Towards Adaptive Radiotherapy through Development of Treatment Response Prediction

Anthony V. Lausch
The University of Western Ontario

Supervisor
Dr. Eugene Wong
The University of Western Ontario

Graduate Program in Medical Biophysics

A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

© Anthony V. Lausch 2015

Follow this and additional works at: http://ir.lib.uwo.ca/etd
Part of the Medical Biophysics Commons

Recommended Citation
http://ir.lib.uwo.ca/etd/3078

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact tadam@uwo.ca.
TOWARDS ADAPTIVE RADIOTHERAPY THROUGH
DEVELOPMENT OF TREATMENT RESPONSE
PREDICTION

by

Anthony V. Lausch

Graduate Program in Medical Biophysics

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

© Anthony V. Lausch 2015
Abstract

Despite modern treatment advances, overall survival (OS) remains poor for many cancers such as liver and brain. Cancer is a fundamentally heterogeneous and adaptable disease and therefore personalized adaptive treatment strategies may be a key towards improving OS. Radiotherapy, a commonly used cancer treatment technique which employs ionizing radiation to kill tumours, holds promise for delivering adaptive treatment. However, effective adaptation requires the ability to assess and predict tumour treatment response. Therefore development of treatment response prediction tools represents a critical first step towards improving patient outcomes via treatment adaptation. The overall goal of this thesis is to develop treatment response prediction methods with a view towards guiding adaptive radiotherapy.

First, we investigated the relationship between radiation dose and local tumour control among patients with primary and metastatic colorectal liver tumours. We established and compared their dose-response relationships and found that 84 Gy and 95 Gy of radiation could provide 90% probabilities of 6-month local control for the primary and metastatic groups respectively.

Tumour control most often cannot be improved simply through escalating the dose to the entire tumour due to increased risk of side effects. However, it may be possible to safely increase the dose to tumour sub-volumes. Therefore, the second and third contributions of this thesis involve development of image-based treatment
response prediction methods which are needed to identify tumour sub-volumes where additional radiation should be deposited to improve tumour control.

Our second contribution involved augmenting a voxel-based method known as parametric response mapping (PRM) to account for image registration error (IRE). The augmented PRM helped to quantify and visualize IRE-related variability. In our third contribution, we further generalized PRM to permit collective analysis of multi-parametric image data. The proposed method was applied to multi-parametric imaging from a patient cohort with glioblastoma and was found to predict OS $\geq 18$ months (median OS) with a sensitivity and specificity of 90% and 78% respectively.

In summary, these contributions provided some of the response assessment groundwork needed to guide adaptive RT. Image-based dose response relationships via the augmented and multi-parametric response maps will facilitate personalization and guidance of adaptive radiotherapy.

Keywords: treatment response prediction, adaptive radiotherapy, parametric response mapping, multi-parametric imaging, image analysis, image registration error, dose response modelling, principal component analysis, liver cancer, glioblastoma
Co-Authorship

All of the research chapters contained within the following thesis were produced in collaboration with multiple authors. All co-authors are credited below with respect to their contributions to each chapter. My individual contributions to each chapter are also detailed.

Chapter 2 was co-authored by Mr. Kevin Sinclair, Dr. Michael Lock, Dr. Barbara Fisher, Dr. Nikolaj Jensen, Dr. Stewart Gaede, Dr. Jeff Chen and Dr. Eugene Wong. All co-authors helped to critically review the manuscript in addition to individual contributions listed below. Mr. Sinclair helped to collect and organize patient data and performed preliminary data analysis. Drs. Lock, Fisher and Gaede assisted in conception of the initial research idea and data collection. Drs. Jensen and Chen provided input on data analysis. Dr. Wong assisted in conception of the research idea, data collection and analysis methods. I collected and organized patient data, designed and performed data analyses, and prepared the manuscript.

Chapter 3 was co-authored by Dr. Jeff Chen, Dr. Aaron D. Ward, Dr. Stewart Gaede, Dr. Ting-Yim Lee, and Dr. Eugene Wong. All co-authors helped to critically review the manuscript in addition to individual contributions listed below. Drs. Chen, Ward, and Wong helped develop the research idea and provided input on underlying algorithms and algorithm testing strategies. Dr. Gaede provided input on algorithm testing strategies and analysis methods. Dr. Lee assisted with data collection and
supporting infrastructure. I conceived the original research idea, designed underlying algorithms, designed and performed algorithm testing strategies, and prepared the manuscript.

Chapter 4 was co-authored by Dr. Timothy P.C Yeung, Dr. Jeff Chen, Mr. Elton Law, Dr. Yong Wang, Dr. Benedetta Urbini, Dr. Filipo Donelli, Dr. Luigi Manco, Dr. Enrico Fainardi, Dr. Ting-Yim Lee, and Dr. Eugene Wong. All co-authors helped to critically review the manuscript in addition to the individual contributions listed below. Drs. Chen, Yeung, and Wong helped to develop the research ideas and provided input on underlying algorithms and evaluation strategies. Drs. Wang and Yeung helped to segment MR images for image analysis purposes. Mr. Elton Law produced the cerebral blood volume maps. Drs. Urbini Donelli, Manco, Fainardi, and Lee assisted with data collection and supporting infrastructure. I conceived of the original research idea, analysis algorithms, testing strategies and prepared the manuscript.
To my mother who unfailingly lifts me to my feet after every fall and to my father for teaching me the work-ethic needed to climb back up every hill.
Acknowledgments

It is said that the purpose of education is to replace an empty mind with an open one. My mind may or may not remain empty, but it certainly has been opened by some incredible people that I have had the privilege to interact with over the past four years. Without them, this thesis would not have been possible and I would like to take this opportunity to acknowledge their significant contributions to both this work and my personal and professional development.

First and foremost I would like to thank my supervisor Dr. Eugene Wong. Almost nine years ago you took me under your wing as an undergraduate co-op student and shared with me your passion for research. Your contagious enthusiasm ignited my own passion for research and started me down the rewarding path to where I am today for which I am ever grateful. Thank-you so much for your continual support and guidance during the past four years and for training me to be an effective independent researcher. Finally, thank-you for encouraging an open and flexible approach to research where curiosity always leads the way.

I am also grateful for the invaluable advice and guidance provided by my supervisory committee members Drs. Jeff Chen, Stewart Gaede and Aaron Ward. Dr. Jeff Chen: Your advice, incredible attention to detail and above all your ever-present positivity and encouragement were instrumental in the completion of this work. Dr. Stewart Gaede: Your always pragmatic advice coupled with your approachable down-
to-earth style provided much-appreciated reality checks in committee meetings, paper revisions, and in casual conversation throughout the past four years. Dr. Aaron Ward: Thank you so much for all of your encouragement, guidance, and critical feedback, all of which was provided with a wonderful sense of humour.

To Drs. Jerry Battista, Grace Parraga, Aaron Ward, and Hanif Ladak: Thank-you for encouraging my involvement in the graduate department. The opportunities that you extended to me provided some of my most treasured experiences in graduate school. These experiences not only contributed to my professional development but also significantly improved my confidence and overall outlook. This in turn had a significant positive impact on my research pursuits for which I am very grateful.

Thank-you to Drs. Ting Lee, Michael Lock, Timothy Yeung, and Enrico Fainardi for acquiring and sharing all of the data that was vital for demonstration and testing of the methods proposed within this thesis. Also, thank-you very much for all of your valuable expertise and feedback related to these works.

I also truly appreciate all of the contributions and support from fellow students within the department. To research group members Dr. Nikolaj Jensen, Tom Hrinivich, Michael Jensen and Niloufar Zarghami: Thank-you so much for all of your support and insightful feedback that you provided at well-over a hundred different group meetings. To labmates Omar El-Sherif, Ilma Xhaferllari, John Patrick, Dr. Brandon Disher and Jeff Kempe: Thank-you for all of your support during challenging times, your advice on everything from what colour to use in a figure to what I should be when I grow up, and above all, thank-you for the laughs.

Last but not least, a tremendous thank-you to all of my friends and family from which support during the last four years was unwavering, compassionate, and dependably hilarious. I cannot overexpress how much you all mean to me and how vital your support has been in making this work possible.
# Contents

Abstract ii  
Co-Authorship iv  
Dedication vi  
Acknowledgments vii  
Contents ix  
List of Tables xiii  
List of Figures xiv  
List of Appendices xvi  
Nomenclature xvii  

1 Introduction 1  
1.1 Overview ......................................................... 1  
1.2 Cancer and the hallmarks of adaptation .......................... 2  
1.3 Common treatment strategies and outcomes ...................... 3  
1.3.1 Treatment strategies ........................................ 3  
1.3.2 Outcomes .................................................... 5  
1.4 Radiotherapy .................................................... 6  
1.4.1 An overview of conventional radiotherapy ................... 6  
1.4.2 Radiotherapy delivery techniques ............................ 7  
1.4.3 Key challenges and the case for adaptive radiotherapy .... 11  
1.5 Strategies for adaptive radiotherapy ............................. 14  
1.5.1 Adapting radiation dose prescriptions to patients .......... 14  
1.5.2 Adapting dose distributions to intra-tumour heterogeneity .. 17  
1.6 Evaluation of treatment response and the need for radiotherapy response prediction ........................................... 14
4 Generalized multi-parametric response mapping using principal component analysis 103
4.1 Introduction ............................................. 103
4.2 Materials and Methods ................................. 105
  4.2.1 Original PRM: single-parameter response map 105
  4.2.2 Multi-parametric response map .................. 107
    4.2.2.1 Reference principal component analysis 108
    4.2.2.2 Target data classification using the reference PCA 111
    4.2.2.3 Response map construction .................. 114
  4.2.3 Demonstration ....................................... 116
    4.2.3.1 Patient and image data ....................... 117
    4.2.3.2 Image registration and segmentation ........ 119
    4.2.3.3 Parametric response mapping ................. 120
    4.2.3.4 Response prediction and analyses ........... 123
4.3 Results .................................................. 125
4.4 Discussion .............................................. 132
4.5 Conclusion .............................................. 138
4.6 Acknowledgements ..................................... 138
References .................................................. 139

5 Conclusion and Future Work 141
5.1 Summary of findings .................................. 141
  5.1.1 Radiation dose-response relationships for liver tumours 141
  5.1.2 Augmented parametric response mapping ........... 142
  5.1.3 Multi-parametric response mapping ................ 143
5.2 Limitations and Future Work ......................... 144
  5.2.1 Radiation dose-response modelling ............... 144
  5.2.2 Augmented parametric response map ................ 145
  5.2.3 Multi-parametric response map .................... 146
  5.2.4 Verifying the predictive utility of A-PRM and MPRM 148
  5.2.5 Evaluating feasibility of response map guided radiotherapy 149
  5.2.6 Treatment response prediction software tools ...... 150
References .................................................. 151

A Ethics Approvals 153
List of Tables

2.1 Patient characteristics and treatment data .......................... 49
2.2 Dose-response parameter sensitivity to \( \alpha/\beta \) .................. 58
3.1 A-PRM and PRM classification accuracy differences for glioma cases 89
3.2 A-PRM and PRM classification accuracy differences for hepatocellular carcinoma cases .................................................. 90
3.3 Median A-PRM classification for misclassified \( T_{95} \) voxels .......... 91
4.1 Summary of parametric response map analyses ........................ 122
4.2 Reference PCA .................................................................... 128
4.3 Summary of PRM and MPRM classification performance ............ 131
List of Figures

1.1 Modern linear accelerator ....................................... 8
1.2 Multi-leaf collimator ............................................. 8
1.3 Example glioblastoma RT plan ................................. 10
1.4 Example NTCP curve ............................................ 16
1.5 Sub-volume boosting and dose-painting ...................... 18
1.6 RECIST measurement example ................................. 21
1.7 Artificial functional image example ........................... 24
1.8 Region based image analysis .................................... 26
1.9 Voxel based image analysis ..................................... 28
1.10 Example ADC and rCBV maps ................................. 30

2.1 Tumour control probability curves ........................... 51
2.2 Dose-control histories .......................................... 53
2.3 Local control Kaplan-Meier curves ............................ 54
2.4 Time to loss of local control vs pre-treatment tumour volume and pre-
scribed dose .......................................................... 55

3.1 Influence of image registration error on PRM analysis .... 72
3.2 Augmented parametric response map voxel classification 76
3.3 Example glioma and hepatocellular carcinoma functional images 79
3.4 Example zero IRE PRMs and A-PRMs for glioma and hepatocellular carcinoma ........................................ 84
3.5 Example non-zero IRE PRMs and A-PRMs for glioma and hepatocellular carcinoma .................................. 86
3.6 PRM and A-PRM classification accuracy as a function of IRE ................................................................. 88
3.7 Influence of under/over estimation of IRE on A-PRM classification accuracy ............................................. 92
3.8 Influence of under/over estimation of IRE on A-PRM classification accuracy ............................................. 92
4.1 Reference principal component analysis .................................................. 110
4.2 Target data classification using the reference PCA ................................................................. 113
4.3 Response map construction ................................................................. 115
4.4 Multi-parametric image data for two glioma cases .................................................. 118
4.5 PRM examples for patients with short and long OS .................................................. 126
4.6 MPRM examples for patients with short and long OS .................................................. 127
4.7 Summary of PRM and MPRM analyses of CEL and PERIPH .................................................. 129
4.8 ROC analyses for prediction of OS .................................................. 130
List of Appendices

Ethics Approvals 153
Copyright Releases 162
Curriculum Vitae 174
## Nomenclature

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D-CRT</td>
<td>3D conformal radiotherapy</td>
</tr>
<tr>
<td>ABF</td>
<td>arterial blood flow</td>
</tr>
<tr>
<td>ADC</td>
<td>apparent diffusion coefficient</td>
</tr>
<tr>
<td>AFP</td>
<td>α-fetoprotein</td>
</tr>
<tr>
<td>A-PRM</td>
<td>augmented parametric response map</td>
</tr>
<tr>
<td>AUC</td>
<td>area under the receiver operating characteristic curve</td>
</tr>
<tr>
<td>BV</td>
<td>blood volume</td>
</tr>
<tr>
<td>CI</td>
<td>confidence interval</td>
</tr>
<tr>
<td>CT</td>
<td>computed tomography</td>
</tr>
<tr>
<td>CTV</td>
<td>clinical target volume</td>
</tr>
<tr>
<td>CEA</td>
<td>carcino-embryonic antigen</td>
</tr>
<tr>
<td>CEL</td>
<td>contrast-enhancing lesion</td>
</tr>
<tr>
<td>DCE-CT</td>
<td>dynamic contrast-enhanced computed tomography</td>
</tr>
<tr>
<td>DSC</td>
<td>dynamic susceptibility contrast-enhanced</td>
</tr>
<tr>
<td>DWI</td>
<td>diffusion-weighted imaging</td>
</tr>
<tr>
<td>EUD</td>
<td>equivalent uniform dose</td>
</tr>
<tr>
<td>FAIR</td>
<td>flexible algorithms for image registration toolkit</td>
</tr>
<tr>
<td>GTV</td>
<td>gross tumour volume</td>
</tr>
<tr>
<td>Gy</td>
<td>Gray, unit of absorbed radiation dose, 1Gy = 1 J/kg</td>
</tr>
<tr>
<td>HCC</td>
<td>hepatocellular carcinoma</td>
</tr>
<tr>
<td>IGRT</td>
<td>image-guided radiation therapy</td>
</tr>
<tr>
<td>IMRT</td>
<td>intensity modulated radiation therapy</td>
</tr>
<tr>
<td>IRE</td>
<td>image registration error</td>
</tr>
<tr>
<td>LQM</td>
<td>linear quadratic model</td>
</tr>
<tr>
<td>MET</td>
<td>metastatic colorectal liver tumour</td>
</tr>
<tr>
<td>MLC</td>
<td>multi-leaf collimator</td>
</tr>
<tr>
<td>MPRM</td>
<td>multi-parametric response map</td>
</tr>
<tr>
<td>MRI</td>
<td>magnetic resonance imaging</td>
</tr>
<tr>
<td>NGF</td>
<td>normalized gradient fields</td>
</tr>
<tr>
<td>NTCP</td>
<td>normal tissue complication probability</td>
</tr>
<tr>
<td>OS</td>
<td>overall survival</td>
</tr>
<tr>
<td>PC</td>
<td>principal component</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>PCA</td>
<td>principal component analysis</td>
</tr>
<tr>
<td>PDF</td>
<td>probability distribution function</td>
</tr>
<tr>
<td>PERIPH</td>
<td>region of interest defined as a 1 cm thick shell around the contrast enhancing lesion</td>
</tr>
<tr>
<td>PET</td>
<td>positron emission tomography</td>
</tr>
<tr>
<td>PRM</td>
<td>parametric response map</td>
</tr>
<tr>
<td>PTV</td>
<td>planning target volume</td>
</tr>
<tr>
<td>PWI</td>
<td>perfusion-weighted imaging</td>
</tr>
<tr>
<td>rCBV</td>
<td>relative cerebral blood volume</td>
</tr>
<tr>
<td>RECIST</td>
<td>response evaluation criteria in solid tumours</td>
</tr>
<tr>
<td>RILD</td>
<td>radiation induced liver disease</td>
</tr>
<tr>
<td>ROC</td>
<td>receiver operating characteristic</td>
</tr>
<tr>
<td>ROI</td>
<td>region of interest</td>
</tr>
<tr>
<td>RT</td>
<td>radiation therapy</td>
</tr>
<tr>
<td>SBRT</td>
<td>stereotactic body radiation therapy</td>
</tr>
<tr>
<td>TCP</td>
<td>tumour control probability</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Overview

Approximately 41% of females and 45% of males in Canada will develop cancer in their lifetime [1]. Cancer is Canada’s leading cause of mortality accounting for approximately 30% of all deaths in 2014 [1]. While the age-adjusted incidence of cancer has not increased in Canada in the past decade, the absolute number of new cases is expected to continue rising due to the increasing age of the Canadian population [1].

Despite modern treatment advances, the overall prognosis remains poor for many types of cancer such as liver, lung, brain, and late-stage disease [1, 2]. The hallmarks of cancer identify it as a fundamentally heterogeneous and adaptable disease [3]. Consequently, the development of personalized treatment strategies which adapt to heterogeneous tumour response may be a key towards improving survival and/or quality of life.

The development of effective treatment response prediction represents a critical first step towards implementing adaptive treatment strategies. If it is known how a patient’s tumour is going to respond during the course of treatment, then this information could be used to adapt treatments to further optimize tumour control.
and/or reduce treatment related side effects.

This thesis describes the development of treatment response prediction methods with a view towards guiding adaptive radiotherapy. The present chapter provides a brief introduction to cancer and its treatment (§1.2-1.4) focusing on strategies for adaptive radiotherapy and image-based treatment response prediction (§1.5-1.7). The overall goal of this work and research objectives are also presented within this chapter (§1.9).

1.2 Cancer and the hallmarks of adaptation

Cancer is a group of related diseases including over one hundred different types [4]. World-wide the most common types of cancer among women are breast, colorectal, lung, cervical, and stomach, and among men the most common types of cancer are lung, prostate, colorectal, stomach, and liver [5]. Cancer has a diverse set of underlying causes and suspected risk factors including but not limited to tobacco and alcohol use, diet and obesity, radiation or carcinogen exposure, viral or bacterial infections and hereditary factors [6, 7]. Cancer may also develop spontaneously without a direct known cause. The development of cancer involves a series of genetic mutations which lead to unrestrained cell growth and the ability of cells to metastasize to distant sites. Ultimately, cancer causes morbidity and death through inhibiting the normal function of afflicted organs or tissues.

In 2000, Hanahan and Weinberg proposed six common hallmarks of cancer with the underlying idea that normal cells progressively evolve towards exhibiting these traits during malignant transformation [3]. These traits included resisting cell death, sustaining proliferative signalling, evading growth suppressors, activating invasion and metastasis, enabling replicative immortality, and inducing angiogenesis. Since
then, Hanahan and Weinberg have proposed two emerging hallmarks, deregulation of cellular energetics and avoiding immune destruction, as well as two hallmark-enabling characteristics, tumour promoting inflammation and genetic instability [8].

In aggregate, these traits identify cancer as a fundamentally heterogeneous and adaptable disease. Sustained proliferative signalling, evasion of growth suppressors, and the ability to grow new vasculature (angiogenesis) to deliver nutrients and remove waste products enable cancer to grow quickly and in a sustained manner. When paired with genetic instability, this unrestrained proliferative capability amplifies and sustains genetic diversity which supports cancer’s ability to adapt to selective pressures from treatment, immune response, or the microenvironment.

1.3 Common treatment strategies and outcomes

1.3.1 Treatment strategies

A diverse array of strategies is available to treat cancer. Common treatment strategies include surgery, chemotherapy, radiotherapy, or some combination of one or more of these treatments.

Surgical removal can be particularly effective for patients that have a solitary well-defined tumour such as in early stage lung, breast, or liver cancer [9–11]. Conversely, surgery is often contraindicated for patients with advanced stage cancer where tumours tend to be large, multi-focal, diffuse, or have metastasized to multiple sites. Surgery is also not an option for blood-borne cancers such as lymphoma and leukemia and for patients who are medically unfit to undergo general surgery.

Chemotherapy typically involves the oral or intravenous administration of chemical agents into the bloodstream or directly to the tumour. Depending on the specific agent, tumour cell kill can be achieved in a number of different ways including but not
limited to causing direct DNA damage and prevention or disruption of cell replication. For example, temozolomide is an orally ingested chemotherapy agent used to treat high-grade glioma that triggers cell death via DNA alkylation [12]. Chemotherapy is used for treating systemic disease such as metastatic and blood-borne cancers and is often used in combination with surgery or radiotherapy to eliminate residual or recurrent disease. Newer targeted drug therapies which are not commonly classified as cytotoxic chemotherapy agents have also been developed. These novel therapies target certain cellular pathways that support tumour growth such as epidermal growth factor and angiogenic pathways [13, 14].

Radiotherapy (RT) involves the use of ionizing radiation to kill tumours and can be delivered externally via the use of linear accelerators (external beam RT) or internally via the use of small radioactive seeds that are temporarily or permanently implanted within the tumour. Radiation causes tumour cell death via direct and indirect action on DNA. Direct action involves the direct interaction between incident radiation and tumour cell DNA which can cause DNA chemical changes or breaks. Indirect action involves the interaction of incident radiation with molecules other than DNA within tumour cells (predominantly water molecules), resulting in the production of reactive oxygen species which cause DNA damage through ionization and the breaking of chemical bonds. Indirect action is estimated to account for approximately two thirds of DNA damage caused by conventional external beam gamma-ray radiation[15]. Since radiotherapy is a targeted therapy it is well-suited towards treatment of solitary tumours which can be visualized via medical imaging such as computed tomography (CT) or magnetic resonance imaging (MRI). RT is a central theme within this thesis and so a more in-depth discussion of RT is provided in the next section.
1.3.2 Outcomes

Survival and quality of life are typical cancer treatment outcomes that are measured. Survival varies widely depending on cancer type. For example, 5-year relative overall survival (OS) rates for two of the most common cancers, breast and prostate, have been estimated to be 89% and 99% respectively independent of the stage at diagnosis. In contrast, the same figures for liver and brain cancer (focuses of this thesis) are 18% and 33% respectively [16]. Outcomes also vary depending on the stage of cancer at diagnosis. For example, the 5-year OS rates for breast and prostate cancer drop to 25.9% and 28.2% respectively for the subset of patients with advanced stage disease (e.g. distant metastases) [16].

Both patients who respond and do not respond to treatment may experience treatment-related side-effects that can have a significant impact on their quality of life. Depending on the cancer type, side effects could include things such as neurocognitive impairment (brain), difficulty breathing (lung), sexual dysfunction (prostate), urinary and fecal incontinence (bladder, colorectal), difficulty swallowing (head and neck), poor cosmesis (breast), and infertility [17]. Side-effects place limits on the amount of treatment that can be safely delivered to the patient without negatively impacting their health or quality of life. For example, the dose of radiation that can be safely prescribed to liver tumours is often constrained by the risk of causing radiation induced liver disease [18].

In summary, despite favourable overall survival for some cancers, there remains a strong need to develop new treatment strategies to improve overall outcomes including survival and treatment related toxicities. In §1.4 and §1.5 we introduce RT in more detail and discuss the case for implementing adaptive RT strategies to improve treatment outcomes.
1.4 Radiotherapy

1.4.1 An overview of conventional radiotherapy

According to the National Cancer Institute (Bethesda MD), approximately half of all cancer patients will receive some form of radiation therapy during their treatment [19]. There are several major steps in the external beam RT workflow which include CT simulation, treatment planning, patient setup and treatment delivery. CT simulation involves acquiring a CT scan of the patient (planning CT) with the patient positioned in a way that simulates a reproducible treatment position. An RT plan detailing the precise amount of radiation to be deposited throughout the patient is then generated based on the planning CT using a treatment planning system. On treatment day, the radiation therapist sets the patient up on the treatment couch so that they match their position on the planning CT. The setup phase prior to treatment involves the use of patient immobilization equipment and medical imaging devices either built in or calibrated with the linear accelerator (e.g. on-board imaging) in order to improve and verify the accuracy of patient positioning. Accurate setup is crucial for ensuring that the RT plan is delivered to the correct location within the body. This prevents geographic miss which would under-treat the tumour and over irradiate surrounding normal tissues, leading to tumour recurrence and increased risks of treatment side effects.

External beam RT uses a specially designed linear accelerator to generate and deliver ionizing radiation. The RT plan is delivered in a series of treatment fractions (e.g. one per day) over the course of several weeks. For example 60 Gray (unit of absorbed radiation dose, 1 Gy = 1 J/kg) of radiation is commonly delivered to patients in 2 Gy fractions for a total of 30 fractions. The patient must return to the cancer clinic for each fraction and be precisely positioned by the therapist each time. This
fractionated approach takes advantage of tumour and normal tissue radiobiology in order to increase tumour cell kill (improving tumour control) and decrease normal tissue cell kill (decreasing side-effects) [15]. Fractionation allows time for the tumour to become re-sensitized to radiation via re-oxygenation of hypoxic regions [15]. It also allows time for some of the tumour cells which were previously in a radiation insensitive phase of the cell-division cycle (e.g. S-phase) to eventually re-distribute into more radiation sensitive phases of the cycle (e.g. G2 and M phases) [15]. Fractionation helps to protect normal tissues by allowing time for normal tissue DNA repair and for repopulation (replication) of normal tissue cells [15]. In §1.4.2 and §1.4.3 we introduce RT in more detail and discuss the case for implementing adaptive RT strategies to improve treatment outcomes.

1.4.2 Radiotherapy delivery techniques

There are a variety of RT delivery techniques available on modern linear accelerators including but not limited to 3D conformal RT (3D-CRT), intensity-modulated RT (IMRT), and stereotactic body RT (SBRT). For each technique, the angle, shape, and intensity of radiation beams are selected in order to deliver a prescribed dose of radiation to the tumour while minimizing the dose received by surrounding normal tissues. Figure 1.1 shows an example of a modern linear accelerator.

Linear accelerators have a pair of moveable tungsten blocks (jaws) in the treatment head which can be used to create rectangular shaped radiation fields. The blocks provide an effective radiation shield due to tungsten’s high density and melting point. The multi-leaf collimator (MLC), which is composed of two large sets of thin independently-controlled tungsten leaves, can then be used to further refine the shape of radiation fields. Figure 1.2 provides a view of the MLC along the direction of the beam. On board imaging (e.g. portal imager, cone-beam CT) is attached to
Figure 1.1: Modern linear accelerator with on board imaging (cone-beam CT).

Figure 1.2: A view of the multi-leaf collimator within the treatment head of the linear accelerator. Each tungsten leaf can be moved independently to create unique shapes blocking all or part of the beam.
the linear accelerator. The entire linear accelerator can be rotated around the patient to permit multiple beam delivery and imaging angles.

Intensity-modulated RT (IMRT) is a more modern technique that uses multiple MLC shapes at each beam angle in order to create complex non-uniform intensity beams. Each shape blocks and exposes different parts of the tumour and surrounding normal tissue. Sophisticated optimization algorithms are used to determine the specific MLC shapes that must be used at each angle in order to meet a set of planning objectives specified by the user. Objectives could include constraints on the uniformity and magnitude of radiation delivered to the tumour and normal tissues. Figure 1.3 illustrates an example IMRT dose distribution for treatment of glioblastoma which is one of the types of cancer that we will address in this thesis. The visible gross tumour volume (GTV) is shown in pink. The GTV is then expanded by a spatial margin to account for disease spread that is not visible on imaging. This expansion is constrained by anatomic boundaries to disease spread such as the skull in the example from 1.3. The resulting region, called the clinical target volume (CTV), encompasses the GTV and is shown in green. For glioblastomas, the CTV is much larger than the GTV since glioblastoma tumours are highly invasive and tumour outgrowths are known to occupy a large region around the visible tumour. Finally, the CTV is expanded by another margin to account for treatment delivery uncertainties such as setup error. This region is called the planning target volume (PTV). Three levels of radiation dose are shown as isodose lines.

The overall goal of RT is to deliver a high uniform dose of radiation to the target (i.e. to the PTV) while limiting the dose received by critical structures. In this example the PTV is immediately adjacent to a critical structure (brainstem). Due to the ability of IMRT to produce conformal dose distributions with steep dose gradients, the brainstem receives substantially less radiation than the PTV. Similarly, the lenses
of the eye, optic nerves, and contralateral hippocampus are spared.

