Deformable Dose Reconstruction to Optimize the Planning and Delivery of Liver Cancer Radiotherapy

by

Michael Velec

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Institute of Medical Science
University of Toronto

© Copyright by Michael Velec 2014
Deformable Dose Reconstruction to Optimize the Planning and Delivery of Liver Cancer Radiotherapy

Michael Velec
Doctor of Philosophy
Graduate Department of Institute of Medical Science
University of Toronto
2014

Abstract

The precise delivery of radiation to liver cancer patients results in improved control with higher tumor doses and minimized normal tissues doses. A margin of normal tissue around the tumor requires irradiation however to account for treatment delivery uncertainties. Daily image-guidance allows targeting of the liver, a surrogate for the tumor, to reduce geometric errors. However poor direct tumor visualization, anatomical deformation and breathing motion introduce uncertainties between the planned dose, calculated on a single pre-treatment computed tomography image, and the dose that is delivered.

A novel deformable image registration algorithm based on tissue biomechanics was applied to previous liver cancer patients to track targets and surrounding organs during radiotherapy. Modeling these daily anatomic variations permitted dose accumulation, thereby improving calculations of the delivered doses. The accuracy of the algorithm to track dose was validated using imaging from a deformable, 3-dimensional dosimeter able to optically track absorbed dose.

Reconstructing the delivered dose revealed that 70% of patients had substantial deviations from the initial planned dose. An alternative image-guidance technique using respiratory-correlated
imaging was simulated, which reduced both the residual tumor targeting errors and the magnitude of the delivered dose deviations. A planning and delivery strategy for liver radiotherapy was then developed that minimizes the impact of breathing motion, and applied a margin to account for the impact of liver deformation during treatment. This margin is 38% smaller on average than the margin used clinically, and permitted an average dose-escalation to liver tumors of 9% for the same risk of toxicity. Simulating the delivered dose with deformable dose reconstruction demonstrated the plans with smaller margins were robust as 90% of patients’ tumors received the intended dose. This strategy can be readily implemented with widely available technologies and thus can potentially improve local control for liver cancer patients receiving radiotherapy.
For my parents.
Acknowledgments

My interest in research and motivation to pursue doctoral studies was without a doubt inspired by my supervisor Kristy Brock. I feel privileged to have worked under the leader in the field who at the same time is a terrific teacher. She challenged me as a researcher and opened me up to numerous new opportunities. I continue to admire her drive as a scientist and passion for research while always remaining friendly and approachable to everyone around her.

The guidance from my thesis committee was also invaluable. Laura Dawson has been a longtime supporter of my pursuit of research, and as a leader in liver radiotherapy she contributed her clinical expertise helping keep this work focused and relevant. Michael Milosevic balanced the committee and contributed some very welcomed and practical ideas.

It has been such a pleasure to work with the diverse members of the MORFEUS lab, each sharing their unique expertise: Adil Al-Mayah, Shannon Hunter, Jenny Lee, Deirdre McGrath, Joanne Moseley, Thao-Nguyen Nguyen, Carolyn Niu, Navid Samavati, and Roxana Vlad. They made for an incredibly friendly and collaborative research team to which I owe a lot and already miss. I am especially grateful to Joanne for writing nearly all of the MATLAB code for my projects.

I am appreciative of help and consultations from: Tim Craig, Andrea Marshall, Tu Huan Tran, and Anand Swaminath. For their often needed technical support I must acknowledge: Stuart Rose, Gavin Disney, Doug Moseley, Michael Sharpe, Kevin Wang, the entire RMP Helpdesk team and most importantly Graham Wilson. Jan-Jakob Sonke and Marcel van Herk from the Netherlands Cancer Institute helped with the 4D cone-beam CT imaging code, and Titania Juang and Mark Oldham from Duke University provided and helped with the dosimeter data. I also would like to thank the Radiation Therapy leadership for giving me freedom to pursue graduate studies while maintaining my clinical appointment.

My research was funded with awards from the Canadian Institutes of Health Research (Fellowship Award, Masters Award, travel grants) and Kristy Brock’s grants from the U.S. National Institutes of Health (5RO1CA124714-02). I was also supported by the Princess Margaret Cancer Foundation (Robert G. Matthews Fund) and the Canadian Association of Medical Radiation Technologists (Member’s Grant). I would like to acknowledge the patients’ clinical trial participation, which provided the majority of the data used for this research.
Finally I wish to thank those closest to me personally. My close circle of friends has always inspired me with their ambition and success in each of their own careers. Erin Petkopoulos has supported me in so many ways and I am thankful for her love, patience and motivation to get me writing! My family, Marijana, Frank, Sabina and Dave, have always been encouraging of all my choices. Hvala lepa, rad te imam. My dad in particular has always encouraged me to constantly strive for bigger and better and we both agree that it definitely doesn’t end here.
# Table of Contents

Acknowledgments ........................................................................................................... v

Table of Contents ........................................................................................................... vii

List of Abbreviations ..................................................................................................... xi

List of Tables ................................................................................................................... xiii

List of Figures ................................................................................................................... xiv

Chapter 1 Introduction and Objectives ......................................................................... 1

1.1 Background and Significance .................................................................................. 2
1.1.1 Cancers in the liver ............................................................................................ 2
1.1.2 Treatment modalities ....................................................................................... 2
1.1.3 Radiotherapy for liver cancer ........................................................................... 4
1.1.4 Technical challenges of liver SBRT .................................................................. 7
1.1.5 Geometric accuracy of image-guided liver SBRT .............................................. 11
1.1.6 Modeling geometric uncertainties into dose calculations ............................... 15
1.1.7 Deformable image registration ....................................................................... 18
1.1.8 DIR validation methods ................................................................................... 20
1.1.9 Applications of DIR in liver radiotherapy ....................................................... 21

1.2 Rationale for Study ................................................................................................ 23

1.3 Hypothesis and Specific Aims .............................................................................. 24

1.4 Outline of Thesis .................................................................................................... 25

Chapter 2 MORFEUS: Framework and Prior Investigations ...................................... 27

2.1 Theory ..................................................................................................................... 28
2.2 Deformable Image Registration ........................................................................... 28
2.3 Deformable Dose Accumulation .......................................................................... 31
2.4 Validation Studies .................................................................................................. 32
Chapter 3 Accumulated Dose in Liver Stereotactic-Body Radiotherapy: Positioning, Breathing and Deformation Effects ................................................................. 36

3.1 Introduction .......................................................................................... 37

3.2 Methods and Materials ......................................................................... 38
  3.2.1 Patients and SBRT planning ............................................................. 38
  3.2.2 Image-guided radiotherapy .............................................................. 40
  3.2.3 DIR and dose accumulation ............................................................. 40
  3.2.4 Comparison of geometric uncertainties ......................................... 42

3.3 Results .................................................................................................. 44
  3.3.1 Geometric uncertainties ................................................................. 44
  3.3.2 Accumulated dose changes ............................................................. 45

3.4 Discussion ............................................................................................ 50

3.5 Conclusions .......................................................................................... 52

Chapter 4 Deformable Dose Reconstruction to Evaluate IGRT Strategies in Free-Breathing Liver SBRT ...................................................................................... 53

4.1 Introduction .......................................................................................... 54

4.2 Methods.................................................................................................. 55
  4.2.1 Image-guidance strategies ............................................................... 55
  4.2.2 Evaluation ...................................................................................... 58

4.3 Results .................................................................................................. 59
  4.3.1 Geometric impact of each IGRT strategy ....................................... 59
  4.3.2 Dosimetric impact of each IGRT strategy overall .......................... 61
  4.3.3 Delivered GTV doses .................................................................... 66
  4.3.4 Delivered normal tissue doses ......................................................... 67
  4.3.5 Relative impact of geometric errors for each strategy ................. 68

4.4 Discussion ............................................................................................ 69
Chapter 5 Preliminary Validation of Biomechanical-Based Deformable Dose Reconstruction

5.1 Introduction .................................................................................................................. 75

5.2 Methods ........................................................................................................................ 76
  5.2.1 Dosimeter characteristics and experimental setup .................................................. 76
  5.2.2 DIR models .............................................................................................................. 77
  5.2.3 Evaluation and dose warping ................................................................................... 79

5.3 Results ............................................................................................................................ 81
  5.3.1 Comparison with the relative optical CT distribution ............................................. 82
  5.3.2 Dosimetric evaluation ............................................................................................. 83

5.4 Discussion ....................................................................................................................... 85

5.5 Conclusions .................................................................................................................... 88

Chapter 6 Simplified Strategies to Determine the Mean Respiratory Position For Liver Radiation Therapy Planning

6.1 Introduction .................................................................................................................... 90

6.2 Methods and Materials ................................................................................................ 91
  6.2.1 Patient and imaging data ....................................................................................... 91
  6.2.2 Deformable image registration ............................................................................. 92
  6.2.3 Strategies to estimate the mean respiratory position ............................................ 92
  6.2.4 PTV margin calculations ..................................................................................... 94

6.3 Results ............................................................................................................................ 96
  6.3.1 Respiratory motion quantification ........................................................................ 96
  6.3.2 4D GTV position versus mean position .............................................................. 97
  6.3.3 Mid-ventilation CT and mid-position CT accuracy .............................................. 98
  6.3.4 Impact of mean position on the required margin ................................................ 99

6.4 Discussion ....................................................................................................................... 99

6.5 Conclusions .................................................................................................................... 101
**List of Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>4D CBCT</td>
<td>Four-dimensional cone-beam computed tomography</td>
</tr>
<tr>
<td>4D CT</td>
<td>Four-dimensional computed tomography</td>
</tr>
<tr>
<td>ABC</td>
<td>Active breathing control</td>
</tr>
<tr>
<td>AP</td>
<td>Anterior-posterior</td>
</tr>
<tr>
<td>B-spline</td>
<td>Basis spline</td>
</tr>
<tr>
<td>CT</td>
<td>Computed tomography</td>
</tr>
<tr>
<td>CTV</td>
<td>Clinical target volume</td>
</tr>
<tr>
<td>CBCT</td>
<td>Cone-beam computed tomography</td>
</tr>
<tr>
<td>COM</td>
<td>Centre-of-mass</td>
</tr>
<tr>
<td>D_{acc}</td>
<td>Accumulated dose</td>
</tr>
<tr>
<td>DIR</td>
<td>Deformable image registration</td>
</tr>
<tr>
<td>DSC</td>
<td>Dice similarity coefficient</td>
</tr>
<tr>
<td>D_{plan}</td>
<td>Planned dose</td>
</tr>
<tr>
<td>D_{pred}</td>
<td>Predicted dose</td>
</tr>
<tr>
<td>DTA</td>
<td>Distance-to-agreement</td>
</tr>
<tr>
<td>FEA</td>
<td>Finite element analysis</td>
</tr>
<tr>
<td>FEM</td>
<td>Finite element method</td>
</tr>
<tr>
<td>FWHM</td>
<td>Full-width at half-maximum</td>
</tr>
<tr>
<td>GTV</td>
<td>Gross tumor volume</td>
</tr>
<tr>
<td>HCC</td>
<td>Hepatocellular carcinoma</td>
</tr>
<tr>
<td>IGRT</td>
<td>Image-guided radiotherapy</td>
</tr>
<tr>
<td>IMRT</td>
<td>Intensity-modulated radiotherapy</td>
</tr>
<tr>
<td>ITV</td>
<td>Internal target volume</td>
</tr>
<tr>
<td>kV</td>
<td>Kilovoltage</td>
</tr>
<tr>
<td>LR</td>
<td>Left-right</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic resonance imaging</td>
</tr>
<tr>
<td>MI</td>
<td>Mutual information</td>
</tr>
<tr>
<td>MV</td>
<td>Megavoltage</td>
</tr>
<tr>
<td>NTCP</td>
<td>Normal tissue complication probability</td>
</tr>
<tr>
<td>PDF</td>
<td>Probability distribution function</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>PEI</td>
<td>Percutaneous ethanol injection</td>
</tr>
<tr>
<td>PMH</td>
<td>Princess Margaret Hospital</td>
</tr>
<tr>
<td>PTV</td>
<td>Planning target volume</td>
</tr>
<tr>
<td>PVT</td>
<td>Portal vein thrombosis</td>
</tr>
<tr>
<td>RFA</td>
<td>Radio-frequency ablation</td>
</tr>
<tr>
<td>RILD</td>
<td>Radiation-induced liver disease</td>
</tr>
<tr>
<td>SBRT</td>
<td>Stereotactic-body radiotherapy</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SI</td>
<td>Superior-inferior</td>
</tr>
<tr>
<td>SSD</td>
<td>Sum of squared differences</td>
</tr>
<tr>
<td>TACE</td>
<td>Trans-arterial chemo-embolization</td>
</tr>
<tr>
<td>TD$_{5/5}$</td>
<td>Tolerance dose for 5% risk at 5 years</td>
</tr>
<tr>
<td>$V_{\text{eff}}$</td>
<td>Effective Volume</td>
</tr>
<tr>
<td>VMAT</td>
<td>Volumetric-modulated arc therapy</td>
</tr>
</tbody>
</table>
List of Tables


1.2 Studies reporting the accuracy of image-guided liver radiotherapy based on various imaging modalities and tumor surrogates. Pages 13, 14.

