about the patient set as a whole, that would be of large interest.

One approach to include the whole data set in one model could be to used Non-Linear Mixed Effect modelling (NMLE). These kind of models includes both fixed effects *and* random effects of a model, giving rise to a deterministic model that can include the effect of variability in the data [30].

6

Conclusion

After filtering the original patient data set, the extraction of tumour properties from data showed reasonable behaviour for most patients with large variety in values. The survival analysis shows that high velocity and diffuse tumours have a negative effect on survival, while mismatch does not have an effect. However, due to the small data set the power of the test is small, questioning the reliability of the significance of the tests. It would be interesting to investigate the effect of invasiveness and mismatch on a larger data set, whose effect is currently unknown.

The comparison between the different methods of calculating velocity and steepness, as well as solving the Fisher equation, shows that in most cases the results are similar in the range of parameter values that was tested. The largest difference in all cases is found for small values of ρ but could be due to an error when using numerical methods. Although safety measures have been taken to avoid this kind of error it would be of interest to check again if the calculations are accurate for all combinations of parameter values. The difference between the solutions of the Fisher equation are due to the curvature of the curves, the PDE solution having more smooth edges than the Analytic Equation, and the sum of squares is small but not constant for all combinations of parameter values. These findings indicates that the different solutions behave differently dependent of the parameters, but that in most cases the curves are similar.

The parameter estimation using the Particle Swarm Optimization algorithm yields accurate numerical estimates of velocity and steepness and in general objective function values close to zero. The exceptions where the PSO has failed to capture the properties measured from data has been due to extreme values of the properties rather than an error in the optimization process and the optima are not affected by the stochasticity in the algorithm. However, improvements of the algorithm could be made and it would be of interest to use another algorithm for the same purpose that could give more information regarding the robustness and sensitivity of the result. The parameter values estimated from the PSO algorithm differs from the analytically calculated parameters, concluding once again that the analytic calculations and the PDE differs. This behaviour is most often seen for extreme values of velocity and steepness, but in general PDE_{pso} and AE_{acp} yields similar results, concluding that the analytic approximations gives accurate results.

Although all goals have been reached, no new insights have been discovered concerning Low Grade Glioma and its behaviour. That the invasiveness was shown to

have an effect on survival is promising but should be investigated on a larger data set. The measure of invisible tumour from PDE_{pso} is questionable, especially since the assumed visibility thresholds for the MR images are probably not accurate, but could be used as an indicator for the doctors to determine if a larger marginal or stronger radiotherapy should be applied during treatment. The combination of the measure of the invisible volume, the invasiveness truly having a significant effect on survival and the visible relation between invisible volume and steepness in figure could thus be interesting to further look into.

Lastly the model of choice, the PDE solution of the Fisher equation in Spherical coordinates, should be questioned especially regarding the estimate of the age of the tumour. A possible expansion to the model could be to include the effect of dependence of its surrounding, but does not takes into consideration the sometimes unreliable data or small number of measurements. For this reason a model based on Non-Linear Mixed Effect modelling could be of interest, that can better use the information from data and include the random effects of the process.

Bibliography

[1] Swanson K, Rostomily R, and Alvord E. A mathematical modelling tool for predicting survival of individual patients following resection of glioblastoma: a proof of principle. *British Journal of Cancer*, 98(1):113–119, 2008.

[2] Corell A., Carstam L., Smits A., Henriksson R., and Jakola A. Age and surgical outcome of low-grade glioma in sweden. *Acta Neurologica Scandinavica*, 138(4):359–368, 2018.

[3] Claus E, Walsh K, Wiencke J, Molinaro A, Wiemels J, and Schildkraut J et.al. Survival and low grade glioma: the emergence of genetic information. *Neurosurg Focus*, 38(1):1–10, 2015.

[4] Forst D, Nahed B, Loeffler J, and Batchelor T. Low-grade gliomas. *Oncologist*, 19(4):403–413, 2014.

[5] Chowdhury R, Wilson I, Lloyd-Jones G, and Rofe C. *Radiology at a Glance [e-book]*. John Wiley & Sons, 2010.

[6] Delgado-López P, Corrales-García E, J Martino, Lastra-Aras E, and Duenas-Polo M. Diffuse low-grade glioma: a review on the new molecular classification, natural history and current management strategies. *Clin Transl Oncol*, 19:931–944, 2017.

[7] Gerin C, Pallud J, Deroulers C, Varlet P, Oppenheim C, and Roux F et.al. Quantitative characterization of the imaging limits of diffuse low-grade oligodendrogliomas. *Neuro-oncology*, 15(10):1379–1388, 2013.

[8] Gerlee P and Nelander S. The impact of phenotypic switching on glioblastoma growth and invasion. *PLOS Computational Biology*, 8(6):1–12, 06 2012.

[9] Gerin C, Pallud J, Grammaticos B, Mandonnet E, Deroulers C, and Varlet P et.al. Improving the time-machine: estimating date of birth of grade ii gliomas. *Cell Proliferation*, 45(1):76–90, 2012.

