Optimization of Decision Making in Personalized Radiation Therapy using Deformable Image Registration

Jason Vickress
The University of Western Ontario

Supervisor
Yartsev, Slav
The University of Western Ontario

Joint Supervisor
Barnett, Rob
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

© Jason Vickress 2018

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

Recommended Citation
https://ir.lib.uwo.ca/etd/5542

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, wlswadmin@uwo.ca.
Abstract

Cancer has become one of the dominant diseases worldwide, especially in western countries, and radiation therapy is one of the primary treatment options for 50% of all patients diagnosed. Radiation therapy involves the radiation delivery and planning based on radiobiological models derived primarily from clinical trials. Since 2015 improvements in information technologies and data storage allowed new models to be created using the large volumes of treatment data already available and correlate the actually delivered treatment with outcomes. The goals of this thesis are to 1) construct models of patient outcomes after receiving radiation therapy using available treatment and patient parameters and 2) provide a method to determine real accumulated radiation dose including the impact of registration uncertainties.

In Chapter 2, a model was developed predicting overall survival for patients with hepatocellular carcinoma or liver metastasis receiving radiation therapy. These models show which patients benefit from curative radiation therapy based on liver function, and the survival benefit of increased radiation dose on survival.

In Chapter 3, a method was developed to routinely evaluate deformable image registration (DIR) with computer-generated landmark pairs using the scale-invariant feature transform. The method presented in this chapter created landmark sets for comparing lung 4DCT images and provided the same evaluation of DIR as manual landmark sets.

In Chapter 4, an investigation was performed on the impact of DIR error on dose accumulation using landmarked 4DCT images as the ground truth. The study demonstrated the relationship between dose gradient, DIR error and dose accumulation error, and presented a method to determine error bars on the dose accumulation process.

In Chapter 5, a method was presented to determine quantitatively when to update a treatment plan during the course of a multi-fraction radiation treatment of head and neck cancer. This method investigated the ability to use only the planned dose with deformable image registration to predict dose changes caused by anatomical deformations.
This thesis presents the fundamental elements of a decision support system including patient pre-treatment parameters and the actual delivered dose using DIR while considering registration uncertainties.

Keywords

Radiation therapy, image-guided radiation therapy, adaptive radiation therapy, decision support, deformable image registration, registration error
Co-authorship statement

This dissertation includes material published in different journals. The permission to the content from each of these articles is located in Appendix A: Permission to reuse published material.

Chapter 2, “A multivariable model for the prediction of survival for patients with hepatic carcinoma or liver metastasis receiving radiotherapy” was published in Future Oncology, 2017; 13 (1), 19-30 by Jason Vickress, Michael Lock, Stewart Gaede, Aaron Leung, Jeff Cao, Rob Barnett, and Slav Yartsev. I was involved in the data acquisition, statistical analysis, programming, and writing of the manuscript. Aaron Leung performed chart review and obtained the needed clinical parameters. All authors reviewed and edited the final manuscript.

Chapter 3, “Automatic landmark generation for deformable image registration evaluation for 4D CT images of lung” was published in the journal Physics in Medicine and Biology, 2016; 61(20), 7236 by Jason Vickress, Jerry Battista, Rob Barnett, John Morgan, and Slav Yartsev. I was responsible for all data acquisition of CT studies, image registration, analysis, programming of landmarking algorithm, and writing of the manuscript. John Morgan was involved with manual landmarking of the 4DCT images. All other authors were involved in the design of the study, and editing of the final manuscript.

Chapter 4, “Representing the Dosimetric Impact of Deformable Image Registration Errors” published in the journal Physics in Medicine and Biology, 2017, 62(17) N391 by Jason Vickress, Jerry Battista, Rob Barnett, and Slav Yartsev. I performed all of the data acquisition of landmarked CT studies, programming, analysis and writing of the manuscript. All authors were involved in the design the study and editing of the final manuscript.

Chapter 5, submitted to the Journal of Applied Clinical Medical Physics by Jason Vickress, Jerry Battista, Rob Barnett, and Slav Yartsev. I was involved in all data analysis, programming, and writing of the final manuscript. All authors were involved in the design of the study and editing of the final manuscript.
Acknowledgments

The work contained in this Ph.D. thesis was not a solo venture but required the mentoring, help and support by many people, including students, supervisors, and family. I would like to express my gratitude to the many people who have helped me over the last 4 years, making this work possible.

I would first like to thank my two supervisors Drs Slav Yartsev and Rob Barnett because without them none of this would have been possible, as an undergraduate co-op student they encouraged me and gave me the opportunity to pursue my Ph.D. Slav, I would like to thank you for all the necessary “whip cracking” over the years keeping me on track always working towards the next publication. I can’t thank you enough for all the reviewing and editing of my manuscripts, applications, posters, and presentations. My writing has improved immensely under your guidance. Rob, thanks for all guidance and inspiration on my projects with TOXIC DB, and image registration, always reminding the clinical importance of everything we were doing. I would also like to thank you for all the 5 pm meetings in your office, debating ideas for the next step in the project.

Next, I would like to thank my honorary supervisor Dr. Jerry Battista for taking part in most of my monthly project meetings, providing the sober second thought and an endless source of analogies. In the radiation biology course, you shared the joys of teaching and all the responsibilities that go along with it.

Throughout my Ph.D. process, I have worked with many excellent high school and undergraduate students who provided invaluable assistance. From the PEL program, I would like to thank Hugh McHenry, Faran Khalid, and Ken Cha, who helped with contouring, data analysis, and performing deformable image registration hundreds if not thousands of times. From the University of Waterloo, I had the pleasure to work with John Morgan who helped me deal with Pinnacle and MIM.

As a graduate student, I have been surrounded by only the best lab mates and fellow students providing wisdom and guidance throughout the graduate school process. Matthew Mouawad, thanks for all the help with MATLAB, Slicer, and the occasional
euchre game. Thanks to Dr. Anthony Lausch for meeting with me early on in my Ph.D. track and helping me to understand deformable image registration and the importance of registration uncertainty. I would like to thank Dr. Kurtis Decker for providing support and guidance and all around great desk neighbour. Lastly, I couldn’t have finished (or started) my Ph.D. without Mike Macfarlane, always there to bounce ideas off, keeping up to date on Medical Physics news, and always being there for a Friday shawarma.

I couldn’t have survived my graduate studies without all the technical help and support from the staff at the LRCP providing me everything I need to succeed over the last 5 years. Thanks to Barb Barons for always looking out for me and helping with all administration, scheduling meetings and finances. Since my research was heavily based on computing, and medical imaging I relied on the IT support from Darrel Basdeo and Jeff Kempe helping me build databases, and providing top of the line computer systems. I would also like to acknowledge Dr. Rashid Dar, for always encouraging collaboration with oncology, and providing the physicians perspective on my work.

During my time at the London Regional Cancer Program (LRCP), I have had the opportunity to work with many of the great medical physicists for my research and to learn more about their clinical responsibilities. Some of these physicists include Drs Scott Karnas, Hatim Fakir, Stewart Gaede, and Kevin Jordan. I would like individually thank Dr. Bryan Schaly for all his help in providing patient images for our adaptive radiotherapy study and teaching me how clinical adaptive radiotherapy operates at the LRCP.

From the beginning of my Ph.D. studies, great advisors have helped me to produce the thesis presented here. My advisory committee included Drs Aaron Ward, Jeff Chen, and Michael Lock. Aaron Ward, thanks for being our registration and image processing expert. Michael Lock, thanks for always having an upbeat attitude reminding me the clinical importance of all the work I was doing, even if I didn’t believe it at the time. Lastly, I would like to thank Jeff Chen for never being afraid to ask the tough questions, relating to the real world importance and practical application of my projects.
Aside from all the help, I had in London at Western University and the LRCP, I also had massive levels of support from my friends and family. After only the first year of my Ph.D. studies, my father had brain surgery leaving him hospitalized for more than a year requiring me to commute from London to hospitals and rehab centres in Toronto every weekend. There is no way I could have completed my Ph.D. thesis if it wasn’t for all my friends and family member helping to take care of my father and help me make the commute from London. I would like to thank my uncles Ian, Wayne, and Neil for always picking me up at the train station, or driving me to Toronto at any time. I would also thank my Aunt Valery and my grandparents for helping to take care of my Mom as she visited my father in the hospital every day. Also, I would like to thank my sister for always being there to take care of my mother and father during the week, so I could focus on my studies in London.

The last people I would like to thank are my mother and father always encouraging me to try hardest at school, and being proud of my every achievement. To my father, thank you for sharing your love of learning and knowledge; I wouldn’t have gotten this far without you. The last person I would like to thank is my mother, you are my rock, without you, I could not have survived both my father’s hospitalization and recovery while completing a Ph.D. Without knowing you had everything under control at home I would not have had a chance. Thanks, Mom, there is no way I could have finished in one piece without you.
# Table of contents

Abstract ........................................................................................................................................ i

Co-authorship statement .............................................................................................................. iii

Acknowledgments ........................................................................................................................ iv

Table of contents ......................................................................................................................... vii

List of figures ............................................................................................................................... xiii

List of tables ................................................................................................................................ xvii

List of appendices ....................................................................................................................... xviii

List of abbreviations .................................................................................................................. xix

Chapter 1 ..................................................................................................................................... 1

1 Introduction ............................................................................................................................... 1

1.1 Overview ............................................................................................................................. 1

1.2 Cancer .................................................................................................................................. 2

1.3 Radiation therapy ................................................................................................................. 3

1.4 Patient outcomes .................................................................................................................. 6

1.5 Image guided radiation therapy ........................................................................................... 7

1.6 Adaptive radiation therapy .................................................................................................. 8

1.7 Deformable image registration ............................................................................................. 10

1.7.1 Background ....................................................................................................................... 10

1.7.2 DIR fundamentals ........................................................................................................... 12

1.7.3 Transformation ............................................................................................................... 12

1.7.4 Image similarity .............................................................................................................. 14

1.7.5 Regularization ................................................................................................................ 15

1.7.6 Different algorithms ....................................................................................................... 16

1.8 Registration error ................................................................................................................. 17
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.9</td>
<td>Decision making in radiation therapy</td>
<td>19</td>
</tr>
<tr>
<td>1.10</td>
<td>Research goals and objectives</td>
<td>22</td>
</tr>
<tr>
<td>1.11</td>
<td>Thesis outline</td>
<td>22</td>
</tr>
<tr>
<td>1.11.1</td>
<td>Modelling of overall survival for hepatic carcinoma or liver metastasis receiving radiotherapy (Chapter 2)</td>
<td>22</td>
</tr>
<tr>
<td>1.11.2</td>
<td>Automatic landmark generation for evaluation of deformable image registration accuracy (Chapter 3)</td>
<td>23</td>
</tr>
<tr>
<td>1.11.3</td>
<td>Dosimetric impact of deformable image registration error in radiation therapy (Chapter 4)</td>
<td>23</td>
</tr>
<tr>
<td>1.11.4</td>
<td>Daily assessment of dose change using deformable image registration without dose-recalculation (Chapter 5)</td>
<td>23</td>
</tr>
<tr>
<td>1.12</td>
<td>References</td>
<td>24</td>
</tr>
</tbody>
</table>

Chapter 2 .............................................................................................................. 35

2 A multivariable model for the prediction of survival for patients with hepatic carcinoma or liver metastasis receiving radiotherapy ........................................... 35

2.1 Introduction ................................................................................................. 35

2.2 Methods ......................................................................................................... 37

2.2.1 Patient population .................................................................................. 37

2.2.2 Radiation treatment ............................................................................... 39

2.2.3 Statistical analysis ................................................................................ 39

2.3 Results .......................................................................................................... 40

2.3.1 Patient demographics .......................................................................... 40

2.3.2 Effect of liver function ..................................................................... 41

2.3.3 Primary HCC vs liver metastasis ...................................................... 42

2.4 Discussion .................................................................................................... 46

2.4.1 Patient demographics ...................................................................... 46
2.4.2 Liver function .......................................................... 47
2.4.3 Nomogram .............................................................. 47
2.4.4 Potential limitations .................................................. 48

2.5 Conclusion .................................................................. 49

2.6 References .................................................................. 49

Chapter 3 ....................................................................... 53

3 Automatic landmark generation for deformable image registration evaluation for 4D CT images of lung ............................................................................ 53

3.1 Introduction .................................................................. 53

3.2 Methods ..................................................................... 55

3.2.1 Manual .................................................................. 55

3.2.2 SIFT ..................................................................... 55

3.2.3 SIFT-manual hybrid .................................................. 56

3.2.4 Landmark displacement evaluation ............................. 57

3.2.5 Deformable image registration evaluation .................... 58

3.3 Results ..................................................................... 58

3.3.1 Mean displacement ................................................... 59

3.3.2 Proximal deformation ................................................ 60

3.3.3 Deformable image registration evaluation .................... 61

3.4 Discussion .................................................................. 61

3.5 Conclusion .................................................................. 64

3.6 References .................................................................. 64

Chapter 4 ....................................................................... 66

4 Representing the dosimetric impact of deformable image registration errors .... 66

4.1 Introduction .................................................................. 66
List of figures

Figure 1-1: CT image study of head and neck region a) axial b) sagittal c) coronal. ........4

Figure 1-2: Image of overlapping intensity modulated radiation therapy (IMRT) fields generated by specific multi-leaf collimator (MLC) configurations. Colour map represents the dose distribution. Image from http://www.ofunachuohp.net/rt/technique.html ..........5

Figure 1-3: a) checkerboard overlap of MRI and CT of female Pelvis [44] B) CT image and CT image with PET colour map showing metabolic activity and highlighting lesion [50]. Permission for both figures is presented in Appendix A.1 & A.2 ................. 11

Figure 1-4: Example of a deformation vector field, shown in the interface from the Varian Velocity software suite. .............................................................................................................................................. 12

Figure 1-5: a) Example of a 1D spline interpolation b) example of demons DVF vector for a small object shrinking from the bottom. Produced in MATLAB. ......................... 13

Figure 1-6: Example of the effect of registration error. Colour map is the radiation dose. Point A) miss-registration but still on the boundary, B) miss-registration still within boundary and C) miss-registration with same dose level. .................................................. 18

Figure 1-7: Schematic for outcome-driven decision making in personalized radiotherapy. .............................................................................................................................................. 21

Figure 2-1: Kaplan--Meir survival curves for patients with HCC and secondary liver metastasis. ........................................................................................................................................... 41

Figure 2-2: Nomogram for the probability of six months and one-year survival for patients with (a) HCC and (b) liver metastasis. ................................................................. 44

Figure 2-3: Kaplan--Meir survival curves for HCC (a) and Mets (b) showing the lower (solid) and higher (dashed) risk groups obtained by the multivariate models. .............. 45

Figure 2-4: Calibration curves for predicting 6-month survival for HCC (a) and Mets (b). The goodness of fit was calculated using the HL test yielding a p-value of 0.1693 and 0.0412 for HCC and Mets respectively. .............................................................................................. 45
Figure 2-5: Calibration curves for predicting 1-year survival for HCC (a) and Mets (b). The goodness of fit was calculated using the HL test yielding a p-value of 0.211 and 0.077 for HCC and Mets respectively.

Figure 3-1: Example of an excluded landmark pair (crosses) produced from SIFT. The pair was excluded because the landmarks were located in different slices of the CT study.

Figure 3-2: Proximal landmark selection with landmarks in the 0% phase (circles) and 50% phase (squares) from the test (white) and gold standard (black) landmark sets including; landmark displacement D, DIR predicted landmark displacement DDIR, distance between 0% positions d and DIR error ε for a test landmark T, and two GS landmarks 1 and 2.

Figure 3-3: Number of landmark pairs per litre located using SIFT and SIFT-M landmark creation methods for 10 4DCT studies. The SIFT-M method excludes some of the SIFT points using visual inspection.

Figure 3-4: Mean lung deformations for 20 lungs: a) Observers 1 and 2 vs GS and b) SIFT and SIFT-M methods vs GS. The black line shows ideal agreement.

Figure 3-5: Landmark displacement for proximal pairs for a) SIFT vs GS and b) SIFT-M vs GS for all ten studies.

Figure 4-1: Flowchart outlining the calculation of the range of dose uncertainty for a single landmark starting in the end-inspiration image.

Figure 4-2: RDU process for a single landmark in end-inspiration phase. The DIR predicted position of the landmark is located in end-expiration image and dose is sampled in a sphere around it. RDU is the maximum and minimum doses within the sampling sphere a) end inspiration b) end respiration c) dose voxel histogram for points within the sampling sphere.

Figure 4-3: The range of dose uncertainty (RDU) for 31 landmarks with the highest doses from a representative study calculated using four different measures of DIR error a) 2mm constant b) 5 mm constant c) DDM d) Actual error. Actual dose (red X) and RDU is shown as a black error bar. Landmarks are ordered in increasing predicted dose difference.
from the actual dose. Dose values are displayed relative to the DIR deformed dose value for each landmark. ........................................................................................................................................ 73

Figure 4-4: Magnitude of dose uncertainty (MDU) averaged across all 3000 landmarks within 10 4DCT image studies, using the actual DIR error, 2 and 5 mm constant error, distance discordance metric (DDM), inverse consistency error (ICE), transitivity error (TE), and the mean actual DIR error per image. Error bars represent the standard deviation across 3000 landmarks........................................................................................................................................ 74

Figure 4-5: Inclusion rate (IR) averaged across all 10 4DCT image studies, using the actual DIR error, 2 and 5 mm constant error, distance discordance metric (DDM), inverse consistency error (ICE), transitivity error (TE), and the mean actual DIR error per image. Error bars represent standard deviation across 10 studies.................................................................................................................. 75

Figure 4-6: Distributions, of the (a) planned, (b) deformed dose, (c) lower bound, (d) upper bound of the range of dose uncertainty calculated using DDM. Contour labels: (light blue) PTV and (dark blue) is gross target volume. Dose prescription was 60 Gy to 95% of the PTV ........................................................................................................................................ 76

Figure 4-7: Cumulative DVH curves for the planned and deformed dose with the upper and lower bounds illustrating the range of dose uncertainty(RDU) calculated using the DDM. Dose prescription was 60 Gy to 95% of the PTV ........................................................................................................................................ 76

Figure 5-1: Illustration of the CBCT\textsubscript{P} method for the evaluation of the need for plan adaptation using the DIR of planning CT to daily CBCT study and the planned dose distribution. DIR – deformable image registration, RR – rigid registration. ..................... 87

Figure 5-2: Schema describing the daily dose estimation using DIR from the planning CT to either the daily CBCT or re CT study (ReCT). Two different dose distributions computed on the PCT or ReCT are transferred to the moving image using a 6 degree of freedom rigid registration. The gold standard method is highlighted in yellow using the ReCT and recomputed dose. Day X is when ReCT was ordered due to observed significant anatomical changes........................................................................................................................................ 88

Figure 5-3: Predicted mean dose using a) CBCT\textsubscript{P} and b) CBCT\textsubscript{R} methods compared to ReCT\textsubscript{R} (gold standard) for 15 parotid glands. Clinical threshold of 26 Gy is shown by solid lines. ReCT\textsubscript{R} is the DIR to ReCT using the recalculated dose. CBCT\textsubscript{P} is the DIR to
daily CBCT using the planned dose. CBCT\(_R\) is the DIR to daily CBCT using the recalculated dose. .......................................................... \(91\)

Figure 5-4: The difference between predicted D95 and the prescribed dose for the PTV for using a) CBCT\(_P\) and b) CBCT\(_R\) methods compared to ReCT\(_R\) (gold standard). Values are presented as the difference from the prescribed dose. ReCT\(_R\) is the DIR to ReCT using the recalculated dose. CBCT\(_P\) is the DIR to daily CBCT using the planned dose. CBCT\(_R\) is the DIR to daily CBCT using the recalculated dose. .......................................................... \(92\)

Figure B-1: Diagram describing the use of a clinical nomogram, using clinical variables to determine the probability of one year and six month survival. ........................................ \(121\)
List of tables

Table 2-1: Patient demographics.................................................................38
Table 2-2: Univariate Cox regression analysis of two CP score groups for OS. ..........42
Table 2-3: Univariate and multivariable Cox regression results for both HCC and metastasis. .........................................................................................................................42
Table 2-4: Model results including β values, units and 6 months / 1-year baseline cumulative hazard ratio.................................................................................................................................43
Table 3-1: ICC values and corresponding 95% confidence intervals for DIR error and landmark displacement for both mean and proximal methods. ........................................61
Table 5-1: Relative voxel-wise dose difference from gold standard (ReCT_R) (RD_j) for ipsilateral and contralateral parotids and spinal cord, averaged over 18 patients. Standard deviation is displayed in brackets.................................................................90
List of appendices

Appendix A: Permission to reproduce published material ........................................... 106

A.1 Permission to reproduce content for Figure 1-3a ........................................... 106
A.2 Permission to reproduce content for Figure 1-3b (same license agreement as A.1) ................................................................. 112
A.3 Permission to reproduce the content of chapter 2 from Future Oncology .......... 114
A.4 Permission to reproduce the content of chapter 3 and 4 from Physics in Medicine and Biology ................................................................. 119

Appendix B: Description of clinical parameters and tools from chapter 2 ............... 120

B.1 Important clinical measures for liver cancer patients .................................... 120
B.2 Interpreting Cox proportional hazard model results ..................................... 120
B.3 Applying a nomogram ..................................................................................... 121

Appendix C: Description of the Scale Invariant Feature Transform (SIFT) ............ 122

C.1 Scale invariance .............................................................................................. 122
C.2 Finding unique points .................................................................................... 122
C.3 Removal of edge points ................................................................................ 123
C.4 Point localization ........................................................................................... 123
C.5 Point descriptors and matching ..................................................................... 123
C.6 Optimization .................................................................................................. 124
C.7 References ..................................................................................................... 124
List of abbreviations

3DCRT – Three-dimensional conformal radiotherapy

4DCT – Four-dimensional computed tomography

AUC – Area under the curve

CBCT – Cone-beam CT

C-index – Concordance index

CP – Child Pugh

CT – Computed tomography

CTV – Clinical target volume

DDM – Distance discordance metric (mm)

DFS – Disease free survival

DIR – Deformable image registration

DVF – Deformation vector field

DVH – Dose-volume histogram

EQD₂ – Biological effective dose in 2 Gy fractions

GS – Gold standard

GTV – Gross target volume

Gy – Units of radiation dose (J/kg)

HCC – Hepatocellular carcinoma

HL – Hosmer-Lemeshow

HR – Hazard ratio
ICC – Intraclass correlation
ICE – Inverse consistency error (mm)
ICRU – International commission on radiation units and measurements
IGRT – Image-guided radiation therapy
IMRT – Intensity modulated radiation therapy
IR – Inclusion rate (%)
ITV – Internal Target Volume
kV – Kilovoltage
MDU – Magnitude of dose uncertainty (Gy)
MI – Mutual information
MLC – Multi-leaf collimator
MRI – Magnetic resonance imaging
MV – Megavoltage
NCC – Normalized cross-correlation
NTCP – Normal tissue complication probability
OAR – Organ at risk
OS – Overall Survival
PCT – Planning computed tomography
PET – Positron emission tomography
PTV – Planning target volume
RDU – Range of dose uncertainty (Gy)
ROI – Region of interest

SBRT – Stereotactic radiation therapy

SIFT – Scale invariant feature transform

SIFT-M – Scale invariant feature transform with manual landmark exclusion

SSD – Sum of squared difference

TCP – Tumour control probability

TE – Transitivity error (mm)

Tomo – Helical tomotherapy

V24 – Volume of normal liver receiving more than 24 Gy

VMAT – Volume modulated arc therapy
Chapter 1

1 Introduction

1.1 Overview

Compared to most other diseases cancer presents a unique problem of unregulated cell growth, resulting from a wide range of cell mutations. These mutations create many different cancer types and pathologies each responding differently to treatment. Radiation therapy has been used since the early 1940s to provide effective treatment for many cancers and is currently one of the primary treatment options for cancer. Through different delivery methods and geometric targeting high levels of radiation can be delivered to cancer cells, while trying to spare surrounding healthy tissue. The challenge is to ensure that the necessary amount of radiation is delivered to a cancerous tumour to treat a patient’s disease.