1.3 Summary of intra- and inter-fraction dose accumulation studies for liver radiotherapy. Page 17.

2.1 Accuracy of Morfeus DIR in the liver. Page 33.


3.2 Population geometric errors (in mm) at planning and during 6-fraction SBRT evaluated with deformable image registration. Page 48.

3.3 Accumulated treatment dose ($D_{acc}$) deviations relative to the planned static dose ($D_{plan}$) and the predicted 4D CT breathing dose ($D_{pred}$). Page 52.

4.1 Residual mean (M), systematic ($\Sigma$), and random (σ) errors following each IGRT strategy, between the planning exhale 4D CT and treatment exhale 4D CBCTs (in mm). Page 60.

4.2 Deviations between the delivered dose and the planning 4D CT-predicted dose. Page 65.

5.1 Field location and shape errors (in mm) between the dose distributions predicted using two DIR models, each compared to the measured distribution of optical density changes measured via optical CT (predicted–measured). Page 83.

6.1 Patient and imaging details. Page 91.

6.2 Example calculation of the inferior PTV margin applied to the GTV on the exhale 4D CT where the breathing amplitude is fully encompassed (ITV-based PTV), or on the mid-position CT where the amplitude is incorporated into the PTV as a random error. Page 96.

6.3 Maximum amplitude on 4D CT (in mm). Page 96.

6.4 Errors in liver GTV position for different planning CT datasets relative to the time-averaged position (in mm). Page 98.

7.1 Planning dosimetry for the dose-probability PTV plans, and its change from the clinical and ITV-based PTV plans. Page 110.

7.2 Increase in the delivered GTV doses with the dose-probability PTV plan compared to the delivered doses with the clinical or ITV-based PTV plans. Page 112.
List of Figures

1.1 The impact of geometric radiotherapy errors on the dose distribution (coloured lines). Figure adapted from [67,68]. Page 8.

1.2 A ‘ground-truth’ correspondence can be established between primary (left) and secondary (middle) images by identifying common features (e.g. contour=GTV, arrow=vessel bifurcation). The actual feature position in the image can then be compared to the position predicted by DIR for validation (right). The difference is equal to the registration error. Page 21.

1.3 Deformation from the nominal planning geometry causes uncertainty in the intended dose distribution (coloured lines), which can be resolved with DIR. Page 22.


3.1 Example of multi-organ deformable registration between exhale 4D CT (left) and exhale 4D CBCT (middle), and the resulting deformation map (right). Page 41.

3.2 Schema of dose accumulation comparisons to extract the contribution of geometric uncertainties on dose deviations: (A) 4D CT-predicted breathing dose (D_{pred}); (B) full SBRT dose accumulation (D_{acc}); (C) deformable inter-fraction accumulation with the 4D CT breathing motion applied; (D) rigid inter-fraction accumulation with the 4D CT breathing motion applied. Abbreviations: Morfeus=DIR. Rigid=liver-to-liver COM registration. Fx=fractions. Page 43.

3.3 Deviations from the accumulated dose (D_{acc}) are shown in a dose difference image (top) and dose-volume histogram (bottom). PTV coverage was compromised inferiorly on the static exhale 4D CT plan (D_{plan}) to spare the large bowel, whereas 4D CT predicted decreased dose (D_{pred}) as these tissues moved inferior away from the high-dose region. Geometric errors seen on 4D CBCT moved these tissues back toward the high-dose region. Page 46.

3.4 Relative contribution of geometric errors to changes between the accumulated (D_{acc}) and predicted (D_{pred}) dose for all patients’ stomachs and bowels. Data is sorted by decreasing magnitude and direction of dose deviation. Page 49.

3.5 An example of a large dose deviation caused primarily by deformation. Left: patient was planned on exhale 4D CT for 30 Gy with abdominal compression (liver=purple; GTV=red). Middle: deformation in the abdomen shown on CBCT was possibly due to a misaligned compression plate and increased stomach contents. Right: dose difference between the accumulated and predicted doses (D_{acc}–D_{pred}) showing delivered decreases (dark blue region) to the minimum GTV of 4.3 Gy, and the maximum duodenum by 11.5 Gy. Page 50.
4.1 Conceptual comparison of the three IGRT strategies. The planning exhale 4D CT image (left column) is common to all strategies while they differ by means of the CBCT (3D/4D), the registration type (rigid/deformable), and the targeted structure (liver (magenta)/GTV (cyan)). The dashed lines indicate the potential for structure alignment (green) or misalignment (red). Page 57.

4.2 Schema of how the delivered doses were compared for each IGRT strategy. Page 59.

4.3 Population residual geometric errors (based on each tissue’s centre-of mass displacement) following three image-guidance strategies, relative to the planning exhale 4D CT position. Pages 62, 63.

4.4 The mean (absolute) total dose deviations (%) between the delivered dose and the 4D CT-predicted dose (square markers), for each tissue and IGRT strategy. The stacked colored bars depict the mean (absolute) impact of the different geometric errors towards the total dose deviation. Page 64.

5.1 A. CT images of the dosimeter with and without the compression applied (with the plates digitally removed). B. Model of the experimental setup. C. The dose distribution planned on the CT in the compressed geometry. Page 77.

5.2 A. Each deformed field was auto-segmented at half the maximum value of the intensity profile across that field. B. The full widths at half the maximum (FWHM) values were measured in the horizontal and vertical directions to quantify the field shape (optical CT is shown). Page 80.

5.3 Cumulative frequency distribution of the node-by-node displacement differences between MorfeusSurfProj and MorfeusPlates. Page 81.

5.4 Relative deformed dose distributions in the dosimeter’s uncompressed geometry measured with optical CT (represented as change in optical density), or predicted using either DIR model. Central cross-sections are shown in the coronal (top), axial (middle), and sagittal (bottom) planes. Page 82.

5.5 Dose profiles along a line bisecting the central row of deformed dose fields at varying coronal cross-sections. Page 84.

5.6 Deformed dose distributions predicted with (A) biomechanical-model based DIR from the present study, and (B) a commercially available intensity-based DIR (Velocity) previously evaluated by Juang et al. (reproduced with permission[151]). Page 86.

6.1 Left: Schema showing the 4D CT tumor motion (solid line, white circles); the possible mid-ventilation CT phase, the mid-position CT (gray circle) along the linear exhale-to-inhale trajectory (dashed line), versus the time-averaged position (black circle). Note that it is possible for multiple phases (e.g. 20% or 80%) to be selected as the mid-ventilation CT if they are equally close to the time-averaged position. Right: the SI diaphragm (white boxes) position on 4D CT. Mid-ventilation CT was selected as the phase with the diaphragm closest to the mean diaphragm position (gray box). The relative SI mean diaphragm position was also applied to the exhale-to-inhale 4D CT deformation map to
generate the mid-position CT. Page 93.

6.2 The local rigid registration volume is around the left diaphragm, and excludes the chest wall and heart. A mutual information algorithm was used, only considering translations. Page 93.

6.3 Workflows to select a mid-ventilation CT or reconstruct a mid-position CT. Page 95.

6.4 Diaphragm and GTV motion over all 4D CT phases for 1 patient (GTV locations shown on the inset figure). Page 97.

7.1 Schema of the different PTV strategies (grey volumes). Clinically, the GTV was delineated on exhale 4D CT (1). DIR of 4D CT allowed for propagation of the exhale contours into the inhale (2) or time-averaged position (3). Figure adapted from [76]. Page 108.

7.2 Workflows to model the delivered dose following CBCT-guidance based on the liver’s exhale or mean breathing position. Page 108.

7.3 PTV size (total SI margin) versus the breathing amplitude. Above 6 mm of motion the difference in total SI expansion exceeds 5 mm between the dose-probability and ITV-based PTVs. Page 109.

7.4 Delivered min GTV dose deviations relative to the planning PTVD95 for (A) all patients combined, and (B) separately for three patients each with a decrease ≥1 Gy to any GTV (Number 2 and 3 each have three GTVs). Page 113
Chapter 1

Introduction and Objectives
1 Introduction and Objectives

1.1 Background and Significance

1.1.1 Cancers in the liver

Primary liver cancers including hepatocellular carcinoma (HCC) and intra-hepatic cholangiocarcinoma are now the 5th most common cancer worldwide and 3rd leading cause of cancer death[1]. Regional incidence varies and is highest in Eastern Asia and Africa relating to the high rates of Hepatitis B infection[1,2]. Western countries have a lower but rising incidence which is additionally associated with liver cirrhosis secondary to alcohol intake, fatty liver diseases, and the simultaneous rising incidence of Hepatitis C infection[3]. The 5 year survival rate for HCC is <5%[4].

The liver is also a common site of metastases of other cancers. For colorectal cancer patients in particular, more than half will develop liver metastases[5,6]. One quarter of colorectal cancer patients will have the spread confined to the liver[7]. If the number and size of the metastases are limited or are confined to one organ, an intermediate metastatic state termed ‘oligometastases’, local therapies could delay progression and even be potentially curative.

1.1.2 Treatment modalities

Surgery is curative in well-selected patients with 5 year survival rates of 43% for HCC[8], and 47% for patients with colorectal liver metastases[9]. Liver transplantation for HCC has also resulted in 5-year survival rates as high as 80%[10], but has no role for metastatic disease. Local recurrence rates unfortunately exceed 50% at 5 years following resection[8]. Furthermore, 80% of patients are inoperable due to high tumor burden, macrovascular tumor invasion termed portal vein thrombosis (PVT), poor liver function and cirrhosis.

Percutaneous and laparoscopic therapies are often employed for liver tumors. Radio-frequency ablation (RFA) raises the tumor temperature to over 60°C through an inserted electrode, destroying the tumor’s cells and vasculature. Five year survival rates are 41–77% for HCC and 18–49% for colorectal metastases, with regional recurrence rates higher than for surgical resection[11,12]. RFA is less effective for tumors >4 cm or near the diaphragm[13], and for PVTs as the large blood vessels act as a heat sink reducing RFA’s mode of action[14]. Percutaneous ethanol injection (PEI) causes dehydration and complete necrosis for tumors <2
cm, is simple to administer, inexpensive, and is well tolerated[15]. PEI for HCC has a 5 year survival rate of nearly 50% with higher recurrence rates than RFA for tumors >2 cm[16].

Local treatment using radioactive isotopes, or brachytherapy, can help manage liver cancers. Yittrium-90 (\(^{90}\text{Y}\)), a beta emitter encapsulated in glass microspheres with a half-life of 2.7 days, has been administered via radioembolization through the hepatic artery. Up to 150 Gy can be deposited into segments of the liver with \(^{90}\text{Y}\) and it is a good choice for diffuse disease, although heterogeneous distribution of doses is a potential technical limitation[17,18]. Median survival ranges from 7.7–17.2 months for HCC, and 4.5–10.5 months for colorectal metastases[19,20]. If \(^{90}\text{Y}\) is contraindicated, the beta emitting Iodine-131 (\(^{131}\text{I}\)) has a half-life of 8.04 days and can be tagged to a fatty contrast agent (e.g. Lipiodol) which remains inside HCC tumors long after administration[21]. Two year survival of HCC following \(^{131}\text{I}\) including treatment of PVTs has been reported at 60%[22]. Response rates are lower however for liver metastases[21].

Systemic chemotherapy is generally poorly tolerated and ineffective for liver cancers[23]. Regional liver administration can be delivered through trans-arterial chemo-embolization (TACE) by interrupting the hepatic artery, the primary blood source for liver tumors. Because the normal liver receives blood from the portal vein primarily, high chemotherapy doses are directed to the tumor. HCC survival at 2 years increases from 11–27% with supportive care to 31–63% with TACE[24,25]. Colorectal metastases survival at 2 years following TACE is 28%[26]. The chance of long-term cure remains low with TACE and it is also ineffective for PVTs.

Systemic biological agents under investigation target the molecular basis of tumor growth. Sorafenib for example inhibits both tumor proliferation and tumor angiogenesis. Llovet et al. demonstrated a median survival of 10.7 months (44% increase) when treating advanced HCC with Sorafenib versus placebo[27]. Combining local therapies like radiation therapy (described in the following section) with sequential novel agents has enhanced treatment efficacy in preclinical data[28]. Phase I trials are underway combining hypo-fractionated liver radiotherapy and Sorafenib, a potential radio-sensitizer. Preliminary outcomes suggest the combination is tolerable for liver metastases cases in particular, and for HCC when the radiation volumes are small[29,30].