Stereotactic body RT (SBRT) is a technique which involves delivering a patient’s entire treatment in a small number of high-dose fractions (e.g. 5-6). SBRT often uses intensity-modulation and requires precise image-guidance in order to ensure that surrounding normal tissues are not exposed to the high doses of radiation that are delivered in each fraction. SBRT is a common technique for treatment of liver and lung and is an important topic discussed further in Chapter 2 where we investigate the relationship between tumour control and radiation dose for liver tumour SBRT.
1.4.3 Key challenges and the case for adaptive radiotherapy

The overall goal of curative-intent cancer treatment is to maximize tumour control while minimizing treatment-related side effects. Conventionally, RT tries to accomplish this goal by delivering a high uniform dose of radiation to the tumour while minimizing the dose received by surrounding healthy tissues. However, there are several key challenges to this approach which support the case for developing personalized adaptive RT strategies. Four challenges are broadly described below and this thesis will address two of these challenges.

1) Determination of ideal dose prescriptions

The relationship between the radiation dose fractionation (particularly SBRT) and the resulting tumour control, as well as the associated treatment side-effects may not always be well-understood. Therefore, in some cases, it can be unclear about the trade-offs between tumour control and normal tissue toxicities. This equivocation is particularly challenging for cancers with poor outcomes or for late-stage disease where patients are declining in health with no curative intent treatment options available. In cases where the disease is localized, RT may be able to substantially assist in survival and improve quality of life. Liver cancer provides an instructive example where there is a risk of radiation induced liver disease (RILD) [18]. The risk of RILD is related to the size of the tumour relative to the rest of the healthy liver. On the other hand, sufficient dose must be given to control the tumour. Therefore, the resulting cost-benefit analysis between tumour control and RILD is patient-specific. Determining the radiation doses required to effectively control liver tumours is one of the focuses of this thesis.
2) Motion management during treatment

Voluntary and involuntary patient motion during treatment can cause radiation to be deposited where it is not intended, decreasing the dose received by the tumour and increasing the dose received by normal tissues. Dose can be delivered to larger regions which encompass the tumour motion, however this also increases the dose received by normal tissues. This in turn may decrease the overall radiation dose that can be prescribed to the tumour. Lung and liver tumours provide examples where motion is a challenge. Lung and liver tumours are influenced by respiratory motion and in two small studies have been shown to move an average of 12 mm ± 2 mm and 9 mm ± 5 mm in the superior-inferior directions respectively [20, 21]. This motion is patient-specific, pseudo-periodic, and may or may not be consistent throughout treatment [20, 21].

A variety of techniques have been developed to manage breathing-related tumour motion. Keall et al. [22] provide a comprehensive review of current clinical motion management strategies. Some of the most common techniques include immobilization, breath-hold, active breathing control, and gating. Adapting radiation delivery to inter and intra-fraction patient motion is a key aspect of current clinical practice and is typically referred to as image-guided RT (IGRT).

3) Geometric adaptation

Tumours may change size and patients may gain or lose weight during treatment. For example, one study found that head and neck tumours shrink by a median of 69.5% in volume during RT [23]. As a result of these intra-treatment changes, the planning CT may no longer accurately represent the anatomy of the patient. Therefore the RT plan, which is based on the pre-treatment planning CT, may become sub-optimal leading to patient-specific changes in the dose received by the tu-
mour and surrounding normal tissues. Adaptive re-planning has been implemented whereby patients undergo one or more planning CT scans throughout treatment and the treatment plan is updated according to large tumour and anatomical changes [24]. Adaptive re-planning based on volumetric tumour and anatomical changes has been investigated for head and neck, lung, bladder, and prostate cancers in the clinic [24–30].

4) *Functional adaptation*

Finally, as discussed in §1.2, cancer is a fundamentally heterogeneous and adaptable disease. Significant genetic variation exists between different tumour types, between tumours of the same type, and even between different cells within a single tumour [31]. Currently, different treatment regimens are applied to different tumour types, addressing a portion of this variability. However, standardized treatment protocols are applied to tumours of the same type and therefore inter-patient variability is not addressed. Moreover, tumours are usually targeted with a uniform dose of radiation which cannot address intra-tumour variability. Different sub-regions within a tumour can be more or less sensitive to radiation depending on factors like cycling hypoxia and local oxygen concentration [32]. These factors can also change throughout the course of treatment. Therefore the use of standardized uniform doses may not maximize tumour control for individual patients. While adaptation due to volumetric changes in tumours is being practiced in the clinic, adaptation to intra-tumour heterogeneity has not yet entered the clinic. This thesis investigates and develops tools to support this goal.
In summary, key challenges in conventional RT relate to the application of the same treatment plan throughout the fractionated treatment course. Cancer represents a heterogeneous and adaptable disease which may necessitate personalized adaptive treatment strategies in order to maximize tumour control and minimize treatment-related side-effects. RT is particularly well suited towards adaptive implementations since treatment fractionation provides multiple opportunities to evaluate and adapt during the treatment course. RT is already an image-guided localized treatment and so has the potential to adapt to local variability. However, personalized adaptive RT requires additional resources and conventional standardized RT already provides favourable outcomes for many cancers. Therefore investigation of adaptive strategies may be best suited to disease contexts with relatively poor outcomes such as short survival or high toxicities.

This thesis focuses on personalized dose-prescription (the first challenge) and using functional imaging to assess tumour response to guide RT adaptation (the fourth challenge). Current and emerging adaptive RT strategies related to these challenges are described in the next section.

1.5 Strategies for adaptive radiotherapy

1.5.1 Adapting radiation dose prescriptions to patients

The full relationship between the amount of radiation delivered and tumour control is less well-understood than for normal tissue complications. However, substantial effort has been undertaken to try to define normal tissue dose limits beyond which unacceptable side-effects are expected to occur. For example, a dose of 59 Gy to 1-10 cc of the brain stem is associated with a 5% rate of permanent cranial neuropathy [33]. Similarly, dose-volume constraints have been defined for many other critical
structures including but not limited to the brain, spinal cord, rectum, bladder, lungs, liver, and heart [33]. These normal tissue dose constraints are then incorporated into treatment planning to ensure that side-effect severity is minimized. Constraint-based planning is a fundamental part of current clinical practice which helps to adapt standardized dose prescriptions to the anatomy of different patients.

Less commonly, normal tissue complication probability (NTCP) models may be integrated into clinical practice. These models seek to predict the probability that a patient will develop a specific side effect as a function of the dose received by a normal tissue of interest. For example, Dawson et al. [18] modelled the probability that a patient would develop radiation induced liver disease (RILD) less than 4 months after RT. The dose prescribed to liver tumours can then be adapted to individual patients based on managing their risk of developing RILD according to the NTCP model [34, 35]. This particular NTCP model plays an important role in Chapter 2 of this thesis. Marks et al. [33] provide an excellent overview of NTCP models used in the clinic. The Lyman-Kutcher-Burman NTCP model [36, 37] is among the most commonly used methods and is described by the following three equations:

$$NTCP = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}} dx$$  \hspace{1cm} (1.1)$$

$$t = \frac{EUD - TD_{50}}{mTD_{50}}$$  \hspace{1cm} (1.2)$$

$$EUD = (\sum_{i} v_i D_i^{\frac{1}{n}})^n$$  \hspace{1cm} (1.3)$$

The $EUD$ (equivalent uniform dose) summarizes the non-uniform normal tissue dose distribution. $TD_{50}$ describes the uniform dose that when received by the normal
tissue of interest would elicit a 50% probability of a side-effect. $m$ controls the slope of the NTCP curve and $n$ describes how much the NTCP is influenced by the volume of irradiated normal tissue. $D_i$ represents the dose received by a fractional volume $v_i$ of the normal tissue of interest. The model is then fit to the dose and complication data amongst a group of patients to determine the $TD_{50}$, $m$, and $n$ model parameters. Figure 1.4 provides an example of a typical sigmoidal NTCP model curve.

![Figure 1.4: Example of a typical sigmoidal NTCP model for a fictional side-effect with $TD_{50} = 35$ Gy.](image)

Similar to normal tissue, tumour control probability (TCP) models have also been developed which seek to predict the probability of tumour response at some time post-RT as a function of the dose received by the tumour. For example, Chang et al. [38] modelled the probability of tumour local control at 1 year post-treatment for patients treated for colorectal liver metastases. Many different models have been investigated and a review is provided by O’Rourke et al. [39]. The logistic TCP model [40] is a commonly used approach and is described by

$$TCP = \frac{1}{1 + (D_{50}/D)^k} \quad \text{(1.4)}$$
where $D$ represents the dose received by the tumour, $D_{50}$ represents the dose required to elicit a 50% probability of tumour control, and $k$ controls the slope of the TCP curve at $D_{50}$. The model parameters $D_{50}$ and $k$ are then acquired by fitting the model to patient tumour dose and control data.

Together, NTCP and TCP model the relationship between radiation dose, tumour control and treatment side-effects. Knowledge of these relationships can then be used prospectively during treatment planning to adapt prescription doses in an effort to maximize the cost-benefit ratio for individual patients. Liver NTCP models have been previously established [18]. However, liver tumour irradiation has only recently been enabled through advances in radiation targeting (e.g. IMRT, image-guidance) which can reduce the risk of radiation-induced liver disease to acceptable levels. Therefore liver TCP is not as well-understood as liver NTCP and is the focus of Chapter 2.

1.5.2 Adapting dose distributions to intra-tumour heterogeneity

Significant genetic variation exists within individual tumours and different tumour sub-regions can be more or less sensitive to radiation depending on temporally varying factors like hypoxia and local oxygen concentration [31, 32]. Sub-volume boosting and dose-painting-by-numbers strategies have been proposed to address this variability in order to improve tumour control and decrease side-effects [41].

Sub-volume boosting typically involves escalating the dose delivered to a suspected radio-resistant tumour sub-region by a fixed amount where resistance is assessed using functional imaging (assessment discussed further below). In contrast, dose-painting-by-numbers involves continuous modulation of the dose throughout the entire tumour according to local radio-sensitivity. Regions suspected to be more radio-resistant re-
ceive more dose and radio-sensitive regions receive less dose. Sub-volume boosting is the simpler of the two techniques as it can be readily implemented using clinically available treatment planning systems. This is because clinical treatment planning systems are already adept at creating plans which meet discrete dose objectives for well-defined regions of interest. In contrast, dose-painting by numbers requires more sophisticated treatment planning optimization algorithms which can meet a continuum of dose objectives within the tumour. This approach requires more precise radiation delivery since the radiation dose is modulated within the tumour on a finer spatial scale and has yet to be tested clinically for its feasibility. For both approaches, tumour radio-resistance could be re-assessed during treatment to monitor and adapt to temporal variation. Figure 1.5 demonstrates the difference between sub-volume boosting and dose painting by numbers using a mock tumour example.

![Mock tumour example](image)

Figure 1.5: Mock tumour example (circle) illustrating the differences between sub-volume boosting and dose-painting by numbers with respect to radiation dose deposition. $D_{\text{min}}$ and $D_{\text{max}}$ refer to the minimum and maximum dose prescribed to the tumour where $D_{\text{min}}$ is the conventionally delivered dose and $D_{\text{max}}$ is the maximum boost dose.

These next-generation adaptive RT techniques represent a personalized and adaptive approach to RT. Due to dose constraints imposed by normal tissues, only a fi-
nite amount of radiation can be delivered to a patient. Therefore the overall goal is to use this radiation as efficiently as possible by exploiting variability in tumour radio-sensitivity. However, a key challenge to this approach is obtaining reliable measures of radio-sensitivity that have been linked to clinical tumour control outcomes. Novel imaging biomarkers are needed which indicate regional tumour sensitivity or predict which parts of the tumour are responding or progressing during treatment. Feasibility studies have focused on the use of various positron emission tomography (PET) image markers of hypoxia such as F18-fluorazomycin arabinoside (FAZA) and F18-fluromisonidazole (FMISO), or markers of cell proliferation such as C11-choline [42–47]. Hypoxia has been used as surrogate for local radio-sensitivity since low oxygen conditions inhibit radiation-induced production of free radicals thereby reducing tumour cell DNA damage and death. Alternatively, cell proliferation has been used as a surrogate for local tumour progression.

Next-generation adaptive RT techniques are resource intensive and require IMRT, image-guidance, functional imaging, image processing and non-commercially available treatment planning algorithms. To date, research into these methods has focused mainly on technical feasibility [42–47]. However a few dose-painting studies have been performed with prospective application to patients with head and neck cancer [48, 49]. Preliminary modelling and feasibility studies as reviewed by Bentzen et al. [41] predict gains of 50% or more in tumour control probability, suggesting that significant patient benefit could be derived from development of next-generation adaptive RT techniques. However, clinical trials are still needed to verify whether targeting resistant or proliferative tumour sub-regions actually provides the predicted benefit in practice.
1.6 Evaluation of treatment response and the need for radiotherapy response prediction

Next generation adaptive RT will be enabled by effective image-based treatment response prediction. Here we discuss current clinical methods for evaluating treatment response and motivate the need for treatment response prediction (a focus of this thesis) to support adaptive RT strategies.

1.6.1 Radiotherapy response evaluation in the clinic

Since RT is a localized treatment technique, evaluation of post-treatment changes in tumour size is a commonly used treatment response metric. Decreases in tumour size may indicate a favourable response and increases in tumour size may indicate tumour progression. The response evaluation criteria in solid tumours (RECIST) were initially proposed in 2000 and revised in 2009 to provide a standardized way of assessing and describing treatment response using changes in tumour size [50].

The RECIST criteria are comprehensive and include many special considerations. The main elements can be summarized as follows. Complete response is defined as the complete disappearance of all targeted lesions. Partial response is defined as at least a 30% decrease in the sum of the diameters of all of the target lesions with respect to pre-treatment baseline measurements. Progressive disease is defined as at least a 20% increase in the sum of the diameters of the target lesions, or by the appearance of new lesions. Stable disease is defined by changes that do not meet the criteria for complete response, partial response, or progressive disease. The tumour diameters are defined by the long-axis of each tumour. That is, the longest straight-line tumour diameters that can be measured are used to assess response. Figure 1.6 provides an example of RECIST measurement for a glioblastoma tumour.
Figure 1.6: Example RECIST measurements for a glioblastoma tumour on baseline and follow-up contrast-enhanced T1-weighted MRI scans. A 33% increase in tumour diameter identifies this patient as having progressive disease according to the RECIST criteria.

Blood-based measurements of specific proteins have also been used to evaluate treatment response and progression. For example, alpha-fetoprotein, carcinoembryonic antigen, and prostate-specific antigen levels are often measured for hepatocellular carcinoma, colorectal cancer, and prostate cancer respectively [51–55]. Decreased levels compared to pre-treatment measurements may indicate response and increased levels may indicate progression or recurrence.

1.6.2 The need for radiotherapy response prediction

The central idea behind next-generation adaptive RT is to adapt treatment to predictions of patient-specific tumour response to radiation. RECIST provides a useful framework for treatment response evaluation however it is not well-suited towards guiding these techniques. RECIST measurements indicate the trajectory of tumour size changes but can be confounded by factors such as radiation injury, edema and transient tumour size increases known as pseudo-progression [56, 57]. RECIST cut-off values are designed to document gross tumour changes (progression, partial response,
complete response), thus limiting descriptive and predictive utility particularly early on in treatment when such changes may not yet have occurred. Finally, and most importantly, RECIST does not indicate which parts of the tumour will respond or progress during or after treatment.

Ideally, techniques are needed which can predict tumour response before or during treatment. Early treatment response prediction could enable efficient adaptation without the need for additional treatment and risk of additional side-effects. On a rudimentary level, TCP modelling attempts to predict overall tumour response prior to treatment based on proposed prescription doses. However, next generation adaptive RT requires the ability to predict response throughout the tumour in order to help define the shape and intensity of radiation beams. Therefore image-based biomarkers of early treatment response are needed for guiding adaptive treatment. Such image-based biomarkers are identified through explicit correlations with treatment outcomes and a variety of different imaging parameters can be investigated for predictive potential. For example, it is well-known that hypoxia limits the effectiveness of RT [32], therefore PET-based hypoxia imaging has been used in previous adaptive RT studies as a direct surrogate for radio-resistance [42–44]. Image processing techniques can also be exploited to mine other types of image data for predictive trends that may not be as intuitive as hypoxia-imaging. Indeed, Bentzen, a key authority on dose-painting research has commented that [41]:

“Validation of a dose painting target does not necessarily require a mechanistic understanding of the relationship between dose response and expression of the target. It is a sufficient and from a bioethical standpoint probably also a necessary condition that an empirical relationship has been demonstrated between target expression and worse local outcome of radiation therapy.”
Here, Bentzen reasoned that prediction of treatment-resistant regions of tumours (i.e. dose painting targets) does not need to be verified through understanding the physiological cause of that resistance. Rather, it must be validated that predicted resistance leads to tumour recurrence in order to ethically justify the costs associated with the delivery of modified or additional radiation treatments to the patient. Therefore treatment response prediction via image-analysis techniques such as those presented in this thesis is well-suited for defining dose-painting targets. Such image analysis driven techniques can yield predictive trends that may not have a known mechanistic explanation. However, the associated predictions are data driven and so efficacy is demonstrated through verifying empirical relationships with treatment outcome.

In summary, image-based biomarkers of early treatment response are a necessary and critical component in the development of next-generation adaptive RT strategies. This is a central concept within this thesis and motivates the investigations in Chapters 3 and 4 of this thesis. In the next section we discuss some of the fundamentals of image based treatment response prediction including image analysis and imaging methods.

1.7 Image-based treatment response prediction

An image allows us to visualize a two-dimensional array of numbers corresponding to some physical property that has been measured by an imaging device. For example, each pixel within a functional image can depict “blood volume” representing the volume of blood supplied (millilitres per unit mass) within that pixel. Each number in the corresponding image array is represented by a different shade of gray or colour within the image permitting visualization of higher and lower blood volume regions within the image. Figure 1.7 provides an artificial example of a blood volume image
Figure 1.7: Artificial blood volume functional image with its corresponding image array. The image array depicts the volume of blood estimated within each pixel and is expressed in millilitres of blood per 100 grams of tissue.

along with its corresponding image array.

Image-based treatment response prediction involves analyzing the arrays of numbers that underlie images. For example, baseline functional images acquired before treatment can be analyzed or comparisons can be made between images acquired before and after treatment. 3-dimensional image volumes can also be analyzed and are defined by 3-dimensional arrays of numbers with each value in the array representing a quantity measured for a small volume of tissue (voxel).

Many different types of imaging have been analyzed for biomarkers of treatment response using a wide spectrum of image analysis techniques. These techniques are often customized to the specific image data being analyzed and can range from direct utilization of image intensity values to more complex methods involving machine learning [57–59]. Consequently, a comprehensive review is beyond the scope of this thesis. However this section will provide a brief overview of the image analysis concepts involved in image-based treatment response prediction. Imaging methods which are actively being investigated for predictive utility will also be discussed. The following discussion is divided into treatment response prediction via region-based and voxel-based image analyses, both of which are employed within this thesis.
1.7.1 Region based image analysis

Region-based analysis involves summarizing the image intensity values (image array numbers) within a region of interest prior to searching for correlations with treatment response. For example, the average image intensity within a tumour could be computed for multiple patients. These values could then be found to be significantly different between responding and non-responding patients and subsequently used to predict outcome. Other descriptive summary statistics could also be investigated such as the standard deviation, median, or inter-quartile range of the tumour intensity values. Figure 1.8 provides an example of a region-based analysis of the artificial blood volume image shown in figure 1.7.

A common approach involves acquisition of multiple images of the patient at different times (i.e. longitudinal imaging). Differences in tumour summary statistics can then be investigated for correlations with outcome. For example images can be acquired before and after treatment and then the change in the average image intensities within the tumour can be determined. Increases or decreases in average image intensity could then be found to relate to treatment response or tumour progression.

As demonstrated in figure 1.8, a potential limitation of region-based image analysis is that global summary measures can be confounded by image intensity heterogeneity within the tumour. For example, blood volume increases in one region of the tumour could be offset by decreases in another region resulting in no change in average tumour blood volume. For the purposes of prediction, such a patient could then be considered to be the same as a patient with no change in tumour blood volume. In such examples, differences in variability measures (e.g. standard deviation) may be more instructive.

Region-based methods can only be used to predict global measures of treatment response such as overall survival and cannot be used to predict which parts of the tumour are responding or progressing. Only global measures which summarize the
Figure 1.8: Example of a region-based image analysis of two longitudinally acquired blood volume images. The parameters $\mu$ and $\lambda$ represent the mean and standard deviation of the image intensity values within the red regions of interest respectively.

detailed voxel-wise image intensity changes within the tumour (e.g. mean values) are available for prediction. These single-valued measures do not specify the image intensity characteristics or changes of each tumour voxel and therefore cannot be used to predict differential response between different parts of the tumour. The focus of the next section is voxel-based image analysis which has potential for predicting differential response within tumours.

1.7.2 Voxel based image analysis

Voxel-based image analysis takes image heterogeneity into account and can therefore be more sensitive to changes than region-based methods. However, voxel-based methods are more complicated and can be influenced by additional sources of uncertainty.

One common approach towards voxel-based image analysis involves acquisition
of longitudinal imaging for each patient which is then spatially aligned using image registration algorithms [60]. After alignment, voxel values in one image can be compared to the intensity of the spatially corresponding voxels in another image. With this approach, localized intra-tumour changes in image intensity can be interrogated for markers of treatment response. However, residual image registration error (alignment error) introduces variability into the analysis [61]. In the presence of image registration error, voxels in one image may be incorrectly associated with voxels in the neighbourhood around the spatially corresponding voxels in the other images. Depending on the heterogeneity in the images, neighbouring voxels may or may not have intensities that are similar to the spatially corresponding voxels and therefore image registration error may introduce errors into subsequent analyses. Figure 1.9 provides a simple example of voxel-based image analysis without image registration error.

More sophisticated analyses can also be performed which involve investigating image texture in one or more images. Many different image texture features can be calculated [62]. Each feature describes a different aspect of the relationship between the intensities of neighbouring voxels and therefore provides a measure of the image heterogeneity. This information can then be fed into machine learning algorithms which build classifiers for treatment response or tumour progression via pattern recognition [58, 63]. The central idea is that progressing or responding tumours may exhibit characteristic image textures which, if identified, can be used to identify patients which may not be responding to treatment.

Voxel-based image analysis is necessary in order to predict which regions within a tumour are responding and which regions are progressing (i.e. voxel-wise treatment response). However to date, these methods have almost exclusively been applied to improve prediction of global measures of treatment response (e.g. overall survival,
Figure 1.9: Example of voxel-based image analysis where patient motion (10 degree rotation) must be corrected using image registration.
RECIST criteria). Current voxel-wise treatment response prediction studies have focused on the use of imaging to define input parameters for mathematical models of tumour growth [64–66]. Therefore prediction of voxel-wise treatment response via direct analysis of imaging represents largely uncharted territory in current treatment response prediction literature.

1.7.3 Imaging

While prediction of voxel-wise treatment response has yet to be pursued, many different types of imaging have been investigated for biomarkers of treatment response of the entire tumour [67]. Typically, imaging which measures some aspect of tumour function or of the tumour micro-environment is analyzed. Changes in these measurements can then be monitored and analyzed over time to predict treatment response.

Diffusion and perfusion-weighted magnetic resonance imaging (MRI) are commonly investigated for predictive potential [67, 68] and are also analyzed in Chapter 4 of this thesis. Diffusion-weighted imaging (DWI) measures the ability of water molecules to diffuse in each voxel of tissue and provides a surrogate measure of local cellular density. Tumours tend to exhibit restricted diffusion due to increased cellular density and so increases in diffusion after treatment could indicate response and decreases could indicate progression. Perfusion-weighted imaging (PWI) provides a measure of hemodynamic quantities such as blood flow and blood volume within each voxel. Tumours may exhibit elevated blood volume compared to normal tissue due to their ability to stimulate angiogenesis (growth of new vasculature) to support rapid growth. Therefore increases in blood volume after treatment could indicate tumour progression and decreases could indicate response. Figure 1.10 provides an example of apparent diffusion coefficient (ADC) and relative cerebral blood volume (rCBV)
maps derived from DWI and PWI imaging respectively. rCBV maps are expressed in relative units with respect to the blood volume measured in healthy white matter. Dynamic contrast-enhanced ultrasound (DCE-US) and CT (DCE-CT) can also be used to measure hemodynamic quantities and have been similarly investigated in treatment response prediction studies [67, 69]. In this thesis we used blood volume (BV) images derived from DCE-CT in Chapter 3 and investigated response prediction using MRI-derived ADC and rCBV images in Chapter 4.

Figure 1.10: Example ADC and rCBV maps derived from diffusion and perfusion weighted MR imaging acquired from a patient with glioblastoma at 1 and 3 months post-radiotherapy. Contrast-enhanced T1-weighted images are also shown to indicate the location of the tumour. ADC maps are expressed in units of $10^{-3}$ mm$^2$/s and rCBV maps are expressed in relative units with respect to the blood volume in white matter.
1.8 Summary

Cancer is heterogeneous and adaptable due to its characteristic genetic instability and unrestrained growth. Genetic variability exists between different tumour types, tumours of the same type and within individual tumours. This heterogeneity in turn leads to variability in treatment efficacy not only between patients but potentially within different sub-regions of individual tumours. Therefore, personalized treatment strategies which can evaluate and adapt to heterogeneity in tumour response may be a key towards improving tumour control and reducing treatment-related side-effects.

RT is a commonly used cancer treatment technique which uses ionizing radiation. RT is well-suited towards adaptive implementations since intensity modulation and image guidance can enable localized modification of treatment plans within the tumour region (e.g. sub-volume boosting). RT is also delivered in multiple fractions providing multiple opportunities to evaluate and adapt to variable tumour response. Consequently, personalized adaptive RT is an active area of research which may have potential for improving outcomes. Development of effective early treatment response prediction constitutes a critical step towards implementation of personalized cancer treatment strategies such as adaptive RT. In particular, image-based treatment response evaluation methods which can predict voxel-wise tumour response are needed to guide next generation adaptive RT techniques such as sub-volume boosting and dose-painting by numbers.
1.9 Thesis objectives

There is a strong theoretical basis for improving patient outcomes via the development of personalized adaptive RT strategies. However, implementation of these strategies directly depends on effective treatment response prediction software tools. The overall goal of this thesis is to contribute to the development of treatment response prediction methods to support future implementation of adaptive RT techniques. There are three main objectives which correspond to Chapters 2, 3, and 4 respectively.

Objective 1 - Investigate and model the relationship between the probability of tumour control and prescribed radiation dose for patients with liver tumours.

Objective 2 - Augment an image-based treatment response prediction method to incorporate uncertainties due to unavoidable image registration error and to facilitate identification of adaptive RT targets.

Objective 3 Develop a generalized image-based treatment response prediction framework which enables improved analysis of multi-parametric image data.
References


Chapter 2

Determination and comparison of radiotherapy dose responses for hepatocellular carcinoma and metastatic colorectal liver tumours *

2.1 Introduction

Hepatocellular carcinoma (HCC) is the fourth most common cancer worldwide [1]. The liver is also the most common site of metastasis from colorectal carcinoma, with the incidence of liver metastases exceeding 25% [2, 3]. Although transplantation and surgical resection offer significant survival benefits among these two groups [4–7], impaired liver function, tumour size and the number of lesions can limit a patient’s eligibility for these treatments. Consequently, there has been a push to develop alternative locoregional therapies for treatment of primary and metastatic liver cancer.

Stereotactic body radiotherapy (SBRT) has shown promise as a new method to

*This chapter has been previously published as Lausch et al. Determination and comparison of radiotherapy dose responses for hepatocellular carcinoma and metastatic colorectal liver tumours. Br. J. Radiol., 2013;85:20130147.
safely and non-invasively treat liver tumours [8–11]. SBRT involves precise image-guided delivery of a high dose of radiation to the tumour in a small number of fractions and usually employs motion suppression or gating techniques. A positive association between outcome and dose has previously been reported [11–13]. However, unlike other tumour sites [14–18], there is currently a paucity of explicit liver tumour dose-response modelling within the literature [11]. An improved understanding of radiation dose response is necessary to help better inform future dose prescriptions.

At our institution, patients with liver lesions are prescribed the highest possible radiation dose while maintaining < 5% normal tissue complication probability (NTCP) with respect to an end point of ≥ grade 3 Radiation Therapy Oncology Group (RTOG) radiation-induced liver disease (RILD) [19]. As a result, patients are treated with different doses depending on their tumour size relative to the size of remaining healthy liver. This provides for a valuable opportunity to investigate the dose response and to determine the radiation dose required to control liver tumours.

Here, our primary goal is to retrospectively determine separate dose-response relationships for patients with HCC and metastatic (MET) colorectal liver tumours using tumour control probability (TCP) modelling. Our secondary goal is to evaluate local control rates between these two groups at our institution.

