Image-informed patient-specific breast tumour models for early prediction of neoadjuvant chemotherapy response

Rose Mathilde Collet

The University of Leeds

Centre for Doctoral Training in Fluid Dynamics

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

February 2025

À Martin et Sara,

Chacun sa quiche.

Intellectual Property Statement

I confirm that the work submitted is my own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Rose Mathilde Collet to be identified as Author of this work has been asserted by Rose Mathilde Collet in accordance with the Copyright, Designs and Patents Act 1988.

Acknowledgements

First and foremost, I'm very grateful to my supervisors Zeike Taylor and David Buckley for their support, guidance and encouragement throughout this project. Zeike, thank you especially for your kindness, patience and good humour. Your support and encouragement in moments of frustration and doubt made this process far more manageable. David, I am especially grateful for your genuine interest in my work and for always taking the time to engage with my ideas. Our meetings where you showed enthusiasm for what I had done made such a difference in keeping me motivated.

I'm deeply grateful to my friends and colleagues at the Université du Luxembourg, who hosted me for two months and several return visits, and without whom this work would not have been possible. I owe a particular debt of gratitude to Stéphane Urcun, whose generosity with his time and expertise made a profound impact on me. I greatly benefitted from his dedication to fostering a tumour modelling community in Europe, and his mentorship has significantly shaped both my personal and professional growth. Camilo Suarez Afanador was incredibly helpful with software installations and HPC queries, and excellent company over many weekends spent in the office. He also kindly provided HPC access. Additionally, I would like to thank Meryem Abbad Andaloussi, Thomas Lavigne, Anne Kerachni, Andreÿ Latyshev, Paulina Mendoza Ayllón and Diego Hurtado Catalifaud for their friendship and support during my time in Luxembourg and beyond.

Thank you to the EPSRC CDT in Fluid Dynamics for the opportunity and funding to undertake this research, and for the generous travel budget. I'm grateful to have been able to speak at so many interesting conferences over the last 3.5 years. Thank you to Tom Ranner for conducting my annual progress reviews with exceptional kindness; to Zyad Almutlaq for helping me get to grips with MRI and the CHERNAC study; to the rest of cohort

7 for locking in and finishing their theses on time, as it was very motivating to think of us all graduating together.

I'm grateful for the many friendships I have made through the CDT. From the very start, I have looked up to and been inspired by Michael MacRaild, who was the first person I spoke to when choosing this project and who has been there to provide guidance at every turn ever since. He notably proofread most of this thesis – a thankless task, given my stubbornness and personal punctuation beliefs. I reluctantly removed several semi-colons in his honour; I did not, however, add half the commas he suggested. I'm also particularly grateful for the consistent support and encouragement of my officemate Alasdair Roy. His kindness, unwavering belief in me, and ability to make me laugh at myself mid-cry have meant more than I can say. My work has greatly benefitted from Michael and Alasdair's input and I couldn't be more grateful for their support and friendship.

Some personal thank-yous to my friends and family, and a sincere apology for being so unavailable (/neurotic when available) over the last few months. Thank you to my parents for never doubting that I could do this. To my sister Sara, my best friend. To Gran for encouraging me to do a good job, Karen for encouraging me to do a good *enough* job, Colin for the pep talks and advice, Ben for being my shoulder most cried on, to Martin, my brother. To George for car talk, Elliot for feeding me while I wrote most of this, Greg, Al and Jose for the company in the office, Emily, Cristina, Olga, Seth, Izzy, Vanessa, Sam, Jocelyne.

Lastly, I would also like to thank my car for breaking down 6 weeks ago – perfect timing to ensure I remained firmly at my desk and completed this thesis.

Abstract

Neoadjuvant chemotherapy (NACT) is a widely used treatment for patients with locally advanced breast cancer. However, some breast cancer subtypes exhibit pathological complete response rates as low as 10% [1]. Non-responsive patients endure NACT side-effects without therapeutic benefit. It is essential to identify non-responsive tumours as early as possible, to enable clinicians to discontinue the unsuccessful NACT and proceed with alternative treatments. Furthermore, recognising exceptional responders could allow clinicians to minimise their unnecessary exposure to NACT.

A promising new approach, gaining traction in breast tumour modelling over the past decade, is the integration of patient-specific imaging data into mathematical models of tumour physiology and treatment response [2]. Among these, predictive models based on reaction-diffusion (RD) equations have been widely explored [3–6]. This thesis presents a family of four mathematical models based on these and tailored to each individual patient using magnetic resonance imaging (MRI) data. These models are designed to predict tumour progression throughout the course of NACT.

Models were implemented in FEniCSx and personalised using diffusion-weighted (DW) and dynamic contrast-enhanced (DCE) MRI from 34 patients, acquired before and after one NACT cycle, to predict response after three cycles. After the first cycle, up to three model parameters are calibrated to each patient: tumour cell proliferation, diffusion and drug efficacy.

The best-performing model was the simplest, one that accounts solely for the effects of chemotherapy without incorporating tumour growth. This model distinguished responders from non-responders after just one cycle of NACT, achieving 82% sensitivity, 75% specificity and an 86% negative predictive value. These findings suggest that tumour cell diffusion is not a critical factor in predicting response to NACT, which aligns with the observation that breast tumours do not exhibit significant growth over NACT timescales.

Contents

Abbreviations

α: chemotherapy-only model (5.1)

2D: two-dimensional, two dimensions

3D: three-dimensional, three dimensions

AI: artificial intelligence

ADC: apparent diffusion coefficient

AIF: arterial input function

BINNs: biology-informed neural networks

BFGS: Broyden–Fletcher–Goldfard–Shanno

CA: contrast agent

C_{drug}^{plasma}: concentration of chemotherapy drug estimated in plasma

C_{drug}^{tissue}: concentration of chemotherapy drug estimated in tumour tissue

COBYLA: Constrained Optimization BY Linear Approximations

CR: complete response (RECIST category)

$D\alpha$: diffusion and chemotherapy model (5.3)

DCE: dynamic contrast enhanced (-MRI)

DICOM: Digital Imaging and COmmunications in Medicine, file format

DI-MC-RD: drug-incorporated mechanically-coupled reaction–diffusion, equations (2.6)–(2.8)

DSC: Dice similarity coefficient

DW: diffusion-weighted (-MRI)

EC90: combined epirubicin and cyclophosphamide chemotherapy protocol

HPC: high performance computing

FD: finite difference

FDM: finite difference method

FE: finite elements

FEM: finite element method

FOV: field of view

Hct: haematocrit

HSR: high spatial resolution

HTR: high temporal resolution

$k\alpha$: proliferation and chemotherapy model (5.2)

$kD\alpha$: proliferation, diffusion and chemotherapy model (5.4)

K^{trans}: volume transfer constant from plasma to tissue

MC-RD: mechanically-coupled reaction–diffusion, equations (2.5)–(2.7)

ML: machine learning

MRE: magnetic resonance elastography

MRI: Magnetic Resonance Imaging

NACT: neoadjuvant chemotherapy

NPV: negative predictive value

ODE: ordinary differential equation

pCR: pathological complete response

PD: progressive disease (RECIST category)

PDE: partial differential equation

PMI: Platform for Research in Medical Imaging

PPV: positive predictive value

PR: partial response (RECIST category)

RECIST: response evaluation criteria in solid tumours

ROI: region of interest

RD: reaction–diffusion

SI: signal intensity

SD: stable disease (RECIST category)

SSD: sum of squared differences

TE: echo time

TR: repetition time

Chapter 1

Introduction

Breast cancer is the most common cancer in women across the globe and a leading cause of death. Since 2007, a moderate but significant increase in incidence has been reported in many countries [7]. Neoadjuvant chemotherapy (NACT) is a common treatment option for patients with locally advanced breast cancer (stage II-III) [8, 9], delivered with the intention of shrinking the tumour before proceeding to surgery [10]. However, pathological complete response (pCR) rates can range from under 10% to over 60% depending on the breast cancer subtype and the treatment regimen [1, 11].

Non-responsive patients suffer the side-effects of their chemotherapy regimen without reaping any benefit. There is a clear need to accurately identify non-responsive tumours as early as possible [6, 9, 12–14], enabling clinicians to replace the unsuccessful NACT with an alternative, potentially more successful therapy. Furthermore, recognising exceptional responders could allow clinicians to minimise their unnecessary exposure to NACT.

Reliable techniques to accurately evaluate patient response early in the course of NACT are currently lacking [6, 12]. Most commonly, response is assessed based on the size of the tumour: morphological changes are gauged by physical exam or conventional medical imaging, and metrics such as the Response Evaluation Criteria in Solid Tumours (RECIST) are used to categorise the response objectively [15]. This approach is fundamentally limited, as these changes in size may not occur until after several cycles of NACT [6, 9, 14, 15]. On the other hand, parameters including tumour cell density, blood flow or tissue volume fractions which can be estimated using quantitative magnetic resonance imaging (MRI) [9, 16] are much quicker to show evidence of response [12]. While these techniques hold greater potential than clinical assessment and conventional scans, imaging data alone are often insufficient to characterise treatment response [17, 18] or offer quantitative forecast.

A promising new approach that has gained traction in breast tumour modelling over the past decade is integrating patient-specific imaging data into mathematical models of tumour physiology and treatment response [3–6]. Among these, a family of predictive models based on reaction-diffusion (RD) equations has been widely explored. RD models describe tumour growth by capturing the interplay between cell proliferation and diffusion throughout the tumour microenvironment. These models have been the subject of extensive development and have shown promise in forecasting tumour response to therapy. A drawback of RD-type models is their reliance on a high number of parameters, which increases the risk of overfitting and can reduce interpretability. A detailed discussion of RD models and their formulation is provided in Chapter 2.

Image-informed models bridge data-driven and knowledge-driven predictions, with imaging data offering a detailed characterisation of the patient's anatomical, functional, and/or pathological state, while the models incorporate underlying biological mechanisms that govern tumour behaviour and treatment response. By combining these

complementary sources of information, rather than relying solely on image-derived or clinical predictors, this integrated approach has the potential to enhance predictive accuracy and improve treatment personalisation. In practice, the model is initialised with imaging data and calibrated to represent a specific patient's tumour, then solved to determine the time evolution of the tumour under the prescribed chemotherapy regimen.

In this thesis, I adopt the recommendation of Brady and Henderling [19]: a quantity of interest that can be measured experimentally (in the clinic or *in vitro*), a minimal number of parameters, parameter initial values extracted from the available data rather than from theoretical literature as much as possible, and the evaluation of the calibrated parameters performed on data distinct from the calibration. To this effect the number of calibrated parameters in the current state-of-the-art, the drug-incorporated mechanically-coupled reaction–diffusion (DI–MC–RD) model [4], is reduced.

This model and other current image-informed tumour modelling approaches are discussed in Chapter 2. The imaging dataset is presented in Chapter 3, along with its analysis and preprocessing for use in such models. Chapter 4 examines the current state-of-the-art, identifying key computational challenges. Chapter 5 builds on these findings, proposing a family of light and efficient mathematical models that replace the spatially-varying calibrated proliferation matrix in the DI–MC–RD with a single calibrated scaling of a perfusion image. Finally, Chapter 6 summarises the findings and discusses limitations and potential future work.

Chapter 2

Background

Contents

This chapter presents a literature review of mechanistic mathematical models of breast tumour growth and response to treatment. Predicting treatment response on a patient-specific basis requires a reliable mathematical model grounded in biophysical reality. Such a model must be personalised using patient data, while parameters that cannot be measured *in vivo* should be calibrated to the individual patient.

We begin by reviewing mathematical models of breast tumour growth in Section 2.1, focusing on those with practical relevance to predicting tumour response in the manner described in Chapter 1, rather than purely theoretical approaches. Next, we examine methods for personalising these models in Section 2.2, including how routinely acquired MRI data can be used for parameterisation. Finally we discuss the practical implementation of these models in Section 2.3, covering essential image preprocessing steps and numerical techniques required for their application.

2.1 Mathematical models of tumour growth

A large number of mathematical models of tumour growth exist, broadly falling into three categories: continuum, discrete and hybrid. Continuum approaches treat the material as a continuum and model variables as smooth fields; they can be based on ordinary differential equations (ODEs, Subsection 2.1.1) or partial differential equations (PDEs, Subsections 2.1.2 and 2.1.3). In contrast, discrete approaches such as agent-based models and cellular automata derive overall prediction from considering the dynamics of single cells [20, 21]. These are thus able to model different types of tumour cells and stochastic events such as cell mutation [21, 22], but this comes at a

prohibitively large computational cost: only a small number of tumour cells (of the order of thousands) can be considered [21, 23, 24]. As even the smallest tumours contain at least tens of millions of cells, this largely prohibits their clinical use. Furthermore, discrete models often require a large number of parameters that can be impossible to measure *in vivo*. On the other hand, hybrid models can be successful but require a breadth of data across multiple scales that is not routinely acquired [25, 26]. For these reasons, we focus on continuum models.

2.1.1 Exponential, Gompertz and logistic models

The earliest attempts at capturing the dynamics of tumour growth mathematically used exponential models to describe the changes in population observed in tumour cell cultures [27]. These models describe the number of tumour cells N at time t as:

$$N(t) = N(0)e^{bt}, \tag{2.1}$$

where b is a constant [28]. As more *in vitro* cancer cell population data was acquired, it was observed that exponential population growth only occurred at the earliest stages of tumour development [29]. Once the population reaches a certain size, the notion of a "carrying capacity", a maximum number of cells that can be sustained by the environment, limits this exponential growth [20, 29, 30]. Consequently, Gompertzian models (such as equation (2.2) [28, 30]) and logistic models (equation (2.3) [31]) were favoured for their ability to reflect this:

$$N(t) = N(0)\exp\Big\{\log\left(\theta/N(0)\right)\cdot\left[1-e^{-bt}\right]\Big\}, \tag{2.2}$$

$$N(t) = \frac{\theta N(0)}{N(0) + (\theta - N(0))\exp^{-kt}}, \tag{2.3}$$

where θ is the carrying capacity and k the proliferation rate. As these equations are not spatially-varying, these models are unable to account for spatial variation and growth. This fundamentally limits their application to tumour growth modelling, where cells are mobile and the spatial heterogeneity of the surrounding environment is known to impact tumour evolution [32].

2.1.2 Reaction–Diffusion models (RD)

Modelling spatio-temporal tumour evolution requires partial differential equation (PDE) models, rather than one-dimensional ordinary differential equations (ODEs) [20]. Models based on the reaction–diffusion (RD) equation are most widely used, as they are well-suited to personalisation (see Section 2.2) and are often more tractable numerically than other PDE models, such as advection–diffusion [21, 33, 34]. Pioneered by Swanson et al. [35] in the brain in the early 2000s, RD models are able to account for the spatial proliferation and diffusion of tumour cells and are particularly appropriate for problems in which interaction with the environment plays an important role in tumour evolution [21, 36]. In the following example [35], the RD equation describes the rate of change of the tumour cell population density as the sum of the diffusion of the tumour cells in the brain and their net proliferation:

$$\frac{\partial N(\mathbf{x},t)}{\partial t} = \nabla \cdot (D(\mathbf{x})\nabla N(\mathbf{x},t)) + kN(\mathbf{x},t), \tag{2.4}$$

where $N(\mathbf{x},t)$ represents the concentration of tumour cells at location $\mathbf{x}$ and time t, k is the proliferation rate of tumour cells. $D(\mathbf{x})$ is a two-valued function controlling the random movement of glioma cells. In this model, its spatial variation is only used to assign different scalar values in grey and white matter, as cells have been observed to move around more freely in white matter than in grey [37].

2.1.3 Mechanically-Coupled Reaction–Diffusion models (MC–RD)

Experiments with cell cultures led Helmlinger et al. [38] to discover that a growing tumour is restricted by the mechanical stress caused as it pushes against the surrounding tissue. This phenomenon is known as the mass effect. Several mathematical models were rapidly able to reproduce this *in vitro* behaviour, but were too simplistic to be relevant *in vivo* [39, 40]. Garg and Miga [41] bridged this gap by extending Swanson's RD model [35] to account for the mechanoinhibitory effect of the surrounding tissue in brain tumour modelling, thus formulating the first such mathematical model for macroscopic tumour growth. They justified this approach by noting that the inhibitory

effect of mechanical stress on glioma growth could be significant, given its expansion within the confined space of the cranium. This is referred to as "mechanically-coupled reaction–diffusion" (MC–RD):

$$\frac{\partial N(\mathbf{x},t)}{\partial t} = \underbrace{\nabla \cdot \Big(D(\mathbf{x},t)\nabla N(\mathbf{x},t)\Big)}_{\text{Random cell diffusion}} + \underbrace{k(\mathbf{x})\left(1-\frac{N(\mathbf{x},t)}{\theta}\right)N(\mathbf{x},t)}_{\text{Logistic growth}}, \tag{2.5}$$

$$D(\mathbf{x},t) = D_0 \exp\{-\gamma\sigma_{vm}(\mathbf{x},t)\}, \tag{2.6}$$

$$\nabla \cdot \sigma_{vm} - \lambda \nabla N = 0. \tag{2.7}$$

Equation (2.5) represents the RD part, describing as before the spatiotemporal change in tumour cell number $N(\mathbf{x},t)$ as the sum of random cell diffusion (the first term on the right hand side of the equation) and logistic growth (the second term) – $k(\mathbf{x})$ is the spatially-varying growth rate, and θ is the carrying capacity. D is the apparent cell diffusion coefficient, this time linked to the surrounding tissue stiffness through equation (2.6) where σ_{vm} is the von Mises stress, D_0 is the diffusion of tumour cells in the absence of stress and γ is an empirically-derived mechanical coupling constant. Equation (2.7) establishes mechanical equilibrium (isotropic and linear elastic) subject to an expansive force due to the change in tumour cell number (the second term on the left hand side) with coupling constant λ.

As Garg and Miga demonstrated that the new MC–RD model [41] provided improved predictions of tumour cell number compared to the original RD formulation [35] in brain tumours, pioneering patient-specific work by Weis et al. [6] subsequently adapted these models to breast cancer and populated them with patient data (see Subsection 2.2.2). This enabled them to evaluate each model against actual pathological outcome, revealing that the MC–RD results were more accurate: the specificity was increased from 55% to 84%, though the sensitivity decreased from 100% to 92%.

2.1.4 Models of the vasculature and angiogenesis

Another key characteristic of cancer is angiogenesis, involving the formation of new blood vessels from existing ones. This process is vital for tumour growth, as it meets the increased demand for oxygen and nutrients necessary for cancer cell proliferation [42]; tumours typically cannot exceed 1-2 mm in diameter without developing new blood vessels [43]. In normal tissues, blood is supplied through a mature and orderly vascular network, on the other hand, tumour vasculature is characterised by disorganisation and dysfunction, resulting in irregularly shaped, tortuous, immature and leaky blood vessels [44]. Although there is an extensive body of work on the mathematical modelling of angiogenesis [45, 46], few studies address this phenomenon at the imaging or tissue scale, or personalise it for individual tumours [47]. Hormuth et al. [48] review the integration of time-resolved imaging data with mathematical modelling to systematically study angiogenesis from the cellular to tissue scale, with the ultimate goal of predicting treatment response.

Hormuth et al. [49] developed a murine model of glioma growth coupled with angiogenesis, where vasculature influences the tumour growth direction and tumour cells, in turn, drive vascular evolution. The vasculature was also linked to the carrying capacity. In contrast, Roque et al. [50] did not explicitly model vascular evolution but used vasculature-derived parameters to initialise model variables and calibrate key parameters.

Wu et al. [51] introduced a computational fluid dynamics model of blood flow in breast tumours, coupling 1D blood flow with 3D tissue flow. DWI and DCE-MRI were employed to identify a vasculature-interstitium geometry, which enabled patient-specific assessment of tumour-associated blood supply and interstitial transport characteristics. They demonstrated that malignant lesions exhibit greater interstitial flow velocity and higher heterogeneity in blood pressure and vascular extraction rate. The model was later used to identify an optimised NACT regimen for each patient, balancing treatment efficacy and toxicity [52].

2.1.5 Drug-Incorporated Mechanically-Coupled Reaction–Diffusion models (DI–MC–RD)

Jarrett et al. [4] extended the MC–RD to include the effect of spatially-varying chemotherapy drug delivery. Rather than explicitly modelling drug transport, a snapshot of drug concentration estimated from DCE-MRI is incorporated directly into the RD equation as an explicit cell death term:

$$\frac{\partial N(\mathbf{x},t)}{\partial t} = \nabla \cdot \Big(D(\mathbf{x},t)\nabla N(\mathbf{x},t)\Big) + k(\mathbf{x})\left(1 - \frac{N(\mathbf{x},t)}{\theta}\right) N(\mathbf{x},t) - \underbrace{\alpha C(\mathbf{x},t) N(\mathbf{x},t)}_{\text{Death}}, \tag{2.8}$$

where C represents the concentration of drug in the tissue and α is a globally-fitted parameter describing the efficacy of the drug. The mechanical coupling and mechanical equilibrium equations (equations (2.6) and (2.7)) are unchanged compared to the MC–RD model. When personalised using patient data (further discussed in Subsection 2.2.2), this drug incorporated MC-RD model (DI-MC-RD) was shown to outperform the earlier MC-RD model, with concordance correlation coefficient increased from 0.85 to 0.99. However, this was only tested on five patients.

2.1.6 Machine learning

Artificial intelligence (AI) broadly refers to the capability of machines to perform tasks autonomously without human intervention. Machine learning (ML), a subset of AI, is a general-purpose approach that identifies patterns and relationships in data without requiring them to be predefined [26]. This data-driven approach has gained widespread adoption across numerous disciplines, including cancer modelling. ML can leverage the numerous datasets available in cancer research to identify statistical correlations between biomarkers and patient survival [53], particularly in breast cancer [54]. However, when applied to patient-specific tumour models, ML faces a fundamental limitation in the lack of sufficient data for any individual patient [55]. Clinical imaging and biomarker data are typically collected at sparse time points, providing only infrequent snapshots of tumour evolution. This makes it impossible to train an ML model on

patient-specific data alone to produce reliable predictive outcomes [56].

One alternative is to train ML models on multi-patient datasets. However, this approach is complicated by the high heterogeneity of tumour imaging: differences in scanners, imaging protocols across institutions and, critically, the variability in tumour characteristics and patient response. ML models often struggle with so-called "out-of-distribution" cases: instances where a patient's data deviates significantly from the patterns observed in the training set. As a result, models trained on population-level data may struggle to capture the unique physiological characteristics of individual patients, potentially leading to inaccurate predictions [26, 55].

A promising avenue lies in hybrid methods, such as biologically-informed neural networks (BINNs), which offer frameworks for combining data-driven and mechanistic modelling [57]. This strategy has shown success in brain tumour modelling, demonstrating improved predictive accuracy on an *in silico* cohort compared to purely mechanistic methods [58]. Further discussions on the so-called "synergistic combination" of ML and mechanistic models can be found in [59].

Beyond direct tumour growth prediction, the most valuable role for ML may lie in complementary tasks. For instance, it can assist with model calibration by integrating population-level insights with individual patient data as in [60]. Additionally, ML techniques could help process mechanistic model outputs to optimise treatment selection [26, 55] and assist with imaging-related tasks such as tumour segmentation and image registration.

2.2 Model personalisation

The foregoing models provide a mathematical description of the hypothesised main physiological phenomena governing tumour dynamics. To be predictive of tumour evolution and treatment response in real subjects, the models must be personalised. Early ODE efforts described in Subsection 2.1.1 were severely limited by their model parameters, which were not based on patient data [61]. These parameters were hard

or impossible to evaluate *in vivo* thus were assigned values from the literature, often from different biological systems. The problem of unreliable parameter values was compounded by a heavy dependence of the model prediction on these values [61]. Without reliable data, biophysical theories of tumour growth and treatment response have very limited clinical utility [62].

Medical imaging has since tremendously improved, and quantitative MRI is presented as a noninvasive tool capable of estimating parameters such as the volume transfer rate (K^{trans}), from which tumour blood flow or vascular permeability can be calculated. These prove much better indicators of treatment response than standard RECIST measures [9]. It is crucial to inform models with data that can be acquired widely and not exclusively in research-oriented hospitals with state-of-the-art facilities. Dependence on data which cannot be obtained in the community hospitals where the majority of patients receive their care severely limits the relevance of a model, and reduces its potential significance to cancer care as a whole [2, 5, 9]. MRI is a strong candidate in this regard, as it is already a routine imaging modality for breast cancer patients. Moreover, it is the only imaging technique consistently used longitudinally throughout the course of treatment, enabling data acquisition at multiple time points as required by these modelling approaches.

2.2.1 Diffusion-weighted (DW-)MRI

Atuegwu et al. [63] were the first to incorporate DW-MRI data into a logistic model of breast tumour growth (equation (2.3)) in order to predict changes in the cellularity of breast tumours during NACT. They used DWI to calculate maps of the apparent diffusion coefficient (ADC) before the initiation of treatment and after one cycle of NACT:

$$ADC = \frac{1}{3} \sum_{i=x,y,z} \frac{\ln S_0/S_i}{b_i}, \tag{2.9}$$

where i is the diffusion-weighting direction, b_i is the amount of diffusion-weighting imparted to the sample, and S_0 and S_i are the measured signal in each voxel [3].

These in turn provide estimates of the number of cells within the tumour, as a strong negative correlation between ADC and cellularity has long been observed [31, 64–66]:

$$N(t) = \theta \left(\frac{ADC_w - ADC(t)}{ADC_w - ADC_{min}} \right), \tag{2.10}$$

where θ is the carrying capacity, ie. the total number of tumour cells that fit within a voxel. This was determined based on the volume of a tumour cell (assumed to be $4189\mu\text{m}^3$, based on a packing density of 74.05% and spherical tumour cells of radius 10 μm) and the volume of each voxel, which depends on MRI acquisition parameters. ADC_w is the ADC of free water at 37°C (3×10^{-3} mm^2/s), $ADC(t)$ is the ADC value measured at time t, and ADC_{min} the lowest observed ADC, which is assumed to correlate to the voxel containing the largest number of cells.

Maps of the tumour cell number, or tumour cellularity, derived in this way are used to parameterise several mathematical models [4, 6, 31, 67], which can then be run forward in time to predict cellularity at the conclusion of NACT. In [63] they were combined with a logistic model of tumour growth to estimate a voxel-wise proliferation rate $k(r)$, which could be negative to account for cell death or positive for proliferation.

2.2.2 Dynamic contrast-enhanced (DCE-)MRI

In the first instances of image-informed patient-specific tumour models, DCE-MRI was simply used to identify the location of the tumour [6, 13, 14, 31]; segmentation techniques to this aim are discussed in Subsection 2.3.1. Most recently, DCE-MRI data were used by Jarrett et al. [2] to describe voxel-wise drug delivery and spatially-varying vasculature (equation (2.8)). They used the extended Kety-Tofts model to estimate the concentration of the contrast agent (CA) in the tissue, and assumed that the drug would extravasate in the tissue in an identical manner:

$$C_{drug}^{tissue}(\mathbf{x}, t) = K^{trans} \int_0^t \left(C_{drug}^{plasma}(u) \exp\left\{ -\frac{K^{trans}}{\nu_e}(t-u) \right\} \right) du + \nu_p C_{drug}^{plasma}(\mathbf{x}, t), \tag{2.11}$$

where C_{drug}^{tissue} and C_{drug}^{plasma} are the concentrations of the drug in the tissue and in the plasma, respectively. K^{trans} is the volume transfer constant from the plasma to

the tissue and ν_p, ν_e are the volume fraction of the plasma and of the extravascular extracellular space respectively. This effect was incorporated directly into the RD equation (see Section 2.1, equation 2.8).

DCE-MRI can also be used to distinguish between different types of tissue in the breast. The MC–RD [6, 13, 14] and DI–MC–RD [2, 4] models classify fibroglandular and adipose (fat) tissue using a two-class $k-$means clustering algorithm. These tissue types are then assigned different Young's moduli to implement the mechanical coupling: 4 kPa for fibroglandular and 2 kPa for adipose tissue, compared to 20 kPa for the tumour tissue. These are population values, i.e. not patient-specific, based on magnetic resonance elastography (MRE) measurements in six breast cancer patients [68]. It would be difficult to acquire these in each patient as MRE is not part of routine imaging for breast cancer.

2.2.3 Parameter calibration methods

Some parameters cannot be reasonably informed by data, either because they cannot be acquired for individual patients or at all, meaning no reasonable population reference values exist in the literature. In such cases a so-called inverse problem is formulated to calibrate said parameters to an individual patient's dataset. Data from at least two time points are generally required; typically, an initial guess is made for the unknown parameters and refined iteratively in an optimisation process, until the model predictions from the first time point to the next align with the observed data. Lorenzo et al. [47] outline the general formulation of inverse parameter estimation and explore standard methods for solving it.

In the MC–RD [6, 13, 14] models a scalar apparent cell diffusion coefficient in the absence of stress is calibrated, D_0. The DI–MC–RD [2, 4] models also calibrate a scalar chemotherapy efficacy constant, α. Both MC–RD and DI–MC–RD models locally calibrate the proliferation rate $k(\mathbf{x})$, assigning a different scalar value at each voxel both within the tumour and within a large region of interest surrounding it.

The k values can be negative at voxels where tumour shrinkage occurs or positive for tumour cell proliferation. Depending on the size of the tumour and refinement of the computational grid the problem is solved on, this can amount to hundreds of calibrated k values.

2.3 Implementation techniques for model personalisation: data pre-processing and numerical methods

This section covers the practical implementation of mechanistic tumour models. We begin with the essential image preprocessing steps of tumour segmentation and image registration. We then discuss the numerical techniques required to solve the models introduced in Sections 2.1 and 2.2, focusing on the finite difference and finite element methods.

2.3.1 Segmentation

Tumour segmentation is an active and extensively studied field. Here, the focus is on techniques specifically used for the image-informed breast tumour models discussed above, with a brief mention of promising ML/AI techniques.

Methods used in image-informed breast tumour models DCE-MRI typically involves acquiring one or more images before contrast agent (CA) injection, followed by a series of images during and after the injection. These postcontrast images capture the CA's uptake and washout dynamics. The CA accumulates in the tumour, making it appear brighter in postcontrast images. A common approach to highlight the enhanced tissue is to subtract the precontrast image (or an averaged precontrast image set, if multiple are available) from the brightest postcontrast image. The resulting subtraction image provides a clear distinction between the tumour and surrounding tissue.

A straightforward approach to tumour segmentation is to simply apply a SI enhancement threshold based on CA uptake: voxels exhibiting enhancement above this

threshold are classified as tumour, while those below it are considered non-tumour. In the early days of image-informed tumour modelling, Atuegwu et al. [63] used an empirical threshold of a 40% SI increase between a single precontrast and postcontrast image. Weis et al. [6, 13, 14] refined this method by computing a subtraction image between the average of three precontrast images and the average of 22 postcontrast images. They adopted an 80% SI increase as their threshold, again based on empirical observations. This threshold was later validated to show a strong correlation with tumour volume measurements from excised tissue obtained during surgery [69].

Jarrett et al. [4] later adapted a method inspired by Stadlbauer et al. [70]. Their approach begins by averaging the precontrast images and the three consecutive postcontrast images showing the highest enhancement. A subtraction image is then computed by subtracting the precontrast average from the postcontrast average. To account for background variability, they also compute the average of the final six postcontrast images and measure the standard deviation of signal intensity (SI) enhancement in the breast which doesn't contain the tumour, i.e. in purely healthy tissue. A conservatively sized bounding box is drawn around the tumour and, within this region, voxels are classified as tumour if their percentage enhancement exceeds twice the standard deviation of healthy tissue enhancement [71].

Another approach adapts a fuzzy c-means algorithm developed in [72] to classify voxels based on probabilistic membership in the tumour ROI. This method has been employed in [5, 51, 73, 74].

Other promising methods AI and ML have demonstrated significant promise in automating breast tumour segmentation in MRI. Deep learning techniques, particularly convolutional neural networks (CNNs), excel in complex image analysis tasks due to their ability to learn intricate patterns from large datasets. In breast MRI, CNN-based models have been developed to automatically detect and segment tumours, achieving high accuracy and efficiency [75–77]. Deep learning models have also been trained to perform fully automated tumour segmentation on DCE-MRI images of triple-negative

breast cancer, achieving a Dice similarity coefficient of 93% and a sensitivity of 96% in independent testing datasets [78]. Some studies extend beyond segmentation to lesion detection: Zhang et al. [79] implemented a deep learning method using the Mask R-CNN algorithm to detect abnormal lesions in breast MRI, reaching an accuracy of 0.86 on a per-slice basis in the training dataset. These AI-driven approaches often outperform traditional methods by leveraging large annotated datasets to learn complex tumour characteristics, thereby improving robustness and accuracy across diverse imaging conditions. The field continues to evolve, with ongoing research exploring various supervised, unsupervised and hybrid models to enhance segmentation precision and clinical applicability [80].

Lastly, image spatial resolution significantly affects segmentation quality for modelling [70]. The slice thickness is typically 3–5 times larger than the in-plane resolution, introducing uncertainty in the $z-$axis segmentation. Even small segmentation errors can substantially impact the estimated tumour size and volume, which can be a problem for modelling.

2.3.2 Registration techniques

Image-informed modelling requires images acquired at different times for model calibration and prediction evaluation. When patients are scanned multiple times on different days, their positioning within the scanner may vary slightly. Therefore, all images of a given patient's tumour must be registered to a shared coordinate system to facilitate image-based modelling. This process ensures that the voxel located at coordinates (x, y, z) corresponds to the same anatomical site across all scans of that patient.

The *elastix* manual [81] provides an introduction to image registration and an overview of some common techniques. A brief explanation of registration is provided here. Two images are involved: the moving image $I_M(\mathbf{x})$, which is deformed to align with the fixed image $I_F(\mathbf{x})$. Both images have dimension d and are each defined on their respective spatial domains, $\Omega_F \subset \mathbb{R}^d$ and $\Omega_M \subset \mathbb{R}^d$. The goal of registration is to find

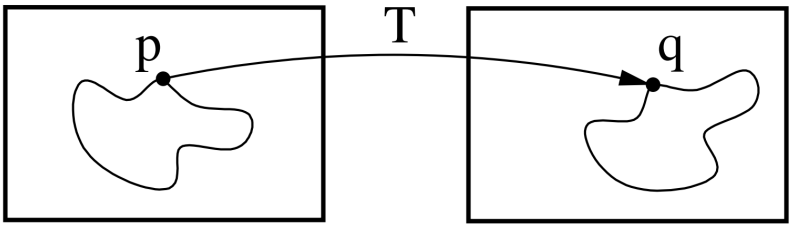


Figure 2.1: With the fixed image I_F on the left and the moving image I_M on the right, image registration is the task of finding a spatial transformation T mapping one image to another. From [81].

a displacement field $u(\mathbf{x})$ such that $I_M(\mathbf{x}+\mathbf{u}(\mathbf{x}))$ aligns with $I_F(\mathbf{x})$. Equivalently, this can be expressed as finding a transformation $\mathbf{T}(\mathbf{x}) = \mathbf{x}+\mathbf{u}(\mathbf{x})$ mapping the fixed image domain to the moving image domain:

$$T : \Omega_F \subset \mathbb{R}^d \to \Omega_M \subset \mathbb{R}^d.$$

The alignment quality is quantified by a similarity measure S, such as the sum of squared differences, correlation ratio or mutual information. Since nonrigid transformations T make the problem ill-posed, a regularisation or penalty term P is often introduced to constrain the transformation. Thus, registration is typically formulated as an optimisation problem, where a cost function $\mathcal{C}$ is minimised with respect to T:

$$\hat{\mathbf{T}} = \arg\min_T \mathcal{C}(\mathbf{T}; I_F, I_M), \tag{2.12}$$

$$\mathcal{C}(\mathbf{T}; I_F, I_M) = -S(\mathbf{T}; I_F, I_M) + \gamma P(\mathbf{T}), \tag{2.13}$$

where γ weighs similarity against regularity [81].

Like segmentation, registration is a well-established field of research; it would be beyond the scope of this work to provide an extensive review of all available methods. Instead, we will focus on the techniques employed in the models mentioned earlier.

In the context of breast imaging, registration poses particular challenges due to the ductile and easily deformable nature of breast tissue. In contrast to brain tumours,

where the skull provides a consistent shape which can be used for guidance, the outline of the breast can be deformed or compressed during MRI scans resulting in varying appearances across different imaging sessions. When the entire breast domain is required for modelling, a deformable registration algorithm is necessary to account for these changes. However, it is crucial to maintain the integrity of the tumour itself during this process. If the tumour is deformed, valuable information regarding its size at different scans may be lost, particularly if it has shrunk; the deformable registration could inadvertently stretch it to match previous time points. To address this, the image-informed breast tumour models which we have reviewed above employ a deformable registration with a "rigidity penalty" on the tumour [2, 4–6, 13, 14, 51, 63, 67, 74, 82]. Developed by Li et al. [83, 84], this approach allows for the registration of surrounding tissue while preserving the tumour's original shape.

Lastly, while intuitive, it is worth emphasising that registration accuracy heavily depends on a precise tumour segmentation. Even small errors, such as missing or adding a few voxels, can significantly alter the tumour's shape at one timepoint, making it difficult to align with its differently shaped counterpart at another timepoint.

2.3.3 Numerical methods

Mechanistic tumour models typically consist of coupled, nonlinear PDEs. There are numerous techniques and approaches for solving these PDEs. Here we provide a brief overview of the methods employed or mentioned later in the thesis: the finite difference method (FDM) and the finite element method (FEM).