Common treatment options for liver cancer, excluding external-beam radiation, are summarized in Table 1.1. The majority of evidence available for loco-regional treatment of metastases is for
colorectal cancer. There is also limited evidence for other primary cancers supporting the hypothesis that improving local control may improve survival for those patients as well.

Table 1.1. Summary of treatment options for liver cancer, excluding external beam radiation.

<table>
<thead>
<tr>
<th>Modality</th>
<th>Potential advantages</th>
<th>Potential disadvantages</th>
<th>Survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resection, transplantation</td>
<td>Curative for early stage disease</td>
<td>&lt;20% of patients are eligible</td>
<td>5 year: 43–80%</td>
</tr>
<tr>
<td>RFA</td>
<td>Potentially curative for small tumors</td>
<td>Ineffective for tumors &gt;4 cm, near large vessels; PVTs are contraindicated</td>
<td>5 year: 18–77%</td>
</tr>
<tr>
<td>PEI</td>
<td>Inexpensive; well tolerated; effective for small tumors</td>
<td>RFA is superior for tumors &gt;2 cm</td>
<td>5 year: 50%</td>
</tr>
<tr>
<td>TACE</td>
<td>Option for intermediate stage disease</td>
<td>PVTs are contraindicated</td>
<td>2 year: 28–63%</td>
</tr>
<tr>
<td>$^{131}$I</td>
<td>Option for multi-focal tumors and PVTs</td>
<td>Ineffective for tumors &gt;6 cm, liver metastases</td>
<td>2 year: 60%</td>
</tr>
<tr>
<td>$^{90}$Y</td>
<td>Option for diffuse disease and PVTs</td>
<td>Heterogeneous radiation dose</td>
<td>Median: 5–17 months</td>
</tr>
<tr>
<td>Sorafenib</td>
<td>Option for advanced stage disease and PVTs; potential radio-sensitizer</td>
<td>Modest survival benefits</td>
<td>Median: 10.7 months</td>
</tr>
</tbody>
</table>

1.1.3 Radiotherapy for liver cancer

*Indications and toxicity*

For patients who have failed or are unsuitable for the previous standard therapies, external-beam radiation therapy is a versatile option. Radiotherapy can be delivered to larger and multi-focal tumors, PVTs, and to both primary and metastatic liver cancers. It has also acted as a bridging therapy to stabilize disease while patients await other options (e.g. transplantation)[31].

Historically, the most common serious toxicity is radiation-induced liver disease (RILD), appearing 2 weeks to 3 months post-radiotherapy. RILD pathology is not precisely known, but it is often described as a veno-occlusive disease and manifests clinically as anicteric hepatomegaly, ascites, and elevated liver enzymes[32]. RILD can resolve spontaneously or be managed with supportive care, or it can progress to liver failure with a 10–20% mortality
rate[33,34]. Toxicities not classically associated with RILD include reactivation of Hepatitis B virus, jaundice, and a decline in liver function[32].

Technological limitations historically resulted in the whole liver being irradiated, leading to rates of RILD of <5% for ≤30 Gy, 10–13% for 30–35 Gy, and up to 44% for >35 Gy (delivered with daily 2 Gy/fraction, or twice daily 1.5 Gy/fraction)[32,35]. For whole liver irradiation the suggested tolerance dose for a 5% risk of RILD at 5 years (TD5/5) is ≤30 Gy for liver metastases and ≤28 Gy for primary cancer, in conventional daily 2 Gy/fraction[32]. These doses may provide palliation but are unlikely to result in disease cure.

Normal tissue complication probability (NTCP) models

The liver’s parallel functioning subunits allow focal damage to be sustained without organ failure. NTCP models for RILD have been developed to understand the liver’s partial volume tolerance to radiation and enable less toxic application of radiotherapy. The most common model, proposed by Lyman, assumes a sigmoid relationship between a dose of uniform partial liver irradiation and NTCP risk[36]. Because radiotherapy delivers non-uniform irradiation, the effective volume (Ve) is often used in the Lyman model. Ve is the liver volume uniformly irradiated to the prescribed dose resulting in an iso-NTCP as the non-uniform irradiation[37].

Increased RILD risk has been associated with treatment factors such as mean liver dose, liver volume receiving >30 Gy, concurrent chemotherapy, and clinical factors such as primary liver cancer, cirrhosis and poor liver function, Hepatitis B, and male patients[38-40]. Using refined Lyman-NTCP model parameters, Dawson et al. demonstrated a strong volume effect where RILD risk was related to mean liver dose and doses >100 Gy could theoretically be delivered to small liver volumes (Ve <30%) for an NTCP <5%[39].

Modern liver radiotherapy

Advances in imaging such as computed tomography (CT), 3D treatment planning, and beam shaping devices on linear accelerators (e.g. multi-leaf collimators) have allowed for 3D conformal dose shaping around tumors and partial sparing of normal liver tissue. Additionally, the University of Michigan pioneered the use of NTCP models to prescribe doses allowing 40–90 Gy delivered in 1.5 Gy/fractions twice-daily to liver tumors[41,42]. Their Phase II trial of conformal radiotherapy with concurrent chemotherapy resulted in a median survival of 15.8
months with <5% RILD[41]. Other series have attempted to minimize normal liver dose, avoid radiation for patients with underlying liver disease, or have based the prescription on the normal liver dose without patient-specific NTCP calculations. Conformal radiotherapy of 25–60 Gy in 1.5–3 Gy/fraction has resulted in median survival of 11–18 months for HCC[43-45], and 14 months for liver metastases[46].

Stereotactic-body radiotherapy (SBRT) aims to deliver potent doses of radiation in typically 10 or less fractions. Hypo-fractionation is potentially advantageous over conventional radiotherapy as it is more convenient for patients and can be radio-biologically more damaging (i.e. ablative). SBRT therefore requires highly conformal tumor doses that rapidly drop off to maximally spare adjacent healthy tissue. Daily imaging for accurate target localization is almost always used. SBRT with doses of 18–55 Gy/1–10 fractions, often to tumors <5 cm, has resulted in median survival rates of 14.5–28.6 months and only rare instances of RILD and liver failure[47-52].

Princess Margaret Hospital (PMH) has approached SBRT with an individualized dose allocation scheme initially building on the Michigan experience. Dose is prescribed using the Lyman NTCP model and liver V_{eff} to maintain a pre-defined risk of RILD[53]. This approach allows for large tumors to be treated. Phase I dose-escalation trials of SBRT delivered doses of 27–60 Gy/6 fractions over 2 weeks, resulting in a median survival of 17.6 and 13.4 months for patients with metastases and primary cancers respectively[54,55]. Phase I and II data combined resulted in a median survival of 17 months for HCC[56]. No dose-limiting toxicities or RILD were observed.

Low rates of other potentially serious SBRT toxicities can include gastrointestinal ulcerations or bleeding, rib fractures, and for primary cancers in particular, a decline in liver function, elevated liver enzymes, and transient biliary obstructions[47,54,55].

Particle-based radiotherapy using protons or carbon-ions is another option offering superior normal-liver sparing, dose conformity, higher relative biological effectiveness, and dose-escalation potential over photon-based radiotherapy[57]. Excellent clinical results have been observed[58], though currently it is far less widely available than conventional radiotherapy delivered with medical linear accelerators.

The wide range of reported doses and fractionations result in uncertainty over the optimal tumor doses and normal tissue dose constraints[59]. There is however a trend for higher doses relating
to improved local control and survival[41,47,52,54], warranting the study of further dose-escalation provided toxicity rates can remain low.

1.1.4 Technical challenges of liver SBRT

*Dose distribution optimization*

Radiotherapy dose distributions are planned and calculated on a single computed tomography (CT) dataset. Individualized liver SBRT plans are often created with 3–10, sometimes non-coplanar, static beams that can be subdivided into 1–4 segments to conform dose to targets and spare normal tissue[60]. The creation and weighting of segments can be manually optimized. In intensity-modulated radiotherapy (IMRT), a series of dosimetric objectives are specified upfront and computer algorithms determine the optimal beam segments (up to 50) and weightings to achieve the desired dose distribution, termed ‘inverse-planning’. Compared to forward-planning, IMRT improves target coverage in 73% of liver SBRT plans and allows for dose-escalation in 35% due to more normal tissue sparing[60]. Mean liver dose can increase with IMRT however when suboptimal objectives are specified, which is subject to observer variability[61]. Incorporating biological parameters (e.g. NTCP and measures of tumor control probability) directly into liver IMRT optimization can allow for further dose-escalation[62]. Volumetric modulated arc therapy (VMAT) is a technique that dynamically delivers IMRT over one or more rotations of the gantry, as opposed to static beams. VMAT reduces liver SBRT delivery time by over 60% and improves target dose conformity compared to IMRT[63].

*Target volumes and margin definitions*

The primary liver SBRT target is the gross tumor volume (GTV), the visible tumor on the planning CT, often enhanced with contrast or a co-registered magnetic resonance imaging (MRI). With intravenous contrast, HCC typically enhances 20–30 seconds post-injection during the arterial phase while metastases enhances 50–60 seconds post-injection during the venous phase. The clinical target volume (CTV) is a 5–8 mm expansion of the GTV within the liver to account for possible microscopic disease. There is no consensus on the optimal CTV expansion. Wang et al. correlated radiological studies with pathology for HCC, finding no microscopic GTV extensions beyond 4 mm, and for solitary small GTVs the maximum extension was 2 mm in 96% of cases[64]. Shrinkage and deformation of the pathology specimens currently limits the accuracy of such studies.
The planning target volume (PTV) is a geometric expansion of the GTV or CTV targets which accounts for the geometric uncertainties in the radiotherapy process. The PTV ensures the primary target receives the prescribed dose during treatment delivery. They are often subdivided into an internal target volume (ITV) which accounts for the size, shape and position of the GTV and CTV due to physiological processes (e.g. breathing), and the setup margin which accounts for daily targeting uncertainties[65]. Therefore PTVs depend on the accuracy of the SBRT technique. Reducing margins is beneficial for liver SBRT as it may allow for higher prescribed doses as more normal tissues are spared.

Uncertainties can result from target delineation errors, organ motion, daily positioning and targeting errors, and machine-related errors. Geometric errors for a single patient can be separated into systematic and random components, computed as the mean and standard deviation of the errors respectively[66]. Each type is typically defined in the left-right (LR), anterior-posterior (AP), and superior-inferior (SI) anatomical planes, or summarized as a single 3D vector length. Figure 1.1 demonstrates how the dosimetric impacts of systematic and random errors fundamentally differ. PTV margin ‘recipes’ have been proposed based on the computation of the standard deviation of random (σ) and systematic (Σ) errors measured in a previous population of patients. Van Herk et al. derived a recipe for conventional prostate radiotherapy based on the probability of desired target dosage[67]. For example, a CTV expansion equal to 2.5Σ + 0.7σ was derived, under simplified assumptions, to result in 90% of patients receiving 95% of the prescribed dose[67]. Applying these dose-probability PTVs proposed by van Herk to liver SBRT may not account for the hypo-fractionation, steeper SBRT dose gradients, often multiple targets, and respiratory motion.

Figure 1.1. The impact of geometric radiotherapy errors on the dose distribution (coloured lines). Figure adapted from [67,68].
Breathing motion management

The mean SI liver amplitude ranges from 11–25 mm during normal respiration[69]. Substantial inter-patient variations exist, as do intra-patient variations due to tumor location[70]. Patient-specific measurements are therefore required for radiotherapy planning. Breathing motion can be quantified using 2D fluoroscopy of either the diaphragm or fiducial markers implanted near the tumor[71]. Respiratory-correlated CT, or ‘4D CT’, is reconstructed from an over-sampled helical CT where the data is retrospectively binned into phases of the respiratory cycle acquired with an external respiratory signal[72]. The resulting CT series show full 3D motion of the liver and internal anatomy across typically 10 respiratory phases. Automated and synchronized contrast-injection is ideally used during 4D CT because liver tumors otherwise partially clear the contrast agents during image acquisition[73]. MRI has superior soft-tissue contrast over CT and when coupled with high-speed acquisition sequences called ‘cine-MRI’, direct tumor motion can be measured in 2D. The uncertainty with imaging surrogates needs to be considered as well. Kirilova et al. showed that fluoroscopy of the diaphragm underestimated tumor motion seen on cine-MRI in half of liver patients[74]. Measuring tumor motion in 2D with cine-MRI may be inaccurate if out-of-plane motion is present. Temporally-resolved 3D images are possible with MRI allowing for 4D acquisition. The current status of 4D MRI is its lower spatial resolution and image quality compared to 4D CT, although it can capture more breathing cycles than 4D CT and does not rely on an external respiratory signal[75].