[10] Swan A, Hillen T, Bowman J, and Murtha A. A patient-specific anisotropic diffusion model for brain tumour spread. *Bulletin of Mathematical Biology*, 80:1259–1291, 2018.

[11] Bogdańska M, Bodnar M, Piotrowska M, Murek M, Schucht P, and Beck J et.al. A mathematical model describes the malignant transformation of low grade gliomas: Prognostic implications. *PLOS ONE*, 12(8):1–24, 08 2017.

[12] Baldock A, Ahn S, Rockne R, Johnston S, and Neal M ans Corwin D et.al. Patient-specific metrics of invasiveness reveal significant prognostic benefit of resection in a predictable subset of gliomas. *PLoS One*, 9(10):263–269, 2014.

[13] Mandonnet E, Pallud J, Clatz O, Taillandier L, Konukoglu E, and Duffau H et.al. Computational modeling of the who grade ii glioma dynamics: principles and applications to management paradigm. *Neurosurgical Review*, 31(3):263–269, 2008.

[14] Murray J. *Mathematical Biology I: An Introduction*. Springer, New York, NY, third edition, 2002.

[15] Murray J. *Mathematical Biology II: Spatial Models and Biomedical Applications*. Springer, New York, NY, third edition, 2003.

[16] Larsen J, Wharton S, McKevitt F, Romanowski C, Bridgewater C, and Zaki H et.al. 'low grade glioma': an update for radiologists. *The British Journal of Radiology*, 90:1–10, 2017.

[17] Mandonnet E, Delattre J, Tanguy M, Swanson K, Carpentier A, and Duffau H et.al. Continuous growth of mean tumor diameter in a subset of grade ii gliomas. *Annals of Neurology*, 53(4):524–528, 2003.

[18] Harpold H, Alvord E, and Swanson K. The Evolution of Mathematical Modeling of Glioma Proliferation and Invasion. *Journal of Neuropathology & Experimental Neurology*, 66(1):1–9, 01 2007.

[19] Chloe Gui, Suzanne E. Kosteniuk, Jonathan C. Lau, and Joseph F. Megyesi. Tumor growth dynamics in serially-imaged low-grade glioma patients. *Journal of Neuro-Oncology*, 139(1):167–175, Aug 2018.

[20] Pallud J, Mandonnet E, Duffau H, Kujas M, Guillevin R, and Galanaud D et.al. Prognostic value of initial magnetic resonance imaging growth rates for world health organization grade ii gliomas. *Annals of Neurology*, 60(3):380–383, 2006.

[21] Amelot A, Mandonnet E, Froelich S, Deroulers C, Badoual M, and Polivka M et.al. Surgical Decision Making From Image-Based Biophysical Modeling of Glioblastoma: Not Ready for Primetime. *Neurosurgery*, 80(5):793–799, 2017.

[22] Kleinbaum D and Klein M. *Survival Analysis*. Springer, New York, NY, 2012.

[23] Dudley W, Wickham R, and Coombs N. An introduction to survival statistics: Kaplan-meier analysis. *J Adv Pract Oncol*, 7(1):91–100, 2016.

[24] Berg S, Danielsson W, Nilsson E, and Orton Sörensen W. Matematisk modellering av hjärntumörer - dataanalys och utvidgning av modell av glioblastom, 2017.

[25] Hadley Wickham and Jennifer Bryan. *readxl: Read Excel Files*, 2019. R package version 1.3.0.

[26] Terry M Therneau. *A Package for Survival Analysis in S*, 2015. version 2.38.

[27] Alboukadel Kassambara and Marcin Kosinski. *survminer: Drawing Survival Curves using 'ggplot2'*, 2018. R package version 0.4.3.

[28] Andréasson N, Evgrafov A, and Patriksson M. *An Introduction to Continuous Optimization*. Studentlitteratur AB, 3 edition, 2016.

[29] Wahde M. *Biologically inspired optimization methods*. Southampton : WIT Press, 2008.

[30] Sébastien Benzekry, Clare Lamont, Afshin Beheshti, Amanda Tracz, John M. L. Ebos, Lynn Hlatky, and Philip Hahnfeldt. Classical mathematical models for description and prediction of experimental tumor growth. *PLOS Computational Biology*, 10(8):1–19, 08 2014.

A

Particle Swarm Optimisation algorithm

Here the Particle Swarm Optimisation algorithm (PSO) will be described more thoroughly to let the reader in detail understand how the algorithm works. Since the algorithm is based on the PSO algorithm described in Mattias Wahdes book *Biologically inspired optimization methods: an introduction*, only the steps that differs from his algorithm are presented here [29, p. 123]. Consider again the flowchart of the PSO algorithm in figure 3.4, where the steps that differs are *Initialise conditions, Evaluate particle* and *Terminate and save results* and will be discussed below. Lastly, a test for checking the robustness of the algorithm was made and the result will be presented below.