### 2.2 Patients and Methods

#### 2.2.1 Patient data

The records of patients treated with radiotherapy (RT) at our institution for HCC or MET colorectal liver tumours from 2004 to 2011 were reviewed. Patients who had previous regional or systemic therapy were included in the analysis as long as their treatment concluded prior to the start of RT. Patients who underwent additional
concurrent therapy were excluded from the study. Patients who received alternative therapy, post RT, were included; however, their follow-up data were censored on initiation of the additional therapy. No limit was placed on the size or the number of target lesions. A total of 36 patients treated for HCC and 26 patients treated for MET colorectal liver tumours were analysed. Follow-up data typically included CT-based measurements of tumour size and measurements of α-fetoprotein (AFP) and carcino-embryonic antigen (CEA) biomarkers. The times of all follow-up observations (i.e. follow-up schedule) were reported with respect to the treatment end date and were typically at 1 month, 3 months, 6 months, 1 year and 2 years after treatment. However, patient-specific scheduling challenges introduced variability into actual follow-up dates and some tumour response measurements were assessed in between the aforementioned time points. Patients’ data were collected, quality assured and analysed in a database approved by our Institutional Review Board (Health Sciences REB#: 16487E).

2.2.2 Radiotherapy

The clinical target volume (CTV) was defined based on four dimensional CT scans. Motion management involved patient immobilisation (Vac-Lok; CIVCO Medical Systems, Orange City, IA) and respiratory gating (Varian Real-Time Position Management; Varian Medical Systems, Inc., Palo Alto, CA) to minimise internal motion. Depending on patient motion and residual motion measurements, an additional 2-10-mm, 3-25-mm and 3-7-mm planning target volume (PTV) margin was added to the CTV in the anterior-posterior, superior-inferior and lateral directions, respectively. Patients were treated with three-dimensional conformal RT, intensity-modulated radiotherapy or tomotherapy.

Diverse dose fractionation regimens were employed in the two patient groups,
including both traditional and hypofractionation schedules. The dose prescribed to each patient was maximised while maintaining < 5% Lyman-Kutcher NTCP for the remaining healthy liver with respect to an end point of grade 3 or higher RTOG RILD [19], and while maintaining the dose to other organs at risk to acceptable levels. The NTCP model employed parameter values reported by Dawson et al. [19], who had separate parameter sets for calculating the probability of liver toxicity between HCC and MET patient groups. Doses were prescribed to the PTV, such that 95% of the PTV received at least 95% of the prescribed dose. The maximum PTV dose was typically limited to no more than 107% of the prescribed dose. Consequently, the PTV mean dose was similar to the prescribed dose. The doses received by the small bowel, lung, heart, stomach, kidneys and spinal cord were also constrained to prevent toxicity.

2.2.3 Tumour control probability

A TCP model was used to investigate the relationship between the radiation dose and the tumour control. 6-month local control was chosen as the end point for the TCP model as detailed 6-month local control data were available for the largest number of patients and were considered clinically important a priori. Patients who had failed locally or died owing to liver disease progression prior to the 6-month time point were considered to be uncontrolled at 6 months. Since the patients included in this study were treated with diverse dose-fractionation regimens, prescription doses were converted to 2 Gy per fraction equivalent doses prior to TCP modelling to ensure biological comparability. Equivalent doses were computed using the standard linear quadratic model (LQM) approach [20].

Briefly, the LQM defines the number of radiation-induced cell-lethal events (i.e. double-strand DNA breaks) due to a single radiation dose $d$ by $n = \alpha d + \beta d^2$. The
number of lethal events among a population of cells (i.e. within a tumour) is proposed to follow a Poisson distribution. The surviving fraction \(SF\) of tumour cells within a tumour after an applied dose \(d\) is then given by \(SF = e^{-(\alpha d + \beta d^2)}\). Parameters \(\alpha\) and \(\beta\) are determined empirically through fitting the \(SF\) equation to cell-survival curves obtained through irradiating in-vitro tumour cell cultures with different doses of radiation followed by counting the surviving fraction of cells after each irradiation.

After administering \(N\) fractions of radiation, each with dose \(d\), the surviving fraction of tumour cells is given by \(SF = e^{-N(\alpha d + \beta d^2)}\). Noting that the total prescription dose is given by \(D = Nd\), this equation can be re-expressed as \(SF = e^{-\alpha D\left(1 + \frac{d_1}{\alpha/\beta}\right)}\).

Let \(D_1\), \(d_1\), and \(SF_1\) represent the prescribed dose, dose/fraction, and the \(SF\) after dose \(D_1\) is administered to the tumour respectively. The 2 Gy/fraction equivalent or iso-effective dose \((D_2)\) can then be obtained by equating \(SF_1\) to \(SF_2\) where \(SF_2\) is the theoretical \(SF\) after administering an equivalent prescription dose \(D_2\) to the tumour in 2 Gy fractions \((d_2 = 2\ \text{Gy})\).

\[
D_2 = \left(\frac{d_1 + \alpha/\beta}{d_2 + \alpha/\beta}\right)D_1
\]

Currently, there is no strong consensus on the LQM parameter \(\alpha/\beta\) for liver tumours with published values ranging from 3.1 Gy to 15 Gy [21]. Consequently, a typical tumour value of \(\alpha/\beta = 10\ \text{Gy}\) was used for this conversion. A sensitivity analysis was also performed to assess the impact of using other \(\alpha/\beta\) values on dose-response parameters. A uniform dose distribution within the target was assumed owing to the dose uniformity constraints described in §2.2.2.

To facilitate comparisons with Chang et al. [11], we elected to use the same logistic TCP model [22] to fit the 6-month tumour control data:
\[ TCP = \frac{1}{1 + (D_{50}/D)^k} \tag{2.2} \]

where \( D_{50} \) is the dose that would result in a 50% probability of local control, \( D \) is the prescription dose and \( k \) controls the slope of the TCP curve. Similar to other response studies [23, 24], the model was fitted to the data using the maximum likelihood method, which can provide both the fitted parameters and the estimates of their standard deviations [25].

In our context, the maximum likelihood method determines the values of \( D_{50} \) and \( k \), which maximise the probability of predicting response within our patient groups using the TCP from equation 2.2. We employed a binary relationship where patients either respond or do not respond to treatment, which we denote for the \( i \)-th patient by \( R_i = 1 \) or \( R_i = 0 \), respectively. Here, response corresponds to observing local control approximately 6 months (±1 month) post RT. The probability that patient \( i \) responds or does not respond given a prescribed dose \( D_i \) is given by

\[ f(D_i, R_i) = p_i^{R_i}(1 - p_i)^{1-R_i} \tag{2.3} \]

where \( p_i \) is equal to the TCP defined by equation 2.2, evaluated at \( D = D_i \). A constrained minimisation procedure from the MATLAB optimisation toolkit (fmincon) (The Mathworks, Inc., Natick, MA) was used to search for the values of \( D_{50} \) and \( k \), which minimise the negative of the log-likelihood function

\[ l = \sum_{i=1}^{N} \log[f(D_i, R_i)] \tag{2.4} \]

where \( N \) is the number of patients in the patient group of interest and \( f(D_i, R_i) \) is defined by equation 2.3.
2.2.4 Local control definition

Tumour response was evaluated in the largest treated lesion in each patient. Local control was defined using radiographic or tumour biomarker information depending on data availability. Using radiographic information, local control corresponded to observing at least stable disease criteria or better using the Response Evaluation Criteria in Solid Tumours v. 1.1 [26]. Tumour diameter measurements used to evaluate RECIST criteria were obtained from radiologist reports. Using biomarker information, local control corresponded to observing a $\leq 20\%$ increase from baseline measurements of AFP for HCC patients and CEA for MET patients.

AFP has recently been shown to be a reliable biomarker of radiological response after locoregional therapy for HCC [27, 28]. CEA has also previously been used as a biomarker for response and metastasis after surgical resection and chemotherapy for colorectal carcinoma [29, 30]. Biomarker levels were required to exceed the normal values either prior to or after RT in order for them to be included in the definition of local control. The normal values used for AFP and CEA levels were 6 ng ml$^{-1}$ and 5 ng ml$^{-1}$, respectively [31, 32]. If a treated lesion met the criteria for radiographic local control but not biomarker local control (or vice versa), then the lesion was still considered to be locally controlled. If a patient did not have follow-up CT data, then local control was assessed using biomarker data and vice versa. Data were censored starting at the time of the last follow-up if a patient was lost to follow-up prior to the loss of local control.

2.2.5 Local control analysis

Local control follow-up data were visualised using what we have termed dose-control history plots. These plots contain horizontal timelines for each patient, which indicate
follow-up history and outcomes. The timelines are displayed in order of increasing prescription dose. Standard Kaplan-Meier analyses and log-rank tests were also performed using SPSS Statistics v. 19 (IBM Corporation, Armonk, NY) to quantitatively assess local control as a function of time.

Potential factors affecting local control, such as the radiation dose and pre-treatment tumour volume, were explored using plots of patient pre-treatment tumour volume vs the prescribed equivalent dose, with colour coding indicating time to loss of local control or time to censor (last follow-up before loss of patient contact). Two-tailed Spearman rank tests were used to test for possible correlations. Censored data (patients lost to follow-up before loss of local control) were included in correlative analyses, provided the data were censored at least 6 months post RT. We expect that RT had an effect if the tumour was controlled for at least 6 months, and, therefore, that time to last follow-up for censored data may be correlated with the radiation dose or pre-treatment tumour volume.

2.3 Results

Patient and treatment characteristics for the full HCC and MET groups included in the local control analysis are summarised in table 2.1. The median follow-up time was 197 days (27-1095 days) and 178 days (51-1101 days) for the HCC and MET groups, respectively.

Sufficient follow-up data were available to assess 6-month local control for 27/36 (75%) HCC and 19/26 (73%) MET cases. 5/27 HCC and 1/19 MET patients died prior to true observation of loss of local control. However, in all six cases, patient charts indicated that there was no extrahepatic progression and that liver disease progression was the cause of death. Therefore, these tumours were considered to be
Table 2.1: Patient characteristics and treatment data. All summary statistics are medians with ranges displayed in brackets. p-values were calculated using the Wilcoxon rank sum test. *Abbreviations*: 3D-CRT, three-dimensional radiotherapy; CTV, clinical target volume; HCC, hepatocellular carcinoma; IMRT, intensity-modulated radiotherapy; MET, metastatic colorectal liver tumours; RFA, radiofrequency ablation; TACE, transarterial chemoembolisation.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>HCC (n = 36)</th>
<th>MET (n = 26)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>74 (22-87)</td>
<td>68 (42-90)</td>
<td>0.06</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male (%)</td>
<td>29 (81)</td>
<td>16 (62)</td>
<td></td>
</tr>
<tr>
<td>Female (%)</td>
<td>7 (19)</td>
<td>10 (38)</td>
<td></td>
</tr>
<tr>
<td>Child-Pugh score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class A (%)</td>
<td>27 (75)</td>
<td>21 (81)</td>
<td></td>
</tr>
<tr>
<td>Class B (%)</td>
<td>8 (22)</td>
<td>5 (19)</td>
<td></td>
</tr>
<tr>
<td>Class C (%)</td>
<td>1 (3)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Previous treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surgical resection (%)</td>
<td>0</td>
<td>8 (31)</td>
<td></td>
</tr>
<tr>
<td>Chemotherapy (%)</td>
<td>5 (14)</td>
<td>23 (88)</td>
<td></td>
</tr>
<tr>
<td>TACE (%)</td>
<td>9 (25)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>RFA (%)</td>
<td>1 (3)</td>
<td>1 (4)</td>
<td></td>
</tr>
<tr>
<td>Number of liver lesions</td>
<td>1 (1-4)</td>
<td>2 (1-5)</td>
<td></td>
</tr>
<tr>
<td>Active extrahepatic disease (%)</td>
<td>11 (31)</td>
<td>12 (46)</td>
<td></td>
</tr>
<tr>
<td>Tumour diameter (%)</td>
<td>6.6 (3.0-18.0)</td>
<td>5.0 (1.0-13.0)</td>
<td>0.03</td>
</tr>
<tr>
<td>CTV (ml)</td>
<td>186 (8-995)</td>
<td>57 (5-1804)</td>
<td>0.02</td>
</tr>
<tr>
<td>Radiotherapy technique</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D-CRT (%)</td>
<td>25 (70)</td>
<td>21 (81)</td>
<td></td>
</tr>
<tr>
<td>IMRT (%)</td>
<td>8 (22)</td>
<td>4 (15)</td>
<td></td>
</tr>
<tr>
<td>Tomotherapy (%)</td>
<td>3 (8)</td>
<td>1 (4)</td>
<td></td>
</tr>
<tr>
<td>Gating (%)</td>
<td>21 (58)</td>
<td>15 (58)</td>
<td></td>
</tr>
<tr>
<td>Number of fractions</td>
<td>15 (6-20)</td>
<td>15 (6-21)</td>
<td>0.37</td>
</tr>
<tr>
<td>Dose per fraction (Gy)</td>
<td>4.0 (2.0-10.0)</td>
<td>3.6 (2.0-13.0)</td>
<td>0.24</td>
</tr>
<tr>
<td>Total dose (Gy)</td>
<td>52 (29-83)</td>
<td>55 (30-80)</td>
<td>0.78</td>
</tr>
<tr>
<td>Equivalent dose (Gy)</td>
<td>63 (33-107)</td>
<td>61 (33-154)</td>
<td>0.88</td>
</tr>
</tbody>
</table>
uncontrolled at 6 months. 16/27 (59%) and 7/19 (32%) patients were locally controlled at approximately 6 months post RT in the HCC and MET groups, respectively.

The logistic TCP model was fitted independently to the response data for each subgroup using the end point of 6 months of local control (Figure 2.1). One outlier within the HCC group was omitted prior to model fitting (107 Gy, uncontrolled at 6 months) owing to its severe contradiction with the trends indicated by the other data. For reference purposes, the observed patient response data used by the maximum likelihood fitting method are also indicated in Figure 2.1 as a function of equivalent dose. These data were then binned and averaged to generate estimates of the observed TCP at various dose levels and to assist in evaluating the quality of the model fits. Horizontal whiskers indicate the range of doses included in each bin.

$D_{50}$ was determined to be 53 Gy ($\sigma = 5.6$ Gy, 95% CI = 51-55 Gy) and 70 Gy ($\sigma = 6.6$ Gy, 95% CI = 67-73 Gy), and the slope parameter $k$ was estimated to be 4.8 ($\sigma = 2.0$, 95% CI = 1.2-2.8) and 7.1 ($\sigma = 3.3$, 95% CI = 5.6-8.6) for the HCC and MET groups, respectively. 2 Gy per fraction equivalent doses of 84 Gy and 95 Gy are predicted to result in 90% 6-month local control rates for patients with HCC and MET colorectal liver tumour, respectively. Uncertainty in the dose-response curves was illustrated in Figure 2.1 by plotting the TCP model using the fitted $D_{50}$ parameters plus or minus two standard deviations while keeping the fitted $k$ parameter constant. The standard deviations of the parameter estimates serve as surrogates for assessing the quality of the model fits, with higher values indicating greater parameter uncertainty.

General treatment-related toxicities (e.g. nausea, abdominal pain, fatigue) were scored using the Common Terminology Criteria for Adverse Events v. 3.0 (National Cancer Institute, Bethesda, MD) for the subset of patients included in TCP analysis. We note that both HCC and MET patient groups were retrospectively determined
Figure 2.1: Tumour control probability curves and patient response data for the (a) hepatocellular carcinoma and (b) metastatic colorectal liver tumours patient groups approximately 6 months post treatment. The uncertainty estimates indicate the TCP model plotted using the fitted $D_{50}$ parameters plus or minus two standard deviations while keeping the fitted $k$ parameters constant.
within this study. For the HCC group after 6 months, there were 5 asymptomatic patients, 9 patients with Grade 1, 10 patients with Grade 2 and 2 patients with Grade 3 complications. For the MET group after 6 months, there were five asymptomatic patients, three patients with Grade 1, eight patients with Grade 2 and three patients with Grade 3 complications. Liver specific toxicities were scored using the same RILD end point as the NTCP model used during planning [19]. No instances of Grade 3 or higher Radiation Therapy Oncology Group RILD were observed among any of the HCC or MET patients.

Figure 2.2 summarizes the local control data acquired for the two patient groups using dose-control history plots. The plots indicate all the radiographic and biomarker local control data that were available for each HCC and MET patient. Between the full HCC (n = 36) and MET (n = 26) cohorts, local control was determined by radiographic data for 21/36 (58%) HCC and 13/26 (50%) MET patients at the time of last follow-up, with local control determined by the AFP and CEA data for the remaining patients. For HCC, only 12/36 patients had both radiographic and biomarker data at the time of last follow-up and both measures were in agreement for 7/12 patients. For MET, only 5/26 patients had both radiographic and biomarker data at the time of last follow-up and both measures were in agreement for 4/5 patients. The median follow-up time was 197 days (range, 27-1095 days) and 178 days (range, 51-1101 days) for the HCC and MET groups, respectively. At the time of last follow-up, 21/36 (58%) HCC and 14/26 (54%) MET patients had lost local control.

Kaplan-Meier analysis was performed on the local control data for the full HCC (n = 36) and MET (n = 26) cohorts (Figure 2.3). The HCC and MET curves were found to be significantly different (log-rank $p = 0.03$). For the HCC group, actuarial 1- and 2-year local control rates were 65% (45-85%) and 48% (23-73%), respectively.
Figure 2.2: Dose-control histories for the (a) hepatocellular carcinoma and (b) metastatic colorectal liver tumours patient groups. Horizontal lines correspond to individual patient histories and are displayed in the increasing order of the radiation treatment dose. Solid and dotted lines are alternated to facilitate visualization of adjacent histories.
Figure 2.3: Local control Kaplan-Meier curves for the (a) hepatocellular carcinoma and (b) metastatic colorectal liver tumours patient groups (log-rank p = 0.03). CI, confidence interval.

For the MET group, actuarial 1- and 2-year local control rates were 32% (6-58%) and 0% (0-42%), respectively. The median time to failure (loss of local control) between HCC and MET groups was estimated to be 543 days [95% confidence interval (CI) 374-711] and 183 days (95% CI 72-294), respectively. To visualise how the data are distributed about the $D_{50}$ parameters, plots of pre-treatment tumour volume vs equivalent prescribed dose with time to failure/censor colour coding were generated (Figure 2.4). Data point radii were also related to time to failure/censor for improved visualisation.

We performed two-tailed Spearman’s rank correlation tests on the subset of data included in Figure 2.4c (n = 27/36) and 2.4d (n = 19/26). Dose was significantly correlated with tumour volume for both the HCC ($\rho = -0.73, p \leq 0.001$) and the MET groups ($\rho = -0.62, p = 0.005$). However, baseline tumour volume was not significantly correlated with time to failure/censor for either of the HCC ($\rho = -0.047, p = 0.82$) or of the MET groups ($\rho = -0.44, p = 0.06$). Dose and time to fail-
ure/censor were significantly correlated for the MET group ($\rho = 0.50, p = 0.03$) but not for the HCC group ($\rho = 0.22, p = 0.26$).

Figure 2.4: Pre-treatment clinical target volume (CTV) vs the prescribed equivalent dose with time to loss of local control/censor colour coding. (a) and (b) Plots for the hepatocellular carcinoma and metastatic colorectal liver tumour patient groups, respectively. (c) and (d) contain the same information except we have omitted the data censored prior to 6 months. Outliers (35 Gy, 1804 ml, 0.3 years), (33 Gy, 1122 ml, 0.14 years) and (154 Gy, 7 ml, 0.14 years) have been omitted from (b) to better visualise trends. $D_{50}$ is indicated by the dashed vertical line.
### 2.4 Discussion

In this study, we have demonstrated 6-month dose-response relationships for patients treated with RT for HCC or MET colorectal liver tumour groups. The heterogeneous doses prescribed to the patients within the HCC and MET groups provided a valuable opportunity to evaluate dose response in the liver. Although dose-response relationships have previously been established for HCC [13] and MET [11] patients, currently, there is a lack of literature on dose response and explicit TCP modelling for the liver. These data are critical if personalised radiobiologically guided dose escalated RT is to be applied to patients. This study employed the maximum likelihood method to fit a TCP model to HCC and MET tumour response data, adding to this body of literature. To our knowledge, this is the first study to explicitly model TCP for HCC patients and the second to do so for MET patients [11].

Park et al. [13] reported a 77% partial response rate, 4-6 weeks post RT, among a subset of patients with HCC (n = 83) who were prescribed a radiation dose ≥50 Gy. In the present study, we found that a 50% response rate at 6 months post RT could be achievable with a higher dose of 53 Gy (σ = 5.6 Gy). This suggests that dose escalation may improve the durability (duration of local control) of tumour response which is consistent with the findings of Park et al. [13].

Chang et al. [11] reported a $D_{50}$ value of 68 Gy (standard error = 6 Gy) with an end point of 12 months of local control among patients (n = 65) treated with SBRT for MET colorectal liver tumours in a multicentre pooled analysis. Here, we found a slightly higher $D_{50}$ value of 70 Gy (standard error = 1.5 Gy) for an end point of 6 months of local control among a similar group of patients, although this difference is within the uncertainties of the two $D_{50}$ estimates. If we had similar patient demographics, we would expect that higher doses should translate to improved
local tumour response durability. Therefore, a 12-month response $D_{50}$ value as in Chang et al. [11] would be expected to be higher than a 6-month response value. However, an important difference between the patient population in Chang et al. [11] and the MET group is the median tumour volume, which is smaller in Chang et al. [11] than in the present study [30 cc, range = (0.5, 3008) vs 57 cc, range = (5, 1804)]. A steeper dose-response relationship was found in our study with a slope parameter of 7.1 (standard error = 0.8) compared with 4.2 (standard error = 1.6) as reported by Chang et al although this difference is within the uncertainties of the two slope parameter estimates.

Prior to TCP modelling, the LQM was used to convert doses to 2 Gy per fraction equivalents. The LQM has been shown to overestimate cell kill for dose fractionations beyond 8-10 Gy per fraction [33]. Although some patients within this study received higher ablative doses, 96% of the HCC patients and 95% of the MET patients included in the TCP analyses received $\leq$8 Gy per fraction, supporting our use of the LQM. None of the included patients received doses exceeding 10 Gy per fraction.

Currently, there is no strong consensus on $\alpha/\beta$ ratios for HCC, and, to our knowledge, there are no published data on ratios for MET colorectal liver tumours. Consequently, a single $\alpha/\beta$ ratio of 10 Gy was used to determine the 2 Gy per fraction equivalent doses for both the HCC and the MET groups prior to dose-response modelling. Therefore, we investigated the sensitivity of our $D_{50}$ and $k$ dose-response parameter values to the $\alpha/\beta$ ratio. Wigg et al. [21] have compiled a short list of reported HCC $\alpha/\beta$ ratios. We used these alternative values to recompute the 2 Gy per fraction equivalent doses and then recalculate the dose-response parameter values (Table 2.2). The parameter variability for values of $\alpha/\beta \geq 7.2$ Gy was much smaller than the uncertainty (standard deviation) in the parameter estimation itself. Therefore, our reported values should be robust to current $\alpha/\beta$ uncertainty provided that
Table 2.2: Influence of using different $\alpha/\beta$ ratios on estimated dose-response parameters. The prescription doses were converted to 2 Gy per fraction equivalent doses using each $\alpha/\beta$ ratio followed by tumour control probability model fitting. Parameter results for $\alpha/\beta = 10$ Gy were included for comparison. Standard deviations are shown in brackets. **Abbreviations:** $D_{50}$, dose that would result in a 50% probability of local control; $k$, slope parameter; HCC, hepatocellular carcinoma; MET, metastatic colorectal liver tumours.

<table>
<thead>
<tr>
<th>$\alpha/\beta$ (Gy)</th>
<th>HCC</th>
<th>MET</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$D_{50}$ (Gy)</td>
<td>$k$</td>
<td>$D_{50}$ (Gy)</td>
<td>$k$</td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>63 (6.7)</td>
<td>4.8 (2.0)</td>
<td>80 (8.4)</td>
<td>6.1 (2.7)</td>
<td></td>
</tr>
<tr>
<td>7.2</td>
<td>55 (5.8)</td>
<td>4.9 (2.0)</td>
<td>72 (7.2)</td>
<td>6.8 (3.1)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>53 (5.6)</td>
<td>4.8 (2.0)</td>
<td>70 (6.6)</td>
<td>7.1 (3.3)</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>51 (5.6)</td>
<td>4.6 (2.0)</td>
<td>67 (6.0)</td>
<td>7.4 (3.5)</td>
<td></td>
</tr>
</tbody>
</table>

Liver tumours are early responding tissues.

In this study, our primary aim was to examine whether dose-response relationships exist for HCC and MET patients. Since limited research has been done on liver tumour dose response, we employed the commonly used logistic TCP model to fit the observed data. This provided a simple way to demonstrate dose response in the liver and facilitated comparison with the only other existing literature. In the future, more refined TCP models will be investigated, which account for non uniform tumour doses and tumour volumes. Furthermore, the predictive utility of the models will then be investigated in independent data sets. Toxicity among the patients included in the TCP analyses was comparable to previous liver irradiation studies [10, 11, 34–36]. The large tumour sizes in the HCC and MET patient groups may have contributed to marginally higher toxicities. However, HCC patients who experienced Grade 3 complications post RT had pre-existing Grade 2 complications before treatment. Similarly, for the MET group, one patient with post-RT Grade 3 complications had pre-RT Grade 3 complications and the remaining two had pre-RT Grade
In addition, no instances of Grade 3 or higher Radiation Therapy Oncology Group RILD were observed among either of the two groups.

We found that the dose and time to loss of local control/censor were correlated for the MET patients but not for the HCC patients. This could be because of the increased percentage of patients in the HCC group who were lost during follow-up prior to observation of loss of local control (18/27, 67% for HCC vs 5/19, 26% for MET). HCC patients tend to lose local control later than MET patients, and therefore patients are more frequently lost to follow-up prior to observation of loss of local control. Although data censored at later times may have a connection to treatment parameters, inclusion of these data in the analysis may weaken correlations in the HCC group since time to last followup (or censor) can be related to factors aside from the dose. Tumour volume was not found to be correlated with time to loss of local control/censor for both the HCC and MET groups, which is in agreement with the studies by Andolino et al. [10] and Chang et al. [11], respectively. Therefore, although the HCC tumours were significantly larger than the MET tumours (table 2.1), tumour size could not be used to explain the significant differences between HCC and MET dose-response parameters or local control rates. Prescription dose was also not significantly different between the two groups (table 2.1). In the future, Cox regression analyses may provide a more robust understanding of the interactions between tumour volume, dose, and time to loss of local control when compared to the correlation tests performed within the present study.

A key difference between the two demographics was the heavily pre-treated nature of the MET group. 88% of MET patients had received previous chemotherapy compared with only 14% of HCC patients. This is consistent with the pattern of referral at our institution, whereby MET patients are usually treated with RT after failing multiple chemotherapy regimens. Consequently, the MET patients tend to be further
along in their disease than the HCC patients. This may explain why the median time to loss of local control for the MET group (183 days) was much lower than for the HCC group (543 days) as well as why higher doses were required to control MET colorectal liver tumours.

The local control rates determined for the HCC and MET patients in this investigation may seem low when compared with the current literature. For example, Andolino et al. [10] found that HCC patients (n = 37) who were ineligible for transplant had a 2-year local control rate of 87% compared with the 48% (23-73%) rate observed in this study. Similarly for the MET group, Chang et al. [11] reported 1- and 2-year local control rates of 67% and 55%, respectively, for patients with MET colorectal liver tumours (n = 65) compared with the 32% (6-58%) and 0% (0-42%) local control rates reported here. Currently, there is a lack of consensus on whether liver tumour size correlates with radiotherapy outcome [10, 11, 35, 36]. However, a notable difference between the present and aforementioned studies is tumour size, which will be summarised in the subsequent paragraph. The lower local control rates observed in this study is the reason for our choice of a 6-month local control TCP end point (instead of 1 year), particularly among the MET patient group whose median time to loss of local control was approximately 6 months. This ensures that patient response within the two groups can be maximally stratified as a function of dose.

For Andolino et al. [10], the median tumour diameter was 3.5 cm [range = (1, 6.5)] compared with 6.6 cm [range = (3, 18)] for our HCC group. In contrast, Tse et al. [34] investigated a group of HCC and intrahepatic cholangiocarcinoma patients (n = 49) with more comparable tumour sizes and reported a 65% 1-year local control rate that is similar to the 1-year local control rate reported here. The median tumour volume in their investigation was 173 cc [range = (9, 1913)] compared with 186 cc [range = (8, 995)] for our HCC group. For Chang et al. [11], the median tumour
volume was 30 cc \([\text{range} = (0.5, 3008)]\) compared with 57 cc \([\text{range} = (5, 1804)]\) for the MET group in this study.

In this study, we used a combination of biomarker and radiographic response data to help define local control. However, for MET patients with untreated metastatic tumour burden, multiple tumours could be contributing to increased CEA levels that hide decreases caused by treatment of the target lesion. We attempted to mitigate this effect by allowing for a \(<20\%\) increase in biomarker levels within the local control definition and by supplementing biomarker data with local radiographic information.