There is no consensus on how to incorporate breathing motion into the PTV. A popular method creates a union of the tumor’s delineated positions on 4D imaging (i.e. an ITV) followed by a linear expansion for the setup margin. Effectively this aims for complete dosimetric coverage of the tumor over the entire breathing cycle. Alternatively, patient-specific breathing motion can be divided into its systematic and random components and incorporated into a margin recipe (e.g. van Herk’s). The ITV method generally results in a larger PTV[76]. Stroom et al. demonstrated that because breathing motion is uncorrelated with setup error it is not necessary to add the two linearly as is the case with the ITV method[77]. For either strategy, elimination or reduction of breathing motion is desirable to avoid prohibitively large PTVs. Methods to manage breathing motion are described below:

Breath-hold techniques: Radiation can be administered under voluntarily breath-hold or with an active breathing control (ABC) device. For liver SBRT, the caregiver suspends the patient’s
breathing with ABC at the exhale phase for 15–30 seconds. The average maximal SI drift in this position is 1.4 mm[78]. The intra-fraction SI reproducibility (σ) of the diaphragm relative to vertebral bodies is 1.5 mm on average (range: 0.6–3.9 mm)[78]. This stability results in a smaller required PTV than for non-breath-hold SBRT. A study by Ten Haken et al. estimated that eliminating PTV portion required for breathing motion would result in a 6–7% increase in tumor control probability with an iso-NTCP dose allocation[79]. ABC breath-hold is the preferred method of treatment at PMH, however approximately 40% of patients are deemed unsuitable (e.g. not reproducible, intolerant of the device, language barriers, etc.))[78].

Gating and tracking: The synchronization of radiation delivery to a specific or narrow range of breathing phases is termed ‘gating’. Gating is often applied to the exhale phase as more time is spent in exhale and it is more stable than inhale[80,81]. Wagman et al. reduced the 90th percentile of SI breathing amplitude (mean±standard deviation) from 23±7 mm to 5±2 mm with gating[82]. This halved the required PTV on average enabling an iso-NTCP dose escalation of 7–27% (median: 21%). With ‘tracking’, the radiation beam dynamically adjusts to follow the tumor motion rather than being triggered on in certain phases as occurs with gating. Real time imaging of the tumor is ideally done via stereoscopic x-ray imaging of implanted markers, as large differences compared to external surrogates (e.g. infrared-tracked markers) have been observed[83,84]. For both methods, near continuous real-time imaging may be necessary to capture irregular breathing patterns[85].

Abdominal compression and free-breathing: Abdominal compression plates can force shallower breathing in liver SBRT patients and limit breathing motion. Wunderink et al. reduced the median SI and AP motion with compression by 62% and 38% respectively compared to free-breathing, measured with fluoroscopy of implanted fiducials[86]. Motion was reduced to <5 mm in all directions for 83% of patients, and was stable within 2 mm during the treatment course. Eccles et al. however showed that tumor motion on cine-MRI increases by ≥3 mm in 8% of patients under compression, and mean tumor motion was reduced by only 2.3 mm and 0.6 mm in the SI and AP directions respectively[87]. Higher compression magnitudes correlate with smaller breathing amplitudes[88], at a potential cost of greater patient discomfort. In general, compression can therefore reduce the required PTV margin compared to free-breathing. Approximately 10% of patients are unsuitable for either ABC or compression and must be treated in free-breathing.
Image-guided radiotherapy (IGRT) modalities

Modern SBRT is often coupled with IGRT devices to target the tumor or surrogate prior to each fraction. Reproducing the patient’s planned geometry daily helps ensure the planned dose is delivered to the patient. At the treatment linear accelerator, orthogonal 2D kilovoltage (kV) or megavoltage (MV) images can be used to verify the patient’s position relative to planning and kV fluoroscopy can quantify the breathing motion. These 2D modalities are only useful for visualizing anatomy with high contrast (e.g. vertebral bodies, lung-liver interfaces) or implanted fiducials. Ultrasound is a non-ionizing alternative for 2D respiratory tracking and gating for liver radiotherapy[89]. CT-on-rails systems have been integrated into treatment rooms with a common couch that can be repositioned between the CT and linear accelerator allowing diagnostic quality imaging for IGRT[90]. Cone-beam CT (CBCT) can also be reconstructed with the patient in the treatment position using kV sources mounted to the treatment gantry or directly using the MV treatment beams[91,92]. CBCT is acquired over 2–4 min causing breathing-induced motion artifacts that manifest as blurring. Respiratory-correlated CBCT, or ‘4D CBCT’, is advancement which retrospectively extracts a breathing signal directly from the raw image data, and then bins the data to reconstruct time-resolved breathing phases, analogous to 4D CT[93]. 4D CBCT allows for specific breathing phases (e.g. exhale) to be aligned between planning and treatment, removing the uncertainty caused by the blurring[94,95].

Significant differences in liver localization have been observed between 2D and 3D imaging. Hawkins et al. compared the 3D liver position on kV CBCT following guidance based on orthogonal MV imaging of the diaphragm and vertebral bodies for breath-hold liver SBRT[96]. In 33% of CBCTs the residual error in liver position was >5 mm in any direction compared to MV imaging, up to a maximum of 12 mm. Although MV imaging was better than no guidance at all, this study demonstrated the benefit of CBCT-based IGRT.

1.1.5 Geometric accuracy of image-guided liver SBRT

Day-to-day or inter-fraction changes can occur relative to the initial planning CT. Intra-fraction changes occurring during irradiation delivery relative to the start of the fraction can be caused by respiratory motion or other changes in patient position. At PMH the clinical goal is to reduce the geometric treatment uncertainties to less than 3 mm. A summary of studies evaluating the accuracy of liver radiotherapy is shown in Table 1.2, and demonstrates two important trends.
First, IGRT is required to ensure large errors from the initial setup based on tattoo and laser alignment are corrected. Second, fiducial markers or soft-tissue targeting is required as IGRT based on bony anatomy is inadequate because baseline shifts in liver position occur relative to bone.

Inter-fraction setup errors can be identified and corrected with couch translations based on daily pre-treatment IGRT. Baseline shifts in liver position relative to vertebral bodies can be a large source of error in liver SBRT, providing a rationale for soft-tissue image guidance. With ABC and IGRT based on the liver, systematic (random) errors are reduced from 5.1 (6.5), 3.1 (4.7) and 3.4 (4.2) mm to 1.3 (2.2), 1.8 (2.7) and 1.6 (2.3) mm, in the SI, AP and RL directions respectively[96]. With free-breathing and abdominal compression, residual systematic (random) errors in exhale liver position relative to bony anatomy on 4D CBCT are 2.9 (3.1), 1.7 (2.4) and 1.5 (1.8) mm in the SI, AP and RL directions respectively[95]. For non-ABC patients, mean absolute inter-fraction changes in breathing amplitude are ≤1.7 mm in all directions[94].

Intra-fraction errors can be detected with 2D fluoroscopy, infrared-tracked external markers, or repeat 3D imaging. For treatment in free-breathing without gating, intra-fraction 3D breathing motion can be estimated with a pre-treatment 4D CBCT. Intra-fraction changes in amplitude were calculated by Case et al. by comparing pre- and post-treatment 4D CBCT over a mean delivery time of 12 min (range: 4–25 min)[94]. Random and systematic amplitude changes were ≤1.9 mm, although one patient in severe pain had a 6.8 mm increase in amplitude during one fraction. Random and systematic intra-fraction baseline liver shifts at the exhale phase were similarly measured to be ≤1.9 mm, significantly smaller than the inter-fraction error [95].

Another source of error is tissue deformation which includes changes in shape or volume. Eccles et al. determined that 22% of the liver volumes deforms by >3 mm on average between CBCT and the planning CT[97]. Romero et al. estimated the daily delivered dose using daily repeat CT for liver SBRT of 37.5 Gy/3 fractions[68]. The mean loss in target volume dose coverage was reduced from 6.8% to 1.7% with IGRT. However, dose deviations in the adjacent gastrointestinal normal tissues ranged from -22.9 to 11.2 Gy resulting from abdominal deformation that could not be corrected with translational couch shifts alone.
Table 1.2. Studies reporting the accuracy of image-guided liver radiotherapy based on various imaging modalities and tumor surrogates.

<table>
<thead>
<tr>
<th>Reference; No. patients; imaging (surrogate)</th>
<th>Comparison(s): statistic(s)</th>
<th>Inter-fraction error (mm)</th>
<th>Intra-fraction error (mm)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dawson 2001[98]; N=8, ABC; kV fluoroscopy (fiducials, diaphragm)</td>
<td>Pre-IGRT diaphragm vs. bones: mean</td>
<td>4.4</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>Pre-IGRT fiducials vs. bones: mean</td>
<td>4.3</td>
<td>2.3</td>
</tr>
<tr>
<td>Balter 2002[99]; N=8, ABC; kV fluoroscopy (diaphragm)</td>
<td>Pre-IGRT: mean</td>
<td>6.7</td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td>Post-IGRT: mean</td>
<td>3.5</td>
<td>2.3</td>
</tr>
<tr>
<td>Dawson 2005[100]; N=20, ABC; orthogonal/cine MV (diaphragm)</td>
<td>Pre-IGRT: $\Sigma (\sigma)$</td>
<td>5.1 (6.5)</td>
<td>3.1 (4.7)</td>
</tr>
<tr>
<td></td>
<td>Post-IGRT: $\Sigma (\sigma)$</td>
<td>1.4 (2.5)</td>
<td>1.9 (2.9)</td>
</tr>
<tr>
<td></td>
<td>$\Delta$ diaphragm motion: mean (SD)</td>
<td>0.6 (1.4)</td>
<td></td>
</tr>
<tr>
<td>Hawkins 2006[96]; N=13, ABC; orthogonal kV/MV (diaphragm), CBCT (liver)</td>
<td>Pre-IGRT diaphragm: $\Sigma (\sigma)$</td>
<td>5.1 (6.5)</td>
<td>3.1 (4.7)</td>
</tr>
<tr>
<td></td>
<td>Post-IGRT diaphragm: $\Sigma (\sigma)$</td>
<td>1.3 (2.2)</td>
<td>1.8 (2.7)</td>
</tr>
<tr>
<td></td>
<td>Post-IGRT liver: $\Sigma (\sigma)$</td>
<td>1.1 (2.7)</td>
<td>1.9 (3.0)</td>
</tr>
<tr>
<td>Guckenberger 2008[101]; N=11, free-breathing; CBCT (liver, diaphragm)</td>
<td>Pre-IGRT liver: $\Sigma (\sigma)$</td>
<td>4.3 (6.4)</td>
<td>4.0 (4.3)</td>
</tr>
<tr>
<td></td>
<td>Pre-IGRT liver vs. bones: $\Sigma (\sigma)$</td>
<td>2.6 (4.2)</td>
<td>3.2 (1.8)</td>
</tr>
<tr>
<td></td>
<td>$\Delta$ diaphragm motion: mean (SD)</td>
<td>1.7 (2.4)</td>
<td>1.1 (3.0)</td>
</tr>
</tbody>
</table>

Continued on next page
Continued from previous page

<table>
<thead>
<tr>
<th>Reference; No. patients; imaging (surrogate)</th>
<th>Comparison(s): statistic(s)</th>
<th>Inter-fraction error (mm)</th>
<th>Intra-fraction error (mm)*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SI</td>
<td>AP</td>
</tr>
<tr>
<td>Wunderink 2010[102]; N=11, compression; kV fluoroscopy (fiducials, diaphragm)</td>
<td>Pre-IGRT fiducials: Σ (σ)</td>
<td>4.2 (2.8)</td>
<td>3.0 (2.0)</td>
</tr>
<tr>
<td></td>
<td>Pre-IGRT diaphragm: Σ (σ)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brock 2008[103]; N=13, ABC; CBCT (tumor via DIR)</td>
<td>Post-IGRT tumor: mean (SD)</td>
<td>1.0 (1.7)</td>
<td>1.0 (1.1)</td>
</tr>
<tr>
<td>Case 2009[95]; N=29, free-breathing/compression; 4D CBCT (exhale liver)</td>
<td>Post-IGRT liver vs. bones: Σ (σ)</td>
<td>2.9 (3.1)</td>
<td>1.7 (2.4)</td>
</tr>
<tr>
<td>Case 2009[94]; N=29, free-breathing/compression; 4D CBCT (exhale liver)</td>
<td>Post-IGRT liver amplitude: Σ (σ)</td>
<td>1.6 (1.5)</td>
<td>1.1 (1.8)</td>
</tr>
</tbody>
</table>

**Abbreviations and symbols:** SI=superior-inferior; AP=anterior-posterior; RL=right-left; SD=standard deviation; N=number (of patients); CBCT=cone-beam CT; DIR=deformable image registration; ∆=change; Σ=SD of population systematic errors; σ= SD of population random errors. Notes: *all intra-fraction errors are post-IGRT measurements.
The PTV margin applied to the GTV on the static planning CT ensures it receives the prescribed dose in the face of these geometric errors. IGRT then allows setup errors, including some caused by deformation, to be approximately corrected about a single point by applying rigid translations to the patient. The PTV concept is robust in that it ensures the GTV dose deviations are minimized at a clear cost: normal tissue irradiation. An additional complication is that the PTV does not account for normal tissue motion. The doses to normal tissues may not be accurately accounted for on the static planning CT because motion (e.g. due to setup error, or respiration) is moving these tissues in and out of the PTV and high dose region. This has prompted the development of tools to better incorporate geometric uncertainties into dose estimates.