A.1 Initialise conditions

The first step in the algorithm is to extract data from the patient of consideration. The data of interest is the age of the patient, the volumes estimated from Magnetic Resonance images with T2-FLAIR, the time difference between the measurements and lastly the volume estimated from MR images using T1 and the time point it was taken. From this data, the velocity and steepness from data can be calculated, together with the mismatch that will be used when examining the result. The velocity is calculated from the linear regression

$$r_{data} = r^0_{data} + c_{data} t_{data} \quad i = 1, 2, ... \tag{A.1}$$

using least squares, and the steepness is calculated as

$$s_{data} = \frac{-(0.16 - 0.80)}{r_{0.16} - r_{0.80}} \tag{A.2}$$

where $r_{0.80}$ is the radius calculated from the volume estimated from MR images using T1 and $r_{0.16}$ is the corresponding radius measured from MR images using T2-FLAIR taken at the same time point as $r_{0.80}$. With the data extracted and velocity and steepness defined, the next step is to define the settings necessary for solving the Fisher equation.

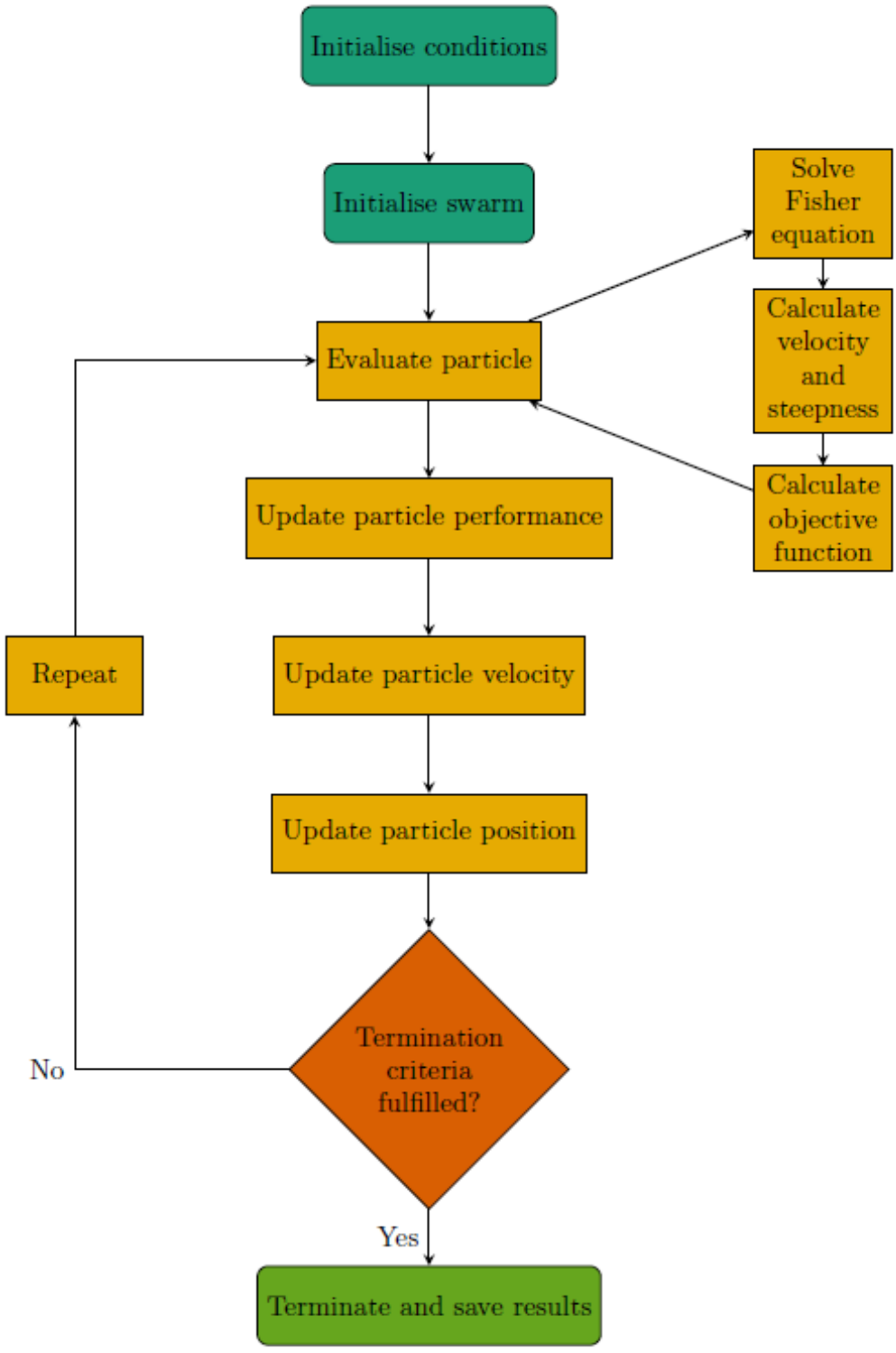


Figure A.1: *Flowchart illustrating in general how the Particle Swarm Optimisation algorithm works. The colors represents different phases of the code: Initialisation phase in dark green, the iteration process in yellow/orange and the termination step in light green*