FDM The finite difference method is one of the simplest numerical techniques for solving PDEs. It involves discretising the continuous domain into a grid or mesh and approximating the derivatives in the equations by replacing them with difference equations, by means of Taylor series expansions [47]. Derivatives at a point t are approximated using values of the function at nearby points. For example, a time-dependent first-order derivative $\frac{df}{dt}$ can be approximated using a central difference scheme, where

h describes the distance from point t to its neighbours $t + h$ and $t - h$:

$$\frac{df}{dt} \approx \frac{f(t+h) - f(t-h)}{2h}, \tag{2.14}$$

a forward scheme:

$$\frac{df}{dt} \approx \frac{f(t+h) - f(t)}{h}, \tag{2.15}$$

or a backward scheme:

$$\frac{df}{dt} \approx \frac{f(t) - f(t-h)}{h}. \tag{2.16}$$

Similarly, spatial derivatives are approximated by using values at neighbouring grid points.

In a typical implementation, the computational domain is divided into a mesh of grid points, and the solution is iteratively updated over time. FDM is particularly advantageous for problems with simple geometries, where a grid structure can be easily defined. Thus it can be well-suited to MRI grids. However, the accuracy of the method is heavily dependent on the grid resolution and the choice of finite difference approximations [85].

FEM In contrast, the finite element method is a more versatile approach, especially well-suited for solving PDEs in complex geometries. FEM involves dividing the computational domain into a collection of smaller, simpler regions called elements, which can take various shapes (e.g., triangles or quadrilaterals in 2D, tetrahedra or hexahedra in 3D). The solution is approximated by constructing piecewise polynomial functions over these elements [86].

The core idea of FEM is to derive a "weak" formulation of the PDE, also known as its variational form, where the equation has solutions only for certain "test functions". The variational problem is the form in which the problem is solved. The steps required to transform a second order PDE into a variational problem are as follows:

1. Multiply the PDE by a "test function" v;

2. Integrate the resulting equation over the domain;

3. Perform integration by parts on the terms containing second-order derivatives.

FEM provides several advantages, including its ability to handle irregular geometries, varying material properties and complex boundary conditions. Additionally, it allows for local refinement of the mesh, enabling higher accuracy in areas where the solution exhibits rapid changes, such as near tumour boundaries.

2.4 Conclusions

To conclude, the foregoing personalised models of tumour growth suggest that it is possible, after only one cycle of NACT, to evaluate whether a patient will respond positively and it is worth pursuing the treatment, or whether the patient is a non-responder and alternative therapy regimens should be sought out. The current state-of-the-art [2] shows that the spatial variation of "fluid" transport within the tumour, manifested in that instance as variation of drug concentration, affects the overall tumour response.

The conclusions from the above literature review are as follows:

- Discrete models of tumour growth are unable to consider enough cells to accurately represent a patient's tumour, due to large computational expense. Continuum models are better suited to clinically-relevant tumour growth modelling.
- Continuum models based on PDEs can be used to model spatio-temporal tumour evolution. Among these, those based on the reaction–diffusion equation are best suited to personalisation due to their dependence on parameters which can feasibly be acquired using routine MRI.
- The diffusion of tumour cells is restricted *in vitro* by the healthy tissue which surrounds it. Taking this into account in modelling via mechanically coupling the tissue stiffness to the diffusion coefficient improves model predictions.
- Current models based on the RD equation often rely heavily on numerous calibrated parameters (such as the locally-calibrated proliferation rates $k(\mathbf{x})$ in [4]),

which can overshadow the influence of the underlying mathematics on predictions. Reducing the number of calibrated parameters may allow for a deeper exploration of the mathematical framework, helping to identify the most influential terms in the solution and ensuring that all included terms are relevant.

- Tumour segmentation and image registration are important steps in the modelling process, and can be difficult to achieve in breast MR. Image slice thickness can introduce errors in segmentation, and the ductile nature of breast tissue complicates registration.

- FDM is generally easier to implement, computationally less intensive and can be applied directly to the MR voxel grid. However, it has limitations in handling complex shapes which may compromise its accuracy near boundaries. While FEM may be more computationally expensive, it offers greater accuracy and robustness as a result, and allows for local refinement of the mesh in areas of interest.

Chapter 3

CHERNAC data

Contents

3.1 Introduction

The CHERNAC study [87] acquired DW-MRI and DCE-MRI data for 37 patients with breast cancer before, during and after a course of NACT. This chapter describes the preprocessing of these images, i.e. the steps required to obtain model input data from them, for the models described in Chapter 5. The required data are spatially-varying maps of the tumour cell volume fraction and of the concentration of chemotherapeutic drug evolving in the tumour tissue over several days following its delivery.

3.2 Patient details

40 female patients with newly diagnosed primary invasive carcinoma of the breast due to undergo NACT were originally recruited to the study, three patients withdrew. Patients who had received previous treatment for breast cancer or had recurrent breast cancer, impaired kidney function or contraindications to MRI were excluded.

3.3 Treatment and scanning schedule

All recruited patients were treated with 90 mg/m^2 epirubicin and 600 mg/m^2 cyclophosphamide (EC90) for the first three cycles of chemotherapy, administered every three weeks. The treatment plan aimed for patients to continue with docetaxel (100 mg/m^2) for an additional three cycles (Figure 3.1). Several patients transitioned to paclitaxel or albumin-bound paclitaxel after just one cycle of docetaxel due to side effects (Figure 3.2). For patients with HER2-positive tumours, docetaxel was combined with trastuzumab and pertuzumab (Figure 3.3). One patient received six cycles of EC90 alone (Figure 3.4). All patients were scanned before the start of chemotherapy ("Scan 0"), after one cycle and after three cycles of EC90 ("Scan 1" and "Scan 3"), and after their last cycle (patients received either six or seven cycles in total).

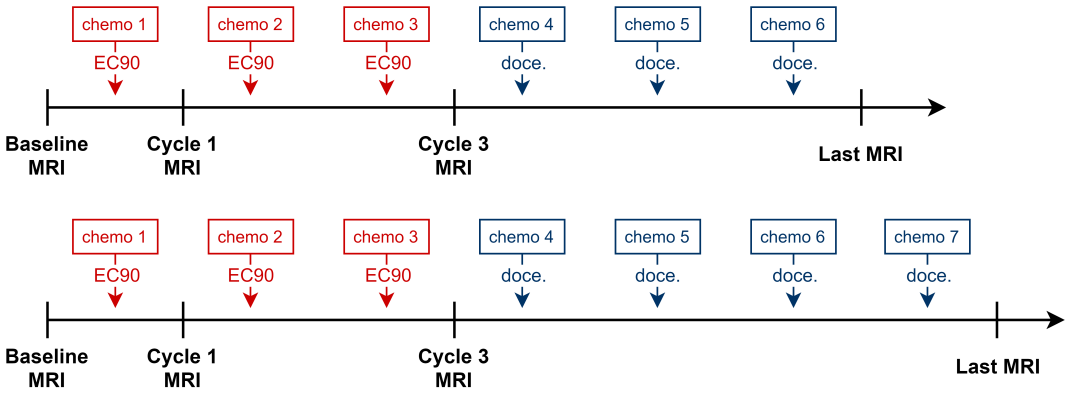


Figure 3.1: Treatment plan received by $n = 27$ patients in the study: three cycles of EC90 followed by three ($n = 25$) or four ($n = 2$) cycles of docetaxel. One patient in this group received an additional cycle of trastuzumab after the third docetaxel. MRI scans were taken before treatment, at two timepoints during and after the EC90 cycles and at the end of treatment.

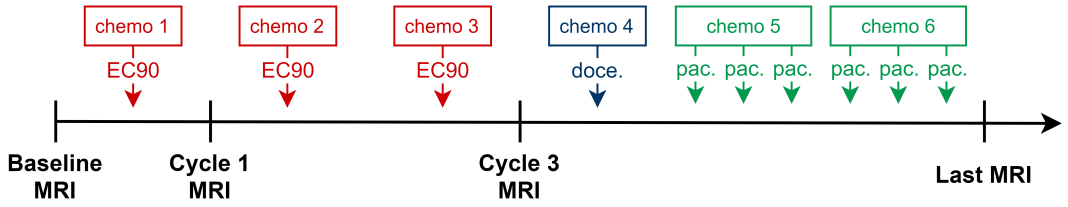


Figure 3.2: Treatment plan received by $n = 4$ patients in the study: three cycles of EC90 followed by one cycle of docetaxel, discontinued due to side effects. For these four patients the docetaxel was replaced by two cycles of weekly paclitaxel, for one further patient the treatment was interrupted completely. MRI scans were taken before treatment, during and after three cycles of EC90 and at the end of treatment.

3.4 MRI data acquisition

The MRI data were acquired several years ago as part of the CHERNAC study [87]. The details of their acquisition are provided here, for completeness.

Blood samples obtained as part of normal standard of care were used to estimate

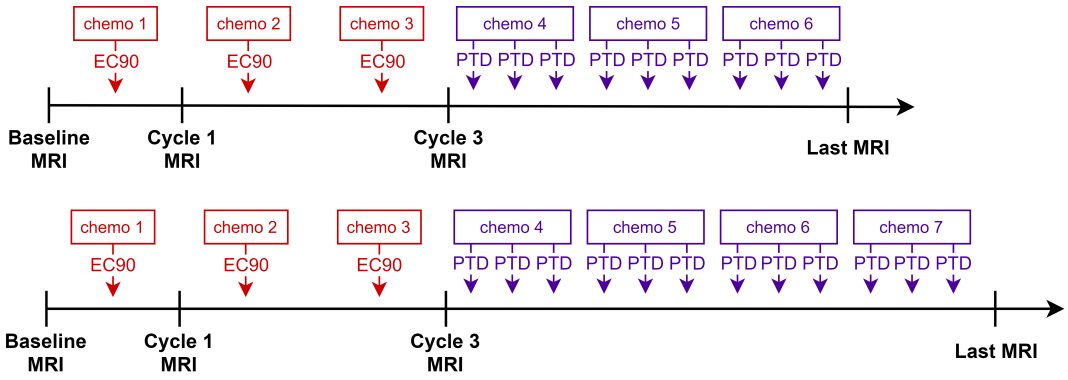


Figure 3.3: Treatment plan received by $n = 4$ patients in the study: three cycles of EC90 followed by three ($n = 1$) or four ($n = 3$) cycles of pertuzumab, trastuzumab and docetaxel (PTD). MRI scans were taken before treatment, during and after three cycles of EC90 and at the end of treatment.

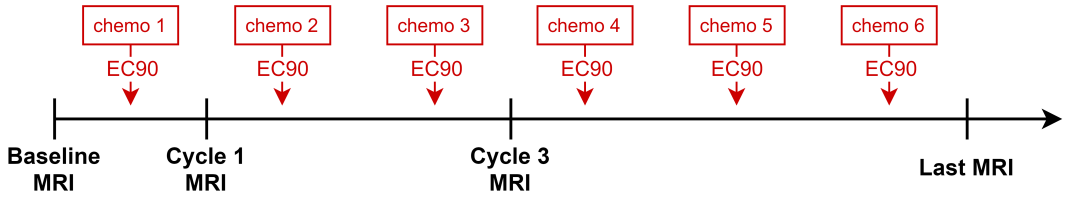


Figure 3.4: Treatment plan received by one patient: six cycles of EC90 (epirubicin (90 mg/m^2) and cyclophosphamide (600 mg/m^2)). MRI scans were taken before treatment and after the first, third and sixth cycles of EC90.

the large vessel haematocrit (Hct) before each scan. All imaging was performed using a 1.5 T MR scanner (Aera; Siemens) with the patients positioned head-first and prone. A dedicated 16-channel bilateral breast coil (Sentinelle; Siemens) was used for signal detection and an additional flexible array coil was placed on the patients' backs to increase the signal from the descending aorta [88].

The MRI protocol included axial T_1-weighted 3D spoiled gradient echo (FLASH), axial T_2-weighted turbo spin echo, DWI, and DCE series. Before the dynamic acquisition,

axial DWI was performed using a spectral attenuated inversion-recovery fat-suppressed, 2D single-shot spin echo-echo planar imaging sequence at six b-values (0, 50, 100, 200, 400 and 800 s/mm^2) with the following parameters: TR/TE: 7200/59 msec, field of view (FOV): 340×136 mm^2, matrix size: 280×116, slice thickness: 4 mm, and GRAPPA parallel factor: 2. The acquisition time was 5 minutes and 31 seconds. Maps of the apparent diffusion coefficient (ADC) were created by the scanner software immediately after the DWI acquisition.

The DCE series consisted of 93 high temporal resolution (HTR) volumes interleaved with eight high spatial resolution (HSR) volumes in the following manner: 10 HTR, 1 HSR, 43 HTR, [1 HSR, 5 HTR] repeated seven times, 5 HTR [89]. The HTR data were obtained with a T_1-weighted 3D spoiled gradient echo sequence (TR/TE: 2.37/0.73 ms, FOV: $340 \times 340 \times 180$ mm^3, matrix size: $128 \times 128 \times 36$, flip angle: 25°, CAIPIRINHA parallel factor: 2×2, acquisition time: 2 s). For the HSR data, a fat-suppressed T_1 weighted 3D FLASH sequence (TR/TE: 4.1/1.2 ms, matrix size $384 \times 384 \times 128$, flip angle: 10°, GRAPPA parallel factor: 3, acquisition time: 36 s) was employed. The volume selected encompassed both breasts, the aortic arch and part of the descending aorta. A dose of 0.1 mmol/kg of the contrast agent Gd-DOTA (Dotarem; Guerbet Laboratories) was administered intravenously via the antecubital vein using an automated power injector (Spectris Solaris EP; Bayer) at the beginning of the eleventh dynamic volume acquisition, followed by 20 mL of saline at a rate of 3 mL/s.

3.5 Pre-processing

A chronological overview of the pre-processing which transformed the raw DICOM files into `.mat` model inputs is provided here and expanded on in each subsection. The steps required were:

1. The location of the tumour, referred to hereafter as the region of interest (ROI), was identified in the scan 0, 1 and 3 HSR and HTR image volumes by a breast

radiologist and provided as a MATLAB (MathWorks, [90]) file registered to the ADC data.

2. The HTR DICOM time series were used in the software PMI (Platform for Research in Medical Imaging, version 0.4) [91] to calculate spatially-varying maps of K^{trans} and v_e at the first two timepoints, scans 0 and 1 (Subsection 3.5.3);

3. K^{trans} and v_e were upsampled and cropped to match the ADC (and ROI) image resolution and field of view in MATLAB (Subsection 3.5.2);

4. When the HSR and HTR ROIs were not coincident (if the patient moved during the longer HTR acquisition), the latter were registered to the former using an *elastix* [81] rigid registration implemented in MATLAB. The resulting transformation was applied to the K^{trans} and v_e image volumes using *transformix* [81]. This concludes the intravisit image alignment.

5. All datasets acquired during or after chemotherapy were registered to the pre-treatment dataset using a rigid registration of the ROIs (Subsection 3.5.2);

6. The model data were extracted from the registered image sets: maps of the tumour volume fraction V_{TC} were derived from the ADC (Subsection 3.5.4) and the concentration of chemotherapy drug in the tumour tissue over several hours post-NACT infusion was estimated from K^{trans} and v_e (Subsection 3.5.3).

3.5.1 Tumour segmentation

The location of the tumour was identified on the subtraction of the HSR precontrast image from the first HSR postcontrast image by a breast radiologist. A tumour ROI was generated by a radiographer using a three-dimensional region-growing algorithm based on a signal intensity threshold [92]. Each enhanced voxel in the subtraction image was compared to its equivalent voxel in the ADC image, to distinguish tumour (dark in ADC) from fluid (bright in ADC) and ensure only tumour voxels were added to the ROI. This was provided as a binary mask (`.mat`) of the ADC data.

3.5.2 Registration

All images of a given patient's tumour must be registered to a shared coordinate system to enable image-based modelling. This ensures that the voxel located at coordinates (x, y, z) represents the same anatomical site across all scans of that patient. As the surrounding breast tissue is not included in the models described in Chapter 5, we only require the tumour ROIs to be registered and not the entire breast – this greatly simplifies the process. Registration is required both intravisit (aligning all images acquired at a given timepoint) and intervisit (aligning the images acquired at timepoints during and after chemotherapy to the pretreatment images). This was carried out in MATLAB.

The process of intravisit registration was relatively straightforward and is described in Figure 3.5. If the patient remained stationary during a single scanning session, the images acquired at different timepoints were already inherently in a common anatomical reference frame. In those cases, all that remained was to reorder the slices by anatomical location rather than acquisition time, then resample and crop the images to a common resolution and FOV. The DCE images (K^{trans} and v_e, see Subsection 3.5.3 for how they were obtained) were upsampled to the ADC (and ROI) resolution and cropped to the narrower ADC FOV to obtain a matching dataset with voxel-to-voxel equivalence. In the few cases where the patient moved during the HTR acquisition, a motion-corrected ROI mask was also provided for the HTR data. This was rigidly registered to the HSR mask using *elastix* [81], and the resulting transformation was applied to the K^{trans} and v_e images using *transformix* [81] prior to their resampling and cropping.

Once the intravisit registration was completed for each timepoint, the Scan 1 and 3 ROIs were rigidly registered to the pretreatment ROI and the resulting transformations applied to all other images.

The Dice similarity coefficient (DSC) quantifies the quality of the registration by assessing the overlap between the tumour volumes v across different scans, e.g. between

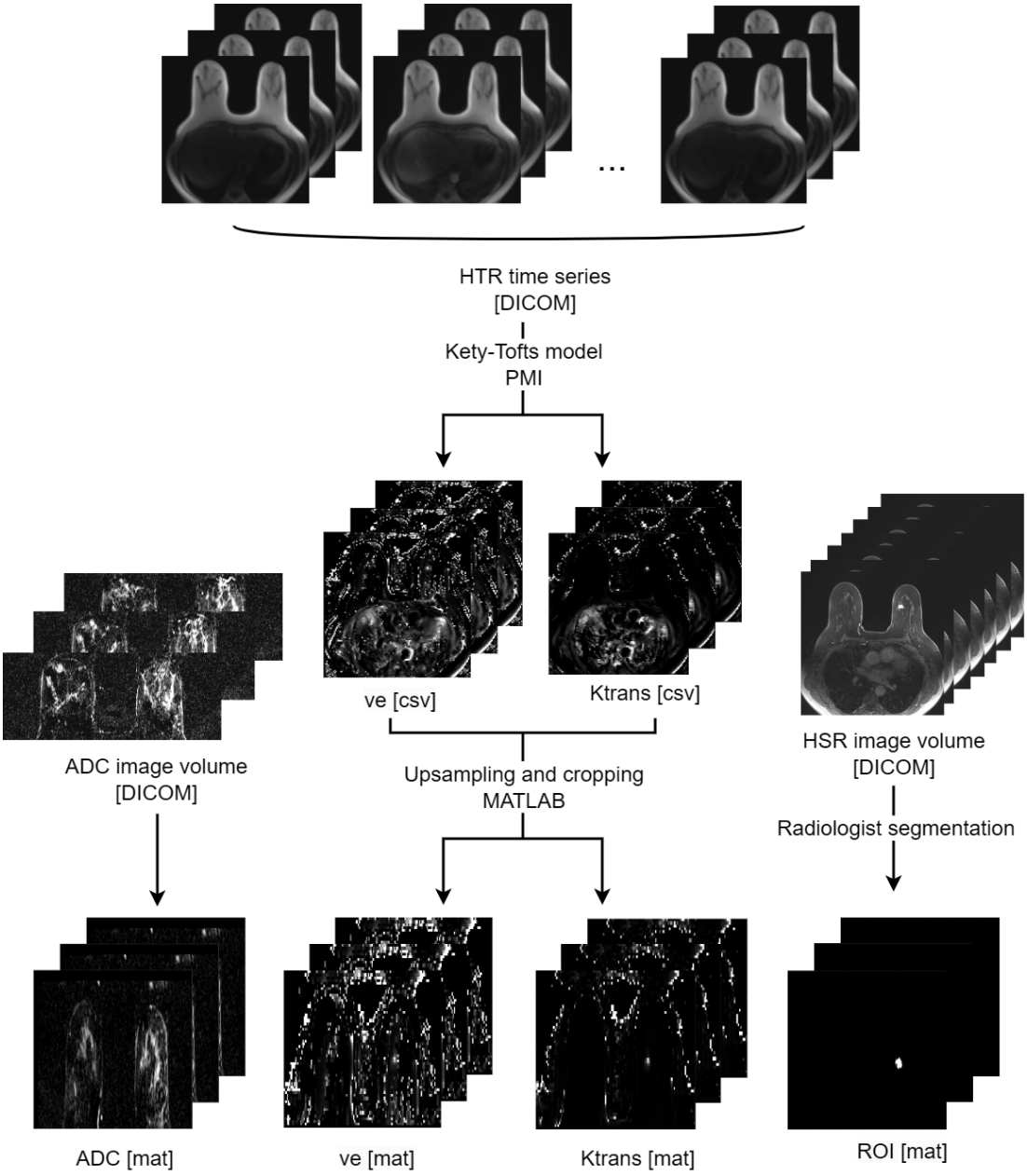


Figure 3.5: Forming a matching dataset of `.mat` files from all DICOM images acquired at a given timepoint. The HTR DICOM images were used to derive spatial maps of K^{trans} and v_e using the software PMI. The tumour ROI was identified on the HSR and HTR image volumes by a breast radiologist and provided as a `.mat` file registered to the ADC data. In cases where the HSR and HTR ROIs were not coincident, the HTR data were registered to the HSR with a rigid registration in *elastix*. In MATLAB, K^{trans} and v_e were upsampled and cropped to match the ADC (and ROI) image resolution and field of view.

Scan 1 and Scan 3:

$$\text{DSC} = \frac{2(\text{v}_{S1} \cap \text{v}_{S3})}{\text{v}_{S1} + \text{v}_{S3}}. \tag{3.1}$$

Table 3.1 presents these values for the overlap between scans 0 and 1 as well as between scans 1 and 3 – the higher the coefficient, the greater the overlap. Note that the DSC will be low if there is a substantial change in volume between two scans, but this does not necessarily indicate poor registration in these cases.

Only patient 28 was deemed to have a failed registration (shown in Figure 3.6) and removed from the study. Patient 19 (also in Figure 3.6) is an example of a subpar registration that was kept in the study.

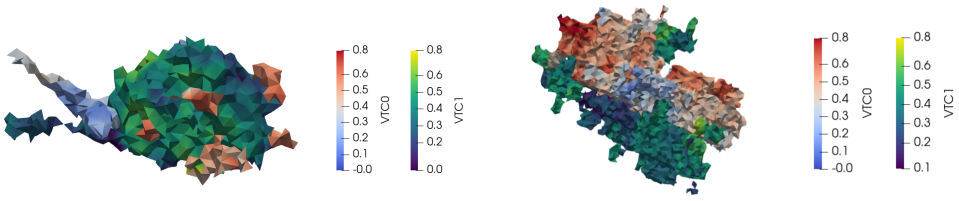


Figure 3.6: Examples of subpar or poor registration between Scan 0 and Scan 1 tumours. For patient 19 (left) the "tail" is clearly misaligned, but this is only a small percentage of the volume. The registration was deemed adequate. For patient 28 (right) the two tumours barely overlap (the volume of the intersection is only 16% of the Scan 0 tumour volume and 29% of the Scan 1 tumour volume), and the DSC is very low (0.2).

3.5.3 Drug concentration data

The HTR DICOM images were processed in PMI to estimate the concentration of chemotherapy drug in each tumour tissue voxel. The images were acquired with a large field of view which includes the aorta.

Estimating K^{trans} and v_e A ROI in the descending aorta was drawn on the 5th to the 14th slice in all cases (where it is usually most visible and to ensure consistency across patients) and trimmed to include only voxels with a peak relative signal intensity

ID	$DSC_{0\cap1}$	$DSC_{1\cap3}$
1	0.65	0.38
3	0.71	0.66
4	0.78	0.65
5	0.77	0.30
6	0.52	0.41
7	0.22	*
8	0.67	0.36
9	0.51	0.49
10	0.73	0.65
11	0.76	0.62
13	0.53	0.77
14	0.49	0.40
15	0.60	0.38
16	0.72	0.52
17	0.66	0.12
18	0.53	0.50
19	0.61	0.66
20	0.50	0.56
21	0.74	0.50
23	0.50	*
24	0.66	0.63
25	0.64	0.43
26	0.67	0.51
28	0.20	0.46
29	0.58	0.24
30	0.45	0.44
31	0.87	0.64
32	0.68	0.62
33	0.51	0.44
34	0.51	0.70
35	0.81	0.39
36	0.26	0.33
37	0.76	0.53
38	0.54	0.50
39	0.68	0.60

Table 3.1: Dice similarity coefficients (DSC) for tumour volumes at scans 0 and 1 ($DSC_{0\cap1}$) and for tumour volumes at scans 1 and 3 ($DSC_{0\cap1}$). Patients 7 and 23 (*) have no tumour present on their third scan.

(SI) change between the 50th and the 95th percentile of the signal change maxima. This thresholding was chosen to minimise inflow, section profile and partial volume artifacts [88]; an example of such a ROI is shown in Figure 3.7. SI time curves were generated from the aorta ROI. The contrast agent (CA) concentration in plasma (C_{CA}^{plasma}) at each voxel was deduced from the CA concentration in whole blood (C_{CA}^{blood}), approximated as proportional to the change in SI induced by its passage, ΔSI, over the SI measured before its injection, SI_0:

$$C_{CA}^{blood} = \frac{1}{r_1} \cdot \frac{\Delta \mathrm{SI}}{\mathrm{SI}_0}, \tag{3.2}$$

$$C_{CA}^{plasma} = \frac{1}{1 - \mathrm{Hct}} \cdot C_{CA}^{blood}. \tag{3.3}$$

The constant of proportionality in equation (3.2) is known as the CA relaxivity. For Gd-DOTA this is $r_1 = 4.2\ \mathrm{s}^{-1}\mathrm{mM}^{-1}$ in whole blood [93]. The change in CA concentration over time is known as the arterial input function (AIF), shown in Figure 3.9 for two patients.

CA concentrations were used with a modified Tofts model (extended Kety–Tofts) to fit $K^{trans}(\mathbf{x})$ (the volume transfer constant of contrast agent from the plasma to the tissue space at position $\mathbf{x}$), $v_e(\mathbf{x})$ (the extravascular extracellular volume fraction at position $\mathbf{x}$) and $v_p(\mathbf{x})$ (the plasma volume fraction at position $\mathbf{x}$):

$$\begin{aligned} C_{CA}^{blood}(\mathbf{x}, t) = K^{trans}(\mathbf{x}) \int_0^t C_{CA}^{plasma}(u) \cdot \exp\left(-\frac{K^{trans}(\mathbf{x})}{v_e(\mathbf{x})}(t-u)\right) du \\ + v_p(\mathbf{x}) C_{CA}^{plasma}(t). \end{aligned} \tag{3.4}$$

Figure 3.8 shows the mean K^{trans} estimated from scans 0 and 1 for each patient.

Concentration of drug in plasma A negative exponential decay law was used to estimate the concentration of each EC90 drug in the patient's plasma immediately following its infusion:

$$C_i^{plasma}(t) = C_i^{\max} e^{-0.693 t / t_i^h}, \tag{3.5}$$

where $C_i^{\max}$ is the maximum concentration and t_i^h the half life of drug i. These are listed in Table 3.2. The plasma concentrations were calculated in each case over the

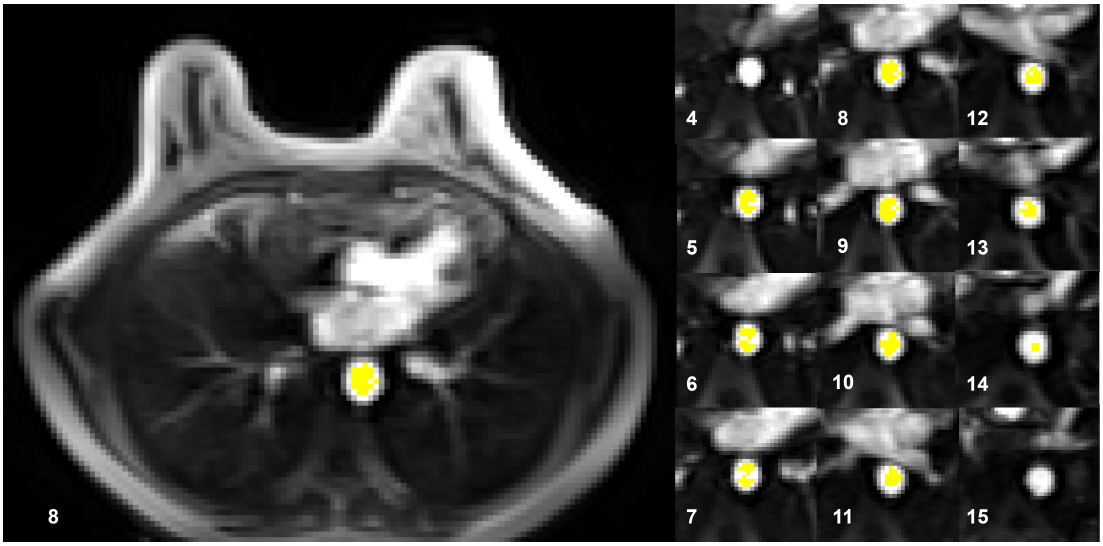


Figure 3.7: Example of an aorta region of interest used to calculate the arterial input function in PMI. The ROI (yellow) was drawn in the descending aorta on slices 5–14 of the HTR time series, and trimmed to include only voxels with a peak relative SI change between the 50th and 95th percentile of the signal change maxima. The left-hand side shows the ROI on the full FOV for slice 8, and the smaller images on the right show slices 4–15, with the ROI drawn on slices 5–14.

length of time required for the concentration to reach 1% of its starting value $C^{\max}$. These are shown in Figure 3.10.

	Half-life t_h (h)	Maximum concentration C_{max} (ng/mL)
Epirubicin	22.4 [94]	3486 [94]
Cyclophosphamide	5 [95]	61320 [95]

Table 3.2: Drug half-life (h) and maximum concentration in plasma (ng/mL) for epirubicin and cyclophosphamide, which all patients received for the first three cycles of their treatment.

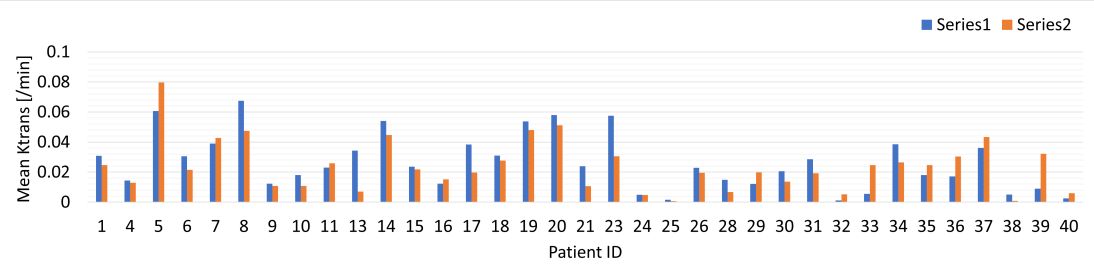


Figure 3.8: Mean K^{trans} estimated at each tumour voxel from scans 0 and 1 for each patient. Patient 3 was removed for visualisation purposes, due to an unusually high mean K^{trans} at Scan 0 (0.37/min). The Scan 1 value was normal (0.02/min).

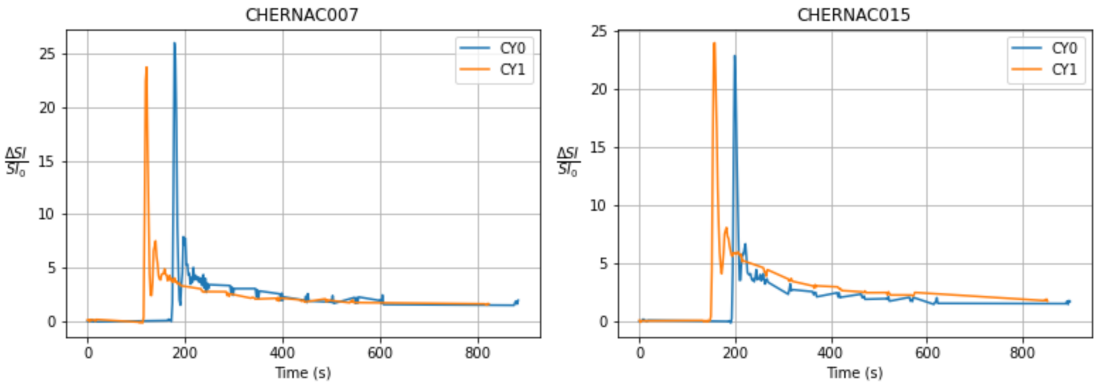


Figure 3.9: Arterial input functions calculated at Scan 0 (blue) and at Scan 1 (orange) for patients 7 and 15.

Concentration of drug in tissue Using the spatial maps of K^{trans} and v_e derived from equation (3.4) and plasma drug concentrations from equation (3.5), the concentration of each drug in the tumour tissue was estimated from the Kety-Tofts model:

$$C_i^{tissue}(\mathbf{x},t) = K^{trans}(\mathbf{x}) \int_0^t C_i^{plasma}(u) \cdot \exp\left(-\frac{K^{trans}(\mathbf{x})}{v_e(\mathbf{x})}(t-u)\right) du, \tag{3.6}$$

for $i \in \{\text{epirubicin (E)}, \text{cyclophosphamide (C)}\}$. Compared to equation (3.4), the term $v_p(\mathbf{x})C^{CA}_{plasma}(t)$ was dropped as it was consistently almost zero over the tumour voxels, so would have no measurable impact on the overall concentration. Furthermore, it represents drug delivery to the blood plasma – i.e., not to the tumour tissue – which is not of interest in this context.

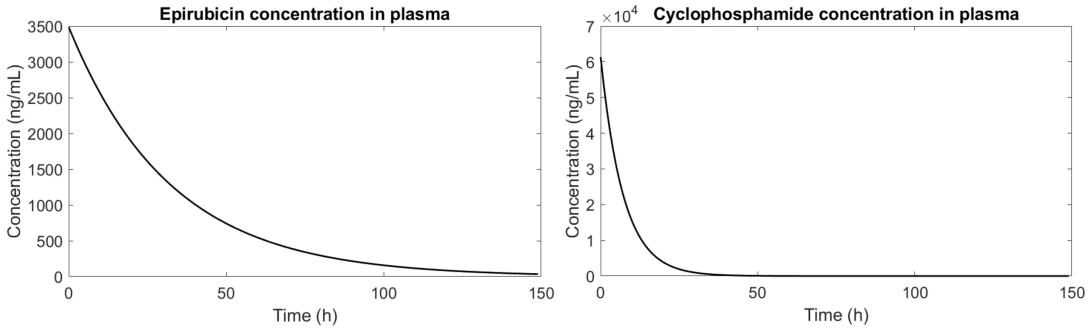


Figure 3.10: Concentration of epirubicin (left) and cyclophosphamide (right) in blood plasma over 150 h immediately following their infusion from equation (3.5). This is the duration required for both of them to drop below 1% of their peak concentration, and the values provided are the mean concentration over all voxels over this duration.

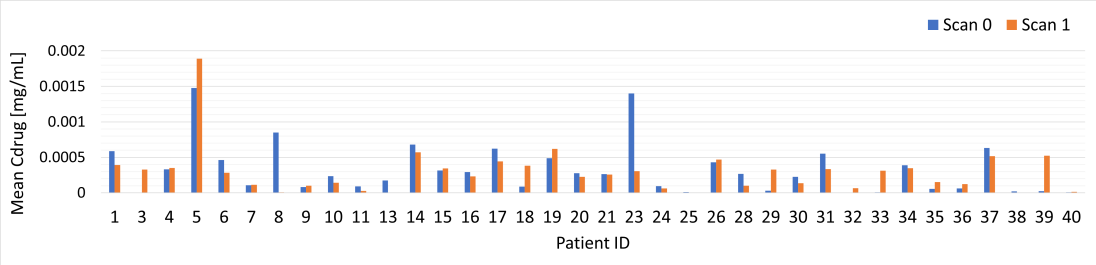


Figure 3.11: Mean drug concentration in tissue averaged across all tumour voxels and chemotherapy timesteps (i.e. from infusion until the concentration falls to less than 1% of its maximum), estimated from scans 0 and 1 in each patient.

From the data collected at scans 0 and 1, the total concentration of EC90 chemotherapy in the tumour tissue was calculated from the relative weight and concentration of each drug at each timepoint:

$$C_{drug}^{tissue}(\mathbf{x},t) = \omega_E C_E^{tissue}(\mathbf{x},t) + \omega_C C_C^{tissue}(\mathbf{x},t), \tag{3.7}$$

where $\omega_E = 90/690$ and $\omega_C = 600/690$ reflect their relative doses. Figure 3.11 shows the mean C_{drug}^{tissue} estimated at scans 0 and 1.

3.5.4 Tumour cell volume fraction

At each timepoint, a tumour cell volume fraction V_{TC} was estimated at each voxel from the normalised ADC maps [65]:

$$V_{TC}(\mathbf{x}) = 0.75 \left(1 - \frac{ADC(\mathbf{x}) - ADC_{min}}{ADC_{max} - ADC_{min}}\right), \tag{3.8}$$

where $ADC(\mathbf{x})$ is the ADC measured at voxel $\mathbf{x}$ and ADC_{min}, ADC_{max} are the minimum and maximum values observed over all voxels at this timepoint. A maximum of 0.75 is imposed, which is the average of the five highest values measured experimentally in [96].