Unlike other treatment modalities, the intended amount of radiation to specific organs or regions can be visualized from a 3D computerized calculation (treatment plan) and be used to optimize disease control and limit toxicity. Radiation therapy has progressed over the last two decades from basic target coverage to “precision” coverage for the benefit of improved tumour control and reduced normal tissue damage. Unfortunately with precision coverage the static treatment plan (created on the patient’s pre-treatment anatomy) becomes less accurate as treatment progresses and the patient’s anatomy changes requiring new tools to characterize changes in accumulated dose during fractionated treatment. Image guidance was introduced to limit positioning errors and monitor for anatomy changes but it could not effectively compensate for anatomical deformations because reducing position errors is insufficient for optimized dose distributions. Even if the plan can be adapted when significant anatomical changes are noticed, it is currently difficult to determine the accumulated dose on a per voxel basis and the correlation with outcomes.
Deformable image registration (DIR) is a graphics tool that maps every point in a reference volume to the corresponding point in another image volume. With DIR and image guidance the dose received in every fraction can be mapped back to the original planning volume determining the cumulative delivered dose. However, DIR currently contains inherent uncertainty introducing quandary to the record of delivered dose.

Currently, in radiation therapy, there is lack of information on the actual delivered dose due to limitations in image registration accuracy preventing effective correlation with patient outcomes. Historically, treatment outcomes were correlated only with patient’s planned dose distribution calculated prior to the start of treatment. The inclusion of detailed information about planned and actually delivered dose would improve the predictive power of treatment outcomes and optimize/personalize the treatment process. Ideally with an accurate record of the distribution of the received dose paired with outcomes and pre-treatment parameters improved personalized medicine can be realized. The following sections will introduce cancer, adaptive radiation therapy and deformable image registration to provide context for later chapters.

1.2 Cancer

Since the beginning of human history, our mortality has been tested from predators, environmental disasters, war and disease. Through cooperation and innovation, we have thrived in our environment and extended the average human lifespan from 30 years in the 1800s to 67 years in the year 2000 [1]. This modern longevity can be attributed to better nutrition and modern medicine. Traditionally medicine sought to cure diseases from foreign sources including bacterial infections and viruses or to repair physical damage or birth defects. As many life-threatening diseases and conditions were eradicated and with increased life expectancy, cancer became much more prevalent especially in developed nations. As a fundamentally different type of disease, cancer required an entirely different approach.

Cancer is not directly caused by any foreign organism or event but occurs when there is a significant mutation in a cell’s DNA, making cancer cells grow uncontrollably without normal growth regulation. The primary symptoms of a cancer cell are rapid and
uncontrolled growth and loss of the cell’s original function. The lethality of cancer is caused by the tumour’s necessity for oxygen, nutrients and space, ultimately starving surrounding healthy tissue as it spreads and grows. Treatment of cancer is difficult since it has few differences from healthy cells, unlike bacteria, and the immune system can no longer effectively combat it. Unlike other diseases the mutations at the root of cancer are a fundamental part of cellular evolution, making cancer inescapable.

Throughout the developed world cancer is the leading cause of death, making its treatment and management paramount. Studies have shown specifically in Canada 1 in 2 people will get cancer in their lifetime [2]. With such a large burden, the management of cancer has become a major component of healthcare costing 4.4 billion dollars in 2008 in both direct health care costs and lost productivity [2]. There are currently three main treatment modalities for cancer management; i) surgery, ii) systemic therapy and iii) radiation therapy. Surgery is the desired option, physically removing the malignancy but only applies to removable tumours and patients healthy enough to endure the procedure and recovery. Systemic therapies use drugs to treat cancer systemically but can create significant toxicities since the effects are typically poorly targeted. Radiation therapy is a non-surgically invasive targeted therapy, capable of treating a majority of patients, producing limited toxicity and easily combined with other therapies. The complexity of cancer diagnosis and treatment (with multiple modalities) requires accurate modelling to determine the optimal treatment technique. Currently, making a decision for treatment is challenging, because outcome prediction is made based on averaged data, while patient’s anatomy or disease may vary significantly. For modelling to be effective, a complete account of all treatment parameters is required.

1.3 Radiation therapy

Radiation therapy employs ionizing radiation to treat cancer, with the goal to deliver sufficient levels of radiation dose to a tumour while sparing surrounding healthy tissue. Quantitatively, the required dose to kill cancerous cells is given by tumour control probability (TCP) models [3-5]. Tolerance doses for organs at risk (OAR) are obtained from normal tissue complication probability (NTCP) models [6-7]. Recently the
QUANTEC publications [8] provided updated recommendations for limiting dose/volume parameters for OARs.

Radiation can be delivered externally (external beam radiation therapy) or from implanted radiation sources (brachytherapy). The primary mechanism of radiation therapy is to fatally damage cancer cells by breaking their DNA leading to cell death. Every radiation therapy treatment plan begins with a CT simulation representing the patient’s anatomy in three dimensions and depicting (and contrasting) the electron density of different tissues required for dose calculation. An example of a CT simulation is presented in Figure 1-1 with the axial, sagittal and coronal cross-sections of a CT volume with the gray scaling illustrating the different electron densities. Contours are produced to identify individual tissue types, OAR and target the tumour.

![CT image study of head and neck region](image.png)

**Figure 1-1:** CT image study of head and neck region a) axial b) sagittal c) coronal.

The radiation treatment plan is optimized for the patient’s contoured CT simulation to select the best combination of radiation beams produced at different angles overlapping on the target region demonstrated in Figure 1-2.
Figure 1-2: Image of overlapping intensity modulated radiation therapy (IMRT) fields generated by specific multi-leaf collimator (MLC) configurations. Colour map represents the dose distribution. Image from http://www.ofunachuohp.net/rt/technique.html.

A radiation treatment plan is generated as the result of applying targeted beams to a tumour while limiting the dose to organs at risk. Historically treatments were delivered using large rectangular fields with wedged beams or lead blocks which could cause significant toxicities from less conformal dose distributions. With higher precision delivery tools developed in the mid-nineties such as multi-leaf collimators (MLC) and Intensity-modulated radiation therapy (IMRT) toxicity was reduced compared to earlier treatment techniques [9]. In modern radiotherapy practice, there are many degrees of freedom, including beam angles, energy and modulation utilizing an MLC that defines the treatment plan. An MLC is a series of tungsten segments that shape the radiation field during treatment shown for each beam in Figure 1-2. Optimization of these parameters to satisfy the prescribed dose constraints creates highly conformal dose distributions, displayed in the centre of Figure 1-2. With further improvements and the combination of arcing, varied dose rate and beam modulation, volumetric arc therapy (VMAT) was developed in 2010 improving plan quality further [10-13] with reduced treatment times. As technology developed, the ability to deliver radiation with pinpoint accuracy has
become possible, but the difficulty lies with adapting to anatomical changes during treatment while maintaining the same precision and accuracy.

1.4 Patient outcomes

Radiation delivery has improved significantly with the introduction of new techniques including IMRT, VMAT, and proton therapy, but though technically superior they have shown limited improvements in patient outcomes especially when related to their cost. Health care costs are increasing unsustainably in the developed world with the USA spending 17.2% of their GDP on healthcare in 2017 as shown from the OECA health statistics [14]. Much of the increase in spending will come from private insurers increasing premiums, funding cuts to other programs and increased taxation. The concept of value in healthcare has been largely avoided, especially within a private health care system like in the US. The value approach has attracted considerable attention in many countries employing a single payer system (i.e. Canada) leading to a near 50% reduction in health care spending per capita compared to the US [14] without compromising health care. The concept of value is simple, benefit (i.e. patient survival or quality of life) divided by cost but in practice, it is more difficult to define. As costs continue to soar with the development of new drugs, imaging and technology the value we are receiving is decreasing and this rate of development may challenge the benefits of randomized clinical trials. To justify higher cost options healthcare service providers must provide evidence of extended quantity and quality of life obtained through impactful outcome research. Outcome research is an integral part of the value discussion helping to define the real-world benefit from new treatments and technologies. Only with a clear definition of both the benefits and the opportunity costs of medical treatments can healthcare become sustainable. This issue of value-centred care has been discussed in radiation oncology [15-17] where patient outcomes and cost can vary greatly.

The difficulty for radiation oncology is how to effectively measure benefits and cost while accounting for both patient-specific parameters and radiation delivery options. As radiation treatments progress there are biological and anatomical changes that impact how the radiation is received making it difficult to compare treatments based only on their initial plans. For example, recent studies predicting patient outcomes avoid including
specifics on how the treatment was delivered [18, 19] and consider using more advanced imaging parameters like Radiomics [20]. Even in studies trying to model radiation effects on TCP and NTCP only the prescribed dose is considered [21, 22]. For a true assessment of value, a better understanding of what treatment was delivered is necessary otherwise any real effect could be misinterpreted.

In chapter 2 we have presented a determination of overall survival for liver cancer patients and include treatment parameters, and demonstrate the poor prognostic value of using only the prescribed dose. A multivariable analysis presented the possibility of the advantage of including additional parameters namely a uniform low dose to surrounding tissue [23]. Previous predictive liver cancer models had not included the effect of radiation dose [24, 25] or did not include all other pertinent liver function and disease parameters [26].

The inclusion of the actual delivered dose that accounts for daily anatomy changes enhances the predictive power. Decision making based on complex data including pre-treatment patient characteristics, planning parameters, delivered dose is technically feasible in the era of integrated electronic records and treatment planning systems as demonstrated in our previous work on the design of a clinical database system [27]. With the complete picture of treatment delivery more clinical trials and machine learning can be performed to obtain a better understanding of the value added by new treatment techniques.

1.5 Image guided radiation therapy

Typical radiation therapy treatments are delivered over multiple fractions ranging from three fractions delivered over the course of a week, to 30 or more. As treatment progresses it is difficult to determine how well the delivered dose matches the prescribed dose without additional imaging and tracking data. Historically as radiation treatment plans became more advanced daily imaging was introduced to ensure patient alignment to the planning position before each treatment and check for significant anatomical changes. This process is called Image-guided radiation therapy (IGRT) and involves patient setup using external markers (tattoos and immobilization devices) followed by imaging
performed with onboard megavoltage (MV) or kilovoltage (kV) imaging systems. Once images have been obtained couch shifts are applied to ensure the patient is in the correct (planned) position relative to the treatment field. Current forms of image guidance include 2D kV, MV CT (used on Tomotherapy), CT on rails and cone beam CT (CBCT).

A radiation therapy treatment plan is optimized to deliver a radiation dose to the gross tumour volume (GTV) with margins added to incorporate potential microscopic disease spread creating a clinical target volume (CTV). Furthermore, to account for set up and radiation delivery errors an additional margin is introduced creating the planning target volume (PTV). These margins were defined by the international commission on radiation units and measurements (ICRU) reports 50 [28], 62 [29] and 83 [30]. With increased margins delivering higher levels of radiation to the PTV became challenging because the PTV can overlap or be in close proximity to sensitive organs. Image-guided radiation therapy improves the accuracy of treatment delivery with smaller PTV margins and allows for target conformal treatments with higher doses.

Clinically IGRT has paved the way for more aggressive treatment options providing better outcomes for patients without the risk of toxicities. A study by Zelefsky et al. demonstrated reduced urinary toxicity, and improved survival for high-risk prostate cancer patients using IGRT, compared to standard IMRT [31]. There have been recent studies describing the improvements brought by IGRT for prostate [32-34], lung [35], brain [36] and head and neck [37]. As the cost of implementing IGRT is large it is important to optimize the IGRT process by making it patient specific and determine the sub-group of patients and disease sites where IGRT has a significant impact on clinical outcomes. IGRT also provides accurate records of the patient’s anatomy throughout treatment.

1.6 Adaptive radiation therapy

Radiotherapy is planned on a patient’s static anatomy provided by CT simulation before the start of treatment. As the treatment fractions are delivered the patient’s anatomy often changes and affects how the dose is deposited. These dose changes to the target and surrounding healthy tissue influence disease control and/or increase the risk of
complications. To counteract anatomical changes the treatment plan can be modified via adaptive radiation therapy to minimize deviation from the original plan. Such modification is only needed for sites where anatomical changes are important based on the impact on the cumulative dose distribution.

Adaptive radiation therapy is the modification of a plan during a multi-fraction radiation treatment in response to changes in the patient’s anatomy and disease. These changes can include weight loss, body functions (bladder/rectum filling/emptying), breathing, setup errors and tumour progression or shrinkage. Multiple studies have shown significant benefits in survival and reduced toxicity by adapting treatments [38, 39]. Unfortunately, there is currently no standard adaptation strategy with different centres and studies having different approaches. Nijkamp et al. adapted treatment after fraction 6 out of 39, using the average PTV measured from the previous 6 treatments [40]. Whereas earlier studies would perform a new CT simulation after a specific amount of dose has been delivered to a tumour [41, 42]. Most institutions use image guidance to monitor the patient’s anatomy, using different image matching techniques and action levels. When large changes are observed the treating physician can require an adaptation to their treatment and he/she is responsible for the final decision. Such a decision is typically based on observations of anatomy changes only, while the clinical assessment should include changes in the dose distribution. Currently, this dose-effect evaluation is achieved only by repeated full CT-simulation with contouring and dose calculation. In chapter 5 a method is presented to provide an automatic daily evaluation of the necessity of plan adaptation for head and neck cancer without additional imaging and dose calculations. Without a standard approach, comparison of treatment outcomes with respect to radiation treatment is difficult since they will be adapted at different points.

Any radiation treatment plan is dependent on the imaging acquired for diagnostic or planning to provide a complete picture of a patient’s anatomy and disease. All modern imaging modalities produce a 3D volume with voxel values representing electron density, cellular structure or cellular processes. Radiation therapy is targeted to voxels suspected of cancer and avoids voxels of sensitive structures typically highlighted by contours. Unfortunately, contoured regions are only present on the planning CT simulation and can
change significantly throughout treatment. No longer can the planned dose to individual voxels or structures be used to correlate with outcomes without accounting for anatomy change and plan adaptation. Currently, the only effective method to track dose changes from anatomy change and to perform plan adaptation is with deformable image registration.

1.7 Deformable image registration

1.7.1 Background

Image registration is the process of pairing two image volumes together that are both derived from the same object. There are multiple objectives for employing image registration, for example, registering images of an object viewed from different points, different image modalities or after experiencing deformation. Each of these scenarios occurs frequently for medical imaging using CT, magnetic resonance imaging (MRI) and positron emission tomography (PET). Independent of the application, the goal of image registration is to map every pixel on every slice (or voxel in 3D) from one image space to another.

In radiotherapy, image registration is primarily used to delineate cancerous regions (GTV) and sensitive structures from diagnostic images (MRI or PET) and propagate the contours to the planning CT study. MRI is commonly used for superior soft tissue contrast compared to CT for target delineation for many treatment sites including prostate [43], cervical [44], brain [45], and head and neck [46]. But MRI does not contain the electron density values required for radiotherapy planning. Image registration can facilitate the transfer of diagnostic information from MRI to CT to facilitate dose calculation and treatment planning. An example of improved target delineation is presented in Figure 1-3a for the bladder, tumour and rectum in the female pelvis [44]. Another important accompanying imaging modality is PET, which can locate and stage disease using its metabolic activity detected by specifically labelled radioisotopes, primarily fluoride 18. PET imaging has greatly improved target delineations [47] especially for thoracic cancers [48, 49]. When a PET image is registered to a planning CT, a “fused” image is generated indicating areas of glucose uptake or cancer aggressiveness as a colour map shown in Figure 1-3b.
As treatment methods improved, IGRT and adaptive radiation therapy became common, traditional rigid registration was no longer always suitable because patient’s anatomy would deform non-linearly throughout treatment. Now multiple image studies are collected at different time points which are anatomically different. To solve this problem deformable image registration (DIR) was introduced [51] to map every point between image spaces non-linearly, producing a deformation vector field (DVF), an example illustrated in Figure 1-4. A DVF can deform contours from the planning image to daily images and after calculating the dose distribution compare and accumulate the dose from each fraction. The final goal is to effectively track and accumulate the radiation dose being delivered from each fraction to correlate real delivered dose with patient outcomes.
1.7.2 DIR fundamentals

The results of a DIR for an individual point can be interpreted as the vector between the point from the original image space \((x,y,z)\) to the new image space \((x',y',z')\). When used for 3D medical image volumes the problem becomes complex by assigning deformation vectors to millions of voxels with different physical properties. DIR has been a long-standing tool in computer vision with its recent surge in popularity due to improvements in data storage and computation [52]. Registration algorithms can be defined by four main characteristics: i) transformation type ii) image similarity metric iii) regularization, and iv) optimization. Many fundamentals of image registration have been explored in the textbooks *Medical Image Registration* by J Hajnal [53] and *Numerical Methods for Image Registration* by J Modersitzki [54]. Other informative journal articles describing the fundamentals behind DIR are by Glocker et al. [55] and Oh et al. [56].

1.7.3 Transformation

When two image volumes (referred to as the moving and final image volumes) are presented the first question is how to match them together or deform the moving image to match the final. The simplest option is to use a combination of affine transformations.
including translation, rotation and scaling producing a “rigid registration” with the shape of objects conserved. As the complexity of the transformation increases more parameters are added to explain the local deformation at different points throughout the image. The deformation of these points can be driven by different methods and linear combinations of spline functions (e.g. B-spline) are commonly used. B-spline methods linearly combine different B-spline functions to deform all points from the moving image to the final image. Interpolation between control points can also be performed following the combined B-spline functions, an example of a B-spline interpolation is shown in Figure 1-5a. Spline-based transformations are used in multiple studies [57, 58] and within the open source Insight segmentation and registration toolkit (ITK) registration software [59]. Splines are a parametric model and can be solved rapidly since each point is not solved individually with a resulting DVF looking similar to Figure 1-4.

Different from spline methods the popular ‘Demons’ method is able to represent the DVF at each voxel of the volume by calculating a demons force at each voxel based on the local intensity gradients. The demons method has been adapted in many commercial systems [60-61] because of high-resolution DVFs but has difficulties driving deformations in regions of low-intensity gradients. A demons transformation can produce fine local deformations as shown in Figure 1-5b for an example of the shrinking of a small structure.

![Figure 1-5: a) Example of a 1D spline interpolation b) example of demons DVF vector for a small object shrinking from the bottom. Produced in MATLAB.](image-url)
To better model physical systems improved algorithms were developed using finite element modelling [62] and Hooke’s law to govern the resistance to deformation and elasticity of biological tissue. Most medical CT and MRI image studies have important organs contoured and can also highlight regions of bones, air and soft tissue all with different intrinsic mechanical properties. Applying the different mechanical properties from contoured images with finite element methods allowed for a biomechanical model for image transformation producing highly accurate results, implemented in the Ray station treatment planning systems developed from RaySearch Laboratories [63]. Though its results have been promising it currently requires contoured images for the best results, which are difficult and time-consuming to produce with daily image guidance. Contoured images are used to assign mechanical properties to specific voxels, and to model the surface deformations of structures.

1.7.4 Image similarity

With a transform applied to the moving image, a measure of image similarity is required to evaluate how accurately the deformed volume matches the final volume. There are two main categories of image similarity metrics based on landmarks and image intensity values. Landmark or feature methods locate important unique regions including corners, edges or object centres and try to match their position between the transformed moving and final image volume. Though simple to understand, landmark methods are rarely used for driving a DIR because of large calculation times and the requirement of the objects to be imaged clearly. Methods based on image intensity are more general and have been the standard for all DIR algorithms requiring only the intensity values at each voxel.

Image similarity metrics based solely on image intensity values are the mathematical relationships between two data sets, arranged in a 2D or 3D matrix. One of the simplest metrics is the sum of squared difference (SSD).

$$SSD = \frac{1}{N} \sum_{i} (X_{Mi} - X_{Fi})^2$$  \hspace{1cm} (1)

Where $N$ is the total number of voxels $i$, $X_M$ and $X_F$ are the voxel intensity values for the deformed moving (subscript M) and final (subscript F) image, respectively. SSD is very
easy to calculate only requiring $N$ operations for any image, but it is limited and has
difficulties with noisy images and any differences in imaging protocol or modality. To
improve upon the SSD other metrics are used including normalized cross correlation
(NCC), equation (2), with standard deviation $\sigma$ and $\bar{X}$ being the mean value per image.

$$NCC = \sum_n \frac{1}{\sigma_M \sigma_F} (X_{Mi} - \bar{X}_M)(X_{Fi} - \bar{X}_F)$$  \hspace{1cm} (2)

For cross modality DIR (for example CT to MRI, or CT to CBCT) both SSD and NCC
metrics can suffer due to non-linear differences in image intensity values in the same
material (i.e. bone or soft tissue).