Assuming that the analytically calculated parameters are close to the optimal values of the parameters, they can be used as an initial guess of the parameters. The parameter values estimated analytically are calculated as

$$\begin{cases} c = 2\sqrt{\rho D} \\ s = \frac{-(0.16-0.80)}{\frac{c}{\rho}\left(\ln\left(\frac{1}{0.16}-1\right)-\ln\left(\frac{1}{0.80}-1\right)\right)} \end{cases} \quad \Leftrightarrow \quad \begin{cases} D = \frac{0.16-0.80}{\ln\left(\frac{1}{0.80}-1\right)-\ln\left(\frac{1}{0.16}-1\right)}\frac{c}{4s} \approx 0.053\frac{c}{s} \\ \rho = \frac{\ln\left(\frac{1}{0.16}-1\right)-\ln\left(\frac{1}{0.80}-1\right)}{0.80-0.16}cs \approx 4.757cs \end{cases} \tag{A.3}$$

and are used to solve the PDE solution of the Fisher equation. From that solution, time and spatial size is decided, in order to minimise the risk of numerical errors. In addition, the maximum radius and time is initially defined as the radius of the last measurements and the time estimated from the linear regression and additional time and space is added according to the solution in order to make sure that the constant wave phase is reached. Using the analytically calculated parameter values, the range of possible values for the PSO is defined.

A.2 Evaluate particle

In this step, every particle with its unique combination of parameter values is evaluated based on its capability to properly estimate the velocity and steepness measured from data. The PDE solution of the Fisher equation is solved using the parameter values from the particle and using this solution, the velocity and steepness is calculated numerically according to the procedure presented in section 3.3. These values are then used in the objective function, defined as

$$f(c_{numeric}, s_{numeric}) = \frac{(c_{numeric} - c_{data})^2}{c_{data}} + \frac{(s_{numeric} - s_{data})^2}{s_{data}} \tag{A.4}$$

where f is the objective function that should be minimized, $c_{numeric}$ $s_{numeric}$ is the velocity and steepness measured using the parameter values D and ρ for each particle and c_{data} and s_{data} is the velocity and steepness measured from data. The value of the objective function is returned to the main algorithm and the particles that minimises the objective function the most is the ones that performs the best and thus leads the rest of the swarm closer to the optimum.

In this step, there are many ways of getting either unreasonable results or even values of the objective function that is Not a Number (NaN). In order to restrict the algorithm and discard unreasonable solutions, the objective function of a particle directly gets a large value, ranging from 1000 to 1000000, if:

1. One or both of the parameters D and ρ have a value below zero,
2. The MATLAB function `pdepe`, calculating the PDE solution of the Fisher equation, gets numerical errors and fails to calculate the proportional cell density for all time steps. If this happens, the code enters a while-loop that decreases the spatial step size and calculates the Fisher equation once again. If the numerical errors remains after ten iterations, the solution is discarded.
3. The solution of the Fisher equation has not reached the constant wave phase,
4. either c or s is NaN.

A.3 Terminate and save results

The algorithm is terminated if the maximum number of iterations is reached or the objective function gets a value below 10^{-8}. After termination, the velocity and steepness profiles over time are calculated and visualised in order to check if the results seems reasonable. If so, the optimum parameter values and the value of the objective function is saved, together with the settings that was used in order to retrieve the result, for reproducibility.

A.4 Robustness test of the PSO algorithm

In order to investigate if the stochastic optimization algorithm PSO gets different values for every run or not, the algorithm was run ten times under the same settings and the result was retrieved for every run. The result is summarised in table A.1 and shows that it is able to find the same optimum for every run. This shows that although the PSO was only run once for many patients, the result would not have been different if averaging over many runs.

Table A.1: *Results from running a PSO using the same settings multiple times. The values varies barely and indicates that the algorithm finds the same optimum every time.*

Run	1	2	3	4	5	6	7	8	9	10	**Mean:**
D_{opt}	0.313	0.313	0.313	0.313	0.313	0.313	0.313	0.313	0.313	0.313	**0.313**
ρ_{opt}	0.608	0.608	0.608	0.608	0.608	0.608	0.608	0.609	0.608	0.608	**0.6081**
Obj.fct $\cdot 10^8$	0.014	14	0.67	0.80	5.00	2.00	2.60	27.0	5.00	5.00	**6.21**

B

Tumour growth over time for all patients in the SC-data

In this Appendix additional plots illustrating tumour growth are visualised, using the subset SC-data of 14 patients in total. The volumetric growth is shown in figure B.1 and the radial growth in figure B.2, illustrating great variety in growth and number of measurements. In figure B.3, the normalised growth is visualised, where the radius at each time point is divided by the initial radius. Figure B.3 thus illustrates at which rate the tumour grows relative its own size. For most patients, the radius of the last measurement is 1.2-1.4 times larger than the initial radius, but for one patient the last radius is approximately 2.6 times larger. For better visibility, the y-axis is given a maximum value of 1.5, thus not visualising the radius mentioned above.