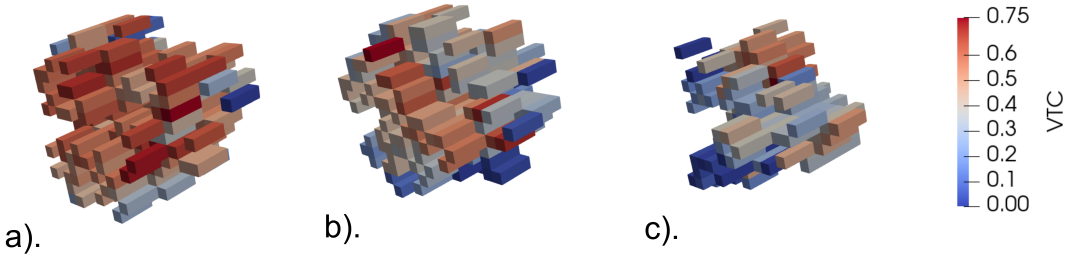


Figure 3.12: Tumour volume fractions calculated per voxel in patient 21 at three timepoints: pretreatment (a), after one cycle of NACT (b) and after three cycles of NACT (c). Voxels are 1.2 mm in the in-plane directions with 4 mm slice thickness.

Chapter 4

Locally-calibrated proliferation rate: a finite difference implementation

Contents

A finite difference (FD) solution for the state-of-the-art mechanically-coupled reaction diffusion (MC–RD) model implemented in MATLAB is available on GitHub [2, 14]. Preliminary experiments indicated that this implementation incurs a high computational cost, with the mechanical coupling component being the primary contributor. Consequently, this chapter examines the role of mechanical coupling in predicting breast tumour response to NACT, to determine whether its inclusion is justified.

Example pre-processed data from one test patient is also provided: tumour cellularities and tissue distributions estimated from one scan before NACT is initiated and one scan after the patient has received one cycle of the treatment. The code provided calibrates the model parameters to the test patient data before extrapolating forward in time to predict the tumour evolution by the third scan. Data from the third scan is not provided, therefore we cannot directly compare the model prediction to the pathological outcome for this particular tumour.

4.1 Implementation and data

The mathematical model is given in equations (2.5)-(2.7). The problem is solved using a fully explicit three-dimensional finite difference (FD) grid, implemented with central differences in space and forward difference in time. The full numerical details can be found in [33]. The FD solution is executed in two steps. The calibration step first uses the scanning schedule and test patient data from two DW-MRI scans (prior to the initiation of treatment and after one cycle of NACT). The mesh is coarsened by a factor of two in the $x-$ and $y-$ planes to reduce the computational time, and once the solution is computed it is reinterpolated to the original grid for evaluation. Figure 4.1 depicts one forward pass of the calibration, which is embedded within a Levenberg-Marquardt optimisation loop [97, 98]. The output is a best estimation of the tumour cell proliferation rate, which can be positive or negative, at each voxel in a large ROI surrounding the tumour ($k(\mathbf{x})$) and of their scalar diffusion coefficient in the absence of stress (D_0). The simulation step then uses these newly calibrated parameters to output a three-dimensional prediction of the tumour cellularity after two to four cycles of NACT.

The data provided are not the MR images themselves but the preprocessed `.mat` files required to run the model. These include the number of tumour cells at each voxel and the spatial distribution of different breast tissues (adipose, fibroglandular and tumour) at cycles 0 and 1, for seven slices of processed MRI data (resolution 171-by-171, voxel

volume 8.45 mm^3). The image acquisition, processing, registration and analysis is described in detail by Jarrett et al. [2].

Figure 4.1: Flowchart depicting one forward pass of the model: equation (2.7) establishes mechanical equilibrium, equation (2.6) is the MC and equation (2.5) the RD. Note that the main model variable N_{TC} represents the cellularity at each voxel, and the carrying capacity θ is a maximum number of tumour cells allowed in a voxel. For the parameter calibration, a Levenberg-Marquardt algorithm is used to output a best estimation of the diffusion coefficient of tumour cells in the absence of stress (D_0) and of numerous proliferation rates calibrated at each voxel in and around the tumour ($k(\mathbf{x})$).

4.2 Computational expense

The calibration code failed to complete in the time available when run locally. It also exceeded the standard 48-hour job limit on the University of Leeds' ARC4 supercomputer. The results presented here were obtained using the University of Luxembourg's more powerful Aion supercomputer [99]; its specifications are detailed in Subsection 5.6.5.

The mechanical coupling was observed to be the most computationally expensive part of the calibration. To investigate its impact, this coupling was removed to solve the RD equation alone. This was achieved by modifying the function which calculates the new, mechanically-coupled diffusion to simply use the value of diffusion calibrated from the patient data instead. The tumour cellularity results predicted by each model are shown in Figure 4.3, along with a subtraction image showing their difference, and the calibration and initialisation data. Qualitatively, the tumour presents the same general shape and comparable cellularity. Results are discussed quantitatively in Section 4.3. Figure 4.4 shows the calibrated proliferation rates from the MC–RD model.

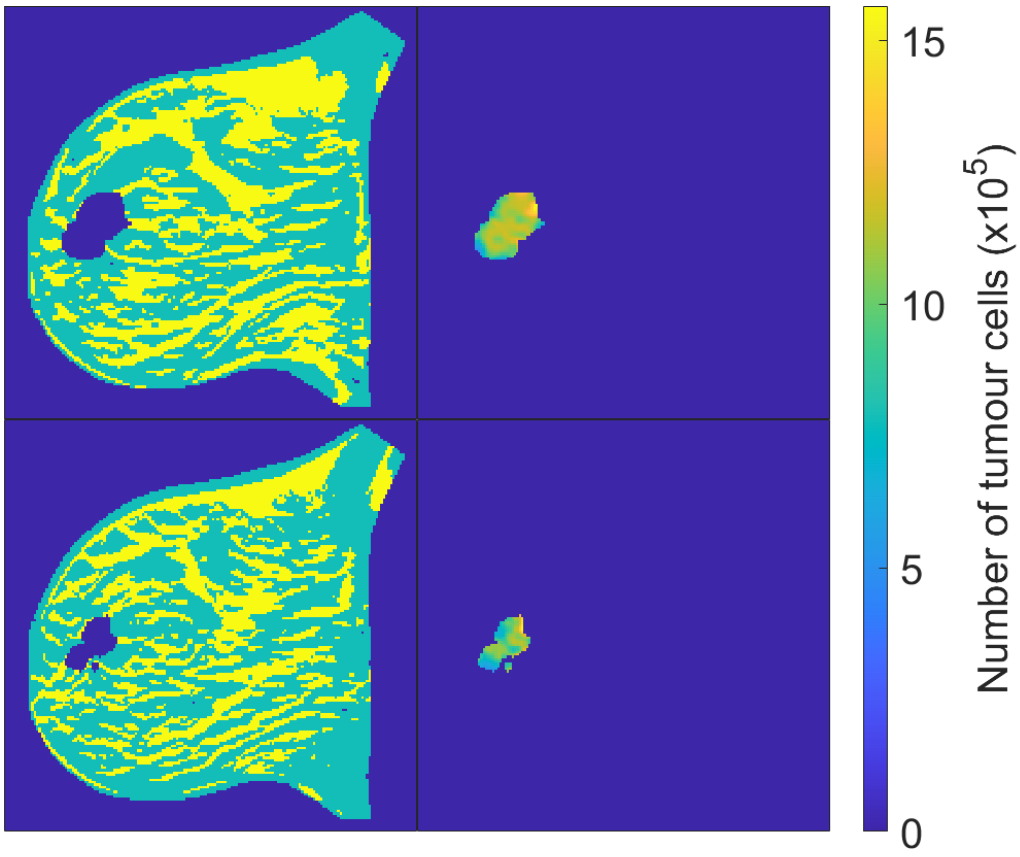


Figure 4.2: Example data: slice 4 shown from Scan 0 (top row) and Scan 1 (bottom row). The breast tissue structure data is shown on the left, where yellow voxels indicate fibroglandular tissue and cyan indicates adipose. The tumour cellularity is shown on the right, with its associated colourbar.

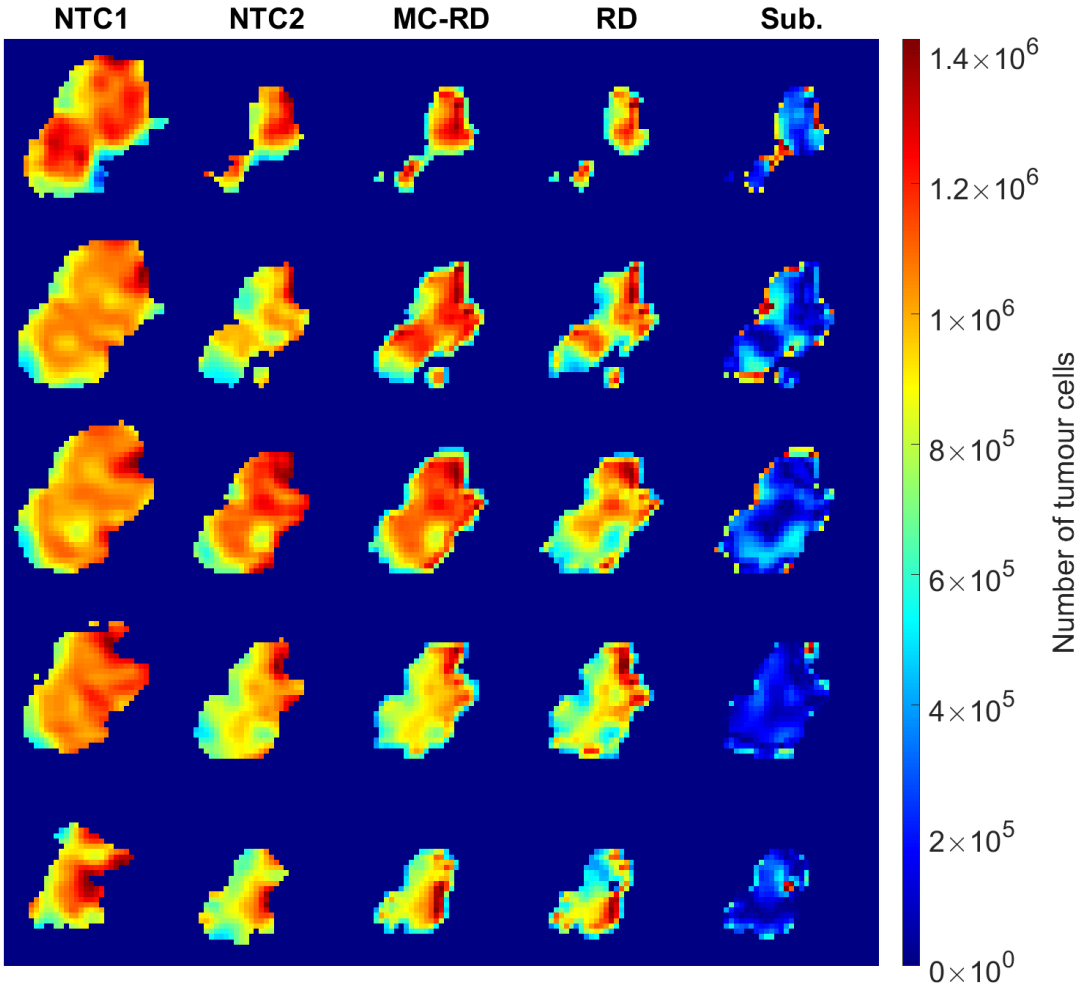


Figure 4.3: Tumour cell number predictions for slices 3–7 of the test patient data, with (MC–RD) and without (RD) mechanical coupling. An image of the absolute difference between each model prediction is also shown (Sub.). The patient data is also included: NTC1 and NTC2 were first used for model calibration, then NTC2 served as the starting point for the prediction. The calibration and simulation were carried out on a grid coarsened by a factor of two compared to what is shown here. The domain is only to a ROI around the tumour for visualisation purposes only; the predictions were computed on the $86 \times 86 \times 7$ mesh. The root mean square error between the predicted tumour cellularities (MC-RD and RD) is 1.8×10^{-4}.

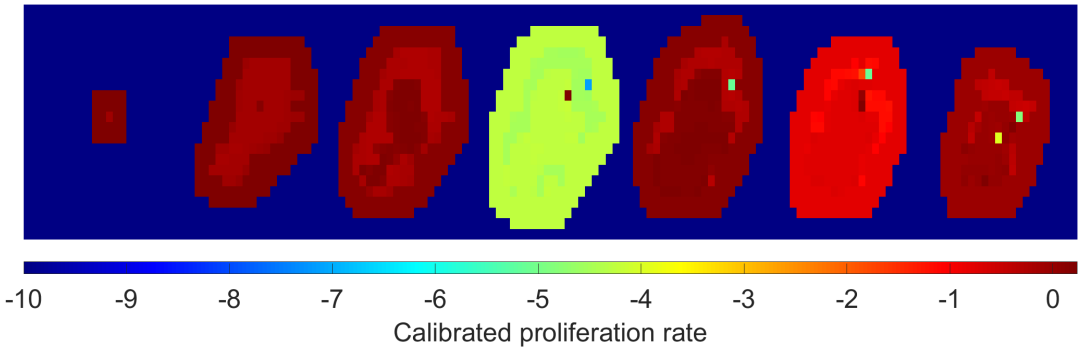


Figure 4.4: Proliferation rates calibrated between scans 1 and 2 (NTC1 and NTC2 shown in Figure 4.3) used for the MC–RD model prediction, shown for each slice of data. Beyond the ROI in which proliferation rates are calibrated, the proliferation rate is set to a fixed value of -10. The calibrated diffusion coefficient was $D_0 = 3.3 \times 10^8$ $\mathrm{m^2 s^{-1}}$.

Table 4.1 summarises the predicted tumour characteristics and calibration durations for each model. The full MC–RD model, as originally formulated, was calibrated in one day, 12 hours and 17 minutes. The optimisation function coarsened the initial $171 \times 171 \times 7$ mesh to $86 \times 86 \times 7$. Removing mechanical coupling allowed the RD model to run in just 9 minutes, highlighting that the computational cost is primarily driven by the MC component. Simulation times were negligible (less than a minute in both cases).

4.3 Conclusions

Since no target data was provided for evaluation, we cannot assess how closely each model prediction aligns with reality. While mechanical coupling has been shown to improve model performance [14], the metrics in Table 4.1 suggest it is not the primary factor influencing the solution. The predicted tumour volumes are similar with or without MC, and the longest axes are nearly identical (3.83 cm versus 3.88 cm, a 1.3% difference). The root mean square error between the predicted tumour cellularities is

	NTC2	RD	MC–RD
Longest axis (cm)	3.82	3.88	3.83
Total cellularity ($\times 10^6$)	1180	931	1104
Tumour volume (cm^3)	10.56	9.72	10.54
Calibration time (h)	–	0.21	36.28

Table 4.1: Comparison of tumour characteristics (longest axis, total cellularity and volume) across the prediction initial data (NTC2) and prediction results with (MC–RD) and without (RD) mechanical coupling.

only 1.8×10^{-4}.

The primary objective of tumour modelling for NACT prediction is to identify responders and nonresponders to treatment. RECIST evaluations, detailed in Subsection 5.8.1, serve as a widely used metric for this differentiation. In the case of the RD and MC–RD predictions discussed here, both models would classify this patient in the same way. This suggests that mechanical coupling does not significantly improve model predictions, while accounting for almost all of the computational time.

Chapter 5

A minimally calibrated, maximally image-informed approach

Contents

5.1 Introduction

In this chapter, a family of minimally calibrated and highly image-informed models for predicting breast tumour response to NACT is introduced. As discussed in Chapter 4, the mechanical coupling mechanism used in current state-of-the-art models [2, 14] does not appear to significantly enhance their performance relative to the clinical standard RECIST criteria. It does, however, substantially increase computational time. For this reason, mechanical coupling is omitted from the present models. Furthermore, as noted in Chapter 2, state-of-the-art models rely on locally optimising proliferation rates at each voxel, resulting in a large number of parameters and a risk of overfitting [100, 101]. In contrast, rather than calibrating numerous spatially varying proliferation rates $k(\mathbf{x})$ as in previous approaches, we propose calibrating a single scalar proliferation rate k. This parameter scales a map of K^{trans}, the volume transfer constant reflecting the efflux

rate of contrast agent from blood plasma into the tissue extravascular extracellular space. With K^{trans} as a proxy for tumour blood flow, we hypothesise that regions with higher K^{trans} receive better blood supply and that this leads to greater tumour cell proliferation. This approach prioritises mathematical modelling over the potential risk of overfitting through extensive calibration. By reducing the number of calibrated parameters, we may also decrease the large computational expense of the optimisation process.

Ultimately, the aim is to predict the tumour's response to chemotherapy after just one cycle, by calibrating the model with scans 0 and 1 then comparing the model prediction at the end of treatment to the final scan. However, our dataset presents some challenges. The full treatment and scanning schedules are described in Chapter 3, Section 3.3. In summary, all patients received antineoplastic drugs for the first three cycles and most were switched to a taxane from cycle 4 onward, with mid-treatment scans only taken after cycles 1 and 3. Since antineoplastic drugs and taxanes work through entirely different mechanisms, predictions based on response to cycle 1 may not be applicable to cycles 4-7, which complicates modelling the full treatment course. There is one exception, as one patient received the same EC90 treatment for all six cycles (Figure 3.4). For this patient, the model can be calibrated with scans 0 and 1 and those parameters can be used to predict the tumour response after six cycles. For the other patients, however, we are limited to predicting the response after three cycles based on the first two scans (Figure 5.1).

5.2 Mathematical model

Starting from a model inspired by the image-informed reaction–diffusion model in [2], we aim to improve clinical efficiency by reducing computational cost. To do so, a family of models presented in increasing order of complexity is proposed. The models contain only one unknown, the tumour cell volume fraction V_{TC}. We hypothesise that the evolution of V_{TC} is dependent on at most three patient-specific scalar parameters: the

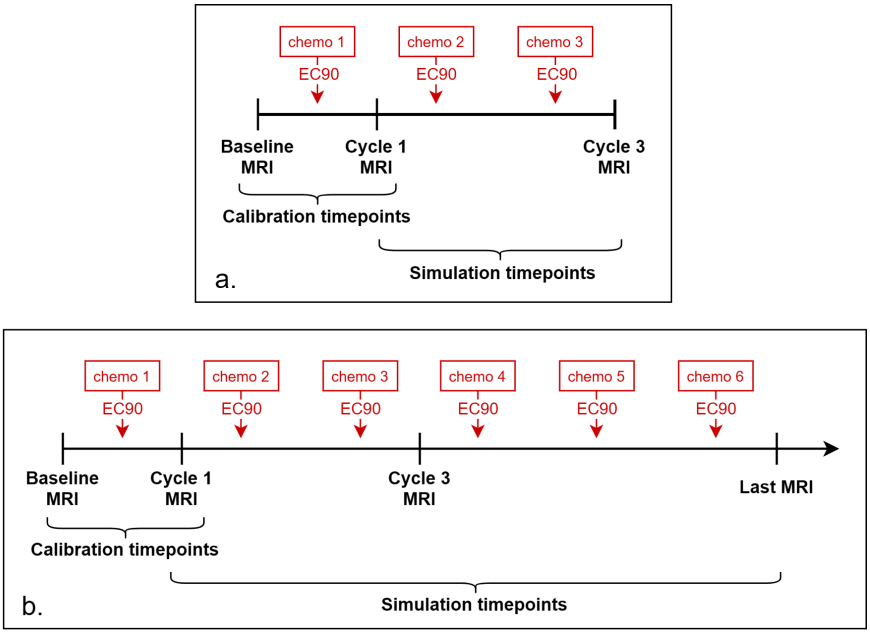


Figure 5.1: Treatment, scanning, calibration and simulation timeline for most patients (a) and for the single patient who received six cycles of the same drug (b).

chemotherapy response α, the diffusion coefficient of the tumour cells D – *i.e.* the ability of the tumour cells to invade surrounding tissue – and the tumour cell proliferation coefficient k.

This leads to the following process of patient-specific model calibrations: first, the tumour evolution during NACT is evaluated only by the chemotherapy response α:

$$\frac{\partial V_{TC}(\mathbf{x},t)}{\partial t} = -\alpha C_{\text{drug}}^{\text{tissue}} V_{TC}(\mathbf{x},t), \tag{5.1}$$

where $C_{\text{drug}}^{\text{tissue}}$ is the tumour tissue drug concentration (defined in Subsection 3.5.3). Then, if this simplified model is not sufficient to predict patient response to NACT, the invasive behaviour of the tumour cells is implemented first via a proliferation term:

$$\frac{\partial V_{TC}(\mathbf{x},t)}{\partial t} = k \cdot K^{\text{trans}} V_{TC}(\mathbf{x},t)\left(1 - \frac{V_{TC}}{\theta}\right) - \alpha C_{\text{drug}}^{\text{tissue}} V_{TC}(\mathbf{x},t), \tag{5.2}$$

where the proliferation rate is informed by a K^{trans} mapping defined over the entire breast domain, and $\theta = 0.75$ is the maximum tumour cell volume fraction permitted

(introduced in Section 3.5.4). Another option is to implement the invasive behaviour of the tumour cells via a diffusion term:

$$\frac{\partial V_{TC}(\mathbf{x},t)}{\partial t} = D\nabla^2 V_{TC}(\mathbf{x},t) - \alpha C_{\text{drug}}^{\text{tissue}} V_{TC}(\mathbf{x},t). \tag{5.3}$$

A final model considers the diffusion, proliferation and chemotherapy response:

$$\frac{\partial V_{TC}(\mathbf{x},t)}{\partial t} = D\nabla^2 V_{TC}(\mathbf{x},t) + kK^{\text{trans}} V_{TC}(\mathbf{x},t)\left(1 - \frac{V_{TC}}{\theta}\right) - \alpha C_{\text{drug}}^{\text{tissue}} V_{TC}(\mathbf{x},t). \tag{5.4}$$

Beyond this apparent simplicity, these equations are informed by an in-depth analysis of the patient imaging: the initial condition of V_{TC} is informed by DW-MRI, and parameters k and α both scale the mappings of K^{trans} and $C_{\text{drug}}^{\text{tissue}}$ obtained from the DCE-MRI analysis.

For clarity, the remainder of this thesis refers to these models by their calibrated parameters: the α (5.1), $k\alpha$ (5.2), $D\alpha$ (5.3) and $kD\alpha$ (5.4) models.

Name	Unit	Meaning	Source	Type
V_{TC}	dimensionless	tumour cell volume fraction	variable initialised with patient data	image
α	[mL mg^{-1} s^{-1}]	drug efficacy	calibrated	scalar
D	[m^{2}s^{-1}]	tumour cell diffusion coefficient	calibrated	scalar
k	dimensionless	proliferation coeff.	calibrated	scalar
$C_{\text{drug}}^{\text{tissue}}$	[mg mL^{-1}]	concentration of drug in tissue	patient data	image
K^{trans}	[s^{-1}]	perfusion	patient data	image
θ	dimensionless	maximum V_{TC}	measured *in vivo* [96]	scalar

Table 5.1: Model variables and parameters.

5.3 Implementation

5.3.1 Numerical methods

All models were implemented in three dimensions (3D) and initialised using patient-specific DW-MRI and DCE-MRI. The calibration and simulation codes were written in the Python-based finite element library FEniCSx [102], with first-order in space Lagrange elements and backward Euler timestepping. The timestep size Δt was gradually

increased from 10 seconds to 30 minutes in 10-second increments over the first 180 timesteps, to ensure stability for the $k\alpha$, $D\alpha$ and $kD\alpha$ models. This was not necessary for the α model, which employed a constant $\Delta t = 30$ mins. Upper and lower thresholds on acceptable tumour cell volume fractions were imposed after each solver iteration: values were capped at $\theta = 0.75$ [96], and those below 10^{-5} were considered nontumour and set to zero [103].

A stability check was implemented to terminate the solver early if the solution remained unchanged for 50 iterations (equivalent to 25 real-time hours, chosen empirically). The α model, which only simulates the duration of the chemotherapy (from administration until the drug concentration decreases to 1% of its maximum), frequently stabilised before reaching the final iteration. This early termination resulted in computational time savings often exceeding 50%. The other models include "growth" phases, where the tumour cells are allowed to diffuse and/or proliferate in the absence of a chemotherapy term: from the first scan until the point at which the drug is delivered, then from the end of drug delivery until the date of the next scan. For these models, stabilisation occurred less frequently; if it did, the computation advanced to the next stage rather than terminating entirely.

Variational formulation For the $kD\alpha$ model, multiplying equation (5.4) by the test function v, V_{TC} and v in the Sobolev space H1, and integrating over the computational domain Ω we have:

$$\int_\Omega \frac{\partial V_{TC}}{\partial t} v \mathrm{d}\Omega = \int_\Omega D\nabla^2 V_{TC} v \mathrm{d}\Omega + \int_\Omega k K^{\mathrm{trans}} V_{TC} \left(1 - \frac{V_{TC}}{\theta}\right) v \mathrm{d}\Omega - \int_\Omega \alpha C v \mathrm{d}\Omega, \quad (5.5)$$

writing $C = C_{\mathrm{drug}}^{\mathrm{tissue}}$ for conciseness. Using integration by parts and the divergence theorem, the diffusion term $\nabla^2 V_{TC}$ can be rewritten as:

$$\int_\Omega \nabla^2 V_{TC} v \mathrm{d}\Omega = \int_{\partial\Omega} \frac{\partial V_{TC}}{\partial n} v \mathrm{d}S - \int_\Omega \nabla V_{TC} \cdot \nabla v \mathrm{d}\Omega, \quad (5.6)$$

where the surface integral $\int_{\partial\Omega} \frac{\partial V}{\partial n} v \mathrm{d}S$ is prescribed to be zero on the domain boundary $\partial\Omega$, which was defined as a box surrounding the tumour with a padding of at least two

voxels in each direction. Equation (5.5) thus becomes

$$\int_\Omega \frac{\partial V_{TC}}{\partial t} v \mathrm{d}\Omega = -\int_\Omega D\nabla V_{TC}\cdot\nabla v \mathrm{d}\Omega + \int_\Omega kK^{\mathrm{trans}} V_{TC}\left(1-\frac{V_{TC}}{\theta}\right) v \mathrm{d}\Omega - \int_\Omega \alpha C v \mathrm{d}\Omega. \tag{5.7}$$

We use a backward Euler scheme for the time discretisation:

$$\int_\Omega \frac{V_{TC}-V_{TC0}}{\Delta t} v \mathrm{d}\Omega = \int_\Omega \left[-D\nabla V_{TC}\cdot\nabla v + kK^{\mathrm{trans}} V_{TC}\left(1-\frac{V_{TC}}{\theta}\right) v - \alpha C v\right] \mathrm{d}\Omega, \tag{5.8}$$

which rearranges to finally give the finite element variational form F_{chemo} used for the chemotherapy timesteps:

$$F_{\mathrm{chemo}} = \int_\Omega \left[(V_{TC}-V_{TC0})v + \Delta t\left(D\nabla V_{TC}\cdot\nabla v - kK^{\mathrm{trans}} V_{TC}\left(1-\frac{V_{TC}}{\theta}\right) v + \alpha C v\right)\right] \mathrm{d}\Omega. \tag{5.9}$$

The variational form for the "growth" timesteps is derived identically, minus the chemotherapy term:

$$F_{\mathrm{nonchemo}} = \int_\Omega \left[(V_{TC}-V_{TC0})v + \Delta t\left(D\nabla V_{TC}\cdot\nabla v - kK^{\mathrm{trans}} V_{TC}\left(1-\frac{V_{TC}}{\theta}\right) v\right)\right] \mathrm{d}\Omega. \tag{5.10}$$

Following the same methodology, the variational form for the α model (5.1) is:

$$F_{\mathrm{chemo}} = \int_\Omega \left[(V_{TC}-V_{TC0})v + \Delta t \alpha C v\right] \mathrm{d}\Omega, \tag{5.11}$$

for the $k\alpha$ model (5.2):

$$F_{\mathrm{chemo}} = \int_\Omega \left[(V_{TC}-V_{TC0})v + \Delta t\left(-kK^{\mathrm{trans}} V_{TC}\left(1-\frac{V_{TC}}{\theta}\right) v + \alpha C v\right)\right] \mathrm{d}\Omega, \tag{5.12}$$

$$F_{\mathrm{nonchemo}} = \int_\Omega \left[(V_{TC}-V_{TC0})v - \Delta t kK^{\mathrm{trans}} V_{TC}\left(1-\frac{V_{TC}}{\theta}\right) v\right] \mathrm{d}\Omega, \tag{5.13}$$

and for the $D\alpha$ model (5.3):

$$F_{\mathrm{chemo}} = \int_\Omega \left[(V_{TC}-V_{TC0})v + \Delta t\left(D\nabla V_{TC}\cdot\nabla v + \alpha C v\right)\right] \mathrm{d}\Omega, \tag{5.14}$$

$$F_{\mathrm{nonchemo}} = \int_\Omega \left[(V_{TC}-V_{TC0})v + \Delta t\left(D\nabla V_{TC}\cdot\nabla v\right)\right] \mathrm{d}\Omega. \tag{5.15}$$

5.3.2 Computational domain and meshing

For each patient, the size of the computational domain was defined as a cuboid encompassing the tumour segmentations with a minimum padding of 2 voxels in all directions. This padding threshold was selected based on the established observation that adenocarcinomas exhibit minimal growth over NACT timescales. Tetrahedral meshes of increasing refinement were generated using Gmsh [104] to determine the resolution at which the numerical solution became independent of the choice of mesh. Figure 5.2 shows the difference between the error from each mesh and the error from the finest mesh (as defined in Subsection 5.6.2), all other patient plots can be found in Appendix A.

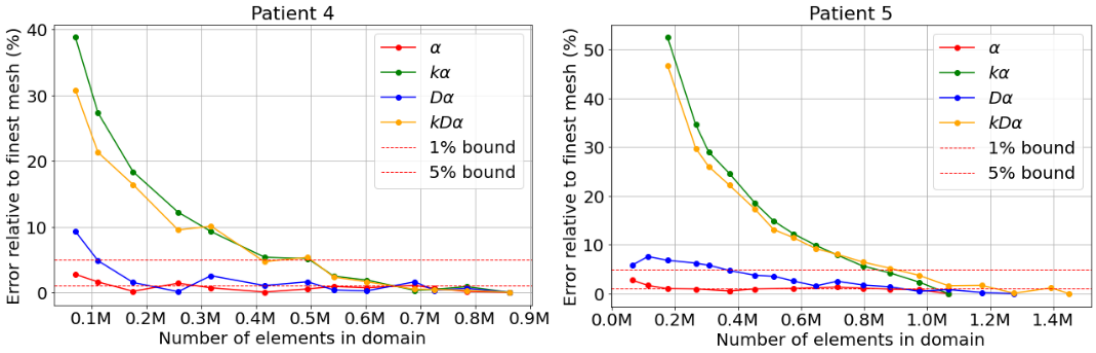


Figure 5.2: Examples of mesh convergence plots for two patients. The solution is considered to be independent of the mesh if the volume of the solution tumour is within 1% of the volume of the finest mesh's solution tumour for the α model, and within 5% for the $k\alpha$, $D\alpha$ and $kD\alpha$ models (to avoid prohibitively long simulation times).

An additional risk associated with overly coarse meshes is the potential for errors to arise when interpolating data between function spaces. Figure 5.3 describes a section of code in which this can be a problem. The optimisation's target data originally provides a single value at the centre of each hexahedral voxel and is first read into FEniCSx via a discontinuous Lagrange function space, which comprises of one value

at the centre of each solution mesh tetrahedron. The target data is subsequently interpolated onto the continuous Lagrange function space where the equations will be solved, so that the solution tumour may be compared with it. This space is comprised of one value at each element vertex, i.e. four values per tetrahedron, and this is the step at which discrepancies may occur. If the mesh is too coarse, significant differences can emerge between the actual tumour volume and the volume computed on the solution grid, which critically impacts the optimisation process when the cost function involves tumour volumes (see Figure 5.11).

Thus the optimal mesh for each patient and each model was selected as the coarsest mesh beyond which further refinement resulted in a change in the solution of no more than 1%. Furthermore, consistency in the volume of the target tumour data between the different function spaces was confirmed.

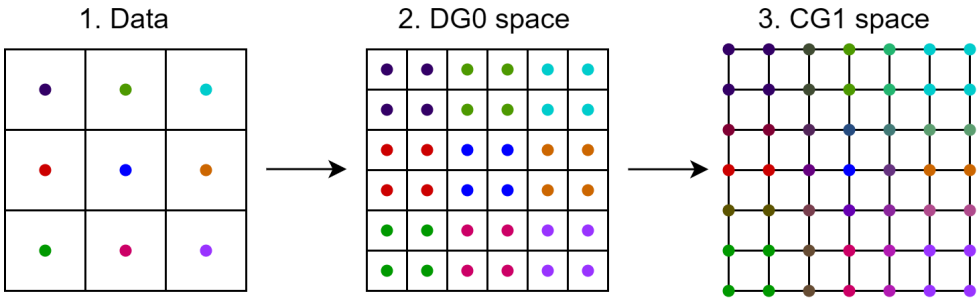


Figure 5.3: Illustration of the process of reading MR data into FEniCSx. Note that in reality the problem is three-dimensional and the solution mesh (shown in steps 2 and 3) is tetrahedral.

5.4 Excluded patients

Eight patients were excluded from the entire study, and a further two only from the prediction step:

- Patients 12 and 40 were excluded due to issues in their MR data. For patient 12, the DCE field of view was cropped to the wrong size and didn't include the

entire aorta. This rendered the DCE analysis described in Chapter 3 impossible. For patient 40, an error occurred during the acquisition of the DCE-MRI.

- Patient 28 was excluded due to unsatisfactory image registration. The overlap volume between the registered scan 0 and scan 1 tumours was only 16% of the scan 0 tumour volume and 29% of the scan 1 tumour volume (with Dice similarity coefficient 0.2), making it impossible to accurately compare the model predictions to the true tumour size and shape.

- Patient 38 was excluded due to the tumour's small size: when the tumour occupies only a few voxels, segmentation errors affecting a single voxel can constitute a significant proportion of the tumour volume, leading to artificially high errors which significantly impede the optimisation process.

- The upper bound of 1×10^{-2} on α (described in Subsection 5.5) is necessary to prevent model instability for the patient with the highest $C_{\text{drug}}^{\text{tissue}}$. This arises when the chemotherapy term $\left(-\alpha C_{\text{drug}}^{\text{tissue}}\right)$ becomes excessively large compared to the other terms. This poses a practical constraint on the models' applicability: when the mean $C_{\text{drug}}^{\text{tissue}}$ is too low, i.e. falls below 10^{-6}mg/mL, the chemotherapy term is insufficiently large to influence the solution meaningfully. As a result, patients 3, 25, 32, and 33 are excluded from this study due to their $C_{\text{drug}}^{\text{tissue}}$ at Scan 0, and patients 8, 13, and 38 are excluded from the simulation step because of their $C_{\text{drug}}^{\text{tissue}}$ at Scan 1. These data are shown in Figure 3.11.

- Lastly, patients 2, 22 and 27 withdrew from the CHERNAC study. To avoid confusion, as the DICOM datasets are identified by the original patient IDs, the remaining patients were not renumbered.

Model calibrations and predictions after three cycles of NACT were still conducted for the seven patients with low $C_{\text{drug}}^{\text{tissue}}$; these results are included in Appendix B. This was not possible for the patients with problematic MRI data. Thus of the 34 patients whose datasets could be used in these models, 27 remained in the final cohort discussed

in this chapter.

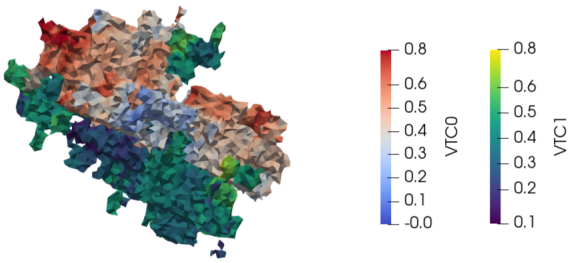


Figure 5.4: Patient 28 was excluded due to poor registration between scans 0 and 1.

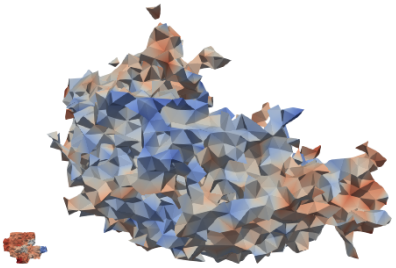


Figure 5.5: Patient 38's tumour at Scan 0 (small; left) next to patient 3's (large; right) for scale.

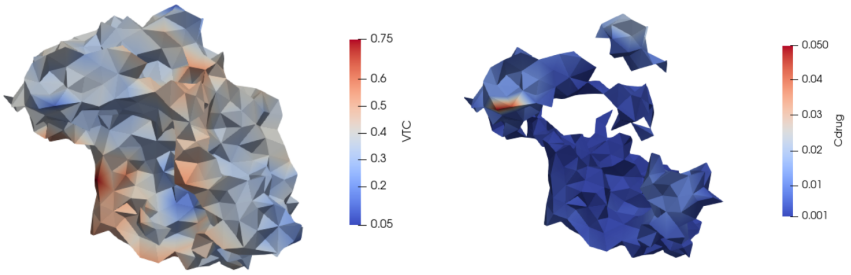


Figure 5.6: Example of a low C_{drug}^{tissue} case: patient 8 at Scan 1. The visualisation of the tumour cell volume fraction on the left shows the full shape of the tumour, with the drug concentration within it shown on the right.