### 2.5 Conclusion

In conclusion, we have found radiation dose-response relationships for patients with HCC and MET groups. \(D_{50}\) was determined to be 53 Gy \((\sigma = 5.6 \text{ Gy})\) and 70 Gy \((\sigma = 6.6 \text{ Gy})\) and the slope parameter \(k\) was estimated to be 4.8 \((\sigma = 2)\) and 7.1 \((\sigma = 3.3)\) for the HCC and MET groups, respectively. 2 Gy per fraction equivalent doses of 84 Gy and 95 Gy were predicted to result in 90\% 6-month local control rates for patients in the HCC and MET groups, respectively. RT for HCC and MET results in significantly different local control rates at our institution, which may warrant an investigation into the effect of earlier RT referral for patients with MET. Improved understanding of the dose-response relationships for patients with primary or metastatic liver cancer will help to inform future dose prescriptions.
2.6 Funding

This research was funded by the Canadian Cancer Society (grant no. 700386). The study sponsors had no involvement in the study design, collection, analysis and interpretation of data and in the writing of the manuscript or in the decision to submit this manuscript for publication. The authors have no financial interest in the submitted work.
References


[28] Beom Kyung Kim, Sang Hoon Ahn, Jin Sil Seong, Jun Yong Park, Do Young Kim, Ja Kyung Kim, Do Youn Lee, Kwang Hoon Lee, and Kwang-Hyub Han. Early α-fetoprotein response as a predictor for clinical outcome after localized


Chapter 3

An augmented parametric response map with consideration of image registration error: towards guidance of locally adaptive radiotherapy *

3.1 Introduction

Traditionally, the aim of radiation therapy (RT) is to deliver a high uniform dose to the tumour while limiting the dose received by surrounding normal tissues. However, an increasing number of studies as reviewed by Dewhirst et al. [1] suggests that there may be a large amount of variability in intra-tumoural radio sensitivity. Consequently there has been a push towards developing new treatment strategies such as sub-volume boosting or dose-painting by numbers to account for this variability [2].

Since the local radio-sensitivity within a tumour can also vary during the course of treatment [1], these personalized strategies may require an adaptive implementation.

[2]. Imaging-based biomarkers of local radio-sensitivity or tumour response could be evaluated one or more times during the treatment course to ensure the continuing optimality of the non-homogeneous dose distribution. In case of a sub-optimal dose distribution, the RT plan could be updated to take advantage of each patient’s unique tumour response to radiation.

A vital step towards personalized adaptive RT is the development of reliable imaging-based biomarkers which can predict the early localized influence of radiation on the tumour. Recently, Galban et al. [3] presented a promising method for predicting treatment response known as the parametric response map (PRM). The PRM is generated through a local voxel-wise analysis of repeat imaging with the objective of predicting a global measure of post-treatment response or progression. The PRM also indicates which tumour voxels are associated with the predicted response or progression. Since the analyzed images are typically acquired prior to and during the treatment course, response could potentially be predicted on a voxel-wise basis during the treatment itself. Therefore the PRM approach may be ideal for guiding adaptive RT strategies prior to or in the absence of observable changes in tumour volume. To date, the method’s predictive utility has been demonstrated for several different pathologies and treatments such as RT for gliomas and head and neck cancers and chemoembolization for hepatocellular carcinomas [3–8].

A major challenge towards guiding RT using voxel-based analyses such as the PRM is their inherent sensitivity to image registration error (IRE) [9–11]. Repeat imaging must be spatially registered to establish correspondences between voxels prior to analysis. However, voxel-based comparisons may be sensitive to IRE with the amount of sensitivity depending on the interplay between the magnitude and direction of the error and the size, orientation and complexity of the image features. Consequently, the sensitivity of voxel-based analyses to IRE is a property of the
analyzed images and may be different for every patient. In a previous work we demonstrated that indeed the local voxel-wise classification accuracy of the PRM method may be prohibitively sensitive to small registration errors [12]. While the overall PRM analysis may still be correlated with a global measure of response, IRE can cause individual responding or progressing tumour voxels to be misclassified as one another, which could lead to targeting errors if PRMs were used to guide RT.

Since the careful management of uncertainty is a critical component of the RT workflow, robust voxel-based analysis methods are needed which permit the user to identify whether IRE-related analysis uncertainty can influence target delineation. Here our primary objective was to address this challenge by extending the original single-threshold PRM method [3] to include a PRM analysis confidence interval related to the influence of IRE (§3.2.2, §3.2.3, §3.2.4). In addition, we incorporate multiple graded classification thresholds into this framework in order to facilitate visual assessment of IRE-related analysis uncertainty and regional tumour response trends (§3.2.5). Risholm et al. [10] and Murphy et al. [11] have previously reported on accounting for IRE uncertainty during RT dose accumulation; however to our knowledge the present study is the first to explore IRE-related image analysis uncertainty within the context of image-guided locally-adaptive RT.

The augmented PRM (A-PRM) method was demonstrated and contrasted with the PRM through application to repeat functional imaging with simulated IRE (§3.3). The new method was shown to help visualize and quantify the influence of registration error on PRM analysis while also providing additional contextual information which could be useful for guiding adaptive RT strategies such as sub-volume boosting (§3.4 and §3.5). Strategies for using the A-PRM method to guide RT in the presence of IRE-related PRM analysis uncertainty are discussed.
3.2 Methods

3.2.1 Original single-threshold parametric response map

PRMs are generated by analyzing the voxel-wise changes between spatially registered baseline and follow-up imaging of some region of interest (e.g. tumour). Since functional imaging is typically analyzed, registration is often performed by applying a transformation derived from the registration of accompanying co-registered anatomical images. The PRM indicates which voxels within the region of interest (ROI) are significantly increasing, decreasing, or not significantly changing in function. This information has been shown to correlate with global treatment response and progression for several datasets [3–8].

Let \( d(x) \) represent the observed voxel-wise change in function for a single voxel at position \( x \) within the baseline image set. The functional change is classified as significantly increasing or decreasing if \( d(x) > T \) or \( d(x) < -T \) respectively, where \( T \) is a single constant threshold value. Voxels where \( |d(x)| \leq T \) are classified as not significantly changing. The threshold \( T \) is determined by the functional changes occurring within an ROI of corresponding healthy tissue. For example, Galban et al. [3] analyzed repeat functional imaging of gliomas and used an ROI on the contralateral side of the brain that was homologous to the tumour ROI in a subset of patients. Follow-up functional image voxel values within the healthy ROI are plotted versus the corresponding voxel values within the baseline functional images. A line is then fit to the plot and \([-T, T] \) corresponds to the 95% confidence interval in the residuals from the linear fit. Henceforth we will refer to the PRM threshold as \( T_{95} \).
3.2.2 Influence of image registration error on PRM analysis

Since the PRM threshold $T_{95}$ is determined from a summative analysis of a large population of voxels, it should be robust to the influence of several sources of uncertainty such as IRE and image noise. More precisely, the calculated threshold value for a patient would not be expected to vary significantly due to these uncertainties. However, IRE introduces a second source of uncertainty into the PRM method. In the presence of IRE, each voxel-wise intensity difference is also uncertain because it is no longer clear which voxel intensity values should be compared to one another in the registered images. In this context, an individual voxel intensity difference takes on a distribution of possible values which could correspond to different classification outcomes when each value is compared to the PRM threshold $T_{95}$.

To illustrate this, we consider two 1D images, $I_1$ and $I_2$. Let $D(x)$ represent the distribution of possible voxel-wise differences for a single pixel at position $x$ within the baseline image, $I_1$. Figure 3.1 provides a motivating example for determining $D(x)$ in the presence of IRE. When $IRE = 0$, the difference between $I_1$ and $I_2$ at position $x$ is given by a single value $d(x) = I_2(x) - I_1(x)$. However when IRE is non-zero it is no longer certain which pixel in $I_2$ corresponds to $I_1(x)$. Therefore $d(x)$ takes on a distribution of potential values $D(x)$ which is produced by taking the difference between $I_1(x)$ and a neighborhood ($\Omega$) of pixels around and including position $x$ in $I_2$. $\Omega$ is defined by all the likely registration errors that could have occurred at position $x$.

If $T_{95} = 2$ in figure 3.1, then the analyzed pixel would be classified by the PRM method as either significantly decreasing or not significantly changing depending on which single pixel intensity difference was actually observed from the distribution $D(x)$. Consequently, IRE could lead to pixel (or voxel) misclassification by the PRM method.
Figure 3.1: 1D example for determining the distribution of possible image differences at a single image position $x$ when there is non-zero image registration error (IRE). $I_1$ and $I_2$ are two 1D images and $D(x)$ is the distribution of possible differences between $I_1$ and $I_2$ at position $x$. In this example, the IRE has a potential value of two pixels in either direction and spans the spatial domain $\Omega$. 

\[ D(x) = \begin{cases} 
I_2(x - 2) - I_1(x) \\
I_2(x - 1) - I_1(x) \\
I_2(x) - I_1(x) \\
I_2(x + 1) - I_1(x) \\
I_2(x + 2) - I_1(x) 
\end{cases} = \begin{Bmatrix} 
-3 \\
-5 \\
-4 \\
0 \\
-2 
\end{Bmatrix} \]
3.2.3 Incorporating IRE uncertainty into PRM analysis

For clarity we return to the 1D example from figure 3.1 however in subsequent sections the proposed methods are applied in 3D.

Instead of classifying a pixel based on a single intensity difference value observed from the distribution $D(x)$, the expected value of this distribution could be used. Let $\{\Delta s_i(x)\}$ represent the set of potential registration errors for a pixel at position $x$ where $\Delta s_i(x)$ is the distance between the position of the $i$-th pixel in $\Omega$ and the position $x$. If $p(x, \Delta s_i(x))$ represents the probability that the IRE is $\Delta s_i(x)$ for the pixel at position $x$ and $D_i(x)$ is the resulting pixel intensity difference then the expected value for $D(x)$ is given by

$$
\overline{D}(x) = E[D(x)] = \sum_i D_i(x)p(x, \Delta s_i(x)) \quad (3.1)
$$

In practice, the probability distribution function (PDF) $p$ for the registration error is unknown and must be approximated. The estimation of $p$ will be discussed in the next section. The uncertainty in $D(x)$ can then be represented by the standard deviation,

$$
\sigma_{IRE} = \sqrt{\sum_i (D_i(x) - \overline{D}(x))^2 p(x, \Delta s_i(x))} \quad (3.2)
$$

The expected value of the PRM analysis can be obtained for each pixel by comparing $\overline{D}(x)$ to the PRM threshold $T_{95}$. Upper and lower bounds of a PRM analysis confidence interval (CI) can then be generated by comparing $|\overline{D}(x)| + a\sigma_{IRE}(x)$ and $|\overline{D}(x)| - a\sigma_{IRE}(x)$ to $T_{95}$ respectively where $a$ is a positive constant that controls the width of the computed confidence interval (e.g. $a = 2$ implies a 95% CI). If $|\overline{D}(x)| - a\sigma_{IRE}(x) < 0$ then the voxel is classified by $||\overline{D}(x)| - a\sigma_{IRE}(x)|$ and the direction of the change (i.e. increasing or decreasing) is reversed.
3.2.4 Estimating the IRE probability distribution function

Accurate voxel-wise estimation of $p$ poses an extremely challenging task and is an open area of research [9, 11, 13–16]. As a first step, and for demonstration purposes, we explore the utility of a simple Gaussian definition of $p$ which is defined by a single conservative estimate of the voxel-wise IRE in each coordinate direction within the ROI. This approach was also adopted as it mirrors uncertainty management in other aspects of the RT workflow such as the use of standardized margins for generating planning target volumes. However in the future it would be straightforward to extend this method to spatially varying uncertainties if more localized IRE information were available.

Let $[\text{IRE}_{LR}, \text{IRE}_{AP}, \text{IRE}_{SI}]$ represent the upper bound of a prediction interval for the spatial misalignment due to IRE within the ROI in the left-right, anterior-poster and superior-inferior directions respectively. We defined $p$ by a 3D spatially invariant ellipsoidal Gaussian distribution where the standard deviation along each axis is given by,

$$[3\sigma_{LR}, 3\sigma_{AP}, 3\sigma_{SI}] = [\text{IRE}_{LR}, \text{IRE}_{AP}, \text{IRE}_{SI}]$$  \hspace{1cm} (3.3)

In practice, this upper bound would need to be estimated and provided by the user. Accordingly, the impact of under and over-estimation of the IRE is also reported in this study.

3.2.5 Augmented parametric response map

The A-PRM combines the ideas presented in §3.2.3 and §3.2.4 with the use of multiple graded classification thresholds [17] in order to enhance visualization of IRE-related analysis uncertainty and regional tumour response trends. Every voxel in the A-
PRM is classified by a specific threshold that indicates the magnitude and sign of the functional change. In contrast to the original single threshold PRM, the A-PRM permits visualization of the full distribution of tumour functional changes instead of only those changes with magnitudes that exceed the $T_{95}$ threshold. This should facilitate identification of responding or progressing tumour sub-volumes that could be targets for adaptive RT and permit visualization of IRE-related analysis uncertainty throughout the ROI when combined with the calculation of the analysis confidence interval.

Let $\{T_i\}$ for $i = 1 \ldots 95$ represent a set of classification thresholds which correspond to the 1%-95% confidence intervals for the linear fit residuals used to determine the original PRM threshold where a normal distribution is assumed for the linear fit residuals. Accordingly, $T_{95}$ represents the threshold used by the original PRM method. The expected value of each voxel-wise change $\overline{D}(x)$ is uniquely classified by the maximum $T_i$ for which $|\overline{D}(x)| \geq T_i$. The A-PRM then corresponds to an image of the ROI where each voxel value is given by $+i$ or $-i$ depending on whether the change was an increase or a decrease respectively. Therefore each A-PRM voxel value indicates the maximum threshold exceeded by the expected magnitude of each voxel-wise functional change as well as the direction of the change. This calculation is repeated for each voxel within the region of interest.

As a first step and for demonstration purposes, the associated confidence interval was chosen to span $\pm 3\sigma_{IRE}$. Other confidence intervals were also investigated and are reported within this study in a sensitivity analysis. To generate the upper bound of the confidence interval, a voxel is classified by the maximum $T_i$ for which $|\overline{D}(x)| + 3\sigma_{IRE}(x) \geq T_i$. To generate the lower bound of the confidence interval, a voxel is classified by the maximum $T_i$ for which $|\overline{D}(x)| - 3\sigma_{IRE}(x) \geq T_i$. Response maps corresponding to these upper and lower bound classifications were also generated and
are denoted by \(A-\text{PRM}^+\) and \(A-\text{PRM}^-\) respectively. If \(|\mathcal{D}(x)| - 3\sigma_{\text{IRE}}(x) < 0\) then the voxel is classified according to the magnitude of the left hand side of the inequality and the direction of the change (i.e. increasing or decreasing) is reversed.

Figure 3.2 provides an illustrative example which contrasts the PRM and A-PRM voxel classification approaches using scatter plots. Without loss of generality, all analyses occur in the spatial frame of reference of the baseline image set (i.e. uncertainty is in follow-up image voxel values).

![Figure 3.2: Voxel-wise change classification using the PRM and A-PRM methods for eight example voxels (A-H). The observed and expected change for each voxel are marked with an \(x\) and a circle, respectively, and the error bars span \(\pm 3\sigma_{\text{IRE}}\). The table indicates the threshold that would be used to classify voxels A-H for each method with the sign indicating whether the change was an increase or a decrease. The range reported for the A-PRM represents classification uncertainty due to IRE (i.e. the confidence interval). A-PRM\(^+\), the upper bound of the interval, classifies voxels by the ends of the error bars furthest away from the unity line in the plot. Conversely, A-PRM\(^-\), the lower bound of the interval, classifies voxels by the ends of the error bars closest to the unity line.](image-url)
3.3 Evaluation

Ideally, the A-PRM method would be evaluated using perfectly registered pairs of functional images. PRM and A-PRM analysis of these images would be uncontaminated by IRE and would serve as a “ground truth” analysis. Controlled amounts of IRE could then be simulated by spatially perturbing the registered images. The PRM and A-PRM analyses would be repeated and the voxel classification results could be compared to the ground truth analyses to determine the relative performance of the two methods in the presence of known IRE.

However in practice the transformation required to perfectly register an image pair is unknown. For the purposes of this study, we approximate this transformation with the transformation output from a non-rigid registration of the functional image pairs. Due to residual IRE, the functional changes conveyed by the resulting “ground truth” PRM and A-PRM analyses may not fully coincide with the true functional changes occurring within the tumours. However, this should serve as a suitable approximation for demonstrating and verifying the behavior of the A-PRM methodology. The evaluation pipeline is summarized below and detailed further in subsequent sections.

(a) non-rigidly register original baseline and follow-up functional images

(b) perform PRM and A-PRM analysis on non-rigidly registered images to produce ground truth analysis

(c) apply known rigid transformations to non-rigidly registered images to simulate controlled rigid IRE

(d) perform PRM and A-PRM analysis on images with simulated rigid IRE and compare to ground-truth analysis
3.3.1 Image data

CT-perfusion scans of four patients with high-grade glioma and four patients with hepatocellular carcinoma (HCC) were used to demonstrate the A-PRM methodology. Pre and post-RT scans were acquired using a 64 slice scanner (Discovery VCT or CT750HD, GE Healthcare, Waukesha, WI). Glioma scans were acquired with a voxel size of 0.5 mm by 0.5 mm by 5 mm and a temporal sampling of 1 s for the first 45 s and 15 s for the remaining 105 s. HCC scans were acquired with a voxel size of 0.7 mm by 0.7 mm by 5 mm and a temporal sampling of 2.8 s. Glioma and HCC image volumes consisted of 512 by 512 by 8 voxels and 512 by 512 by 16 voxels respectively. The specific imaging protocols and perfusion analysis methods used for the glioma and HCC groups are further described in Yeung et al. [18] and Jensen et al. [19] respectively.

Blood volume (BV) and arterial blood flow maps (ABF) from the CT-perfusion analyses were the functional images analyzed for the glioma and HCC cases respectively (figure 3.3). BV maps were investigated for the glioma cases since MRI derived relative BV has previously been shown to have predictive utility for brain tumours using the PRM method [3]. ABF maps were investigated for the HCC cases since arterial phase enhancement in contrast-enhanced CT scans may have the potential to identify viable liver tumour using the PRM method [8]. These functional data are sufficient for evaluating the A-PRM methodology and are promising candidates for response analysis; however, verifying their predictive utility is outside the scope of the present study.
Figure 3.3: Example slices from the functional image data used for all simulations. (a) Pre and post-RT blood volume maps for four different gliomas. (b) Pre and post-RT arterial blood flow maps for four different hepatocellular carcinomas. Tumour functional maps were non-rigidly registered and superimposed on the pre-RT average dynamic contrast-enhanced computed tomography (DCE-CT) images for each patient.
3.3.2 Image registration

The Flexible Algorithms for Image Registration (FAIR) toolkit was used to non-rigidly register the functional images [20]. Pre and post-RT functional images were registered by applying a transformation derived from non-rigidly registering accompanying co-registered anatomical CT images. Registration was performed using a multi-resolution elastic registration algorithm based on normalized gradient fields (NGF). The NGF edge parameter was set to 5. The linear elastic Lame parameters $\mu$ and $\lambda$ were set to 1 and 0 respectively and the regularization weighting parameter was set to 500. The maximum deformation grid considered by the multi resolution framework consisted of 256 by 256 by 32 control points. Effective registration parameter values were determined through trial and error.

3.3.3 Simulating image registration error

Each set of registered follow-up functional images was spatially perturbed by 1000 randomly selected rigid transformations. Translations and rotations were sampled from normal distributions with zero means and 1 mm and 1° standard deviations respectively which produced a spectrum of IREs within previously reported ranges for brain and liver [21, 22]. The centre of rotation was defined to be the centre of each image volume. The spatial locations of all voxels within the functional images were known before and after each simulation permitting exact calculation of the simulated IRE throughout the investigated ROIs.

3.3.4 Response maps

PRMs and A-PRMs were first generated for each patient from the non-rigidly registered functional image data without simulated IRE. Patient-specific sets of thresholds
\{T_i\} for \( i = 1 \ldots 95 \) were determined from normal tissue ROIs defined for each patient. For the HCC cases, this ROI was defined for each patient as the entirety of the non-tumour bearing liver. For glioma cases, normal tissue ROIs were defined as tumour-homologous regions on the contralateral side of the brain with respect to the tumour.

A thousand different PRMs and A-PRMs were then generated for each patient corresponding to analysis of the registered functional images that contained a thousand different magnitudes of simulated rigid IRE. For each IRE simulation, the 99th percentile of the known tumour voxel displacements in each coordinate direction was used to define \([IRE_{LR}, IRE_{AP}, IRE_{SI}]\) and the IRE PDF \( p \) defined in §3.2.4. This definition should ensure that the width of the gaussian IRE probability distribution function \( p \) encompasses the full range of simulated IREs. The sensitivity of the A-PRM method to different IRE estimates is also reported.

### 3.3.5 Analysis

Voxel classification outcomes from the PRM and A-PRM analyses with simulated IRE were compared to the PRM analyses without simulated IRE for each patient. Therefore the ground-truth analyses within this study were defined by the PRMs generated from the non-rigidly registered image data prior to IRE simulation. Three different investigations were performed.

#### 3.3.5.1 Voxel classification performance

First, the true classification rates for the PRM and A-PRM methods were plotted as a function of median tumour voxel IRE for each patient. Analysis was performed with respect to the original \( T_{95} \) threshold and separate plots were generated for the decreasing, no change, and increasing classes. The true classification rate represents
the fraction of the total voxels that belong to a given class that are correctly identified as such by the classification method. The purpose of this analysis was to compare how well the ground truth PRM information is preserved in the PRMs and A-PRMs when both are influenced by IRE. This analysis was also performed on the maps corresponding to the bounds of the A-PRM confidence interval (A-PRM⁻ and A-PRM⁺).

True classification rates were also computed for the original PRM method after performing a spatial expansion of the regions identified as significantly increasing and decreasing within the IRE-contaminated PRMs. The purpose of this additional analysis was to identify if the simple spatial margin approach used to manage patient position uncertainty in RT dose planning can be used to manage IRE uncertainty in PRM analysis. Regions classified as increasing and decreasing were expanded by $[IRE_{LR}, IRE_{AP}, IRE_{SI}]$ in each coordinate direction. If regions classified as increasing and decreasing overlapped with each other after expansion, then the overlapping region was classified as no-change. Spatially expanded PRMs were denoted by PRM\text{exp}.

3.3.5.2 Sensitivity to IRE estimation

Second, we assessed the impact of under or over-estimation of $[IRE_{LR}, IRE_{AP}, IRE_{SI}]$ on the A-PRM method. Outside of a simulation context this information is unknown and would need to be estimated to define the IRE PDF $p$. Therefore for all IRE simulations, another set of A-PRM analyses were performed for each patient corresponding to six different estimations of $[IRE_{LR}, IRE_{AP}, IRE_{SI}]$. Each IRE estimate was defined by increasing or decreasing the known 99th percentile of the tumour voxel displacements in each coordinate direction by 1 mm, 3 mm, and 5 mm. If decreasing the known 99th percentile of the voxel displacements along a coordinate direction
resulted in a negative IRE, then the IRE estimate was set to zero along that direction.

Voxel classification accuracy with respect to the $T_{95}$ threshold was re-computed in each case to investigate A-PRM sensitivity to the IRE estimate.

### 3.3.5.3 Generalized A-PRM analysis

Third, an analysis was performed on a generalized version of the A-PRM method corresponding to classification based on the inequality $|\bar{D}(x)| + a\sigma_{IRE}(x) \geq T_i$. The constant $a$ was varied from $-6$ to 6 to produce a spectrum of different A-PRMs for each magnitude of simulated IRE. Accordingly, maps generated using $a = -3$, $a = 0$, and $a = 3$ correspond to the originally proposed A-PRM$^-$, A-PRM, and A-PRM$^+$ maps respectively.

### 3.4 Results

Figure 3.4 depicts representative slices from the 3D PRM and A-PRM analyses of the non-rigidly registered functional images from figure 3 prior to IRE simulation. Therefore, figure 3.4 PRMs (first rows) represent slices from the ground truth analyses used throughout this study. Figure 3.4 A-PRMs (second rows) were generated with IRE set to 0 since we used the non-rigidly registered functional images as a zero-IRE reference within this study. Consequently, there is no IRE-related analysis uncertainty and the A-PRM confidence interval is not displayed as it has zero width. Since IRE is set to zero, the figure 3.4 A-PRMs solely demonstrate the impact of adding multiple graded thresholds to the original PRM method as in Lausch et al. [17].

Figure 3.4 A-PRMs contain all of the same information as the PRMs but also provide additional contextual information by using multiple graded thresholds to classify the entire distribution of functional changes within the tumour. For example, distinct
Figure 3.4: PRM (first row) and A-PRM (second row) analyses of the functional image data from figure 3.3(a) (gliomas) and 3.3(b) (HCCs), respectively, with $IRE = 0$. Response maps have been superimposed on the corresponding pre-RT average DCE-CT images for each patient. Notable sub-volumes of functional increase are highlighted for Glioma-3 and HCC-3 (dashed ellipses).
sub-volumes of functional increase and decrease became more clearly observable in the A-PRMs with notable examples indicated by dashed ellipses in the Glioma-3 and HCC-3 A-PRMs.

For figure 3.4, there was no IRE-related analysis uncertainty. In contrast, figure 3.5 demonstrates how the figure 3.4 analyses for glioma-1 and HCC-3 would change if 3 mm of isotropic IRE was assumed after non-rigid registration. Maps produced from the original PRM method (§3.2.1) and the A-PRM method which incorporates both IRE uncertainty and graded thresholds into the PRM (§3.2.5) are shown for comparison. Since the original PRM method does not take IRE into consideration, the two original PRMs shown in figure 3.5 are the same as those shown in figure 3.4. Uncertainty analysis can also be incorporated into the original PRM method as discussed in section §3.2.3. PRMs which incorporate IRE uncertainty (but without graded thresholds) are also shown in 3.5 for further comparison. To facilitate visualization of differences between the upper and lower bounds and the the expected value map, magnitude difference images were also included in figure 3.5. These images were generated by subtracting the magnitude of the expected value maps from the upper and lower bound maps.

The figure 3.5 A-PRMs contain all of the same information as the figure 3.5 PRMs that incorporate IRE uncertainty but also classify the functional changes occurring in the rest of the ROI using multiple graded thresholds. The non-zero IRE produces a non-zero confidence interval reflecting potential IRE-related analysis uncertainty. For the A-PRM method, the lower bound estimate (A-PRM⁻) classifies voxel-wise changes with lower magnitude thresholds while the upper bound estimate (A-PRM⁺) classifies voxel-wise changes with higher magnitude thresholds. This can be observed in the magnitude difference images. The upper and lower bounds maps correspond to the bounds of the analysis confidence interval. For the A-PRM, these bounds cor-
Figure 3.5: Example parametric response maps (PRMs) and augmented PRMs when the uncertainty due to 3 mm of isotropic IRE is incorporated into the PRM and A-PRM analysis of (a) glioma-1 and (b) HCC-3 from figure 3.4.
respond to A-PRM$^+$ and A-PRM$^-$ respectively. The magnitude difference images in figure 3.5 also provide a representation of the width of the confidence interval for each voxel in the A-PRM analysis. Regions of high IRE-related analysis uncertainty are reflected by non-zero values in the magnitude difference images. High uncertainty regions appeared in the vicinity of A-PRM feature boundaries like the interface between the decreasing (blue) and increasing (red) rims in figure 3.5(a).

### 3.4.1 Voxel classification performance

The PRM and A-PRM true classification rates with respect to the $T_{95}$ threshold were plotted as a function of the simulated IRE for all patients (figure 3.6). To produce each curve, the true classification rate values resulting from the thousand IRE simulations per patient were first binned according to median tumour voxel IRE (1-mm bins). The median true classification rates within each bin were then plotted against the median tumour voxel IRE for the A-PRMs as well as for the maps corresponding to the bounds of the confidence intervals (A-PRM$^-$ and A-PRM$^+$).

True classification rates for the PRM and A-PRM decreased with increasing IRE for all patients but at different rates suggesting the need for patient-specific analysis of PRM uncertainty due to IRE. A-PRM$^+$ maps improved identification of increasing and decreasing voxels at the cost of decreased identification of no change voxels.

Conversely, A-PRM$^-$ maps provided improved identification of no-change voxels at the cost of decreased identification of increasing and decreasing voxels. True increasing rates were found to be more sensitive to IRE than the true decreasing or true no change rates for the investigated image data.
Figure 3.6: True classification rates with respect to the $T_{95}$ threshold for the original PRM method and the A-PRM method plotted as a function of simulated rigid IRE for the four glioma and HCC cases. The true classification rate represents the fraction of the total voxels that belong to a given class that are correctly identified as such by the classification method.
Table 3.1: Median difference between A-PRM and PRM true classification rates (A-PRM rates subtract PRM rates) for all glioma simulations, binned by median tumour voxel image registration error. The indicated range contains 95% of all values within each bin. Median differences between PRM_{exp} and the PRM are also shown for comparison.