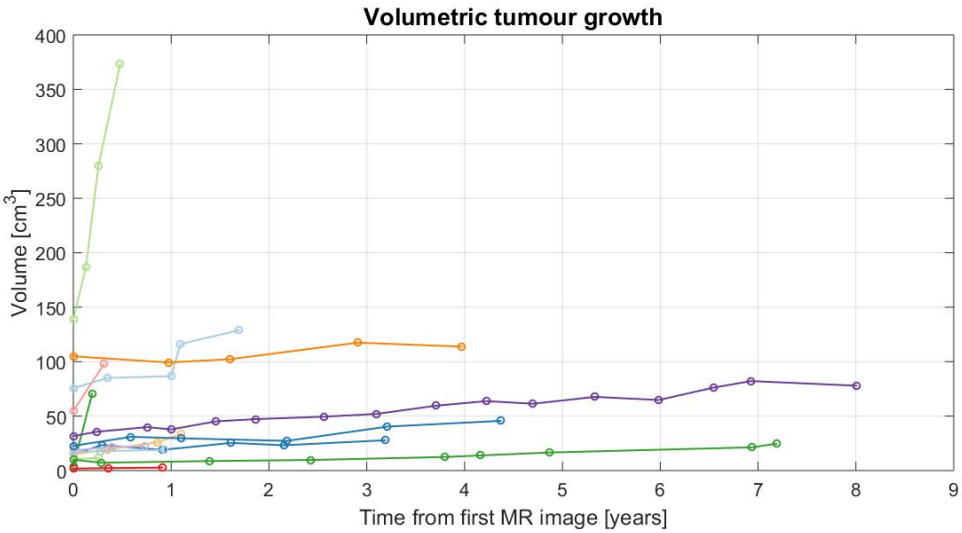


Figure B.1: *Volumetric growth over time, using the data subset that is also used in the parameter estimation, SC-data.*

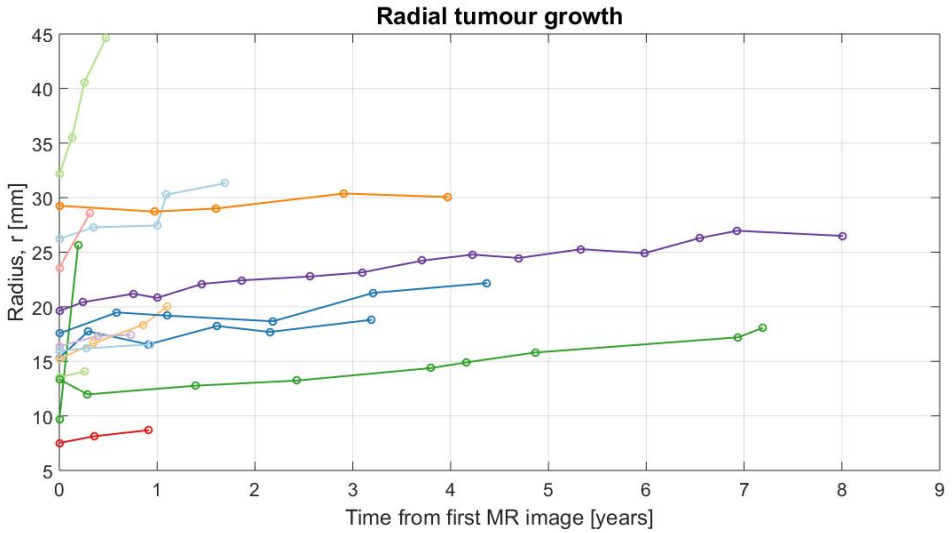


Figure B.2: *Radial growth over time, using the data subset that is also used in the parameter estimation, SC-data.*