5.5 Parameter space exploration

To determine appropriate bounds for the parameters requiring calibration, simulations from Scan 0 to Scan 1 were carried out for each model using various parameter combinations. The lowest values were chosen to ensure no observable impact, while the upper values were increased until they caused the solver to fail. These simulations were carried out for five patients with different treatment outcomes to guide the choice of parameter bounds for the optimisation process.

Figure 5.7 shows an overview of the $kD\alpha$ model parameter space. Nine hundred parameter combinations were investigated: the (dimensionless) proliferation coefficient k was varied between 10^{-14} and 10^{-6}, the diffusion coefficient D between 10^{-16} and 10^{-8} m^2/s and the drug efficacy α between 10^{-9} and 1 mL/(mg·s). Each dot represents a simulation carried out with parameters (k, D, α) given by its coordinates, where the relative error between the numerical tumour volume and the data tumour volume is indicated by the simulation error colourbar. The V_{TC} colourbar is relevant to the numerical results visualisations; it shows the tumour volume fraction V_{TC} at each element of the domain where $V_{TC} > 1e-5$ (below which the element is considered non-tumour) [103]. Seven areas of the parameter space were identified, corresponding to seven solution behaviours; one example solution is shown for each zone.

Simulation errors are lowest in zone A, corresponding to the region where the numerical tumour exhibits qualitatively "reasonable" behaviour: it may diffuse slightly, shrink, or disappear entirely. This is consistent with behaviours observed throughout treatment. Beyond $D = 10^{-12}$ m^2/s, the tumour cell diffusion becomes excessive, leading to errors of at least 100% (zone C, where the tumour expands to occupy most of the domain) or higher (zone F, where the tumour fully invades the domain). Similarly, when $k = 10^{-6}$, tumour cell proliferation becomes excessive and results in a tumour that occupies too much of the domain (zone G). Beyond $\alpha = 10^{-2}$ mL/(mg·s) the chemotherapy parameter is four to fourteen orders of magnitude larger than the others, resulting in quantitative errors above 500% and qualitatively unphysical behaviour

(zone B). Lastly, when $D < 10^{-14}$ m^2/s, $k < 10^{-8}$ and $\alpha < 10^{-6}$ mL/(mg·s), these parameters have no meaningful impact on the solution.

Further examples of simulation results with parameters in the "reasonable" range (zone A) are shown in Figure 5.8 for the α model (equation (5.1)) and in Figure 5.10 for the $D\alpha$ model (equation (5.3)), which shows cross-sections of the tumour.

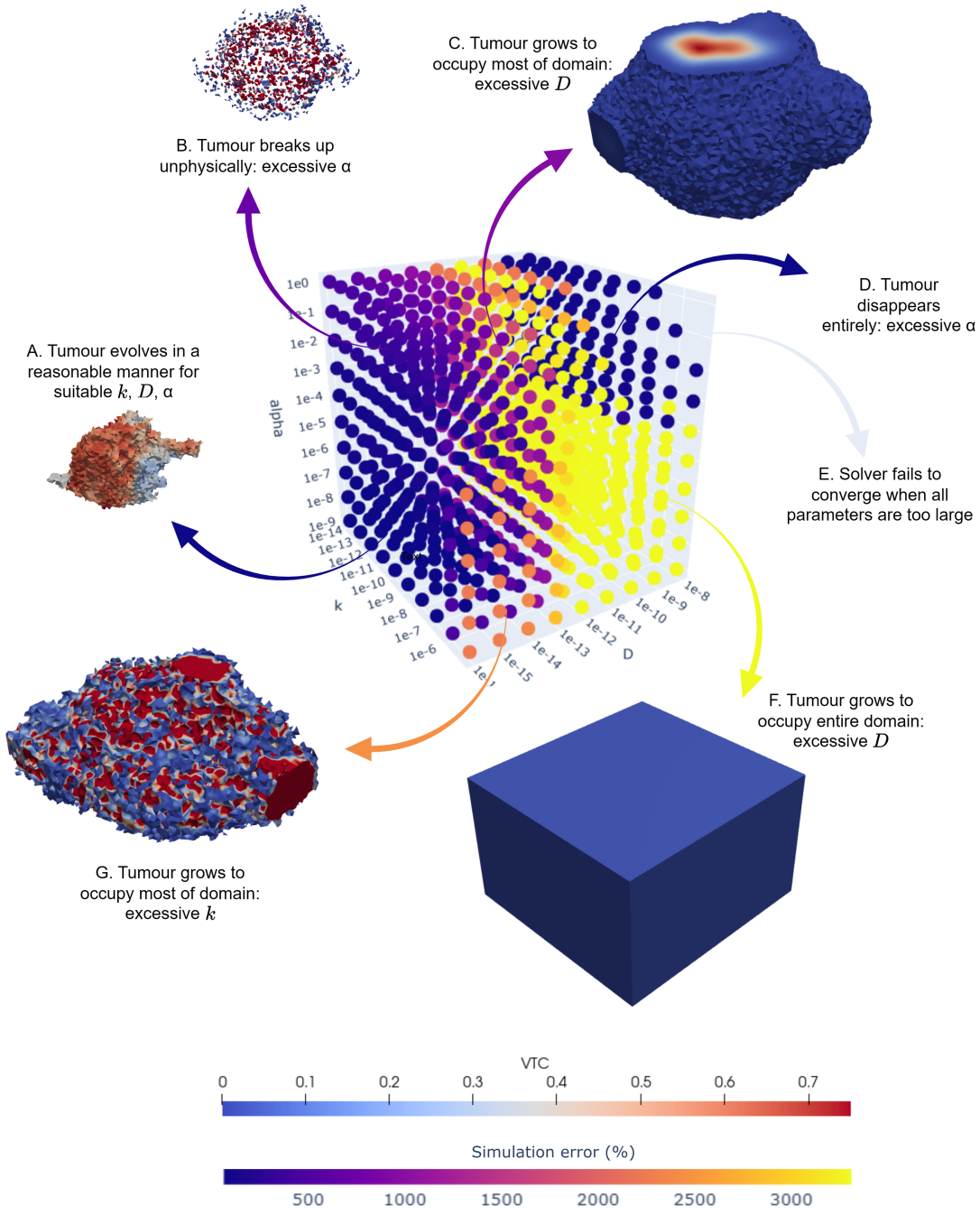


Figure 5.7: $kD\alpha$ model simulation results for one patient with 900 parameter combinations: $k \in (10^{-14}, 10^{-6})$, $D \in (10^{-16}, 10^{-8})$ m^2/s and $\alpha \in (10^{-9}, 1)$ mL/(mg·s). The scale of the error is given by the lower colourbar. Seven zones are identified (A-G) and an example of a solution tumour is provided for each of the zones.

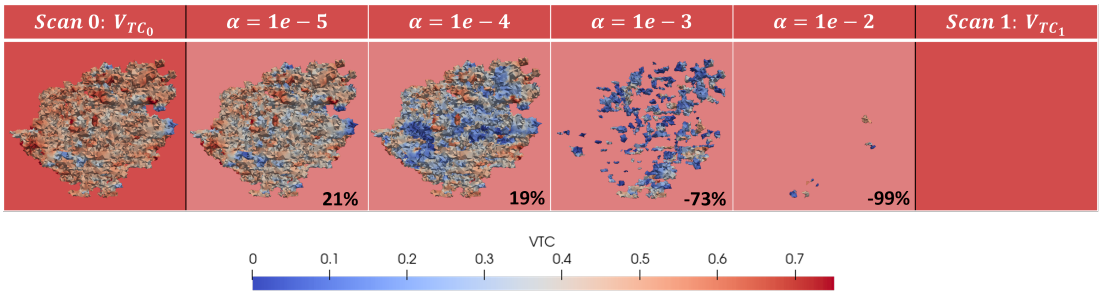


Figure 5.8: Result of four simulations from Scan 0 (shown in the first column) to Scan 1 (last column) for one patient using the α model with $\alpha = 10^{-5}$, $\alpha = 10^{-4}$, $\alpha = 10^{-3}$ and $\alpha = 10^{-2}$ mL/(mg·s). The error shown is the percentage difference between the volumes of the numerical tumour and the Scan 1 tumour.

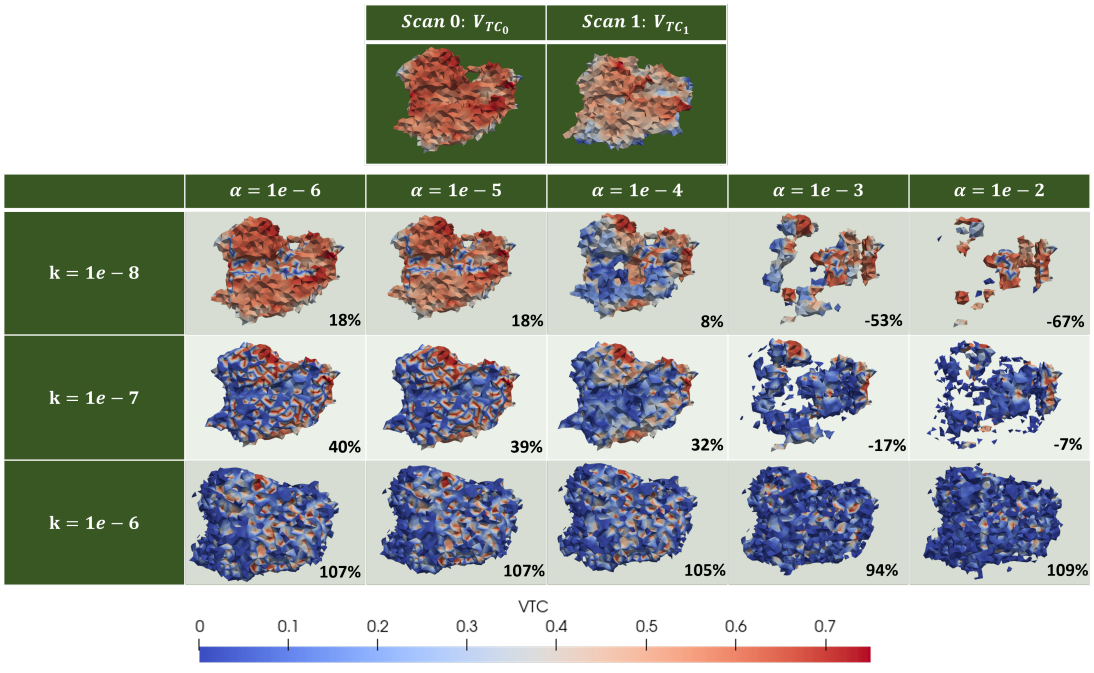


Figure 5.9: Result of 15 simulations from Scan 0 to Scan 1 for one patient using the $k\alpha$ model with $k \in (10^{-8}, 10^{-6})$ m^2/s and $\alpha \in (10^{-5}, 10^{-2})$ mL/(mg·s). The error shown is the percentage difference between the volumes of the numerical tumour and the Scan 1 tumour.

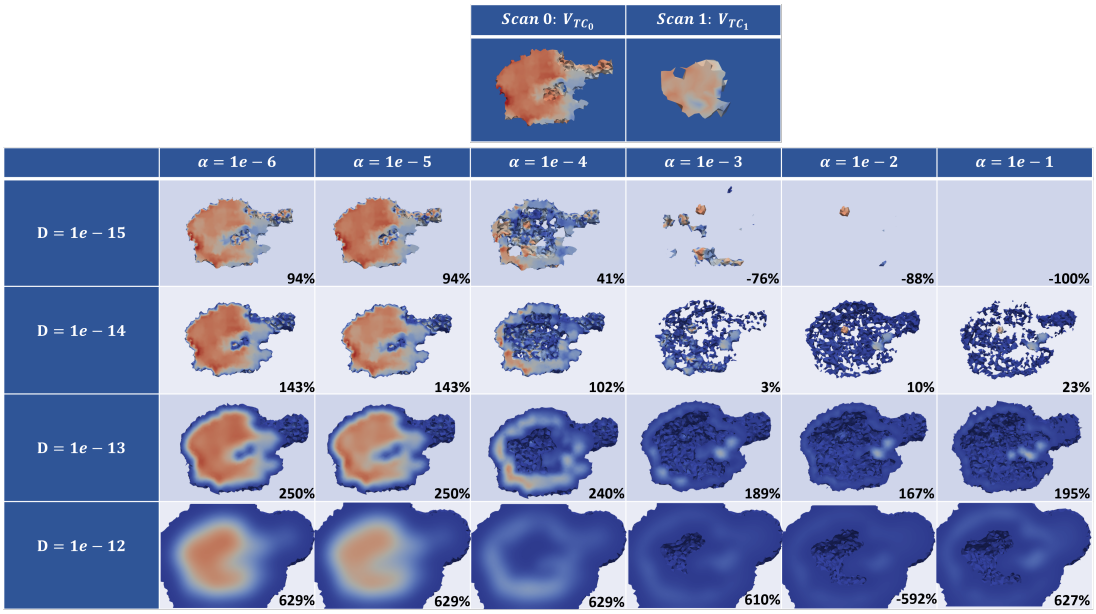


Figure 5.10: Result of 24 simulations from Scan 0 to Scan 1 for one patient using the $D\alpha$ model with $D \in (10^{-15}, 10^{-2})$ m^2/s and $\alpha \in (10^{-6}, 10^{-1})$ mL/(mg·s). The error shown is the percentage difference between the volumes of the numerical tumour and the Scan 1 tumour.

5.6 Model calibration

5.6.1 Overview

Figure 5.11 describes the calibration process, which was handled by an optimisation algorithm. A first guess chosen randomly from Zone A in the parameter space (see Figure 5.7) is made for the parameters to be calibrated (k, D, α). These values are used with the tumour volume fraction measured before the start of NACT (scan 0) to yield the model-predicted tumour volume fraction after one cycle of NACT (scan 1). This prediction is compared to tumour cell number measured after one cycle. The algorithm iteratively adjusts the parameter values until suitable agreement is reached (as defined in Subsection 5.6.2). Once the best set of parameters is identified for the dataset, the forward evaluation of the model to predict further timepoints can proceed.

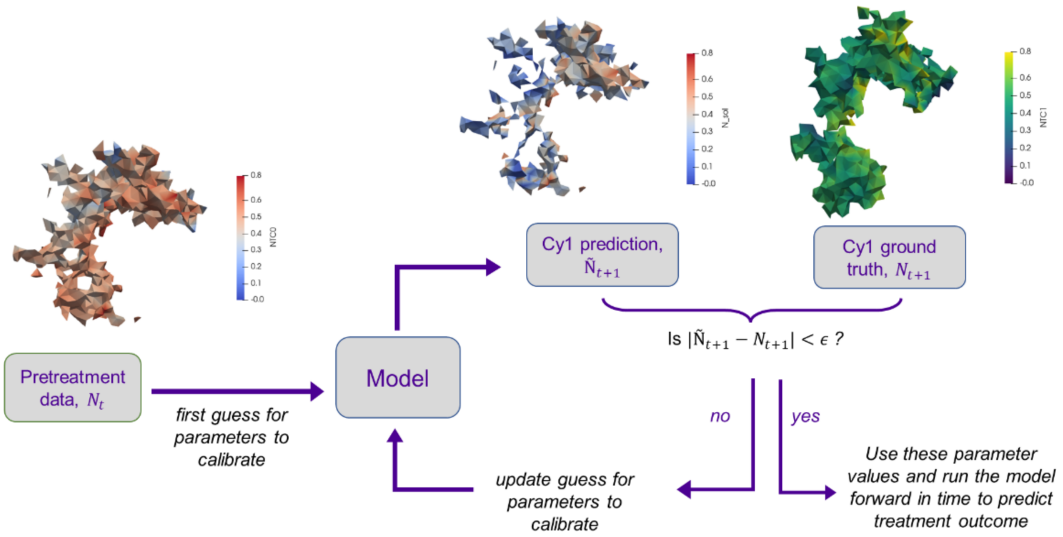


Figure 5.11: Parameter calibration process.

5.6.2 Cost function design

Choosing an appropriate cost function, the error measure used to evaluate the accuracy of each Scan 1 prediction within the optimisation loop, is crucial to achieving a successful optimisation. We explored three cost functions. The first simply minimised the percentage difference between the solution tumour volume v_{sol} and the data tumour volume v_{data}:

$$E_1 = \frac{|\mathrm{v}_{sol} - \mathrm{v}_{data}|}{\mathrm{v}_{data}}. \tag{5.16}$$

The second applied to same approach to the average tumour cell volume fraction, A:

$$E_2 = \frac{|A_{sol} - A_{data}|}{A_{data}}. \tag{5.17}$$

The third was designed to measure the spatial overlapping of the predicted tumour and the target data tumour in an element-by-element manner (Figure 5.12). This cost function considered the proportion of the solution tumour that is made up of elements inaccurately identified as part of the tumour (false positives) and the proportion of the

data tumour which was not identified by the solution (false negatives):

$$E_3 = \frac{1}{2}\left(\frac{\sum_{i\in T_{data}} V_i - \sum_{i\in T_{data}\cap T_{sol}} V_i}{\sum_{i\in T_{data}} V_i} + \frac{\sum_{i\in T_{sol}} V_i - \sum_{i\in T_{data}\cap T_{sol}} V_i}{\sum_{i\in T_{sol}} V_i}\right), \tag{5.18}$$

where V_i is the volume of voxel i and T_{data} and T_{sol} are the data and the solution tumours.

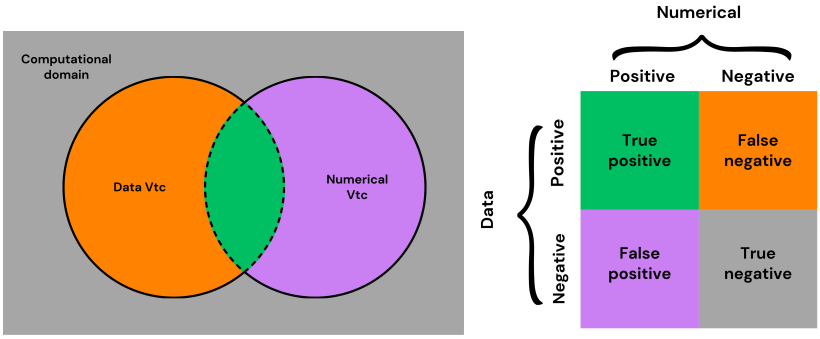


Figure 5.12: Error based on volume of overlap between solution tumour and target data tumour.

Minimising the percentage difference between the solution tumour volume and the data tumour volume proved to be by far the most effective approach, and thus this was used as the cost function for all the results presented subsequently. The average tumour cell volume fraction was not deemed a relevant measure of tumour evolution, while the voxel-by-voxel error was too stringent: since the models struggle to predict the exact tumour location at the voxel level, the best outcome using this metric often resulted in no change to the tumour.

5.6.3 Choice of optimisation algorithm

Three optimisation algorithms were implemented using the NLopt library [105]: COBYLA (Constrained Optimization BY Linear Approximations) [106], Nelder–Mead Simplex [107] and Low-storage BFGS (Broyden–Fletcher–Goldfard–Shanno) [108]. These were identified as those with the most potential for a reasonably quick optimisation process,

from a larger subset of algorithms initially tested. Each algorithm was used to calibrate all four mathematical models for the same twelve patient datasets, to identify the best-performing. Figure 5.13 compares the results from each algorithm in terms of optimisation time and lowest error identified. Each dot represents one patient dataset, with the ellipses centred around the average and stretched by the standard deviation in each direction. BFGS was frequently the fastest to converge but rarely successful in identifying the lowest possible error. Cobyla and Nelder–Mead tended to identify very similar best errors, with the latter being slightly faster for the α model and significantly faster for the $kD\alpha$ model. Therefore, Nelder-Mead was chosen as the best optimisation algorithm for generalising to the rest of the patient cohort.

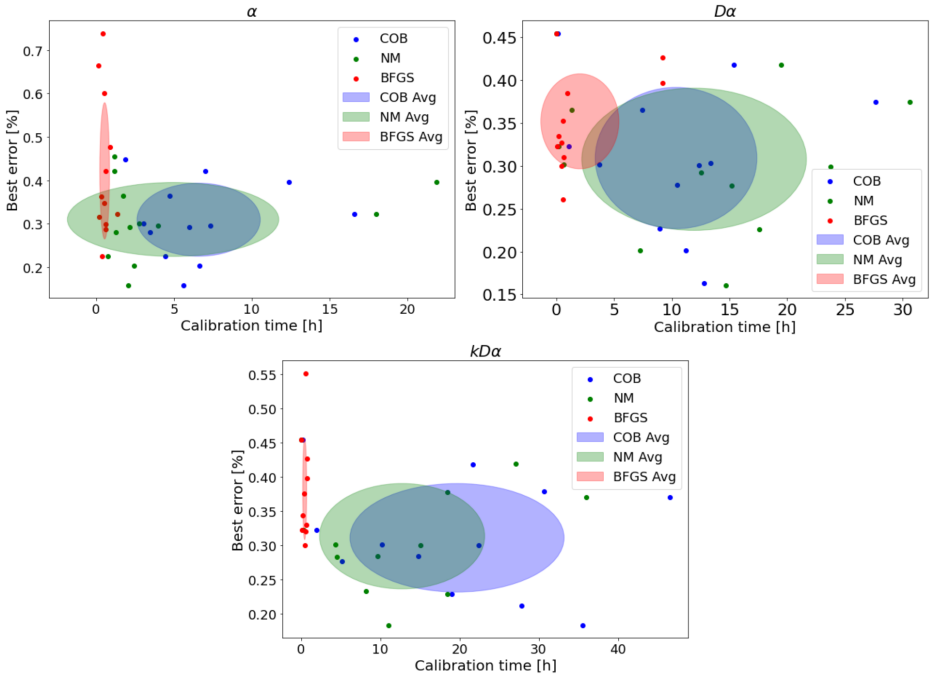


Figure 5.13: Comparing optimisation algorithms: best error against parameter calibration time for a model with one, two or three parameters, tested on 12 patient datasets for the three optimisation candidates: COBYLA, Nelder-Mead and Low-storage BFGS. The ellipses are centred around the mean and stretched by the standard deviation in each direction.

5.6.4 Optimisation settings

A stopping criterion for the cost function was set at 1% for the α model and 5% for the $k\alpha$, $D\alpha$ and $kD\alpha$ models. Relative tolerances of 10^{-6} were imposed on the parameters and on the cost function, below which the optimisation process was terminated. Additionally, the optimisation was constrained to a maximum of 200 iterations or a total runtime of 30 days. This constraint was imposed firstly due to practical limitations on our computational resources, as each job is restricted to a 48-hour runtime before requiring relaunch. It is also important to note that the timeframe of the optimisation represents only 3 weeks of real-time tumour evolution. Table 5.2 lists the parameter bounds and initial values used for optimisation in each case, determined empirically from the parameter space exploration described in Section 5.5. In the $D\alpha$ and $kD\alpha$ models, the initial values for α are set significantly higher to ensure the effect of chemotherapy is represented despite the opposing influence of diffusion. The lower bounds on k, D and α in models (5.2)–(5.4) were chosen such that these parameters would always have an effect on the solution, i.e., to prohibit the $k\alpha$ model from reducing to the α model with $k = 0$, etc.

Model	Parameter	Bounds	Starting point
(5.1)	α	$[10^{-5}, 10^{-2}]$	10^{-4}
(5.2)	k	$[10^{-8}, 10^{-6}]$	5×10^{-8}
	α	$[10^{-5}, 10^{-2}]$	1×10^{-4}
(5.3)	D	$[10^{-14}, 10^{12}]$	1.5×10^{-14}
	α	$[10^{-5}, 10^{-2}]$	4×10^{-3}
(5.4)	k	$[10^{-8}, 10^{-7}]$	5×10^{-8}
	D	$[10^{-14}, 10^{-12}]$	1.5×10^{-14}
	α	$[10^{-5}, 10^{-2}]$	4×10^{-3}

Table 5.2: Upper bounds, lower bounds and starting points for each parameter in each model based on the parameter space exploration (Section 5.5).

5.6.5 Hardware

The code was executed on the University of Luxembourg's Aion supercomputer [99], a cluster of x86-64 AMD-based compute nodes. Each node consists of 8 (virtual) processors with 16 cores each, physically hosting 2 physical sockets of AMD Epyc ROME 7H12 processors with 64 cores each, for a total of 128 cores per node. Each node has 256GB of DDR4 3200MT/s memory (8x16GB DIMMs per socket, 16 DIMMs per node). However, the version of FEniCSx installed on Aion is not currently parallelisable, thus each job was only able to be run on one node.

5.7 Calibration results

This section explores several key aspects of the optimisation results: we compare optimisation durations and fit quality across models, investigate why certain patients prove more challenging to calibrate, analyse the values of the calibrated parameters and compare each model's performance in distinguishing between responders and non-responders. Though calibration results are not inherently predictive, any insights derived solely from scans 0 and 1 are highly valuable as the ultimate goal is to predict tumour response after just one cycle of chemotherapy. For instance, if the calibrated parameters for responders consistently fall within a specific range, this information could enable the identification of potential responders to NACT immediately following calibration.

For the purposes of this discussion, the models were categorised into "chemo-heavy" (α and $k\alpha$) and "diffusion-heavy" ($D\alpha$ and $kD\alpha$) groups. This is because the inclusion of diffusion drives the model predictions towards tumour growth, as simulating the outward diffusion of cells gives the appearance of expansion. This has a significant impact on the solution which is inherently absent from the chemo-heavy models. The focus is primarily on volume errors rather than the other two metrics (longest axis and average tumour cell volume fraction), as volume is considered the most clinically relevant measure for assessing prediction quality. Using the longest axis simplifies a

three-dimensional problem into two dimensions – an unnecessary reduction given our capability to estimate volumes. Meanwhile, the average tumour cell volume fraction lacks a direct correlation with assessments of tumour response to NACT.

Lastly, the RECIST guidelines [109] are used to differentiate between responders and nonresponders as determined from the data collected at Scan 3. Briefly, RECIST categorises response into four groups: complete response (CR), partial response (PR), stable disease (SD) and progressive disease (PD). CR and PR are considered responders, SD and PD nonresponders. RECIST classifications are discussed in greater detail in Subsection 5.8.1.

5.7.1 Duration of the optimisations

Figure 5.14 shows the minimum error achieved during the optimisation process and the time required to attain it for each model. These values are also detailed in the ϵ_{opt} and t_{opt} columns of Tables 5.10–5.13.

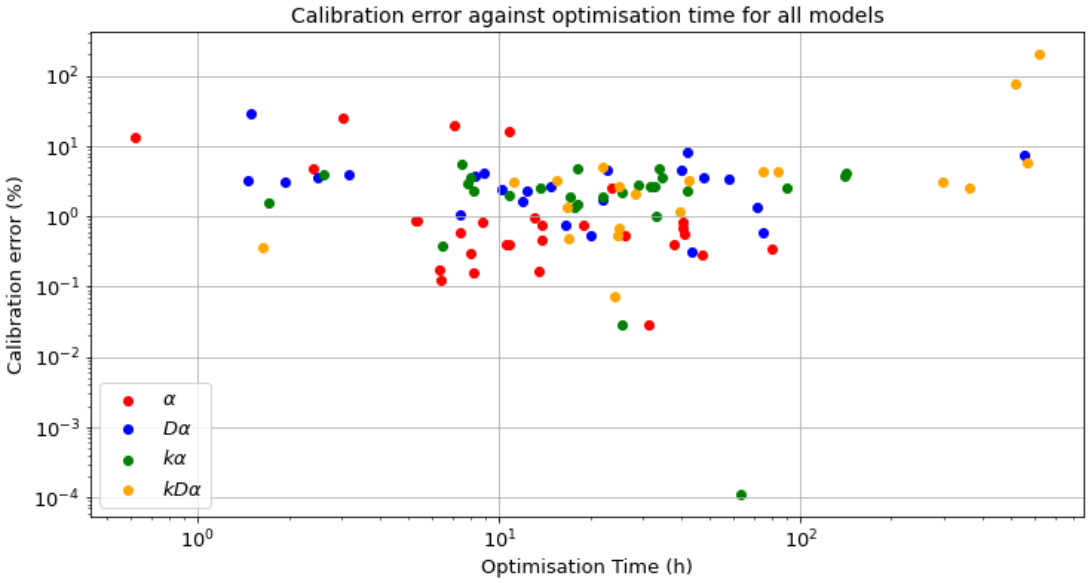


Figure 5.14: Minimum error (tumour volumes) attained during the optimisation process for each model. Note the logarithmic scales in both axes. Each dot represents an individual patient dataset.

The α model was on average the quickest, with the majority of the optimisations (21 out of 29) completing in under 24 h and seven of the remaining eight within 48 h. The median optimisation time was 10.8 h with an interquartile range of $[6.75, 28.65]$.

The $k\alpha$ model generally required more time to converge than the α model, reflecting its increased complexity. Despite this, 13 cases completed within 24 hours and an additional nine within 48 hours. Two cases were unable to converge to a solution in 30 days: patient 9 and patient 14. Both had meshes exceeding one million elements, resulting in extremely long runtimes for a single iteration of the optimisation process: 29 hours 47 minutes for patient 9 and 33 hours 43 minutes for patient 14. These times are significantly longer than the average for the other patients, 6 hours and 13 minutes. These cases, along with the 30 day limit, are discussed further in Appendix C. For the cases that did converge, the median optimisation duration was 22.1 h with an interquartile range of $[8.2, 34.0]$.

The $D\alpha$ model had yet more difficulty achieving convergence, with six cases failing to complete within a month (patients 7, 17, 18, 20, 21 and 31). While patients 18 and 20 experienced extremely long runtimes per forward pass (22 hours 58 minutes and 24 hours 17 minutes, compared to the average 3 hours and 34 minutes of the other patients), the other problematic cases did not have unusually long runtimes. For patients 7 and 17, this could be explained by the significant shrinkage experienced by the tumour between scans 0 and 1, with volume reductions of 66.9% and 48.9% respectively (see Table 5.7). However, comparable reductions in volume occurred for patients 36, 23 and 24 which the $D\alpha$ model successfully calibrated in 1 hour 29 minutes, 11 hours 52 minutes and 22 hours 2 minutes respectively. Patient 17 eventually converged (in just under 23 days), thus those results were included here. Among the 24 converged cases, 16 completed within 24 hours and an additional four finished within 48 hours. The median optimisation duration for all completed cases was 15.65 hours with an interquartile range of $[7.625, 43.125]$.

Finally, the $kD\alpha$ model had the lowest success rate, with only 51.9% of the optimisations (14/27) converging within a month and a further five within 30 days. For

those that did converge, the median optimisation time was 26.55 h with an interquartile range of $[18.225, 243.25]$.

All instances of non-convergence across models are discussed in Appendix C.

5.7.2 Quality of the optimisations

For the α model, 24 of the 29 optimisations achieved an error smaller than 1% and terminated for that reason. Of the remaining six cases, four terminated due to the relative change in parameters falling below the specified 10^{-6} tolerance (patients 14, 19, 20 and 34) and two terminated due to the relative change in the cost function meeting the same criterion (patients 11 and 39). As a result, the optimisation error was higher for these patients.

Five of these cases involve tumours whose volume appeared to increase between scans 0 and 1: patient 14's tumour "grew" by 1.6%, patient 39 by 5.8%, patient 20 by 19.5%, patient 34 by 23.4% and patient 19 by 34.3%. These figures are listed in Table 5.7. Given that the α model assumes a chemotherapy-only scenario, which accounts solely for tumour shrinkage, it is reasonable that the error is higher in these cases. This hypothesis is supported by the observation that none of the tumours successfully optimised with an error below 1% exhibited volume growth between scans 0 and 1. Furthermore, the tumours with the largest volume increases correspond to the highest optimisation errors, as follows: $\epsilon_{opt} = (2.6\%, 4.9\%, 16.2\%, 20.1\%, 24.7\%)$ in the same order as listed above. This topic is further explored in the prediction results discussion (Subsection 5.9.2).

The error for patient 19 could further be attributed to registration error. This patient was not excluded as the overlap between registered tumours was still significant – 54% of the cycle 0 tumour volume and 71% of the cycle 1 tumour volume, with a Dice similarity coefficient between the tumour volumes of 0.61. However, it is possible that the chemotherapy-only model would struggle to match the volumes, given the misalignment in the "tail" of the tumour in the visualisation in Figure 5.15.

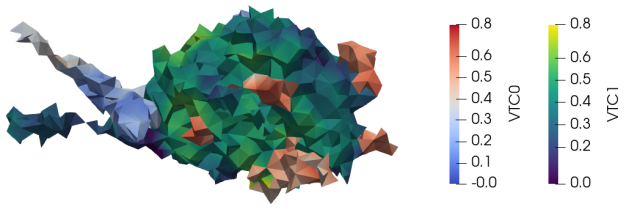


Figure 5.15: Example of a poor registration between data from scans 0 and 1 (patient 19). This patient is an outlier in the α model optimisation process, with unusually high error (24%).

Lastly, the higher-than-usual error for patient 11 (13.6%) does not have an immediately apparent explanation. The optimisation algorithm obtains an identical volume error for the first two α values tested and consequently terminates. This may simply be attributed to an unfavourable starting point for this particular patient.

To avoid excessively long optimisation times, the stopping criterion for the cost function evaluation was relaxed from 1% to 5% for the subsequent models. Only one case exceeded this threshold for the $k\alpha$ model: patient 39, with a calibration error of 5.5%. Most of the other patients for whom the α model had struggled to find optimal parameters were largely "fixed" under this model, achieving errors of 2.0% (patient 11), 2.6% (patient 29), 4.8% (patient 20) and 1.9% (patient 34). This can likely be attributed to the model's additional capabilities of representing growth.

Similarly, the $D\alpha$ model exhibited fewer outliers in calibration error compared to the α model. Two cases exceeded the 5% tolerance: patient 36 and patient 17, with errors of 29.2% and 7.4% respectively. These slightly higher errors are likely related to the shrinkage that these tumours underwent between scans 0 and 1, a behaviour that a diffusion-heavy model may struggle to accurately represent. Patient 36 experienced the most significant shrinkage across all patients (67.6%, detailed in Table 5.7). Patient 7 underwent comparable shrinkage (66.9%) and failed to converge to a solution within the allotted time, making it an "unfinished" case for this model. However, patient 23

was able to converge successfully and achieve a calibration error of just 1.6% despite undergoing a comparable shrinkage of 66.5% between scans 0 and 1. A similar pattern was noticed for the $kD\alpha$ model, where optimisation errors are high for patients 7 and 36 (206.0% and 76.5%) and patient 17 failed to converge.

Lastly, Tables 5.3-5.6 present the median and interquartile ranges for the optimisation error across the models. Despite the outliers discussed earlier, the α model achieves the lowest errors overall – but a bias was introduced in constraining this model to reach errors below 1%, rather than the 5% threshold applied to the other models. The α model calibration errors are very slightly lower for responders (median value 0.55%) compared to non-responders (median value 0.7%). Although admittedly a very marginal difference, this aligns with expectations for a model that exclusively accounts for tumour shrinkage and lacks any growth term. The $k\alpha$ model, which includes a proliferation but no diffusion term, exhibits a similar pattern with a 2.3% median error for responders and 2.6% for non-responders. The trend continues with the $kD\alpha$ model, presenting perhaps surprisingly a median error of 2.6% for responders and 3.1% for nonresponders. The diffusion-heavy models may have been expected to perform better for the nonresponders. This is the case for the $D\alpha$ model, with errors of 3.1% for nonresponders compared to 3.25% for responders.

5.7.3 Calibrated parameters

All the calibrated parameters for each model and for each patient are listed in the first few columns of Tables 5.10–5.13. In addition to the calibrated parameters themselves, Tables 5.3–5.6 include the median and interquartile range of the calibrated parameters scaled by the relevant data in the model equations: $k \cdot \overline{K^{trans}}$ for models incorporating proliferation and $\alpha \cdot \overline{C_{drug}^{tissue}}$ for the chemotherapy term. The mean values are used here as both K^{trans} and C_{drug}^{tissue} are spatially varying quantities.

The values in Tables 5.3–5.6 are reported for all patients as well as for responder and non-responder groups, determined based on RECIST guidelines evaluated at Scan 3 [109]. Figures 5.16–5.18 illustrate the scaled calibrated parameters for each dataset

with individual data points colour-coded according to their RECIST classification. (The figure for the $kD\alpha$ model was omitted as the three-dimensional data was too difficult to visualise in two dimensions.) These visualisations provide further insight into the distribution of parameters across the responder and non-responder groups in a patient-by-patient manner.

Comparing the α model results in Figure 5.16 with Figure 3.11, we observe that patients with the highest calibrated α values tend to have particularly low drug concentrations, while the low-α outliers exhibit higher drug concentrations. This suggests that α can be calibrated to compensate for low C_{drug}^{tissue} values when the tumour is observed to reduce in volume. Comparing α values across models shows that α can also be adjusted to account for growth effects, as it presents higher values in the diffusion-heavy models than in the chemo-heavy models. Similarly the diffusion coefficient D tends to be calibrated to slightly lower values in the $kD\alpha$ model, where proliferation also contributes to simulating growth, than in the $D\alpha$ model where growth is driven by diffusion alone.

Differentiating between responders and nonresponders, responders exhibited higher calibrated α values for the α model (median 15.9 mL/(mg·s), interquartile range $[5.8, 119.9]$) compared to nonresponders (median 10.5 mL/(mg·s), interquartile range $[2.35, 17.73]$). For the $k\alpha$ model, both K^{trans} and C_{drug}^{tissue} were systematically higher in responders, which makes sense as these parameters highlight regions of increased blood flow, facilitating greater chemotherapy delivery. The dimensionless proliferation coefficient k was slightly higher in nonresponders, which aligns with expectations. Surprisingly, higher α values were not indicative of positive response until scaled by C_{drug}^{tissue}. The $D\alpha$ model reveals a slightly higher diffusion coefficient and significantly lower α value in nonresponders, which are both consistent with expectations. As in the $k\alpha$ model, no conclusion is drawn from the α values until scaled by C_{drug}^{tissue}, no conclusion from D, and slightly more proliferation for the nonresponders than the responders.