A different approach is the concept of mutual information (MI), evaluating not the image
intensities but analyzing the “information” present in both images. MI is calculated by
looking at the probability that intensity values from the moving image match with values
from the final images. For example, if the probability of value $X$ from the moving image
matching value $Y$ from the final image is very high, the agreement is consistent. If the
probability of other value pairs is high, you can assume that the registration is accurate. A
more in-depth explanation of MI can be found in Hill et al [64]. Any DIR with multi-
modal imaging (e.g. CT-MRI) exclusively uses MI to overcome the conceptually different
voxels values [65].

1.7.5 Regularization

Applying different image transformations to maximize image similarity can produce
various solutions that “best” match two images together. Unfortunately, image similarity
metrics fail to consider the physical limitations of the objects, producing solutions that
deform across large distances or collapse many voxels down to a single point. To
constrain the registration process, an additional regularization term is used in conjunction
with the image similarity. As the image similarity drives the registration towards
matching both images, the regularization term will penalize certain transformations based
on the application. For example, registrations of the brain and skull will impose strict
volume conservation, unlike registrations of the abdomen or thorax which would apply
regularization to ensure smoothness. An example of a regularization term
\( R(T) = \alpha \int \| \nabla T(x) \| ^2 dx \) (3)

promotes smoothness by penalizing first order derivatives, where \( T \) is the transform (or DVF) and \( \alpha \) is a weighting parameter.

1.7.6 Different algorithms

Through various combinations of transformations, image similarity metrics and regularization different DVFs will yield significantly different results. Depending on the final application or software environment, different commercial vendors have implemented their own unique algorithms. Popular commercial products for radiation therapy include RayStation (RaySearch laboratories, Stockholm, Sweden), MIM Software (Cleveland, OH, USA), Velocity (Varian Medical Systems, Palo Alto, CA). For the treatment planning system Raystation the DIR algorithm named ANACONDA (explained in more detail by O. Weistrand et al [63]) is incorporated, using a correlation coefficient similarity metric and two forms of regularization ensuring smoothness and shape based regularization. Shape-based regularization assumes that anatomical regions are defined by 3D meshes and the DIR must deform the mesh in a consistent and physical manner. Also, the ANACONDA algorithm can include contour information both as a penalty term in the regularization, or simply to focus the registration to important regions.

Unlike Raystation, the algorithm from MIM software is less documented but is an intensity-based free-form algorithm, using the sum of the squared difference as the similarity metric [66]. Velocity uses an intensity based b-spline algorithm, with mutual information as a similarity metric [67]. Currently, commercial vendors of DIR do not report the specifics of their propriety algorithms, resulting in most DIR products being implemented as “black box” systems. Even with the most advanced DIR algorithm inaccuracies will be present when attempting to match millions of voxels between image studies and the problem gets worse when algorithm specifics are locked within black box systems. These inaccuracies can be harmless or lead to miss-registering a tumour to a sensitive structure.
1.8 Registration error

Deformable image registration has a vast scope working with deformed image sets, comparing time points, or different imaging modalities. Because DIR at first glance can appear to accurately register two images, registration errors can be very difficult to spot making additional evaluation necessary. For this reason, many studies have been published producing different methods or metrics to evaluate the “quality” or accuracy of the DIR. The goal of these studies is to provide the user with a means to compare and evaluate different registrations and provide information where a particular registration algorithm is most accurate.

The two most common methods for evaluating DIR are comparing contours or segmentation and propagated manual landmarks. Contour propagation has always been an attractive feature of DIR and comparing the generated contours to expert manual contours has been the standard for registration benchmarking. The common metric for contour comparison is the DICE coefficient measuring the overlap of two contours and it has been used in the majority of studies [68-71]. Unfortunately, contour methods fail to evaluate registration performance within structure boundaries which is essential for accurately determining accumulated radiation dose. Another common method is comparing the propagation of landmark sets produced by experts [72]. This method uses landmarks located in two image sets and finds the distances between actual landmarks and DIR predicted landmarks in the final image. Landmarks provide discrete measurements of DIR accuracy but they need to be located manually. In order to overcome this problem in Chapter 3, an automatic method using the scale-invariant feature transform is presented and tested against manual and expert landmarks.

Some early examples of DIR accuracy measures were the inverse consistency or transverse consistency errors describing how a DVF mapped the moving image to the final image and vice versa [73]. In principle consistency measures are good “warning” signs of unrealistic registrations, but can’t adequately predict the magnitude of error as shown by Bender et al. with poor correlation to registration errors [73]. Methods looking at mechanics-based metrics such as unstable energy by Li et al. [74] yielded improved correlation with registration error. Building on the idea of consistency metrics, Saleh et al.
developed the distance discordance metric (DDM) [75] that uses registrations between images within the same longitudinal study to measure consistency and demonstrated higher correlation than the original consistency metrics alone. Other studies attempted to use different statistical measures to determine registration uncertainty using bootstrapping [76] and the Cramer-Rao bound method [77]. More advanced methods were developed to specifically evaluate a DIR with respect to the physical deformation of tissue. For example, Zong et al [78] used finite element analysis to simulate the elastic properties of soft tissue to measure the internal forces.

Predominantly DIR evaluation methods are derived to compare and contrast different algorithms to best determine the most accurate and efficient algorithm for specific applications. This approach has been followed in multiple studies comparing different algorithms by comparing their DICE scores, or landmark registration errors for given test sets [79, 80]. Work by Kirby et al. constructed realistic 3D phantoms for the head and neck and male pelvis to evaluate several DIR algorithms [81, 82]. What needs to be determined is how this error affects the final results for a dose accumulation and if uncertainty bounds can be added to any final result. Figure 1-6, an example of using DIR to match a single point on the boundary of the rectum from two CT studies shows how the dose results can vary significantly depending on the registration error.

![Figure 1-6: Example of the effect of registration error.](image)

Figure 1-6: Example of the effect of registration error. Colour map is the radiation dose. Point A) miss-registration but still on the boundary, B) miss-registration still within boundary and C) miss-registration with same dose level.

These errors may not be significant for a single registration, but the errors can cumulate as the radiation dose is accumulated from multiple studies. A study by Risholm et al. demonstrated a DIR algorithm that inherently produced the uncertainty of every vector
within the DVF [83]. Using this algorithm, uncertainty bounds could be added to any final results. For example, in a later study, they added uncertainty bounds to a dose accumulation showing their impact on the final accumulated dose distribution [84]. Though the work by Risholm et al. demonstrated the possibility of applying uncertainty information, it was only possible with their custom DIR algorithm that was not designed for clinical practice. What is required for clinical practice is for commercial DIR algorithms to provide DIR uncertainty and accuracy information for each DVF to provide realistic results? In chapter 4 a method of incorporating predicted DIR uncertainty into a commercial DIR algorithm is demonstrated. Having a characterization of the uncertainty of DIR with deformable dose accumulation can provide physicians with useful information for making a decision on plan adaptation and dose prescriptions. Therefore, including registration uncertainty in the dose accumulation process can provide a more accurate account of radiation delivery throughout a multi-fraction treatment, producing more powerful predictive models of disease control and normal tissue complications. With more powerful predictive models treatment options can be determined uniquely for each patient and their desired outcome creating personalized medicine in radiotherapy.

1.9 Decision making in radiation therapy

Since 1951, when the first Co-60 radiation treatment unit was used in London, Ontario for cancer care, tremendous progress in technology has occurred. The huge amount of clinical evidence has been obtained through clinical trials with patient participation, laboratory experiments and various comparative studies of treatment outcomes reported in scientific journals. In practice, treating physicians are often deciding on the best possible treatment of patients quickly based primarily on the physician’s education and personal experience. Yartsev and Mackie [85] proposed a scheme for decision making that involves creating a model for radiation therapy based on a large and constantly growing database of patient’s personal features, parameters of treatment, clinical outcomes and patient’s preferences. An important aspect of such a model is quality data accounting for all elements of treatment to facilitate machine learning methods and allow the models to incorporate incoming patient data. In this thesis, we consider several key
components of the decision-making process, schematically presented in Figure 1-7 highlighting the contributions from each chapter of the thesis.

This schematic describes the connection of the planning data (left side) and the patient data (right side) and the different processes required to link them together described through specific chapters. The first link is the modelling of patient data to determine optimal treatment plans for the desired patient outcomes. In Chapter 2 a multivariable model is produced linking patient, treatment and outcomes data producing two clinical nomograms to help guide physicians with treatment decisions. The second link is between the planning image and treatment images provided through deformable image registration and image guidance (IGRT). DIR is a key element of this link since all daily changes are deformable. Though DIR is an excellent and necessary tool, it has unknown inaccuracies that need to be considered when used clinically. In Chapter 3 an automatic DIR evaluation method was presented allowing DIR to be evaluated routinely whenever it is applied.

The third link is determining the real delivered dose from the planned dose using daily images and deformable image registration for dose accumulation. Similar to the second link deformable image registration is required and its accuracy needs to be considered in the final dose accumulation. Chapter 4 created a method to incorporate DIR uncertainty into the dose accumulation process, to determine what dose was actually delivered. The final link is between the original treatment plan and the adapted treatment plan initiated after significant anatomy change was observed. In Chapter 5 a method for quantitatively evaluating the necessity of plan adaption is presented in order to provide more accurate and consistent plan adaptation strategies. The thesis chapters each provide different necessary elements required to incorporate patient data to determine patient outcomes from their cancer treatment.
Figure 1-7: Schematic for outcome-driven decision making in personalized radiotherapy.
1.10 Research goals and objectives

The research outlined in this thesis focuses on two main goals i) to construct a method of including radiation treatment parameters for prediction of clinical outcomes ii) to provide the methods of tracking accumulated dose in multi-fraction treatment with known registration uncertainty. These goals were achieved through the following objectives including the modelling of patient survival in radiation therapy and the practical applications of DIR:

1. Find the significance of treatment and pre-treatment parameters in modelling patient’s overall survival and construct nomograms.
2. Produce and evaluate an automatic landmark generation method and test against both amateur and expert manual landmarks.
3. Determine the effect of registration error on measuring the daily dose using DIR.
4. Evaluate the ability to decide when to adapt a treatment plan for head and neck cancer patients using DIR without a dose recalculation.

1.11 Thesis outline

1.11.1 Modelling of overall survival for hepatic carcinoma or liver metastasis receiving radiotherapy (Chapter 2)

There are many parameters influencing treatment success in liver cancer including patient and treatment parameters. In this study, patients diagnosed with hepatic carcinoma and liver metastasis were analyzed with multivariable Cox regression of overall survival. The results were two nomograms for primary hepatic and metastatic patients demonstrating the impact of radiation prescription and other patient parameters on overall survival. This chapter is based on the publication in the Journal of Future Oncology titled “A multivariable model to predict survival for patients with hepatic carcinoma or liver metastasis receiving radiotherapy” by Jason Vickress, Michael Lock, Simon Lo, Stewart Gaede, Aaron Leung, Jeff Cao, Rob Barnett, and Slav Yartsev.
1.11.2 Automatic landmark generation for evaluation of deformable image registration accuracy (Chapter 3)

DIR has become a pivotal tool in radiation therapy for contour propagation, anatomical assessment and dose accumulation even though errors are common and poorly characterized. Different DIR evaluation techniques are used in practice including contour evaluation and landmark matching but both require significant user input and are difficult to apply routinely. We developed a method of automatically generating landmarks using the scale-invariant feature transform and tested it against both amateur and expert landmark sets. The results of this work were published in the Journal of Physics in Medicine and Biology titled “Automatic landmark generation for deformable image registration evaluation for 4D CT images in lung” by Jason Vickress, Jerry Battista, Rob Barnett, John Morgan, and Slav Yartsev.

1.11.3 Dosimetric impact of deformable image registration error in radiation therapy (Chapter 4)

DIR has been introduced to track anatomical changes between daily fractions to accumulate the actual radiation dose. Unfortunately, inherent errors in DIR can lead to significant changes in dosimetric analysis and need to be accounted for in any analysis. In this study, we derived a method to incorporate predicted registration uncertainty into the final dosimetric analysis using expert landmarked lung 4DCT studies. This chapter is based on the publication in the Journal of Physics in Medicine and Biology titled “Representing the dosimetric impact of deformable image registration errors” by Jason Vickress, Jerry Battista, Rob Barnett, and Slav Yartsev.

1.11.4 Daily assessment of dose change using deformable image registration without dose-recalculation (Chapter 5)

Adaptive radiation therapy has been successful for many treatment sites where anatomical changes are frequent and can cause the dose distribution to deviate from what was planned. Currently plan adaptation is triggered by manual evaluations of daily CBCT images by therapists, physicists and oncologists leading to either plan adaptation or to continue the original plan. Using DIR and the planned dose distribution an automatic method for determining the necessity of plan adaptation was developed and tested against
the gold standard of using a new CT study with a new dose calculation. The results of this method have been submitted to the Journal of Applied Clinical Medical Physics titled “Online daily assessment of dose change in head and neck radiotherapy without dose-recalculation” by Jason Vickress, Jerry Battista, Rob Barnett, and Slav Yartsev.

1.12 References

1 Rising life expectancy: a global history, James C. Riley, 2001 University of Cambridge, Cambridge, UK


K. Takeda et al, "Treatment outcome of high-dose image-guided intensity-modulated radiotherapy using intra-prostate fiducial markers for localized prostate


51 W. Lu et al, "Deformable registration of the planning image (kVCT) and the daily images (MVCT) for adaptive radiation therapy," *Physics in Medicine and Biology*, vol. 51, (17), pp. 4357-4374, 2006.


Chapter 2

2 A multivariable model for the prediction of survival for patients with hepatic carcinoma or liver metastasis receiving radiotherapy

This chapter has been previously published as “A multivariable model for the prediction of survival for patients with hepatic carcinoma or liver metastasis receiving radiotherapy” published in *Future Oncology*, 13 (1), 19-30 by Jason Vickress, Michael Lock, Stewart Gaede, Aaron Leung, Jeff Cao, Rob Barnett and Slav Yartsev. Permission to reproduce this article was provided by the publisher and presented in Appendix A.3. Additional explanation of clinical tools and parameters is presented in Appendix B.

2.1 Introduction

Liver cancer, as of 2013, has the sixth highest incidence rate with the third worst survival rate worldwide [1]. Historically, radiation therapy has played a limited role in the treatment of liver cancer because of the normal liver’s inherent radiosensitivity. However, studies on the radiobiological effects of radiation on the liver provided the ability to determine suitable radiation dose levels, and on how to select patients appropriate for radiation have resulted in an increasingly important role for radiation [2].

The advent of stereotactic body radiotherapy (SBRT) made possible the delivery of high radiation doses to the target while sufficiently sparing the surrounding healthy tissues, which is critical for the radiotherapy of liver cancers. Studies have shown significant benefits to local control and survival [3-7] using more intensive treatments at the cost of potential toxicity [8,9]. Recently, two comprehensive summaries of the literature have established the current knowledge on hepatic radiation. Klein et al. reviewed hepatocellular carcinoma radiotherapy demonstrating a local control rate of 80-90% for selected patients with smaller tumours and no portal venous thrombosis [10]. This thorough review also confirms that low rates of adverse events have been found.
However, the review also confirms the wide variation in treatment technique and patient selection criteria for radiation. This makes interpretation and comparison of trials difficult. For example, comparison of trial A that included patients with a better outcome than trial B would result in the inappropriate conclusion that trial A treatment technique was better. Nomograms and risk stratification play an important role in allowing a comparison of study outcomes and selection of the most appropriate patients for future trials.

For metastatic disease, a similar review performed by Hoyer et al. plus an international survey demonstrated similar results and made similar recommendations to the hepatocellular review. For metastatic lesions, 1-year local control rates ranged widely from 23% to 95% likely due to a lack of consensus on the selection of patients [11, 12]. Despite multiple well-performed studies, radiation is not included in many of the major guidelines used to treat liver lesions [13]. Why is this? First, large-scale multicentre trials in this patient group have not been performed despite being a cancer with one of the highest death rates and a research priority. Trial accrual is often difficult due to the rapid progression of the disease, variability in treatment availability and the patient’s reluctance to be randomized to a treatment arm with a known poor outcome or one with evidence collected primarily from relatively small case series. Second, treatment-related factors have not been properly established. Therefore, individual centres are performing liver radiation using a wide variety of strategies especially variable doses, fractionation and patient selection [12]. Third, the patient population is very heterogeneous with respect to presentation, previous treatments, the extent of disease, and location of disease requiring larger sample sizes. A larger number of patients need to be investigated to be able to statistically reveal important information from such a heterogeneous population. Lastly, current analyses have not always accessed the vast amount of physics and dosimetric data available and have previously focused on clinical and imaging parameters, such as alpha-fetoprotein (AFP) or tumour size. Furthermore, patient numbers in individual centres are insufficient to provide the higher number of data points required to assess the potentially numerous critical parameters such as appropriate dose [14].
To address these factors, our ability to properly combine and analyze data collected from multiple institutions is critical; this will lead to practical treatment strategies based on information from comparable groups and clinical trials where outcomes can be assessed as better or worse than previously published trials. The use of a predictive tool will assist in standardizing patient selection based on risk group, thereby providing a greater ability to accrue to multicenter trials and compare outcomes. This methodology may also promote the inclusion of physics and dosimetric parameters not usually included in nomograms. Finally, a tool to provide consistent outcome data in this heterogeneous population will provide information a) on which patients may benefit from radiotherapy to referring physicians, and b) for administrators, clinical trialists and guideline developers on the expected outcome of patients receiving radiation.

Using large collections of patient data, clinical models of survival and recurrence have been successfully developed for multiple treatment sites in radiation therapy [15, 16], including liver cancer [17]. In this publication, we develop a multivariable prognostic model for predicting overall survival after radiation therapy for patients with primary and metastatic liver cancer based on data from one of the largest published liver databases as proof of principle of this methodology.

2.2 Methods

2.2.1 Patient population

The patient population consisted of 195 patients receiving radiation treatment for hepatocellular carcinoma (HCC) or liver metastasis (Mets) in a comprehensive tertiary cancer facility from 2004 to 2015, with demographics shown in Table 2-1. Patients were considered for the study if they were ineligible for standard treatment (TACE, surgery or RFA) and received radiation treatment, but were excluded if they received low dose palliative treatment. Sorafenib was the principle chemotherapy provided. However, many of the patients were entered into trial prior to the release of sorafenib. Therefore, chemotherapy included the previous standard which was doxorubicin. Study endpoints, such as overall survival (OS) and disease-free survival (DFS), were collected from patient charts and electronic records. All data collection and analysis was approved by the local ethics board.
Table 2-1: Patient demographics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>HCC (N = 66)</th>
<th>Mets (N = 129)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>70 ± 12</td>
<td>65 ± 12</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>56 (85%)</td>
<td>67 (52%)</td>
</tr>
<tr>
<td>Female</td>
<td>10 (15%)</td>
<td>62 (48%)</td>
</tr>
<tr>
<td>tumour volume (cm³)</td>
<td>384±447</td>
<td>351±870</td>
</tr>
<tr>
<td># of lesions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>45 (68%)</td>
<td>65 (50%)</td>
</tr>
<tr>
<td>2</td>
<td>7 (11%)</td>
<td>32 (25%)</td>
</tr>
<tr>
<td>3</td>
<td>10 (15%)</td>
<td>19 (15%)</td>
</tr>
<tr>
<td>&gt;3</td>
<td>4 (6%)</td>
<td>13 (10%)</td>
</tr>
<tr>
<td>CP Grade</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>42 (64%)</td>
<td>99 (77%)</td>
</tr>
<tr>
<td>B</td>
<td>22 (33%)</td>
<td>28 (22%)</td>
</tr>
<tr>
<td>C</td>
<td>2 (3%)</td>
<td>2 (1.5%)</td>
</tr>
<tr>
<td>unknown</td>
<td>0 (0%)</td>
<td>2 (1.5%)</td>
</tr>
<tr>
<td>Bilirubin</td>
<td>21.6±29</td>
<td>17±32.7</td>
</tr>
<tr>
<td>Serum Albumin</td>
<td>35±6</td>
<td>38±6</td>
</tr>
<tr>
<td>Ascites</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>19 (29%)</td>
<td>12 (9%)</td>
</tr>
<tr>
<td>N</td>
<td>47 (71%)</td>
<td>117 (91%)</td>
</tr>
<tr>
<td>primary site (N = 130)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>colorectal carcinoma</td>
<td>52 (40%)</td>
<td></td>
</tr>
<tr>
<td>cholangiocarcinoma</td>
<td>22 (17%)</td>
<td></td>
</tr>
<tr>
<td>neuroendocrine carcinoma</td>
<td>10 (8%)</td>
<td></td>
</tr>
<tr>
<td>breast carcinoma</td>
<td>10 (8%)</td>
<td></td>
</tr>
<tr>
<td>non-small cell lung carcinoma</td>
<td>10 (8%)</td>
<td></td>
</tr>
<tr>
<td>pancreatic carcinoma</td>
<td>7 (5%)</td>
<td></td>
</tr>
<tr>
<td>renal cell carcinoma</td>
<td>3 (2%)</td>
<td></td>
</tr>
<tr>
<td>metastatic melanoma</td>
<td>3 (2%)</td>
<td></td>
</tr>
<tr>
<td>gastric carcinoma</td>
<td>3 (2%)</td>
<td></td>
</tr>
<tr>
<td>esophageal</td>
<td>3 (2%)</td>
<td></td>
</tr>
<tr>
<td>other</td>
<td>6 (4%)</td>
<td></td>
</tr>
<tr>
<td>previous treatments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chemo embolization</td>
<td>17 (26%)</td>
<td>7 (5%)</td>
</tr>
<tr>
<td>I-131 lipiodal</td>
<td>7 (11%)</td>
<td>6 (5%)</td>
</tr>
<tr>
<td>radiofrequency ablation</td>
<td>3 (5%)</td>
<td>9 (7%)</td>
</tr>
<tr>
<td>chemotherapy</td>
<td>6 (9%)</td>
<td>95 (74%)</td>
</tr>
<tr>
<td>liver resection</td>
<td>5 (8%)</td>
<td>32 (25%)</td>
</tr>
<tr>
<td>previous abdomen radiation</td>
<td>0</td>
<td>9 (7%)</td>
</tr>
</tbody>
</table>

HCC - hepatocellular carcinoma, Mets - metastatic disease, CP - Child-Pugh
2.2.2 Radiation treatment

Patients received radiation treatment using Volumetric Modulated Arc Therapy (VMAT), fixed beam Intensity Modulated Radiation Therapy (IMRT), three-dimensional conformal radiotherapy (3DCRT), or helical Tomotherapy (Tomo). Radiation dose prescriptions varied between 15 and 88 Gy in 5 to 35 fractions using a radiobiologically-guided prescription program similar to the radiobiological NTCP dose calculation published by Dawson [18, 19]. Patients receiving 5 or 6 fractions were irradiated every other day, and the remaining patients were treated daily for 10-35 fractions. Most patients were treated using a respiratory gated and image-guided technique. The remainder of patients, who were either unable to hold their breath or had irregular breathing, were treated with full respiratory motion using a complete internal target volume (ITV). Gross tumour volume (GTV) delineation was performed on the primary planning contrast-enhanced 4DCT images.