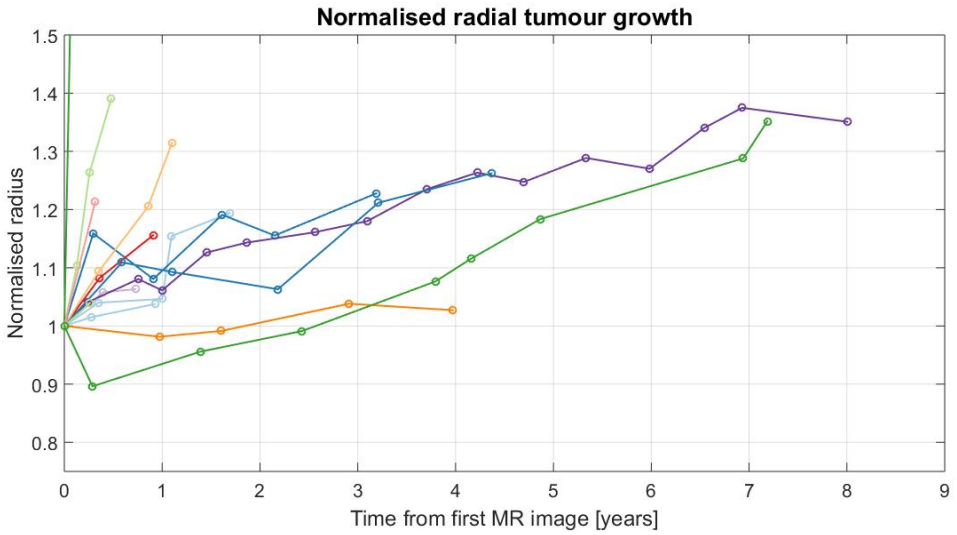


Figure B.3: *Normalised radial growth over time, using the data subset that is also used in the parameter estimation, SC-data. The radii is divided by the initial volume in order to visualise how fast the tumours grow relative their own size.*

C

Complete tables of data and results

In section 4.1.1 and 4.3.2, only summaries of the data and values of parameters and properties have been presented. In this Appendix, the complete tables of all summaries will be visualised. The tables in consideration are table 4.1, 4.2 and 4.6.

For the summary of tumour properties, table 4.1 and 4.2, the full table can be found in table C.1, where the rows colored in green represents the patients also included in the subset SC-data. The abbreviations used is patient index (Index), age at diagnosis (Age_d [$years$]), time to surgery (Time_s [$years$]), tumour volume at diagnosis (V_d [cm^3]), tumour radius at diagnosis (r_d [mm]), velocity (c [$mm/year$]), steepness (s [mm^{-1}]) and mismatch (MM).

For the results from the parameter estimation using Particle Swarm Optimization, table 4.6, the full table is presented in table C.2. The abbreviations is patient index (Index), age at diagnosis (Age_d), velocity (c [$mm/year$]), steepness (s [mm^{-1}]), diffusion coefficient analytically calculated (D_A [$mm^2/year$]), growth rate analytically calculated (ρ_A [$year^{-1}$]), diffusion coefficient calculated from PSO (D_{PSO} [$mm^2/year$]), growth rate calculated from PSO (ρ_{PSO} [$year^{-1}$]), value of objective function (ObjFct), estimated time before diagnosis (Time_d [$years$]), invisible tumour volume (V_I [cm^3]) and invisible radius (r_I [mm]). The horizontal lines is only to divide the table into smaller sections, for better readability.

Table C.1: *Complete table of the tumour properties for all patients. The rows coloured in green represents the patients that are also used in the parameter estimation and all abbreviations are described in the text above.*

Index	Age_d	Time_s	V_d	r_d	c	s	MM
1	67	1.693	75.600	26.231	2.950	0.071	0.636
2	63	3.189	15.042	15.313	0.811	0.088	0.767
3	62	3.986	66.330	25.112	1.222	-0.896	-0.073
4	57	0.416	20.270	16.914	-0.070	1.027	0.107
5	50	0.729	0.000	0.000	34.206	-0.209	-0.415
6	49	0.375	29.800	19.233	-0.831	0.096	0.723
7	48	0.230	42.432	21.637	-0.210	0.630	0.135
8	21	0.479	138.968	32.132	26.730	0.047	0.668
9	66	0.192	3.757	9.644	83.341	0.051	0.869
10	53	0.315	54.853	23.571	15.959	0.211	0.338
11	46	0.482	18.390	16.374	-3.378	-0.239	-0.651
12	32	0.148	81.240	26.867	-5.508	-1.214	-0.062
13	56	0.910	1.779	7.517	1.267	1.205	0.172
14	40	0.622	113.840	30.066	-3.815	-2.205	-0.032
15	25	1.471	14.800	15.231	4.156	0.552	0.161
16	69	3.973	104.956	29.262	0.341	1.700	0.037
17	48	0.367	3.693	9.589	0.349	-1.921	-0.106
18	41	0.455	2.700	8.638	1.218	-0.435	-0.712
19	37	0.726	18.414	16.381	1.456	0.169	0.520
20	31	0.208	31.270	19.544	-0.879	0.410	0.223
21	43	0.244	10.510	13.588	-0.463	0.229	0.502
22	58	0.258	81.933	26.944	-2.564	0.110	0.518
23	56	8.055	31.611	19.615	0.873	1.300	0.056
24	63	1.085	17.000	15.951	0.630	0.538	0.239
25	42	0.247	2.509	8.430	-0.761	-0.516	-0.603
26	40	4.367	22.681	17.560	0.911	0.155	0.461
27	67	0.263	10.410	13.545	2.030	0.347	0.344
28	51	7.186	9.980	13.356	0.743	2.369	0.046

Table C.2: *Complete table of the result from the parameter estimation using Particle Swarm Optimization, abbreviations are described in the text above.*