An important point to note is that in both diffusion-heavy models the diffusion coefficient D is frequently calibrated to its lowest allowable value, 1×10^{-14} m^2/s. This

	$\times 10^{-5}$ α [mL/mg·s]	$\times 10^{-3}$ $\alpha \overline{C_{drug}^{tissue}}$ [/s]	ϵ_{opt} (%)
All patients	11.7 [4.95, 23.25]	3.82 [1.46, 6.75]	0.6 [0.3, 0.9]
Responders	15.9 [5.8, 119.9]	8.17 [4.43, 15.41]	0.55 [0.325, 0.775]
Nonresponders	10.5 [2.35, 17.73]	2.04 [0.58, 3.52]	0.7 [0.3, 9.25]

Table 5.3: Median and quartile ranges for the calibrated α, calibrated α multiplied by the mean drug concentration measured in the tissue at Scan 0 and optimisation error of the α model. The responder and nonresponders categories are determined by the RECIST status evaluated at Scan 3.

is the lower bound imposed on D during the optimisation, chosen because at this level it has no measurable effect on the solution on its own. Although diffusion is permitted to reach an upper limit of 1×10^{-12} m^2/s, the highest calibrated values observed in practice were 2.6×10^{-14} m^2/s for the $D\alpha$ model and 2.5×10^{-14} m^2/s for the $kD\alpha$ model. This is an important point which suggests that diffusion may not be a valuable mechanism to include in these models for this specific application.

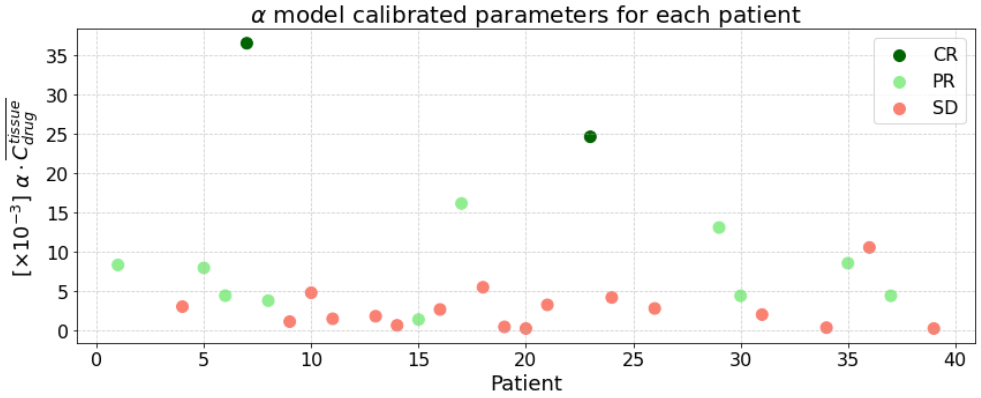


Figure 5.16: Calibrated α values multiplied by the mean drug concentration estimated in the tissue from Scan 0, for each patient dataset. The colours correspond to the RECIST evaluation at Scan 3: CR = complete response, PR = partial response, SD = stable disease.

	$\times 10^{-8}$ k [–]	$\times 10^{-6}$ $k\overline{\mathrm{K}^{trans}}$ [/s]	$\times 10^{-5}$ α [mL/mg·s]	$\times 10^{-3}$ $\alpha\overline{C_{drug}^{tissue}}$ [/s]	ϵ_{opt} (%)
All patients	5 [1, 5.8]	11.42 [3.84, 18.04]	13.8 [10, 38.7]	4.45 [3.60, 7.64]	2.6 [1.5, 3.7]
Responders	5 [1.5, 5.0]	11.79 [5.44, 16.65]	10 [10, 27.75]	6.63 [4.16, 15.76]	2.3 [1.25, 3.7]
Nonresponders	5.30 [1.0, 8.0]	7.49 [2.17, 22.96]	15.3 [10, 53.90]	4.23 [2.59, 4.87]	2.6 [1.9, 4.0]

Table 5.4: Median and quartile ranges for $k\alpha$ model parameters grouped by responder status, as defined by the RECIST evaluation at Scan 3. The parameters include the calibrated values (k and α), their products with the mean of the data they scale in the model equation (5.2), $k\overline{\mathrm{K}^{trans}}$ and $\alpha\overline{C_{drug}^{tissue}}$, and the optimisation error of the $k\alpha$ model.

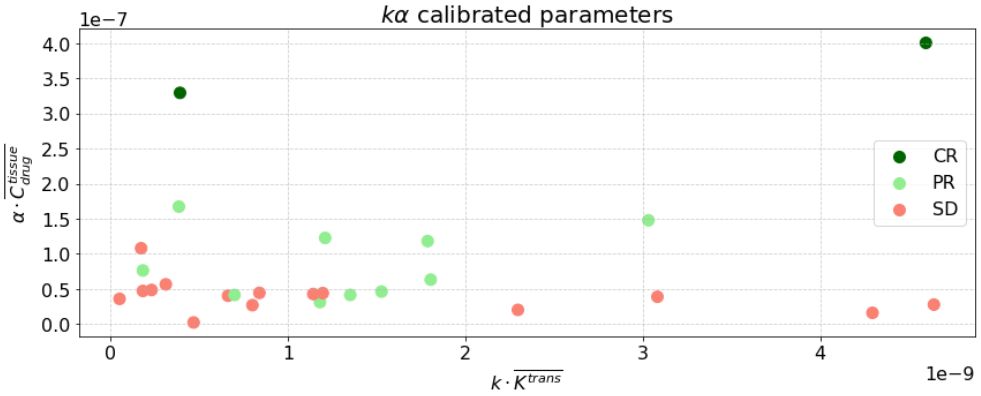


Figure 5.17: Calibrated k and α values multiplied with the mean of the data they scale in the $k\alpha$ equation, $k\overline{\mathrm{K}^{trans}}$ and $\alpha\overline{C_{drug}^{tissue}}$. The colours correspond to the RECIST evaluation at Scan 3: CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease.

	$\times 10^{-14}$ D [m^2/s]	$\times 10^{-5}$ α [mL/mg·s]	$\times 10^{-3}$ $\alpha\overline{C_{drug}^{tissue}}$ [/s]	ϵ_{opt} (%)
All patients	1.25 [1.03, 1.8]	649.75 [400, 774.63]	186.45 [59.67, 293.89]	3.15 [1.38, 3.98]
Responders	1.2 [1, 1.9]	764.38 [462.44, 774.63]	272.02 [125.55, 551.41]	3.25 [0.975, 5.225]
Nonresponders	1.5 [1.1, 1.6]	540.48 [400, 680.97]	127.15 [51.25, 200.21]	3.15 [1.525, 3.675]

Table 5.5: Median and quartile ranges for $D\alpha$ model parameters grouped by responder status, as defined by the RECIST evaluation at Scan 3. The parameters include the calibrated values (D and α), the product of α with the data it scales in the $D\alpha$ equation, $\alpha\overline{C_{drug}^{tissue}}$, and the optimisation error of the $D\alpha$ model. Note that the scaled data refers to the mean value of the spatially-varying data.

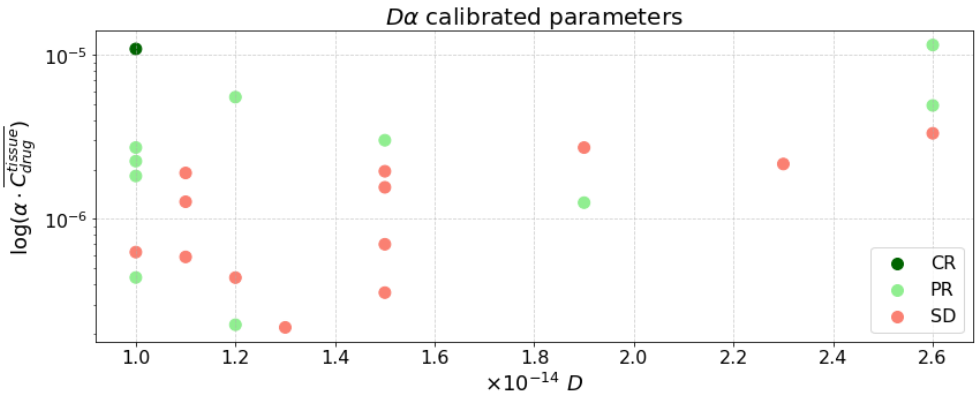


Figure 5.18: Calibrated α values multiplied by the mean drug concentration $\overline{C_{drug}^{tissue}}$ against calibrated D values for each patient dataset using the $D\alpha$ model. Note the log scale in the y-axis. The colours reflect the RECIST evaluation at Scan 3: CR = complete response, PR = partial response.

5.8 Prediction results: scan 3

After calibrating with data from scans 0 and 1, the models were populated with the calibrated parameters and the data from Scan 1 to generate a prediction at Scan 3. Due to the low drug concentration data measured for patients 8 and 13 at Scan 1, these datasets could not be used for the prediction step. This limitation is discussed

	$\times 10^{-8}$ k [–]	$\times 10^{-5}$ $k\mathrm{K}^{trans}$ [/s]	$\times 10^{-14}$ D [m^2/s]	$\times 10^{-5}$ α [mL/(mg·s)]	$\times 10^{-3}$ $\alpha \cdot (C_{drug}^{tissue})$ [/s]	ϵ_{opt} (%)
All pats.	5.15 [2.4, 7.2]	73.4 [46.71, 191.65]	1 [1, 1.8]	649.8 [400, 972.8]	167.18 [59.51, 237.90]	2.9 [0.8, 4.4]
Resp.	4.2 [2.5, 7.25]	75.96 [60.84, 230.89]	1 [1, 1.6]	649.8 [358.55, 970.7]	186.53 [81.0, 236.2]	2.6 [0.5, 4.65]
Nonresp.	5.3 [2.3, 7.2]	52.97 [37.94,171.32]	1 [1, 1.9]	649.8 [400, 981]	86.20 [58.65, 243.58]	3.1 [1.4, 4.3]

Table 5.6: Median and quartile ranges for $kD\alpha$ model parameters grouped by responder status, as defined by the RECIST evaluation at Scan 3. The parameters include the calibrated values (k, D and α), their products with the data they scale in the model equation (5.4), $k\overline{\mathrm{K}^{trans}}$ and $\alpha\overline{C_{drug}^{tissue}}$, and the optimisation error of the $kD\alpha$ model. Note that the scaled data refers to the mean values of the spatially-varying data.

in Section 5.4, in Section 5.10 and in Appendix B.

For the rest of the cohort, a RECIST evaluation [109] was performed for each model prediction and compared to the RECIST assessment derived from the data, providing a qualitative measure of each model's ability to predict response or non-response to NACT (Subsection 5.8.1). In addition, the accuracy of the model predictions was quantitatively evaluated using tumour volume and longest axis as key measures of tumour size. Two evaluation approaches were employed: comparing predicted values to measurements at Scan 3 (Subsection 5.8.2), and assessing the predicted tumour evolution over three cycles of NACT against the measurements from Scan 3 (Subsection 5.8.3).

5.8.1 RECIST evaluation

RECIST evaluations are ordinarily based on changes in the tumour's longest axis [109]. A tumour is classified as achieving complete response (CR) if it disappears entirely, partial response (PR) if its longest axis decreases by at least 30% and progressive disease (PD) if it increases by at least 20%. Tumours that exhibit neither sufficient shrinkage nor growth to qualify as PR or PD are classified as stable disease (SD). A well-known limitation of this approach is that the longest axis is not the most reliable measure of treatment response [110]. It provides only a two-dimensional assessment of a three-dimensional problem and, in clinical practice, is often measured on MR slices

ID	Scan 0 to 1	Scan 1 to 3	Ratio
1	−43.43	−70.57	0.62
4	−14.22	−49.00	0.29
5	−23.52	−78.53	0.30
6	−27.82	−67.44	0.41
7	−66.86	−100.00	0.67
8	−38.26	−77.57	0.49
9	−11.53	23.04	−0.50
10	−31.27	−39.15	0.80
11	−12.36	−54.27	0.23
13	−39.46	−17.53	2.25
14	1.68	−5.13	−0.33
15	−5.97	−73.45	0.08
16	−15.35	−22.24	0.69
17	−48.94	−92.61	0.53
18	−28.66	16.54	−1.73
19	34.25	8.06	4.25
20	19.47	−27.62	−0.71
21	−13.62	−34.30	0.40
23	−66.49	−100.00	0.66
24	−47.82	15.70	−3.05
26	−2.95	−48.80	0.06
29	−34.19	−81.54	0.42
30	−44.19	−70.73	0.62
31	−5.56	−17.83	0.31
34	23.35	16.69	1.40
35	−17.06	−75.31	0.23
36	−67.64	100.98	−0.67
37	−16.28	−61.26	0.27
39	5.78	−48.24	−0.12

Table 5.7: Percentage change in tumour volume for each patient from Scan 0 to Scan 1, from Scan 1 to Scan 3, and the ratio of the first change over the second. Negative values represent decrease and positive values represent increase in volume.

– potentially overlooking the possibility that the longest axis lies in the slice direction. A volume-based approach is used here, applying equivalent criteria for CR, PR, SD and PD based on volume [92]. Assuming spherical tumours, their volume is given by $\frac{4\pi}{3}\left(\frac{\text{LA}}{2}\right)^3$, where LA is the longest axis. A 20% increase in longest axis thus yields the following change in volume:

$$\frac{4\pi}{3}\left(\frac{\text{LA}}{2}+\frac{0.2\times\text{LA}}{2}\right)^3 = 1.73\times\frac{4\pi}{3}\left(\frac{\text{LA}}{2}\right)^3, \tag{5.19}$$

and tumour volume must increase by 73% for a patient to be marked as PD. Following the same reasoning, tumour volume must reduce by at least 66% for PR.

Table 5.9 shows the volume-based RECIST evaluations performed for each model prediction and those derived from the Scan 3 data. The results in terms of the original RECIST criteria, based on the longest axis, are shown in Appendix D.

Given that our models are intended to identify and cease treatments predicted to be ineffective, and that incorrectly discontinuing an effective treatment due to a misclassification would be highly detrimental, the negative predictive value (NPV) is particularly important. It is determined by the number of true negatives (TN; correctly identified nonresponders) and false negatives (FN; responders misidentified as nonresponders), calculated as:

$$\text{NPV} = \frac{|TN|}{|TN|+|FN|}. \tag{5.20}$$

Denoting correctly identified responders as true positives (TP) and nonresponders misidentified as responders as false positives (FP), we can also calculate positive predictive values (PPV), sensitivities (Sens.) and specificities (Spec.) for each model:

$$\text{PPV} = \frac{|TP|}{|TP|+|FP|}, \tag{5.21}$$

$$\text{Sens.} = \frac{|TP|}{|TP|+|FN|}, \tag{5.22}$$

$$\text{Spec.} = \frac{|TN|}{|TN|+|FP|}. \tag{5.23}$$

These values are reported for each model in Table 5.8.

Eleven patients of the calibrated 27 were determined to be responders (two CR and nine PR) and 16 were nonresponders (all SD) at Scan 3. Looking at each model's predictions:

- The α model identified nine of the 11 responders, including one of two CR and seven of nine PR. The misclassifications were limited to one CR case predicted as PR (so, still a responder predicted as a responder), and SD cases predicted as PR or vice versa. It also correctly identified 12 of the 16 SD. Overall, the model achieved an accuracy of 21/27 (77.8%) when categorising patients as either responders or nonresponders and 19/27 (70.4%) when distinguishing between all four categories, and a negative predictive value of 85.7%.

- Two of the SD patients were failed optimisations for the $k\alpha$ model, leaving 14 SD patients in this cohort. The model identified eight of the 11 responders, including one of two CR and six of nine PR. It also correctly identified ten of 14 SD. Again, the misclassifications were limited to one CR case predicted as PR (still a responder predicted as a responder), and SD cases predicted as PR or vice versa. The model achieved an overall accuracy of 18/25 (72.0%) when categorising patients as either responders or nonresponders, and 17/25 (68.0%) when distinguishing between all four categories, and a negative predictive value of 76.9%.

- The failed optimisations of the $D\alpha$ model were one CR case and four SD, leaving ten responders (one CR, nine PR) and twelve SD. The model identified six of the ten responders, including zero of one CR and five of eight PR. It also correctly identified ten of twelve SD. Again, the misclassifications were limited to the CR case predicted as PR, and SD cases predicted as PR or vice versa. Overall, the model achieved an accuracy of 16/21 (76.2%) when categorising patients as either responders or nonresponders and 15/21 (71.4%) when distinguishing between all four categories, and a negative predictive value of 71.4%.

- Lastly, the failed optimisations of the $kD\alpha$ model were one CR case, one PR

and six SD, leaving nine responders (one CR, eight PR) and ten SD. The model identified only one of the nine responders, incorrectly marking the other eight as SD. It correctly identified all ten of the SD cases, thus achieving an accuracy of 11/19 (57.9%) whether using two or four categories, and a negative predictive value of 55.6%. It is worth pointing out that, except for patient 30 which was correctly identified as PR, all other patients are predicted to be SD by this model.

In conclusion, all models except the $kD\alpha$ model demonstrated reasonable accuracy in predicting RECIST status. The α model exhibited the best predictive performance followed by the $k\alpha$, $D\alpha$ and $kD\alpha$ models. This holds true both for distinguishing between responders and non-responders as well as for classifying CR, PR, SD and PD. To ensure a fair comparison, we can focus solely on the subset of patients that successfully calibrated across all models. These are patients 1, 4, 5, 6, 11, 15, 16, 24, 29, 30, 34, 35, 36, 37 and 39, among which there are eight PR and seven SD. Within this group, the α model identifies 6/8 PR and 5/7 SD; the $k\alpha$ model identifies 5/8 PR and 5/7 SD; the $D\alpha$ model identifies 4/8 PR and 6/7 SD; and the $kD\alpha$ model identifies 1/8 PR and 7/7 SD. These results indicate that diffusion-heavy models perform slightly better for predicting non-responders, whereas chemo-heavy models are more effective at identifying responders.

	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
$\boldsymbol{\alpha}$	81.8	75.0	69.2	85.7
$\boldsymbol{k\alpha}$	72.7	71.4	66.7	76.9
$\boldsymbol{D\alpha}$	60.0	83.3	75.0	71.4
$\boldsymbol{kD\alpha}$	11.1	100.0	100.0	55.6

Table 5.8: Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for each model, evaluated based on the RECIST evaluation at Scan 3. The values are based on classifying patients as either responders (CR and PR) or nonresponders (SD only, as there were no PD cases).

Patient	α	$k\alpha$	$D\alpha$	$kD\alpha$	Scan 3	RCB (Scan 6)
1	PR	PR	PR	SD	PR	0
4	SD	SD	SD	SD	SD	3
5	PR	SD	SD	SD	PR	2
6	PR	PR	PR	SD	PR	1
7	CR	CR	–	SD	CR	0
9	SD	–	SD	–	SD	3
10	PR	PR	PR	–	SD	1
11	SD	SD	SD	SD	SD	1
14	SD	–	SD		SD	2
15	SD	SD	SD	SD	PR	2
16	SD	SD	SD	SD	SD	3
17	PR	PR	PR		PR	1
18	PR	PR	–	SD	SD	2
19	SD	SD	SD	SD	SD	1
20	SD	SD	–	–	SD	3
21	SD	SD	–	–	SD	2
23	PR	PR	PR	–	CR	0
24	PR	PR	SD	SD	SD	1
26	SD	SD	SD	–	SD	3
29	PR	PR	SD	SD	PR	2
30	PR	PR	PR	PR	PR	0
31	SD	SD	–	SD	SD	1
34	SD	SD	SD	SD	SD	2
35	PR	PR	PR	SD	PR	3
36	PR	PR	PR	SD	SD	1
37	SD	SD	SD	SD	PR	2
39	SD	SD	SD	SD	SD	2

Table 5.9: Volume-based RECIST evaluation of each model prediction at Scan 3 (α, $k\alpha$, $D\alpha$, $kD\alpha$), for the data at Scan 3, and residual cancer burden determined after the completion of NACT (RCB). Missing values (–) indicate failed optimisations for that model.

5.8.2 Tumour characteristics

As well as the calibrated parameters discussed in Section 5.7, Tables 5.10-5.13 list the optimisation error (defined as the percentage difference between the numerical tumour volume and the Scan 1 tumour volume), optimisation duration and percentage difference between the Scan 3 data and model predictions for each error measure. All model predictions and corresponding measurements from Scan 3 are provided in Appendix E.

5.8.3 Tumour evolution

Figure 5.19 shows the percentage change in tumour volume and tumour longest axis predicted by each model against the observed percentage change from Scan 0 to Scan 3. The models all provide a more accurate estimation of tumour volume than the longest axis. Notably, most points fall within either the lower-left or upper-right quadrant for the volume error, indicating that when shrinkage (or growth) is observed, the models generally predict shrinkage (or growth) accordingly.

5.9 Discussion

5.9.1 Effect of registration

Figure 5.20 shows the volume prediction error for each model against the Dice similarity coefficient (DSC), defined in Chapter 3, Subsection 3.5.2. The DSC for each patient between scans 0 and 1 and scans 1 and 3 were listed in Table 3.1. The results indicate that patients with a high degree of tumour overlap between Scans 1 and 3 generally exhibit lower prediction errors, suggesting that the quality of the registration plays an important role in achieving accurate predictions. This explains the particularly high errors found in the predictions for patient 5 (DSC = 0.3), patient 29 (DSC = 0.24), patient 17 (DSC = 0.12) and patient 36 (DSC = 0.33) across all models. We see that for these cases, the error generally gets worse as the model complexity increases.

ID	α ($\times 10^{-5}$)	ϵ_{opt} (%)	t_{opt} (hrs)	ϵ_{vol} (%)	ϵ_{LA} (%)
1	14.2	0.4	10.8	−12.3	17.4
4	9.2	0.8	5.3	0.9	−1.0
5	5.4	0.9	13.0	68.2	48.8
6	9.6	0.8	8.8	3.1	22.2
7	340.8	0.7	40.6	*	*
8	4.5	0.6	7.4	†	†
9	14.2	0.3	47.0	−58.9	−10.7
10	20.5	0.7	19.0	−24.2	−21.9
11	16.8	13.6	0.6	123.0	23.0
13	10.5	0.3	8	†	†
14	1.0	2.6	23.5	4.8	−4.1
15	4.5	0.2	13.5	176.2	109.2
16	9.2	0.2	8.2	−29.2	26.0
17	26.0	0.7	13.9	29.1	74.6
18	64.0	0.4	38.1	−93.2	−89.2
19	1.0	24.7	3.0	−11.1	34.7
20	1.0	16.2	10.8	36.3	23.3
21	12.5	0.9	5.2	−20.6	22.7
23	17.6	0.5	26.0	*	*
24	45.4	0.4	10.5	−63.3	−19.5
26	6.6	0.6	40.9	44.3	9.5
29	442.4	0.3	80.5	−62.7	43.5
30	19.7	0.5	13.8	3.0	56.9
31	3.7	0.1	6.4	18.1	−14.2
34	1.0	20.1	7.1	−15.3	10.6
35	151.2	0.2	6.4	67.0	35.2
36	168.6	< 0.1	31.3	−98.8	−81.8
37	7.0	0.8	40.8	33.7	60.2
39	11.7	4.9	2.4	−29.8	10.0

Table 5.10: Results from the α model for each patient, including the calibrated α value, optimisation error (ϵ_{opt}), and optimisation duration in hours (t_{opt}). Simulation errors are reported as the percentage difference between the predicted and actual tumour volumes (ϵ_{vol}) and longest axes (ϵ_{LA}). Missing values (–) denote failed optimisations (see Appendix C), daggers (†) cases unsuited to prediction (see Appendix B), and stars (*) indicate tumours no longer detectable by Scan 3. Negative errors indicate underprediction and positive errors indicate overprediction relative to actual values.

ID	k ($\times 10^{-8}$)	α ($\times 10^{-5}$)	ϵ_{opt} (%)	t_{opt} (hrs)	ϵ_{vol} (%)	ϵ_{LA} (%)
1	5.8	20.1	2.6	31.8	18.2	18.3
4	5.8	13.4	2.6	13.6	9.9	3.6
5	5.0	10.0	3.7	8.0	165.4	40.9
6	5.0	10.0	1.2	1.7	4.7	24.7
7	1.0	307.2	4.2	140.8	0.0	0.0
8	2.0	4.9	< 0.1	111.1	†	†
9	–	–	–	–	–	–
10	1.0	20.1	< 0.1	25.5	26.9	−6.9
11	1.0	53.9	2.0	10.8	110.7	20.6
13	6.7	11.6	2.9	7.8	†	†
14	–	–	–	–	–	–
15	5.0	10.0	0.4	6.5	156.0	87.2
16	5.4	13.8	2.8	28.9	16.2	31.5
17	1.0	26.9	2.3	42.2	23.6	80.5
18	1.0	65.7	2.6	90.0	41.9	−60.6
19	8.0	3.3	2.6	32.7	5.1	33.8
20	8.0	10.0	4.8	34.0	14.6	
21	5.0	16.8	1.9	17.2	36.0	
23	8.0	28.6	3.8	140.5	*	*
24	1.0	38.7	4.0	2.6	62.7	−22.3
26	5.0	10.0	1.5	18.2	49.3	10.2
29	10	414.2	2.2	25.6	148.7	94.9
30	3.4	18.5	3.7	34.5	17.3	58.0
31	2.8	4.9	1.0	33.0	15.2	−11.4
34	8.0	10.0	1.9	22.1	38.9	15.1
35	1.0	134.7	1.4	17.8	66.5	28.1
36	1.0	172.2	4.9	18.1	98.6	−81.8
37	5.0	10.0	2.3	8.2	58.9	60.2
39	5.2	10.0	5.5	7.5	30.9	3.3

Table 5.11: Results from the $k\alpha$ model for each patient, including the calibrated values (k and α), optimisation error (ϵ_{opt}) and optimisation duration in hours (t_{opt}). Simulation errors are reported as the percentage difference between the predicted and actual tumour volumes (ϵ_{vol}) and longest axes (ϵ_{LA}). Missing values (–) denote failed optimisations (see Appendix C), daggers (†) cases unsuited to prediction (see Appendix B), and stars (*) indicate tumours no longer detectable by Scan 3. Negative errors indicate underprediction and positive errors indicate overprediction relative to actual values.

ID	D ($\times 10^{-14}$)	α ($\times 10^{-5}$)	ϵ_{opt} (%)	t_{opt} (hrs)	ϵ_{vol} (%)	ϵ_{LA} (%)
1	1.0	462.4	4.5	22.7	23.2	22.4
4	2.3	649.8	3.6	2.5	61.4	11.8
5	2.6	774.6	2.7	14.7	272.9	67.6
6	1.5	649.8	4.0	8.9	17.2	31.2
7	–	–	–	–	–	–
8	1.2	649.8	0.6	75.3	†	†
9	1.2	540.5	4.6	40.2	−27.6	−8.2
10	1.1	540.5	3.6	47.8	−23.2	−21.1
11	1.1	649.8	0.7	16.6	97.0	29.0
13	1.5	400	3.2	1.5	†	†
14	1.9	400.0	1.3	71.5	1.9	11.7
15	1.9	400.0	2.4	10.2	244.0	90.1
16	1.1	649.8	2.3	12.4	25.2	38.0
17	1.0	293.4	7.4	551.2	222.8	77.5
18	–	–	–	–	–	–
19	1.5	400.0	3.9	3.2	−27.7	40.9
20	–	–	–	–	–	–
21	–	–	–	–	–	–
23	1.0	774.6	1.6	11.9	*	*
24	1.5	382.2	1.7	22.0	−5.2	−10.4
26	2.6	774.6	0.5	20.2	75.2	17.5
29	1.2	764.4	0.3	43.5	221.1	105.4
30	1.0	1000.0	8.3	42.0	36.6	54.3
31	–	–	–	–	–	–
34	1.5	400.0	3.1	1.9	−38.0	19.3
35	1.0	774.6	3.8	8.3	21.4	33.0
36	1.0	1000.0	29.2	1.5	−62.8	−48.1
37	2.6	774.6	1.1	7.4	140.9	70.3
39	1.3	930.7	3.5	57.4	67.2	13.1

Table 5.12: Results from the $D\alpha$ model for each patient, including the calibrated values (D and α), optimisation error (ϵ_{opt}) and optimisation duration in hours (t_{opt}). Simulation errors are reported as the percentage difference between the predicted and actual tumour volumes (ϵ_{vol}) and longest axes (ϵ_{LA}). Missing values (–) denote failed optimisations (see Appendix C), daggers (†) cases unsuited to prediction (see Appendix B), and stars (*) indicate tumours no longer detectable by Scan 3. Negative errors indicate underprediction and positive errors indicate overprediction relative to actual values.

ID	k ($\times 10^{-8}$)	D ($\times 10^{-14}$)	α ($\times 10^{-5}$)	ϵ_{opt} (%)	t_{opt} (hrs)	ϵ_{vol} (%)	ϵ_{LA} (%)
1	2.3	1.0	317.1	2.6	364.7	134.7	21.1
4	2.3	2.5	733.0	1.4	16.8	59.9	8.2
5	7.3	1.9	100.0	< 0.1	24.2	369.5	51.8
6	6.5	1.1	566.5	0.5	17.0	107.3	44.7
7	1.7	1.0	999.9	206.0	618.1	*	*
9	–	–	–	–	–	–	–
10	–	–	–	–	–	–	–
11	7.2	1.0	649.8	2.7	25.0	88.9	29.4
13	5.0	1.5	400.0	0.4	1.6	†	†
14	–	–	–	–	–		
15	7.2	1.0	649.8	1.2	39.5	255.3	101.3
16	3.1	1.0	981.0	4.3	83.5	42.2	38.3
17	–	–	–	–	–	–	–
18	1.7	1.0	999.9	3.1	296.5	−10.7	−32.0
19	7.3	1.9	400.0	3.2	15.5	9.8	43.3
20	–	–	–	–	–	–	–
21	–	–	–	–	–	–	–
23	–	–	–	–	–	–	–
24	7.2	1.0	667.7	2.1	28.1	4.7	−7.8
26	–	–	–	–	–	–	–
29	3.7	1.3	786.8	3.3	42.3	265.2	106.1
30	2.7	1.0	974.9	4.4	74.9	30.4	53.4
31	7.3	1.9	400.0	0.7	25.0	15.9	−5.7
34	1.1	1.0	999.9	5.9	567.7	−38.7	19.9
35	4.2	1.0	966.5	4.9	21.9	68.9	23.2
36	5.3	1.0	563.2	76.5	514.3	−38.0	12.0
37	7.3	1.9	400.0	0.5	24.5	138.6	70.9
39	5.9	1.4	504.1	3.1	11.1	84.9	1.4

Table 5.13: Results from the $kD\alpha$ model for each patient, including the calibrated values (k, D and α), optimisation error (ϵ_{opt}) and optimisation duration in hours (t_{opt}). Simulation errors are reported as the percentage difference between the predicted and actual tumour volumes (ϵ_{vol}) and longest axes (ϵ_{LA}). Missing values (–) denote failed optimisations (see Appendix C), daggers (†) cases unsuited to prediction (see Appendix B), and stars (*) indicate tumours no longer detectable by Scan 3. Negative errors indicate underprediction and positive errors indicate overprediction relative to actual values.

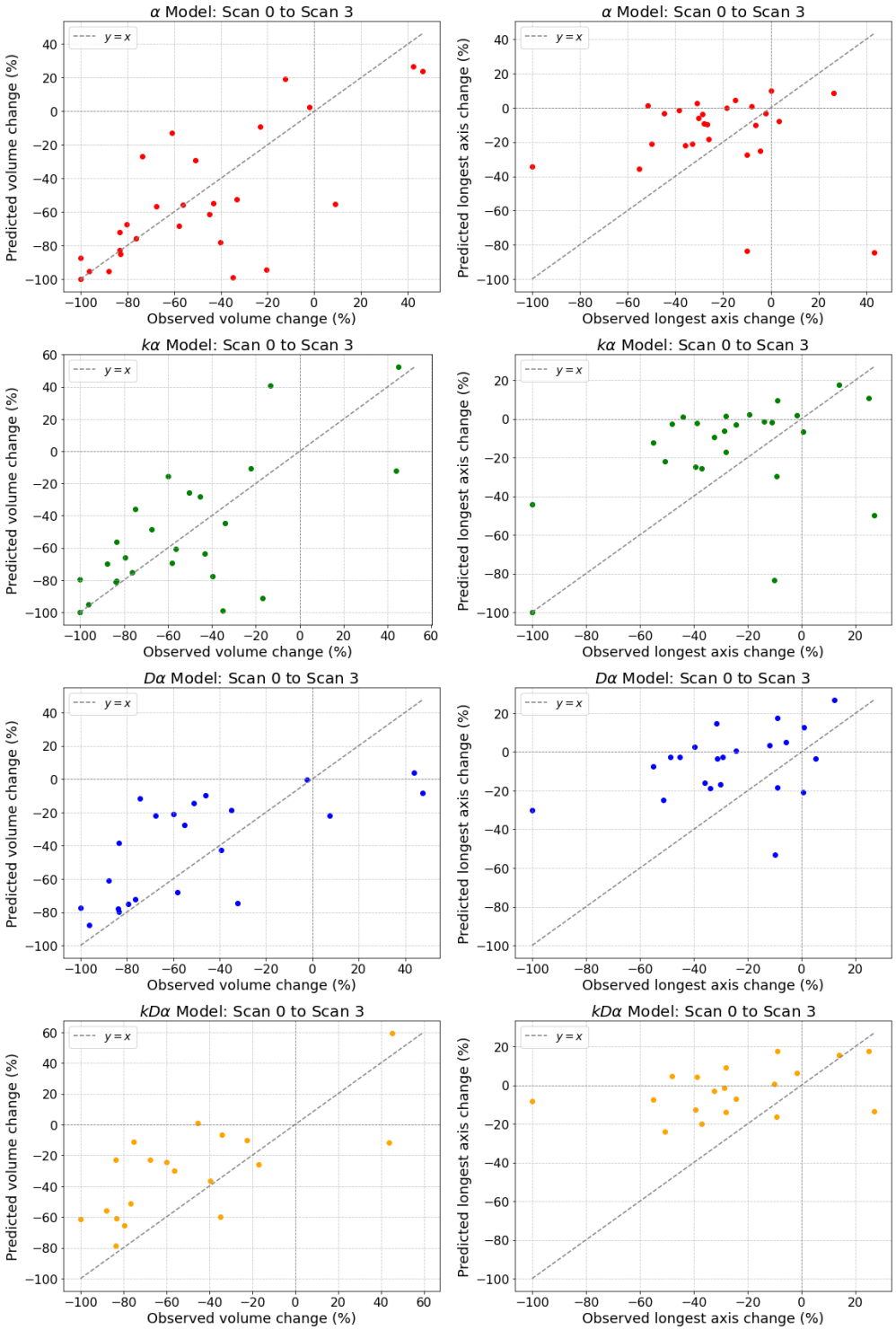


Figure 5.19: Percentage change in tumour volume (left column) and longest axis (right column) from Scan 0 to Scan 3 predicted by each model, against observed percentage change in tumour volume from scan 0 to scan 3. Negative change represents shrinkage; positive change represents growth.

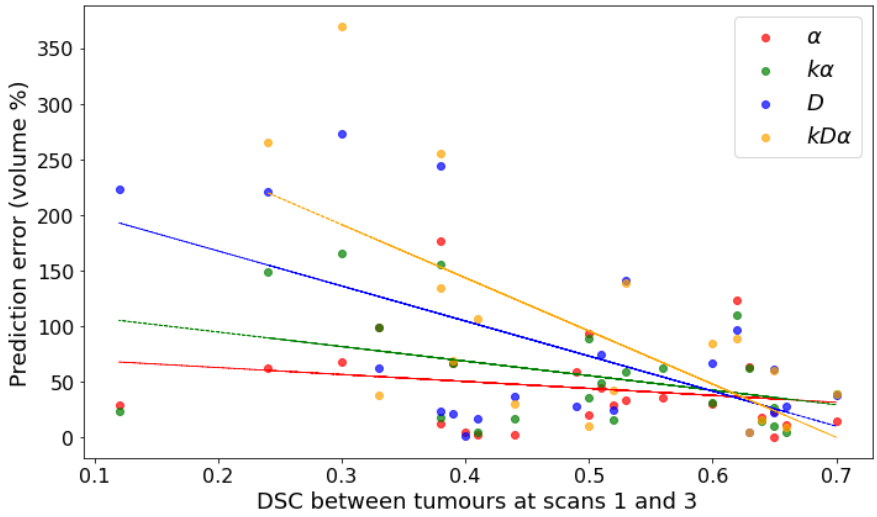


Figure 5.20: Prediction error (percentage difference in tumour volumes) for each model against Dice similarity coefficient (DSC) of the Scan 1 and Scan 3 overlap. The lines are linear best-fits for each model.

5.9.2 Types of tumour response

In the cohort of calibrated patients we observe several distinct patterns of tumour evolution across scans 0, 1 and 3. In this section, we evaluate how well each model predicts tumour volume for each category of behaviour.