2.2.3 Statistical analysis

The effect of each variable on OS was evaluated using the Cox proportional hazards model. Continuous variables entered into the model were biologically effective dose to the tumour ($\alpha/\beta = 10$ Gy) in 2 Gy fractions (EQD$_2$), volume of normal liver receiving more than 24 Gy total dose (V24), GTV, bilirubin count, serum albumin, prothrombin, Child-Pugh (CP) score and age. Dichotomous variables; ascites, prior chemotherapy, and colorectal primary disease, were assigned values of 0 or 1. Hepatic encephalopathy was not entered into the model because only 3 patients exhibited symptoms. For univariate analysis, patients were stratified into several subgroups based on (i) liver function (CP A vs CP B+C) and (ii) primary HCC vs metastasis (Mets). Hazard ratios (HR) were defined by dividing the partial hazard of the third quartile by the first quartile to get an effect across the spectrum of clinical values.

Identification of the most useful variables was performed with univariate Cox proportional hazard modelling, followed by multivariable analysis performed using a stepwise selection with a significance level of 0.05. Once the most significant variables were identified, nomograms predicting OS at one year were assembled for both HCC and Mets subgroups. Nomograms were evaluated by a concordance index (C-index) [20].
using four-fold cross-validation with 200 iterations to reduce bias. The C-index measures the discrimination of the data and the values range from 0.5 (no discrimination) to 1.0 (perfect discrimination). Model calibration was performed by grouping patients according to their nomogram-predicted probabilities and comparing them against the observed Kaplan Meir Overall Survival estimate, presented as calibration curves. The goodness of fit of the calibration curves are evaluated by the Hosmer-Lemeshow (HL) test, with a value greater than 0.05 having a strong fit.

2.3 Results

2.3.1 Patient demographics

The heterogeneous patient population used in this study allowed us to investigate the effect of a large number of parameters on the outcome of radiation therapy (Table 2-1). The majority of patients had liver metastases (n = 129), but the number of patients with primary HCC (n = 66) was still sufficient for statistical comparison. Because of developing technologies within 2004-2015, the radiation treatment techniques evolved from 3DCRT (n = 58) to IMRT (n = 68), to VMAT (n = 62) and Tomo (n=7). The prescribed planning tumour volume (PTV) doses had a wide range of 15 – 88 Gy (median 45 Gy). Treatment schedules were either treated every other day for 5-6 fractions (58% of patients) or daily treatment for 10-35 fractions (42% of patients). Patients had varying liver functions with CP scores ranging from 5 to 10 (median 6). The overall survival for patients with HCC or liver metastasis is described by the Kaplan Meir curve in Figure 2-1 with 12/66 and 33/129 missing events for HCC and liver metastasis groups respectively with a median follow up of 37 months.
2.3.2 Effect of liver function

To investigate the impact of liver function before radiation treatment, patients were stratified into two groups based on their overall pre-treatment liver function, as defined by CP score A vs score B+C. Cox regression analysis was performed on both patient groups and the effect of radiation treatment parameters and liver function on OS is shown in Table 2-2. For patients with a CP score A, increased radiation dose (EQD2) showed a significant survival benefit. In both CP A and CP B+C groups, poor liver function parameters (CP Score and Serum Albumin) and larger GTV size had a negative effect on overall survival.
Table 2-2: Univariate Cox regression analysis of two CP score groups for OS.

<table>
<thead>
<tr>
<th>Factor</th>
<th>CP-A HR</th>
<th>95% CI</th>
<th>p-value</th>
<th>CP B/C HR</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GTV</td>
<td>1.18</td>
<td>1.03 - 1.34</td>
<td><strong>0.014</strong></td>
<td>1.19</td>
<td>1.07 - 1.32</td>
<td><strong>0.001</strong></td>
</tr>
<tr>
<td>EQD2Gy</td>
<td>0.61</td>
<td>0.46 - 0.8</td>
<td>&lt; <strong>0.001</strong></td>
<td>0.84</td>
<td>0.49 - 1.45</td>
<td>0.538</td>
</tr>
<tr>
<td>V24</td>
<td>0.94</td>
<td>0.70 - 1.25</td>
<td>0.654</td>
<td>1.20</td>
<td>0.81 - 1.79</td>
<td>0.366</td>
</tr>
<tr>
<td>Bilirubin</td>
<td>1.14</td>
<td>0.87 - 1.5</td>
<td>0.326</td>
<td>1.12</td>
<td>1.06 - 1.18</td>
<td>&lt; <strong>0.001</strong></td>
</tr>
<tr>
<td>Serum Albumin</td>
<td>0.54</td>
<td>0.37 - 0.79</td>
<td><strong>0.001</strong></td>
<td>0.41</td>
<td>0.27 - 0.63</td>
<td>&lt; <strong>0.001</strong></td>
</tr>
<tr>
<td>Ascites</td>
<td>1.34</td>
<td>0.58 - 3.08</td>
<td>0.487</td>
<td>1.54</td>
<td>0.84 - 2.82</td>
<td>0.160</td>
</tr>
<tr>
<td>Prothrombin</td>
<td>1.22</td>
<td>0.92 - 1.61</td>
<td>0.167</td>
<td>0.87</td>
<td>0.74 - 1.02</td>
<td>0.080</td>
</tr>
<tr>
<td>CP Score</td>
<td>2.65</td>
<td>1.18 - 5.97</td>
<td><strong>0.018</strong></td>
<td>3.95</td>
<td>2.02 - 7.75</td>
<td>&lt; <strong>0.001</strong></td>
</tr>
<tr>
<td>Age</td>
<td>1.06</td>
<td>0.84 - 1.35</td>
<td>0.617</td>
<td>0.77</td>
<td>0.48 - 1.21</td>
<td>0.252</td>
</tr>
<tr>
<td># of Lesions</td>
<td>0.92</td>
<td>0.79 - 1.08</td>
<td>0.33</td>
<td>0.91</td>
<td>0.74 - 1.11</td>
<td>0.35</td>
</tr>
<tr>
<td>Extrahepatic</td>
<td>1.33</td>
<td>0.91 - 1.95</td>
<td>0.14</td>
<td>1.63</td>
<td>0.89 - 2.98</td>
<td>0.11</td>
</tr>
<tr>
<td>Chemo</td>
<td>0.95</td>
<td>0.65 - 1.39</td>
<td>0.795</td>
<td>0.82</td>
<td>0.45 - 1.51</td>
<td>0.532</td>
</tr>
</tbody>
</table>

GTV – gross tumour volume, EQD2Gy – equivalent dose in 2 Gy fractions, V24 – the volume of normal liver receiving 24 Gy
CP – child pugh

2.3.3 Primary HCC vs liver metastasis

Based on our analysis, two nomograms were constructed for the OS of patients treated for a primary HCC (n = 66) or liver metastasis (n = 129). The results of the univariate variable selection and multivariate modelling are presented in Table 2-3. The final model variables are presented in the multivariable columns of Table 2-3, with the full model described in Table 2-4.

Table 2-3: Univariate and multivariable Cox regression results for both HCC and metastasis.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Univariate</th>
<th>Multivariate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HCC</td>
<td></td>
</tr>
<tr>
<td>Factor</td>
<td>HR</td>
<td>95% CI</td>
</tr>
<tr>
<td>-----------------</td>
<td>------------</td>
<td>-------------</td>
</tr>
<tr>
<td>GTV</td>
<td>1.68</td>
<td>1.3 - 2.17</td>
</tr>
<tr>
<td>EQD2Gy</td>
<td>0.51</td>
<td>0.34 - 0.78</td>
</tr>
<tr>
<td>V24</td>
<td>1.28</td>
<td>0.9 - 1.81</td>
</tr>
<tr>
<td>Bilirubin</td>
<td>1.34</td>
<td>1.15 - 1.56</td>
</tr>
<tr>
<td>Serum Albumin</td>
<td>0.39</td>
<td>0.25 - 0.62</td>
</tr>
<tr>
<td>Ascites</td>
<td>2.18</td>
<td>1.16 - 4.09</td>
</tr>
<tr>
<td>Prothrombin</td>
<td>0.95</td>
<td>0.81 - 1.12</td>
</tr>
<tr>
<td>CP Score</td>
<td>2.25</td>
<td>1.49 - 3.41</td>
</tr>
<tr>
<td>Age</td>
<td>1.01</td>
<td>0.74 - 1.39</td>
</tr>
<tr>
<td>Factor</td>
<td>Univariate HR</td>
<td>95% CI</td>
</tr>
<tr>
<td>---------------------</td>
<td>---------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>GTV</td>
<td>1.12</td>
<td>1.05 - 1.19</td>
</tr>
<tr>
<td>EOD2Gy</td>
<td>0.66</td>
<td>0.5 - 0.87</td>
</tr>
<tr>
<td>V24</td>
<td>0.86</td>
<td>0.61 - 1.22</td>
</tr>
<tr>
<td>Bilirubin Serum</td>
<td>1.15</td>
<td>1.1 - 1.2</td>
</tr>
<tr>
<td>Albumin</td>
<td>0.43</td>
<td>0.31 - 0.58</td>
</tr>
<tr>
<td>Ascites</td>
<td>1.00</td>
<td>1 - 1</td>
</tr>
<tr>
<td>Prothrombin</td>
<td>1.04</td>
<td>0.97 - 1.12</td>
</tr>
<tr>
<td>CP Score</td>
<td>1.64</td>
<td>1.35 - 1.99</td>
</tr>
<tr>
<td>Age</td>
<td>1.08</td>
<td>0.83 - 1.39</td>
</tr>
<tr>
<td># of Lesions</td>
<td>0.87</td>
<td>0.77 - 1.04</td>
</tr>
<tr>
<td>Extrahepatic</td>
<td>1.79</td>
<td>1.18 - 2.71</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>0.90</td>
<td>0.58 - 1.4</td>
</tr>
<tr>
<td>Colorectal</td>
<td>0.67</td>
<td>0.44 - 1.01</td>
</tr>
</tbody>
</table>


Table 2-4: Model results including β values, units and 6 months / 1-year baseline cumulative hazard ratio.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Units</th>
<th>β</th>
<th>6 month λ0</th>
<th>1 year λ0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilirubin count</td>
<td>mg/dl</td>
<td>0.0203</td>
<td>5.91</td>
<td>11.80</td>
</tr>
<tr>
<td>Serum Albumin</td>
<td>mg/dl</td>
<td>-0.104</td>
<td>0.00131</td>
<td></td>
</tr>
<tr>
<td>GTV</td>
<td>cm³</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Metastasis Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Units</th>
<th>Beta</th>
<th>6 month λ0</th>
<th>1 year λ0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilirubin count</td>
<td>mg/dl</td>
<td>0.0154</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serum Albumin</td>
<td>mg/dl</td>
<td>-0.0889</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colorectal primary site</td>
<td>yes (1), no (0)</td>
<td>-0.4774</td>
<td>11.83</td>
<td>23.07</td>
</tr>
<tr>
<td>GTV</td>
<td>cm³</td>
<td>0.000296</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

λ₀ indicates cumulative baseline hazard ratio, GTV – gross tumour volume
The nomogram predicting OS at six months and at one year for patients with a primary HCC is displayed in Figure 2-2a. The greatest effect on the patient’s survival is predicted by liver function parameters (bilirubin and serum albumin). Similar trends are found in the nomogram for metastases in Figure 2-2b, with a lower overall survival than HCC. For model evaluation, the concordance values were found to be 0.74 for the HCC model and 0.68 for the metastasis group after four-fold cross-validation, performed 200 times showing good discrimination.

![Nomogram for the probability of six months and one-year survival for patients with (a) HCC and (b) liver metastasis.](image)

Figure 2-2: Nomogram for the probability of six months and one-year survival for patients with (a) HCC and (b) liver metastasis.

The calibration curves for both models are shown in Figures 2-3 and 2-4 (with the result of the HL test) for the HCC and metastasis models respectively. Calibration curves yielded good calibration abilities for both six months and one-year survival for the HCC model (Figures 2-3a and 2-4a) and the Mets model (Figures 2-3b and 2-4b).
Figure 2-3: Kaplan–Meir survival curves for HCC (a) and Mets (b) showing the lower (solid) and higher (dashed) risk groups obtained by the multivariate models.

Figure 2-4: Calibration curves for predicting 6-month survival for HCC (a) and Mets (b). The goodness of fit was calculated using the HL test yielding a p-value of 0.1693 and 0.0412 for HCC and Mets respectively.

Kaplan Meir curves between high and low-risk patients (as determined by the model separated by the median hazard ratio) presented in Figure 2-5, showed significant separation for both the HCC and Mets models.
2.4 Discussion

2.4.1 Patient demographics

Liver cancer affects a wide range of patients with different underlying pathologies, treated with different treatment strategies and with varied functional status creating a very heterogeneous patient population that is typically seen in the literature. Though heterogeneous patient data causes problems with generalizability and requires large sample sizes to determine predictive variables; when we have sufficient patient numbers, heterogeneous treatment data provides an opportunity to clarify questions such as what dose should be prescribed as many different prescriptions are used in clinics. Heterogeneous data in this study enabled the analysis of HCC and metastasis patients with different liver functions (CP-A, or B+C) in one of the largest published datasets in the literature. This dataset included varying dose prescriptions (20 to 70 Gy), and fractionations (5 to 25 fractions), which created different biological effective doses and allowed for the analysis of the benefits of different dose prescriptions.
2.4.2 Liver function

A patient’s liver function before the start of radiation therapy is known to be a strong predictor of OS [21]. The analysis from the Cox regression for OS shows a significant difference between patients with a Child-Pugh score of A vs B+C. Patients with CP A had a significant survival advantage with increased radiation dose to a tumour, whereas patients with grade B+C did not. Previous work on patients with CP B+C has shown similar results regarding the poor response to radiation therapy [21], with similar effects from tumour size and CP score. Increased CP score post radiation therapy has also been shown as a strong predictor of overall survival [22]. Poor liver function and large tumour size correlated with decreased survival in all cases. However, only patients with better liver function showed a significant increase in survival due to higher radiation dose. Therefore, these data suggest that if patients have a poor liver function (defined as CP B+C), clinicians would provide better care by not dose escalating and unnecessarily risking toxicity. For those with good liver function, dose escalation is an important factor for improved survival. Further research is required to determine the appropriate dose.

2.4.3 Nomogram

The nomogram in Figure 2-2a for patients with HCC describes the effect of liver function and tumour size as major factors in determining patient survival. The nomogram demonstrates that liver function and tumour size are much more significant compared to an increase in radiation dose. With an average C-index of 0.74, which is substantial given the heterogeneous nature of the dataset the model for HCC is accurate in predicting OS. For example, the six months and one year OS probability would be 28% and 9%, respectively for an HCC patient with a bilirubin count of 40 mg/dl, serum albumin 30 g/dl, GTV volume of 500 cm³. For a patient with a bilirubin count of 30 mg/dl, serum albumin 30 g/dl and GTV volume of 300 cm³ six months and one year OS would be 50% and 26%. After reviewing models predicting OS for HCC patients after non-radiation treatments, this radiotherapy nomogram falls within a similar C-index range with chemoembolization 0.84 [23], immunotherapy 0.698 [24] and liver resection surgery 0.66 [25]. All these models share the same large weightings on the patient’s liver function pre-treatment. Our model also has similar results to the prognostic score ALBI (validated on a larger cohort of 2599 patients) [26] demonstrating the significance of Serum Albumin and
Bilirubin in predicting overall survival for HCC patients, without the other factors in the Child-Pugh score. We are unaware of any nomograms for liver cancer radiotherapy and believe our model is the first to be created for patients receiving radiation therapy for hepatocellular carcinoma.

The metastasis model in Figure 2-2b with an average C-index of 0.68 is less accurate than the HCC model which is expected given the more complex and heterogeneous nature of the metastatic disease. This fact is further emphasized from the HL test on the calibration curves with values close to 0.05 showing a worse fit than for the HCC model. Our analysis showed that increased radiation dose is not a significant factor in predicting patient survival. A previous study [27] has created a nomogram predicting OS for liver metastasis with an area under the curve (AUC) value of 0.83, but it included different liver parameters and had no treatment information. The trial focused on a select and relatively homogenous subpopulation of liver metastases with external validation. The trial assessed only the subpopulation of colorectal patients treated with selective internal radiation therapy. Our model is more generalizable and attempts to include a broader range of parameters. In particular, the effect of treatment parameters, especially prescription dose, to help determine when aggressive treatment is warranted and what the patient prognosis may be with radiation. We believe our nomogram is the first to be created for liver metastases patients treated with external beam radiation.

2.4.4 Potential limitations

Although our model was shown to have comparable accuracy to other models created for the treatment of liver malignancies, the heterogeneity of the data results in decreased accuracy and ability to address specific subgroups. Therefore, despite being based on one of the largest published databases, this model would benefit from more data. Data accrued from multiple centres using different dose regimens and constraints would provide the heterogeneity in data and sufficient sample size required for detailed analysis of subgroups. Using patient sample sizes of 66 for HCC and 129 for Mets can create useful models for predicting survival using only four variables. However, there may be other undiscovered variables that could improve the predictive ability of these models. With larger sample sizes, we can also develop nomograms for specific subgroup such as
cholangiocarcinomas or those patients heavily pretreated prior to radiation versus those referred early in their natural history. In this model, they are grouped together within the Mets group. Validation was successfully performed using a four-fold cross-validation, performed in 200 different random combinations, but external validation with a different data set is recommended. Determining the true accuracy of this model will require external validation, from an equivalent population of HCC and Mets patients, which is an aim for future work since the c-indexes for both models are likely to be lower in a validation set.

2.5 Conclusion

Two models in the form of practical nomograms were constructed for HCC and liver metastasis patients to help predict overall survival. The nomograms may help clinicians better select appropriate treatment strategies and provide prognostic information for patients. The nomograms provide more practical and effective use of available data from patients treated for liver cancer. Our conclusion that dose escalation is beneficial only for patients with good liver function (CP score A) but not CP B+C patients is also useful clinical information. Overall this study presents new tools for the selection of appropriate treatment options for patients diagnosed with liver malignancies.

2.6 References


Chapter 3

3 Automatic landmark generation for deformable image registration evaluation for 4D CT images of lung

This chapter has been previously published as “Automatic landmark generation for deformable image registration evaluation for 4D CT images of lung” published in the journal Physics in Medicine and Biology, 61(20), 7236 by Jason Vickress, Jerry Battista, Rob Barnett, John Morgan and Slav Yartsev. Permission to reproduce this article was provided by the publisher and presented in Appendix A.4. Additional details about the Scale Invariant Feature Transform (SIFT) is presented in Appendix C.

3.1 Introduction

Medical imaging has become a large focus in health care across both fields of diagnosis and therapy. Extensive longitudinal imaging studies require reliable and accurate image registration for scientific analysis. In diagnostic applications, changes in tissue distribution over time helps assess the progression or regression of disease prior to and following treatment. Another example is the need to track radiation dose accumulation quantitatively across a multi-fraction treatment where sharp dose gradients transform small geometric displacements into potentially large dose errors in the cumulative dose. With advancements in computing, many image registration algorithms and software platforms have been developed to perform such tasks. Unfortunately, each commercial product has hidden features and variable performance creating registration inaccuracies which impact the quality of any geometric or dosimetric analysis [1, 2]. In modern clinical practice, simple rigid registration methods are being replaced by more sophisticated deformable image registration (DIR) creating a large demand for reliable evaluation techniques that assess the accuracy of the deformation field and its effect on the radiation dose distribution.
Evaluation of image registration methods is difficult since it requires a “ground truth” benchmark derived from biophysical modelling, mechanical phantoms or landmark sets established by expert interpretation of paired images. The landmark method is currently the most popular, but it is very time-consuming as indicated in a study by Castillo et al. [3] where an average of 12 hours per computed tomography (CT) study was required to create high-quality landmarks in the lung. Another important factor is the accuracy or appropriateness of the landmark which can be limited by observer bias and/or poor image quality.

Many methods that locate landmark pairs between images automatically have been developed [4, 5] with the most promising results from the Scale Invariant Feature Transform (SIFT) algorithm [6]. Although SIFT algorithms have been shown to accurately locate landmarks in 2D images they have been less successful for 3D medical image studies [7]. Some of these problems include matching similar, non-unique landmarks and localization of the specific points within the 3D image space. Even with demonstrated improvements in the accuracy of SIFT [8, 9], limitations persist, especially in crossing imaging modalities such as matching diagnostic CT and cone-beam CT images acquired on radiotherapy machines. To improve both the resource-intensive nature of manual landmark selection and SIFT pitfalls and limitations, we present a hybrid solution involving straightforward manual editing of results obtained initially by a rapid SIFT algorithm. As DIR becomes more popular in the field of clinical evaluation, quality assurance, and “big data” mining, validation techniques need to be developed to ensure that matching landmarks are truly homologous and located correctly. Automatic landmark creation approach is an attractive option for DIR and its evaluation [10] if it is shown to have comparable accuracy and usability to manual landmarks. In this study, we compare an automatic landmark method to conventional manual landmarking, and high quality (gold standard) landmark sets to demonstrate that automated algorithms can replace manual landmarks for the validation of DIR. An example using 4DCT datasets of the lung at different respiration phases is used to illustrate the principles and advantages.
3.2 Methods

Ten thoracic 4DCT studies were obtained from the “dir-labs” group (http://www.dir-lab.com/) for use in this study. Three methods of landmark selection on the 0% and 50% phases of respiration were tested: manual, SIFT, and SIFT with manual correction (SIFT-M). Each landmarking method was compared to a landmark set produced by the dir-labs group for these images, which was considered as the gold standard (GS), having been validated in multiple publications [3]. The GS landmark set contained 300 landmark pairs manually selected between the 0% and 50% respiratory phases.

3.2.1 Manual
An in-house software tool was created in MATLAB® (version r2015a, MathWorks) to facilitate manual landmark selection and matching across both phases of a 4DCT study (0% and 50% phases). The tool allows users to search through slice images and, after applying a pre-defined window and level, to locate unique matching pairs of anatomical points across both images. Two landmark sets were created independently by two observers, to mitigate effects of observer bias. Three rules were outlined for the selection and definition of unique landmark pairs: (1) pairs must be at unique locations in three dimensional spaces, (2) pairs must be surrounded by lung parenchyma tissue and not be near or defined by an edge, (3) a unique CT feature can only be landmarked once. For practical reasons, each observer was limited to a maximum of 300 landmarks per 4DCT study, and four hours of work per study.