1.1.6 Modeling geometric uncertainties into dose calculations

Dose accumulation methods attempt to cumulatively account for geometric uncertainties over a period of time such as over the breathing cycle (intra-fraction), or entire treatment course (inter-fraction). Table 1.3 summarizes published dose accumulation studies for liver radiotherapy.

*Intra-fraction breathing dose accumulation*

Dose convolution is a simple approach to model breathing motion into dose distributions. The static dose distribution calculated on one CT image is convolved according to a probability distribution function (PDF) characterizing the likelihood of different positions in the breathing cycle from occurring[104]. Lujan et al. used a one dimensional (1D) PDF, based on one patient’s liver fluoroscopy motion, demonstrating point dose differences up to 26% between the static distribution and convolved distributions that approximate the breathing motion[80]. Simulating amplitudes changes >5 mm resulted in greater dosimetric changes compared to changes of 3 mm or variations in the breathing pattern (i.e. time spent in each breathing phase). The inclusion of the latter only minimally effects the breathing motion model[105]. Convolution of the beams’ fluences, instead of the dose distribution, is a useful extension that more accurately accounts for spatial heterogeneity differences in densities, such as the lung-liver interface[106]. Discrepancies between fluence- and dose-convolution are ±2% on average[106].

A number of studies have shown a trend for dose distributions planned on exhale CT. After modeling the breathing motion, normal tissues superior to the PTV increase in dose as they move inferior towards the inhale position and into the beam, while inferior tissues decrease in
dose as they move out of the beam[107-109]. The effect on $V_{\text{eff}}$ and liver NTCP is highly variable depending on the location of target within the liver[106,107]. Assuming the PTV fully encompasses the breathing motion and is irradiated with a fairly uniform dose, the GTV dose should not change with this motion. It should be noted that SBRT doses to the PTV are rarely uniform.

A complication of modeling breathing motion with IMRT plans termed ‘interplay’ has been observed. Interplay is the combined effect of variable breathing motion relative to the ‘motion’ of beam segments during IMRT delivery. Bortfeld et al. reported that interplay caused maximum IMRT point dose variations $<$2% over 30 fractions, but were up to 18% for 5 fractions[104]. The predominant effect of breathing motion however was blurring of the IMRT dose. Kuo et al. studied interplay in 8 liver IMRT patients treated over 25 fractions, reporting that changes in CTV were $\leq$2.5% for 1.8 cm of motion, versus $<$1% for under 1 cm of motion[110]. Interplay effects have not yet been quantified for hypo-fractionated liver SBRT. However, using simple IMRT with fewer segments can achieve similar dose conformality versus more complex IMRT[111], potentially reducing the impact of interplay effects.

Inter-fraction dose accumulation

For liver radiotherapy, modeling inter-fraction errors into the dose distributions has been studied less than breathing motion. Rosu et al. used dose-convolution to model both the breathing dose (i.e. intra-fraction) based on a 1D PDF, and also the delivered treatment dose based on a Gaussian distribution of random inter-fraction setup errors in 43 liver patients[108]. Relative to the planned NTCP on the static CT, the overall delivered NTCP changes ranged from -12% to 12%. The contribution from breathing motion (range: -11% to 15%) was larger than the random setup errors (range: -6% to 2%). Balter et al. compared modeling either population or patient-specific setup errors for 8 liver patients[112]. The delivered liver $V_{\text{eff}}$ decreased by up to 2% and 4% with the population and patient-specific models respectively. Unfortunately, patient-specific random errors would not be known upfront at planning but could be quantified retrospectively to improve correlation of outcomes (e.g. improve liver NTCP models). Both of these studies assumed systematic setup errors to be nil with the use of IGRT, however systematic errors have been shown in other disease sites to have greater impact on tumor control than either random errors or breathing motion[113].
Table 1.3. Summary of intra- and inter-fraction dose accumulation studies for liver radiotherapy.

<table>
<thead>
<tr>
<th>Reference; No. of patients, plan modality</th>
<th>Intra-/Inter-fraction accumulation: methods</th>
<th>Major findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lujan 1999[80]; N=1, conformal</td>
<td>Intra-fraction: dose convolution + 1D breathing PDF</td>
<td>Max point dose difference 26% vs. static dose</td>
</tr>
<tr>
<td>Lujan 2003[105]; N=1, conformal</td>
<td>Intra-fraction: dose convolution + 1D breathing PDF</td>
<td>Effect of breathing motion &gt; amplitude/pattern changes</td>
</tr>
<tr>
<td>Balter 2005[112]; N=8, conformal</td>
<td>Inter-fraction: dose convolution + 3D population/patient-specific σ</td>
<td>Effect of patient-specific σ &gt; population σ</td>
</tr>
<tr>
<td>Chetty 2003[106]; N=1, conformal</td>
<td>Intra-fraction: fluence convolution + Monte Carlo/dose convolution + 1D breathing PDF</td>
<td>Fluence convolution ≈ dose convolution (±2%)</td>
</tr>
<tr>
<td>Rosu 2003[108]; N=40, conformal</td>
<td>Intra-fraction: dose convolution + 1D breathing PDF</td>
<td>Effect of breathing motion &gt; random setup errors (σ)</td>
</tr>
<tr>
<td></td>
<td>Inter-fraction: dose convolution + 3D population σ</td>
<td></td>
</tr>
<tr>
<td>Kuo 2007[110]; N=8, IMRT</td>
<td>Intra-fraction: fluence convolution + 1D breathing PDF</td>
<td>Breathing motion blurs IMRT dose</td>
</tr>
<tr>
<td>Coolens 2005[114]; N=1, IMRT</td>
<td>Intra-fraction: B-spline DIR of 4D CT</td>
<td>Max target under-dosage 30%</td>
</tr>
<tr>
<td>Brock 2003[107]; N=4, conformal</td>
<td>Intra-fraction: linear elastic DIR/ rigid registration of 4D CT</td>
<td>Rigid-based accumulation ≠ deformable accumulation</td>
</tr>
<tr>
<td>Wu 2008[115]; N=5, conformal/ arcs/IMRT</td>
<td>Intra-fraction: thin-plate spline DIR of 4D CT</td>
<td>Effect of breathing motion &gt; for IMRT/larger motion</td>
</tr>
<tr>
<td>Velec 2011[109]; N=21, conformal RT/IMRT</td>
<td>Intra-fraction: biomechanical DIR of 4D CT</td>
<td>Max normal tissues dose difference 25% vs. static dose</td>
</tr>
</tbody>
</table>

Abbreviations: DIR=deformable image registration; IMRT=intensity-modulated radiotherapy; PDF=probability distribution function.
1.1.7 Deformable image registration

Image registration is the geometric alignment of two or more images of the same anatomy. It is largely limited in the clinic setting to rigid global transformations, allowing for only 6 degrees of freedom (i.e. translations and rotations). Deformable image registration (DIR) increases the degrees of freedom by means of nonlinear transformations between images. This allows for a voxel-to-voxel mapping between two image sets, limited only by the accuracy of the algorithm to identify this correspondence. Therefore, DIR can account for the inevitable patient motion and complex deformation that limits the accuracy of rigid registration.

A multi-institution study evaluated the accuracy of DIR algorithms in the liver on common patient datasets: exhale to inhale 4D CT, and CT to MRI[116]. Seven of eighteen DIR algorithms reported liver 4D CT mean absolute registration errors <2.5 mm, similar to the image slice thickness. Only three reported results for the liver CT and MRI (mean absolute error ranged from 1–5 mm), identifying a challenge with multi-modality DIR.

Registration algorithms differ primarily by how they measure similarity (or dissimilarity) between images, and by how they model the non-linear transformations. Optimization algorithms work within each DIR loop between altering the registration transformation and then recalculating the similarity.

Similarity metrics

DIR algorithms require a means to evaluate objectively how well the images are aligned. They can be broadly classified as intensity-based, relying directly on the voxel intensities, or geometry-based, relying on extracted features common in both images (e.g. organ contours). The former is more commonly encountered.

A simple and computationally efficient intensity-based similarity metric is the sum of the squared differences (SSD). SSD is analogous to an intensity difference image plot with the aim to minimize any differences. It is only applicable to mono-modality registration (e.g. CT to CT) where identical intensities exist and any differences are assumed to be caused by image noise. With contrast administration SSD cannot be entirely minimized. SSD could potentially perform poorly with large intensity differences in even a small number of voxels[117].
Where a linear intensity relationship does not exist (e.g. CT and MR), mutual information (MI) can be used as a metric. Rather than using the direct intensity values, MI describes how the information content in one image describes the information content in another image. Entropy is a measure of information content and relates the probability of an intensity value being in the images both jointly and independently[118]. MI is maximized when the images are registered. MI is considered the optimal metric for multi-modality registration however it is more computationally complex than SSD.

**Transformation and regularization models**

A well-behaved DIR uses a transformation model that realistically approximates the true anatomical deformation. In many scenarios a regularization term is additionally applied to smooth the deformation map. Parametric transformation models align a subset of corresponding control points while the displacements at the remaining voxels are given by the interpolation of basis-functions about the control points[119]. One class, termed ‘splines’, are analogous to a bending metal strip about fixed knots but are not physically based. For thin-plate splines, control points are often irregularly spaced around an organ contour each having a global effect on the deformation map[120]. For basis- or B-splines, control points are regularly spaced throughout the image and the basis-functions act locally allowing for more complex deformations to be modeled versus thin-plate splines[121]. For regularization of splines a term may be added to the similarity metric (e.g. minimizing the total bending energy of the system[119]) to avoid unrealistic deformations. An example of a non-parametric transformation model is called optical flow, or the ‘Demons’ algorithm. These model deformation as a physical diffusion process based on the displacement of intensity gradients in the secondary image being pushed by local ‘demons’ to match the primary image at each voxel[122]. For low gradient regions, the addition of bi-directional forces can improve efficiency and convergence[123]. To avoid non-continuous singularities in the deformation map caused for example by image noise, a regularization term (e.g. Gaussian filter) can also be applied[123]. A potential disadvantage of B-spline and Demon’s algorithms are they are not biophysical models of deformation.

**Biomechanical model-based DIR**

A novel DIR algorithm developed in-house, termed ‘Morfeus’, performs registration by modeling the biomechanical motion and changes of tissues between images. Surfaces of organs
(e.g. liver) delineated on each image are aligned by guided-surface projections, acting as similarity metric for the transformation model. This classifies Morfeus as a geometry-based algorithm. The internal tissues away from the surfaces (e.g. including the GTV) are deformed according to the modeled biomechanical material properties. Multi-organ registration with Morfeus has been validated with an accuracy of ≤2 mm[124], within the typical slice thickness of images used in liver SBRT. Morfeus is described in detail in the following Chapter.

An advantage of biomechanical DIR is that multi-modality images can be registered without establishing a relationship between image intensities. It can also potentially model complex anatomical motion, such as the sliding of the organs against lung in the chest wall[125]. Potential disadvantages include the requirements for contours and their associated uncertainties, and uncertainties in the assigned material properties.

### 1.1.8 DIR validation methods

A challenge of DIR validation in clinical images is that the underlying ground truth is often only sparsely known. A common approach compares the residual geometric positions of features such as anatomic landmarks or fiducial markers between the DIR-predicted image and the actual secondary image (Figure 1.2). The availability of these point-based measures is often limited, particularly in low contrast images and anatomic sites such as the liver. Contours can also be compared by means of their residual centroid positions, surface distance-to-agreement (DTA), or overall volumes via overlap measures (e.g. Dice similarity coefficient [DSC]). Overlap measures do not inform of the accuracy inside the contoured region however. Both point and contour-based measurements are limited by the accuracy and precision by which they can be identified.

Physical phantoms have been constructed and imaged before and after applying known physical deformations, providing a ground truth for validation[126]. Similarly, mathematical phantoms have been constructed by applying known deformations to clinical images to generate an artificial secondary image[127]. DIR is then performed and the predicted displacements can be compared to the actual displacements applied on a voxel-by-voxel basis. A challenge for both phantoms is to realistically mimic the complexity of motion and deformation seen clinically.
1.1.9 Applications of DIR in liver radiotherapy

*Tumor delineation*

Liver deformation between multi-modality imaging sessions confound the ability to account for the GTV’s size and shape during delineation. These uncertainties manifest as systematic errors in liver SBRT planning. Voroney et al. compared GTVs between MRI and CT in 26 liver SBRT patients\[128\]. To correct for the deformation between images, a geometrically resolved CT was created using Morfeus. Compared to rigid registration of vertebral bodies with MI, Morfeus improved the median GTV concordance volume between CT and MRI from 74% (range 15–86%) to 78% (range 26–88%). The geometrically resolved view demonstrated median GTV surface area differences >3 mm of 55% (range 28–93%) between CT and MR. These differences in GTV representation (i.e. due to deformation, or inherent CT-MRI differences) are potentially large enough to alter SBRT planning and dose-allocation based on liver NTCP.