Patients with consistent tumour shrinkage from scans 0 to 1 and scans 1 to 3. These are patients 1, 7, 10, 16, 17, 23 and 30. In this category, the Scan 0 to Scan 1 shrinkage is always at least 50% of the Scan 1 to Scan 3 shrinkage – hence "consistent". In most of these cases, volume errors are all below 50%. This low error aligns with expectations, as the models are calibrated on a stable pattern of behaviour that persists in the prediction phase. The exceptions are patient 17 with the $D\alpha$ model (222.8%) and patient 1 with the $kD\alpha$ model (134.7%). In both cases this can be attributed to poor registration, as mentioned above (Subsection 5.9.1).

Patients with consistent tumour growth across scans 0 to 1 and scans 1 to 3. This group includes patients 19 and 34. As with the consistent shrinkage group, volume errors in these cases remain below 50%, again as expected since the tumour's behaviour during prediction matches the calibrated trend. Note that actual tumour growth over the NACT timescale is unlikely, and this apparent growth is likely to result from segmentation errors or uncertainty.

Patients with minimal change between scans 0 and 1, followed by significant change between scans 1 and 3. We defined this category as a Scan 0 to Scan 1 volume change less than 45% of the Scan 1 to Scan 3 volume change. This group comprises patients 4, 5, 6, 11, 15, 21, 26, 29, 31, 35 and 37. Most patients in this group exhibit volume errors exceeding 50%, with the exceptions of patients 21 and 31. These higher errors are expected as the models are calibrated using scans where little change is observed. When a substantial change occurs between scans 1 and 3, it is not surprising that the models struggle to predict it accurately.

Patients with initial tumour growth followed by shrinkage. Patients 14, 20 and 39 fall into this category. For patient 39 the chemo-heavy models perform well, while the diffusion-heavy models perform poorly. This is consistent with the calibration process, as the diffusion-heavy models are biased towards predicting growth. When the tumour unexpectedly shrinks these models struggle to adapt, resulting in higher errors. For patients 14 and 20, only two models converge (α and $k\alpha$ for patient 20; α and $D\alpha$ for patient 14) and these models achieve low errors.

Patients with initial tumour shrinkage followed by apparent growth. This group includes patients 9, 18, 24 and 36. These cases consistently produce the highest errors, particularly for the chemo-heavy models. This result is expected as growth is inherently more challenging for diffusion-less models to predict – especially when the calibration process led the models to anticipate continued shrinkage.

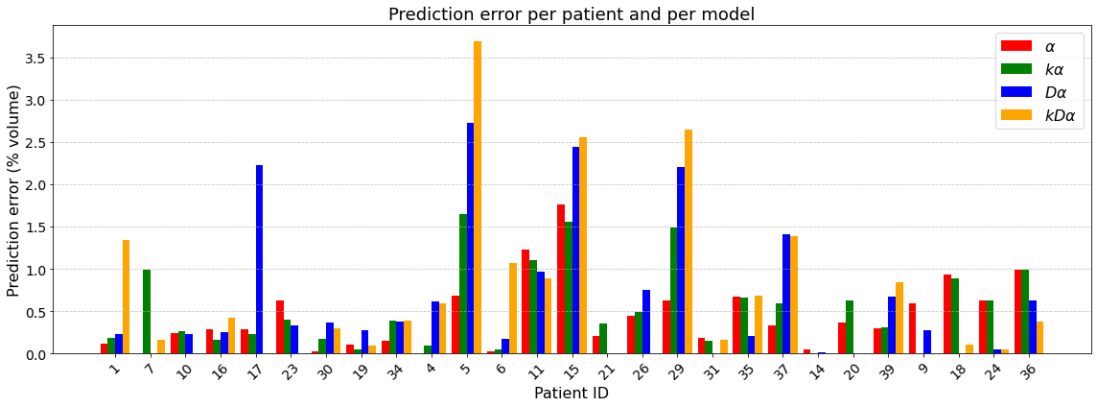


Figure 5.21: Volume error for each patient from each model's prediction at Scan 3. Patients are grouped by type of response, as described in Subsection 5.9.2.

5.10 Conclusions

The key findings from this chapter are summarised here. To avoid immediate repetition, the limitations of our modelling approach are saved for the overall discussion of the thesis, in Chapter 6.

- The α model, which exclusively models the effect of chemotherapy, is significantly faster to calibrate and predict treatment response than the models which include proliferation ($k\alpha$), diffusion ($D\alpha$) or both ($kD\alpha$). It is also the only model capable of doing so for every patient in our cohort in a "workable" timeframe.

- The α model can be more challenging to calibrate than the other models, reflecting that it is the model with the least features. Diffusion-heavy models struggle to capture tumour shrinkage, and the α model struggles to capture growth. In contrast, the $k\alpha$ model is able to handle both scenarios.

- However, there is little correlation between how well a model is calibrated and how accurately it predicts outcomes. That is, a case can have a low prediction error despite a high calibration error, and vice versa. This is shown in Figure F.1 in Appendix F.

- We cannot reliably determine whether a patient will be a responder based solely on the values of the calibrated parameters.
- In terms of distinguishing responders from nonresponders, the α model significantly outperforms all other models, achieving an accuracy of 77.8% and a negative predictive value of 85.7%. This compares to 72.0% and 76.9% for $k\alpha$, 76.2% and 71.4% for $D\alpha$, and 57.9% and 55.6% for $kD\alpha$.
- The single greatest factor in whether a model prediction will be accurate or not is the consistency of the tumour behaviour between scans 0 and 1 and scans 1 and 3. If the tumour behaves consistently between 0–1 and 1–3, the models can make a reasonably good prediction. However, if the tumour's behaviour changes across these intervals – such as growing then shrinking, shrinking then growing, or remaining stable before a sudden change – prediction becomes much harder.

The conclusion from this chapter is that the single-parameter, single-term α model is sufficient for predicting breast tumour response to NACT. It is the fastest, generally the most accurate, and can be calibrated for all patients within a workable time frame. Even with advancements in computing power allowing for the $kD\alpha$ model to complete, its results do not offer improvement. Our data and results strongly suggest that for the specific application of predicting breast tumour response to NACT, incorporating diffusion is not justified.

Chapter 6

Discussion and Conclusion

In this chapter, the key contributions of this work and their conclusions are outlined in Section 6.1, we discuss the limitations of both the data and modelling approaches in Section 6.2 and finally explore perspectives and potential directions for future work in Section 6.3.

6.1 Contributions

The key contributions and conclusions from this work are summarised and discussed here. In short, I propose to reduce the number of calibrated parameters in RD models of breast tumour response to NACT, forego any mechanical coupling, and consider a chemotherapy-only –or, at most, chemotherapy and proliferation-only– approach to breast tumour models for predicting early response to chemotherapy.

6.1.1 Removing the locally calibrated proliferation rates

The primary contribution of this thesis was to remove the local calibration of numerous proliferation rates in the state-of-the-art breast tumour model for treatment outcome prediction [4]. A single scalar proliferation coefficient was introduced instead, scaling a map of K^{trans} and together defining a new proliferation rate for tumour cells. Due to the change in chemotherapy regimen after cycle 4 in the CHERNAC dataset, predicting the final treatment outcome was not possible – a limitation that prevents direct comparison with previous models. Regardless of whether these proposed models outperform previous efforts in terms of accuracy, they offer a viable alternative to the risk of overfitting through matrix calibration [100, 101]. While extensive parameter tuning might still yield the most precise treatment outcome predictions, it does so at the expense of interpretability, making it difficult to discern which of the underlying mechanisms (tumour cell diffusion, proliferation or chemotherapy effects) are driving the results. In contrast, the approach proposed here enables a critical evaluation of each term in the mathematical model, ensuring that only biophysically relevant components are retained.

6.1.2 Ablation study of the DI–RD

Given the above, analysing the four models proposed in this work is effectively an ablation study on our alternative to the DI–MC–RD model: the $kD\alpha$ model. This approach allowed for the assessment of which terms in the equation were most meaningful, concluding that this is the chemotherapy response. Indeed, the simplest α model significantly outperformed all others across all metrics except for the PPV (see Table 5.8). However, the PPV values for the diffusion models were artificially high, as they tended to class nearly all cases as nonresponders. This resulted in very few false positives, making any correctly identified responder a large proportion of the total (see the PPV equation 5.21).

These findings suggest that tumour cell diffusion is not a critical factor in predicting

response to NACT. If tumour growth is driven by a combination of proliferation and diffusion, this aligns with the observation that breast tumours do not exhibit significant growth over the modelled timescales [111, 112]. By the principle of parsimony [101], if tumours do not grow over NACT timescales then a growth term need not be included. However, it is worth noting that for the rare patient whose tumour does grow, omitting a growth term may pose a risk of underestimating progression.

6.1.3 Removing the mechanical coupling

This work posits that mechanical coupling may not be crucial for breast tumour models. Chapter 4 demonstrated that mechanical coupling is not the dominant factor in the test patient dataset provided by the authors [2]. While Weis et al. [6] reported improved predictions using an MC–RD model compared to a simple RD model, neither of these models accounted for drug delivery, and other RD models have not experimented with removing the MC since then. When considering enhancements that introduce challenges to practical applicability, such as the significant computational burden of mechanical coupling, it is important to weigh the potential benefits against the associated costs to determine whether the improvement justifies the added complexity.

Beyond the performance of MC–RD models, mechanical coupling equations in breast tissue models are based on assumptions that may not accurately reflect the complex behaviour of soft tissues. These models assume that tumour tissue behaves as a linearly elastic material, yet soft tissues exhibit nonlinear behaviour with stiffness varying based on strain rate and loading conditions [113–115].

Moreover, mechanical coupling relies on parameters that cannot yet be routinely acquired in a patient-specific manner yet exhibit high patient-specific variability [114]. While MRE can estimate tissue stiffness, it is not widely used and exhibits patient-to-patient variability [68], furthering the need for patient-specific values rather than a population value from previous MRE studies. Other techniques for measuring the Young's modulus of living tissue present further challenges, as standard mechanical testing relies on excised tissue samples, which may not accurately reflect the properties

of tissue *in vivo* [114].

6.2 Limitations

This study faced several technical limitations related to data acquisition and the inherent nature of the data, as well as limitations associated with the mathematical modelling.

6.2.1 Data limitations

Uncertainty quantification in MRI: There are challenges in quantifying uncertainty in MRI data, particularly concerning the applicability of measurements across different hospitals. The consistency of values obtained from varying institutions remains uncertain [116, 117].

Resolution constraints of MRI data impacting segmentation and registration quality: The very high temporal resolution at which DCE-MRI was acquired for the CHERNAC study comes at the cost of spatial resolution. In this context, the spatial resolution of the MRI presents a significant limitation: 4 mm in the slice direction ($z-$axis) and 1.3 mm in the $x-$ and $y-$axes. This resolution leads to considerable uncertainty in tumour segmentation along the $z-$axis, particularly since the longest axes of tumours often fall within this range. For instance, Table E.1 indicates that the median longest axis at Scan 3 measures 25.4 mm – so a segmentation error concerning only one voxel could introduce a 25% error in this tumour's size. Additionally, missing tumour features in the segmentation complicates registration, further affecting the accuracy of our models. Prediction accuracy was shown to be correlated to registration quality (Figure 5.20).

Impact of drug changes in the dataset: A significant limitation arises from the nature of our dataset, in particular from the drug change at cycle 4 whose effect was not captured by a scan. A dataset featuring the same drug or type of drug throughout the

entire chemotherapy regimen would be more relevant for predicting treatment outcome – rather than "mid-point of treatment". Alternatively, an additional scan after the drug change would allow for models to be recalibrated to that drug. As it is, we are only able to predict the response after three cycles and not the outcome at the end of treatment, which limits our ability to compare the performance of the proposed models to others in the literature.

Factors affecting ADC measurements: Although there is compelling evidence supporting a linear inverse correlation between the ADC and tumour cell volume fraction or density, two issues persist. First, numerous factors can influence the ADC values measured, including cell membrane permeability, cell size, tissue tortuosity and gradient nonlinearity [4]. Second, each voxel of tissue contains components beyond just tumour cells. We attempted to mitigate this by imposing a maximum tumour cell volume fraction of 0.75, which reflects the average of the five highest volume fractions recorded in [96].

Sample size: Lastly, it is worth noting that the dataset comprises only 37 patients. While this sample size is comparable to those of other studies (33 in [13], 5 in [4], 18 in [5]), it nonetheless limits the generalisability of our findings. A larger cohort would strengthen the conclusions drawn from this research.

6.2.2 Modelling limitations

Tumour tissue heterogeneity A key simplification in our model is the assumption of a homogeneous tumour tissue. In reality, tumours are composed of heterogeneous subpopulations of cells, each exhibiting distinct growth dynamics and treatment responses [118]. However, MRI data is routinely acquired at the tissue scale and therefore does not give us information on cell populations. While hybrid approaches are promising, they are not compatible with routinely acquired data as of yet [25, 26].

Low C_{drug}^{tissue} cases The variation in the estimated tissue drug concentration poses a significant challenge for modelling. In certain patients (e.g. patients 5 and 23), C_{drug}^{tissue} is so high that the upper bound on α must be restricted to 10^{-2} to prevent the code from crashing. This constraint means that for patients where C_{drug}^{tissue} is below 10^{-6} the chemotherapy efficacy α cannot be calibrated to a sufficiently large value for it to take effect, resulting in little to no chemotherapy reaching the tumour in the simulation. Perhaps instead of excluding the patients with low concentration, we could have excluded the patients with high concentration and raised the upper bound on α. This would have increased our cohort by four patients.

Tumour behaviour between Scans 0–1 and 1–3 The models struggle to predict cases where the tumour behaviour between Scans 0 and 1 doesn't match the behaviour between Scans 1 and 3. This includes cases where minimal change occurs between Scans 0 and 1 followed by a substantial change between Scans 1 and 3, and cases where growth is followed by shrinkage or *vice versa*.

Registration and segmentation: As discussed in Subsection 6.2.1, the accuracy of registration and segmentation is crucial for model performance. The cost function relies on tumour volume measurements, meaning that errors in segmentation that lead to uncertainty in tumour volume will have a profound impact on the model prediction. Furthermore, artificial tumour growth resulting from imprecise segmentation was frequently observed in our dataset. This was observed to negatively affect both the model calibration and performance (see Subsection 5.9.2). Improvements in registration and segmentation techniques for breast tumours would therefore improve the accuracy of our modelling by providing more precise spatial information, leading to better-informed predictions.

Starting points of the optimisations It is possible that some model optimisations converged to local minima due to the choice of initial parameter values. Several starting points were tested for each model, but only those presented in Section 5.5 were docu-

mented formally in the thesis. The intent was to start each parameter at a value where it would have a measurable impact on the solution, to avoid calibrating such a small value of D that the $D\alpha$ model would reduce to the α model. A more comprehensive analysis of the parameter space, possibly using randomised starting points, could help mitigate this issue.

Mesh refinement While our current FEM meshes are of approximately uniform density, adaptive meshing could potentially significantly enhance efficiency. Refining the mesh specifically at the tumour–breast boundary would conserve accuracy in regions where fine resolution is most critical, while reducing computational cost in less relevant regions. This targeted approach could ultimately lower computational costs, an ever-relevant concern when predicting early response to therapy. Although this approach was explored, integrating it into the workflow proved impractical within the available time frame and was ultimately abandoned.

Computational resources Computational resources posed a significant limitation. The cluster I used allows each user to run only two "long" jobs (up to seven days) at any given time. With a dataset of 35 patients, four models to evaluate and additional jobs for testing different starting points and optimisation algorithms, this restriction was unworkable. The two long jobs were effectively useless given the hundreds of simulations required, leaving me reliant on the "normal" jobs capped at 48 hours. This required relaunching jobs every two days, with no guarantee of obtaining the same type of node. This makes it difficult to compare runtimes as the same simulation could be allocated different nodes every 48 hours, leading to fewer iterations in one period compared to another.

Furthermore, FEniCSx is not currently parallelised on the Aion supercomputer. Implementing this model in legacy FEniCS or in another library that can be parallelised would significantly improve computational times.

The deal.II library [119] may be an ideal framework for future implementations of these models. This C++ software library facilitates the development of parallelised FE

codes and can automatically refine the mesh in critical regions, directly addressing two of the implementation limitations discussed above.

6.3 Perspectives

Several suggested directions for future work emerge from Section 6.2. These include conducting further experiments with the models proposed herein, as well as exploring alternative modelling approaches.

It would be interesting to apply our models to a dataset with consistent drug administration throughout, or to one that includes additional scans after a drug change to enable recalibration. However, in the latter case, this would negatively impact the 'early' prediction of response as we would need to wait longer for the subsequent scan. Additionally, as previously mentioned, a more thorough investigation of the parameter space could be beneficial, including starting optimisations from randomised initial points. Reimplementing the models in deal.II [119], which appears more capable than FEniCSx, could also accelerate computational times through parallelisation and adaptive mesh refinement.

To improve and modify the models further, a multiphase poromechanics approach could be considered to better account for the inherent heterogeneity within tumour tissue. Such an approach would treat each element as a mixture of different cell types – introducing, however, a heavy computational burden.

The most significant improvement I envision for this work is the inclusion of chemotherapy diffusion, which I was unable to complete within the timeframe of my project. The tissue drug concentration is modelled as its own variable $C(\mathbf{x}, t)$, initialised by the values measured as in Subsection 3.5.3. The following set of coupled equations is proposed:

$$\frac{\partial V_{TC}(\mathbf{x},t)}{\partial t} = D\nabla^2 V_{TC}(\mathbf{x},t) + kK^{\text{trans}} V_{TC}(\mathbf{x},t)\left(1 - \frac{V_{TC}}{\theta}\right) - \alpha C(\mathbf{x},t) V_{TC}(\mathbf{x},t), \quad (6.1)$$

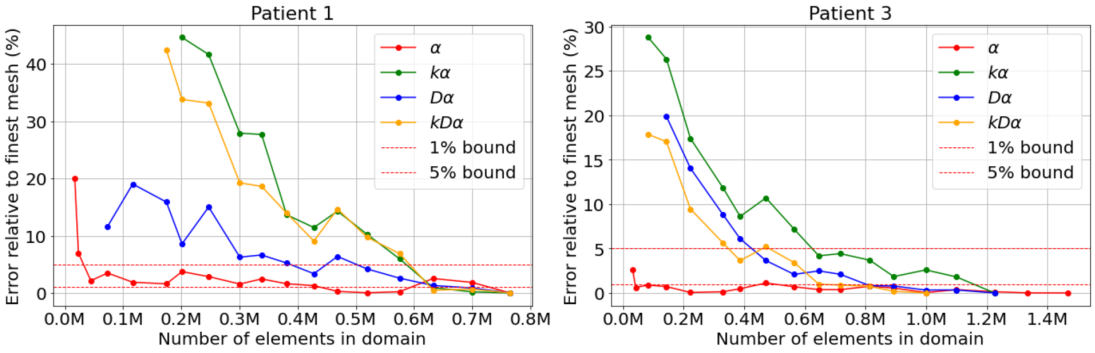
$$\frac{\partial C(\mathbf{x},t)}{\partial t} = D_C \nabla^2 C(\mathbf{x},t) + K^{trans}(\mathbf{x}) \int_0^T C_{drug}^{plasma}(u) \cdot e^{-\frac{K^{trans}(\mathbf{x})}{v_e(\mathbf{x})}(t-u)} du - s, \quad (6.2)$$

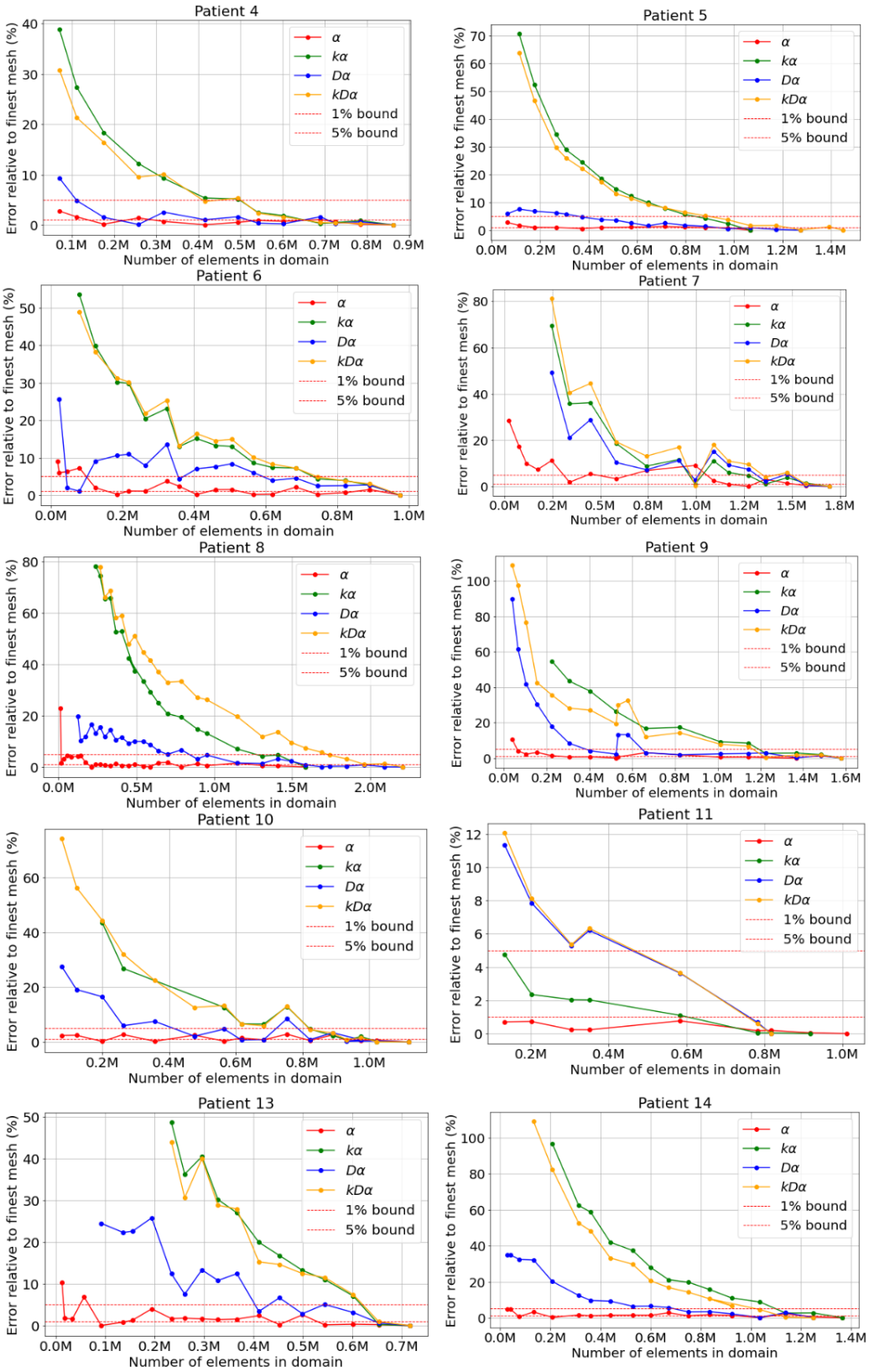
where the second term is a source term as in equation (3.4), with $[0, T]$ the fixed interval over which the plasma drug concentration reduces to 1% of its maximum as described in Subsection 3.5.3. D_C is a diffusion coefficient for the chemotherapy (available in the literature for each drug), and lastly s is a decay term yet to be determined more specifically, possibly informed by the drug half-life.

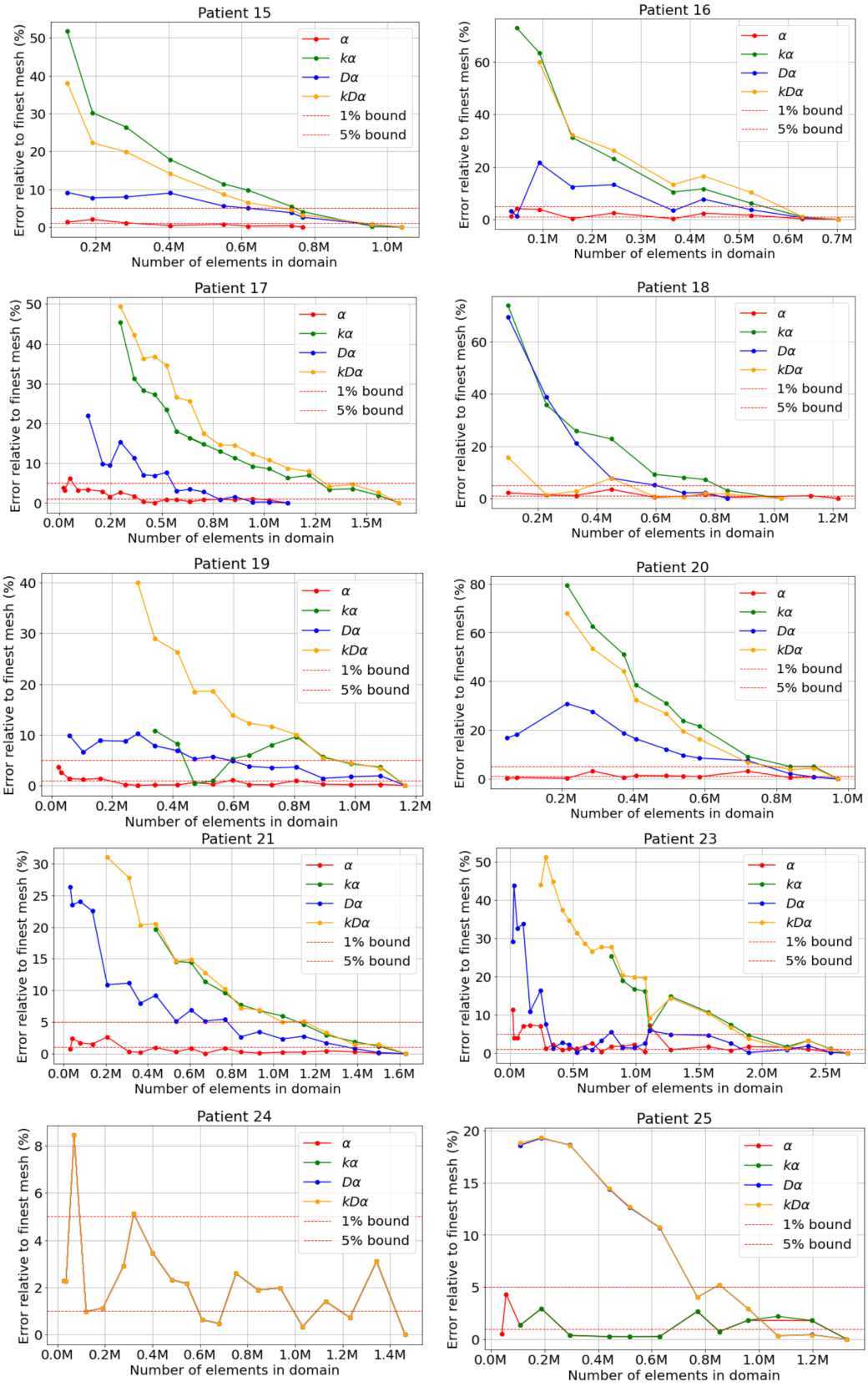
Appendix A

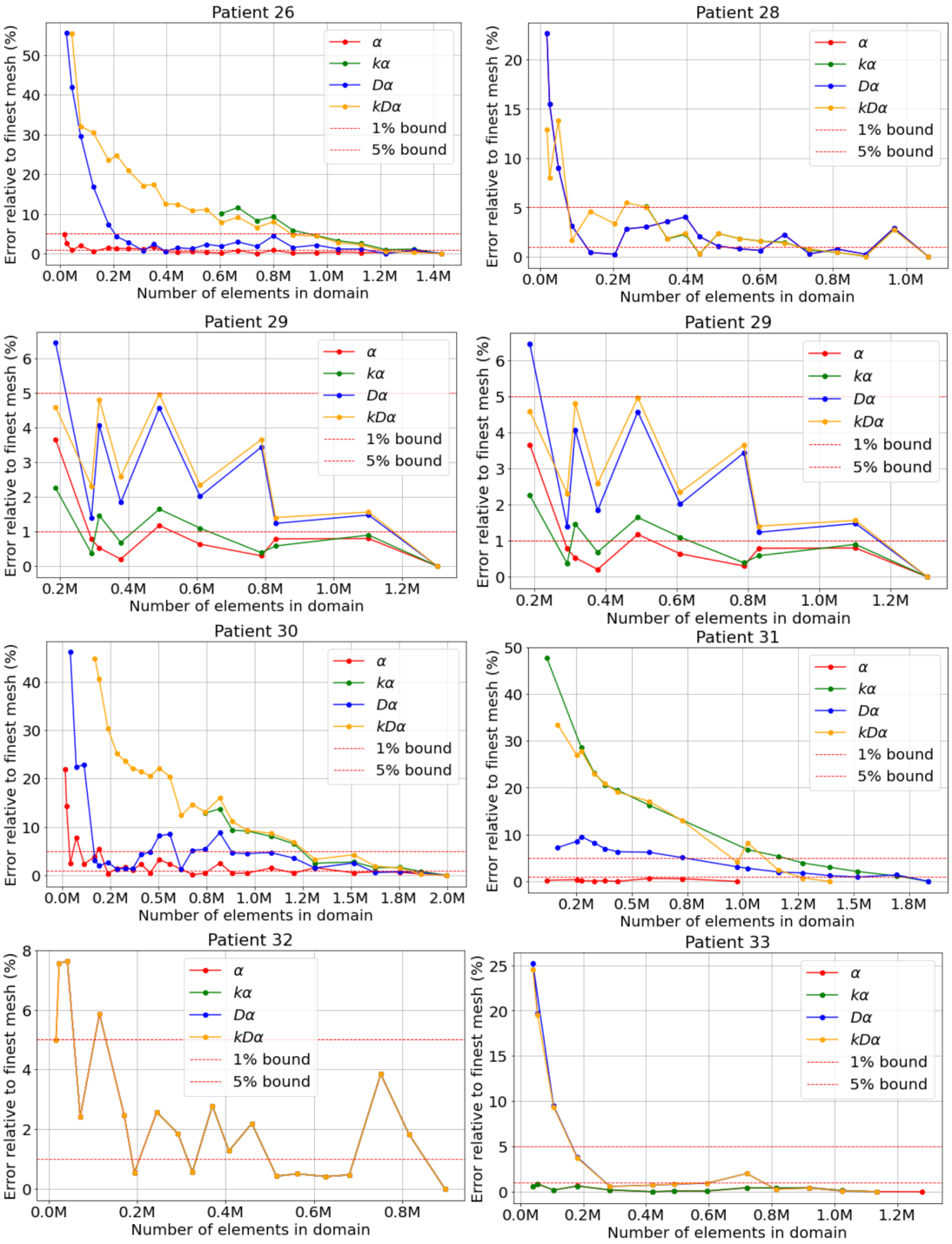
Mesh Independence Study

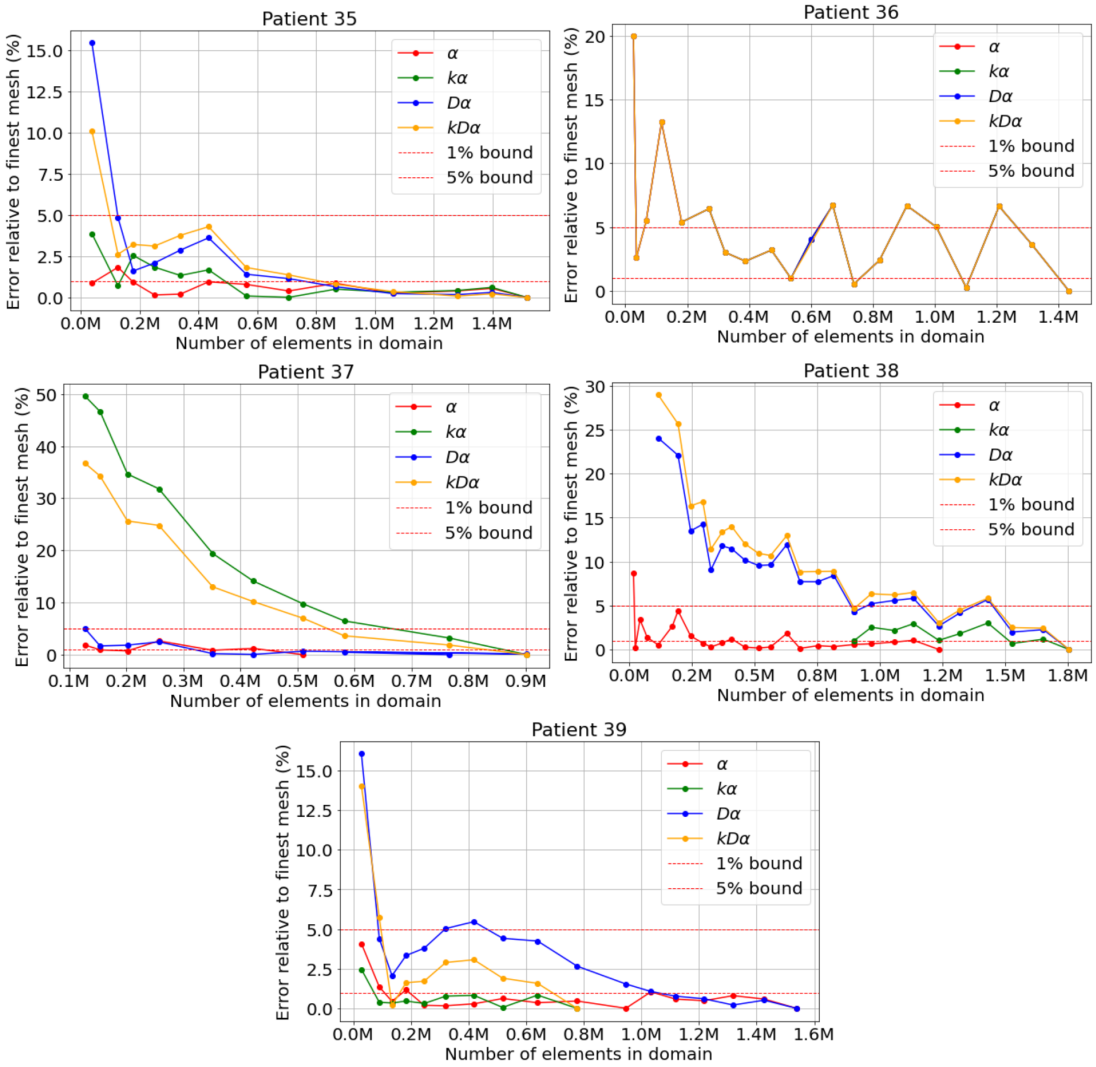
This appendix presents the mesh independence study conducted for each patient and model. The plots show the relative error between simulations run on different meshes and the corresponding simulation on the finest mesh, against the number of elements in each mesh. The coarsest mesh for which the error remained within a specified tolerance relative to the finest mesh was selected: 1% for the α model and 5% for the other models.


Patient 1
Error relative to finest mesh (%)
Number of elements in domain
α
kα
Dα
kDα
1% bound
5% bound
0.0M
0.1M
0.2M
0.3M
0.4M
0.5M
0.6M
0.7M
0.8M
Patient 3
Error relative to finest mesh (%)
Number of elements in domain
0.0M
0.2M
0.4M
0.6M
0.8M
1.0M
1.2M
1.4M


Patient 4
Patient 5
Patient 6
Patient 7
Patient 8
Patient 9
Patient 10
Patient 11
Patient 13
Patient 14
Error relative to finest mesh (%)
Number of elements in domain
α
kα
Dα
kDα
1% bound
5% bound


Patient 15
Patient 16
Patient 17
Patient 18
Patient 19
Patient 20
Patient 21
Patient 23
Patient 24
Patient 25
Error relative to finest mesh (%)
Number of elements in domain
α
kα
Dα
kDα
1% bound
5% bound


Patient 26
Patient 28
Patient 29
Patient 29
Patient 30
Patient 31
Patient 32
Patient 33
Error relative to finest mesh (%)
Number of elements in domain
α
kα
Dα
kDα
1% bound
5% bound


Patient 35
Error relative to finest mesh (%)
Number of elements in domain
α
kα
Dα
kDα
1% bound
5% bound
Patient 36
Patient 37
Patient 38
Patient 39

Appendix B

Patients excluded due to low C_{drug}^{tissue}

Results for the patients whose drug concentration in tissue was measured below 10^{-6} are provided in this Appendix. Patients 3, 25, 32 and 33 had low C_{drug}^{tissue} at Scan 0 and were therfore excluded entirely. Patients 8, 13 and 38 had low C_{drug}^{tissue} at Scan 1, so they were included in the optimisation results but excluded from the prediction. Patient 38 was excluded anyway due to small tumour volume.

ID	$[\times 10^{-5}]\ \alpha$	ϵ_{opt} (%)	t_{opt} (hrs)	ϵ_{vol} (%)	ϵ_{LA} (%)
3	1	11.2	20.4	−6.5	−0.19
25	16.8	74.8	1.4	40.1	−0.39
32	16.8	46.0	0.7	−32.8	0.17
33	16.8	54.3	0.4	−48.3	−0.16
8	4.5	0.6	7.4	333.7	0.40
13	10.5	0.3	8	204.1	0.27
38	13.4	33.2	0.5	105.4	0.59

Table B.1: Results from the α model for each patient, including the calibrated α value, optimisation error (ϵ_{opt}), and optimisation duration in hours (t_{opt}). Simulation errors are reported as the percentage difference between the predicted and actual tumour volumes (ϵ_{vol}) and longest axes (ϵ_{LA}). Negative errors indicate underprediction and positive errors indicate overprediction relative to actual values.