3.2.2 SIFT
Automatic landmark selection was performed using a modified version of the SIFT algorithm based on the original work of Lowe [6] for 2D images, starting with a double size image with three octaves each having three levels per octave. We adapted the SIFT algorithm to 3D medical CT images using the methods suggested by Allaire et al. [8] not including full landmark re-orientation. Specific parameters selected for the algorithm included $\sigma = 1.1$ for the Gaussian scaling, $H_{\text{max}} = 7$ for edge removal, $r_{\text{thresh}} = 0.01$ for contrast selection. Parameter selection was based on the work of Allaire et al., with minor modifications to improve landmark selection for the lung 4DCT studies.
Keypoint descriptors were made of 4×4×4 subgroups each containing a 2D angular histogram with 8 azimuthal directions (between 0 and 360 degrees) and 4 elevation directions (between 0 and 180 degrees) creating a 2,048 element array. Matching a key-point in the initial image is performed by finding a key-point in the final image that has the lowest Euclidian descriptor distance calculated as the distance between two 2,048 element vectors. A pair is considered unique if the second most similar point (second lowest Euclidian distance) has a Euclidian distance 1.66 times larger than the most similar point (lowest Euclidian descriptor distance). All computation was performed on a 3.7 GHz quad-core processor running a single-threaded program in MATLAB®. To reduce computation time and memory limitation, landmarking was only performed within the lung region.

3.2.3 SIFT-manual hybrid

The hybrid approach selects landmarks initially identified by the SIFT algorithm as previously described followed by a review of each landmark pair. Each landmark pair was presented to the reviewer (located in the axial plane of the image) and accepted or rejected based on visual similarity. The visual similarity criterion involves a quick evaluation of each landmark, to determine if they are both located in the same region of the image. An example of a rejected landmark pair is shown in Figure 3-1, where it is clear that the landmarks are in different slices of the image study. All landmark reviews were performed before any analysis took place, making the reviewer blind to true erroneous landmarks.

Figure 3-1: Example of an excluded landmark pair (crosses) produced from SIFT. The pair was excluded because the landmarks were located in different slices of the CT study.
3.2.4 Landmark displacement evaluation

The accuracy of each method was determined by comparison of the generated landmarks to the GS landmark set using two metrics: the mean (global) landmark displacement and proximal (local) landmark displacement.

Mean landmark displacement was calculated as the mean displacement for all landmarks within a structure and was compared to the mean displacement of the gold standard landmark set within the same structure. Such evaluation provides a global sense of magnitude agreement between displacement vectors. The structures of interest were the left and right lungs separately, creating 20 test cases in total for all ten 4DCT studies. The final results were evaluated by an Intra Class Correlation (ICC) test between each landmarking method and the GS across all 20 lung structures.

Proximal landmark evaluation was introduced to compare the displacement of a test landmark pair to the displacement of the closest GS landmark pair. To find a proximal landmark pair a test landmark pair is first selected, then a GS landmark pair is located that has the shortest Euclidian distance between their initial positions in the 0%-phase image study. Figure 3-2 illustrates the proximal landmark selection with GS1 being selected as the proximal landmark pair because of \( d_1 < d_2 \). Then using the deformation vector field produced from the DIR, the approximate magnitude of deformation for the test and closest GS landmark is determined. If the difference in magnitude of deformation \( |D_{DIRT} - D_{DIR1}| \) from Figure 3-2) is greater than the smallest voxel dimension the test landmark pair is excluded. This final criterion is required to ensure that the proximal landmark initial positions are similar enough for direct comparison. The result is a set of test landmark pairs with the corresponding proximal GS landmark pairs. The landmark displacements in both sets were compared using an ICC.
Figure 3-2: Proximal landmark selection with landmarks in the 0% phase (circles) and 50% phase (squares) from the test (white) and gold standard (black) landmark sets including: landmark displacement $D$, DIR predicted landmark displacement $D_{DIR}$, distance between 0% positions $d$ and DIR error $\varepsilon$ for a test landmark $T$, and two GS landmarks 1 and 2.

3.2.5 Deformable image registration evaluation

The DIR error was calculated by the algorithm from MIM software (version 6.5 MIMVista) for each landmark as the Euclidian distance between the DIR predicted and actual positions of the landmark in the 50% phase shown as $\varepsilon$ in Figure 3-2. The evaluation of DIR error was done for the manual, SIFT, SIFT-M and GS landmark sets using mean DIR error and proximal DIR error metrics. These two metrics were defined by the same method as the mean and proximal displacements but using the DIR error for each landmark pair ($\varepsilon$), instead of the landmark displacement ($D$).

3.3 Results

Both observers 1 and 2 were able to select 300 landmarks in each CT study pair (0% and 50% phases) taking 1100 and 1700 minutes, respectively, resulting in an average generation time per landmark pair of 22 and 34 seconds. The SIFT algorithm produced on average 128 landmarks per study and the SIFT-M method produced on average 120 landmarks per study. To normalize for different lung volumes, the number of landmarks per study was divided by the total lung volume determined from manual lung contours, resulting in landmarks per litre values are shown in Figure 3-3. The calculation time for SIFT was 47 minutes for all 10 studies, yielding an average time per landmark of 2.3
seconds. The hybrid SIFT-M method excluded on average 8 landmarks per study, with most rejected landmarks located in regions of distorted image contrast caused by motion artifacts. The landmark review by the SIFT-M method took on average 3.3±1.2 minutes per study, resulting in an overall average time of 4.1 seconds per landmark.

3.3.1 Mean displacement
Comparison of the mean lung displacement for manual landmarks versus GS is shown in Figure 3-4a with an ICC value of 0.86 for both observers. The SIFT algorithm’s mean lung displacement is compared to the GS method in Figure 3-4b, with an ICC value of 0.85, and the results for the SIFT-M method were similar (shown in Figure 3-4b) with an ICC of 0.82.
3.3.2 Proximal deformation

The proximal deformation comparison between manual landmarks and GS gave an ICC of 0.97 and 0.98 for observer 1 and 2 respectively. When compared to the GS, the SIFT algorithm had an ICC value of 0.91 with the results from all ten studies presented in Figure 3-5a. With the removal of visually dissimilar landmarks using SIFT-M, the results became much more reliable (Figure 3-5b) with an ICC of 0.96.

The summary of landmark comparisons is presented in Table 1.
3.3.3 Deformable image registration evaluation

For the mean DIR error (per lung) the accuracy varied significantly between manual and SIFT methods with ICC values of 0.97, 0.89 and 0.65 for observer 1, observer 2 and SIFT-M, respectively. For the proximal DIR error, the ICC values were 0.94, 0.96 and 0.90 for observer 1, observer 2 and SIFT-M respectively. The results for DIR error and landmark displacement for both proximal and mean methods are shown in Table 1 along with the 95% confidence intervals, across both users, SIFT and SIFT-M techniques.

Table 3-1: ICC values and corresponding 95% confidence intervals for DIR error and landmark displacement for both mean and proximal methods.

<table>
<thead>
<tr>
<th></th>
<th>Proximal displacement</th>
<th>Mean Lung displacement</th>
<th>Proximal DIR Error</th>
<th>Mean DIR Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIFT</td>
<td>0.91 (0.90-0.92)</td>
<td>0.84 (0.67-0.94)</td>
<td>0.80 (0.77-0.82)</td>
<td>0.72 (0.43-0.88)</td>
</tr>
<tr>
<td>SIFT M</td>
<td>0.96 (0.96-0.97)</td>
<td>0.82 (0.61-0.93)</td>
<td>0.90 (0.89-0.92)</td>
<td>0.65 (0.30-0.84)</td>
</tr>
<tr>
<td>Observer 1</td>
<td>0.97 (0.97-0.97)</td>
<td>0.86 (0.69-0.94)</td>
<td>0.94 (0.94-0.95)</td>
<td>0.97 (0.93-0.99)</td>
</tr>
<tr>
<td>Observer 2</td>
<td>0.98 (0.98-0.98)</td>
<td>0.86 (0.69-0.94)</td>
<td>0.96(0.95-0.96)</td>
<td>0.89 (0.76-0.96)</td>
</tr>
</tbody>
</table>

3.4 Discussion

Producing a landmark set on medical image studies for the purpose of DIR evaluation is a laborious and tedious task that can take several hours for even the most experienced individual. Automatic landmarking tools are very attractive if they can provide a reliable set of landmarks to properly evaluate a deformation vector field produced by a commercial DIR algorithm. In this study, we have shown a possibility of a six-fold reduction in time required. With optimization using parallel computing a further reduction by 20 times compared with purely manual methods appears feasible. In the future, automatic landmark placement could make near-real-time DIR evaluation a reality, providing users with more confidence in the quality of their DIR, whenever it is used.

One of the biggest issues with automatic landmarking is producing a sufficient number of relevant landmarks throughout the image space. In this study we limited the location of landmarks within the left and right lungs, obtaining 128 landmarks per study using SIFT,
compared to the 300 landmarks per study used in the gold standard. The number of landmarks per image varies drastically between images due to the difference in image quality as shown by the low landmark count per litre in studies #5 and #6 (Figure 3-2) which exhibited the worst CT image quality.

Too few landmarks per structure can significantly alter the regional analysis of deformation and DIR error as demonstrated by the lower correlation between GS and SIFT for mean displacement and mean DIR error compared to the proximal values. This relatively low correlation for mean lung deformation values was also present when comparing GS landmark pairs to the ones selected by each observer, demonstrating that even 300 landmarks were not sufficient to effectively represent the overall range of deformation in both lungs. The distribution of landmarks can significantly affect any regional analysis, especially in structures exhibiting a large amount of deformation or points lying in high dose gradients for dose accumulation studies. Study #8 had the largest lung deformation and the largest discrepancy in mean lung deformation from the GS as shown by the largest outlier in Figure 3-4. Higher quality images allow for a more reliable identification (both automatically and manually) of a greater number of unique landmark pairs per study. For example, study #3 had superior image quality and fewer motion artifacts than study #6, resulting in twice as many landmarks located per litre of lung tissue.

Apart from investigating global deformation properties of an entire structure, another important factor in deformation analysis is the accuracy of each individual landmark pair in describing the local deformation. Landmarks created by different methods or observers very rarely lie on the same point, making it practically impossible to directly compare two landmark pairs and their impact on evaluating a specific DIR algorithm. Currently, landmark evaluation is done with manual inspection of each landmark, to determine the mismatch [7, 4]. In this study, we proposed a new method of comparing landmarks pairs that are in close proximity. Using this method to compare the landmark displacement between SIFT-M and the GS yielded an ICC greater than 0.95. Even with the added uncertainty of evaluating only proximal landmark pairs between each method and the GS, the reliability of the new landmark sets was maintained. Comparison of the two sets of
measurements was performed using an ICC because it evaluates the relative agreement between two different landmarking methods. For this reason, these ICC values cannot be directly compared to other landmarking methods produced from other image sets, because they could have a different range of landmark displacements.

Aside from the obvious efficiency benefits of automatic landmarking software, there are often hidden pitfalls from overlooked erroneous landmark pairs. A SIFT algorithm matches uniquely contrasted regions (i.e. neighbourhood pixel intensities) in image pairs, but it has difficulty in non-unique regions allowing occasional significant feature mismatch. In spite of attempts to improve the landmarks by employing several program iterations and setting different parameters, between 5%-7% erroneous landmarks were still produced per image. To discard any incorrect landmarks, a simple manual review and removal process was implemented, improving the ICC value of proximal displacement from 0.91 to 0.96, while only doubling the time per landmark generation. The manual review process was performed using 2D images in the axial plane which created a bias in the landmark selection due to anisotropy but avoided the timely task of manually inspecting multiple planes. Even with the added landmark review process, the SIFT-M method was still six times faster than a purely manual approach.

The proximal landmark evaluation has demonstrated that the SIFT method performed slightly worse than the manual landmark method when compared to the GS baseline set. Through the quick removal of obvious erroneous landmarks using SIFT-M, the landmark accuracy was indistinguishable from the performance of manual users. These results demonstrate that semi-automatic SIFT landmarks can be used as a replacement for common manual landmarks in the evaluation of DIR. One important caveat is that any landmarking algorithm can create incorrect landmark pairs, but using a simple “sanity check” and ergonomic display software, these landmarks can be effectively removed through visual inspection. In the future, SIFT algorithms could be developed to incorporate further consistency checks on these landmarks to further reduce user involvement and correction.
3.5 Conclusion

Manually selected image landmarks have long been a standard for guiding image registration but require a large time investment. We provide a practical evaluation method of automatic landmarks as a replacement for manual landmarks with a six-fold time reduction. The SIFT-M method was significantly faster and required only minimal post-hoc user intervention to achieve comparable accuracy. Therefore, automatic landmarking methods can provide a standardized method for guiding and evaluating DIR algorithms without lengthy landmarking sessions. The remaining challenge for both automatic and manual landmarking methods is the selection of the optimal/sufficient number or spatial distribution of landmarks to accurately characterize the deformation of a large heterogeneous structure such as the lung.

3.6 References

1 S. Wognum et al, "Validation of deformable image registration algorithms on CT images of ex vivo porcine bladders with fiducial markers," Medical Physics, vol. 41, (7), pp. 071916-n/a, 2014.


Chapter 4

Representing the dosimetric impact of deformable image registration errors

This chapter has been previously published as “Representing the Dosimetric Impact of Deformable Image Registration Errors” published in the journal *Physics in Medicine and Biology* 62(17) N391 by Jason Vickress, Jerry Battista, Rob Barnett and Slav Yartsev. Permission to reproduce this article was provided by the publisher and presented in Appendix A.4.

4.1 Introduction

Deformable image registration (DIR) is becoming a common tool for all image-guided procedures including radiation therapy, making it possible to track changes in a patient’s anatomy using longitudinal image studies. Cumulative dose calculation across multiple fractions has been introduced for a number of treatment sites including cervical [1-3] and prostate [4] cancer. The result of a DIR is a deformation vector field (DVF) which maps every point between pairs of images and may be used to accumulate dose in a reference image. The veracity of DIR dose accumulation depends on the accuracy of the DVF, translating registration errors to potentially significant errors in dose.

Improvements in radiation delivery techniques allow for highly conformal treatments with steep dose gradients. While high dose gradients are attractive to improve patient outcomes, they may amplify DIR errors into large dose errors. Many vendors of treatment planning software have produced their own unique algorithms each functioning with different similarity metrics and regularizers. A range of different DIR results was demonstrated using 12 algorithms on a deformable pelvic phantom [5] and two algorithms using a head and neck phantom [6]. Other studies have quantified the DIR error of various commercial products [7, 8].
One way of visualizing the impact of DIR error is via the addition of error bounds to a planned dose volume histogram (DVH) curve, as proposed by Risholm et al. [9] using a probabilistic registration to account for the DIR uncertainty. From the literature, the common evaluation of DIR error for commercial algorithms was done by reporting an expected global DIR error [10-12]. However, clinically it is more important to evaluate the local dose uncertainty from the DIR error at each point in the image.

Currently, landmarks serve to match point pairs and constitute the most effective method of local DIR evaluation, measuring the registration error at individual sample voxels within an image volume. Such landmarks are useful for determining the dosimetric impact at single voxels, but cannot be used for evaluating the dosimetric impact throughout an entire image volume since they are limited to a discrete set of voxels. To fill this role several methods have been developed to predict DIR error in every point of an image volume, allowing the evaluation of the specific DVF used for dose accumulation [13-15].

In this study, a set of well-defined landmarked image pairs were used to characterize DIR error at multiple voxels, with the aim of exploring its dosimetric impact. With the dosimetric impact determined at each landmark as the gold standard, the effectiveness of different measures of DIR error can be evaluated and if found promising propagated to the entire volume. To the best of our knowledge, this is the first study to investigate the impact of DIR uncertainty on dosimetric analysis using landmarks as reference data.

4.2 Methods

4.2.1 4DCT studies and deformable image registration

Ten thoracic 4DCT images with standardized landmarks were obtained from “dir-lab” (www.dir-lab.com) [16] and 300 landmarks were provided for each image study pair matching the end-inspiration and end-expiration phases, matching every point in the end-inspiration study with a corresponding point in the end-expiration study. Landmarks were located across both the left and right lung with no significant bias and were not limited to only the clinically important regions within the lung. All images were reconstructed axially with an in-plane resolution between 0.97 mm and 1.16 mm with a matrix size of
256 × 256 (5 studies) and 512 × 512 (5 studies) and common slice thickness of 2.5 mm. Actual DIR error was calculated for each landmark pair by calculating the Euclidian distance between the DIR predicted landmark position in the end-expiration phase and the actual landmark position. Mean DIR error per study was calculated as the average landmark error of all 300 landmarks per study. In this work, all DIR procedures were performed using the standard DIR algorithm within the MIM Maestro software package (version 6.5 MIM, MIM Software Inc, Cleveland, OH, USA) with no additional refining of the DVF. Dose distributions were planned on the inspiration phase of each 4DCT study with standard five-field IMRT plans of a 60 Gy prescription dose to 95% of the planning target volume (PTV) with maximum doses ranging 69-73 Gy. Each study had a single target (no nodes) with gross tumour volumes (no ITV) ranging 1 – 43 cm³, with 5 located in the upper left lobe, 3 in lower left lobe and 2 in the lower right lobe. All dose calculations were performed on both inspiration and expiration phases using the plan from the inspiration phase using Pinnacle treatment planning system software version 9.10 (Philips Healthcare, Fitchburg, WI, USA).

4.2.2 DIR error prediction

We investigated three DIR error prediction methods: inverse consistency error (ICE), transitivity error (TE) and the distance discordance metric (DDM). Each method produced a value (pm) in millimetres proportional to the DIR error (DE) using the actual landmark displacement error as the gold standard. \( DE \) was calculated by Eq. (1), where \( k \) is the slope and \( C \) is a constant (in millimetres) found by the linear regression of the data from each prediction method vs actual DIR error for all 3000 landmarks in the collection of ten 4DCT image studies.

\[
DE = pm \times k + C
\]  

(1)

4.2.2.1 Inverse consistency error

The ICE is calculated as the disagreement between the forward and backward DVFs between the end-inspiration (0%) and end-expiration (50%) phase studies. Calculation of ICE was done with equation 2 and 3, with the coordinates in the 0% as \((x,y,z)\) and the forward and backward DVFs as \(DVF_{0.50}\) and \(DVF_{50.0}\) respectively.
(x_l, y_l, z_l) = (x, y, z) + DVF_{0-50}(x, y, z) + DVF_{50-0}((x, y, z) + DVF_{0-50}(x, y, z))  \tag{2}

ICE = \sqrt{(x_l - x)^2 + (y_l - y)^2 + (z_l - z)^2}  \tag{3}

4.2.2.2 Transitivity error

The TE is calculated in a similar manner to the ICE but includes the disagreement between three DVF’s across three different image studies (the 0%, 30% and 50% image studies). Calculation of TE was done with equation 4 and 5 with DVF_{0-30} and DVF_{30-50} denoting the DVF’s between the 0% and 30%, and 30% and 50% respectively.

(x', y', z') = (x, y, z) + DVF_{0-30}(x, y, z) + DVF_{30-50}((x, y, z) + DVF_{0-30}(x, y, z)) \tag{4}

(x_T, y_T, z_T) = (x', y', z') + DVF_{50-0}(x', y', z')  \tag{5}

TE = \sqrt{(x_T - x)^2 + (y_T - y)^2 + (z_T - z)^2}  \tag{6}

4.2.2.3 Distance discordance metric

The DDM is a simplified version from Saleh et al. [14] using ten phases (0, 10, 20, 30, 40, 50, 60, 70, 80 and 90%) of the 4DCT images to evaluate the DVF between the 0% and 50% phase. Calculation is performed using equations 7 and 8 shown below where P_i is one of the 8 additional phases (10, 20, 30, 40, 60, 70, 80, 90%) and (x_{0-50}, y_{0-50}, z_{0-50}) is the position in the 50% from the 0% phase using the DVF_{0-50}.

(x_i, y_i, z_i) = (x, y, z) + DVF_{0-P_i}(x, y, z) + DVF_{P_i-50}((x, y, z) + DVF_{0-P_i}(x, y, z)) \tag{7}

DDM = \frac{1}{8} \sum_{i=0}^{8} \sqrt{(x_i - x_{0-50})^2 + (y_i - y_{0-50})^2 + (z_i - z_{0-50})^2}  \tag{8}

4.2.3 The range of dose uncertainty

The range of dose uncertainty (RDU) was calculated for each individual landmark in the end-inspiration phase as a function of the DIR error and the dose distribution calculated on the end-expiration phase study. A flowchart describing the process of calculating the RDU is shown in Figure 4-1, with a corresponding illustration in Figure 4-2. The first step used the DVF to find the DIR-predicted location of the landmark in the end-expiration phase of the image by adding the corresponding deformation vector to the initial landmark position in the end-inspiration phase. A sphere of surrounding voxels was
created around the predicted landmark position in the end-expiration phase, with the sphere’s radius equal to one of the measures of DIR error for that landmark (2D illustration in Figure 4-2). Multiple measures of DIR error included (i) the actual DIR error, (ii) the mean DIR error per study, (iii) ICE, (iv) TE, (v) DDM prediction methods, (vi) 2 mm, and (vii) 5 mm constant values (representing low and high errors from the average error of 3.5 ± 5.5mm). A spherical sampling volume was created from an 11×11×11 voxel cube with the exclusion of voxels with \( r > (1.1 \times \text{side-length}/2) \). To accommodate a sphere with a radius equal to the measure of DIR error the cube’s side length was set at twice this measure. The result is an approximate sampling sphere of 739 voxels with an approximate radius equal to the measure of DIR error. 11×11×11 dimension was selected as a reasonable compromise between a representative number of voxels and limitations of memory consumption and computation time. A histogram of dose values throughout the sphere shown in Figure 4-2c was calculated using linear interpolation from the dose grid values. The RDU in this sphere was defined by the maximum and minimum dose values within the dose voxel histogram. The RDU was designed to evaluate the possible doses given the limits of known registration error. The RDU calculated using the actual DIR error as the radius of the sampling sphere characterizes the true dose uncertainty which is used to compare with other estimation.