Intensity-based DIR has also been investigated to improve liver GTV delineation. Two groups have applied optical flow-based DIR to register all 4D CT phases to a common dataset (e.g. exhale phase, or time-averaged position). Combining the partial contrast from all phases typically reduced the image noise by a factor of three\[129,130\], resulting in more easily identified GTVs and potentially reducing delineation uncertainties.
**DIR-based image-guidance**

Brock et al. investigated the potential role of DIR on GTV localization for 12 liver SBRT patients treated with ABC, following CBCT localization of vertebral bodies and fluoroscopic localization of the diaphragm[103]. The breath-hold CBCT liver position was retrospective reconstructed. Morfeus then deformed the liver from the planning CT to each CBCT, modeling the residual GTV position not appreciated on the low-contrast CBCTs. One third of patients had ≥1 fractions with residual GTV displacements exceeding the clinical IGRT tolerance of 3 mm. Maximum displacements of 3.4, 6.5 and 9.7 mm occurred in the LR, AP, and SI directions respectively. These residual errors may cause significant GTV under dosage for hypofractionated liver SBRT, indicating a potential role for DIR-based IGRT.

**Deformable dose accumulation**

Central to this thesis is the inherent ability of DIR to track deforming tissues across imaging studies making it ideally suited for dose accumulation. A potential limitation of the convolution-based methods previously described is they effectively use rigid-body motion models. However, tissues deform during radiotherapy causing deviations in the dose received by those tissues relative to their dose at planning. DIR can provide a voxel-to-voxel mapping between a common reference geometry (e.g. the planning CT) and the subsequent images of deformed tissues. Resolving the geometric uncertainty caused by deformation then allows for doses to be accumulated over multiple instances and evaluated (Figure 1.3).

![Figure 1.3. Deformation from the nominal planning geometry causes uncertainty in the intended dose distribution (coloured lines), which can be resolved with DIR.](image-url)
The majority of published studies have focused on breathing (i.e. intra-fraction) dose accumulation. Generally the dose distribution is independently calculated on each 4D CT phase, DIR tracks each voxel and the dose received at each phase, and finally these doses are warped back to the initial planning CT phase for accumulation and evaluation. It would be additionally possible with deformable dose accumulation to include changes in tissue density for organs such as the lung, occurring during respiration.

The impact of modeling breathing motion on dose distributions using deformable models was first reported by Brock et al. on liver patients[107]. Biomechanical DIR of 4 patient’s livers between exhale and inhale CT was performed, and the dose to normal liver tissue was accumulated from simulated plans and tumor locations. From a planned iso-NTCP level of 15% calculated on the static exhale CT, the mean changes to allowable prescribed doses to tumors in the superior, middle and inferior liver regions were 4.0, -3.6 and -14.5 Gy respectively (range: -27.0 to 5.3 Gy). Differences in breathing dose accumulation between rigid-body and deformable models ranged from -2.4 to 7.3 Gy, highlighting the importance of including deformation into dose calculations. This study used simulated tumors and plans, omitting the complexities seen in clinical plans where strict dose constraints must be respected for all normal tissues.

Wu et al. used thin-plate splines and MI for 5 liver SBRT patients to accumulate the breathing dose over all 4D CT phases[115]. Plans using conformal static beams, conformal arcs, and IMRT were compared. Dose changes to 99% of the CTV between the DIR-predicted breathing doses relative to the static dose calculations ranged from -0.2 to -2.2% for the conformal plan, -0.3 to -4.7% for conformal arcs, and -1.0 to 9.3% for IMRT plans. Changes to non-liver normal tissues were negligible in this study, possibly a result of all patients’ tumors being centrally located in the liver avoiding any PTV-gastrointestinal organ overlap. This study highlighted that IMRT dose distributions are more prone to accumulated dose deviations than conformal planning.

1.2 Rationale for Study

Summarizing the clinical data, there is substantial room to improve the clinical outcomes following the current treatments for liver cancers. Although SBRT it is a versatile treatment option, the technical details play an important role in how dose is allocated in clinic. Reducing geometric uncertainties would permit the use of smaller PTV margins, and potentially allow for
higher dose allocation or improved normal-tissue sparing. Furthermore, the combination of a
dose-response effect for liver tumors and the dose-limiting effects of the planned dose to the
surrounding normal tissues requires research into better understanding what the delivered doses
are. This may improve our current understanding of dose-effect relationships.

A major assumption of studies incorporating breathing motion into the planning dose
distributions with DIR and 4D CT is that they more accurately estimate the true doses delivered.
This is based on the work showing breathing motion to be an important factor in liver SBRT
planning. However, whether these breathing dose distributions improve estimates of the true
delivered dose better than the current static doses used clinically is not known. In practice,
IGRT is used to target tumor surrogates with the intention of ensuring what is planned is
delivered to the patient. An opportunity thus exists to extend DIR to the images acquired during
IGRT. This would allow for direct tumor targeting, and incorporating the information from
these images to improve estimates of the delivered dose, termed ‘dose reconstruction’. Understanding
the geometric and dosimetric uncertainties in both planning and delivery
techniques will allow better SBRT to be designed to optimally treat liver tumors.

1.3 Hypothesis and Specific Aims

The general hypothesis of this thesis is that DIR can improve the design and application of
radiotherapy for liver cancer through reconstruction of the delivered dose, improved tumor
targeting and dose allocation. This hypothesis was rigorously tested through three Specific
Aims.

Specific Aim 1: Evaluate the impact of DIR and deformable dose reconstruction to reduce
uncertainties in the delivered doses for liver SBRT.

Aim 1.1: Apply DIR to track tissue and accumulate dose both at planning using 4D CT, and at
each treatment using 4D CBCT. Determine if the 4D CT-predicted breathing dose is a better
estimate of the reconstructed delivered dose compared to the clinical static dose that does not
model breathing motion. Quantify the relative impact of geometric uncertainties occurring in
SBRT delivery on the reconstructed delivered dose relative to the 4D CT-predicted breathing
dose.
Aim 1.2: Retrospectively resolve tumor targeting errors observed with the clinical IGRT strategy, based on rigid liver localization of non-respiratory correlated CBCT. Investigate IGRT strategies using either a simple rigid liver registration of 4D CBCT to account for gross liver-to-liver targeting errors, or directly targeting the tumor on the non-contrast 4D CBCT using DIR. Evaluate the geometric and dosimetric impacts of these two techniques by modeling the delivered dose with deformable dose reconstruction.

**Specific Aim 2: Validate the accuracy of biomechanical-based deformable dose reconstruction.**

Aim 2.1: Predict the delivered dose in an irradiated deformable dosimeter using biomechanical DIR of low-contrast CT images. Quantify the sensitivity of the DIR model’s boundary conditions to differences in the predicted displacements throughout the dosimeter. Evaluate the accuracy of the DIR-predicted dose against the experimental measurements of the delivered dose from optical CT data.

**Specific Aim 3: Optimize the therapeutic ratio in liver SBRT to permit dose escalation while minimizing the toxicity risk to normal tissues.**

Aim 3.1: Develop a liver SBRT planning technique using the mean respiratory position of the liver that enables the application of dose-probability margins to reduce the irradiation of normal tissue while ensuring tumor coverage. Quantify the improved normal tissue sparing in liver SBRT with dose-probability PTV versus those based on the tumor’s ITV. Compare the uncertainties in establishing the mean respiratory position at planning using simple rigid liver registration of 4D CT relative to that predicted using DIR.

Aim 3.2: Develop iso-toxic SBRT plans using ITV-based PTVs at the exhale liver position, and dose-probability PTVs the mean liver position on 4D CT. Model IGRT based on rigid liver alignment at the exhale and mean liver position on 4D CBCT for these strategies respectively, and then estimate the delivered dose with deformable dose reconstruction. Evaluate the robustness of the two strategies to deliver the planned dose to liver tumors.

**1.4 Outline of Thesis**

The current chapter, Chapter 1, has summarized the clinical outcomes for liver cancer and current state and challenges of liver SBRT. DIR and its applications to liver radiotherapy were introduced leading to the Hypothesis and Aims for this thesis. Chapter 2 details the framework
of Morfeus DIR for dose reconstruction, summarizes the validation studies done to date and a previous investigation of breathing dose accumulation in liver SBRT. The original work beginning in Chapter 3 addresses Aim 1.1 by performing a comprehensive analysis of the delivered SBRT doses from previously treated patients using deformable dose reconstruction. Chapter 4 then investigates the potential geometric and dosimetric improvement in using DIR for direct image-guidance of liver tumors (Aim 1.2). Chapter 5 seeks to validate the accuracy of the biomechanical-based deformable dose reconstruction used throughout the thesis, using data from a deformable dosimeter (Aim 2.1). Returning to clinical data in Chapters 6 and 7 (Aims 3.1 and 3.2 respectively), the development and evaluation of a new liver SBRT strategy is proposed that enables iso-toxic, dose-escalated treatment plans without increasing the risk of under-dosing tumors. Finally, a general discussion of the thesis is presented in Chapter 8, followed by a discussion of future opportunities and an overall conclusion.
Chapter 2

MORFEUS: Framework and Prior Investigations
2 Morfeus: Framework and Prior Investigations

2.1 Theory

Biomechanical-model based DIR algorithms rely on a numerical approach termed the finite element method (FEM). To solve the deformation of a complex structure such as the liver, the deforming tissue is first discretized into a finite number of many much smaller pieces called elements. Tetrahedral elements comprised of 4 connected nodes at each vertex are used, and adjacent elements are tied together at shared nodes resulting in a volumetric, continuous tetrahedral mesh-model, representing the tissues of interest. Linear-elastic material properties are then assigned to the nodes of the mesh comprising each different tissue, resulting in a 3D biomechanical model. Young’s Modulus (E) is a measure of stiffness and is defined by the ratio of stress to strain along the direction of the applied stress. Linear stress-strain responses are assumed to be valid over a small range of stresses, however it is also possible to model more complex hyper-elastic responses. Poisson’s ratio (ν) is a measure of compressibility and is defined by the ratio of the percent-extension normal to the applied stress over the percent-contraction in the direction of the stress. Known displacements for a subset of nodes, typically those comprising an organ surface, are required to act as boundary conditions. A system of equations is generated that approximates how the nodes with known displacements (i.e. via the boundary conditions) relate to the remaining nodes via their biomechanical properties. This system, termed the constitutive equations, is solved using a technique called finite element analysis (FEA). The solution provides a displacement for every node and element, and by extension tissue, in the model.

2.2 Deformable Image Registration

The Morfeus DIR and dose accumulation research environment is a MATLAB-based application. It performs biomechanical-based registration by integrating clinical, FEM and FEA software programs as follows (Figure 2.1)[124,131]:

1) As part of the clinical radiotherapy process, clinicians delineate all normal tissues and GTVs on a single planning dataset (e.g. exhale 4D CT) in the treatment planning system (Pinnacle³ v6.2–9.2, Philips Medical Systems, Madison, USA). Throughout this thesis the planning dataset serves as the primary image which is deformed to other secondary images. On the secondary datasets a subset of organs are additionally contoured: the external body, liver, and
spleen. These are easily visualized on most clinical images (e.g. 4D CT, 4D CBCT). Each contour is exported as a binary mask in the resolution of the CT, typically $1 \times 1 \times 2.5\text{–}3 \text{ mm}^3$. The binary mask volume preserves the 3D shape and volume of the contour, but replaces the image texture with a uniform intensity value of 1 internally and 0 externally.

2) Inside Morfeus the masks of the following structures are converted into 3D surface meshes consisting of triangular elements using a modified ‘marching cubes’ algorithm (IDL v6.3 Research Systems Inc., USA): external body, liver, GTV(s), spleen, kidneys and stomach. The surface meshes are smoothed while preserving the topology of the structure delineated by the clinicians[124].

Figure 2.1. Morfeus-DIR framework.
3) The FEM model is constructed using an FEM pre-processor (HyperMesh v9.1-11, Altair Engineering, Troy, USA). Each structure from the primary image is automatically filled with volumetric tetrahedral elements based on the size of their triangular surface element, beginning with the innermost structure (e.g. the GTV) and moving outwards. Nodes at surface mesh-tetrahedral mesh interfaces (e.g. GTV-liver, liver-body) are common to both tissues resulting in a tied connection. Masks for remaining tissues including the heart and luminal gastrointestinal organs are interpolated onto the body’s 3D mesh, and the overlapping elements are re-assigned to those respective tissues. Element volumes are approximately 0.8 mm$^3$ on average, smaller than the image and dose grid voxel sizes. The total number of elements in the model depends on the scan length of the planning CT (average: 2.4×10$^5$ elements).