Index	Age_d	c	s	D_A	D_{PSO}	ρ_A	ρ_{PSO}	ObjFct	Time_d	V_I	r_I
1	67	2.950	0.071	2.171	1.705	1.002	1.019	$1.19 \cdot 10^{-6}$	19	118.654	7.582
2	63	0.811	0.088	0.482	0.413	0.341	0.350	$2.83 \cdot 10^{-9}$	50	40.694	6.570
8	21	26.730	0.047	30.123	25.452	5.930	6.043	$1.83 \cdot 10^{-9}$	3	421.824	12.504
9	66	83.341	0.051	86.256	67.320	20.131	20.148	$2.20 \cdot 10^{-6}$	1	160.220	13.258
10	53	15.959	0.211	3.969	3.914	16.043	16.219	$4.40 \cdot 10^{-10}$	2	22.234	2.152
13	56	1.267	1.205	0.055	0.047	7.261	7.344	$4.61 \cdot 10^{-9}$	6	0.381	0.382
15	25	4.156	0.552	0.395	0.388	10.923	11.063	$8.50 \cdot 10^{-9}$	4	4.230	0.805
16	69	0.341	1.700	0.011	0.004	2.754	2.770	$6.34 \cdot 10^{-9}$	87	3.575	0.312
19	37	1.456	0.169	0.452	0.411	1.172	1.206	$4.40 \cdot 10^{-9}$	18	12.322	2.767
23	56	0.873	1.300	0.035	0.029	5.401	5.439	$5.33 \cdot 10^{-9}$	23	3.218	0.359
24	63	0.630	0.538	0.062	0.061	1.612	1.636	$4.89 \cdot 10^{-9}$	28	2.988	0.826
26	40	0.911	0.155	0.309	0.276	0.672	0.692	$4.54 \cdot 10^{-10}$	32	20.509	2.920
27	67	2.030	0.347	0.308	0.305	3.349	3.405	$8.39 \cdot 10^{-9}$	9	3.547	1.301
28	51	0.743	2.369	0.017	$2.34 \cdot 10^{-7}$	8.37	7.402	$3.21 \cdot 10^{-2}$	18	1.227	0.294

D

Results from the Particle Swarm Optimization algorithm for all patients

In this Appendix the results of the PSO is presented for all patients that were not mentioned in section 4.3.1. Here only the table presenting the summary of the performance is shown with a short comment regarding the result.

Patient 2

The summary is found in the table below and the solutions differs slightly.

Table D.1: *Summary of the velocity c, steepness s, mismatch MM and the parameter values D and ρ for Patient 2 (63 years), calculated either from Data, using Analytic calculations or estimated using PSO.*

	$c\ [mm/year]$	$s\ [mm^{-1}]$	MM	$D\ [mm^2/year]$	$\rho\ [year^{-1}]$
Data	0.811	0.088	0.767	-	-
Analytic	-	-	-	0.482	0.341
PSO	0.811	0.088	0.838	0.413	0.35

Patient 8

The summary is found in the table below and the solutions differs slightly.

Table D.2: *Summary of the velocity c, steepness s, mismatch MM and the parameter values D and ρ for Patient 8 (21 years), calculated either from Data, using Analytic calculations or estimated using PSO.*

	$c\ [mm/year]$	$s\ [mm^{-1}]$	MM	$D\ [mm^2/year]$	$\rho\ [year^{-1}]$
Data	26.73	0.047	0.668	-	-
Analytic	-	-	-	30.123	5.93
PSO	26.728	0.047	0.739	25.452	6.043

Patient 9

The summary is found in the table and figures below. Here, the results differs quite much. Also, when running the PSO, numerical errors occurred on regular basis even in the neighbourhood of the analytic solution, which is used as a initial guess. Since particles with parameter values that fails to calculate the proportional cell density u for all r and t automatically gets a objective function value of 1000000, see Appendix A, the true optimum may have been discarded.

Table D.3: *Summary of the velocity c, steepness s, mismatch MM and the parameter values D and ρ for Patient 9 (66 years), calculated either from Data, using Analytic calculations or estimated using PSO.*

	$c\ [mm/year]$	$s\ [mm^{-1}]$	MM	$D\ [mm^2/year]$	$\rho\ [year^{-1}]$
Data	83.341	0.051	0.869	-	-
Analytic	-	-	-	86.256	20.131
PSO	83.342	0.05	0.892	67.32	20.148

Patient 10

The summary is found in the table below and the solution using the result from the PSO and the analytic calculations are very similar.

Table D.4: *Summary of the velocity c, steepness s, mismatch MM and the parameter values D and ρ for Patient 9 (53 years), calculated either from Data, using Analytic calculations or estimated using PSO.*

	$c\ [mm/year]$	$s\ [mm^{-1}]$	MM	$D\ [mm^2/year]$	$\rho\ [year^{-1}]$
Data	15.959	0.211	0.338	-	-
Analytic	-	-	-	3.969	16.043
PSO	15.959	0.211	0.335	3.914	16.219

Patient 13

The summary is found in the table below and the solutions differs slightly.