<table>
<thead>
<tr>
<th>Class</th>
<th>Median tumour voxel IRE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-2 (mm)</td>
</tr>
<tr>
<td>-------</td>
<td>----------</td>
</tr>
<tr>
<td>A-PRM</td>
<td></td>
</tr>
<tr>
<td>inc</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>[0.06, 0.28]</td>
</tr>
<tr>
<td>nc</td>
<td>-0.05</td>
</tr>
<tr>
<td></td>
<td>[-0.11, -0.01]</td>
</tr>
<tr>
<td>dec</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>[0.04, 0.15]</td>
</tr>
<tr>
<td>A-PRM^+</td>
<td>inc</td>
</tr>
<tr>
<td></td>
<td>[-0.06, 0.02]</td>
</tr>
<tr>
<td>nc</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>[0, 0.05]</td>
</tr>
<tr>
<td>dec</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>[-0.01, 0.03]</td>
</tr>
<tr>
<td>A-PRM^-</td>
<td>inc</td>
</tr>
<tr>
<td></td>
<td>[-0.30, -0.07]</td>
</tr>
<tr>
<td>nc</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>[0.01, 0.12]</td>
</tr>
<tr>
<td>dec</td>
<td>-0.10</td>
</tr>
<tr>
<td></td>
<td>[-0.14, -0.06]</td>
</tr>
<tr>
<td>PRM_{exp}</td>
<td></td>
</tr>
<tr>
<td>inc</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>[0.03, 0.36]</td>
</tr>
<tr>
<td>nc</td>
<td>-0.15</td>
</tr>
<tr>
<td></td>
<td>[-0.47, -0.03]</td>
</tr>
<tr>
<td>dec</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>[0.02, 0.21]</td>
</tr>
</tbody>
</table>

Patient simulation results were then pooled together (independently for each group) and the median differences in true classification rates between the A-PRM and PRM methods (A-PRM subtract PRM) were computed within three IRE bins: 0-2 mm, 2-4 mm and 4-6 mm. Tables 3.1 and 3.2 summarize the median differences by IRE bin for the glioma and HCC groups respectively. For comparison, median differences were also computed between the PRM_{exp} and PRM true classification rates within each IRE bin.

A priori, all table 3.1 and 3.2 median differences were expected to be statistically significant due to the large number of simulations performed (1,000). To confirm this, repeated Wilcoxon signed-rank tests with a Bonferroni correction for multiple comparisons were performed to test for significance. Indeed, all median differences in tables 3.1 and 3.2 were found to be significant for both patient groups (p < 0.001).
Table 3.2: Median difference between A-PRM and PRM true classification rates for all HCC simulations, binned by median tumour voxel image registration error. The indicated range contains 95% of all values within each bin. Median differences between PRM_{exp} and PRM are also shown for comparison.

<table>
<thead>
<tr>
<th>Class</th>
<th>Median tumour voxel IRE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-2 (mm)</td>
</tr>
<tr>
<td>A-PRM+</td>
<td>inc 0.16 [0.06, 0.21]</td>
</tr>
<tr>
<td></td>
<td>nc -0.03 [-0.09,-0.01]</td>
</tr>
<tr>
<td></td>
<td>dec 0.06 [0.03, 0.08]</td>
</tr>
<tr>
<td>A-PRM</td>
<td>inc 0 [0, 0.03]</td>
</tr>
<tr>
<td></td>
<td>nc 0.01 [0.01, 0.03]</td>
</tr>
<tr>
<td></td>
<td>dec 0 [-0.02, 0.01]</td>
</tr>
<tr>
<td>A-PRM-</td>
<td>inc -0.17 [-0.29,-0.10]</td>
</tr>
<tr>
<td></td>
<td>nc 0.03 [0.01, 0.09]</td>
</tr>
<tr>
<td></td>
<td>dec -0.06 [-0.09,-0.04]</td>
</tr>
<tr>
<td>PRM_{exp}</td>
<td>inc 0.17 [0, 0.33]</td>
</tr>
<tr>
<td></td>
<td>nc -0.07 [-0.35, 0.02]</td>
</tr>
<tr>
<td></td>
<td>dec 0.03 [0, 0.08]</td>
</tr>
</tbody>
</table>

with the exception of table 3.2 A PRM increasing class for the 0-2 mm IRE bin and the table 3.2 A-PRM decreasing class for the 2-4 mm IRE bin ($p > 0.05$).

For the figure 3.6 simulations, the median A-PRM classification (median value of $i$) was computed for the increasing and decreasing voxels that were misclassified due to IRE. Results were aggregated for each patient group and binned by IRE (table 3.3). The purpose of this sub-analysis was to assess the magnitude of A-PRM misclassification for voxels that should have been classified by the original PRM threshold ($T_{95}$) but were misclassified by other $T_i$ due to IRE.

### 3.4.2 Sensitivity to IRE Estimation

The median A-PRM true classification rates resulting from different under or over-estimations of the known tumour voxel IRE were then computed for each IRE bin.
Table 3.3: Median A-PRM classification for voxels that should have been classified by the original PRM threshold ($T_{95}$) but were misclassified due to IRE. Simulation results were aggregated within each group and binned by IRE. The indicated range contains 95% of all values within each bin.

<table>
<thead>
<tr>
<th>True Class</th>
<th>Median A-PRM classification</th>
<th>Median tumour voxel IRE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0-2 (mm)</td>
</tr>
<tr>
<td>Gliomas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>68</td>
<td>[-84, 94]</td>
</tr>
<tr>
<td>-95</td>
<td>-87</td>
<td>[-94, 40]</td>
</tr>
<tr>
<td>HCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>78</td>
<td>[-22, 94]</td>
</tr>
<tr>
<td>-95</td>
<td>-88</td>
<td>[-94, 95]</td>
</tr>
</tbody>
</table>

Results were aggregated for all patients within each patient group and are plotted in figure 3.7. Each line corresponds to a different IRE bin. Each IRE estimate was defined by increasing or decreasing the known 99th percentile of the voxel IRE in each coordinate direction by the amount specified by the horizontal axis.

3.4.3 Generalized A-PRM analysis

For the analysis of the generalized A-PRM ($|D(x)| + aσ_{IRE}(x) ≥ T_i$), the simulations for the patients within each group were again pooled together and then binned by IRE. The median true classification rates within each IRE bin were computed for each definition of the generalized A-PRM corresponding to different values of the constant $a$. The true classification rates with respect to the $T_{95}$ threshold were then plotted versus $a$ independently for each IRE bin (figure 3.8).
Figure 3.7: Influence of under/over estimation of the IRE on the A-PRM true classification rates aggregated for the four glioma and HCC cases. Each line corresponds to a different IRE bin and the horizontal axis indicates the amount of under/over estimation of the IRE in each coordinate direction. True classification rates were calculated with respect to the PRM threshold ($T_{95}$).

Figure 3.8: Influence of the parameter $a$ on the true classification rates for the generalized A-PRM method. Results for $a = 3$, $a = 0$, and $a = 3$ correspond to the originally proposed A-PRM$^-$, A-PRM, and A-PRM$^+$ maps respectively. For reference, the median performance of the PRM (circle marker) and PRM$_{exp}$ (x marker) are indicated for each IRE bin.
3.5 Discussion

The PRM shows potential as a method for guiding personalized locally adaptive treatment strategies such as sub-volume boosting or dose-painting by numbers. However, in the present study we have shown that IRE can lead to patient-specific tumour voxel misclassification which could pose challenges for the reliable guidance of RT. Since uncertainty management is a critical component of the RT workflow, our goal was to extend the PRM to permit identification of IRE-related analysis uncertainty with a view towards determining if it has the potential to influence target delineation. Our approach involved using a simple Gaussian IRE model to determine the expected response value of each voxel (A-PRM) as well as estimate a confidence interval to reflect the analysis uncertainty caused by IRE (A-PRM, A-PRM+). Multiple graded classification thresholds were also incorporated to facilitate visual assessment of IRE-related analysis uncertainty and regional tumour response trends (§3.2.5). Here we will discuss the potential utility of the A-PRM for guiding adaptive RT followed by a discussion of the study limitations.

First, through the use of multiple graded thresholds, the A-PRM provided additional contextual information related to the regional functional changes occurring in the cases shown in figure 3.4. For example, in glioma-3 and HCC-3, distinct sub-volumes of functional increase and decrease became apparent. While voxels within these regions may be classified by thresholds less than \( T_{95} \) (the threshold typically correlated with response), the graded threshold information may be useful because it highlights the location of relatively homogeneous regions of response or progression which could be used to define adaptive RT targets. In addition, the table 3.3 results suggest that the A-PRM misclassifies voxels due to IRE in a graded fashion (i.e. voxels are misclassified by slightly lower or higher thresholds). This is in contrast to
the PRM method where misclassification (due to IRE) is more severe. For example, an increasing voxel can only be misclassified as no change or decreasing within the original PRM.

The homogeneous regions of response visualized using the graded thresholds in the A-PRM were also observed to be less susceptible to the influence of IRE. Qualitatively, this can be appreciated upon inspecting the magnitude differences between the A-PRM and A-PRM$^+$ and A-PRM$^-$ maps shown in figure 3.5. Analysis uncertainty was the greatest for voxels that were close to feature boundaries in the A-PRM, underscoring the relationship between increased analysis heterogeneity and IRE influence. For example, non-zero magnitude differences in the glioma case were located along the boundaries of the outer blue and red rims in the A-PRM. In contrast, the analysis uncertainty was the least (i.e. small magnitude differences) for voxels that were far away from feature boundaries. This can be observed in the center of the tumour where there are no distinct features and A-PRM analysis is more homogeneous.

In practice, if potential targets would be substantially different based on using A-PRM$^+$ versus A-PRM$^-$ maps then it could be decided that the IRE related analysis uncertainty was too high to permit reliable guidance of RT for that patient. However within this context, several different approaches for adapting the radiotherapy target could be taken using the different A-PRM maps. For example, A-PRM$^+$ could be used to ensure that progressing tumour regions are not missed due to IRE. A-PRM$^+$ was designed to improve detection of increasing and decreasing voxels at the cost of misclassifying no change voxels as changing (figure 3.6 and tables 3.1 and 3.2). Therefore if within organ at risk dose limits, radiation dose could be escalated to regions identified as progressing with the understanding that non-progressing tumour voxels might also receive a higher dose.

This approach may be well-suited to the specific patient image data investigated
within this study due to their high sensitivity to IRE with respect to classification of increasing voxels (figure 3.6). This increased sensitivity was due to the lower number and more spatially diffuse nature of the increasing voxels in the investigated image data when compared to the decreasing and no change voxels (i.e. decreased feature size relative to the IRE). The simple spatial margin approach used to create PRM_{exp} produced results similar to A-PRM^+. However, tables 3.1, 3.2 and figure 3.8 indicated that as IRE was increased, the true no-change rate for PRM_{exp} dropped off much faster than for A-PRM^+, suggesting limited utility for the spatial margin approach.

Conversely, A-PRM^- was designed to improve detection of no-change voxels at the cost of misclassifying increasing or decreasing voxels as no-change (figure 3.6 and tables 3.1 and 3.2). This reduces the likelihood that voxels are misclassified as increasing or decreasing due to IRE, but reduces the total number of the actual increasing and decreasing voxels detected. Therefore dose could be modified using the A-PRM^- with the understanding that treatment only deviates away from conventional non adaptive dose distributions where the image analysis is known to be less influenced by IRE.

Ultimately, the ability to locally adapt RT is limited by the size of the regions to be adapted relative to the radiation targeting capabilities of the treatment machine. With this constraint in mind, the delivery of boost doses to progressing tumour sub-volumes will be an important first step towards locally adapted RT. Previously we discussed how increased analysis heterogeneity increases the influence of IRE. Regions of homogeneous analysis that are large enough to be treated with a sub-volume boost in practice (e.g. 1-2 cm) may be fairly robust to the influence of moderate magnitudes of IRE (e.g. 2-4 mm). If they exist, these regions are easily identified on the A-PRM due to the use of the multiple graded classification thresholds. Therefore, the A-PRM
should be well suited towards defining sub-volume boost targets within contexts where standard PRM analysis has been correlated with a measure of treatment response. A retrospective RT planning study of the technical feasibility and potential benefit of radiation delivery to these sub-volumes constitutes an important next step towards implementing locally adaptive radiotherapy using the A-PRM.

A key limitation of the current A-PRM method is the use of the simple Gaussian error model to define the IRE probability distribution function. There are two major drawbacks to this approach.

First, the user must provide an estimate of the IRE anticipated within the tumour. While this is feasible through identification of anatomically corresponding fiducials in the registered images (e.g. De Silva et al. [23]), it can present a labor-intensive task. In practice a single suitably conservative estimate may need to be determined which represents an upper bound for the IRE anticipated for the majority of patients scanned and registered with the same technique. The standard RT workflow adopts a similar approach with respect to managing uncertainty in the delineation and position of planning target volumes. However, in figure 3.7 we observed that IRE overestimation reduces detection of increasing and decreasing voxels while increasing detection of unchanging voxels. Therefore depending on the images, use of a single over-estimated value of the IRE may obscure some of the significantly changing tumour regions classified by $T_{95}$. Despite this effect, regional response trends should still be observable in the A-PRM due to the use of multiple graded classification thresholds.

Second, the IRE PDF is defined to be independent, constant and symmetric about a zero mean IRE for all voxels within a pair of registered image volumes. This constitutes a first order approximation since the true but unknown IRE PDF will vary for each voxel depending on the proximity of the voxel to image features that are salient to the algorithm used for registration. The IRE PDFs for adjacent voxels may also be
highly inter dependent due to the regularization of the deformation field. Since a single upper estimate of the voxel-wise IRE is used to define the IRE PDF, the influence of IRE will be over-estimated for many voxels. Other error models (statistical distributions) apart from a Gaussian model could be considered. However, such models would exhibit similar limitations such as spatial invariance and under-representation of IRE complexity. Other models would also introduce different assumptions about the underlying IRE distribution (e.g. non-symmetric, direction-bias) that may or may not be valid throughout the analyzed ROI.

An advantage of this simple model-based approach is that it provides the user with a highly accessible method to rapidly probe the influence of IRE on their analysis. The A-PRM can be updated in near real time as the user tests the effects of different anticipated IRE magnitudes. In addition, the simple Gaussian error model can easily be applied to voxel based analyses other than the A-PRM in order to estimate the potential influence of IRE. Despite its simplicity, this first order approach was also demonstrated by 3.1, 3.2 and the generalized A-PRM analysis to provide some small improvements to classification performance for decreasing and no change voxels.

Nonetheless, improved local estimation of the IRE PDF should further improve A-PRM classification in the presence of IRE. Several promising registration algorithms have been proposed which seek to estimate an optimal transformation as well as its associated uncertainty [9, 11, 13–16]. In particular, the probabilistic registration framework proposed by Risholm et al. [9] may be an ideal candidate for improved local estimation of the IRE PDF.

While the present study focuses on voxel-wise classification accuracy, it would be interesting to investigate if improved voxel-wise classification also improves prediction of global treatment response measures like overall survival. The A-PRM classifies each voxel based on the expected value of a distribution of intensity dif-
ferences estimated for each voxel using a Gaussian error model. In contrast, the original PRM method classifies each voxel based on a single observed difference from this distribution. Therefore, provided that the Gaussian error model is a reasonable approximation to the underlying IRE distribution, we hypothesize that the A-PRM would more accurately reflect the true voxel intensity differences in the presence of IRE. However, as previously discussed, the Gaussian error model may not accurately reflect the true IRE distribution. Furthermore, global treatment response prediction using the PRM, and by extension A-PRM, involves an aggregate analysis of a large number of voxels (e.g. computing fractional volumes) and may already be robust to IRE. Therefore we hypothesize that the A-PRM may not significantly improve global treatment response prediction.

Our testing strategy involved comparing results against a “ground truth” PRM analysis performed on the initial non-rigidly registered data. Therefore resulting analyses may have been influenced by residual IRE from the non rigid registration. An alternative approach would involve artificially creating the follow-up functional images by simulating treatment related functional changes in the baseline images. While this would ensure perfect alignment, it would be challenging to simulate realistic treatment related tumour functional changes since they are not yet well understood. Therefore the former method was adopted as the closer approximation. While rigid IRE simulation results were reported here to demonstrate the efficacy of the A-PRM method, we have also performed non-rigid IRE simulations in parallel and observed very similar classification results. Therefore for the sake of clarity, we focused on rigid IRE within the present proof-of-principle study. However, owing to the complexity of non-rigid IRE simulation, a follow-up study investigating non-rigid IRE may be appropriate.
3.6 Conclusion

The A-PRM was demonstrated to help visualize and quantify the influence of IRE on parametric response map analysis. The augmented method provided additional contextual information including image registration uncertainties which could facilitate the definition of targets for locally adaptive radiotherapy (e.g. sub volume boosts). Upper and lower bound response map estimates show promise as useful tools for probing the local influence of anticipated IRE on analysis. The A-PRM method should facilitate reliable use of PRM analysis for guiding adaptive RT by allowing the user to identify in advance whether a patient’s unique IRE-related PRM analysis uncertainty has the potential to influence RT target delineation.

3.7 Acknowledgments

We would like to thank Dr. Glenn Bauman and Dr. Michael Lock for providing invaluable assistance in acquiring the patient data used in this study. This work was funded by the Ontario Research Fund for the Ontario Consortium for Adaptive Interventions in Radiation Oncology (OCAIRO).
References


A generalized approach towards multi-parametric response mapping using principal component analysis *

4.1 Introduction

The development of early treatment response prediction methods represents an important step towards improving cancer care via personalized adaptive treatment strategies. Once identified, ineffective treatments could be adapted, halted, or alternative therapies could be considered to improve patient outcomes or reduce toxicities. Such an approach may be of particular interest for aggressive cancers such as glioblastoma multiforme where there is a limited therapeutic window and treatments may be associated with significant toxicity [1].

Parametric response mapping (PRM) has emerged as a powerful image-analysis technique which can be used to predict early treatment response [2]. Typically, PRM analysis involves a voxel-wise comparison of longitudinally acquired and spatially-

*This chapter is adapted from the manuscript entitled “A generalized approach towards multi-parametric response mapping using principal component analysis.” This manuscript is in preparation for submission to the journal Medical Physics.
aligned functional images. Tumour voxels are classified as increasing, decreasing, or not changing in function. The fraction of the tumour volume associated with one of these classes may then be found to correlate with a global measure of response such as overall survival [2]. To date, the predictive utility of the PRM method has been demonstrated for a variety of different pathologies such as glioblastoma, head and neck cancer, breast cancer, hepatocellular carcinoma and chronic obstructive pulmonary disease (COPD) [2–7]. Due to the voxel-wise nature of PRM analysis, this technique may also have potential for guiding personalized locally adaptive interventions such as adaptive radiotherapy (RT).

There are several key advantages to the PRM method. First, it involves voxel-wise image analysis and incorporates spatial correspondences to describe the spatial heterogeneity in intra-tumour response. Second, the PRM enables intuitive visualization of this heterogeneity. Third, PRM analysis is straightforward to implement and interpret, providing a uniquely accessible and effective means of probing image data for treatment response biomarkers.

However to date the PRM method has been almost exclusively applied to longitudinally acquired pairs of single-parameter image data. Galban et al. [8] previously investigated the benefit of combining two independent single-parameter analyses for a cohort of patients diagnosed with glioblastoma. Independent PRM analyses were performed on diffusion and perfusion image data and then the resulting scalar-valued PRM biomarkers were combined via multivariate logistic regression. This two-variate approach was shown to improve prediction of overall survival when compared to PRM analysis of diffusion or perfusion data alone, highlighting the potential of multi-parametric response prediction. However, since the two PRM analyses were performed independently, spatial correspondences between the diffusion and perfusion data were not included in the analysis. Consequently, spatial-heterogeneity in multi-parametric
treatment response cannot be taken into account nor visualized. Incorporation of spatial correspondences between functional imaging at two time points was a key driver of the success of the original PRM method. Therefore, an analogous multi-parametric approach may have potential for improving treatment response prediction and assist in finding new treatment response biomarkers. A fully voxel-wise approach towards multi-parametric treatment response prediction using PRM may also facilitate investigation into voxel-wise treatment response prediction for the purposes of guiding locally adaptive interventions.

Here our goal was to develop a unified multi-parametric response analysis framework which extends the key advantages of the single-parameter PRM method to analysis of multi-parametric data. We introduce a generalizable $N$-dimensional approach towards multi-parametric response mapping (MPRM) which takes spatial heterogeneity in multi-parametric response into account and supports visualization and interpretation of this heterogeneity. For preliminary demonstration, the proposed method is applied to a multi-parametric image dataset acquired from a group of patients treated for high-grade glioma. Representative MPRMs and prediction of overall survival are demonstrated with comparison to single-parameter PRM analyses.

4.2 Materials and Methods

4.2.1 Original PRM: single-parameter response map

The original single-parameter PRM is generated through analysis of temporal voxel-wise intensity changes between two spatially registered image volumes within some region of interest (e.g. tumour). Longitudinally acquired functional images of the same type are typically analyzed. Voxel-wise changes are classified as increasing, decreasing, or not changing in intensity according to a single threshold value $T$ (de-
termination of $T$ discussed further below). If $\Delta I(x)$ represents the voxel-wise intensity change for a voxel located at position $x$ within the registered pair of image volumes, then the intensity change is classified as increasing or decreasing if $\Delta I(x) > T$ or $\Delta I(x) < -T$ respectively. Voxels where $|\Delta I(x)| \leq T$ are classified as not changing.

The PRM itself corresponds to an image of the region of interest (ROI) which indicates the classification of each voxel. The fraction of the ROI volume associated with one of the classes may then be found to correlate with a measure of global response (e.g. overall survival) among a patient group [2].

The cut-off threshold $T$ has been determined in several different ways. For example, Galban et al. [2] determined $T$ among a cohort of patients with glioblastoma by analyzing the functional changes in a region of normal tissue on the contralateral side of the brain with respect to the tumour. Follow-up functional image voxel values within the normal tissue ROIs were plotted against the spatially corresponding baseline functional image voxel values. $T$ was then defined as the 95% confidence interval (CI) in the residuals from a linear fit to this plot. Other studies have successfully used similar approaches towards threshold determination [3, 5, 8]. Receiver-operator characteristic (ROC) analysis has also been performed to find thresholds which maximize the predictive utility of the PRM method with respect to the outcome of interest [4]. More recently, the variability from single-session test-retest imaging as well as principal component analysis (PCA) of image intensity histograms have been used for threshold determination [7, 9].

The PCA-based method proposed within the present study represents a new approach towards response mapping when compared with the use of PCA in Zha et al. [9]. In Zha et al., PCA was applied to histograms of lung CT image intensities from a cohort of patients to derive single-parameter PRM thresholds to classify patients with airways disease and or emphysema. In contrast, we apply PCA directly to
co-registered multi-parametric image data (rather than image intensity histograms) within a normal tissue reference ROI establishing reference principal components (PC). We then classify candidate image data within a target ROI through projection onto the principal components established from the normal tissue reference PCA. The use of PCA within our proposed method is further discussed in section §4.2.2.1-4.2.2.3.

4.2.2 Multi-parametric response map

There are three major components in the MPRM method; i) application of PCA to image data within a normal tissue reference ROI, ii) classification of data within a target ROI (e.g. tumour) using the reference PCA, and iii) construction of the response maps.

The reference PCA represents the MPRM analogue of determining a set of PRM classification thresholds. The classification of the target data using the reference PCA corresponds to the PRM step where tumour voxels are classified according to the threshold determined from normal tissue. Finally, construction of the response maps involves creating visualizations of the classifications for each voxel. Details for each of these components follow in §4.2.2.1, §4.2.2.2, and §4.2.2.3 respectively.

The MPRM method can be applied to any multi-parametric image dataset. In practice, each patient’s data would likely include multiple pairs of longitudinally acquired multi-parametric imaging as a direct extension of the original PRM approach. For example, an eligible dataset could consist of diffusion and perfusion weighted MRI acquired before, during, and/or after treatment for each patient. However this is not an algorithm constraint and any set of images could potentially be analyzed provided that the same imaging (type, acquisition time points, protocol) is available for each patient. Similarly there are no algorithm-imposed constraints on the definition of the reference and target ROIs. However, for clarity in subsequent discussion of methods,
we will associate the reference ROI with normal tissue and target ROI with tumour or peritumoral tissue.

Image volumes must be spatially co-registered and resampled to the same array and voxel size prior to analysis. Since PCA is sensitive to the numerical range of the input image intensity values, all reference and target image data must be scaled to a common domain prior to MPRM analysis. Here, we linearly rescale each image so that its range of voxel values fall between 0 and 1. The potential limitations of this rescaling and PCA are explored in the discussion section.

4.2.2.1 Reference principal component analysis

The first step in the MPRM method is to use PCA to summarize the typical multi-parametric voxel values and variance within a normal tissue reference ROI. Figure 4.1 provides an example of a reference PCA as applied to MRI-derived apparent diffusion coefficient (ADC) and relative cerebral blood volume (rCBV) functional images acquired 1 month and 3 months post-RT for a patient treated for glioblastoma.

The reference PCA is analogous to the threshold determination step in the original single-parameter PRM method. For the original PRM method, follow-up scan voxel values were plotted against the spatially corresponding baseline scan voxel values within normal tissue ROIs. Each voxel is then represented by a point in a 2-dimensional Cartesian space. A line of best fit was used to describe the typical voxel value pairings and the 95% CI in the fit residuals was used to represent the typical variance about this line.

For multi-parametric data consisting of $N$ spatially registered image volumes, voxel intensity values from each of the different image volumes are similarly plotted along coordinate axes in an $N$-dimensional Cartesian space. PCA [10] can then be used to summarize the multi-parametric image data within a reference ROI. In the
PCA, each of the $N$ images is considered to be a variable and each $N$-tuple of spatially corresponding voxel intensity values within the reference ROI is considered to be a single observation of these $N$ variables. If $M$ is the number of voxels within the reference ROI, then the data matrix input into the PCA has size $M \times N$ where each column has had its mean value subtracted from it. The output of the PCA is a set of $N$ eigenvectors $\{v_i\}$ for $i = 1 \ldots N$ (principal components) which describe the primary directions of variance within the reference data plotted in the $N$-dimensional Cartesian space. A set of $N$ corresponding eigenvalues $\{\lambda_i\}$ for $i = 1 \ldots N$ represents the variance explained by each principal component (PC).

Figure 4.1 shows an example with $N = 4$ images (two images 1-month post-RT and two images 3-months post-RT). The reference data would normally be analyzed in a 4-dimensional Cartesian space with four PCs output from the reference PCA. It is possible to plot the data and PCs in 4D and then project into 3D for visualization. However, a clearer 3D visualization is obtained when we apply PCA to the first three images in figure 4.1 and plot the resultant data in 3D. Consequently, this is the approach taken in figure 4.1. In practice, PCA would be applied to all $N$ images.

When positioned at the $N$-dimensional mean of the reference data, the first PC corresponds to an $N$-dimensional linear fit to the plotted reference data, analogous to the line of best fit used to compute classification thresholds in the original PRM method. We denote this first PC by $v_1$, a vector in the direction of the line of best fit. The second PC ($v_2$) is orthogonal to this linear fit and the associated eigenvalue is equal to the variance of the data about the fit in the direction of the second PC. In effect, the second eigenvalue $\lambda_2$ is the sum of squares of the linear fit residuals in the direction of the second PC. This variance can then be used to determine classification thresholds similar to the original PRM method where the linear fit residuals were also used to define a threshold.
Figure 4.1: Example of a reference PCA as applied to MRI-derived apparent diffusion coefficient (ADC) and relative cerebral blood volume (rCBV) data acquired at 1 and 3-months post-RT for a patient treated for glioblastoma. The target ROI (T) is the tumour and the reference ROI (R) is normal tissue on the contralateral side of the brain. Spatially corresponding voxel values in R are plotted in a Cartesian space and then PCA is used to summarize typical multi-parametric voxel values and variance within R. To permit 3-dimensional visualization within this figure, only the first three images are analyzed and so there are only three PCs denoted by $v_1$, $v_2$, and $v_3$ respectively. The ellipses in the central plot span $\pm 2\sqrt{\lambda_1}$, $\pm 2\sqrt{\lambda_2}$, and $\pm 2\sqrt{\lambda_3}$ in the direction of the first, second and third PCs respectively (i.e. two standard deviations).
Higher order PCs \( (v_3, v_4, \ldots v_N) \) also describe a proportion of the variance of the data about the linear fit. However, due to the nature of PCA the proportion of the data variance described by higher order PCs drops off when there are correlations in the input data. Therefore thresholds determined from the second PC may summarize the majority of the variability in the reference data about the \( N \)-dimensional linear fit. Determination of classification thresholds is discussed further in section §4.2.2.2 and §4.2.2.3.

In aggregate, the output of the reference PCA consists of the set of PCs \( \{v_i\} \) positioned at the multi-dimensional mean of the reference data and the corresponding eigenvalues \( \{\lambda_i\} \) for \( i = 1 \ldots N \). Similar to the original PRM method, patient-specific or population-based thresholds can be determined by including the data from one or more patients in the reference PCA.

4.2.2.2 Target data classification using the reference PCA

The second step in the MPRM method is to use the normal tissue reference PCA to classify the multi-parametric data within a candidate patient’s target ROI. This corresponds to the step in the original PRM method where tumour voxels are classified by the threshold determined from normal tissue. Here, the goal is to classify the target data in terms of the variability observed along each reference PC. Figure 4.2 provides an overview of target data classification, expanding on the example considered in figure 4.1.

First, the multi-parametric voxel values within the patient’s target ROI are plotted in the same \( N \)-dimensional Cartesian space occupied by the reference data. The target data are then projected onto the reference PCs to determine the component of each data point along each of the PC axes. After projection, each target voxel is represented by a new set of coordinates in a rotated \( N \)-dimensional space spanned by the PCs.
and centered about the $N$-dimensional mean of the reference data. Since the origin of
the space is the $N$-dimensional mean of the reference data, a target voxel’s coordinate
along a PC axis describes the difference between its multi-parametric value and the
reference data mean (in the direction of the PC axis). Therefore higher magnitude
coordinates identify target voxels as increasingly different from the reference data.