4) Tissue-specific material properties are assigned to the nodes comprising each tissue using the FEM pre-processor. These parameters were adapted by Brock et al. from existing literature and subsequently refined to optimize the geometric accuracy of Morfeus[124]. All tissues are modeled as nearly incompressible ($\nu$ range: 0.4–0.499). The stiffness ($E$) ranges from 78 kPa for the GTVs to 1.5 kPa for the non-specific body tissues. The stomach is intentionally modeled stiffer ($E = 500$ kPa) than published values of 1–100 kPa to account for its non-anatomic stomach contents present in the images[124].

5) The boundary conditions are determined by explicitly modeling individual organ surface deformation for tissues with contours on the secondary image (i.e. the external body, liver, and spleen). In the FEM pre-processor the two triangular surface meshes for the primary and secondary organ are converted into smooth surfaces. A guided-surface projection is then performed between the two surfaces (HyperMorph, Altair Engineering, Troy, USA). This effectively projects the nodes from the primary organ representation, normal to that surface, onto the surface of the corresponding secondary representation. Note that this is not an explicit point-to-point mapping between the surface contours, but rather a point-to-surface mapping that does not expect the user to identify corresponding surface points. This alignment is not based on tissue biomechanics but serves as an automated method to establish a correspondence between images. The vector projections for the surface nodes serve as boundary conditions for the multi-organ FEM model.

6) Using the biomechanical FEM model based on the primary image anatomy (step 3), and the boundary conditions partially describing the nodal displacements to the secondary image (step 5), FEA is then performed (Abaqus v6.9, Dassault Systèmes Simulia Corp., Providence,
USA). FEA implicitly solves the stress and strain of each element and displacements for all remaining nodes away from the boundary conditions. In this manner, displacements for all tissues such as the GTV are modeled without explicit identification on the secondary image.

### 2.3 Deformable Dose Accumulation

Doses are calculated in Pinnacle³ on the available 4D CT phases with isotropic grid resolutions of 2–2.5 mm³ which are used clinically. All patients have end-exhale and end-inhale 4D CT providing dose grids at the two density extremes. Intermediate 4D CT phases are not routinely archived for liver SBRT patients, therefore a previously developed simplification was performed to account for doses at these positions[107,109]. Dose accumulation using the extracted dose grids was performed in Morfeus. DIR provides the location and size of every element in the model at exhale and inhale. Four intermediate positions are linearly interpolated for each element in lieu of the intermediate phase CTs. Non-linear breathing motion trajectories, termed hysteresis, of >1 mm occur in approximately 20% of liver patients although the mean deviation from the linear trajectory ~1 mm (maximum ~2 mm)[70].

Next, each dose grid is interpolated onto the elements for all 6 steps where the dose contribution is weighted to the element’s position along the breathing trajectory. For example, the exhale grid contributes greater dose to an element closer to exhale. Steps are additionally time-weighted based on population fluoroscopic liver motion data reported by Lujan et al. [80]. For the data reported in Chapters 3 and 4, the accumulated element dose at exhale coordinates \(D(x,y,z)\) is summed across the exhale-to-inhale trajectory as:

\[
D(x,y,z) = \sum_{\phi=0}^{1} \left[ D_{exh}(x+\phi \Delta_x, y+\phi \Delta_y, z+\phi \Delta_z)[1-\phi] + \right. \\
\left. D_{inh}(x+\phi \Delta_x, y+\phi \Delta_y, z+\phi \Delta_z)[\phi] \right] T_\theta 
\]

(1)

where \(D_{exh}\) and \(D_{inh}\) are the dose contributions from the exhale and inhale dose grids respectively. The six steps (\(\phi\)) and corresponding time weights (\(T\)) begin at 0 (0.48) the exhale position, progress through 0.2 (0.13), 0.4 (0.09), 0.6 (0.08), and 0.8 (0.10) and end in the inhale position 1 (0.12)[107]. The exhale-to-inhale translations, \(\Delta\), are determined from DIR. In Chapter 8 an additional CT dataset representing the mean breathing position was retrospectively reconstructed, allowing for the incorporation of a 3rd dose grid. The accumulated element dose was modified to be summed across each exhale-to-inhale deformation map as:
where each dose grid $D$ is weighted with factor $T$, the time in the breathing cycle spent closest to that that grid. This was summed across 3 positions (denoted by subscripts): exhale (exh), time-averaged exhale-to-inhale position (mid), and inhale (inh). The weighting factors were also refined to be patient-specific, measured from the average liver-motion over six 4D CBCTs. The mean weightings from the 20 patients investigated in Chapter 8 were $T_{exh} = 0.39\pm0.06$, $T_{mid} = 0.32\pm0.05$ and $T_{inh} = 0.29\pm0.03$.

Equations 1 and 2 accumulate the breathing dose over one 4D imaging session via DIR. To reconstruct the delivered dose, breathing dose accumulation is first done using each fraction’s 4D CBCT and then additionally summed across all six fractions. Note that this approach effectively assumes that the original planning dose grid is shift- and deformation-invariant. A dose reconstruction study of prostate IMRT showed that accumulation using only the planning dose grids produces per-voxel mean difference of 0.01±1.56% for shifts up to 10 mm, compared to grid re-calculation using repeat CTs\[132\]. Studies conducted in the lung have also shown that breathing dose accumulation does also not require dose grids from all 4D CT datasets. Rosu et al. demonstrated that accumulation using all 10 phases can be approximated with just exhale and inhale 4D CT to within 2%\[133\]. Mexner et al. and Glide-Hurst et al. demonstrated that a single grid from the mean 4D CT (i.e. the average over 10 phases), or the mid-ventilation CT (i.e. one phase representing the mean breathing position) respectively results in 4D dose accumulation errors <2% compared to using all 10 phases\[134,135\]. The abdomen and liver region typically have less density variations than the lung, and therefore can be assumed to be even less sensitive to the discrepancies in using fewer dose grids.

### 2.4 Validation Studies

The published geometric validation studies are summarized in Table 2.1. The accuracy based on vessel bifurcations within the liver is $\leq 1.7$ mm in each direction, within the image voxel sizes for these studies. Brock et al. additionally quantified the surface accuracy for organs that are implicitly deformed without surface-projections (e.g. kidneys and stomach)\[124\]. The registration error was (average±standard deviation) 1.1±1.1 mm LR, 1.3±1.2 mm AP and 0.8±0.9 mm SI.
Table 2.1. Accuracy of Morfeus DIR in the liver.

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Image data</th>
<th>Accuracy evaluation (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brock (2005)[124]</td>
<td>5 volunteers, MRI to MRI</td>
<td>7–10 bifurcations,</td>
</tr>
<tr>
<td>Brock (2006)[131]</td>
<td>5 patients, CT to MRI</td>
<td>4–5 bifurcations, average±SD (range): 3D: 4.2±1.4 (4.0–4.4)</td>
</tr>
<tr>
<td>Voroney (2006)[128]</td>
<td>17 patients, CT to MRI</td>
<td>5 bifurcations, average±SD: 3D: 4.2±1.7</td>
</tr>
<tr>
<td>Al-Mayah (2009)[125]</td>
<td>5 patients, CT to CT</td>
<td>~5 bifurcations,</td>
</tr>
</tbody>
</table>

Al-Mayah et al. experimented with a frictionless contact-surface between the liver and its surroundings (e.g. chest wall, stomach etc.) to permit the liver to more realistically slide during breathing[125]. The mean absolute registration error differed by <0.5 mm compared to the surface projection methods, but required greater computation time. Therefore contact-surfaces were not applied to the clinical data in this thesis.

The accuracy of Morfeus is largely independent of the image intensities, but is limited to some extent by the uncertainty in the contours driving the DIR. Inter-observer variability, based on two observer’s liver contours from five (i.e. non-4D) CBCT images acquired under ABC, previously impacted the Morfeus-predicted GTV centre-of-mass by <1 mm on average[103]. The impact of inter-observer variability on delivered dose was assessed for one patient treated with ABC. Re-contouring the six breath-hold CBCTs prior to dose accumulation was repeated by a radiation oncology fellow, Anand Swaminath, changing the delivered minimum (0.5 cm³) GTV dose by 0.5 Gy. Intra-observer variability was assessed by re-contouring exhale and inhale 4D CBCTs for one random fraction each for 10 patients prior to accumulating the breathing dose for that fraction. For the re-contoured fraction, the mean±SD difference in GTV centre-of-mass exhale position was 0.0±1.5 mm LR, 1.5±1.9 mm AP, and -0.5±1.6 mm SI, while the difference to the accumulated minimum (0.5 cm³) GTV doses were 0.0±0.4 Gy. These observer variations resulted in geometric differences similar in magnitude to the accuracy of Morfeus, while the dosimetric differences are lower than the threshold of changes later reported to be of potential clinical significance (e.g. 5% or 1 Gy).
Finally, Niu et al. developed a deformable gel dosimeter that was irradiated during breathing-like motion to experimentally validate deformable dose accumulation[136]. The Morfeus-predicted dose distribution was compared to the physical dose distribution in the dosimeter reconstructed with MRI. Greater than 95% of the voxels passed the ~5% dose difference and 3 mm DTA acceptance criteria, indicating a high degree of agreement to the measured data.

### 2.5 Prior Dose Accumulation Study of Liver SBRT

The largest study investigating deformable dose accumulation in 21 free-breathing liver SBRT patients was performed using Morfeus[109]. Morfeus registered exhale to inhale 4D CT to incorporate multi-organ breathing motion into the dose distribution of clinical plans initially optimized on static exhale CT. Compared to the static dose, the 4D CT-predicted dose (DIR and 4D CT) resulted in changes of ≥5% to the relevant GTV and normal tissue dose criteria in 12 patients (57%). The 4D CT-predicted minimum GTV dose decreased on average by -0.9% (range: -14.2 to 8.0%) compared to the static dose. The 3D GTV motion modeled with DIR exceeded the total clinical PTV (ITV plus a setup margin) by 1–10 mm in 9 (43%) patients. This suggests that complex deformation-motion seen with DIR and 4D CT may aid in PTV design which was clinically based on 2D cine-MRI GTV motion, 2D diaphragm fluoroscopy motion, or rigid 3D liver motion on 4D CT. Compared to the static dose, changes in the maximum dose to normal tissues ranged from -25 to 13% with the 4D CT-predicted dose. In cases where the PTV dose coverage was compromised to spare normal tissues on the static dose, those normal tissue doses change by -13% to 9% with the 4D CT-predicted dose. For example, the PTV dose was compromised in one patient to spare the bowel which was near the planned dose limit on static exhale CT. The maximum bowel dose decreased by 13% on the 4D CT-predicted distribution as DIR modeled the bowel motion outside of high-dose region. With this knowledge, this patient potentially could have been planned to receive higher tumor doses. Deformable dose accumulation was additional compared to rigid accumulation based on 3D COM motion of the liver. Differences up to 8.3% to the minimum GTV doses, and 7.2% to the maximum normal tissue doses were observed highlighting the limitation in using rigid motion models for dose accumulation. In summary, this study demonstrated that the 4D CT-predicted breathing doses distribution may have a substantial clinical impact on the planning of liver SBRT patients.
2.6 Summary

An in-house, biomechanical model-based DIR algorithm called Morfeus has been previously developed and validated. The accuracy of Morfeus is within the voxel size of typical clinical images, and is image-modality independent. The application of biomechanical DIR to dose accumulation has been previously demonstrated for liver SBRT planning. Extension to the images used for SBRT guidance would permit breathing dose accumulation at each fraction, and retrospective reconstruction of the delivered SBRT dose.
Chapter 3
Accumulated Dose in Liver Stereotactic-Body Radiotherapy: Positioning, Breathing and Deformation Effects

3 Accumulated Dose in Liver SBRT: Positioning, Breathing and Deformation Effects

3.1 Introduction

Stereotactic body radiotherapy (SBRT) is a promising treatment for primary and metastatic liver cancer patients ineligible for other localized treatment. SBRT planning uses individualized, highly conformal dose distributions aimed at reducing treatment margins and sparing normal tissue dose and related toxicity. Liver normal tissue complication probability (NTCP) models can help estimate SBRT toxicity and allocate or escalate dose\(^1\). Trials have shown high local control rates with acceptable toxicity\(^2\), whereas lower doses have been associated with poorer survival or disease control\(^3\), suggesting that further dose escalation may be beneficial provided toxicity rates remain low.