![Flowchart](image)

**Figure 4-1:** Flowchart outlining the calculation of the range of dose uncertainty for a single landmark starting in the end-inspiration image.
4.2.4 Dose uncertainty evaluation

Evaluation of the calculated dose uncertainty was performed using two methods: the magnitude of dose uncertainty (MDU) and the inclusion rate (IR). The MDU is simply the size of the RDU in units of Gy. The MDU was calculated for each individual landmark in the end-inspiration phase and averaged over all 300 landmarks across all ten 4DCT studies. Another important characteristic is determining if the actual dose value lies within the RDU. For an individual study (containing 300 landmarks) the IR is the percent of landmarks where the actual dose value lies within the RDU “error bars”. The IR was calculated for each individual study and averaged over all 10 studies. The sensitivity of the proposed methods was tested using a total of 3000 unique landmarks experiencing various ranges of deformation error within different local dose distributions. Both the MDU and IR were calculated for RDUs using each measure of DIR error (actual DIR error, mean actual DIR error per study, 2 and 5 mm, ICE, TE and DDM). Comparisons between different methods MDU and IR values were performed using a two-tailed paired T-tests for a significance level below 0.05 for IR since it was calculated for each study and averaged, and below 0.01 for MDU because it was averaged across all 3000 landmarks.

Figure 4-2: RDU process for a single landmark in end-inspiration phase. The DIR predicted position of the landmark is located in end-expiration image and dose is sampled in a sphere around it. RDU is the maximum and minimum doses within the sampling sphere a) end inspiration b) end respiration c) dose voxel histogram for points within the sampling sphere.
4.2.5 Dose uncertainty distribution

The uncertainty of a deformed dose distribution was produced by calculating the RDU at each voxel using a given initial dose distribution and DVF derived from MIM Maestro. Unlike calculating the RDU at specific landmarks, calculating it at every voxel within a volume requires an estimation of registration uncertainty for each voxel performed with DDM. The RDU at each voxel is defined by two values: the upper and lower bounds of the deformed dose. The deformed dose distribution is where every voxel in the end-inspiration study is mapped to a dose value from the dose distribution calculated on the end-expiration study. Calculating the RDU at every voxel in the end-inspiration study creates two dose distributions for the upper and lower bounds of the deformed dose distribution representing all possible levels of dosimetric impact. Visualization of the dose uncertainty distribution was done by displaying the two dose distributions for the upper and lower bounds of the deformed dose. Cumulative DVHs for the PTV were computed for the planned dose and deformed dose containing the upper and lower bounds.

4.3 Results

4.3.1 DIR error prediction

Linear regression was performed for each DIR prediction method (ICE, TE and DDM) vs the actual DIR error. Linear regression factors for the DDM method were \( k = 2.05 \) (2.02-2.08) and \( C = 1.10 \) (1.01-1.18) mm with a Pearson correlation of \( R^2 = 0.71 \) (p < 0.001), for the ICE method \( k = 0.79 \) (0.76-0.82), \( C = 1.86 \) (1.73 – 1.99) mm and \( R^2 = 0.34 \) (p < 0.001), and for the TE method \( k = 0.81 \) (0.77-0.84), \( C = 1.76 \) (1.64-1.89) mm and \( R^2 = 0.37 \) (p < 0.001).

4.3.2 The range of dose uncertainty

The RDU was calculated for 300 landmarks from the end-inspiration phase on each of the ten 4DCT images. The results for 31 landmarks receiving the largest radiation dose (from a single study) are shown in Figure 4-3 using a constant error of 2 mm and 5 mm, the DDM, and the actual DIR error. For a constant error of 2 mm (Figure 4-3a), the RDU
values encompass the actual dose value 71% of the time (i.e. the inclusion rate for this subset of 31 landmarks) because the ranges are not large enough. A 5 mm constant error (Figure 4-3b) had ranges that were too large because it more frequently overestimated the dosimetric error than using the constant error of 2mm. For the DDM (Figure 4-3c) the RDU is more accurate than using 2 and 5 mm errors but not as accurate as using the actual DIR error (Figure 4-3d).

Figure 4-3: The range of dose uncertainty (RDU) for 31 landmarks with the highest doses from a representative study calculated using four different measures of DIR error a) 2mm constant b) 5 mm constant c) DDM d) Actual error. Actual dose (red X) and RDU is shown as a black error bar. Landmarks are ordered in increasing predicted dose difference from the actual dose. Dose values are displayed relative to the DIR deformed dose value for each landmark.

4.3.3 Dose uncertainty evaluation

The RDU was evaluated for 3000 landmarks in 10 studies and the average results for seven prediction methods are shown in Figures 4-4 and 4-5 for the MDU and IR, respectively. Using the actual DIR error produced the highest average IR of 97%, while its average MDU value of 2.5 Gy was similar to that of the DDM, ICE and TE methods.
which were 2.49, 2.56, 2.56 Gy, respectively, and the difference was not statistically significant (p > 0.01). The IR of DDM, ICE and TE were all significantly lower than using the actual DIR error (p < 0.05) but similar to each other (p > 0.05) with values 86, 88, 86%, respectively. The 2 mm constant error produced the smallest IR of 75% because it underestimates the DIR error. The 5 mm constant error has the highest average MDU of 3.5 Gy because it overestimates the DIR error the most. MDU’s from DDM, ICE, TE and mean error was significantly lower than using a constant 5 mm (p < 0.01). IR using constant 2 mm was significantly lower (p < 0.05) than those obtained by DDM, ICE and TE methods. IR for the mean error per study was significantly lower (p < 0.05) than those resulting from DDM and ICE methods.

Figure 4-4: Magnitude of dose uncertainty (MDU) averaged across all 3000 landmarks within 10 4DCT image studies, using the actual DIR error, 2 and 5 mm constant error, distance discordance metric (DDM), inverse consistency error (ICE), transitivity error (TE), and the mean actual DIR error per image. Error bars represent the standard deviation across 3000 landmarks.
Figure 4-5: Inclusion rate (IR) averaged across all 10 4DCT image studies, using the actual DIR error, 2 and 5 mm constant error, distance discordance metric (DDM), inverse consistency error (ICE), transitivity error (TE), and the mean actual DIR error per image. Error bars represent standard deviation across 10 studies.

4.3.4 Dose uncertainty distribution
The dose uncertainty distributions are illustrated in Figure 4-6 for a representative study using the DDM as the measure of DIR error. Both the upper and lower bounds of the RDU have worse coverage than the planned dose distribution. The cumulative DVH curves of the PTV for the deformed dose with the RDU and the planned dose are shown in Figure 4-7 showing the degradation in target coverage.
Figure 4-6: Distributions of the (a) planned, (b) deformed dose, (c) lower bound, (d) upper bound of the range of dose uncertainty calculated using DDM. Contour labels: (light blue) PTV and (dark blue) is gross target volume. Dose prescription was 60 Gy to 95% of the PTV.

Figure 4-7: Cumulative DVH curves for the planned and deformed dose with the upper and lower bounds illustrating the range of dose uncertainty (RDU) calculated using the DDM. Dose prescription was 60 Gy to 95% of the PTV.
4.4 Discussion

DIR accuracy can have a large potential impact on evaluations of radiotherapy treatments because the dose cumulates across multiple fractions introducing significant anatomical changes. Unlike other applications of DIR, radiotherapy dose accumulation is performed at the individual voxel level and errors in DIR can lead to untreated cancer or over-irradiated sensitive tissue. Earlier studies of DIR in radiotherapy investigated contour propagation and dose accumulation. For contour propagation the global accuracy of the DIR was evaluated, focusing on the ability to define contour edges typically measured with a DICE coefficient. Kumarosiri et al. evaluated multiple DIR algorithms for propagating contours for Head and Neck cancer using the DICE coefficient [17]. While contour propagation is a very useful tool for IGRT, it fails to evaluate a DIR throughout the interior (and exterior) of the contoured region. Studies presenting dose accumulation applications are useful in demonstrating the potential of DIR in radiotherapy but fail at providing the ground truth and range of uncertainty. Abe et al. showed the difference in DVH parameters using DIR compared to simple DVH parameter addition [18] to demonstrate benefits of using deformed dose accumulation.

Several studies have compared the results of dose accumulation from different DIR algorithms [19, 20] revealing that the results are dependent on the specific DIR algorithm. Typically the investigation of DIR performance and uncertainty is undermined by the lack of ground truth. In many publications (including an earlier work by Risholm et al. [21]) colour maps and graphs are used to demonstrate the potential DIR uncertainty but do not consider the actual DIR error. Li et al. used landmarks for the direct correlation between a registration evaluation method and the actual DIR error but did not consider the dose implications [22]. Following their approach, we used landmarks to obtain the true DIR error and impact on dosimetric analysis, across multiple sets of image studies. Though this does limit the analysis to specific regions of the image, it could also be used with automatic landmarking software (23, 24) to provide real-time evaluation of DIR uncertainty. We previously showed the efficacy of using the scale-invariant feature transform for landmark generation in DIR evaluation and demonstrated that such landmarks can be a good surrogate for manually selected landmarks [25].
Error bounds are typically used to demonstrate the level of uncertainty for any measured or computed value with known limitations or from a sample with a defined statistical distribution. In this study, we investigated error bars representing the RDU at individual landmarks using different DIR error prediction methods. The RDU error bars are a function of both the DIR uncertainty and the local gradient of the dose distribution. For example, a landmark with large DIR error in a homogenous region of the dose distribution will have a very low MDU. In Figure 4-2b, the RDU varies greatly even when using a constant DIR error of 5 mm, because the dose gradient varies greatly throughout the image volume. The large range in RDU values is the result of using the maximum and minimum values from each sampling sphere. The extreme values were selected instead of other descriptors (i.e. 5th and 95th percentile) because it yielded the highest IR. The MDU and IR metrics were introduced to represent the two most important properties of an error bar: its size and if a “true” value lies within its bounds. The ideal RDU would be as small as possible minimizing MDU, while always containing the true value with an IR = 100%. The RDU calculated from the actual DIR error was considered the gold standard representing the true dose uncertainty, with the highest IR while maintaining a minimum MDU. The IR while using the actual DIR error was below 100% because of the limitations in the interpolation of dose values near the boundary of the sampling sphere. We evaluated different DIR prediction methods on their capability to represent the RDU.

Testing the impact of DIR error for a number of test voxels is limited to regions where the landmarks are located and those locations may not be in clinically relevant regions. Determining the impact of DIR error in all parts of the image would require a method to quantify or predict DIR error for every point of the image volume. Several studies created different tools to predict the magnitude of DIR error given specific properties of the image volume and its deformation [13-15]. In our study, we selected the DDM for DIR error prediction because it could be tested with 4DCT image studies, considering each of the 10 phases as separate studies. The other voxel-wise predictors of DIR error (ICE and TE) performed similarly to DDM with differences in IR of less than 1% and differences in MDU within 0.1 Gy. The small differences in RDU performance were remarkable given that DDM had the best correlation with actual DIR error having a Pearson
correlation coefficient of 0.71 compared to 0.34 for ICE and 0.37 for TE. This remarkable result is caused by the MDU’s and IR’s insensitivity to small differences in registration errors produced by different measures because they present the average values across 3000 landmarks with different local dose distributions, whereas the correlation compares individual landmark errors. The RDU created using the DDM was comparable to the RDU created using the actual DIR error defined by landmarks. The results using DDM were also significantly better than the RDU derived from constant error values of 2 and 5 mm having a significantly higher IR and lower MDU respectively. Even the mean DIR error for each study performed worse than the DDM demonstrating that the heterogeneous distribution of DIR error throughout a lung volume makes global measures of DIR error inadequate and should be avoided for this application.

Presently commercial DIR software is being used to calculate the cumulative dose distributions without considering the impact of registration uncertainty on the risks associated with poorer tumour control, increased toxicities, and long-term cancer. Currently in radiation therapy many uncertainties in the treatment planning, delivery and alignment are accounted for when developing a patient’s treatment plan, but DIR uncertainty is normally neither included nor propagated across dose fractions. Ideally, the dose uncertainty should be defined as the dose between the DIR predicted and reality, but without landmarks located throughout the image space, this is not feasible. Utilizing different measures of DIR error such as DDM allowed the calculation of a range of potential dose error throughout the whole volume. By taking the maximum and minimum dose values within the sphere, the method quantifies the most conservative situation when the registration error is always in the direction of the largest dose difference. Such information is valuable for the radiation oncologist by indicating possible areas of concern. In radiation therapy practice the knowledge of the RDU provides information on the span of possible dose values, indicates the reliability of the applied DIR algorithm and can prompt further investigation, including contouring or further imaging. The RDU evaluation proposed in this work can be included to account for DIR uncertainty limits when applied to dose accumulations, and allow DIR based dose accumulation to become standard in radiotherapy clinics. Our future work plans to apply the RDU approach to other cancer sites using clinical treatment plans to demonstrate how it can impact real
treatment decisions. The goal is to introduce the RDU evaluation so that it becomes commonplace whenever DIR is used to assess the impact on a cumulative radiation dose distribution.

4.5 Conclusion

DIR error is inevitable for any application of repeated image guidance that requires highly accurate DVF solutions. In our study, we determined the impact of DIR error on the dosimetric analysis and found that surrogates of DIR uncertainty can provide error bars for the cumulative dose distribution. Using the landmarks actual displacement error as the gold standard we evaluated the dose-predictive power of various measures of DIR error. The RDU evaluation was developed to represent the range in the deformed dose distribution. Calculation of the RDU required a measure of DIR error and it was shown that voxel-wise predictions of DIR error (such as DDM) performed best, second only to using the actual DIR error. It was also confirmed that fixed global measures of DIR errors are not adequate for determining the RDU when compared to voxel-wise methods. The RDU calculated using the DDM provided the upper and lower dose limits throughout an entire deformed dose distribution, not being limited to specific landmarks. Ultimately the RDU can provide a useful representation of the dosimetric impact of registration uncertainty for any commercial DIR algorithm, providing clinicians with a more realistic analysis of how accurately a radiation treatment is converging to the prescribed treatment.

4.6 References


Chapter 5

5 Online daily assessment of dose change in head and neck radiotherapy without dose-recalculation

This chapter has been accepted for publication in the *Journal of Applied Clinical Medical Physics* by Jason Vickress, Jerry Battista, Rob Barnett and Slav Yartsev. The article will be open access.

5.1 Introduction

Radiation therapy is a standard treatment option for a variety of cancers where the precise geometric targeting of tumours can be exploited for achieving better tumour control while limiting healthy tissue damage. The specific targeting and attenuation of radiation are unique to the patient’s anatomy at the time of the planning-CT simulation (PCT), but these conditions are difficult to maintain throughout an entire course of treatment due to changes in anatomy [1-4]. To account for changes in patient anatomy, plan modification may be required during the treatment course to ensure accurate targeting. Plan adaptation has been shown to improve treatment outcomes by promoting better tumour control and limiting toxicities [5, 6], but this procedure entails additional costs of re-imaging, re-planning, and additional quality assurance. Though the potential benefits of plan adaptation are obvious, no guidelines on decision-making and optimal time for re-planning are available.

Plan adaptation has been reported for various treatment sites including lung [7], prostate [8-10], and head and neck cancers [11, 12]. Across all treatment sites, adaptation is necessary due to tumour shrinkage, weight loss or other significant anatomical changes that impact the dose distribution (e.g. lung collapse or re-inflation). Specifically for head and neck cancers, large volume changes are common and often detected by external examination or through poor fitting of immobilization devices, but minor changes can go unnoticed. However, relatively minor anatomy changes may still have a significant effect
on the dose distribution and are more difficult to discern by visual inspection of anatomy alone.

More precise and conformal radiation treatments available with modern techniques may need more plan adaptations to provide consistent target dose coverage and healthy tissue sparing with a changing anatomy. For making a decision on the necessity of plan adaptation in clinical practice, efficient daily evaluation of the delivered dose distribution on the modified anatomy is required. Different methods have been presented on detecting volume changes [13] and landmark movements [14] but most rely solely on visual inspection by clinicians. These visual inspections may not be consistent as shown by inter-observer studies [15]. Several groups have presented adaptation strategies and schedules throughout treatment [16, 17]. A recent study using the same dataset as in this study has produced a method of detecting anatomical differences to flag consideration of plan re-evaluation without considering the dose distribution [18].

Currently, cone beam CT (CBCT) imaging is routinely used for patient alignment and anatomy monitoring but can also be used for dosimetric assessment of actual radiation delivery. Dose calculations on CBCTs are possible with the results varying between reported studies [19–21] because of inferior image quality and tissue densitometry. Performing reliable analysis of the dose to the target and organs at risk would require contouring of relevant structures on the daily CBCT image. An attractive alternative is to employ deformable image registration (DIR) to transfer contour information from the planning CT study for analysis. DIR has been shown to produce a variety of results depending on the algorithm used, original contouring accuracy and imaging modalities (i.e. CT simulation, MRI or CBCT). Unfortunately, registration between different imaging modalities has been shown to have worse accuracy [22] especially for CBCT images due to limited image quality and artifacts.

There are two primary effects of anatomical deformations on a radiation treatment: 1) movement of voxels and regions of interest (ROI) relative to the planned dose distribution and 2) change of the dose distribution itself due to re-arrangement of voxels or density changes therein. The current gold standard (GS) for determining whether to adapt a treatment plan involves a new CT simulation (ReCT), dose calculation and DIR to
map contours from the PCT. This procedure is time-consuming and expensive but accounts for both effects of anatomical deformation and is applied when gross anatomical changes are suspected.

The best alternative without a new CT simulation involves using DIR to warp the planning CT to match the daily anatomy from the CBCT and perform dose calculation as proposed by Veiga et al [23] and accounts for both effects of anatomical deformations. However, the dose recalculation practically can be difficult and time-consuming. It is usually performed off-line which limits its routine daily use at the treatment unit. What if you could determine the necessity of plan adaptation without a new CT scan and dose calculation? Without the re-computation of the dose, only the movement of voxels and ROI’s relative to the planned dose distribution are considered, but not the change to the dose distribution. The dose distribution is assumed to be robust and only mildly affected by the re-arrangement of the voxels. In this study, we explore the results of using the CBCT without a dose calculation and a CBCT with a dose calculation and compare both to the current gold standard. The goal is to see if assessing the movement of ROI relative to the planned dose distribution provides enough dose information to properly trigger the plan adaptation process, when compared to current clinical practice of visual inspection.

5.2 Methods

5.2.1 Patient studies

For this study, 18 patients who received multi-fractionated radiotherapy for head-and-neck cancer and had plan adaptation during treatment course were selected. Each patient had a CT scan taken before treatment (range 4-30 days) and used for planning (i.e. PCT), daily pre-treatment CBCT studies and another CT re-taken during treatment (ReCT) when anatomy changes were deemed significant (day “X”). Significant changes included sensitive structures moving into high dose regions, tumour moving out of this region or excessive weight loss by the patient. Both PCT and ReCT studies were obtained on a 120 keV Phillips Big Bore CT scanner (Philips Healthcare, Fitchburg, WI, USA) with a 512 x 512 image size, 0.9-1.2 mm resolution, and 3 mm slice thickness. CBCT scans were performed every 1 to 5 fractions with the onboard imaging available on Varian iX and True Beam treatment units (Varian Medical Systems, Palo Alto, CA, USA)
using 100 keV with a 512 x 512 or 384 x 384 image size, 0.5 – 0.65 mm resolution and 2.5 – 2 mm slice thickness. Treatment plans had prescribed doses ranging 50-70 Gy to the planning target volume (PTV) in 30 to 35 fractions using volumetric arc therapy (VMAT) with two 360° arcs and included 1 or 2 target volumes. Specifically, 15 patients had only one target, 3 had two targets and all patients had a larger nodal volume overlapping all targets prescribed to a lower dose. Treatment planning and dose calculations were performed on a Pinnacle treatment planning system (version 9.10, Philips Healthcare, Fitchburg, WI, USA) using Pinnacle’s collapsed cone convolution superposition algorithm [24] using a dose grid of 3×3×3 mm. All image registrations (both rigid and deformable) were performed with software from MIM Maestro (version 6.5 MIM Software Inc., Cleveland, OH, USA) using the default DIR algorithm applying an intensity based free form algorithm, with a sum of squared differences similarity metric [25]. The mean registration error using MIM Maestro between two kVCT’s was shown to be 1.7 mm by Kirby et al. [22] using a deformable Head and Neck phantom.

5.2.2 Dose distribution estimation
To determine the necessity of plan adaptation, an estimation of the dose distribution “of the day” was required and three estimation methods are presented and compared to the current gold standard which requires a re-planning CT. The first method (CBCT_P) used DIR to map the contours from the PCT to the daily CBCT with the planned dose distribution rigidly registered to the daily CBCT as shown in figure 1.

Figure 5-1: Illustration of the CBCT_P method for the evaluation of the need for plan adaptation using the DIR of planning CT to daily CBCT study and the planned dose distribution. DIR – deformable image registration, RR – rigid registration.
The second method (CBCT<sub>R</sub>) used the DIR to map the contours from the PCT to the daily CBCT with the recalculated dose (from the ReCT) rigidly registered to the daily CBCT. The third method (ReCT<sub>P</sub>) used the DIR to map the contours from planning CT to the ReCT with the planned dose distribution rigidly registered to the ReCT. The gold standard method (ReCT<sub>R</sub>) applied DIR to map contours from the PCT to ReCT with the recalculated dose on the ReCT. Both dose distributions (planned and recalculated) were obtained using the original treatment plan parameters and beam; the plan was not re-optimized. The rigid registration process used 6 degrees of freedom and simulated the alignment of the CBCT study to PCT (or ReCT) performed by the radiation therapists in the clinic before each fraction. In total, four separate methods estimated the daily dose distribution using the CBCT or ReCT as the secondary CT study, with the planned or recomputed dose. For clarity, each method was referred to by the secondary image used (CBCT or ReCT) and if the planned or ReCT dose was used, denoted by subscript P or R, respectively. All dose estimation methods are illustrated in figure 2, showing all four investigated combinations.

Figure 5-2: Schema describing the daily dose estimation using DIR from the planning CT to either the daily CBCT or re CT study (ReCT). Two different dose distributions computed on the PCT or ReCT are transferred to the moving image using a 6 degree of freedom rigid registration. The gold standard method is highlighted in yellow using the ReCT and recomputed dose. Day X is when ReCT was ordered due to observed significant anatomical changes.
5.2.3 Voxel-to-voxel dose comparison

The clinically relevant comparison of the dose results obtained by different estimations requires evaluation on a voxel-to-voxel basis. Every voxel in the PCT study can have a different dose value in fraction X (when ReCT was ordered), depending on the secondary CT study for image registration and the dose distribution. Comparison with any other method is done by calculating the relative dose difference to the GS (RD$j$) for a specific structure $j$ across each individual voxel $i$:

$$RD_j = \frac{1}{\sum_{p}^{18}} \sum_{i}^{N_j} \frac{D_{(GS)}^{i,p} - D_{(T)}^{i,p}}{D_{(GS)}^{i,p}} \times 100\% \quad (1)$$

between a test method ($T$) and GS averaged over all $N_j$ voxels within all 18 patients $p$.