To facilitate comparison of these differences with the variability observed in the
reference data, the coordinates are then normalized by the standard deviation of the
reference data along each PC axis. That is, if $x_i$ represents the coordinate of a voxel
along the $i$-th PC then, $x_i \rightarrow x_i / \sqrt{\lambda_i}$ where $\lambda_i$ represents the $i$-th PC’s associated
eigenvalue (variance) and $\sqrt{\lambda_i}$ is the standard deviation of the reference data along the
$i$-th PC. The normalized coordinates describe the difference between the target data
and the mean of the reference data in units of the reference data standard deviation.
For example, normalized voxel coordinates of $(1, -2, 3)$ would indicate that the voxel
is $1, 2,$ and $3$ reference data standard deviations away from the reference data mean
in the direction of the first, second and third PCs respectively. The negative sign on
the second coordinate indicates that the voxel is positioned on negative side of the
PC axis with respect to the reference data mean.
Figure 4.2: Classification of the target data (red) using the reference PCA (ellipses) from the glioblastoma example from figure 4.1. The first, second, and third reference PCs are denoted by $v_1$, $v_2$, and $v_3$ respectively. The multi-parametric target data are plotted in the same space as the reference data, projected onto the reference PCs and then re-expressed in terms of PIs along each PC. Ellipses in the first two plots span $\pm 2\sqrt{\lambda_1}$, $\pm 2\sqrt{\lambda_2}$, and $\pm 2\sqrt{\lambda_3}$ in the direction of the first, second and third reference PCs respectively (i.e. two standard deviations).
After normalizing by the standard deviation of the reference data along each PC, the coordinates are re-expressed in terms of reference data prediction intervals (PIs). Using the previous example, the coordinate (1, -2, 3) can be re-expressed as (66.8, 95, 99.7) indicating that the target voxel’s multi-parametric intensity values fall within the central 66.8%, 95% and 99.7% of the reference data variation along each PC under the assumption of normally distributed reference data. The PIs can be computed by using the standardized target data coordinates to sample a unit normal distribution. Finally, we set the signs of the new PI-based coordinates to be equal to the signs of the earlier standardized coordinates to convey whether the voxel is positioned on the positive or negative side of each PC axis with respect to the reference data mean. That is, (66.8, 95, 99.7) → (66.8, -95, 99.7). The second coordinate, -95, indicates that the target voxel is within the 95% PI of the reference data along the second PC and is positioned on the negative side of the PC axis with respect to the reference data mean (i.e. with respect to the origin of the space).

4.2.2.3 Response map construction

The final step in the MPRM method is to construct response maps which visualize the target data that has been classified (re-expressed) in terms of the PIs from the reference PCA. In principle, \( N \) separate MPRMs could be constructed which visualize the PI classifications of the target voxels along each PC. That is, a separate MPRM could be constructed from the first coordinate of the re-expressed target data, the second coordinate of the re-expressed target data, and so on. Each of these MPRMs would classify and visualize the full distribution of the target data since every voxel is identified by its own PI-based coordinates. Figure 4.3 illustrates the construction of response maps expanding upon the example in figures 4.1 and 4.2.

For simplicity, a single MPRM should be proposed for prediction. We will discuss
Figure 4.3: Response map construction from the re-expressed target data in figure 4.2. Three separate MPRMs can be constructed, MPRM\(_{v_1}\), MPRM\(_{v_2}\) and MPRM\(_{v_3}\), which visualize the target voxel coordinates (PI classifications) along the principal component axes \(v_1\), \(v_2\), and \(v_3\) respectively. MPRMs are shown superimposed on the patient’s T1-weighted MRI scan.

candidates MPRM\(_{v_1}\), MPRM\(_{v_2}\) and maps built from higher order PCs. MPRM\(_{v_1}\) is built from target data coordinates along the first PC of the reference PCA. For each target voxel, it describes the distance between the position of that voxel and the mean of the normal tissue reference data within the \(N\)-dimensional Cartesian space (in the direction of the first PC). Therefore differentiation of responding and non-responding patients using MPRM\(_{v_1}\) would require one group’s target data to be clustered significantly closer or further away from the other group with respect to a single point (mean of the reference data) in the Cartesian space. However, early treatment response prediction is challenging precisely because responding and non-responding patients do not exhibit significant differences in the mean tumour voxel image intensities. Therefore the predictive potential of MPRM\(_{v_1}\) may be limited.

To address this challenge, the original single-parameter PRM method instead compared the variability in tumour functional images to variability observed in normal tissue as described by the cut-off threshold \(T\) (95% CI in linear fit residuals)[2]. An MPRM built from the second PC of the reference PCA provides an analogous
approach to the PRM. For each target voxel, $\text{MPRM}_{v_2}$ describes the distance between the position of that voxel (in the direction of the second PC) and the $N$-dimensional linear least squares fit defined by the first PC from the reference PCA. That is, $\text{MPRM}_{v_2}$ describes the variation of the target data about a linear fit to the reference data. Similar to the original PRM, $\text{MPRM}_{v_2}$ expresses (and classifies) this target data variation in terms of PIs for the residuals of the reference data about this linear fit. Therefore $\text{MPRM}_{v_2}$ provides a comparison of the variability of the target and reference data about the first PC (in the direction of the second PC).

MPRMs built from higher order PCs (i.e. $\text{MPRM}_{v_3} \ldots \text{MPRM}_{v_N}$) also compare the variation of the target and reference data about the first PC. However provided that the input data is not completely uncorrelated, the proportion of the reference data variance explained by higher order PCs drops off due to the nature of PCA. Consequently, $\text{MPRM}_{v_2}$ may represent a comparison of the target data variability to the majority of the reference data variability eliminating the need to consider higher order MPRMs. Therefore within this study we chose to investigate $\text{MPRM}_{v_2}$ for the purpose of treatment response prediction. We discuss the limitations of this approach and propose alternative methods incorporating multiple higher order PCs in the discussion section. Example MPRMs built from other PCs are also provided for the purposes of demonstrating the MPRM method.

4.2.3 Demonstration

For demonstration, MPRM analysis was applied to a multi-parametric image dataset acquired from a group of 19 patients treated for high-grade glioma. MRI-derived ADC and rCBV maps acquired at 1 and 3 months post-treatment were analyzed. Representative MPRMs and prediction of overall survival are demonstrated with comparison to single-parameter PRM analyses.
4.2.3.1 Patient and image data

This study was performed in compliance with institutional research ethics boards and informed consent was obtained from all patients. Multi-parametric imaging from a total of 19 patients diagnosed with World Health Organization grade IV glioblastoma (n = 17), grade III anaplastic oligodendroglioma (n = 1) and grade III anaplastic astrocytoma (n = 1) were analyzed. All patients were treated with surgery followed by radiotherapy (RT) with concurrent and adjuvant temozolomide chemotherapy. For RT, 18 patients received 60 Gy in 30 fractions while one patient received 45 Gy in 15 fractions with a 24 Gy stereotactic boost delivered in 3 fractions. The median age at the time of diagnosis was 65 years (range, 31-81 years) and median overall survival (OS) was 18.2 months (range, 4.7-41.6 months).

MR imaging acquired at 1 and 3 months post-radiotherapy was available for each patient. Scans were acquired with a 1.5 T Signa HDXT (GE Healthcare, Milwaukee, WI) or a 1.5 T Achieva scanner (Philips Medical Systems, Best, The Netherlands). Imaging at both time points included post-gadolinium axial T1-weighted, diffusion weighted, and dynamic-susceptibility contrast-enhanced (DSC) scans. ADC maps were generated from the diffusion-weighted images which were acquired with b-values of 0 and 1000 s/mm². CBV maps were generated from the DSC scans using the method by Boxerman et al. which corrects for contrast agent extravasation [11]. Each CBV map was converted to a rCBV map via normalization by the mean intensity within 15 white-matter ROIs placed on the contralateral side of the brain with respect to the tumour. Each white-matter ROI was circular and approximately 1 centimeter in diameter. Figure 4.4 shows example imaging for patients with short and long OS.
Figure 4.4: Post-gadolinium T1-weighted MRI, ADC, and rCBV maps at 1 and 3 months post-RT for a patient with (a) OS = 11 months and (b) OS = 41 months. The outline of PERIPH is indicated for each patient. The outline of the CEL is also implicitly shown by PERIPH. ADC maps are expressed in units of $10^{-3}\text{mm}^2/\text{s}$.
4.2.3.2 Image registration and segmentation

For each patient, all T1, T2, DWI, and DSC images were rigidly registered using 3D Slicer (BRAINS module)[12]. The scans for each patient were also resampled to a common image array and voxel size to facilitate voxel-wise analyses. The final image array size varied between patients depending on their individual scan protocol; however, voxel sizes were resampled consistently across patients with a final size of 1 mm by 1 mm by 5 mm in the left-right, anterior-posterior and superior-inferior directions respectively. ADC and rCBV maps were registered indirectly by applying the transformations output from registration of the source images from which they were derived.

After registration, the post-gadolinium T1-weighted images were used to delineate the contrast-enhancing lesion (CEL) at the 1 and 3-month post-RT time points. CEL contours were delineated by AL, revised by an experienced glioblastoma imaging researcher (TY), and then verified by a radiologist with 8 years of experience (YW). As shown in figure 4.4, an ROI encompassing a 1 cm isotropic margin around the periphery of the CEL (PERIPH) was also generated. The original CEL was excluded from the expanded ROI. PERIPH ROIs were then manually adjusted to ensure ventricles and regions outside of the brain were not included in the final ROIs. Finally, a spherical ROI (2 cm diameter) containing white matter was defined on the contralateral side of the brain with respect to the tumour for each patient. This ROI was defined independently of the white matter ROIs used to produce the rCBV maps.
4.2.3.3 Parametric response mapping

Table 4.1 provides a summary of the parametric response mapping performed for each patient. A total of eight PRMs and four MPRMs were generated per patient, including special analyses related to mitigation of image registration error (IRE) related analysis uncertainty. Discussion within this section is organized according to the headings of the table from left to right.

For all PRM and MPRM analyses, population based PRM thresholds and MPRM reference PCAs were generated using image data from the normal white matter ROIs delineated on the contralateral side of the brain with respect to the tumour. These ROIs were selected to be the MPRM reference regions since they included relatively homogeneous voxel intensities (plus noise). This increases the normality of the data input into the reference PCA when compared to alternative regions containing multiple tissue types with non-homogeneous image intensity values. While data normality is not a strict requirement of PCA, it improves the ability of PCA to describe the input data. Furthermore, we compute prediction intervals to classify target data based on an assumption of normality in the reference data residuals about the reference PCs.

Two target ROIs were independently investigated per patient. The spatially intersecting CELs defined at 1 and 3 months post-RT were first compared to the white matter ROI followed by the spatially intersecting PERIPH ROIs. PERIPH was investigated in addition to the CEL since increased rCBV values in the peritumoral region have previously been associated with decreased OS [13].

For each of the target ROIs, separate PRM analyses were performed based on the longitudinally-acquired pairs of ADC and rCBV image data. That is, a PRM analysis was performed on the 1 and 3-month ADC data (ADC\(_1\), ADC\(_3\)) followed by a separate PRM analysis of the 1 and 3-month rCBV data (rCBV\(_1\), rCBV\(_3\)). MPRM
analysis was performed using all of a patient’s 1 and 3-month ADC and rCBV images. Finally, for each combination of method, target, and image data, we investigated the effect of using Gaussian blurring to combat image registration error (IRE) related uncertainty [14]. Blurring was performed using an isotropic 3-dimensional Gaussian kernel where $3\sigma = \text{IRE}_{max}$ along each coordinate direction. For simplicity in the present discussion, we selected $\text{IRE}_{max} = 3$ mm which corresponds to an assumption that 99.7% of the IRE along each coordinate direction is within 3 mm or alternatively that 99.7% of the total IRE is within approximately 5 mm.
Table 4.1: Summary of parametric response mapping analyses performed. Subscripts 1 and 3 on CEL, PERIPH, ADC, and rCBV abbreviations denote 1 and 3-month post-RT respectively. CEL denotes the contrast enhancing lesion and PERIPH denotes a 1-cm-thick shell in the peritumoral region immediately around the CEL. To facilitate subsequent discussion, PRM analyses are labelled P1 through P8 and MPRM analyses are labelled M1 through M4.

<table>
<thead>
<tr>
<th>Method</th>
<th>Reference</th>
<th>Target</th>
<th>Image Data</th>
<th>IRE blurring (mm)</th>
<th>Analysis label</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRM</td>
<td>white matter</td>
<td>CEL$_1 \cap$ CEL$_3$</td>
<td>ADC$_1$, ADC$_3$</td>
<td>0</td>
<td>P1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rCBV$_1$, rCBV$_3$</td>
<td>3</td>
<td>P2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PERIPH$_1 \cap$ PERIPH$_3$</td>
<td>ADC$_1$, ADC$_3$</td>
<td>0</td>
<td>P3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rCBV$_1$, rCBV$_3$</td>
<td>3</td>
<td>P4</td>
</tr>
<tr>
<td>MPRM</td>
<td>white matter</td>
<td>CEL$_1 \cap$ CEL$_3$</td>
<td>ADC$_1$, ADC$_3$, rCBV$_1$, rCBV$_3$</td>
<td>0</td>
<td>M1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>M2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PERIPH$_1 \cap$ PERIPH$_3$</td>
<td>ADC$_1$, ADC$_3$, rCBV$_1$, rCBV$_3$</td>
<td>0</td>
<td>M3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>M4</td>
</tr>
</tbody>
</table>
4.2.3.4 Response prediction and analyses

Similar to previous studies [2–4, 6, 7], the fractional volume of each patient’s target ROI classified as increasing, decreasing, or not changing in function was computed for each of the PRM and MPRM analyses. As discussed in §4.2.2.3, only MPRMs built from the second PC of the reference PCAs were investigated for predictive potential.

For PRM analyses, the threshold used to classify voxels as increasing, decreasing or not changing in function was determined similar to Galban et al. [2], where the 95% CI in the linear fit residuals defined a single value of $T$ for the entire patient group. However, in addition we employed a leave-one-out strategy to find the CI-based threshold which maximized the Pearson correlation of PRM-derived fractional volume measures with overall survival. This threshold was then used to classify voxels as increasing, decreasing, or not changing in order to generate PRMs for each of the 19 patients. Since a leave-one-out analysis was used, PRM classification thresholds were patient-specific. The algorithmic details of the leave-one-out method used to determine the classification threshold for each patient are as follows:

1. Define threshold training set to include $n = 18/19$ patients leaving out one patient.

2. Generate 100 PRMs for each patient in the training set using the 0%, 1% . . . 99% CIs in the linear fit residuals as thresholds.

3. Compute the fractional volume classified as increasing, decreasing, or not changing for each patient’s 100 PRMs. Each threshold is then associated with an $18$ (number of patients) by $3$ array (number of PRM classes) array of fractional volume values.
4. For each threshold, compute the Pearson correlation between the fractional volume measures and overall survival within the training set (independently for each class) producing a 100 (number of thresholds) by 3 (number of classes) array of correlation values.

5. For each PRM class, select the threshold that elicited the highest correlation with overall survival (i.e. three thresholds selected). Each of the selected thresholds, when used to generate PRMs in the training set, maximizes the correlation of either the increasing, decreasing, or no change fractional volume measures with overall survival.

6. Compute three PRMs for the patient left out of the training set based on the three classification thresholds selected in step 5.

7. Repeat steps 1-6 leaving a different patient out of the training set each time resulting in 3 sets of 19 PRMs. Each set maximizes the correlation of either the increasing, decreasing, or no change fractional volume values with overall survival.

The n = 19 patients were then split into two groups, responders (n = 10) and non-responders (n = 9) according to whether OS $\geq 18.2$ months (median OS). One of the three sets of PRMs produced in step 7 was chosen for further analysis. The set which maximally separated the responding and non-responding patients’ fractional volume values according to repeated Wilcoxon rank-sum tests was selected.

To predict OS $\geq 18.2$ months, a simple threshold approach was employed whereby patients were classified as responders or non-responders according to whether their fractional volume measures were above or below a cut-off value which was varied in an ROC analysis. Classification sensitivity, specificity, accuracy and area under
the curve (AUC) values were computed for each analysis. All ROC analyses were performed using SPSS (IBM SPSS, version 20.0, Chicago, IL).

MPRMs were converted to 3-class maps (i.e. increasing, no change, decreasing) by selecting a single PI along the 2nd PC as a cut-off threshold. Similar to the PRM algorithm detailed above, a leave-one-out optimization was performed on the MPRMs to find a threshold which maximized correlation of MPRM-derived fractional volume measures with overall survival. Within the context of these MPRM analyses, increasing or decreasing implies that a voxel’s multi-parametric voxel values are located on the positive or negative side of the 2nd PC axis with respect to the reference data mean. Prediction of OS ≥ 18.2 months using the MPRMs was then performed in the same way as PRM-based prediction (i.e. simple threshold classification and ROC analysis).

4.3 Results

Multiple PRMs and MPRMs were generated from the ADC and rCBV data for the n = 19 glioma patients according to the analyses detailed in section 4.2.2.3 and table 4.1. Figure 4.5 illustrates examples of single-parameter PRMs for the two patients shown in figure 4.4. For demonstration, the 95% CI in the linear fit residuals was used to define a significance cut-off threshold for this figure. Separate analyses were performed in the intersection of the 1 and 3-month post-RT CEL (first row) and the intersection of the 1 and 3-month post-RT PERIPH (second row). The intersection of the 1 and 3-month CELs in figure 4.5a appears to exclude centrally located CEL. However, after careful inspection (e.g. windowing and leveling, comparison to adjacent slices) the radiologist judged this central region to be non-enhancing based on their experience.
Figure 4.5: Original single-parameter PRMs generated from the ADC and rCBV image data from figure 4.4 for two patients with (a) short and (b) long OS. Analysis was performed within $\text{CEL}_1 \cap \text{CEL}_3$ (first row) and $\text{PERIPH}_1 \cap \text{PERIPH}_3$ (second row) and superimposed on the 1 month post-RT T1-weighted scans.

For all MPRM analyses, the multi-parametric image data within white matter ROIs defined on the contralateral side of the brain of all 19 patients were used to perform population-based reference PCAs. Table 4.2 summarizes each of the PCs and associated eigenvalues from this reference PCA.

MPRMs generated for the two patients from figure 4.5 are shown in figure 4.6. Similar to the didactic example from figure 4.3, maps generated from the PI classifications along each PC of the reference PCA are shown for demonstration. $\text{MPRM}_{v_1}$, $\text{MPRM}_{v_2}$, $\text{MPRM}_{v_3}$, and $\text{MPRM}_{v_4}$ correspond to the maps generated from classification along the first through fourth PCs of the reference PCA respectively. Qualitatively, $\text{MPRM}_{v_2}$, and $\text{MPRM}_{v_4}$ appear notably different in figure 4.6 between the patients with short and long OS.

The fractions of the target ROIs (CEL and PERIPH) classified as increasing, decreasing, or not changing were computed for each of the PRM and $\text{MPRM}_{v_2}$ analyses.
Figure 4.6: MPRMs generated from the 1 and 3-month post-RT ADC and rCBV image data from 4.4 with (a) short and (b) long OS. Analysis was performed within $CEL_1 \cap CEL_3$ (first row) and $PERIPH_1 \cap PERIPH_3$ (second row) and superimposed on the 1 month post-RT T1-weighted scans.
Table 4.2: Reference PCA principal component vectors \((v_1 \ldots v_4)\) and corresponding eigenvalues.

<table>
<thead>
<tr>
<th></th>
<th>(v_1)</th>
<th>(v_2)</th>
<th>(v_3)</th>
<th>(v_4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADC(_1)</td>
<td>0.47</td>
<td>-0.57</td>
<td>-0.12</td>
<td>-0.66</td>
</tr>
<tr>
<td>ADC(_3)</td>
<td>0.47</td>
<td>-0.45</td>
<td>-0.12</td>
<td>0.75</td>
</tr>
<tr>
<td>rCBV(_1)</td>
<td>0.55</td>
<td>0.63</td>
<td>-0.54</td>
<td>-0.05</td>
</tr>
<tr>
<td>rCBV(_3)</td>
<td>0.50</td>
<td>0.26</td>
<td>0.82</td>
<td>-0.02</td>
</tr>
<tr>
<td>(\lambda) ((10^{-2}))</td>
<td>3.14</td>
<td>1.51</td>
<td>0.72</td>
<td>0.22</td>
</tr>
<tr>
<td>variance explained</td>
<td>56%</td>
<td>27%</td>
<td>13%</td>
<td>4%</td>
</tr>
</tbody>
</table>

with the different image and target ROI combinations detailed in table 4.1. Figure 4.7 shows the distribution of PRM and MPRM\(_{v_2}\) fractional volumes among responding (R, OS \(\geq\) 18.2 months) and non-responding (NR, OS < 18.2 months) patients. Only the class (i.e. increasing, decreasing, no change) which was found to maximally separate the NR and R patients groups is shown for each of the 12 analyses.

Fractional volume measures were found to be significantly different \((p < 0.05, \text{Wilcoxon rank sum test})\) between R and NR patient groups for the P7 (PRM\(_{rCBV}\), PERIPH, no blurring), M2 (MPRM, CEL, blurring), M3 (MPRM, PERIPH, no blurring), and M4 (MPRM, PERIPH, blurring) analyses. \(p\)-values for analyses P1 (PRM\(_{ADC}\), CEL, no blurring), P5 (PRM\(_{rCBV}\), PERIPH, no blurring), and M1 (MPRM, CEL, no blurring) were found to be 0.08, 0.07, and 0.07 respectively while \(p \geq 0.1\) for the remaining analyses. However, it should be noted that a total of three statistical tests were performed for each of the analyses shown in figure 4.7 in order to select and subsequently show the class (increasing, decreasing, no change) which maximally separated the responding and non-responding patients. Upon applying a Bonferroni correction to correct the significance level for multiple comparisons, fractional volume measures were only found to be significantly different for analyses M2, M3, and M4.
Figure 4.7: Summary of (a) CEL and (b) PERIPH analyses from table 4.1. The distribution of fractional target volumes classified as increasing (red boxplots), decreasing (blue boxplots) or not changing (green boxplots) for the non-responding (NR, OS < 18.2 months) and responding (R, OS ≥ 18.2 months) patient groups are shown. Only the class which maximally separated the NR and R patients groups is shown for each analysis. Error bars encompass the central 95% of each distribution. Each pair of boxplots is labelled according to the analyses described in table 4.1.

ROC analyses were applied to each of the distributions in figure 4.7 for which p < 0.1 in order to predict OS ≥ 18.2 months. For prediction, a simple threshold approach was employed whereby patients were classified as responders or non-responders according to whether their fractional volume measures were above or below a cut-off value which was varied in the ROC analysis. Figure 4.8 shows the ROC curves and indicates the area under the curve (AUC) values for each analysis.

Similar to the analysis of the fractional volume distributions in figure 4.7, only P7, M2, M3, and M4 ROC analyses were found to be significantly predictive of OS ≥ 18.2 months (p < 0.05). p-values for analyses P1, P5, and M1 were found to be 0.07, 0.06, and 0.06 respectively. Table 4.3 provides an overview of the classification performance for the four significant ROC analyses. Maximum sensitivity, specificity, and accuracy values are reported in table 4.3 and correspond to the points on the figure 4.8 ROC curves that are closest to the top left corner of the ROC plots. Fractional volume cut-
off thresholds for analyses M3 and M4 were both 1% suggesting that non-responding patients do not tend to exhibit decreases in multi-parametric intensity values within PERIPH as evaluated in the direction of the second reference PC.

Figure 4.8: Prediction of OS ≥ 18.2 months based on (a) CEL and (b) PERIPH fractional volume distributions from figure 4.7. * indicates AUC values that are significantly different from 0.5 at the p = 0.05 level.
Table 4.3: Summary of classification performance for the statistically significant ROC analyses in figure 4.8. The fractional volume cut-off value (fx vol cut-off) which produced the reported maximum classification sensitivity, specificity and accuracy is also shown.

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Method</th>
<th>Data</th>
<th>Target</th>
<th>IRE blurring</th>
<th>Class</th>
<th>Fx vol cut-off</th>
<th>Sens (%)</th>
<th>Spec (%)</th>
<th>Acc(%)</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>P7</td>
<td>PRM</td>
<td>rCBV</td>
<td>PERIPH</td>
<td>no</td>
<td>dec</td>
<td>0.50</td>
<td>70</td>
<td>89</td>
<td>79</td>
<td>0.78</td>
</tr>
<tr>
<td>M2</td>
<td>MPRM</td>
<td>all</td>
<td>CEL</td>
<td>yes</td>
<td>nc</td>
<td>0.93</td>
<td>90</td>
<td>78</td>
<td>84</td>
<td>0.82</td>
</tr>
<tr>
<td>M3</td>
<td>MPRM</td>
<td>all</td>
<td>PERIPH</td>
<td>no</td>
<td>dec</td>
<td>0.01</td>
<td>80</td>
<td>78</td>
<td>79</td>
<td>0.83</td>
</tr>
<tr>
<td>M4</td>
<td>MPRM</td>
<td>all</td>
<td>PERIPH</td>
<td>yes</td>
<td>dec</td>
<td>0.01</td>
<td>90</td>
<td>78</td>
<td>84</td>
<td>0.84</td>
</tr>
</tbody>
</table>
4.4 Discussion

The development of effective treatment response prediction is an essential step on the path towards personalized cancer treatment. PRM analysis has been shown to be predictive of early treatment response within many contexts [2–7] and may be advantageous for guiding locally adaptive treatments such as sub-volume boosting. However, the PRM method has been almost exclusively applied to longitudinally acquired single-parameter image data. Here, our goal was to present a PRM analysis framework which enables multi-parametric response prediction while maintaining the key advantages of the original PRM method. PCA was first applied to the multi-parametric data within a normal tissue reference ROI. The multi-parametric data within a target ROI (e.g. tumour) were then projected onto the reference PCs in order to classify the target data as increasing, decreasing, or not changing according to the multi-parametric variance within normal tissue reference ROIs.

The key advantages of the original PRM method are that it i) takes into consideration the spatial heterogeneity in tumour response through voxel-wise analysis, ii) enables intuitive visualizations of this heterogeneity and iii) provides an accessible and effective means of probing image data for treatment response biomarkers. The MPRM was designed to reproduce these key elements while extending PRM to the multi-parametric context. Employing PCA, the MPRM classifies each target voxel according to the combination of all $N$ multi-parametric image intensity values that are associated with that voxel. That is, each target voxel is classified according to its unique position with an $N$-dimensional coordinate space that is defined and normalized by the normal tissue reference PCA. In this way the MPRM takes into consideration the spatial heterogeneity in multi-parametric tumour response. As shown in figure 4.6, the spatial heterogeneity in multi-parametric tumour response
can then be intuitively visualized and interpreted via maps generated from each of the reference PCs. Each map expresses the target data heterogeneity in terms of the reference data heterogeneity along each PC using PIs. The MPRM algorithm has also been made publically available via the MATLAB file exchange to mirror the accessibility of the original method.

The MPRM acts as a filter which re-expresses, distils, and visualizes the $N$-dimensional target data in terms of a standardized context of interest. This is advantageous since it can be difficult to directly conceptualize and compare $N$-dimensional data. An MPRM built from the first PC of the reference PCA demonstrates the difference between the multi-parametric target data and the $N$-dimensional mean of the reference data along the direction of the first PC. An MPRM built from the second PC of the reference PCA compares the variability of the target data to variability of the reference data about the first PC. Further intuition may be obtainable by investigating the image weights which compose the reference PCs (e.g. table 4.2). For example, if a reference PC is dominated by a specific image volume and MPRMs built from that reference PC are predictive, then prediction is largely driven by variation in the dominant image volume.

While PCA has been discussed and used for different purposes in previous PRM studies, to our knowledge it has never been employed for the purposes of computing and applying classification thresholds to multi-parametric image data as in the present study. Galban et al. [6] applied PRM analysis to registered inspiration and expiration lung CT scans to evaluate the extent of functional small airways disease and emphysema among a cohort of patients with COPD. PCA was subsequently applied to the registered CT scans to demonstrate that airways disease precedes emphysema in the progression of COPD, however it was not used to compute thresholds for the PRM analysis. Boes et al. [15] previously suggested that PCA could be used
to derive single-parameter PRM thresholds from single-session test-retest image data though this strategy was not employed. Finally, Zha et al. [9] computed a PRM threshold from PCA of voxel value frequency histograms obtained from registered inspiration and expiration lung CT scans. In contrast, the MPRM method applied PCA directly to multi-parametric image data within reference ROIs rather than to image frequency histograms in order to determine classification thresholds. These thresholds were then applied by projecting and re-expressing the target data in terms of the normal tissue variability along the reference PCs. Consequently, target data variation is defined and classified relative to and in the direction of the reference PCs. Furthermore, the target data are classified by a series of PI-based thresholds rather than a single cut-off threshold.

To demonstrate preliminary efficacy, we applied PRM and MPRM analysis to ADC and rCBV data acquired from a group of 19 patients treated for high-grade glioma. The MPRM with IRE blurring was found to be the only significant predictor of OS within the CEL. The MPRM was also found to be a significant predictor of OS for PERIPH analyses both with and without IRE blurring. This result supports the idea that functional changes in peritumoral regions may be predictive of treatment response. In comparison, the PRM was only found to be significantly predictive of OS when applied to rCBV data within PERIPH and without IRE blurring. The MPRM provided improved prediction of OS when compared to PRM for analysis of both the CEL and PERIPH ROIs as shown by the increased AUC values in Figure 4.8 and classification statistics in table 4.3. This further demonstrates the potential benefit of multi-parametric response prediction and indicates some preliminary utility for the MPRM method.