ID	$[\times 10^{-8}]\ k$	$[\times 10^{-5}]\ \alpha$	ϵ_{opt} (%)	t_{opt} (hrs)	ϵ_{vol} (%)	ϵ_{LA} (%)
3	–	–	–	–	–	–
25	8.0	16.8	72.7	5.5	40.5	−0.38
32	8.0	16.8	48.3	0.2	−39.7	−0.07
33	8.0	16.8	55.9	1.8	−27.1	−0.06
8	2.0	4.9	108.9	111.1	344.8	0.55
13	6.7	11.6	2.9	7.5	21.3	0.31
38	100.0	2.3	11.7	23.1	104.8	0.67

Table B.2: Results from the $k\alpha$ model for each patient, including the calibrated values (k and α), optimisation error (ϵ_{opt}) and optimisation duration in hours (t_{opt}). Simulation errors are reported as the percentage difference between the predicted and actual tumour volumes (ϵ_{vol}) and longest axes (ϵ_{LA}). Negative errors indicate underprediction and positive errors indicate overprediction relative to actual values. Missing values (–) indicate a failed optimisation.

ID	$[\times 10^{-14}]$ D	$[\times 10^{-5}]$ α	ϵ_{opt} (%)	t_{opt} (hrs)	ϵ_{vol} (%)	ϵ_{LA} (%)
3	1.0	999.9	18.8	327.4	−67.7	−0.19
25	1.0	1000.0	111.1	46.5	104.5	−0.32
32	1.0	1000.0	74.7	1.8	−29.5	−0.06
33	2.2	1.0	4.9	17.0	70.5	−0.02
8	1.2	650	0.6	75.3	396.4	0.77
13	1.5	400	3.2	1.5	21.6	0.32
38	1.0	1.0	17.8	7.2	102.8	0.65

Table B.3: Results from the $D\alpha$ model for each patient, including the calibrated values (D and α), optimisation error (ϵ_{opt}) and optimisation duration in hours (t_{opt}). Simulation errors are reported as the percentage difference between the predicted and actual tumour volumes (ϵ_{vol}) and longest axes (ϵ_{LA}). Negative errors indicate underprediction and positive errors indicate overprediction relative to actual values.

ID	$[\times 10^{-5}]$ k	$[\times 10^{-14}]$ D	$[\times 10^{-5}]$ α	ϵ_{opt} (%)	t_{opt} (hrs)	ϵ_{vol} (%)	ϵ_{LA} (%)
3	–	–	–	–	–	–	–
25	9.2	1.0	1000.0	111.4	140.0	104.5	−0.34
32	6.6	1.0	1000.0	74.8	5.5	−31.4	−0.05
33	10.0	2.5	1.0	2.5	24.5	77.8	0.06
8	–	–	–	–	–	–	–
13	5.0	1.5	400.0	0.4	1.6	21.3	0.31
38	9.8	1.0	1.0	15.1	23.6	104.8	0.67

Table B.4: Results from the $kD\alpha$ model for each patient, including the calibrated values (k, D and α), optimisation error (ϵ_{opt}) and optimisation duration in hours (t_{opt}). Simulation errors are reported as the percentage difference between the predicted and actual tumour volumes (ϵ_{vol}), longest axes (ϵ_{LA}) and average tumour cell volume fractions ($\epsilon_{V_{TC}}$). Negative errors indicate underprediction and positive errors indicate overprediction relative to actual values.

Appendix C

Lengthy optimisations

In this appendix, we examine the optimisations for which the $k\alpha$, $D\alpha$ or $kD\alpha$ models were unable to converge to a solution in 30 days. We explore potential explanations for these outcomes and provide justification for terminating these optimisations and classifying them as unsuccessful.

C.1 Excessive length of one iteration

The most common cause of an interminable optimisation is that a single iteration of the model is extremely time consuming, which can happen when the mesh contains a very large number of elements. However, the minimum number of elements necessary to ensure a mesh-independent solution was selected. The average duration of one iteration for all patients and for all completed patients (i.e. the patients not discussed in this Appendix) are listed for each model in Table C.1.

	$k\alpha$	$D\alpha$	$kD\alpha$
All patients $\bar{t}_{iter}$ (h)	7.72	4.22	8.72
Completed cases $\bar{t}_{iter}$ (h)	6.23	5.29	3.88

Table C.1: Average duration in hours of one forward pass of the optimisation, for each model. The α model is not included here as it was able to complete optimisations for every patient.

Excessively long iterations were responsible for both failed $k\alpha$ optimisations: one iteration took 29 hours and 47 minutes for patient 9, and 33 hours and 43 minutes for patient 14. Although there is an example of a similarly lengthy iteration that successfully converged – patient 20, with a runtime of 35 hours and 53 minutes per iteration – this optimisation completed in just two iterations, quickly landing on an acceptable error (4.8%). This shows that excessively slow iterations can succeed, but only if the initial starting point is particularly advantageous for the dataset. This was unfortunately not the case for patients 9 and 14; as shown in Figure C.1, their errors continued to increase over 22 days without improvement. Given the computational resource constraints, namely the 48-hour timeout on jobs, it was deemed appropriate to terminate these optimisations and classify them as failed.

For the $D\alpha$ model, excessively long iteration times also explain the failure of patients 3, 18 and 20, who required 13 hours and 17 minutes, 22 hours and 58 minutes and 24 hours and 17 minutes per iteration respectively. As shown in Figure C.2, which plots the error at each iteration over 28 days, it is evident that these optimisations were not converging. An exception to this trend was patient 14, who required 24 hours and 38 minutes per iteration but managed to complete in just two iterations. In contrast, the failed cases of patients 7, 17, 21 and 31 had relatively unremarkable durations for one iteration: 7 hours and 58 minutes, 2 hours and 49 minutes, 3 hours and 31 minutes, and 4 hours and 17 minutes respectively. The average iteration time for all other patients using this model was 4 hours and 13 minutes.

Similarly, the $kD\alpha$ model encountered iteration time issues for patients 3, 8, 9, 14,

18, 20, 23, 26, and 34, with one iteration for each of those cases requiring 11 hours and 8 minutes, 11 hours and 46 minutes, 30 hours and 54 minutes, 34 hours and 56 minutes, 8 hours and 44 minutes, 34 hours and 44 minutes, 16 hours and 13 minutes, 18 hours and 33 minutes, and 11 hours and 29 minutes. By comparison, the average iteration time for all other patients under this model was 3 hours and 53 minutes. While patient 34 eventually finished in just over 23 days, it is clear from the evolution of the errors in Figure C.3 that the other optimisations in this group are not going to improve.

C.2 Tumour behaviour between scans 0 and 1

Another key reason for lengthy optimisation times is the tumour behaviour observed between scans 0 and 1. As discussed in Chapter 5, Section 5.9.2, diffusion-heavy models often struggle to capture significant tumour shrinkage as they are inherently biased towards representing growth. Table 5.7 (in Ch.5, S.5.9.2) lists the percentage change in tumour volume for each patient; highlighting the most significant shrinkages: patients 36, 7, 23, 17, 24, 30, and 1, with reductions of 67.6%, 66.9%, 66.5%, 48.9%, 47.8%, 44.2% and 43.4% respectively. With the exception of patients 24 and 30 – who were successfully calibrated by the $D\alpha$ model in 22 and 42 hours, and by the $kD\alpha$ model in 28 and 75 hours – these patients were unable to complete optimisation for at least one of the diffusion-heavy models.

C.3 Unexplained cases

Table C.2 summarises the optimisation cases that were unfinished for each model, and the theorised reasons why. This leaves patients 21 and 31 unexplained for the $D\alpha$ model and patients 10 and 21 unexplained for the $kD\alpha$ model.

ID	$k\alpha$		$D\alpha$		$kD\alpha$		Big shrink
	Slow opt.	Slow iters.	Slow opt.	Slow iters.	Slow opt.	Slow iters.	
9	✓	✓			✓	✓	
14	✓	✓			✓	✓	
3			✓	✓	✓	✓	
7			✓		✓		✓
17			✓		✓		✓
18			✓	✓	✓	✓	
20			✓	✓	✓	✓	
21			✓		✓		
31			✓				
1					✓		✓
8					✓	✓	
10					✓		
23					✓	✓	✓
26					✓	✓	
34					✓	✓	
36					✓		✓

Table C.2: Summary of lengthy optimisations ("Slow opt." ticked) for each model and the associated reasons. Possible reasons include long iteration times ("Slow iters.") and significant tumour shrinkage between scans 0 and 1 ("Big shrink").

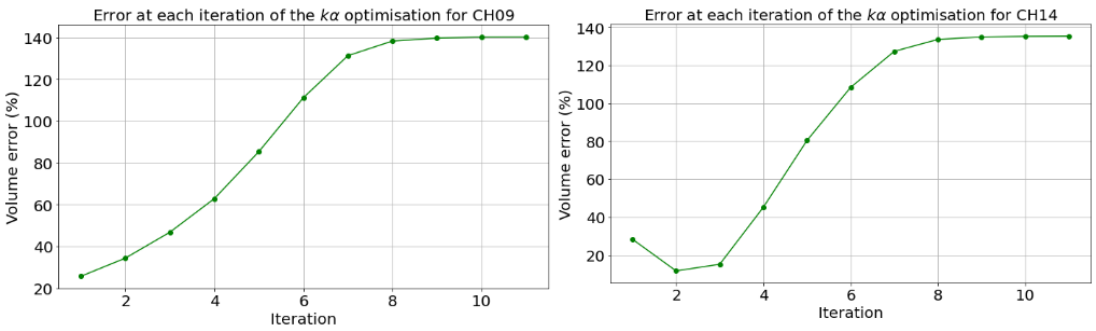


Figure C.1: Volume error evaluated at each iteration of the $k\alpha$ optimisation process over 30 days for the lengthy optimisations, patients 9 and 14.

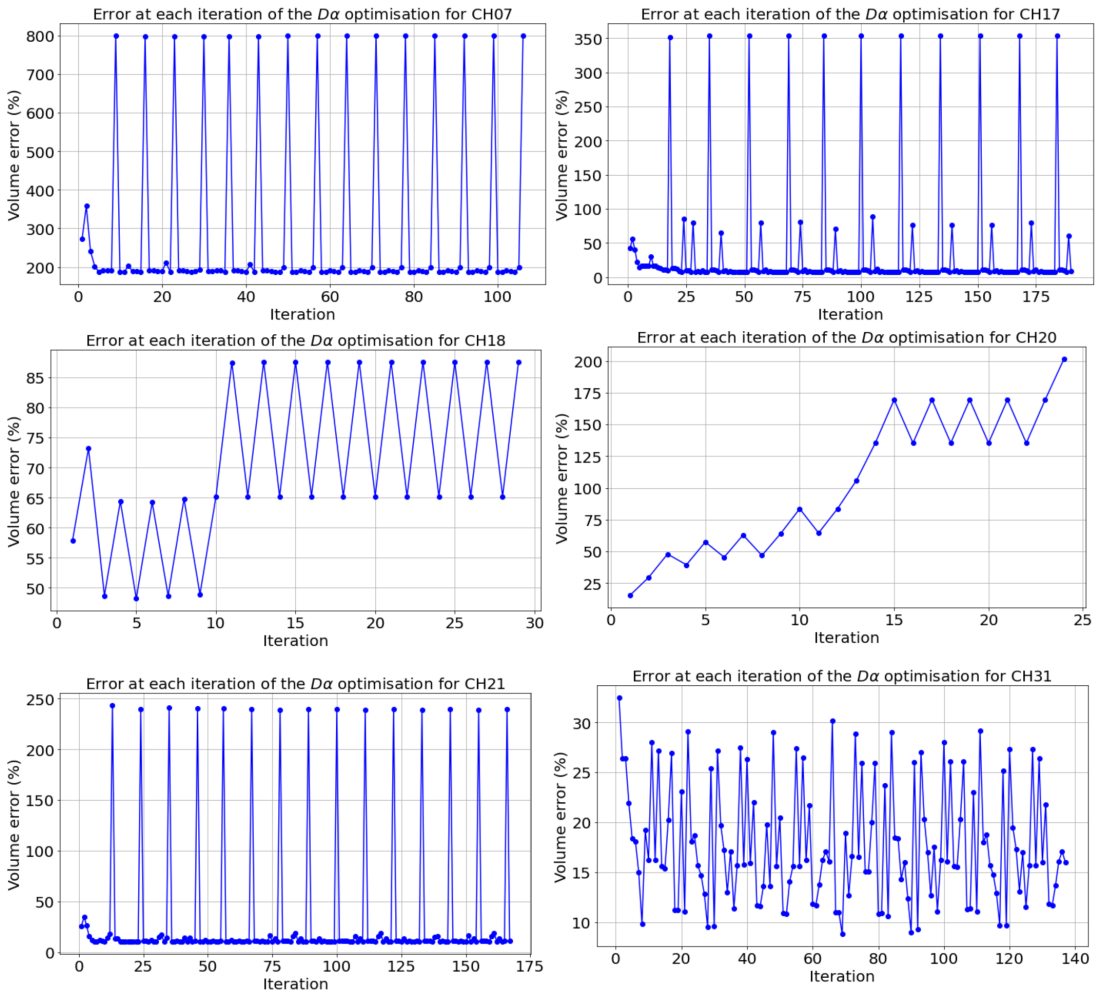


Figure C.2: Volume error evaluated at each iteration of the $D\alpha$ optimisation process over 30 days for the lengthy optimisations: patients 7, 17, 18, 20, 21 and 31.

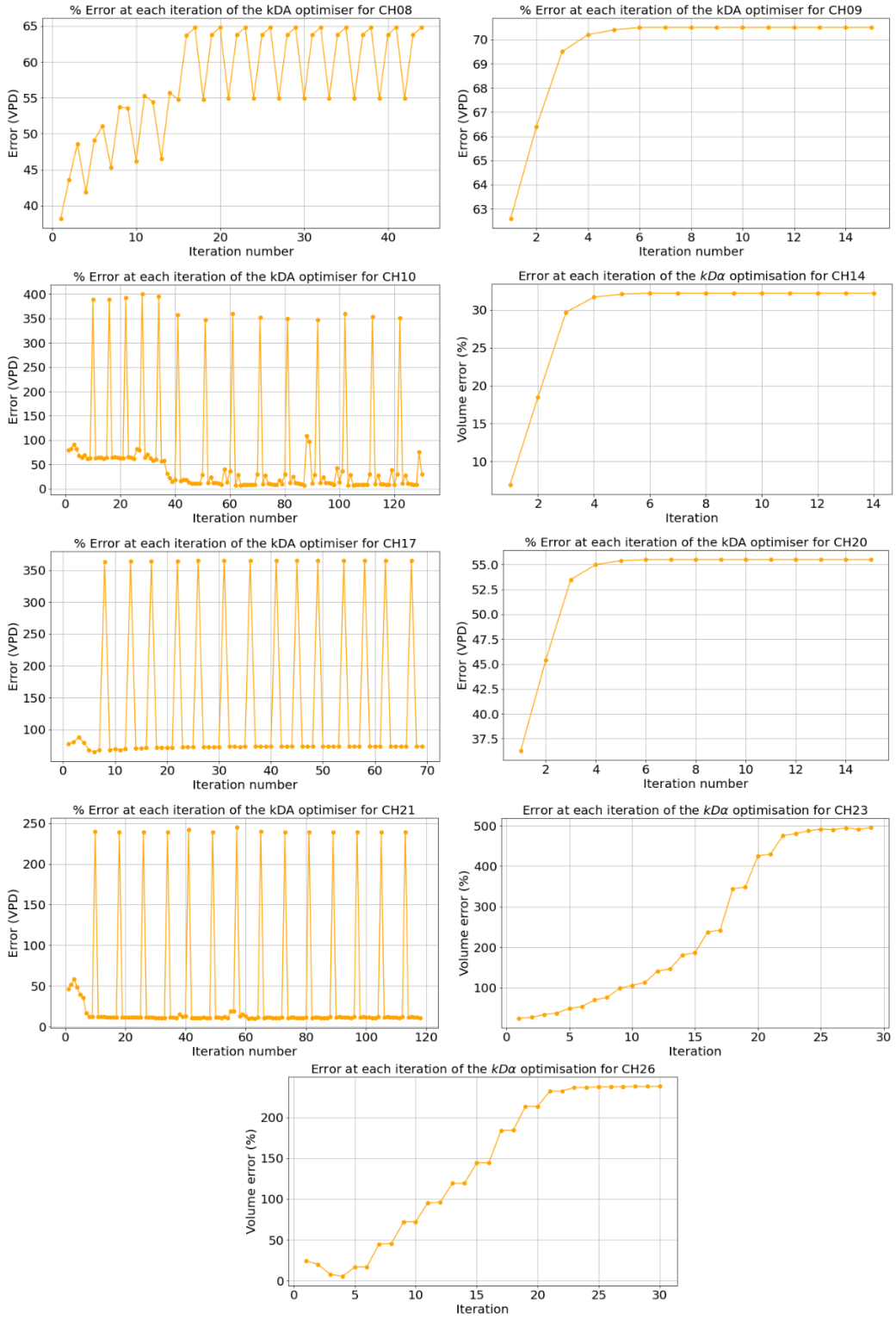


Figure C.3: Volume error evaluated at each iteration of the $kD\alpha$ optimisation process over 30 days for the lengthy optimisations: patients 8, 9, 10, 14, 17, 20, 21, 23 and 26.

Appendix D

Longest axis RECIST

Conventionally, RECIST guidelines are determined based on longest axis measurements. In the main body, we used a volume based RECIST evaluation system. Here we provide the original RECIST evaluations of our model predictions and of the scan 3 data.

Patient	α	$k\alpha$	$D\alpha$	$kD\alpha$	Scan 3	RCB (Scan 6)
1	SD	SD	SD	SD	PR	0
4	SD	SD	SD	SD	SD	3
5	SD	SD	SD	SD	SD	2
6	SD	SD	SD	SD	PR	1
7	CR	CR	–	–	CR	0
9	SD	–	SD	–	SD	3
10	SD	SD	SD	–	SD	1
11	SD	SD	SD	SD	SD	1
14	SD	–	SD	–	SD	2
15	SD	SD	SD	SD	PR	2
16	SD	SD	SD	SD	SD	3
17	SD	SD	SD	–	PR	1
18	PR	PR	–	SD	PD	2
19	SD	SD	SD	SD	PR	1
20	SD	SD	–	–	SD	3
21	SD	SD	–	–	SD	2
23	PR	PR	PR	–	CR	0
24	SD	SD	SD	SD	SD	1
26	SD	SD	SD	–	SD	3
29	PR	SD	SD	SD	PR	2
30	SD	SD	SD	SD	PR	0
31	SD	SD		SD	PD	1
34	SD	SD	SD	SD	SD	2
35	SD	SD	SD	SD	SD	3
36	PR	PR	PR	SD	SD	1
37	SD	SD	SD	SD	PR	2
39	SD	SD	PD	SD	SD	2

Table D.1: RECIST LA evaluation for each model prediction at scan 3 (α, $k\alpha$, $D\alpha$, $kD\alpha$), for the data at scan 3 (Scan 3), and residual cancer burden determined after the completion of NACT (RCB). Missing values indicate failed optimisations.

Appendix E

Scan 3 measured and predicted tumour characteristics

Tables E.1-E.4 present the tumour characteristics predicted by the α, $D\alpha$, $k\alpha$ and $kD\alpha$ models, compared to measurements from Scan 3: volume, longest axis and tumour cell volume fraction.

ID	vol_P [$\times 10^2$ mm^3]	vol_M [$\times 10^2$ mm^3]	LA_P [mm]	LA_M [mm]	V_{TC_P}	V_{TC_M}
1	8.3	9.5	28.4	24.2	0.20	0.47
4	23.1	22.9	30.2	30.5	0.25	0.43
5	34.1	20.3	37.5	25.2	0.32	0.37
6	5.4	5.3	18.7	15.3	0.13	0.34
7	0.0	0.0	0.0	0.0	0.00	0.00
9	49.0	119.3	59.0	66.1	0.32	0.46
10	6.0	8.0	15.0	19.2	0.35	0.27
11	51.5	23.1	31.6	25.7	0.53	0.54
14	287.1	273.9	65.3	68.1	0.40	0.46
15	31.0	11.2	41.0	19.6	0.27	0.44
16	16.3	23.0	32.0	25.4	0.35	0.51
17	1.6	1.3	21.3	12.2	0.23	0.38
18	21.3	311.4	10.2	94.8	0.26	0.36
19	26.1	29.4	31.8	23.6	0.26	0.39
20	332.4	243.8	83.7	67.9	0.38	0.43
21	9.2	11.6	22.7	18.5	0.25	0.38
23	22.7	0.0	26.9	0.0	0.28	0.00
24	1.9	5.2	13.2	16.4	0.46	0.35
26	216.6	150.1	63.6	58.1	0.29	0.51
29	10.2	27.5	44.9	31.3	0.37	0.45
30	3.5	3.4	20.4	13	0.20	0.37
31	51.2	43.4	34.5	40.2	0.29	0.34
34	95.1	112.3	37.5	33.9	0.33	0.42
35	121.4	72.7	53.3	39.4	0.21	0.38
36	0.2	19.4	4.27	23.4	0.17	0.29
37	75.1	56.2	47.9	29.9	0.26	0.46
39	36.6	52.2	43.0	39.1	0.23	0.41
Median	23.1	23.0	31.8	25.4	0.27	0.39
Q. range	[6.0, 51.5]	[8.0, 72.7]	[20.4, 44.9]	[18.5, 39.4]	[0.23, 0.35]	[0.35, 0.46]

Table E.1: Predicted ($_P$) and measured ($_M$) tumour characteristics at Scan 3 using the α model. Results include tumour volume (vol), longest axis (LA) and tumour cell volume fraction (V_{TC}).

ID	vol_P [$\times 10^2$ mm^3]	vol_M [$\times 10^2$ mm^3]	LA_P [mm]	LA_M [mm]	V_{TC_P}	V_{TC_M}
1	10.8	9.1	28.4	24.0	0.20	0.47
4	20.2	22.5	31.5	30.4	0.23	0.42
5	57.5	21.7	38.6	27.4	0.22	0.37
6	5.5	5.3	18.7	15.0	0.12	0.33
7	0.0	0.0	0.0	0.0	0.00	0.00
10	5.7	7.9	19.0	20.4	0.36	0.28
11	45.0	21.4	31.6	26.2	0.46	0.54
15	29.8	11.6	42.5	22.7	0.22	0.44
16	18.7	22.4	32.6	24.8	0.29	0.51
17	1.6	1.3	22.2	12.3	0.23	0.37
18	35.4	330.2	38.1	96.7	0.19	0.36
19	30.5	29.1	32.1	24.0	0.36	0.39
20	437.1	268.1	92.8	81.0	0.20	0.43
21	8.0	12.5	24.4	19.3	0.25	0.38
23	36.8	0.0	26.6	0.0	0.48	0.00
24	1.8	4.9	12.2	15.7	0.46	0.34
26	230.5	154.3	63.6	57.7	0.26	0.51
29	68.3	27.5	61.0	31.3	0.16	0.45
30	4.1	3.5	20.7	13.1	0.19	0.37
31	46.3	40.2	34.3	38.7	0.30	0.34
34	63.6	104.0	38.8	33.7	0.19	0.41
35	104.4	62.7	49.7	38.8	0.22	0.38
36	0.3	19.4	4.27	23.4	0.17	0.29
37	87.8	55.3	47.9	29.9	0.23	0.46
39	66.7	50.9	44.0	42.6	0.17	0.41
Median	30.5	21.7	32.1	24.8	0.22	0.38
Q range	[5.6, 65.1]	[6.6, 53.1]	[21.5, 43.3]	[17.5, 36.2]	[0.19, 0.29]	[0.34, 0.45]

Table E.2: Predicted ($_P$) and measured ($_M$) tumour characteristics at Scan 3 using the $k\alpha$ model. Results include tumour volume (vol), longest axis (LA) and tumour cell volume fraction (V_{TC}).

ID	vol_P [$\times 10^2$ mm^3]	vol_M [$\times 10^2$ mm^3]	LA_P [mm]	LA_M [mm]	V_{TC_P}	V_{TC_M}
1	10.6	8.6	29	23.7	0.17	0.47
4	37.6	23.3	33.1	29.6	0.02	0.43
5	74.5	20.0	42.4	25.3	0.03	0.37
6	6.2	5.3	20.6	15.7	0.00	0.34
9	86.9	120.1	59.5	64.8	0.08	0.46
10	6.0	7.9	16.1	20.4	0.20	0.28
11	42.1	21.4	33.8	26.2	0.18	0.54
14	273.9	268.8	77.4	69.3	0.03	0.46
15	41.1	11.9	42.4	22.3	0.03	0.45
16	28.1	22.4	34.5	25	0.05	0.51
19	19.8	27.3	32.4	23	0.03	0.39
23	37.9	0.0	26.4	0.0	0.23	0.00
24	5.0	5.3	14.6	16.3	0.22	0.35
26	259.6	148.2	62.5	53.2	0.01	0.51
29	88.1	27.5	64.3	31.3	0.06	0.45
30	4.8	3.5	19.9	12.9	0.00	0.37
34	63.2	101.9	37.1	31.1	0.05	0.41
35	76.1	62.7	51.6	38.8	0.24	0.38
36	8.6	23.2	12.1	23.3	0.07	0.28
37	135.3	56.2	49.9	29.3	0.04	0.46
39	76.9	46.0	48.3	42.7	0.04	0.42
Median	41.0	23.2	34.5	25.3	0.05	0.42
Q range	[9.6, 81.9]	[8.2, 59.4]	[23.5, 50.8]	[21.4, 35.1]	[0.03, 0.17]	[0.36, 0.46]

Table E.3: Predicted ($_P$) and measured ($_M$) tumour characteristics at Scan 3 using the $D\alpha$ model. Results include tumour volume (vol), longest axis (LA) and tumour cell volume fraction (V_{TC}).

ID	vol_P [$\times 10^2$ mm^3]	vol_M [$\times 10^2$ mm^3]	LA_P [mm]	LA_M [mm]	V_{TC_P}	V_{TC_M}
4	35.9	22.45	32.9	30.4	0.02	0.42
5	101.8	21.68	41.6	27.4	0.07	0.37
6	10.9	5.27	21.7	15	0.02	0.33
11	40.4	21.37	33.9	26.2	0.19	0.54
15	41.3	11.6	45.7	22.7	0.03	0.44
16	31.8	22.4	34.3	24.8	0.05	0.51
19	31.9	29.1	34.4	24	0.06	0.39
24	5.1	4.9	14.47	15.7	0.19	0.34
29	100.2	27.5	64.5	31.3	0.08	0.45
30	4.6	3.5	19.6	13.1		0.37
31	46.6	40.2	36.5	38.7	0.04	0.34
35	105.9	62.7	47.8	38.8	0.17	0.38
37	131.9	55.3	51.1	29.9	0.07	0.46
39	94.1	51.0	43.2	42.6	0.07	0.41
Median	40.4	22.4	34.4	26.2	0.06	0.41
Q range	[10.9, 100.2]	[9.1, 40.2]	[21.7, 45.7]	[22.7, 31.3]	[0.04, 0.10]	[0.37, 0.46]

Table E.4: Predicted ($_P$) and measured ($_M$) tumour characteristics at Scan 3 using the $kD\alpha$ model. Results include tumour volume (vol), longest axis (LA) and tumour cell volume fraction (V_{TC}).

Appendix F

Optimisation error against prediction error

Figure F.1 shows the optimisation error against the prediction error for each model, given as the percentage difference between the numerical and data tumour volumes. The colours indicate the RECIST status at Scan 3.

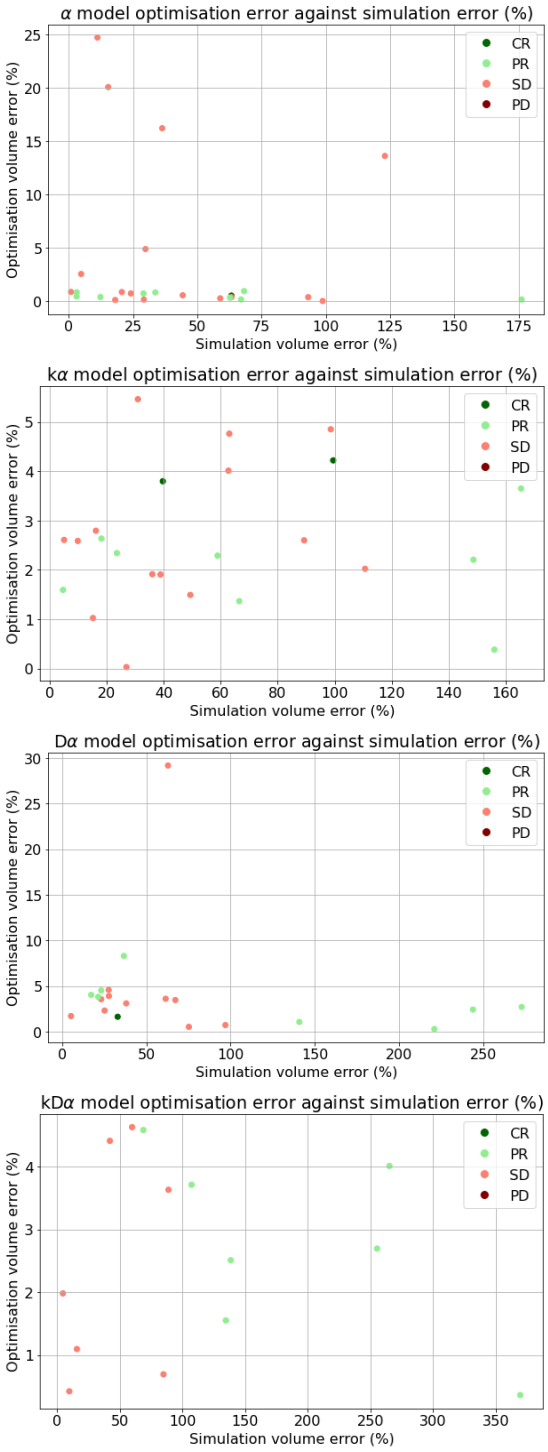


Figure F.1: Optimisation error against prediction error for each model. Both errors are given as the percentage difference in volume.

Bibliography

[1] Fangyuan Zhao, Eric Polley, Julian McClellan, Frederick Howard, Olufunmilayo I. Olopade, and Dezheng Huo. Predicting pathologic complete response to neoadjuvant chemotherapy in breast cancer using a machine learning approach. *Breast cancer research : BCR*, 26(1):148, 12 2024. ISSN 1465542X. doi: 10.1186/s13058-024-01905-7.

[2] Angela M Jarrett, Anum S Kazerouni, Chengyue Wu, John Virostko, Anna G Sorace, Julie C DiCarlo, David A Hormuth, David A Ekrut, Debra Patt, Boone Goodgame, Sarah Avery, and Thomas E Yankeelov. Quantitative magnetic resonance imaging and tumor forecasting of breast cancer patients in the community setting. *Nature Protocols*, 2021. doi: 10.1038/s41596-021-00617-y. URL https://doi.org/10.1038/s41596-021-00617-y.

[3] N C Atuegwu, D C Colvin, M E Loveless, L Xu, J C Gore, and T E Yankeelov. Incorporation of diffusion-weighted magnetic resonance imaging data into a simple mathematical model of tumor growth. *Phys Med Biol*, 57(1):225–240, 1 2012. ISSN 1361-6560 (Electronic); 0031-9155 (Print); 0031-9155 (Linking). doi: 10.1088/0031-9155/57/1/225.

[4] Angela M Jarrett, David A Hormuth, Stephanie L Barnes, Xinzeng Feng, Wei Huang, and Thomas E Yankeelov. Incorporating drug delivery into an imaging-driven, mechanics-coupled reaction diffusion model for predicting the response of breast cancer to neoadjuvant chemotherapy: theory and preliminary clinical

results. *Phys Med Biol*, 63(10):105015, 5 2018. ISSN 1361-6560 (Electronic); 0031-9155 (Print); 0031-9155 (Linking). doi: 10.1088/1361-6560/aac040.

[5] Angela M. Jarrett, David A. Hormuth, Chengyue Wu, Anum S. Kazerouni, David A. Ekrut, John Virostko, Anna G. Sorace, Julie C. DiCarlo, Jeanne Kowalski, Debra Patt, Boone Goodgame, Sarah Avery, and Thomas E. Yankeelov. Evaluating patient-specific neoadjuvant regimens for breast cancer via a mathematical model constrained by quantitative magnetic resonance imaging data. *Neoplasia*, 22(12):820–830, 12 2020. ISSN 14765586. doi: 10.1016/j.neo.2020.10.011. URL https://linkinghub.elsevier.com/retrieve/pii/S1476558620301664.

[6] Jared A Weis, Michael I Miga, Lori R Arlinghaus, Xia Li, A Bapsi Chakravarthy, Vandana Abramson, Jaime Farley, and Thomas E Yankeelov. A mechanically coupled reaction-diffusion model for predicting the response of breast tumors to neoadjuvant chemotherapy. *Phys Med Biol*, 58(17):5851–5866, 9 2013. ISSN 1361-6560 (Electronic); 0031-9155 (Print); 0031-9155 (Linking). doi: 10.1088/0031-9155/58/17/5851.

[7] Hyuna Sung, Jacques Ferlay, Rebecca L. Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal, and Freddie Bray. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, 71(3):209–249, 5 2021. ISSN 0007-9235. doi: 10.3322/caac.21660.

[8] Dimitrios Zardavas and Martine Piccart. Neoadjuvant therapy for breast cancer. *Annual Review of Medicine*, 66:31–48, 1 2015. ISSN 1545326X. doi: 10.1146/annurev-med-051413-024741.

[9] John Virostko, Anna G Sorace, Kalina P Slavkova, Anum S Kazerouni, Angela M Jarrett, Julie C DiCarlo, Stefanie Woodard, Sarah Avery, Boone Goodgame, Debra Patt, and Thomas E Yankeelov. Quantitative multiparametric MRI predicts response to neoadjuvant therapy in the community setting. *Breast*

Cancer Research, 23(1):110, 2021. doi: 10.1186/s13058-021-01489-6. URL https://doi.org/10.1186/s13058-021-01489-6.

[10] Anna G Sorace, Sara Harvey, Anum Syed, and Thomas E Yankeelov. Imaging Considerations and Interprofessional Opportunities in the Care of Breast Cancer Patients in the Neoadjuvant Setting. *Semin Oncol Nurs*, 33(4):425–439, 11 2017. ISSN 1878-3449 (Electronic); 0749-2081 (Print); 0749-2081 (Linking). doi: 10.1016/j.soncn.2017.08.008.

[11] Beatriu Reig, Laura Heacock, Alana Lewin, Nariya Cho, and Linda Moy. Role of MRI to Assess Response to Neoadjuvant Therapy for Breast Cancer. *Journal of Magnetic Resonance Imaging*, 52(6), 12 2020. ISSN 15222586. doi: 10.1002/jmri.27145.

[12] Richard G Abramson, Lori R Arlinghaus, Jared A Weis, Xia Li, Adrienne N Dula, Eduard Y Chekmenev, Seth A Smith, Michael I Miga, Vandana G Abramson, and Thomas E Yankeelov. Current and emerging quantitative magnetic resonance imaging methods for assessing and predicting the response of breast cancer to neoadjuvant therapy. *Breast Cancer (Dove Med Press)*, 2012(4):139–154, 10 2012. ISSN 1179-1314 (Print); 1179-1314 (Electronic); 1179-1314 (Linking). doi: 10.2147/BCTT.S35882.

[13] Jared A. Weis, Michael I. Miga, Lori R. Arlinghaus, Xia Li, Vandana Abramson, A. Bapsi Chakravarthy, Praveen Pendyala, and Thomas E. Yankeelov. Predicting the response of breast cancer to neoadjuvant therapy using a mechanically coupled reaction-diffusion model. *Cancer Research*, 75(22):4697–4707, 11 2015. ISSN 15387445. doi: 10.1158/0008-5472.CAN-14-2945.

[14] Jared A Weis, Michael I Miga, and Thomas E Yankeelov. Three-dimensional image-based mechanical modeling for predicting the response of breast cancer to neoadjuvant therapy. *Computer Methods in Applied Mechanics and Engineering*, 314:494–512, 2017. ISSN 0045-7825. doi: https://doi.org/10.1016/

j.cma.2016.08.024. URL https://www.sciencedirect.com/science/article/pii/S0045782516310167.

[15] David A Hormuth II, Angela M Jarrett, Ernesto A B F Lima, Matthew T Mckenna, David T Fuentes, and Thomas E Yankeelov. Mechanism-Based Modeling of Tumor Growth and Treatment Response Constrained by Multiparametric Imaging Data. *JCO Clinical Cancer Informatics*, 3, 2019. doi: 10.1200/CCI.18. URL https://doi.

[16] David A Hormuth II, Anna G Sorace, John Virostko, Richard G Abramson, Zaver M Bhujwalla, Pedro Enriquez-Navas, Robert Gillies, John D Hazle, Ralph P Mason, C Chad Quarles, Jared A Weis, Jennifer G Whisenant, Junzhong Xu, and Thomas E Yankeelov. Translating preclinical MRI methods to clinical oncology. *Journal of Magnetic Resonance Imaging*, 50(5):1377–1392, 2019. doi: https://doi.org/10.1002/jmri.26731. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jmri.26731.

[17] Ratain MJ and Eckhart SG. Phase II studies of modern drugs directed against new targets: if you are fazed, too, then resist RECIST. *J Clin Oncol*, 22(22): 4442–5, 2004.

[18] Rabiya S Tuma. Sometimes Size Doesn't Matter: Reevaluating RECIST and Tumor Response Rate Endpoints. *JNCI: Journal of the National Cancer Institute*, 98(18):1272–1274, 9 2006. ISSN 0027-8874. doi: 10.1093/jnci/djj403. URL https://doi.org/10.1093/jnci/djj403.