Table D.5: *Summary of the velocity c, steepness s, mismatch MM and the parameter values D and ρ for Patient 13 (56 years), calculated either from Data, using Analytic calculations or estimated using PSO.*

	$c\ [mm/year]$	$s\ [mm^{-1}]$	MM	$D\ [mm^2/year]$	$\rho\ [year^{-1}]$
Data	1.267	1.205	0.172	-	-
Analytic	-	-	-	0.0553	7.261
PSO	1.267	1.205	0.173	0.0471	7.344

Patient 19

The summary is found in the table below and the solution using the result from the PSO and the analytic calculations are very similar.

Table D.6: *Summary of the velocity c, steepness s, mismatch MM and the parameter values D and ρ for Patient 19 (37 years), calculated either from Data, using Analytic calculations or estimated using PSO.*

	c $[mm/year]$	s $[mm^{-1}]$	MM	D $[mm^2/year]$	ρ $[year^{-1}]$
Data	1.456	0.169	0.52	-	-
Analytic	-	-	-	0.452	1.172
PSO	1.456	0.169	0.533	0.411	1.206

Patient 23

The summary is found in the table below and the solutions differs slightly. For this patient the properties are well estimated from data, since in total 17 measurements were used which is much larger than the average number of measurements per patient. This could imply that the Fisher equation *can* catch tumour behaviour for this specific patient, since the properties are accurately estimated by the PSO and the patient data is reliable, assuming that the Fisher equation is a good model.

Table D.7: *Summary of the velocity c, steepness s, mismatch MM and the parameter values D and ρ for Patient 23 (56 years), calculated either from Data, using Analytic calculations or estimated using PSO.*

	c $[mm/year]$	s $[mm^{-1}]$	MM	D $[mm^2/year]$	ρ $[year^{-1}]$
Data	0.873	1.3	0.056	-	-
Analytic	-	-	-	0.0353	5.401
PSO	0.874	1.3	0.056	0.0286	5.439

Patient 24

The summary is found in the table below and the solution using the result from the PSO and the analytic calculations are very similar.

Table D.8: *Summary of the velocity c, steepness s, mismatch MM and the parameter values D and ρ for Patient 24 (63 years), calculated either from Data, using Analytic calculations or estimated using PSO.*

	$c\ [mm/year]$	$s\ [mm^{-1}]$	MM	$D\ [mm^2/year]$	$\rho\ [year^{-1}]$
Data	0.63	0.538	0.239	-	-
Analytic	-	-	-	0.0615	1.612
PSO	0.63	0.538	0.24	0.0607	1.636

Patient 26

The summary is found in the table below and the solutions differs slightly. This patient has the same problem as Patient 16, meaning that the estimated age of the tumour is very unreasonable. Patient 26 is 40 years old while the estimated tumour age is 32 years, meaning that the PSO result implies that the patient got its tumour at 8 years of age.

Table D.9: *Summary of the velocity c, steepness s, mismatch MM and the parameter values D and ρ for Patient 26 (40 years), calculated either from Data, using Analytic calculations or estimated using PSO.*

	$c\ [mm/year]$	$s\ [mm^{-1}]$	MM	$D\ [mm^2/year]$	$\rho\ [year^{-1}]$
Data	0.911	0.155	0.461	-	-
Analytic	-	-	-	0.309	0.672
PSO	0.911	0.155	0.462	0.276	0.692

Patient 27

The summary is found in the table below and the solution using the result from the PSO and the analytic calculations are very similar.

Table D.10: *Summary of the velocity c, steepness s, mismatch MM and the parameter values D and ρ for Patient 27 (67 years), calculated either from Data, using Analytic calculations or estimated using PSO.*

	$c\ [mm/year]$	$s\ [mm^{-1}]$	MM	$D\ [mm^2/year]$	$\rho\ [year^{-1}]$
Data	2.03	0.347	0.344	-	-
Analytic	-	-	-	0.308	3.349
PSO	2.03	0.347	0.348	0.305	3.405

Patient 28

The summary is found in the table below. What is interesting with this patient is that the Fisher equation in Spherical coordinates fails to capture the steepness from the data. The PDE is unable to be as steep as the data suggests, leading to a solution that differs from the analytic calculations and a much larger value of the objective function (0,0321), relative the other patients (median: $5.111 \cdot 10^{-9}$).

Table D.11: *Summary of the velocity c, steepness s, mismatch MM and the parameter values D and ρ for Patient 28 (years), calculated either from Data, using Analytic calculations or estimated using PSO.*

	$c\ [mm/year]$	$s\ [mm^{-1}]$	MM	$D\ [mm^2/year]$	$\rho\ [year^{-1}]$
Data	0.743	2.369	0.046	-	-
Analytic	-	-	-	0.0165	8.377
PSO	0.743	2.093	0.052	$2.34 \cdot 10^{-7}$	7.402