The AUC values within this study also compare favourably to Galban et al. [2] where PRM analysis of rCBV data within the intersecting CEL among 44 patients
with glioblastoma yielded an AUC of 0.754. Interestingly, in the present study the fractional volume classified as decreasing in rCBV by the original PRM method was predictive of OS within the PERIPH region but not within the intersecting CEL as was the case in the study by Galban et al. [2]. This difference is likely due to our analysis of 1 and 3-month post-RT data versus the pre-RT and 1-week mid-treatment data analyzed by Galban et al.[2].

Overall, IRE blurring was found to preferentially improve MPRM analyses while deteriorating PRM analyses. IRE between image volume pairs of the same image type (PRM analyses) may be smaller than the IRE between multiple image volumes of different types (MPRM analyses) and so the IRE blurring kernel may have over-compensated for the influence of IRE on PRM analyses reducing predictive efficacy. Alternatively, the use of a zero-mean isotropic blurring kernel (i.e. IRE probability distribution function) may be better suited for application to multiple images. The cumulative distribution of IREs among four co-registered images may be less biased in a single direction (i.e. more isotropic) than for a single pair of images since there are 6 (4 choose 2) unique pairs of spatial correspondences each with their own IRE distribution.

There are several limitations to the present study. First, the MPRM is subject to the limitations of PCA such as sensitivity to data scaling. The first few PCs are dominated by the primary sources of variation within the data. However large variations in a specific type of functional image data may or may not be as important for prediction as small variations in other types of functional image data. Therefore the MPRM may not be able to capitalize on image data with low variation but high predictive potential when analyzed alongside other high-variation image data. As a first step towards minimizing this effect, we first normalized each image volume so intensity values fell between 0 and 1. While this reduces the differential variance due
to differing imaging modalities, differential variance will still remain. An alternative scaling approach could involve standardizing each image volume using the mean and standard deviation of the image intensity values within the reference ROI (i.e. computing z-scores). This would ensure that the reference data in each image volume has zero mean and unit standard deviation. In a follow-up analysis we used this alternative scaling method and found that MPRMs built from the fourth PC were predictive of OS with AUC values of 0.83 and 0.79 for CEL and PERIPH analyses respectively.

For clarity we primarily investigated the predictive utility of MPRMs built from the second PC of the reference PCA. However, MPRMs built from higher order PCs could also be useful for response prediction. For example in figure 4.6, MPRM_{v4} also appeared notably different between the patients with long and short OS suggesting that MPRM_{v4} may have possessed some predictive utility. Our follow-up analyses which used the alternative scaling procedure also suggested that maps built from the fourth reference PC were predictive. Ultimately, PCA is a purely descriptive technique and does not reveal the directions of multi-parametric image variance which maximally distinguish between responding and non-responding patients. Therefore, there is no a-priori reason to expect that MPRMs built from different reference PCs will be more or less predictive than one another. A maximally predictive direction could potentially be obtained through regressing the reference PCs on the outcome of interest (e.g. overall survival) using a technique called principal component regression. This direction (vector), defined by a linear combination of the reference PCs, could then be used to produce a single MPRM for each patient. The incorporation of principal component regression into the MPRM framework will be the primary focus of future development.
In addition to comparison with the original PRM method, we could have compared the predictive utility of our method to the bivariate logistic regression approach that was used by Galban et al. [8] to combine independent PRM response biomarkers. However, our investigations in figures 4.7 and 4.8 revealed that there were no statistically significant PRM response biomarkers for analysis of the CEL and only one significant PRM biomarker for PERIPH. Therefore we omitted a comparison with this bivariate logistic regression approach as there were no pairs of statistically significant PRM biomarkers to combine within the present study.

Finally, our preliminary demonstration of MPRM efficacy involved analysis of a relatively small number of patients (n = 19). However, this patient number is consistent with another preliminary PRM study [4] and we also employed leave-one-out cross-validation during predictive analyses to mitigate over-fitting. Nonetheless, the predictive efficacy of the MPRM method should still be verified among larger patient populations.
4.5 Conclusion

We have proposed a generalized approach towards multi-parametric response mapping using principal component analysis which preserves the key advantages of the original PRM method within the multi-parametric context. The algorithm takes into consideration the spatial heterogeneity in multi-parametric response, supports intuitive visualizations and interpretation, and was also made publically available to improve accessibility. The MPRM was shown to be significantly predictive of overall survival amongst a group of patients treated for high-grade glioma and also offered improved prediction compared to single-parameter PRM analyses suggesting the potential benefit of multi-parametric response prediction using the MPRM.

4.6 Acknowledgements

This work was funded by the Ontario Research Fund for the Ontario Consortium for Adaptive Interventions in Radiation Oncology (OCAIRO).
References


Chapter 5

Conclusion and Future Work

5.1 Summary of findings

5.1.1 Radiation dose-response relationships for liver tumours

Radiation-dose response relationships define the probability of tumour control as a function of prescribed radiation dose. An improved understanding of these relationships could help to adapt dose prescriptions to patients (i.e. personalize) in order to improve tumour control or reduce side-effects. Chapter 2 contributed to this body of knowledge by defining dose-response relationships for patients treated with SBRT for hepatocellular carcinoma (HCC) and colorectal liver metastases (MET). To our knowledge, this study was the first to explicitly model dose-response relationships for HCC patients and the second to do so for MET patients [1]. 50% and 90% probabilities of 6-month local control were estimated to be achievable by 2 Gy per fraction equivalent doses \((\alpha/\beta = 10\text{Gy})\) of 53 Gy and 84 Gy for the HCC group and 70 Gy and 95 Gy for the MET group, respectively. Results for the MET group also helped to support a previously reported MET dose-response relationship [1]. Overall, we found that a higher radiation dose was required to control MET tumours when compared to HCC and that RT provided improved tumour control for HCC patients when
compared to MET patients at our institution.

5.1.2 Augmented parametric response mapping

Guidance of next generation locally adaptive radiotherapy techniques such as sub-volume boosting requires image-based treatment response prediction. The PRM is an image-based method for prediction of overall treatment outcome (e.g. overall survival) which shows promise as a tool for guiding personalized locally adaptive radiotherapy (RT). However, image registration error (IRE) introduces uncertainty into this voxel-wise analysis technique which may limit its use for guiding RT. Chapter 3 proposed an augmented PRM method (A-PRM) to address this challenge. The original PRM method was extended to include an IRE-related PRM analysis confidence interval and also incorporated multiple graded classification thresholds to facilitate visualization. PRM and A-PRM analyses of CT-perfusion functional images with known simulated IRE were compared to analysis without simulated IRE to investigate the two methods in the presence of controlled IRE. The A-PRM was shown to help visualize and quantify IRE-related analysis uncertainty. The use of multiple graded classification thresholds also provided additional contextual information which could be useful for visually identifying adaptive RT targets (e.g. sub-volume boost regions). The A-PRM should facilitate reliable PRM guided adaptive RT by allowing the user to identify if a patient’s unique IRE-related PRM analysis uncertainty has the potential to influence target delineation.
5.1.3 Multi-parametric response mapping

Voxel-wise analysis of functional imaging acquired at two time points using the PRM has been shown to be an effective tool for early prediction of cancer treatment outcomes (e.g. overall survival) and may also be well-suited towards guiding personalized locally adaptive RT. However, the PRM method has been designed for and almost exclusively applied to analysis of longitudinally acquired pairs of single-parameter image data. Chapter 4 proposed a novel approach towards multi-parametric response mapping (MPRM) to address this challenge. The overall objective was to improve global treatment response prediction (e.g. overall survival) and facilitate future investigations into voxel-wise response prediction for guidance of locally adaptive RT. MPRM analysis was applied to a multi-parametric dataset acquired from a group of \( n = 19 \) patients treated for high-grade glioma with comparisons to original single-parameter PRM analysis. Separate PRM and MPRM analyses of the contrast-enhancing lesion (CEL) and 1 cm of peripheral tissue (PERIPH) were performed. The original single-parameter PRM was found to be significantly predictive of median overall survival only when applied to rCBV data within the PERIPH ROI (\( AUC_{PERIPH} = 0.78, p < 0.05 \)). The MPRM was found to be significantly predictive of median overall survival for both CEL and PERIPH analyses and offered improved prediction (\( AUC_{CEL} = 0.82, AUC_{PERIPH} = 0.84, p < 0.05 \)) suggesting the benefit of multi-parametric response prediction using the MPRM. The significant predictions of OS for PERIPH analyses also supported the idea that functional changes in peritumoral regions can be predictive of treatment response. In summary, the proposed algorithm accounted for spatial heterogeneity in multi-parametric response, supported intuitive visualizations, and was found to improve prediction of overall survival. To mirror the accessibility of the original PRM method, the MPRM algorithm was also made publically available via the MATLAB file exchange.
5.2 Limitations and Future Work

5.2.1 Radiation dose-response modelling

There are several key limitations to the dose-response study in Chapter 2. First, tumour control probability was modelled with respect to an endpoint of 6-month local tumour control. Due to patient-specific variability in follow-up scheduling, the tumour control endpoint was evaluated within a 6-month ± 1 month interval introducing uncertainty into our dose-response analysis.

This 6-month endpoint was appropriate for the patients with colorectal liver metastases since their median time to loss of local control was found to be approximately 6 months (183 days). 1-year local control has also previously been reported to be approximately 60% among this demographic [1]. However for HCC tumours, studies have reported 1-year local control rates of greater than 90% suggesting SBRT provides 6-month local control for most patients [2]. Therefore a modelling endpoint of 1-year local control or greater may be more beneficial to the community for informing HCC SBRT dose prescriptions. In our cohort of HCC patients, a 1-year local control rate of 65% was observed, lower than those reviewed by Dawson et. al. [2]. This could be due to the fact that more advanced HCC patients were referred for radiotherapy when tumours occupied a larger fraction of the liver, limiting our ability to prescribe higher dose without risking treatment-related toxicities.

While we demonstrated in Chapter 2 that there is a relationship between radiation dose and 6-month tumour control for primary and metastatic liver tumours, patients included in the study had a range of dose-fractionations and were not uniformly treated with SBRT type doses. Dose fractionations ranged from 2-10 Gy/fraction with a median of 4 Gy/fraction. Therefore the dose-response relationships determined in Chapter 2 could differ from those among a patient demographic treated with more
uniform dose-fractionations. Nonetheless, this study helps to motivate future dose-response studies.

Another limitation in Chapter 2 is that when imaging data were unavailable, tumour control was assessed using blood-based measurements of alpha-fetoprotein for HCC and carcinoembryonic antigen for colorectal cancer as surrogates for tumour response. While the utility of these measurements for assessing tumour response has been previously demonstrated [3, 4], direct imaging based measurements would have been preferable to access local control information in order to maintain consistency within the study.

The incidence of hepatocellular carcinoma in Canada and the United States is rapidly increasing [5, 6]. Therefore, investigation of radiation dose-response relationships for HCC could have an increasingly positive impact on cancer treatment. While the study in Chapter 2 provided some insight into these relationships, larger studies are needed to verify the single-institution trends that were observed. That is, multi-institution dose-response studies such as the one performed by Chang et al. [1] for colorectal liver metastases need to be undertaken to validate the dose-response relationship for HCC.

5.2.2 Augmented parametric response map

One of the primary limitations of the A-PRM method proposed in Chapter 3 relates to the use of a spatially invariant 3-dimensional Gaussian distribution to model the image registration error probability distribution function. First, the probability of IRE is lower near salient image features (e.g. edges, high-contrast regions) and higher away from these features and so IRE is spatially variable. Therefore in the future the A-PRM would ideally incorporate a spatially varying IRE PDF to more accurately describe this uncertainty. A first approach could involve spatially modulating the
width of the Gaussian IRE PDF according to multiple point estimates of the IRE throughout the analyzed ROI obtained from inspecting the positions of corresponding fiducials.

Second, while the width of the Gaussian IRE PDF could be customized, ultimately it represents a simplified first order approach towards describing IRE. IRE can vary in complex non-rigid ways and cannot be fully captured by this approach. In the future, incorporation of empirically determined IRE PDFs may be beneficial. Bootstrapping and probabilistic image registration frameworks which simultaneously estimate a registration solution and registration uncertainty [7, 8] could potentially be used to improve definition of the IRE PDF in this way. This approach may have the added benefit of not requiring the user to provide an explicit estimate of an IRE upper bound for each patient as was the case in Chapter 3.

The demonstration of the A-PRM involved measurements of the differential impact of simulated rigid IRE on voxel classification between the PRM and A-PRM methods. However in the future, the A-PRM should be applied to larger image datasets with verified predictive utility using PRM analysis. Both rigid and non-rigid IRE would be simulated in the image data and then the impact of IRE on both voxel classification and prediction of global treatment outcomes such as overall survival would be compared between the PRM and A-PRM methods. This approach would provide a more complete assessment of the impact of integrating IRE uncertainty into PRM analysis.

5.2.3 Multi-parametric response map

There are several important limitations to the MPRM study and associated areas for future work that should be discussed. First, our demonstration of efficacy involved analysis of a multi-parametric image dataset acquired from patients treated
for glioblastoma at 1 and 3-months post-RT. While predictive utility was demonstrated using this data, ideally the MPRM would be applied to a dataset which has already been shown to have predictive potential using the original PRM method. For example, application to ADC and rCBV functional images acquired before and during RT may better elucidate the potential benefit of MPRM prediction compared to the original PRM. Moreover, since the ultimate goal of this work is to provide tools for guiding locally adaptive RT, the MPRM should be applied to pre and intra-treatment image data to demonstrate response prediction during treatment. We did not perform a similar analysis in Chapter 4 because pre-treatment MRI was acquired prior to surgery and radiotherapy rather than after surgery but prior to radiotherapy.

Second, the MPRM is subject to the limitations of PCA. For example, PCA is sensitive to the scaling of input data. The first few PCs are dominated by the largest sources of variation within analyzed multi-parametric dataset. Therefore the MPRM cannot take advantage of image data with high predictive potential but low variability when analyzed alongside other image data with high variability. In the future, alternative scale-invariant forms of PCA could be integrated into the MPRM method [9]. However, in a follow-up analysis we found that standardizing each image by the mean and standard deviation of the reference ROI image intensities could offer a simpler approach. PCA also does not add value for analysis of highly uncorrelated image data since resulting PCs would fall along the original axes defined by each of the input images. Therefore the MPRM may be best suited towards analysis of longitudinally acquired image data where two or more images of the same type (i.e. highly correlated images) are included in analysis.

Finally, only MPRMs built from classification along the second PC of the reference PCA were primarily investigated for predictive utility in Chapter 4. MPRMs built from other PCs could have predictive value and should also be investigated.
Ideally, the MPRM should be improved in the future to produce a single response map or include an algorithm for selecting a single map most likely to have predictive potential. One approach towards building a single map could involve combining classifications along multiple reference PCs. For example, each voxel could be classified by the L2-norm (square root of the sum of each vector component squared) of its PI-based coordinates along multiple PCs. Alternatively, a single map could be selected for investigation by performing a second PCA on the target data. MPRMs could be built from the reference PC that most closely matches the direction of the first PC from the target PCA. This would ensure that MPRMs are built from the reference PC which maximally stratifies target data variation. However, both of these alternative approaches would not guarantee that the resulting MPRMs were built in a direction which maximally distinguishes responding and non-responding patients. Such a direction could potentially be found using principal component regression to determine a linear combination of the reference PCs along which MPRM analysis would maximally separate responders and non-responders.

5.2.4 Verifying the predictive utility of A-PRM and MPRM

The A-PRM and MPRM methods need to be applied to other functional image datasets to verify broader predictive utility and search for new biomarkers of treatment response. For example, a previous study has reported PRM predictive utility based on analysis of repeat arterial phase contrast-enhanced CT imaging of patients treated with trans-arterial chemoembolization for HCC [10]. A similar study could be performed using A-PRM analysis of arterial-phase imaging for patients treated with SBRT for HCC.

Parametric response maps predict global measures of response like overall survival based on voxel-wise image analysis. PRMs also indicate the tumour voxels associated
with this prediction. However, it has not yet been verified whether tumour voxels associated with a prediction of global response actually respond in a local sense. That is, do tumour sub-regions associated with prediction of favourable global response disappear after treatment? Conversely, do tumour sub-regions associated with prediction of unfavourable global response remain or grow after treatment? These relationships may need to be confirmed prior to using PRM methods (including the A-PRM and MPRM) to guide locally adaptive treatments such as sub-volume boosting or dose-painting by numbers. Such a study would require functional imaging of a patient group with which to make predictions and then frequent anatomical imaging during and after treatment to monitor changes in tumour size and shape.

5.2.5 Evaluating feasibility of response map guided radiotherapy

Once predictive utility has been verified, studies should be performed to investigate the technical feasibility of delivering radiation to adaptive RT targets defined by parametric response maps. These studies would be similar to the planning feasibility studies performed for hypoxia-guided RT dose escalation [11]. Adaptive RT plans would be generated for a cohort of patients based on a set of PRMs which have been shown to be predictive of treatment response. The planned dose to conventional targets, adaptive targets, and normal tissue would be reported for each patient to assess whether PRM-guided adaptive RT is feasible using current treatment planning and delivery systems. Depending on availability, TCP and NTCP models could also be used to estimate the impact of adaptive treatment on tumour control and normal tissue toxicity.
5.2.6 Treatment response prediction software tools

User-friendly image-based treatment response prediction software should be developed to encourage the broader use and clinical translation of the methods discussed in Chapters 2 and 3. The A-PRM and MPRM methods could be integrated into existing open-source frameworks such as 3D Slicer [12].

For example, an A-PRM module could incorporate slider bars in each of the three coordinate directions to allow the user to easily adjust the amount of IRE they anticipate within their analysis. Such a feature would be accompanied by real-time updates to A-PRM visualizations. This would allow the user to quickly probe patient-specific analysis uncertainty due to IRE and take steps to account for its impact on delineation of adaptive RT targets. The MPRM could be similarly incorporated into an analysis module.

As discussed in §4.4, the MPRM acts as a filter which re-expresses, distills, and visualizes $N$-dimensional multi-parametric data in terms of a standardized context of interest (i.e. the reference PCA). Therefore a user-friendly software implementation of the MPRM could also be advantageous for general exploration and description of multi-parametric data. An open-source library of reference PCAs (i.e. eigenvectors and eigenvalues) could be populated from MPRM analyses of different combinations of multi-parametric imaging, cancer types, and reference ROIs. Other investigators could then use these reference PCAs to build MPRMs for the purposes of analyzing individual patient’s multi-parametric data within standardized frames of reference.
References


Appendix A

Ethics Approvals
RESEARCH OFFICE REVIEW NO.: R-09-506

PROJECT TITLE: Hepatic lesion database

PRINCIPAL INVESTIGATOR: Dr. M Lock

DATE OF REVIEW BY CRIC: November 24, 2009

Health Sciences REB#: 16487E

Please be advised that the above project was reviewed by the Clinical Research Impact Committee and the project:

Was Approved

PLEASE INFORM THE APPROPRIATE NURSING UNITS, LABORATORIES, ETC. BEFORE STARTING THIS PROTOCOL. THE RESEARCH OFFICE NUMBER MUST BE USED WHEN COMMUNICATING WITH THESE AREAS.

Dr. David Hill
V.P. Research
Lawson Health Research Institute

All future correspondence concerning this study should include the Research Office Review Number and should be directed to Sherry Paiva, CRIC Liaison, LHSC, Rm. C210, Nurses Residence, South Street Hospital.

cc: Administration
Use of Human Subjects - Ethics Approval Notice

Principal Investigator: Dr. B. J. Fisher
Review Number: 15557
Review Date: April 21, 2009
Protocol Title: Feasibility Study of Intensity Modulated Radiotherapy for Treatment of Liver Metastasis or Hepatocellular Carcinoma with Assessment of Response by Dynamic Contrast-Enhanced Scanning
Department and Institution: Oncology, London Health Sciences Centre
Sponsor:
Ethics Approval Date: June 23, 2009
Expiry Date: January 31, 2015
Documents Reviewed and Approved: UWO Protocol and Letter of Information and Consent Form dated 06 May 2009
Documents Received: Information:

This is to notify you that The University of Western Ontario Research Ethics Board for Health Sciences Research Involving Human Subjects (HSREB) which is organized and operates according to the Tri-Council Policy Statement: Ethical Conduct of Research Involving Humans and the Health Canada/CIHI Good Clinical Practice Practices: Consolidated Guidelines; and the applicable laws and regulations of Ontario has reviewed and granted approval to the above referenced study on the approval date noted above. The membership of this REB also complies with the membership requirements for REB's as defined in Division 5 of the Food and Drug Regulations.

The ethics approval for this study shall remain valid until the expiry date noted above assuming timely and acceptable responses to the HSREB's periodic requests for surveillance and monitoring information. If you require an updated approval notice prior to that time you must request it using the UWO Updated Approval Request Form.

During the course of the research, no deviations from, or changes to, the protocol or consent form may be initiated without prior written approval from the HSREB except when necessary to eliminate immediate hazards to the subject or when the change(s) involve only logistical or administrative aspects of the study (e.g. change of monitor, telephone number). Expected review of minor change(s) in ongoing studies will be considered. Subjects must receive a copy of the signed information/consent documentation.

Investigators must promptly also report to the HSREB:
- a) changes increasing the risk to the participant(s) and/or affecting significantly the conduct of the study;
- b) all adverse and unexpected experiences or events that are both serious and unexpected;
- c) new information that may adversely affect the safety of the subjects or the conduct of the study.

If these changes/adverse events require a change to the information/consent documentation, and/or recruitment advertisement, the newly revised information/consent documentation, and/or advertisement, must be submitted to this office for approval.

Members of the HSREB who are named as investigators in research studies, or declare a conflict of interest, do not participate in discussion related to, nor vote on, such studies when they are presented to the HSREB.

Chair of HSREB: Dr. Joseph Gilbert

Ethics Officer to Contact for Further Information
[Box with names and contact information]

This is an official document. Please retain the original in your files.
"A primary brain tumours registry as a framework for promoting clinical research in Neuro-Oncology: Project of Emilia-Romagna region on Neuro-Oncology (PERNO). Valutazione delle capacità diagnostiche e prognostiche della tomografia computerizzata di perfusione."

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ditta/sponsor:</td>
<td>Programma di ricerca Regione-Università 2007-2009 (Area 1° - Innovative Research)</td>
</tr>
<tr>
<td>Codice identificativo del piano clinico generale:</td>
<td>Perno</td>
</tr>
<tr>
<td>Specialità medicinale (nome o sigla):</td>
<td>/ /</td>
</tr>
<tr>
<td>Principio/i attivo/i:</td>
<td>/ /</td>
</tr>
<tr>
<td>Codice CAS (ove disponibile):</td>
<td>/ /</td>
</tr>
<tr>
<td>Classe farmacologica di appartenenza:</td>
<td>/ /</td>
</tr>
<tr>
<td>Codice ATC proposto (secondo codifica OMS):</td>
<td>/ /</td>
</tr>
<tr>
<td>Codice ICD:</td>
<td>/ /</td>
</tr>
<tr>
<td>Fase della sperimentazione clinica:</td>
<td>/ /</td>
</tr>
</tbody>
</table>

191
COMITATO ETICO DELLA PROVINCIA DI FERRARA

Indicazione proposta: si tratta di uno studio volto alla valutazione delle capacità diagnostiche della Tomografia Computerizzata (TC) e a verificare se la TC di Perfusion (TCP) può essere utile per distinguere i tumori primitivi cerebrali da altri tipi di lesione, per guidare la biopsia stereotassica e la resezione chirurgica, per distinguere fra la recidiva tumorale e il tessuto reso non più vitale dalla radioterapia, per predire la evoluzione del tumore, per monitorare la risposta alla terapia radiante e alla chemioterapia. La ricerca è anche orientata a chiarire se la TCP può essere di aiuto nella programmazione pre- e post operatoria, nel caso in cui correli con i dati clinici e, quando possibile, con i reperti Risonanza Magnetica Perfusionale (MR-PWI).

Forma farmaceutica: / /
Via di somministrazione: / /
Durata dello studio: 3 anni.

Schema dello studio e posologia: il disegno dello studio prevede due fasi:

1) la fase pre-operatoria riguarderà i pazienti con lesione expansiva soprattentoriale isolata, indicativa di tumore cerebrale intra-assiale: in questa prima fase saranno inclusi tutti pazienti con masse solitarie cerebrali soprattentoriali ammesse nel centro di Ferrara con indagini TC e/o MRI convenzionali che suggeriscano la diagnosi di sospetto tumore primitivo cerebrale. Tutti i pazienti inclusi in questa parte della ricerca verranno sottoposti ad uno studio TCP e ad un’indagine MR-PWI subito dopo l’ingresso. In tutti pazienti sottoposti a rimozione chirurgica del tumore verrà eseguito uno studio TC standard con un somministrazione endovenosa di mezzo di contrasto iodato entro 48 ore dall’operazione, per verificare la presenza di eventuali residui tumorali.

2) La fase post-operatoria interesserà i pazienti con gliona a basso ed alto grado a sede sovratentoriale, confermato istologicamente dopo resezione chirurgica o biopsia stereotassica. In tutti pazienti ammessi alla seconda parte dello studio, il deterioramento clinico o l’aumento della dimensione della lesione riscontrato con risonanza magnetica, saranno ritenuti indicatori di recidiva tumorale.

Tutti i pazienti saranno sottoposti a controlli TCP e/o MR-PWI sequenziali nel tempo, programmati:
a) ogni 3 mesi dall’intervento chirurgico per il primo anno;
b) ogni 6 mesi nei due anni successivi.
COMITATO ETICO DELLA PROVINCIA DI FERRARA

In tutti i pazienti con gioma ad alto grado sottoposti a radioterapia e/o chemioterapia post-operatoria verranno effettuati controlli con TCP e/o MR-PWI longitudinali nel tempo programmati: a) a 1 e a 3 mesi dal termine del trattamento; b) ogni 3-6 mesi successivamente.

In fase sia pre-operatoria che post-operatoria la performance verrà valutata in tutti i pazienti selezionati per lo studio utilizzando le scale di valutazione clinica correntemente accettate e, in particolare, il Karnofsky performance status.

Gli esami TC e MRI standard saranno valutati da due medici neuroradiologi che non saranno a conoscenza dei risultati della TCP e della MR-PWI, mentre le mappe di TCP e di MR-PWI saranno generate da medici che non saranno a conoscenza dei dati clinici e del quadro radiologico dei pazienti. Al termine dello studio tutti i risultati emersi dalle indagini TCP verranno confrontati e correlati con i dati provenienti dagli esami PWI e con quelli clinico-epidemiologici, molecolari e neurofisiologici acquisiti nelle altre unità. Le indagini TCP e MR-PWI saranno eseguite in accordo con protocolli precedentemente convalidati nella letteratura scientifica specifica.

In tutti i pazienti reclutati nei diversi stadi della ricerca gli studi di TCP saranno realizzati secondo il metodo di “first pass bolus-tracking”, che analizza l’effetto prodotto dal primo passaggio di un bollo di mezzo di contatto attraverso il letto vascolare cerebrale, utilizzando un apparecchio TC a scansione singola o multipla equipaggiato con un software dedicato per TCP. Per ogni paziente, in ciascuno degli intervalli temporali stabiliti in fase pre- e/o post-operatoria, verranno generate mappe perfusionali di flusso ematico cerebrale (CBF) di CBV, di tempo di transit medio (MTT) e di flusso di permeabilità di superficie (PSF) per misurare la permeabilità microvascolare.

Studi MR-PWI - in tutti i pazienti inclusi nelle varie fasi del progetto gli studi di MR-PWI verranno eseguiti secondo la metodologia del “Dynamic Susceptibility Contrast” (DSC), che misura ancora una volta l’effetto determinato dal primo passaggio di un bollo di mezzo di contrasto attraverso il letto vascolare cerebrale, impiegando una unità di risonanza magnetica da 1.5 Tesla equipaggiata con un software dedicato per la PWI. Per ogni paziente, in ciascuno degli intervalli temporali stabiliti in fase pre- e/o post-operatoria, verranno generate mappe perfusionali di CBV. Se possibile, negli stessi pazienti e nelle medesime fasi dello studio, verranno anche create mappe di “volume transfer constant” (Ktrans), utili per la valutazione della permeabilità microvascolare, usando il metodo del “Dynamic Contrast-Enhanced (DCE)”

In ogni sezione in cui il tessuto tumorale senza o con accentuazione contrastografica sarà visibile, i parametri TCP e MR-PWI saranno misurati in due differenti regioni di interesse disegnate a mano libera sulla singola sezione della TC o MRI di base: 1) il tessuto tumorale con o senza accentuazione contrastografica; 2) un’area di tessuto nervoso apparentemente sano situata nell’emisfero controlaterale, disposta simmetricamente alla neoplasia e corrispondente alla zona lesionale.