[19] Renee Brady and Heiko Enderling. Mathematical Models of Cancer: When to Predict Novel Therapies, and When Not to. *Bulletin of Mathematical Biology*, 81(10):3722–3731, 10 2019. ISSN 15229602. doi: 10.1007/s11538-019-00640-x.

[20] Angela M. Jarrett, Ernesto A.B.F. Lima, David A. Hormuth, Matthew T. McKenna, Xinzeng Feng, David A. Ekrut, Anna Claudia M. Resende, Amy Brock,

and Thomas E. Yankeelov. Mathematical models of tumor cell proliferation: A review of the literature, 12 2018. ISSN 17448328.

[21] M. Kuznetsov, J. Clairambault, and V. Volpert. Improving cancer treatments via dynamical biophysical models, 12 2021. ISSN 15710645.

[22] Yuri Mansuri, Mark Kimura, Jose Lobo, and Thomas Deisboeck. Emerging Patterns in Tumor Systems: Simulating the Dynamics of Multicellular Clusters with an Agent-based Spatial Agglomeration Model. *Journal of Theoretical Biology*, 219 (3):343–370, 2002. ISSN 0022-5193. doi: https://doi.org/10.1006/jtbi.2002.3131. URL https://www.sciencedirect.com/science/article/pii/S0022519302931310.

[23] Thomas S. Deisboeck, Zhihui Wang, Paul MacKlin, and Vittorio Cristini. Multiscale cancer modeling. *Annual Review of Biomedical Engineering*, 13:127–155, 8 2011. ISSN 15239829. doi: 10.1146/annurev-bioeng-071910-124729.

[24] J. S. Lowengrub, H. B. Frieboes, F. Jin, Y. L. Chuang, X. Li, P. MacKlin, S. M. Wise, and V. Cristini. Nonlinear modelling of cancer: Bridging the gap between cells and tumours. *Nonlinearity*, 23(1), 2010. ISSN 09517715. doi: 10.1088/0951-7715/23/1/R01.

[25] Mohammad Mamunur Rahman, Yusheng Feng, Thomas E. Yankeelov, and J. Tinsley Oden. A fully coupled space–time multiscale modeling framework for predicting tumor growth. *Computer Methods in Applied Mechanics and Engineering*, 320:261–286, 6 2017. ISSN 00457825. doi: 10.1016/j.cma.2017.03.021.

[26] Chengyue Wu, Guillermo Lorenzo, David A. Hormuth, Ernesto A. B. F. Lima, Kalina P. Slavkova, Julie C. DiCarlo, John Virostko, Caleb M. Phillips, Debra Patt, Caroline Chung, and Thomas E. Yankeelov. Integrating mechanism-based modeling with biomedical imaging to build practical digital twins for clinical oncology. *Biophysics Reviews*, 3(2):021304, 6 2022. doi: 10.1063/5.0086789.

[27] H Skipper and F Schabel. Quantitative and cytokinetic studies in experimental

tumor systems. In J F Holland and E Frei III, editors, *Cancer Medicine*, pages 636–648. Lea and Febiger, Philadelphia, 2 edition, 1982.

[28] Larry Norton. A Gompertzian Model of Human Breast Cancer Growth. *Cancer Research*, 48(24 Part 1):7067, 12 1988. URL http://cancerres.aacrjournals.org/content/48/24_Part_1/7067.abstract.

[29] W V Mayneord. On a law of growth of Jensen's rat sarcoma. *Amer J Cancer*, 1932.

[30] A K Laird. Dynamics of Tumor Growth. *British journal of cancer*, 13(3):490–502, 9 1964. ISSN 0007-0920. doi: 10.1038/bjc.1964.55. URL https://pubmed.ncbi.nlm.nih.gov/14219541https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2071101/.

[31] Nkiruka C Atuegwu, Lori R Arlinghaus, Xia Li, A Bapsi Chakravarthy, Vandana G Abramson, Melinda E Sanders, and Thomas E Yankeelov. Parameterizing the Logistic Model of Tumor Growth by DW-MRI and DCE-MRI Data to Predict Treatment Response and Changes in Breast Cancer Cellularity during Neoadjuvant Chemotherapy. *Transl Oncol*, 6(3):256–264, 6 2013. ISSN 1936-5233 (Print); 1936-5233 (Electronic); 1936-5233 (Linking). doi: 10.1593/tlo.13130.

[32] Alex Viguerie, Malu Grave, Gabriel F. Barros, Guillermo Lorenzo, Alessandro Reali, and Alvaro L. G. A. Coutinho. Data-driven simulation of Fisher-Kolmogorov tumor growth models using Dynamic Mode Decomposition. 2 2022. URL http://arxiv.org/abs/2202.13860.

[33] David A Hormuth II, Stephanie L Eldridge, Jared A Weis, Michael I Miga, and Thomas E Yankeelov. Mechanically Coupled Reaction-Diffusion Model to Predict Glioma Growth: Methodological Details. *Methods in molecular biology (Clifton, N.J.)*, 1711:225–241, 2018. ISSN 1940-6029. doi: 10.1007/978-1-4939-7493-1{_}11. URL https://pubmed.ncbi.nlm.nih.gov/29344892https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6530463/.

[34] Ken C.L. Wong, Ronald M. Summers, Electron Kebebew, and Jianhua Yao. Pancreatic Tumor Growth Prediction with Elastic-Growth Decomposition, Image-Derived Motion, and FDM-FEM Coupling. *IEEE Transactions on Medical Imaging*, 36(1):111–123, 1 2017. ISSN 1558254X. doi: 10.1109/TMI.2016.2597313.

[35] K R Swanson, E C Alvord, and J D Murray. Virtual brain tumours (gliomas) enhance the reality of medical imaging and highlight inadequacies of current therapy. *British Journal of Cancer*, 86:14–18, 2002. doi: 10.1038/sj/bjc/6600021.

[36] Robert Gatenby and E Gawlinski. Reaction-diffusion model of cancer invasion. *Cancer research*, 56:5745–5753, 1 1997.

[37] A Giese, L Kluwe, B Laube, H Meissner, M E Berens, and M Westphal. Migration of human glioma cells on myelin. *Neurosurgery*, 38(4):755–764, 1996. ISSN 0148-396X. doi: 10.1227/00006123-199604000-00026. URL http://europepmc.org/abstract/MED/8692396https://doi.org/10.1097/00006123-199604000-00026https://doi.org/10.1227/00006123-199604000-00026.

[38] Gabriel Helmlinger, Paolo A Netti, Hera C Lichtenbeld, Robert J Melder, and Rakesh K Jain. Solid stress inhibits the growth of multicellular tumor spheroids. *Nature Biotechnology*, 15:778–783, 1997. URL http://www.nature.com/naturebiotechnology.

[39] D. Ambrosi and F. Mollica. The role of stress in the growth of a multicell spheroid. *Journal of Mathematical Biology*, 48(5):477–499, 5 2004. ISSN 03036812. doi: 10.1007/s00285-003-0238-2.

[40] Tiina Roose, Paolo A. Netti, Lance L. Munn, Yves Boucher, and Rakesh K. Jain. Solid stress generated by spheroid growth estimated using a linear poroelasticity model. *Microvascular Research*, 66(3):204–212, 2003. ISSN 00262862. doi: 10.1016/S0026-2862(03)00057-8.

[41] Ishita Garg and Michael I Miga. Preliminary investigation of the inhibitory

effects of mechanical stress in tumor growth. San Diego, 2 2008. Proceedings of the Society of Photo-Optical Instrumentation Engineers.

[42] Aristotle Bamias and Meletios A. Dimopoulos. Angiogenesis in human cancer: Implications in cancer therapy, 2003. ISSN 09536205.

[43] Robert S. Kerbel. Inhibition of tumor angiogenesis as a strategy to circumvent acquired resistance to anti-cancer therapeutic agents. *BioEssays*, 13(1):31–36, 1 1991. ISSN 0265-9247. doi: 10.1002/bies.950130106.

[44] Jamie Berger, Ashley Horton, Grand Joldes, Adam Wittek, and Karol Miller. Coupling Finite Element and Mesh-free Methods for Modelling Brain Deformation in Response to Tumour Growth. Technical report, 2008. URL www.mech.uwa.edu.au/ISML.

[45] M. Scianna, C. G. Bell, and L. Preziosi. A review of mathematical models for the formation of vascular networks, 9 2013. ISSN 00225193.

[46] Guillermo Vilanova, Ignasi Colominas, and Hector Gomez. Computational Modeling of Tumor-Induced Angiogenesis. *Archives of Computational Methods in Engineering*, 24(4):1071–1102, 11 2017. ISSN 18861784. doi: 10.1007/s11831-016-9199-7.

[47] Guillermo Lorenzo, David A. Hormuth, Angela M. Jarrett, Ernesto A. B. F. Lima, Shashank Subramanian, George Biros, J. Tinsley Oden, Thomas J. R. Hughes, and Thomas E. Yankeelov. Quantitative in vivo imaging to enable tumor forecasting and treatment optimization. 2 2021. URL http://arxiv.org/abs/2102.12602.

[48] David A. Hormuth, Caleb M. Phillips, Chengyue Wu, Ernesto A.B.F. Lima, Guillermo Lorenzo, Prashant K. Jha, Angela M. Jarrett, J. Tinsley Oden, and Thomas E. Yankeelov. Biologically-based mathematical modeling of tumor vasculature and angiogenesis via time-resolved imaging data, 6 2021. ISSN 20726694.

[49] David A. Hormuth, Angela M. Jarrett, Xinzeng Feng, and Thomas E. Yankeelov. Calibrating a Predictive Model of Tumor Growth and Angiogenesis with Quantitative MRI. *Annals of Biomedical Engineering*, 47(7):1539–1551, 7 2019. ISSN 15739686. doi: 10.1007/s10439-019-02262-9.

[50] Thais Roque, Laurent Risser, Veerle Kersemans, Sean Smart, Danny Allen, Paul Kinchesh, Stuart Gilchrist, Ana L. Gomes, Julia A. Schnabel, and Michael A. Chappell. A DCE-MRI Driven 3-D Reaction-Diffusion Model of Solid Tumor Growth. *IEEE Transactions on Medical Imaging*, 37(3):724–732, 3 2018. ISSN 1558254X. doi: 10.1109/TMI.2017.2779811.

[51] Chengyue Wu, David A Hormuth, Todd A Oliver, Federico Pineda, Guillermo Lorenzo, Gregory S Karczmar, Robert D Moser, and Thomas E Yankeelov. Patient-Specific Characterization of Breast Cancer Hemodynamics Using Image-Guided Computational Fluid Dynamics. *IEEE Trans Med Imaging*, 39(9):2760–2771, 9 2020. ISSN 1558-254X (Electronic); 0278-0062 (Print); 0278-0062 (Linking). doi: 10.1109/TMI.2020.2975375.

[52] Chengyue Wu, David A. Hormuth, Guillermo Lorenzo, Angela M. Jarrett, Federico Pineda, Frederick M. Howard, Gregory S. Karczmar, and Thomas E. Yankeelov. Towards Patient-Specific Optimization of Neoadjuvant Treatment Protocols for Breast Cancer Based on Image-Guided Fluid Dynamics. *IEEE Transactions on Biomedical Engineering*, 69(11):3334–3344, 11 2022. ISSN 15582531. doi: 10.1109/TBME.2022.3168402.

[53] Konstantina Kourou, Themis P. Exarchos, Konstantinos P. Exarchos, Michalis V. Karamouzis, and Dimitrios I. Fotiadis. Machine learning applications in cancer prognosis and prediction. In *Computational and Structural Biotechnology Journal*, volume 13, pages 8–17. Elsevier B.V., 2015. doi: 10.1016/j.csbj.2014.11.005.

[54] Mitra Montazeri, Mohadeseh Montazeri, Mahdieh Montazeri, and Amin Beigza-

deh. Machine learning models in breast cancer survival prediction. *Technology and Health Care*, 24, 2016.

[55] Guillermo Lorenzo, Angela M. Jarrett, Christian T. Meyer, Julie C. DiCarlo, John Virostko, Vito Quaranta, Darren R. Tyson, and Thomas E. Yankeelov. A global sensitivity analysis of a mechanistic model of neoadjuvant chemotherapy for triple negative breast cancer constrained by in vitro and in vivo imaging data. *Engineering with Computers*, 2023. ISSN 14355663. doi: 10.1007/s00366-023-01873-0.

[56] Russell C Rockne, Andrea Hawkins-Daarud, Kristin R Swanson, James P Sluka, James A Glazier, Paul Macklin, David A Hormuth, Angela M Jarrett, Ernesto A B F Lima, J Tinsley Oden, George Biros, Thomas E Yankeelov, Kit Curtius, Ibrahim Al Bakir, Dominik Wodarz, Natalia Komarova, Luis Aparicio, Mykola Bordyuh, Raul Rabadan, Stacey D Finley, Heiko Enderling, Jimmy Caudell, Eduardo G Moros, Alexander R A Anderson, Robert A Gatenby, Artem Kaznatcheev, Peter Jeavons, Nikhil Krishnan, Julia Pelesko, Raoul R Wadhwa, Nara Yoon, Daniel Nichol, Andriy Marusyk, Michael Hinczewski, and Jacob G Scott. The 2019 mathematical oncology roadmap. *Phys Biol*, 16(4):41005, 6 2019. ISSN 1478-3975 (Electronic); 1478-3967 (Print); 1478-3967 (Linking). doi: 10.1088/1478-3975/ab1a09.

[57] John H. Lagergren, John T. Nardini, Ruth E. Baker, Matthew J. Simpson, and Kevin B. Flores. Biologically-informed neural networks guide mechanistic modeling from sparse experimental data. *PLoS Computational Biology*, 16(11), 12 2020. ISSN 15537358. doi: 10.1371/journal.pcbi.1008462.

[58] Pietro Mascheroni, Symeon Savvopoulos, Juan Carlos Lopez Alfonso, Michael Meyer-Hermann, and Haralampos Hatzikirou. Improving personalized tumor growth predictions using a Bayesian combination of mechanistic modeling and machine learning. *Communications Medicine*, 1(1), 12 2021. ISSN 2730664X. doi: 10.1038/s43856-021-00020-4.

[59] John Metzcar, Catherine R. Jutzeler, Paul Macklin, Alvaro Köhn-Luque, and Sarah C. Brüningk. A review of mechanistic learning in mathematical oncology, 2024. ISSN 16643224.

[60] Casey E. Stowers, Chengyue Wu, Zhan Xu, Sidharth Kumar, Clinton Yam, Jong Bum Son, Jingfei Ma, Jonathan I. Tamir, Gaiane M. Rauch, and Thomas E. Yankeelov. Combining Biology-based and MRI Data-driven Modeling to Predict Response to Neoadjuvant Chemotherapy in Patients with Triple-Negative Breast Cancer. *Radiology: Artificial Intelligence*, 7(1), 1 2025. ISSN 26386100. doi: 10.1148/ryai.240124.

[61] Thomas E Yankeelov, Nkiruka Atuegwu, David Hormuth, Jared A Weis, Stephanie L Barnes, Michael I Miga, Erin C Rericha, and Vito Quaranta. Clinically relevant modeling of tumor growth and treatment response. *Sci Transl Med*, 5 (187):187ps9, 5 2013. ISSN 1946-6242 (Electronic); 1946-6234 (Print); 1946-6234 (Linking). doi: 10.1126/scitranslmed.3005686.

[62] David A Hormuth II, Angela M Jarrett, Guillermo Lorenzo, Ernesto A B F Lima, Chengyue Wu, Caroline Chung, Debra Patt, and Thomas E Yankeelov. Math, magnets, and medicine: enabling personalized oncology. *Expert Review of Precision Medicine and Drug Development*, 6(2):79–81, 2021. doi: 10.1080/23808993.2021.1878023. URL https://doi.org/10.1080/23808993.2021.1878023.

[63] Nkiruka C Atuegwu, Lori R Arlinghaus, Xia Li, E Brian Welch, Bapsi A Chakravarthy, John C Gore, and Thomas E Yankeelov. Integration of diffusion-weighted MRI data and a simple mathematical model to predict breast tumor cellularity during neoadjuvant chemotherapy. *Magn Reson Med*, 66(6):1689–1696, 12 2011. ISSN 1522-2594 (Electronic); 0740-3194 (Print); 0740-3194 (Linking). doi: 10.1002/mrm.23203.

[64] Y Hayashida, T Hirai, S Morishita, M Kitajima, R Murakami, Y Korogi, K Makino, H Nakamura, I Ikushima, M Yamura, M Kochi, J.-I Kuratsu, and

Y Yamashita. Diffusion-weighted Imaging of Metastatic Brain Tumors: Comparison with Histologic Type and Tumor Cellularity. Technical report. URL www.ajnr.org.

[65] A W Anderson, J Xie, J Pizzonia, R A Bronen, D D Spencer, and J C Gore. Effects of cell volume fraction changes on apparent diffusion in human cells. Technical report, 2000.

[66] Heidi Lyng, Olav Haraldseth, and Einar K. Rofstad. Measurement of cell density and necrotic fraction in human melanoma xenografts by diffusion weighted magnetic resonance imaging. *Magnetic Resonance in Medicine*, 43(6):828–836, 2000. ISSN 07403194. doi: 10.1002/1522-2594(200006)43:6⟨828::AID-MRM8⟩3.0.CO;2-P.

[67] Chengyue Wu, Angela M. Jarrett, Zijian Zhou, Nabil Elshafeey, Beatriz E. Adrada, Rosalind P. Candelaria, Rania M.M. Mohamed, Medine Boge, Lei Huo, Jason B. White, Debu Tripathy, Vicente Valero, Jennifer K. Litton, Clinton Yam, Jong Bum Son, Jingfei Ma, Gaiane M. Rauch, and Thomas E. Yankeelov. MRI-Based Digital Models Forecast Patient-Specific Treatment Responses to Neoadjuvant Chemotherapy in Triple-Negative Breast Cancer. *Cancer Research*, 82 (18):3394–3404, 9 2022. ISSN 15387445. doi: 10.1158/0008-5472.CAN-22-1329.

[68] Alexia L McKnight, Jennifer L Kugel, Phillip J Rossman, Armando Manduca, Lynn C Hartmann, and Richard L Ehman. MR Elastography of Breast Cancer: Preliminary Results. *American Journal of Roentgenology*, 178(6):1411–1417, 6 2002. ISSN 0361-803X. doi: 10.2214/ajr.178.6.1781411. URL https://doi.org/10.2214/ajr.178.6.1781411.

[69] Xia Li, Lori R. Arlinghaus, Gregory D. Ayers, A. Bapsi Chakravarthy, Richard G. Abramson, Vandana G. Abramson, Nkiruka Atuegwu, Jaime Farley, Ingrid A. Mayer, Mark C. Kelley, Ingrid M. Meszoely, Julie Means-Powell, Ana M. Grau, Melinda Sanders, Sandeep R. Bhave, and Thomas E. Yankeelov. DCE-MRI ana-

lysis methods for predicting the response of breast cancer to neoadjuvant chemotherapy: Pilot study findings. *Magnetic Resonance in Medicine*, 71(4):1592–1602, 2014. ISSN 15222594. doi: 10.1002/mrm.24782.

[70] Andreas Stadlbauer, Ewald Moser, Stephan Gruber, Rolf Buslei, Christopher Nimsky, Rudolf Fahlbusch, and Oliver Ganslandt. Improved delineation of brain tumors: An automated method for segmentation based on pathologic changes of 1H-MRSI metabolites in gliomas. *NeuroImage*, 23(2):454–461, 10 2004. ISSN 10538119. doi: 10.1016/j.neuroimage.2004.06.022.

[71] Nelly Gordillo, Eduard Montseny, and Pilar Sobrevilla. State of the art survey on MRI brain tumor segmentation, 10 2013. ISSN 0730725X.

[72] Weijie Chen, Maryellen L. Giger, and Ulrich Bick. A fuzzy c-means (FCM)-based approach for computerized segmentation of breast lesions in dynamic contrast-enhanced MR images. *Academic Radiology*, 13(1):63–72, 1 2006. ISSN 10766332. doi: 10.1016/j.acra.2005.08.035.

[73] Chengyue Wu, Federico Pineda, David A. Hormuth, Gregory S. Karczmar, and Thomas E. Yankeelov. Quantitative analysis of vascular properties derived from ultrafast DCE-MRI to discriminate malignant and benign breast tumors. *Magnetic Resonance in Medicine*, 81(3):2147–2160, 3 2019. ISSN 15222594. doi: 10.1002/mrm.27529.

[74] Angela M Jarrett, David A Hormuth, Vikram Adhikarla, Prativa Sahoo, Daniel Abler, Lusine Tumyan, Daniel Schmolze, Joanne Mortimer, Russell C Rockne, and Thomas E Yankeelov. Towards integration of 64Cu-DOTA-trastuzumab PET-CT and MRI with mathematical modeling to predict response to neoadjuvant therapy in HER2 + breast cancer. *Scientific Reports*, 10(1): 20518, 2020. doi: 10.1038/s41598-020-77397-0. URL https://doi.org/10.1038/s41598-020-77397-0.

[75] Mohammed Benjelloun, Mohammed El Adoui, Mohamed Amine Larhmam, and

Sidi Ahmed Mahmoudi. Automated Breast Tumor Segmentation in DCE-MRI Using Deep Learning. Technical report.

[76] Mohammed El Adoui, Sidi Ahmed Mahmoudi, Mohamed Amine Larhmam, and Mohammed Benjelloun. MRI breast tumor segmentation using different encoder and decoder CNN architectures. *Computers*, 8(3), 9 2019. ISSN 2073431X. doi: 10.3390/computers8030052.

[77] Jun Zhang, Ashirbani Saha, Zhe Zhu, and Maciej A. Mazurowski. Hierarchical Convolutional Neural Networks for Segmentation of Breast Tumors in MRI With Application to Radiogenomics. *IEEE Transactions on Medical Imaging*, 38(2): 435–447, 2 2019. ISSN 1558254X. doi: 10.1109/TMI.2018.2865671.

[78] Zhan Xu, David E. Rauch, Rania M. Mohamed, Sanaz Pashapoor, Zijian Zhou, Bikash Panthi, Jong Bum Son, Ken Pin Hwang, Benjamin C. Musall, Beatriz E. Adrada, Rosalind P. Candelaria, Jessica W.T. Leung, Huong T.C. Le-Petross, Deanna L. Lane, Frances Perez, Jason White, Alyson Clayborn, Brandy Reed, Huiqin Chen, Jia Sun, Peng Wei, Alastair Thompson, Anil Korkut, Lei Huo, Kelly K. Hunt, Jennifer K. Litton, Vicente Valero, Debu Tripathy, Wei Yang, Clinton Yam, and Jingfei Ma. Deep Learning for Fully Automatic Tumor Segmentation on Serially Acquired Dynamic Contrast-Enhanced MRI Images of Triple-Negative Breast Cancer. *Cancers*, 15(19), 10 2023. ISSN 20726694. doi: 10.3390/cancers15194829.

[79] Yang Zhang, Siwa Chan, Vivian Youngjean Park, Kai Ting Chang, Siddharth Mehta, Min Jung Kim, Freddie J. Combs, Peter Chang, Daniel Chow, Ritesh Parajuli, Rita S. Mehta, Chin Yao Lin, Sou Hsin Chien, Jeon Hor Chen, and Min Ying Su. Automatic Detection and Segmentation of Breast Cancer on MRI Using Mask R-CNN Trained on Non–Fat-Sat Images and Tested on Fat-Sat Images. *Academic Radiology*, 29:S135–S144, 1 2022. ISSN 18784046. doi: 10.1016/j.acra.2020.12.001.

[80] Richard Adam, Kevin Dell'Aquila, Laura Hodges, Takouhie Maldjian, and Tim Q. Duong. Deep learning applications to breast cancer detection by magnetic resonance imaging: a literature review, 12 2023. ISSN 1465542X.

[81] Stefan Klein, Marius Staring, Keelin Murphy, Max A. Viergever, and Josien P.W. Pluim. Elastix: A toolbox for intensity-based medical image registration. *IEEE Transactions on Medical Imaging*, 29(1):196–205, 1 2010. ISSN 02780062. doi: 10.1109/TMI.2009.2035616.

[82] Angela M. Jarrett, Danial Faghihi, David A. Hormuth, Ernesto A.B.F. Lima, John Virostko, George Biros, Debra Patt, and Thomas E. Yankeelov. Optimal control theory for personalized therapeutic regimens in oncology: Background, history, challenges, and opportunities, 5 2020. ISSN 20770383.

[83] Xia Li, Benoit M. Dawant, E. Brian Welch, A. Bapsi Chakravarthy, Darla Freehardt, Ingrid Mayer, Mark Kelley, Ingrid Meszoely, John C. Gore, and Thomas E. Yankeelov. A nonrigid registration algorithm for longitudinal breast MR images and the analysis of breast tumor response. *Magnetic Resonance Imaging*, 27(9): 1258–1270, 11 2009. ISSN 0730725X. doi: 10.1016/j.mri.2009.05.007.

[84] Xia Li, Hakmook Kang, Lori R. Arlinghaus, Richard G. Abramson, A. Bapsi Chakravarthy, Vandana G. Abramson, Jaime Farley, Melinda Sanders, and Thomas E. Yankeelov. Analyzing spatial heterogeneity in DCE-and DW-MRI parametric maps to optimize prediction of pathologic response to neoadjuvant chemotherapy in breast cancer. *Translational Oncology*, 7(1):14–22, 2014. ISSN 19365233. doi: 10.1593/tlo.13748.

[85] Randall J LeVeque. *Finite Difference Methods for Ordinary and Partial Differential Equations*. Society for Industrial and Applied Mathematics, 2007. doi: 10.1137/1.9780898717839. URL https://epubs.siam.org/doi/abs/10.1137/1.9780898717839.

[86] Hans Petter Langtangen. *A Primer on Scientific Programming with Python*, volume 6 of *Texts in Computational Science and Engineering*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. ISBN 978-3-642-02474-0. doi: 10.1007/978-3-642-02475-7. URL http://link.springer.com/10.1007/978-3-642-02475-7.

[87] William Stevens, Isabelle M Farrow, Leonidas Georgiou, Andrew M Hanby, Timothy J Perren, Laura M Windel, Daniel J Wilson, Nisha Sharma, David Dodwell, Thomas A Hughes, Barbara JG Dall, and David L Buckley. Breast tumour volume and blood flow measured by MRI after one cycle of epirubicin and cyclophosphamide-based neoadjuvant chemotherapy as predictors of pathological response. *The British Journal of Radiology*, 94(1123), 2021.

[88] Leonidas Georgiou, Daniel J. Wilson, Nisha Sharma, Timothy J. Perren, and David L. Buckley. A functional form for a representative individual arterial input function measured from a population using high temporal resolution DCE MRI. *Magnetic Resonance in Medicine*, 81(3):1955–1963, 3 2019. ISSN 15222594. doi: 10.1002/mrm.27524.

[89] Leonidas Georgiou, Nisha Sharma, David A. Broadbent, Daniel J. Wilson, Barbara J. Dall, Anmol Gangi, and David L. Buckley. Estimating breast tumor blood flow during neoadjuvant chemotherapy using interleaved high temporal and high spatial resolution MRI. *Magnetic Resonance in Medicine*, 79(1):317–326, 1 2018. ISSN 15222594. doi: 10.1002/mrm.26684.

[90] The MathWorks Inc. MATLAB version: 9.13.0 (R2022b), 2022. URL https://www.mathworks.com.

[91] S Sourbron, A Biffar, M Ingrisch, Y Fierens, and R Luypaert. PMI: Plaform for Research in Medical Imaging. *Magn Reson Mater Phy*, 22(1):539, 2009.

[92] Zyad M. Almutlaq, Sarah E. Bacon, Daniel J. Wilson, Nisha Sharma, Tatendashe Dondo, and David L. Buckley. The relationship between parameters meas-

ured using intravoxel incoherent motion and dynamic contrast-enhanced MRI in patients with breast cancer undergoing neoadjuvant chemotherapy: a longitudinal cohort study. *Frontiers in Oncology*, 14, 2024. ISSN 2234943X. doi: 10.3389/fonc.2024.1356173.

[93] Martin Rohrer, Hans Bauer, Jan Mintorovitch, Martin Requardt, and Hanns-Joachim Weinmann. Comparison of Magnetic Properties of MRI Contrast Media Solutions at Different Magnetic Field Strengths. *Investigative Radiology*, 40(11), 2005. URL https://journals.lww.com/investigativeradiology/fulltext/2005/11000/comparison_of_magnetic_properties_of_mri_contrast.5.aspx.

[94] D. Rischin, S. P. Ackland, J. Smith, M. B. Garg, S. Clarke, M. J. Millward, G. C. Toner, and J. Zalcberg. Phase I and pharmacokinetic study of docetaxel in combination with epirubicin and cyclophosphamide in advanced cancer: Dose escalation possible with granulocyte colony-stimulating factor, but not with prophylactic antibiotics. *Annals of Oncology*, 13(11):1810–1818, 11 2002. ISSN 09237534. doi: 10.1093/annonc/mdf305.

[95] Liu Yang, Xiao chen Zhang, Su feng Yu, Hua Qing Zhu, Ai ping Hu, Jian Chen, and Peng Shen. Pharmacokinetics and safety of cyclophosphamide and docetaxel in a hemodialysis patient with early stage breast cancer: A case report. *BMC Cancer*, 15(1), 11 2015. ISSN 14712407. doi: 10.1186/s12885-015-1932-3.

[96] Takumi Sugitani, Shin-ichi Kubota, Shin-ichiro Kuroki, Kenta Sogo, Koji Arihiro, Morihito Okada, Takayuki Kadoya, Michihiro Hide, Miyo Oda, and Takamaro Kikkawa. Complex permittivities of breast tumor tissues obtained from cancer surgeries. *Applied Physics Letters*, 104(25):253702, 6 2014. ISSN 0003-6951. doi: 10.1063/1.4885087. URL https://doi.org/10.1063/1.4885087.

[97] K Levenberg. A method for the solution of certain non-linear problems in least squares. *Q J Appl Math*, II(2):164–168, 1944.

[98] Donald W Marquardt. An Algorithm for Least-Squares Estimation of Nonlinear

Parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11 (2):431–441, 1963. ISSN 03684245. URL http://www.jstor.org/stable/2098941.

[99] S. Varrette, P. Bouvry, H. Cartiaux, and F. Georgatos. Management of an Academic HPC Cluster: The UL Experience. 2014. ISBN 9781479953134.

[100] G. M. Cox, J. M. Gibbons, A. T.A. Wood, J. Craigon, S. J. Ramsden, and N. M.J. Crout. Towards the systematic simplification of mechanistic models. *Ecological Modelling*, 198(1-2):240–246, 9 2006. ISSN 03043800. doi: 10.1016/j.ecolmodel.2006.04.016.

[101] Douglas M. Hawkins. The Problem of Overfitting, 1 2004. ISSN 00952338.

[102] Martin S Alnaes, Anders Logg, Kristian B Olgaard, Marie E Rognes, and Garth N Wells. Unified Form Language: A Domain-Specific Language for Weak Formulations of Partial Differential Equations. *ACM Trans. Math. Softw.*, 40(2), 3 2014. ISSN 0098-3500. doi: 10.1145/2566630. URL https://doi.org/10.1145/2566630.

[103] Alexander J.J. Smits, J. Alain Kummer, Peter C. De Bruin, Mijke Bol, Jan G. Van Den Tweel, Kees A. Seldenrijk, Stefan M. Willems, G. Johan A. Offerhaus, Roel A. De Weger, Paul J. Van Diest, and Aryan Vink. The estimation of tumor cell percentage for molecular testing by pathologists is not accurate. *Modern Pathology*, 27(2):168–174, 2 2014. ISSN 08933952. doi: 10.1038/modpathol.2013.134.

[104] Christophe Geuzaine and Jean-Francois Remacle. Gmsh: a three-dimensional finite element mesh generator with built-in pre-and post-processing facilities. Technical report, 2009.

[105] Steven G Johnson. The NLopt nonlinear-optimization package, 2021. URL https://github.com/stevengj/nlopt.

[106] M J D Powell. A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation. In Susana Gomez and

Jean-Pierre Hennart, editors, *Advances in Optimization and Numerical Analysis*, pages 51–67. Springer Netherlands, Dordrecht, 1994. ISBN 978-94-015-8330-5. doi: 10.1007/978-94-015-8330-5{_}4. URL https://doi.org/10.1007/978-94-015-8330-5_4.

[107] J A Nelder and R Mead. A simplex method for function minimization. *The Computer Journal*, 7:308–313, 1965.

[108] Jorge Nocedal. Updating Quasi-Newton Matrices With Limited Storage. Technical Report 151, 1980. URL http://www.ams.org/journal-terms-of-use.

[109] E. A. Eisenhauer, P. Therasse, J. Bogaerts, L. H. Schwartz, D. Sargent, R. Ford, J. Dancey, S. Arbuck, S. Gwyther, M. Mooney, L. Rubinstein, L. Shankar, L. Dodd, R. Kaplan, D. Lacombe, and J. Verweij. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *European Journal of Cancer*, 45(2):228–247, 1 2009. ISSN 09598049. doi: 10.1016/j.ejca.2008.10.026.

[110] Nola M Hylton, Jeffrey D Blume, Wanda K Bernreuter, Etta D Pisano, Mark A Rosen, Elizabeth A Morris, Paul T Weatherall, Constance D Lehman, Gillian M Newstead, Sandra Polin, Helga S Marques, Laura J Esserman, and Mitchell D Schnall. Locally advanced breast cancer: MR imaging for prediction of response to neoadjuvant chemotherapy-results from ACRIN 6657/I-SPY TRIAL. *Radiology*, 263(3):663–672, 6 2012. ISSN 1527-1315 (Electronic); 0033-8419 (Print); 0033-8419 (Linking). doi: 10.1148/radiol.12110748.

[111] D. Förnvik, K. Lång, I. Andersson, M. Dustler, S. Borgquist, and P. Timberg. Estimates of breastcancer growth rate from mammograms and its relation to tumour characteristics. *Radiation Protection Dosimetry*, 169(1):151–157, 2016. ISSN 17423406. doi: 10.1093/rpd/ncv417.

[112] Johannes DM Otten, Guido van Schoor, Petronella GM Peer, Gerard J den Heeten, Roland Holland, Mireille JM Broeders, and AndrÃ© LM Verbeek.

Growth rate of invasive ductal carcinomas from a screened 50–74-year-old population. *Journal of Medical Screening*, 25(1):40–46, 3 2018. ISSN 0969-1413. doi: 10.1177/0969141316687791.

[113] Ana Margarida Teixeira and Pedro Martins. A review of bioengineering techniques applied to breast tissue: Mechanical properties, tissue engineering and finite element analysis, 2023. ISSN 22964185.

[114] Nilza G. Ramião, Pedro S. Martins, Rita Rynkevic, Antonio A. Fernandes, Maria Barroso, and Diana C. Santos. Biomechanical properties of breast tissue, a state-of-the-art review. *Biomechanics and Modeling in Mechanobiology*, 15(5):1307–1323, 10 2016. ISSN 16177940. doi: 10.1007/s10237-016-0763-8.

[115] Zhaonan Sun, Bronislaw D. Gepner, Patrick S. Cottler, Sang-Hyun Lee, and Jason R. Kerrigan. In Vitro Mechanical Characterization and Modeling of Subcutaneous Adipose Tissue: A Comprehensive Review. *Journal of Biomechanical Engineering*, 143(7), 7 2021. ISSN 0148-0731. doi: 10.1115/1.4050286.

[116] Wei Huang, Yiyi Chen, Andriy Fedorov, Xia Li, Guido H. Jajamovich, Dariya I. Malyarenko, Madhava P. Aryal, Peter S. Laviolette, Matthew J. Oborski, Finbarr O'sullivan, Richard G. Abramson, Kourosh Jafari-Khouzani, Aneela Afzal, Alina Tudorica, Brendan Moloney, Sandeep N. Gupta, Cecilia Besa, Jayashree Kalpathy-Cramer, James M. Mountz, Charles M. Laymon, Mark Muzi, Paul E. Kinahan, Kathleen Schmainda, Yue Cao, Thomas L. Chenevert, Bachir Taouli, Thomas E. Yankeelov, Fiona Fennessy, and Xin Li. The impact of arterial input function determination variations on prostate dynamic contrast-enhanced magnetic resonance imaging pharmacokinetic modeling: A multicenter data analysis challenge, part ii. *Tomography*, 5(1):99–109, 3 2019. ISSN 2379139X. doi: 10.18383/j.tom.2018.00027.

[117] Stephane Urcun, Guillermo Lorenzo, Davide Baroli, Pierre-Yves Rohan, Giuseppe Sciumè, Wafa Skalli, Vincent Lubrano, and Stephane P A Bordas. Oncology

and mechanics: landmark studies and promising clinical applications. Technical report, 2022.

[118] Gokhan Bahcecioglu, Gozde Basara, Bradley W. Ellis, Xiang Ren, and Pinar Zorlutuna. Breast cancer models: Engineering the tumor microenvironment, 4 2020. ISSN 18787568.

[119] Pasquale C Africa, Daniel Arndt, Wolfgang Bangerth, Bruno Blais, Marc Fehling, Rene Gassmöller, Timo Heister, Luca Heltai, Sebastian Kinnewig, Martin Kronbichler, Matthias Maier, Peter Munch, Magdalena Schreter-Fleischhacker, Jan P Thiele, Bruno Turcksin, David Wells, and Vladimir Yushutin. The deal.II library, Version 9.6. 32(4):369–380, 2024. doi: doi:10.1515/jnma-2024-0137. URL https://doi.org/10.1515/jnma-2024-0137.