Voxel-to-voxel analysis was performed for the right and left parotids because they were present in all image studies, incurred significant deformation and are frequently positioned close to the target volume. The analysis was also performed for the spinal cord because it is a clinically important structure.

5.2.4 Test for the necessity of adaptation

In practice, the estimations of dose distribution changes would be used to determine if a current plan delivery is not within clinical dose tolerances and needs adaptation. For our dose distribution estimation methods, adaptations were considered necessary if the following dose tolerances were exceeded: mean dose to parotid equal or above to 26 Gy, max dose to spinal cord equal or above 50 Gy or dose to 95% of the PTV below the prescription dose. Using two of the methods, CBCT$P$ and CBCT$R$, dose values were calculated and compared to clinical tolerances to see which method would accurately trigger plan adaptation, when compared to the gold standard (ReCT$R$). Any parotid with a planned mean dose equal or above 26 Gy was excluded from this test since the organ was already planned to receive greater than the tolerated dose. The spinal cord is an important organ for head and neck radiotherapy planning, but was not considered for the necessity of adaptation test since the threshold for a max dose of 50 Gy was only crossed by one patient.
For the PTV the dose was determined at each voxel using CBCT\textsubscript{P} and CBCT\textsubscript{R} methods. The threshold criterion for adaptation was for 95% of the volume (D95) to be below the prescribed dose. The D95 parameters were selected because following recommendations for evaluating the target coverage [26]. Only the primary PTV was analyzed for each patient.

Evaluation of the clinical decision to adapt or not relied on the compliance with both parameters: the parotid mean dose and D95 to the PTV. To simulate a conservative treatment situation the mean parotid dose and max spinal cord dose was rounded to the nearest integer, for example, 25.6 Gy is rounded to 26 Gy. The results were reported as the number of unnecessary adaptations (adapting, when within tolerance) and missed adaptations (not adapting, when tolerances were exceeded).

5.3 Results

5.3.1 Voxel-wise dose comparison
The relative dose difference $R_{Dj}$ given by equation (1) for each method are shown in table 1 for the ipsilateral and contralateral parotids and spinal cord. The error caused by only the changed dose distribution is presented by the ReCT\textsubscript{P} row and the CBCT\textsubscript{R} row represents the error caused only by the DIR between different imaging modalities. CBCT\textsubscript{P} row represents the error when both effects were present.

\begin{table}[h]
\centering
\begin{tabular}{lccc}
\hline
Secondary image and dose distribution & Ipsilateral Parotid & Contralateral Parotid & Spinal cord \\
\hline
ReCT\textsubscript{P} & 8 % (5.7 %) & 7.9 % (5 %) & 3.8 % (1.6 %) \\
CBCT\textsubscript{P} & 12.7 % (9.5 %) & 13.5 % (7.8 %) & 5.7 % (2.4 %) \\
CBCT\textsubscript{R} & 7.5 % (4 %) & 7.7 % (4.5 %) & 4 % (2 %) \\
\hline
\end{tabular}
\caption{Relative voxel-wise dose difference from gold standard (ReCT\textsubscript{R}) ($R_{Dj}$) for ipsilateral and contralateral parotids and spinal cord, averaged over 18 patients. Standard deviation is displayed in brackets.}
\end{table}
5.3.2 Test for the necessity of adaptation

The parotid mean dose estimates using CBCT\textsubscript{P} and CBCT\textsubscript{R} are compared relative to the 26 Gy threshold to the gold standard (ReCT\textsubscript{R}) in figures 3a and 3b, respectively. The number of parotids that were incorrectly labelled as either greater than or less than 26 Gy, out of 15 tested parotids was five for CBCT\textsubscript{P} and one for CBCT\textsubscript{R}. The D95 estimates relative to the dose prescription for CBCT\textsubscript{P} and CBCT\textsubscript{R} are compared to the gold standard in figures 4a and 4b, respectively. The number of patients where the CBCT-based prediction was different from the gold standard on their PTV D95 parameter (out of 18 patients) was one for both CBCT\textsubscript{P} and CBCT\textsubscript{R}.

![Figure 5-3: Predicted mean dose using a) CBCT\textsubscript{P} and b) CBCT\textsubscript{R} methods compared to ReCT\textsubscript{R} (gold standard) for 15 parotid glands. Clinical threshold of 26 Gy is shown by solid lines. ReCT\textsubscript{R} is the DIR to ReCT using the recalculated dose. CBCT\textsubscript{P} is the DIR to daily CBCT using the planned dose. CBCT\textsubscript{R} is the DIR to daily CBCT using the recalculated dose.](image)
Figure 5-4: The difference between predicted D95 and the prescribed dose for the PTV for using a) CBCT$_P$ and b) CBCT$_R$ methods compared to ReCT$_R$ (gold standard). Values are presented as the difference from the prescribed dose. ReCT$_R$ is the DIR to ReCT using the recalculated dose. CBCT$_P$ is the DIR to daily CBCT using the planned dose. CBCT$_R$ is the DIR to daily CBCT using the recalculated dose.

To simulate a clinical decision-making situation, the results for both parotids and PTVs were combined to determine whether to adapt or not based on the dose predictions from CBCT$_P$ or CBCT$_R$ methods. Clinically, the plans for all 18 patients were adapted using a conservative approach based on anatomical changes alone but according to our dose analysis, only 7 were outside of tolerance leaving 11 potentially unnecessary adaptations. Using the CBCT$_P$ method there would have been only 4 unnecessary adaptations without missing any required adaptations. For CBCT$_R$ (with the recomputed dose) there would have been 2 unnecessary adaptations while also not missing any required adaptations.

5.4 Discussion

In a standard workflow, the only dose distribution always available is the one calculated using the initial CT simulation for planning purposes. Theoretically, the dose gradients from the planned dose distribution indicate what dose differences may occur due to specific anatomical changes. Dose gradients are mainly defined by the original beam geometry relative to the planned iso-centre, which is not affected by deformation. Without extensive deformation, these gradients can be maintained and could predict dose change, when combined with a deformation field. However, with large volume or density reductions within the beams path significant changes to the dose distribution can result,
which could lead to missed adaptations, if no dose calculation is performed. But very large volume/density changes are clearly visible by visual inspection on imaging and would be flagged and trigger adaptation by a therapist.

The average relative dose differences $RD_j$ for each organ presented in table 1 show that for both parotids and spinal cord the $RD_j$ from the CBCT$_R$ method (which is a result of DIR error alone) is similar to the results from the ReCT$_P$ method, which is the error from using the plan instead of the recomputed dose. The CBCT$_P$ $RD_j$ includes both sources of error but is less than the sum of errors in ReCT$_P$ and CBCT$_R$ methods.

It has been shown that DIR error is specific to the algorithm used [27, 28] and image quality [22]. In this study, only one commercial algorithm was used to evaluate the utility of applying DIR to CBCT studies using an unmodified commercial product. More accurate dose estimations could be performed if registration error was known and accounted for as demonstrated in our previous work [29]. Typical plan adaptation strategies revolve around re-planning on the CBCT study using DIR to propagate contours and evaluate dose [30-32]. Two publications by Veiga et al. have evaluated the process of using DIR to CBCT studies for determining daily dose [23] and accumulated dose with different DIR algorithms [28], but in both cases dose computations are needed for each fraction. In this study, DIR of the daily CBCT study is proposed to evaluate anatomical changes without a re-scan of the patient or dose calculation using a commercial DIR algorithm.

Practically speaking, DIR procedures can help physicians to decide when to adapt their radiation treatment plans. From the results presented in figures 3 and 4, the clinical decision to re-plan all 18 of these cases was not necessary, with 11 of the original patient plans still within clinical tolerances. Clinical decisions of plan adaptation were made before the re-scan using personal experience, which explains the discrepancy in adaptation rates between our GS and what was decided clinically. Our results have shown that both methods using daily CBCT studies (CBCT$_R$ and CBCT$_P$) yielded very conservative results and missed no required adaptations. If the simplest prediction method (CBCT$_P$) was used, only four patients would have been unnecessarily re-scanned and adapted. This demonstrates that using the DIR to the CBCT of the day without a dose
calculation in CBCT$_P$ method can determine when to adapt a treatment plan better than what was done clinically avoiding a number of unnecessary CT simulations and re-planning efforts. Performing an additional dose calculation in CBCT$_R$ caught two additional unnecessary plan adaptations at the cost of additional computation time, while without a dose computation the procedure can be completed within one minute allowing for an efficient “adapt or not” decision online.

5.5 Conclusion

Improvements in IGRT and conformal radiation delivery have made adaptive radiation therapy a reality, but steps need to be taken to ensure its efficiency. Practical implementation requires an efficient method of daily evaluation and decision-making to determine when plan adaptation is truly necessary. The method of dose evaluation using on-board CBCT imaging alone is limited by the necessity for dose calculation, contouring and image registration. We have shown that the daily CBCT image mapped back to the planning CT without a dose calculation can provide sufficient information for the important decision of when to re-plan. The goal is to prevent the use of unnecessary additional CT simulations and dose computations with a quick online evaluation. Further research needs to be performed with more patients and other treatment sites including abdomen and thorax and for treatment techniques that will produce a different landscape of dose gradients.

5.6 Acknowledgement

The authors thank Dr. Bryan Schaly for supplying CT data and discussion.

5.7 References


26 K. K. Ang et al, "Randomized phase III trial of concurrent accelerated radiation plus cisplatin with or without cetuximab for stage III to IV head and neck


Chapter 6

6 Conclusion and future work

6.1 Summary of findings

This thesis has demonstrated (1) the significance of using patient data to guide physician decision making in radiation therapy, (2) the importance of including registration uncertainty, when applying DIR for dose accumulation, and (3) the need for outcome-driven decision making schematically presented in Figure 1-7. In this chapter, the major findings of the four studies comprising this thesis and future work are summarized.

6.1.1 Data-driven decision making in radiation therapy

In the past, treatment decisions in radiation therapy, including dose prescriptions, have been made based on the physician’s experience including findings from clinical trials. With rapidly evolving technology and development of new treatment techniques, clinical trials cannot keep pace and are only attempted when large payoffs are expected. Unfortunately, current clinical trials cannot effectively optimize treatment prescription and delivery parameters for specific disease pathologies and patient characteristics. To fill this void and provide truly personalized medicine, real-time data-driven decision making is required, utilizing the large quantities of data available in this era of “Big data” [1].

Data-driven decision making requires large amounts of patient data to construct models predicting patient outcomes, including survival, disease control, and toxicity based on patient characteristics and available treatment options. Chapter 2 tackled this problem for patients suffering from hepatocellular carcinoma and metastatic liver cancer to predict overall survival and determine the candidacy for curative vs palliative treatment. The results of chapter 2 were two nomograms predicting overall survival for both patient groups, including the patient’s disease burden and liver function. Nomograms act as predictive models to guide a physician’s decision on how to best treat based on the patient’s pre-treatment information providing the link between patient data and treatment
decisions as shown in Figure 1-7. Chapter 2 demonstrated that using routinely collected patient data can help guide physicians with treatment decisions, and highlight new areas of research, that will improve patient care. For example, the prescribed radiation dose was not found to be a significant variable predicting overall survival with a p-value of 0.1 for the HCC model. Prescribed radiation dose only describes what was planned, not necessarily how the radiation was actually delivered. It can be speculated that using the actual delivered dose (acquired through DIR based accumulated dose) should improve the models of survival, which led to the work with DIR in chapters 3 and 4. With improved predictive models patients can receive treatments personalized for their unique biology.

Aside from using patient data for pre-treatment decision making, it can also improve mid-treatment decision making as to when to adapt a treatment plan. Currently, all mid-treatment decisions are made based on the visual inspection of alignment devices (i.e. masks) or daily imaging (i.e. CBCT or 2D kilovoltage) by physicians. The method presented in Chapter 5 is not replacing clinical decisions but providing physicians with the dosimetric changes automatically and reliably in less than one minute before every treatment fraction. Additionally, the results of the study in chapter 5 demonstrated that the decision for plan adaptation could be made without the need for manual contours or a dose calculation. Typically, accurate dose analysis would benefit from a new dose calculation on the modified anatomy, but the results from chapter 5 demonstrate that in order to predict significant changes to a dose distribution, it is not always required. The conclusion from chapter 5 demonstrated that applying data-driven decision making can provide physicians with the daily dosimetric changes and make adaptive radiation therapy more reliable and cost-effective by initiating a re-planning procedure only when necessary.

6.1.2 Including registration uncertainty of DIR for dose accumulation
Deformable image registration is becoming a common tool in radiation therapy allowing for the propagation of contours between different image sets and the accumulation of dose between sets of images. Unfortunately, DIR inherently contains variable levels of accuracy creating significant uncertainty in any results it produces. Originally only rigid registration (including translations and rotations) was used in radiation therapy for
alignment of treatment images and target delineation, and any error could be visualized and quantified as a single value. DIR is defined by thousands of 3D deformation vectors, each having its own associated error vector. When DIR is used in conjunction with another 3D quantity such as absorbed dose the effect of registration error is magnified. For DIR to be effectively used in radiation therapy, the routine evaluation of DIR error and its impact on cumulative dose is required.

In chapter 3, a new novel method for the evaluation of DIR was presented using an auto-landmarking method based on the scale invariant feature transform (SIFT). Landmarking methods find corresponding points in a pair of images and tests how well the DIR can match the same landmark points. Previously landmark pairs could only be located manually requiring significant human resources, reserving their use to algorithm benchmarking [2, 3] and research [4]. The method presented in chapter 3 created landmarks comparable to high-quality gold standard landmarks and reproduced local deformation values with an ICC value of 0.96. The primary limitation of automatic landmarks is the number and distribution which are heavily influenced by the type and quality of images used. For frequent routine evaluation, only a small number of landmarks representative of the anatomy and the dose distribution are required, compared to the large number of landmarks required for DIR algorithm benchmarking. With the method from chapter 3, the routine evaluation of DIR is possible enabling the accurate daily use of DIR in the clinic. Accurate DIR with IGRT is the best method to determine the cumulative dose distribution providing the link between what was planned and delivered as shown in Figure 1-7.

One of the primary uses of DIR in radiation therapy is the accumulation of dose throughout a multi-fraction treatment using the CBCT studies provided by IGRT. Dose accumulation uses the DVF to map the dose from individual fractions back to the original treatment plan to determine the actual delivered dose. The accuracy of a dose accumulation is determined by the accuracy of each vector within the deformation field multiplied by the local dose gradient. Previously, the only option was to apply the best available DIR algorithm, with no account of the registration error. The dose accumulation error cannot be easily included because of the distribution of registration errors
throughout an image volume. If the registration error is high in regions with a large dose gradient, there will be a large dose accumulation error and a homogeneous dose region with a low dose gradient will have a low dose accumulation error. Chapter 4 presented a method of incorporating the distribution of registration error into the dose accumulation process to provide the range of accumulated doses. The method in chapter 4 was also shown to work with DIR uncertainty values calculated from statistical, or consistency metrics. Currently, the best option is to cumulate radiation dose assuming the registration is 100% accurate, but with the method presented in chapter 4, a realistic range of cumulative dose values can be obtained. With a realistic range of delivered dose values, an accurate account of radiation treatment can be obtained and fed into an outcome-driven decision-making process as shown in Figure 1-7. With the real delivered dose stronger predictive models can be produced to guide physicians in their treatment decision for each new patient.

6.2 Future work

6.2.1 Automatic DIR evaluation of different disease sites.

In chapter 3 an auto-landmarking algorithm was applied and evaluated for lung CT studies, demonstrating its ability to quickly evaluate DIR. To increase the viability of auto-landmarking methods, tests with other treatment sites and imaging modalities should be performed. Important disease sites exhibiting large amounts of deformation during radiation therapy include head and neck, breast, and cancers within the abdomen and pelvis (i.e. prostate). This land-marking method should also be tested on different imaging modalities including CBCT and MRI, investigating the method’s performance between like (i.e. CBCT-CBCT) and different modalities (i.e. MRI-CBCT). Effective testing of a landmarking method requires manual gold standard landmark sets similar to the gold standard landmarks used in chapter 3. With gold standard landmarks for each treatment site and imaging modality, the presented algorithm can be tested and optimized with the goal of being introduced for routine DIR evaluation. It is also important to test other land-marking algorithms and compare them to those generated by SIFT. For example, other land-marking algorithms include an edge based region detector [5] or Speeded-up robust feature detection (SURF) [6]. There is also potential to improve upon
the current SIFT implementation to improve landmark yield and quality as shown in Yang et al. [7]. With the routine evaluation of DIR, it can be used freely to assess daily anatomy or determine the cumulative dose distribution and be correlated with outcomes.

6.2.2 Clinical application of DIR with uncertainty.

DIR continues to grow in popularity in radiation therapy, being included in almost every treatment planning system and medical imaging software. From the methodologies and data in chapters 3 and 4, the accuracies of DIR algorithms vary significantly and they need to be routinely evaluated, and the inaccuracies incorporated into any resulting analysis. To achieve this, the auto-landmarking methods presented in this thesis can be included in any DIR package, and the resulting error on each vector can be used to interpret the final error in the cumulative dose analysis. The final result of any DIR should provide error bars on the cumulative dose distribution to better understand what was actually delivered and how it correlates with patient outcomes.

6.2.3 Modelling treatment outcomes with more realistic delivered dose.

Traditionally, effects of radiation therapy have been modelled based on planned dose parameters, for example, the dose to 95% of the target volume. Historically, only important dose parameters from a radiation treatment plan were recorded and linked radiation effects and treatment. With newer technologies and techniques, detailed data are available describing the specific spatial distribution of radiation dose, relative to different tissue types and organs. The spatial distribution of the delivered radiation dose could then help to further explain the effects of the radiation dose, including radiation-induced toxicity and tumour control. The next step is to include daily image guidance and DIR to know the actual delivered dose to patients and provide the most detailed account of a radiation treatment.

Current forms of DIR are not accurate enough to be used for dose accumulation without accounting for the registration uncertainty. In this thesis, two methods have been presented to allow for a more accurate use of DIR through the incorporation of registration uncertainty. When DIR is applied for dose accumulation it can be evaluated using the auto-landmarking method presented in chapter 3 and determine the spatial distribution of registration error. With the known spatial distribution of registration error,
the dose accumulation error can be calculated using the method from chapter 4 on each individual treatment fraction. Information about the dose accumulation error can determine which patients have an accurate account of their delivered radiation treatment and can be correlated with outcomes. Although dose uncertainties are vital for developing tools for a physician’s decision making, clinical significance comes down to absolute cumulative dose to regions of interest for both TCP and NTCP assessment. Further studies can produce predictive models between the “real” delivered dose and patient outcomes to be implemented in outcome-driven decision making as shown in Figure 1-7. The ultimate goal is to use the record of real radiation dose to further improve predictive models and work towards personalized medicine using the already available patient data.

The goal of this thesis is to present methods for the incorporation of outcome-driven decision making in radiation therapy and facilitate personalized treatments for patients based on their own specific biology. It is hoped that this work will provide the roadmap to implement outcome-driven decision making in radiation therapy clinics and utilize the large amount of available clinical data to significantly improve patient care.

6.3 References


Appendix A: Permission to reproduce published material

A.1 Permission to reproduce content for Figure 1-3a
INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol./edition number, Author(s), Title of article / title of chapter, Pages No., Copyright Year, with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit: "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material. Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC’s Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC’s Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC’s Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

https://s100.copyright.com/AppDispatchServlet
9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.
10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.
11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.
12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).
13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated hereinafter), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.
14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:
15. Translation: This permission is granted for non-exclusive world English rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.
16. Posting licensed content on any Website: The following terms and conditions apply as follows: Licensing material from an Elsevier journal. All content posted to the web site must maintain the copyright information line on the bottom of each image. A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com; Central Storage. This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book. A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com. All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The website must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.
17. For journal authors: the following clauses are applicable in addition to the above:
Preprints:
A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting,
copyright, technical enhancement etc.). Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

**Accepted Author Manuscripts:** An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- **immediately**
  - via their non-commercial person homepage or blog
  - by updating a preprint in arXiv or RePEc with the accepted manuscript
  - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
  - directly by providing copies to their students or to research collaborators for their personal use
  - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement

- **After the embargo period**
  - via non-commercial hosting platforms such as their institutional repository
  - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

**Published journal article (JPA):** A published journal article (JPA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment. Policies for sharing publishing journal articles differ for subscription and gold open access articles:

**Subscription Articles:** If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

https://s100.copyright.com/AppDispatchServlet
Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect. Please refer to Elsevier's posting policy for further information.

18. For book authors the following clauses are applicable in addition to the above:
Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. Posting to a repository: Authors are permitted to post a summary of their chapter only in their institution's repository.

19. Thesis/Dissertation: If your license is for use in a thesis dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

Elsevier Open Access Terms and Conditions
You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our open access license policy for more information.

Terms & Conditions applicable to all Open Access articles published with Elsevier:
Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated. The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure that reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license:
CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by/4.0.

CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at http://creativecommons.org/licenses/by-nc-sa/4.0.

CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by-nc-nd/4.0.
Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee. Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies

20. **Other Conditions**:

v1.9
A.2 Permission to reproduce content for Figure 1-3b (same license agreement as A.1)
INTRODUCTION
1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS
2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.
3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:
"Reprinted from Publication title, Vol./edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."
4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.
5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.
6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.
7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.
8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.
9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.
A.3 Permission to reproduce the content of chapter 2 from Future Oncology
<table>
<thead>
<tr>
<th><strong>Author of portion(s)</strong></th>
<th>Jason Vickress, Michael Lock, Stewart Gaede, Aaaron Leung, Jeff Cao, Rob Barnett and Slav Yartsev</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Volume of serial or monograph</strong></td>
<td>13</td>
</tr>
<tr>
<td><strong>Page range of portion</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Publication date of portion</strong></td>
<td>Sept 1 2016</td>
</tr>
<tr>
<td><strong>Rights for</strong></td>
<td>Main product</td>
</tr>
<tr>
<td><strong>Duration of use</strong></td>
<td>Life of current edition</td>
</tr>
<tr>
<td><strong>Creation of copies for the disabled</strong></td>
<td></td>
</tr>
<tr>
<td><strong>With minor editing privileges</strong></td>
<td>no</td>
</tr>
<tr>
<td><strong>For distribution to</strong></td>
<td>Worldwide</td>
</tr>
<tr>
<td><strong>In the following language(s)</strong></td>
<td>Original language of publication</td>
</tr>
<tr>
<td><strong>With incidental promotional use</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Lifetime unit quantity of new product</strong></td>
<td>Up to 499</td>
</tr>
<tr>
<td><strong>Title</strong></td>
<td>Optimization of decision making for personalized radiation therapy using deformable image registration.</td>
</tr>
<tr>
<td><strong>Instructor name</strong></td>
<td>Slav Yartsev</td>
</tr>
<tr>
<td><strong>Institution name</strong></td>
<td>University of Western Ontario</td>
</tr>
<tr>
<td><strong>Expected presentation date</strong></td>
<td>Jul 2018</td>
</tr>
</tbody>
</table>

**Note:** This item will be invoiced or charged separately through CCC’s RightsLink service. More info

$ 0.00
This is not an invoice.