Eventuale terapia concomitante: //

AIC in Italia: //
all’estero: //

Indicazioni all’AIC, posologia, vie di somministrazione e forme farmaceutiche autorizzate: //
COMITATO ETICO DELLA PROVINCIA DI FERRARA

Precedenti approvazioni/autorizzazioni alla sperimentazione per la stessa indicazione proposta: / / 

Obiettivo/i dello/degli studio/i:
Primari:
- verificare se l’approccio multiparametrico con TCP può servire per:
  1- distinguere i tumori primitivi cerebrali da altri tipi di lesione che simulano il tumore, quali le infezioni, le lesioni demielinizzanti tumescenti e le lesioni ischemiche;
  2- discriminare i gliomi da altre lesioni neoplastiche solitarie intra-assiali cerebrali come i linfomi e le metastasi;
  3- differenziare i gliomi ad alto grado da quelli a basso grado;
  4- guidare la biopsia stereotassica e la resezione chirurgica verso la porzione più aggressiva del tumore;
  5- diversificare fra la recidiva tumorale e la necrosi da radiazioni;
  6- predire l’evoluzione del tumore;
  7- monitorare la risposta alla terapia radiante e alla chemioterapia.

Secondari:
1. chiarire se la tecnologia TCP può essere di aiuto nella programmazione pre-chirurgica e nella pianificazione post-operatoria;
2. comprendere se esistano correlazioni fra i dati TCP e quelli clinici, molecolari e MR-PWI.

Tipologia dei soggetti da arruolare (specificare se pazienti o volontari sani):
- saranno arruolati pazienti di entrambi i sessi, maggiorenni. Nella prima fase pre-operatoria dello studio saranno inclusi tutti pazienti con masse solitarie cerebrali sopratenziali ammessi nel nostro centro con indagini TC e/o la MRI convenzionali che suggeriscono la diagnosi di sospetto tumore primitivo cerebrale. Nella seconda fase post-operatoria dello studio verranno inclusi tutti i pazienti con glioma a basso ed alto grado confermato istologicamente dopo resezione chirurgica o biopsia stereotassica.

Numero dei soggetti da arruolare:
lo studio prevede l’inclusione di almeno 100 pazienti nella prima fase pre-operatoria e di almeno 50 pazienti con glomi ad alto e basso grado di malignità e di almeno 30 pazienti con glomi ad alto grado di aggressività sottoposto a radioterapia in quella post-operatoria.
COMITATO ETICO DELLA PROVINCIA DI FERRARA

Informazione al candidato: mediante scheda informativa, nella quale si riportano notizie sulla natura, durata e scopo dello studio, nonché il rapporto rischio/beneficio.
L’informazione del paziente, in virtù della propedeuticità di tale fase, dovrà essere fornita in un momento formalmente distinto dal recepimento del consenso.
Si raccomanda che l’avvenuta informazione venga formalizzata su cartella clinica o su scheda personale del paziente (in alternativa su modulo che ne faccia parte integrante), riportando contestualmente data e firma del medico sperimentatore e dell’arruolando stesso.

Recepimento del consenso: mediante apposito modulo
Si raccomanda che il recepimento del consenso/dissenso avvenga in un momento formalmente distinto dalla fase informativa e ad essa successivo e venga formalizzato su cartella clinica o su scheda personale del paziente (oppure su modulo che ne faccia parte integrante).

Criteri di inclusione/esclusione:

 criteri di inclusione:
Fase pre-operatoria: in questa prima fase verranno arruolati nello studio tutti i pazienti che soddisferanno i seguenti criteri di inclusione:
- età maggiore o uguale a 18 anni;
- assenza di instabilità dei parametri clinici (agitazione, disorientamento, coma ed altri segni di grave ipertensione endocranica);
- assenza di segni riferibili a gravanza in atto;
- assenza di controindicazioni alla TC ed alla MRI comprese quelle riguardanti la somministrazione intravenosa di contrasto iodato o di Gadolinio chelato;
- disponibilità dei dati clinici, compresa una valutazione della severità della malattia mediante il Karnofsky performance status.
Fase post-operatoria: in questa seconda fase verranno inclusi:
- tutti i pazienti con glioma a basso ed alto grado confermato istologicamente dopo resezione chirurgica o biopsia stereotassica.

 criteri di esclusione:
Fase pre-operatoria: in questa prima fase verranno esclusi dallo studio tutti i pazienti con le seguenti caratteristiche:
- età minore di 18 anni;
- presenza di instabilità dei parametri clinici (agitazione, disorientamento, coma ed altri segni di grave ipertensione endocranica);
- presenza di segni riferibili a gravanza in atto;
- presenza di controindicazioni alla TC ed alla MRI comprese quelle riguardanti la somministrazione intravenosa di contrasto iodato o di Gadolinio chelato;
- mancata disponibilità dei dati clinici, compresa una valutazione della severità della malattia mediante il Karnofsky performance status.
COMITATO ETICO DELLA PROVINCIA DI FERRARA

Fase post-operatoria: in questa seconda fase verranno inclusi:
- tutti i pazienti con glioma a basso ed alto grado confermato istologicamente dopo resezione chirurgica o biopsia stereotassica.

Sorveglianza clinica: / /

Modalità di sospensione: è prevista l'interruzione dello studio qualora si verifichino eventi avversi, su decisione del medico sperimentatore e, comunque, per libera scelta e disposizione dell'arruolato.

Il Comitato Etico esprime parere favorevole allo studio proposto, ove siano soddisfatti i prerequisiti etici nei termini innanzi richiesti.
Si ricorda al proponente responsabile la necessità di comunicare alla Segreteria Tecnico-Scientifica del Comitato Etico la fine dello studio, nonché di trasmettere copia di eventuale/i pubblicazione/i ovvero del report finale.

Si dà atto che il Comitato Etico ha preso visione della seguente documentazione:
- n. 1 copia di lettera d'intenti;
- n. 1 copia di protocollo di studio;
- n. 1 copia di sinossi del protocollo di studio;
- n. 1 copia di foglio informativo per il paziente;
- n. 1 copia di modulo di consenso;
- n. 1 copia di lettera per il medico d base;
- n. 1 copia di approvazione definitiva del progetto di ricerca da parte dell'Agenzia Sanitaria Regionale;
- n. 1 copia di elenco dei centri partecipanti.

Il Presidente del Comitato Etico
(Prof. Adalberto Ciaccia)
Appendix B

Copyright Releases
Licence to Publish

Notice: You are the person first named as ‘Corresponding Author’ at ‘Add/Edit/Remove Author’ of the online process to submit a new manuscript (referred to as “You/Your”). We are the British Institute of Radiology (BIR) of 131–151 Great Titchfield Street, London W1W 5BB, UK (referred to as “We/Us/Our”).

You have indicated Your intention to upload the article more fully detailed at ‘Enter Title’ of the online submission process (the “Article”) to be considered for publication by Us. In order to publish Your article We need You to grant Us a licence to publish. This licence also sets out Your rights regarding use of Preprints, Author Generated Postprints and the Definitive Published Version of the Article (as defined below). Please read the terms of this licence carefully before uploading the Article.

Clicking on the button marked ‘I accept’ next to the question, "I have read and agree to abide by the terms and conditions of the British Institute of Radiology’s Licence to Publish" at the ‘Submissions Waiting for Approval by Author’ stage of the online submission process will be taken as assurance that You have read, agree to grant and have the right to grant this licence.

Definitions

“Preprints” – the un-refereed version of the Article;

“Author Generated Postprint” – Your personal copy of the revised version of the Article as accepted by Us;

“Definitive Published Version” – the citable version of the Article produced by Us after peer review, copy editing and print and electronic production.

1. By clicking the button marked ‘I accept' next to the question "I have read and agree to abide by the terms and conditions of the British Institute of Radiology’s Licence to Publish" at ‘Submissions Waiting for Approval by Author’ stage of the online submission process, and uploading the Article, You grant to Us for the full term of copyright in the Article and any extensions thereto:
   a) grant the BIR exclusive rights to reproduce and/or distribute this Article (in full or in part) and to publish the same throughout the world in any format and in all languages including without limitation in printed, electronic or other media now or hereafter known or developed;

   b) the non-exclusive right throughout the world to edit, adapt, translate, publish, reproduce, distribute and display the accompanying electronic supplementary material (“ESM”) in printed, electronic or any other medium and format whether now known or yet to be developed, and to authorise others (including reproduction rights organisations) to do the same.
2. You warrant that:

a) the article is your original work, has not been previously published and is not under consideration by another journal;

b) in the case of a multi-authored Article, You have obtained written authorisation from all the co-authors of the Article (if any) to grant this licence to Us on their behalf as their agent, and You will supply a copy of the same to Us if We so request;

c) in respect of any material included in the Article in which copyright is not owned by the Authors, You have made proper acknowledgement in the Article, and You have obtained all necessary written authorisations for the reproduction and distribution of these materials as part of the Article throughout the world, in all languages and in all media and formats whether now known or yet to be developed and You will supply a copy of the same to Us if We so request;

d) if copyright in the Article is owned by any third party, whether Your employer or someone to whom You have assigned Your rights, You have obtained written authorisation from such copyright owner to grant this licence to Us on their behalf as their agent and will supply a copy of the same to Us if We so request;

e) all the Authors have made substantive contributions to the Article and assume full responsibility for its content.

f) You have read and have complied with the current ‘BJR Instructions to Authors’;

g) You have declared or will accurately declare all competing interests;

h) there has been no undisclosed computer manipulation of radiological images included in the Article;

i) We shall be entitled to examine raw data from which information contained in the Article has been derived; and

j) the Article does not contain anything which is obscene, defamatory, libellous, infringes any right of privacy or any intellectual property right (including without limitation rights in patents, copyright or trade marks) or any other rights of any person or entity, or is otherwise unlawful.

3. You assert Your moral right to be identified as the Author or co-author of the Article (as applicable). If Your Article is published, We will provide You with a PDF copy of the published Article.
4. If You wish the Article to be made available online, free of charge, to the public immediately on publication, under a Creative Commons licence* through our BIR Open option,

You shall pay to Us the relevant fee and We shall make the Article so available from the later of the date of receipt of the relevant fee or the date of first publication of the Article.

5. You retain the right to use the Article in the following non-commercial ways, provided that you acknowledge the Definitive Published Version of the article by placing the full bibliographic reference and URL of the relevant journal homepage close to the title of the Article:

a) In relation to the Preprint, Author Generated Postprint and Definitive Published Version of the Article, You are free to: make copies for Your own personal use; use the Article for the internal teaching purposes of Your own institution or company; and make and distribute copies (including through e-mail) of the Article to research colleagues, for personal use by such colleagues on a non-commercial, non-systematic basis.

b) In relation to the Preprint version only, You are free to post it on web sites, including electronic preprint servers. When the Definitive Published Version of the article is published the Author must acknowledge it by placing the full bibliographic reference and URL of the relevant journal homepage close to the title of the Article.

c) In relation to the Author Generated Postprint only, You are free to: post it on Your personal or institutional web site, load it onto an institutional or not for profit repository provided that a link to the Definitive Published Version is included, no sooner than 6 months after Publication; use it in printed compilations of Your work subsequent to publication of the Definitive Published Version of the Article, expand the Article into book-length form, and/or otherwise re-use portions of the Author Generated Postprint of the Article in other works as defined in Clause 6. You are also free to present the Article at a meeting or conference and to disseminate copies of such Article to the delegates attending such meeting or conference and/or to use the Author Generated Postprint in a thesis or dissertation (provided that this is not to be published commercially).

6. You may use individual figures or tables and continuous sections of the text (under 350 words) without permission in other publications by a third party. Use of more than this number will require express permission from Us.

7. You retain copyright in the Article and any accompanying ESM. However, You authorise Us to act on Your behalf to defend Your copyright in the Article should anyone infringe it, and to retain half of any damages awarded after deducting Our costs.
8. We are entitled to assign Our rights under this licence to any third party without giving notice to You.

9. No change or modification of this licence will be valid unless confirmed in writing by Us.

10. Failure or delay by Us to exercise any right or remedy under this Agreement shall not be deemed to be a waiver of that right or remedy, or prevent Us from exercising that or any other right or remedy on any occasion.

11. This licence is governed by English law and the parties hereby submit to the non-exclusive jurisdiction of the English courts.

12. This license is terminated in the case that the article is rejected for publication or the author withdraws the article for consideration for publication before publication has occurred.

*http://creativecommons.org/licenses/by-nc/3.0 or http://creativecommons.org/licenses/by/3.0 as applicable*
Assignment of copyright and publication agreement

IOP Publishing Limited ("IOP") agrees, on behalf of the Institute of Physics and Engineering in Medicine ("IPEM") to publish:

Manuscript Title: An augmented parametric response map with consideration of image registration error: towards guidance of locally adaptive radiotherapy (the "Article") written by

Names of all authors: Lausch, Anthony; Chen, Jeff; Ward, Aaron; Gaede, Stewart; Lee, Ting-Yim; Wong, Eugene ("the Named Authors") in the following journal Physics in Medicine and Biology ("the Journal")

Name of Institution(s) (if applicable – see Important Information above):
("the Institution")

IOP Ref: PMB-101004

Part 1 - Subscription Copyright Assignment

Assignment of copyright
1.1 Upon acceptance for publication, in consideration for publication of the Article by IOP, the Named Authors of the Article shall assign to IPEM with full title guarantee the entire copyright in all original material published as part of the Article (which expression includes but is not limited to the text, abstract, tables, figures and graphs, video abstracts and other multimedia content but excludes any other item referred to as supplementary material) throughout the world for the full term of copyright (including any extensions or renewals thereof) for all media and formats, whether known or unknown. For the avoidance of doubt, copyright does not subsist in any fundamental data underlying the Article and nothing in this Agreement is intended to limit access to or use of such data.

1.2 If any of the Named Authors are Government employees who produced the Article in the course of their duties, in consideration for publication such Named Authors and the relevant originating department or agency shall grant IPEM a royalty-free non-exclusive worldwide licence for the full term of copyright (including any extensions or renewals thereof) for all media and formats, whether known or unknown, to do in relation to the Article all acts restricted by copyright worldwide. This shall include, but not be limited to, including the Article (or part thereof) on IOP’s own websites (such as insights and labtalk) and/or on third party websites.

1.3 In consideration for publication, the Named Authors shall grant IPEM and IOP a royalty-free non-exclusive worldwide freely transferrable licence for the full term of copyright (including any extensions or renewals thereof) to do in relation to any supplementary material not deemed to be part of the Article all acts restricted by copyright worldwide. This shall include, but not be limited to, including the supplementary material (or part thereof) on IOP’s own websites (such as insights and labtalk) and/or on third party websites.

1.4 Each of the Named Authors consents to the publication and processing by IPEM and IOP of their email addresses.

Representations and warranties

2.1 The Corresponding Author and/or the Named Authors (as appropriate) represent and warrant that:
2.1.1 the Article is the original work of the Named Authors;

2.1.2 the Article has not been published previously in any form, other than as part of the Named Authors' research theses or dissertations (which fact has been notified to IOP in writing) or as a pre-print, for example on the arXiv.org service;

2.1.3 each of the Named Authors has made a material contribution to the conception and/or writing of the Article, has received the final version of the Article, has agreed to its submission on the terms contained herein and takes responsibility for it and submission has been approved as necessary by the authorities at the establishment where the research was carried out;

2.1.4 the Corresponding Author completes and returns this agreement as authorised agent for and on behalf of all the Named Authors and has the full power to enter into this agreement and to make the grants and assignments it contains;

2.1.5 the Article has not been and shall not be submitted to another publisher prior to withdrawal or rejection by IOP;

2.1.6 the Article does not infringe any third party rights, it contains nothing libellous or unlawful, all factual statements are to the best of the Named Authors' knowledge true or based on valid research conducted according to accepted norms and all required permissions have been obtained in writing; and

2.1.7 the Article expressly acknowledges any third party funding and/or potential conflicts of interest.

2.2 The Named Authors indemnify and will keep indemnified IPEM, IOP and their affiliates, parent companies, subsidiaries, directors, officers, servants, employees, agents and shareholders against all actions, claims, costs, demands, expenses and other liabilities suffered or incurred by IPEM or IOP as a result of and/or arising from IPEM and/or IOP's exploitation (or exploitation by third parties authorised by IPEM or IOP) of the Article where such claim arises out of any breach by any of the Named Authors of the representations and/or warranties in this section 2.

The Named Authors’ rights

3.1 IPEM grants the Named Authors the rights specified in paragraphs 3.2, 3.3 and 3.4. All such rights must be exercised solely for non-commercial purposes. Where possible, any use should display citation information and IPEM's copyright notice, and, for electronic use, best efforts must be made to include a link to the online abstract in the Journal. Exercise of the rights in paragraphs 3.3 and 3.4 must not use the final published IOP format but the Named Authors’ own format (which may include amendments made following peer review but not any editing, typesetting or other changes made by IOP) (the “Accepted Manuscript”) and must be accompanied by the following statement of provenance:

'This is an author-created, un-copyedited version of an article accepted for publication in Physics in Medicine and Biology. Neither IPEM nor IOP Publishing Ltd is responsible for any errors or omissions in this version of the manuscript or any version derived from it. The Version of Record is available online at [insert DOI].

'
3.2 The rights are:

3.2.1 To make copies of the Article (all or part) for teaching purposes;

3.2.2 To include the Article (all or part) in a research thesis or dissertation;

3.2.3 To make oral presentation of the Article (all or part) and to include a summary and/or highlights of it in papers distributed at such presentations or in conference proceedings; and

3.2.4 To include figures and text falling within the quota outlined in the STM Permissions Guidelines (http://www.stm-assoc.org/permissions-guidelines/) at the relevant time in force from the Article in new works created or co-created by any of the Named Authors.

For the avoidance of doubt, the Named Authors retain all proprietary rights in the Article other than copyright.

3.3 Additional rights of the Named Authors are to:

3.3.1 Use the Accepted Manuscript (all or part) without modification in personal compilations of the Named Authors’ own works (provided not created by a third party publisher); and

3.3.2 Include the Accepted Manuscript (all or part) on the Named Authors’ own personal websites.

3.4 In addition to the above, following the expiry of a period of 12 months from the date of publication of the Article by IOP (the "Embargo Period"), the Named Authors may:

3.4.1 Include the Accepted Manuscript (all or part) on websites of the institution (including its repository) where the Named Authors worked when research for the Article was carried out; and

3.4.2 Include the Accepted Manuscript (all or part) on third party websites including e-print servers, but not on other publishers’ websites. Further information is included on the web page http://authors.iop.org/atom/help.nsf/LookupJournalSpecific/WebPermissionsFAQ~**

3.5 Notwithstanding the provisions of paragraph 3.4 above, and until any change to IOP policy, which change will be included within the copyright information on IOP’s website, an exception to the Embargo Period is that the Named Authors may at any time, including prior to publication by IOP of the Article, submit the Accepted Manuscript to the e-print service at arXiv.org subject to a non-exclusive, perpetual licence.

4. To the extent that there are moral rights in the Article, all the Named Authors expressly reserve and assert their moral rights to be identified as the authors of the Article.

5. The Named Authors shall execute such further documents, and take such actions and do such things, as may be requested by IOP at IOP's reasonable expense to give full effect to the terms of this agreement.

6. For the avoidance of doubt, the grants and assignment envisaged herein shall become effective only upon acceptance by IOP of the Article for publication. In the event that the Article is withdrawn prior to acceptance, or is rejected, this agreement shall have no effect and no party shall be bound by it.

7. This agreement shall be governed by English Law and subject to the non-exclusive jurisdiction of the English Courts.

8. By typing the Corresponding Author’s name into the box at Part 3 below and clicking “Submit”, the Named Authors agree to these terms.
Part 2 - Open Access Copyright Assignment

Assignment of copyright

1.1 Upon acceptance for publication, in consideration for publication of the Article by IOP, the Named Authors of the Article shall assign to IPEM with full title guarantee the entire copyright in all original material published as part of the Article (which expression includes but is not limited to the text, abstract, tables, figures and graphs, video abstracts and other multimedia content but excludes any other item referred to as supplementary material) throughout the world for the full term of copyright (including any extensions or renewals thereof) for all media and formats, whether known or unknown. For the avoidance of doubt, copyright does not subsist in any fundamental data underlying the Article and nothing in this agreement is intended to limit access to or use of such data.

1.2 If any of the Named Authors are Government employees who produced the Article in the course of their duties, in consideration for publication such Named Authors and the relevant originating department or agency shall grant IPEM a royalty-free non-exclusive worldwide licence for the full term of copyright (including any extensions or renewals thereof) for all media and formats, whether known or unknown, to do in relation to the Article all acts restricted by copyright worldwide. This shall include, but not be limited to, including the Article (or part thereof) on IOP’s own websites (such as insights and labtalk) and/or on third party websites.

1.3 In consideration for publication, the Named Authors shall grant IPEM and IOP a royalty-free non-exclusive worldwide freely transferrable licence for the full term of copyright (including any extensions or renewals thereof) to do in relation to any supplementary material not deemed to be part of the Article all acts restricted by copyright worldwide. This shall include, but not be limited to, including the supplementary material (or part thereof) on IOP’s own websites (such as insights and labtalk) and/or on third party websites.

1.4 Each of the Named Authors consents to the publication and processing by IPEM and IOP of their email addresses.

Representations and warranties

2.1 The Corresponding Author and/or the Named Authors (as appropriate) represent and warrant that:

2.1.1 the Article is the original work of the Named Authors;

2.1.2 the Article has not been published previously in any form, other than as part of the Named Authors’ research theses or dissertations (which fact has been notified to IOP in writing) or as a pre-print, for example on the arXiv.org service;

2.1.3 each of the Named Authors has made a material contribution to the conception and/or writing of the Article, has received the final version of the Article, has agreed to its submission on the terms contained herein and takes responsibility for it and submission has been approved as necessary by the authorities at the establishment where the research was carried out;

2.1.4 the Corresponding Author completes and returns this agreement as authorised agent for and on behalf of all the Named Authors and has the full power to enter into this agreement and to make the grants and assignments it contains;

2.1.5 the Article has not been and shall not be submitted to another publisher prior to withdrawal or rejection by IOP;

2.1.6 the Article does not infringe any third party rights, it contains nothing libellous or unlawful, all factual statements are to the best of the Named Authors’ knowledge true or based on valid research conducted according to accepted norms and all required permissions have been obtained in writing; and

2.1.7 the Article explicitly acknowledges any third party funding and/or potential conflicts of interest.
2.2 The Named Authors indemnify and will keep indemnified IPEM, IOP and their affiliates, parent companies, subsidiaries, directors, officers, servants, employees, agents and shareholders against all actions, claims, costs, demands, expenses and other liabilities suffered or incurred by IPEM or IOP as a result of and/or arising from IPEM and/or IOP's exploitation (or exploitation by third parties authorised by IPEM or IOP) of the Article where such claim arises out of any breach by any of the Named Authors of the representations and/or warranties in this section 2.

3.1 The Named Authors and all third parties will have the rights to use the Article as described in the Creative Commons Attribution 3.0 Unported licence (http://creativecommons.org/licenses/by/3.0/). These rights allow users to copy, distribute and display the published version of the Article and create derivative works, subject to appropriate attribution.

3.2 Where the Article is used in accordance with 3.1 above, the following attribution shall be included:

Article title
Named Authors
DOI
“This article was first published in Physics in Medicine and Biology, which maintains the Version of Record”.

Miscellaneous

4. To the extent that there are moral rights in the Article, all the Named Authors expressly reserve and assert their moral rights to be identified as the authors of the Article.

5. The Named Authors shall execute such further documents, and take such actions and do such things, as may be requested by IOP at IOP's reasonable expense to give full effect to the terms of this agreement.

6. For the avoidance of doubt, the grants and assignment envisaged herein shall become effective only upon acceptance by IOP of the Article for publication. In the event that the Article is withdrawn prior to acceptance, or is rejected, this agreement shall have no effect and no party shall be bound by it.

7. This agreement shall be governed by English Law and subject to the non-exclusive jurisdiction of the English Courts.

Confirmation

8. By typing the Corresponding Author’s name into the box at Part 3 below and clicking “Submit”, the Named Authors agree to all these terms.

Part 3 – Confirmation and Execution

Please tick the box below to confirm the basis upon which you, as the Corresponding Author, on behalf of all the Named Authors, are submitting the Article to IOP.

Subscription Copyright Assignment – Part 1 Applies

Type your name here:
Anthony Lausch (the “Corresponding Author”)
By clicking “Submit“ and typing your name above, you shall be assumed to have read and understood all of the terms and conditions of the relevant part of this agreement and you will be agreeing to all of the terms and conditions and assignment (as the case may be) detailed above.
Appendix C

Curriculum Vitae
Anthony Lausch, BSc, MSc

a. EDUCATIONAL BACKGROUND

<table>
<thead>
<tr>
<th>Degree</th>
<th>Specialty / University</th>
<th>Research</th>
<th>Years Active</th>
</tr>
</thead>
<tbody>
<tr>
<td>PhD, MSc-CL CAMPEP</td>
<td>Medical Biophysics, University of Western Ontario (UWO)</td>
<td>Image-based treatment response prediction; towards guidance of locally adaptive radiotherapy. Supervisor: Dr. Eugene Wong</td>
<td>2011-2015</td>
</tr>
<tr>
<td>MSc</td>
<td>Medical Biophysics, University of Toronto</td>
<td>Deformable image registration for dynamic contrast-enhanced MRI. Supervisor: Dr. Anne Martel</td>
<td>2009 - 2011</td>
</tr>
<tr>
<td>BSc</td>
<td>Mathematical Physics, Co-operative Program, University of Waterloo</td>
<td>Projects summarized in section f. Research assistantships</td>
<td>2004 - 2009</td>
</tr>
</tbody>
</table>

b. PUBLICATIONS

i. Refereed journal articles


S. Wesolkowski, **Lausch A** (2010) Validation of the stochastic fleet estimation model. DRDC-CORA TM 2010-042
ii. Refereed conference proceedings


iii. Conference abstracts


iv. Government Reports


c. SCHOLARSHIPS, AWARDS & HONOURS

Anthony Lausch

Nominee, Macklin Fellowship for Teaching and Research Award 2014
Selected by graduate chair to be Dept. Medical Biophysics nominee for faculty-wide award in recognition of excellence in teaching and research

“Seeing teaching more clearly”, Rapport magazine, p. 11-12 2014
Schulich School of Medicine and Dentistry Alumni Magazine feature on BIOPHYS 9520B curriculum development initiatives

“Top Canadian Cancer Society funded research stories of 2013” 2013
Online CCS news article, Lausch et al. Br J Radiol 2013;86:20130147 recognized as one of the top CCS funded Research stories of 2013.

Western Graduate Thesis Research Award 2013
$1 500 (internal university grant competition)

Ontario Graduate Scholarship in Science and Technology 2009
$10 000

d. COMMITTEES AND OUTREACH

Graduate Student Representative, Departmental Meetings 2014-2015
Dept. of Medical Biophysics, UWO

Graduate Student Representative 2014-2015
CAMPEP PhD MSc-CL Program Steering Committee
Dept. of Medical Biophysics, UWO

Member, Graduate Student Recruitment Committee 2014-2015
Dept. of Medical Biophysics, UWO

Member, Communications Committee 2013
Dept. of Oncology, UWO

e. TEACHING EXPERIENCE

i. Curriculum Development


Summary: Member of eight student team who developed graduate department’s new medical imaging lab course based on desktop CT, US, and MRI systems. On a weekly basis, team also researched and discussed effective pedagogy under the direction of a faculty member.
Summary: Developed and presented initiatives for improving the graduate seminar program and creation of a seminar course. Organized, chaired, and contributed to subsequent development meetings. Created new seminar evaluation forms. Course initiated in Sept 2013.

ii. Graduate Teaching Assistantships

f. RESEARCH ASSISTANTSHIPS

Physics Research Assistant, Physics Department
London Regional Cancer Program, London, ON
Supervisor: Dr. Eugene Wong
Project/Task Summary: Continued development/testing of a computer assisted target definition algorithm for prostate bed radiotherapy.

Defence Research Assistant, Operational Research
Department of National Defence, Ottawa, ON
Supervisor: Dr. Slawomir Wesolkoski
Project/Task Summary: Developed/tested optimization algorithms for assignment of air assets to military transport missions.

Angiography Research Assistant, Imaging Research
Sunnybrook Health Sciences Centre, Toronto, ON
Supervisor: Dr. Normand Robert
Project/Task Summary: Simulated presence of a pre-patient x-ray attenuator on fluoroscopic images to determine feasibility of using attenuator to reduce exposure during fluoroscopic procedures. Verified estimated reduction by performing ion chamber measurements in a fluoroscopy suite during exposure of an anthropomorphic phantom with attenuator in place.

Physics Research Assistant, Physics Department
London Regional Cancer Program, London, ON
Supervisor: Dr. Eugene Wong
Project/Task Summary: Investigated effects of input parameter variation on toxicity prediction models for prostate radiotherapy. Initiated development of a computer assisted target definition algorithm for prostate bed radiotherapy.

9. OTHER EXPERIENCE AND TRAINING

Canadian Good Clinical Practice Online Training Course, 2015
CITI Training Program, Lawson Health Research Unit/University of Miami
Performed basic quality assurance checks and output measurements for Varian Clinac iX and TrueBeam linear accelerators (Sun Nuclear MapCheck, water tank, water-equivalent plastic phantom)

PINNACLE³, Philips Treatment Planning System
Experience with 3D-CRT and IMRT treatment planning (prostate)

MATLAB/3D Slicer
Extensive experience with modeling, optimization, image registration and processing