Total order items: 1

Order Total: 0.00 USD
TERMS AND CONDITIONS

The following terms are individual to this publisher:

None

OTHER TERMS AND CONDITIONS:

STANDARD TERMS AND CONDITIONS

1. Description of Service; Defined Terms. This Republication License enables the User to obtain licenses for republication of one or more copyrighted works as described in detail on the relevant Order Confirmation (the "Work(s)"). Copyright Clearance Center, Inc. ("CCC") grants licenses through the Service on behalf of the rightsholder identified on the Order Confirmation (the "Rightsholder"). "Republication", as used herein, generally means the inclusion of a Work, in whole or in part, in a new work or works, also as described on the Order Confirmation. "User", as used herein, means the person or entity making such republication.

2. The terms set forth in the relevant Order Confirmation, and any terms set by the Rightsholder with respect to a particular Work, govern the terms of use of Works in connection with the Service. By using the Service, the person transacting for a republication license on behalf of the User represents and warrants that he/she/it (a) has been duly authorized by the User to accept, and hereby does accept, all such terms and conditions on behalf of User, and (b) shall inform User of all such terms and conditions. In the event such person is a "freelancer" or other third party independent of User and CCC, such party shall be deemed jointly a "User" for purposes of these terms and conditions. In any event, User shall be deemed to have accepted and agreed to all such terms and conditions if User republicates the Work in any fashion.

3. Scope of License; Limitations and Obligations.

3.1 All Works and all rights therein, including copyright rights, remain the sole and exclusive property of the Rightsholder. The license created by the exchange of an Order Confirmation (and/or any invoice) and payment by User of the full amount set forth on that document includes only those rights expressly set forth in the Order Confirmation and in these terms and conditions, and conveys no other rights in the Work(s) to User. All rights not expressly granted are hereby reserved.

3.2 General Payment Terms: You may pay by credit card or through an account with us payable at the end of the month. If you and we agree that you may establish a standing account with CCC, then the following terms apply: Remit Payment to: Copyright Clearance Center, 2911B Network Place, Chicago, IL 60673-1291. Payments Due: Invoices are payable upon their delivery to you (or upon our notice to you that they are available to you for downloading). After 30 days, outstanding amounts will be subject to a service charge of 1-1/2% per month or, if less, the maximum rate allowed by applicable law. Unless otherwise specifically set forth in the Order Confirmation or in a separate written agreement signed by CCC, invoices are due and payable on a "net 30" terms. While User may exercise the "net 30" terms. While User may exercise the "net 30" terms, User must, immediately upon issuance of the Order Confirmation, the license is automatically revoked and is null and void, as if it had never been issued, if complete payment for the license is not received on a timely basis either from User directly or through a payment agent, such as a credit card company.

3.3 Unless otherwise provided in the Order Confirmation, any grant of rights to User (i) is "one-time" (including the editions and product family specified in the license), (ii) is non-exclusive and non-transferable and (iii) is subject to any and all limitations and restrictions (such as, but not limited to, limitations on duration of use or circulation) included in the Order Confirmation or invoice and/or these terms and conditions. Upon completion of the licensed use, User shall either secure a new permission for further use of the Work(s) or immediately cease any new use of the Work(s) and shall render inaccessible (such as by deleting or by removing or severing links or other locators) any further copies of the Work (except for copies printed on paper in accordance with this license and still in User's stock at the end of such period).

3.4 In the event that the material for which a republication license is sought includes third party materials (such as photographs, illustrations, graphs, inserts and similar materials) which are identified in such material as having been used by permission, User is responsible for identifying, and seeking separate licenses (under this Service or otherwise) for, any of such third party materials; without a separate license, such third party materials may not be used.

3.5 Use of proper copyright notice for a Work is required as a condition of any license granted under the Service. Unless otherwise provided in the Order Confirmation, a proper copyright notice will read substantially as follows: "Published with permission of [Rightsholder’s name], from [Work’s title, author, volume, edition number and year of copyright]; permission conveyed through Copyright Clearance Center, Inc." Such notice must be provided in a reasonably legible font size and must be placed either immediately adjacent to the Work as used (for example, as part of a by-line or footnote.
but not as a separate electronic link) or in the place where substantially all other credits or notices for the new work containing the republication Work are located. Failure to include the required notice results in loss to the Rightsholder and CCC, and the User shall be liable to pay liquidated damages for such failure equal to twice the use fee specified in the Order Confirmation, in addition to the use fee itself and any other fees and charges specified.

3.6 User may only make alterations to the Work if and as expressly set forth in the Order Confirmation. No Work may be used in any way that is defamatory, violates the rights of third parties (including such third parties' rights of copyright, privacy, publicity, or other tangible or intangible property), or is otherwise illegal, sexually explicit or obscene. In addition, User may not conjoin a Work with any other material that may result in damage to the reputation of the Rightsholder. User agrees to inform CCC if it becomes aware of any infringement of any rights in a Work and to cooperate with any reasonable request of CCC or the Rightsholder in connection therewith.

4. Indemnity. User hereby indemnifies and agrees to defend the Rightsholder and CCC, and their respective employees and directors, against all claims, liability, damages, costs and expenses, including legal fees and expenses, arising out of any use of a Work beyond the scope of the rights granted herein, or any use of a Work which has been altered in any unauthorized way by User, including claims of defamation or infringement of rights of copyright, publicity, privacy or other tangible or intangible property.

5. Limitation of Liability. UNDER NO CIRCUMSTANCES WILL CCC OR THE RIGHTSHOLDER BE LIABLE FOR ANY DIRECT, INDIRECT, CONSEQUENTIAL OR INCIDENTAL DAMAGES (INCLUDING WITHOUT LIMITATION DAMAGES FOR LOSS OF BUSINESS PROFITS OR INFORMATION, OR FOR BUSINESS INTERRUPTION) ARISING OUT OF THE USE OR INABILITY TO USE A WORK, EVEN IF ONE OF THEM HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. In any event, the total liability of the Rightsholder and CCC (including their respective employees and directors) shall not exceed the total amount actually paid by User for this license. User assumes full liability for the actions and omissions of its principals, employees, agents, affiliates, successors and assigns.

6. Limited Warranties. THE WORK(S) AND RIGHT(S) ARE PROVIDED "AS IS". CCC HAS THE RIGHT TO GRANT TO USER THE RIGHTS GRANTED IN THE ORDER CONFIRMATION DOCUMENT. CCC AND THE RIGHTSHOLDER DISCLAIM ALL OTHER WARRANTIES RELATING TO THE WORK(S) AND RIGHT(S), EITHER EXPRESS OR IMPLIED, INCLUDING WITHOUT LIMITATION IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. ADDITIONAL RIGHTS MAY BE REQUIRED TO USE ILLUSTRATIONS, GRAPHICS, PHOTOGRAPHY, ARTWORKS, INFRS OR OTHER PORTIONS OF THE WORK (AS OPPOSED TO THE ENTIRE WORK) IN A MANNER CONTEMPLATED BY USER; USER UNDERSTANDS AND AGREES THAT NEITHER CCC NOR THE RIGHTSHOLDER MAY HAVE SUCH ADDITIONAL RIGHTS TO GRANT.

7. Effect of Breach. Any failure by User to pay any amount when due, or any use by User of a Work beyond the scope of the license set forth in the Order Confirmation and/or these terms and conditions, shall be a material breach of the license created by the Order Confirmation and these terms and conditions. Any breach not cured within 30 days of written notice thereof shall result in immediate termination of such license without further notice. Any unauthorized (but licitly used) use of a Work that is terminated immediately upon notice thereof may be liquidated by payment of the Rightsholder’s ordinary license price therefor; any unauthorized (and unlicensable) use that is not terminated immediately for any reason (including, for example, because materials containing the Work cannot reasonably be recalled) will be subject to all remedies available at law or in equity, but in no event to a payment of less than three times the Rightsholder’s ordinary license price for the most closely analogous licensable use plus Rightsholder’s and/or CCC’s costs and expenses incurred in collecting such payment.

8. Miscellaneous.

8.1 User acknowledges that CCC may, from time to time, make changes or additions to the Service or to these terms and conditions, and CCC reserves the right to send notice to the User by electronic mail or otherwise for the purposes of notifying User of such changes or additions; provided that any such changes or additions shall not apply to permissions already secured and paid for.

8.2 Use of User-related information collected through the Service is governed by CCC’s privacy policy, available online here: http://www.copyright.com/content/cc3/en/tools/footer/privacypolicy.html.

8.3 The licensing transaction described in the Order Confirmation is personal to User. Therefore, User may not assign or transfer to any other person (whether a natural person or an organization of any kind) the license created by the Order Confirmation and these terms and conditions or any rights granted hereunder; provided, however, that User may assign such license in its entirety on written notice to CCC in the event of a transfer of all or substantially all of User’s rights in the new material which includes the Work(s) licensed under this Service.

8.4 No amendment or waiver of any terms is binding unless set forth in writing and signed by the parties. The Rightsholder and CCC hereby object to any terms contained in any writing prepared by the User or its principals, employees, agents or affiliates and purporting to govern or otherwise relate to the licensing transaction described in the Order Confirmation, which terms are in any way inconsistent with any terms set forth in the Order Confirmation and/or in these terms and conditions or CCC’s standard operating procedures, whether such writing is prepared prior to, simultaneously with or subsequent to the Order Confirmation, and whether such writing appears on a copy of the Order Confirmation or in a separate instrument.

8.5 The licensing transaction described in the Order Confirmation document shall be governed by and construed under the law of the State of New York, USA, without regard to the principles thereof of conflicts of law. Any case, controversy, suit, action, or proceeding arising out of, in connection with, or related to such licensing transaction shall be brought, at CCC’s sole discretion, in any federal or state court located in the County of New York, State of New York, USA, or in any federal or state court whose geographical jurisdiction covers the location of the Rightsholder set forth in the Order Confirmation. The parties expressly submit to the personal jurisdiction and venue of each such federal or state court. If you have any comments or questions about the Service or Copyright Clearance Center, please contact us at 978-750-8400 or send an email to info@copyright.com.
A.4 Permission to reproduce the content of chapter 3 and 4 from Physics in Medicine and Biology

From the publisher's website https://publishingsupport.iopscience.iop.org/permissions-faqs/

“Upon transfer of copyright, IOP and/or the copyright owner grants back to authors a number of rights. These include the right to include the Final Published Version of the article in your research thesis or dissertation. Please include citation details and, for online use, a link to the Version of Record. IOP’s permission will be required for commercial use of an article published as part of your thesis. IOP does not allow ProQuest to publish or sell the article as part of your dissertation.”
Appendix B: Description of clinical parameters and tools from chapter 2

B.1 Important clinical measures for liver cancer patients

In the diagnosis and treatment of liver cancer there are a number of important clinical measures that are used to describe a patient’s overall liver function. Bilirubin is a protein found in the blood and is created in the liver with high levels being indicative of liver disease. Serum albumin is the most abundant protein found in human blood, and low levels of serum albumin are also indicative of liver disease. Ascites describes a patient retaining water within their liver and another indicator of liver disease. Child Pugh score is a common measure of a patient’s liver function ranging from A to C indicating good to poor liver function respectively. Child Pugh score is calculated from the combination of different clinical factors including bilirubin, serum albumin and ascites.

B.2 Interpreting Cox proportional hazard model results.

The Cox proportional hazard model is used to find the association between specific variables (either continuous or binary) and overall survival. The Cox proportional hazard model is a linear regression solving for equation 1, where \( \lambda(t) \) is the hazard rate, \( \lambda_0(t) \) is the baseline hazard function, \( X_1-X_n \) are the different variables and \( \beta_1-\beta_n \) are the model parameters determined for each variable.

\[
\lambda(t) = \lambda_0(t)\exp(\beta_1X_1 + \cdots + \beta_nX_n)
\] (1)

Hazard rate describes the risk of death and a higher hazard rate describes lower survival. The primary results of a Cox proportional hazard model are the \( \beta \) values, hazard ratio (with 95% confidence interval) and p-values. \( \beta \) values describe the relationship each variable has on survival for example a positive \( \beta \) value indicates a higher hazard rate and lower survival. Hazard ratios also describe the effect each variable has on overall survival, hazard ratio > 1 describes poor survival and hazard ratio < 1 describes better survival. A p-value measures the significance of the variable on overall survival, with a p-value < 0.05 indicating a significant result.
Cox proportional hazard model results are presented for both univariate and multivariate analysis. Univariate analysis describes the individual effect of each variable on survival. Multivariate analysis describes the effect each variable has on survival when combined with other significant variables chosen for the final model.

B.3 Applying a nomogram

Clinical nomograms visualize predictive models created using a cox proportional hazard model for application in the clinic. A diagram describing the process of a nomogram is presented in figure B-1, describing how the score is summed and converted to a one year and six month survival probability.

Figure B-1: Diagram describing the use of a clinical nomogram, using clinical variables to determine the probability of one year and six month survival.
Appendix C: Description of the Scale Invariant Feature Transform (SIFT)

C.1 Scale invariance

The Scale Invariant Feature Transform (SIFT) was engineered to match points between two images regardless of relative scale (scale invariant). If two images are taken with the same imaging system, they can be compared and matched based on the voxel intensities directly. When the imaging systems are different direct voxel comparison is no-longer valid because of scaling differences. In order to compare two images together regardless of scale the difference of Gaussian images is selected as presented in Lowe et al [1].

The difference of Gaussian images is computed by convolving an image with the Gaussian filter with a defined scaling $\sigma$ (equation 1) and subtracting it from an image convolved with a Gaussian filter with different scaling $k\sigma$ (equation 2).

$$L(x, y, z, \sigma) = \frac{1}{\sqrt{4\pi\sigma}^2} e^{-(x^2+y^2+z^2)/2\sigma^2} * I(x, y, z)$$ (1)

$$D(x, y, z, k\sigma) = L(x, y, z, k\sigma) - L(x, y, z, \sigma)$$ (2)

The result is an image describing the gradients within the image, regardless of scale. Also the Laplacian of the Gaussian can be approximate from the difference of Gaussian images, using a modified heat equation using $\sigma$ in place of $t$ (shown in equation 3).

$$\sigma\nabla^2 G = \frac{\partial G}{\partial \sigma} \approx \frac{G(x,y,z,k\sigma) - G(x,y,z,\sigma)}{k\sigma - \sigma}$$ (3)

The result is an approximate Laplacian of the Gaussian at every point in the image volume and at different scaling values construct a 4D $(x,y,z,\sigma)$ matrix used to locate unique points.

C.2 Finding unique points

Once the 4D matrix $(x,y,z,\sigma)$ of the Laplacian of the Gaussian is constructed the next step is to find local maximum and minimum values. This is performed by finding the local
maximum and minimum voxels when compared to all the nearest neighbors across all 4 dimensions (80 neighboring points).

C.3 Removal of edge points

With all the unique points located the next step is to remove points that are located on an edge because they can have strong difference of Gaussians while not actually being unique. This is performed by looking at the magnitude of the primary and secondary eigen-vectors (direction with the highest rate of change). If the magnitude of the primary eigen-vector is $H_{max}$ times greater than the secondary eigenvector it is located on an edge and excluded as a unique point. The parameter $H_{max}$ is tuned for the specific imaging system being used.

C.4 Point localization

All landmark locations are limited by the resolution of the image volume being used, but perhaps the true unique point could lie in between 3D voxels. To interpolate between voxels and localize the approximate maximum or minimum a 2nd order Taylor expansion is used (shown in equation 4) with $x = (x,y,z,\sigma)$. To locate the maximum and minimum the derivative of the Taylor series is set to zero and solved for the localized point $v$ (shown in equation 5).

$$D(x) = D + \frac{\partial D^T}{\partial x} x + 0.5x^T\frac{\partial^2 D}{\partial x^2} x$$  \hspace{1cm} (4)$$

$$v = -\frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x}$$  \hspace{1cm} (5)$$

Once the point is localized in 4D, the new value is measured from the Taylor series and if it is below $r_{thresh}$ it will be excluded. The $r_{thresh}$ parameter is selected to remove points with low contrast.

C.5 Point descriptors and matching

With all unique points located in an image study the next step is to create descriptors describing each point. The first step is to subdivide the surrounding voxels into 4x4x4 sub volumes each containing 4x4x4 voxels, totaling 16x16x16 voxels centered on the unique
point being described. For each voxel the gradient magnitude and direction (both azimuthal and elevation directions) is computed. Now each sub volume includes the magnitude and direction for 4x4x4 voxels and the gradient magnitude is binned into an 8 by 4 matrix based on their azimuthal (8 directions between 0 and 360) and elevation (4 directions between 0 and 180) directions respectively. To have more distinct feature matching the gradient magnitudes can be scaled based on their distance from the center of the sub volume. The final descriptor is a 4x4x4x8x4 matrix describing each unique point.

With each unique point in two image volumes having a descriptor the next step is to match like points between each image. The fastest matching method is to convert each descriptor (4x4x4x8x4) into a 2,048 dimensional vector and finding the Euclidian distance between every possible unique point pair. For example for one unique point you would calculate the Euclidean distance between the 2,048 dimensional vectors from all points in the other image and find the point with the smallest “Euclidian distance”. If the next closest point has a Euclidian distance 1.66 times greater than they are matched together. The distance to the next most similar point (1.66) is chosen to find the balance between the number and quality of landmarks, a smaller multiplier would yield more points with potentially less similarity. The end results are landmarked points between two 3D volumes.

C.6 Optimization

When implementing a SIFT algorithm there is a balance between the number and quality of matched points. This balance is changed by changing the $k \sigma, H_{max, r_{\text{thresh}}}$ and distance to next most similar point. Each of these parameters affects how many unique points are considered or the matching criteria and they need to be optimized for the specific type of images (etc. CT or MRI) or what is being imaged (etc. thorax vs brain).

C.7 References

Curriculum Vitae of Jason Robert Vickress

Jason Robert Vickress

Citizenship  Canadian

Languages  English

Degrees
2014/9 – Present  Doctorate, Medical Biophysics, University of Western Ontario
   Degree Status: In Progress
2008/09 – 2013/06  Bachelor’s, Applied Science, Honours Nanotechnology Engineering, Co-operative Program, University of Waterloo
   Degree Status: Completed, with distinction

Recognitions
2017/5-2019/4  NSERC – Alexander Graham Bell Canada Graduate Scholarship-Doctoral
   CGS D Scholarship – 70,000 (Canadian dollar)
   The Natural Sciences and Engineering Research Council (NSERC)
   Distinction
2015/9 – 2016/8  CIHR Strategic Training Program in Cancer Research and Technology Transfer (CaRTT) – 100 (Canadian dollar)
   Canadian Institute of Health Research (CIHR)
   Distinction
2015/9 – 2016/8  CIHR CGSM Scholarship – 17,500 (Canadian dollar)
   Canadian Institute of Health Research (CIHR)
   Distinction
2014/9 – 2015/8  Ontario Graduate Scholarship (OGS) – 15,000 (Canadian dollar)
   Government of Ontario
   Distinction

Employment
2018/1- 2018/4  Teaching Assistant, a course in Radiobiology taught by Dr. Scott Karnas
   The University of Western Ontario, London, Ontario, Canada
2017/1- 2017/4  Teaching Assistant, a course in Radiobiology taught by Dr. Jerry Battista
   The University of Western Ontario, London, Ontario, Canada
2016/1- 2016/4  Teaching Assistant, a course in Radiobiology taught by Dr. Jerry Battista
   The University of Western Ontario, London, Ontario, Canada
2013/5 - 2014/8  Physics Assistant
   London Regional Cancer Program, London, Ontario, Canada
2012/1 - 2012/8  
Physics Assistant  
London Regional Cancer Program, London, Ontario, Canada

2010/9 - 2011/4  
System Support Specialist  
Just Energy Inc, Mississauga, Ontario, Canada

2010/2 - 2010/4  
Research Assistant  
Department of Chemistry, University of Waterloo, Waterloo, Ontario Canada

Publications


1. Data inventory for cancer patients receiving radiotherapy for outcome analysis and modeling  

Presentations

Imaging Network of Ontario, 2018 March 28 & 29, Toronto, ON, Canada  
*Online assessment of dose changes in head and neck radiotherapy without dose re-computation using deformable image registration.* (oral presentation)  
**Vickress J**, Barnett R, Battista J, Yartsev S.

Imaging Network of Ontario, 2017 March 15 & 16, London, ON, Canada  
*Determining dosimetric impact of deformable image registration error in lung radiotherapy* (poster)  
**Vickress J**, Barnett R, Battista J, Yartsev S.

Medical Imaging Summer School, 2016 August 1 – 5, Favignana, Italy  
*Prediction of deformable image registration error in lung 4DCT* (poster)  

Imaging Network of Ontario, 2016 March 30 & 31, Toronto, ON, Canada  
*Prediction of the spatial distribution of deformable image registration error in lung 4DCTs* (poster)  

World Congress on Medical Physics & Biomedical Engineering, 2015, June 7-12  
*Developing predictive models using retrospective study of liver cancer patients treated with radiation therapy* (poster)  
**Vickress J**, Lock M, Wong E, Barnett R, Yartsev S.
Imaging Network of Ontario, 2015 March 30 & 31. London, ON, Canada
Evaluation of deformable image registration accuracy using virtual phantoms (poster)
Vickress J, Barnett R, Battista J, Yartsev S

Department of Oncology – Research & Education Day – 2014, June 20, 2014, London, ON, Canada
Data inventory for cancer patients receiving radiotherapy for outcome analysis and decision support (poster)
Vickress J, Barnett R, Yartsev S

Awards and Certification

Dean’s Honour List, University of Waterloo, 2009/1