Minimizing geometric uncertainties is necessary for SBRT. Respiratory motion can be negated using active breathing control (ABC) devices allowing gated beam delivery during breath holds or reduced with an abdominal compression plate\(^4\). These may allow for smaller margins, normal tissue sparing, and higher tumor doses, but many patients are ineligible and are treated free-breathing. Incorporating breathing motion into liver SBRT dose calculations can potentially impact tumor and normal tissue doses and margin design\(^5\). Whether these techniques improve estimates of the delivered dose better than static dose calculations is presently unknown. Image-guided radiotherapy (IGRT) can potentially identify and correct baseline shifts in liver position (relative to bone), breathing motion, or deformation before treatment\(^6\). Direct tumor visualization is typically not possible, so IGRT methods for liver SBRT involve imaging fiducial markers\(^7\), or the liver and diaphragm as soft tissue surrogates using two-dimensional fluoroscopy or three dimensional (3D) CBCT in the presence of breathing motion\(^8\).

Méndez Romero et al. estimated daily dose deviations in a liver SBRT trial using rigidly registered repeat CT, finding that IGRT did not on average improve the daily dose to normal tissues, owing to anatomic deformations\(^9\). Rigid registration is unable to accumulate dose over multiple fractions in the presence of these changes. Deformable image registration (DIR) applies spatially variable transformations during registration to more accurately track tissues between two or more imaging sessions (i.e., SBRT fractions). Janssens et al. found that
intensity-based DIR significantly improved inter-fraction dose accumulation on deforming phantoms over rigid registration, noting that its accuracy is highly dependent on both image quality and contrast[137]. They also reported that sharp dose gradients, required by liver SBRT plans to spare normal tissue, can exacerbate dose accumulation errors caused by DIR errors. Soft-tissue contrast is generally poor on CBCT, making IGRT and DIR challenging. Brock et al. applied a biomechanical model-based DIR algorithm on the daily CBCT of liver SBRT patients, revealing residual errors in the tumor position that exceeded the setup tolerance in 15% of fractions[103]. The dosimetric impact of these uncertainties is not well understood.

Previous related work indicated that performing deformable dose accumulation incorporating breathing motion from the 4D CT resulted in substantial deviations in the estimated dose to the tumor and normal tissues compared with the static plan[109]. The work presented here expands on this, evaluating how well the planning 4D CT predicts for the best estimate of delivered dose, using deformable dose accumulation over each fraction’s 4D CBCT. The aim was to accumulate dose using DIR of CBCT over 6-fraction SBRT in free-breathing liver patients. This was compared with both the static dose on the planning CT and the breathing dose predicted from the planning 4D CT, to assess which method better predicts the accumulated dose. The second aim was to investigate the effect of different geometric uncertainties on dose deviations. Characterizing uncertainties in current SBRT techniques may enable robust planning development, enable safe escalation of SBRT doses in future trials, and improve interpretation of clinical outcomes.

3.2 Methods and Materials

3.2.1 Patients and SBRT planning

Thirty patients previously treated on institutional review board approved Phase I and II trials of dose-escalated, hypo-fractionated liver SBRT from February 2006 to April 2010 were investigated. Patient and planning details are summarized in Table 3.1. All were ineligible for active breathing control treatment owing to intolerance or small breathing amplitudes (<5 mm), and thus were treated free-breathing (with or without abdominal compression). Planning was done on end-exhale 4D CT. Inhale 4D CT liver motion, diaphragm fluoroscopy motion, and cine-MRI tumor motion aided in breathing motion characterization for designing individualized planning target volumes (PTV). Delineation and static plan optimization was done on exhale CT in a commercial treatment planning system (Pinnacle; v7.6–8.0; Philips Medical Systems,
Asymmetric PTVs were designed to account for the patient-specific breathing motion observed on the imaging studies, with a minimum PTV of 5 mm required. Dose was individually prescribed for 6 fractions in 2 weeks by determining the risk of radiation-induced liver disease from a Lyman-NTCP model[53]. The primary planning goal was that the minimum dose to the GTV and PTV received a minimum of 95% of the prescribed dose to 0.5 cm, while respecting normal tissue constraints. The maximum allowable doses to the luminal gastrointestinal organs ranged from 30–36 Gy to 0.5 cm³. Volume, margin generation, and NTCP calculations have been detailed by Dawson et al.[53].

Table 3.1. Patient characteristics and SBRT treatment details.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age [median (range)] (years)</td>
<td>72 (45 – 83)</td>
</tr>
<tr>
<td>Gender: male/female (n)</td>
<td>19/11</td>
</tr>
<tr>
<td>Diagnosis (n):</td>
<td></td>
</tr>
<tr>
<td>Primary liver cancer</td>
<td>15</td>
</tr>
<tr>
<td>Liver metastases</td>
<td>15</td>
</tr>
<tr>
<td>No. GTV [median (range)] (n)</td>
<td>2 (1 – 5)</td>
</tr>
<tr>
<td>GTV volume [mean (range)] (cm³)</td>
<td>162 (4 – 707)</td>
</tr>
<tr>
<td>Motion management (n):</td>
<td></td>
</tr>
<tr>
<td>Free-breathing</td>
<td>11</td>
</tr>
<tr>
<td>Abdominal compression plate</td>
<td>19</td>
</tr>
<tr>
<td>Planned fluoroscopy motion [mean (range)] (mm)</td>
<td>9 (2 – 17)</td>
</tr>
<tr>
<td>PTV [mean (max)] (mm):</td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>6 (13)</td>
</tr>
<tr>
<td>Right</td>
<td>5 (7)</td>
</tr>
<tr>
<td>Anterior</td>
<td>7 (20)</td>
</tr>
<tr>
<td>Posterior</td>
<td>5 (8)</td>
</tr>
<tr>
<td>Superior</td>
<td>5 (6)</td>
</tr>
<tr>
<td>Inferior</td>
<td>12 (23)</td>
</tr>
<tr>
<td>No. beams, excluding segments [median (range)] (n)</td>
<td>6 (4 – 15)</td>
</tr>
<tr>
<td>Liver NTCP [mean (range)] (%)</td>
<td>1.8 (0 – 15.2)</td>
</tr>
<tr>
<td>Prescribed dose in 6 fractions [mean (range)] (Gy)</td>
<td>39 (27 – 60)</td>
</tr>
</tbody>
</table>

Abbreviations: GTV=gross tumor volumes; PTV=planning target volume; NTCP=normal tissue complication probability (biologically corrected to 2 Gy/fraction).
3.2.2 Image-guided radiotherapy

Patients were treated with daily IGRT in free-breathing on linear accelerators equipped with kilovoltage CBCT and fluoroscopy (Synergy; Elekta, Stockholm, Sweden). Before treatment, anterior-posterior fluoroscopy was acquired to assess the maximum exhale diaphragm position and breathing amplitude compared with the planning images. In addition, a 360° CBCT was acquired and rigidly registered to the planning exhale 4D CT using a 3D liver-to-liver alignment. Breathing artifacts blurred the CBCT, therefore the liver match was biased to the superior part of the blur closer to planned exhale CT position. This soft-tissue IGRT strategy accounted for baseline liver shifts relative to bony anatomy. Tolerance for image registration was 3 mm and 5°, and any corrections were verified with repeat imaging.

For offline research analysis a CBCT was acquired in the final treatment position before SBRT. These were retrospectively sorted into 10 respiratory-correlated phases (4D CBCT)[93], and the end-exhale and end-inhale phases were extracted.

3.2.3 DIR and dose accumulation

Dose accumulation requires that tissues be accurately tracked between images. DIR has more degrees of freedom over rigid registration to realistically map soft-tissue motion and deformation. This study used Morfeus, a multi-organ biomechanical model-based DIR algorithm, previously described in detail[124]. Briefly, a base model is first created by converting the exhale 4D CT planning contours into 3D surface meshes and filled with tetrahedral elements. The liver, external body, and spleen meshes are deformed into their corresponding secondary surface meshes, created from additional contouring on secondary images (e.g. inhale 4D CT, 4D CBCT) via guided surface projections (HyperMorph; Altair Engineering, Troy, USA). All elements, including GTV and organs without secondary contours, are implicitly deformed by tissue biomechanics and solved using finite element analysis (Figure 3.1). The ability of Morfeus to accurately track the entire volume of the organs of interest (i.e., the liver, including the interior volume) has been previously quantified using visible anatomic landmarks within the liver. This accuracy is 2 mm[124].
Figure 3.1. Example of multi-organ deformable registration between exhale CT (left) and exhale CBCT (middle), and the resulting deformation map (right).

For this study, dose was accumulated and compared in the Morfeus environment. The static clinical plan was calculated on both the exhale and inhale planning CT in the treatment planning system, providing two extreme dose grids occurring during breathing, and imported into Morfeus. Deformable image registration provides the location of all elements in the model between exhale and inhale images on 4D CT or 4D CBCT. To accumulate a breathing dose the dose grids are interpolated onto each element’s position at exhale, inhale, and four linearly interpolated intermediate phases. Each phases’ contribution is weighted according to the element’s position in the breathing cycle, as well as time spent in that phase[80]. Element dose was summed across the breathing cycle as reported in Chapter 2[109].

Three doses were calculated, as follows:

1) **Planned dose** ($D_{\text{plan}}$) is the static clinical plan, replicated by interpolating the exhale 4D CT dose grid onto the initial mesh model at exhale 4D CT. No breathing motion, setup errors, deformation or dose accumulation is considered.

2) **Predicted dose** ($D_{\text{pred}}$) is the breathing dose predicted by the planning 4D CT. Exhale 4D CT is deformed to inhale 4D CT. The exhale and inhale 4D CT dose grids are interpolated onto this deformation map to accumulate breathing dose. Inter-fraction setup errors, deformations, and breathing changes over the course of SBRT are not considered.
3) **Accumulated dose** ($D_{acc}$) is the dose accumulated over the course of SBRT. Exhale 4D CT is first deformed to each fraction’s exhale 4D CBCT, to account for all baseline inter-fraction changes (e.g. setup errors, baseline liver shifts, and deformation) and subsequently deformed to each inhale 4D CBCT to account for the daily breathing motion. The exhale and inhale 4D CT dose grids are interpolated onto each fraction’s exhale-to-inhale 4D CBCT deformation map to accumulate breathing dose, and subsequently the doses from all 6 fractions are then summed. Note that daily dose grids were not recalculated using each CBCT in order to avoid the CBCT intensity number inaccuracies. Rather, DIR was used to track anatomic motion and deformation within the dose grids calculated on the initial planning 4D CT and accumulate the dose therein.

The $D_{acc}$ was primarily compared with $D_{pred}$ for reporting changes in accumulated dose ($D_{acc}$-$D_{pred}$), because the clinical plans did not model breathing motion. However, differences between the accumulated and planned static doses ($D_{acc}$-$D_{plan}$) were also computed to investigate whether 4D calculations ($D_{pred}$) are superior in predicting $D_{acc}$ compared with $D_{plan}$. Minimum dose (to 0.5 cm$^3$) to GTV(s), mean dose to liver and kidneys, and maximum dose (to 0.5 cm$^3$) to all other organs were investigated. Dose deviations are described as a percentage of the prescribed dose, and large deviations ($\geq 5\%$) are assumed to be potentially clinically significant. Liver NTCP changes were also investigated.

Rigid geometric uncertainties during treatment were calculated for each tissue’s center-of-mass (COM) displacement as the group mean ($M$), systematic ($\Sigma_{COM}$), and random ($\sigma_{COM}$) errors[66]. Deformation geometric uncertainties were calculated using the method proposed by van Mourik et al.[138], with the liver COM subtracted from each tissue before calculating the group systematic ($\Sigma_{DEF}$) and random ($\sigma_{DEF}$) errors. Errors were calculated between the planned position on exhale 4D CT and daily exhale 4D CBCT (which accounts for baseline liver shifts relative to vertebral bodies) and between exhale and inhale 4D CBCT. Results are in the left-right (LR), anterior-posterior (AP), and superior-inferior (SI) directions.

### 3.2.4 Comparison of geometric uncertainties

Identification of the cause of dose deviations was a secondary study aim. All deformable registrations started from exhale 4D CT, preserving the number and sequencing of elements in the base model. This facilitated tracking and dose accumulation and allowed the displacement
map from one registration (exhale 4D CT deformed to inhale 4D CT) to be applied to the results of another registration (exhale 4D CT deformed to exhale 4D CBCT). This feature was exploited to investigate the relative importance of different geometric errors on dose deviations ($D_{acc}$-$D_{pred}$) by comparing several scenarios (Figure 3.2), as follows.

**Figure 3.2.** Schema of dose accumulation comparisons to extract the contribution of geometric uncertainties on dose deviations: (A) predicted 4D CT breathing dose ($D_{pred}$); (B) full SBRT dose accumulation ($D_{acc}$); (C) deformable inter-fraction accumulation with the 4D CT breathing motion applied; (D) rigid inter-fraction accumulation with the 4D CT breathing motion applied. **Abbreviations:** Morfeus=DIR. Rigid=liver-to-liver COM registration. Fx=fractions.

**Breathing variations:** These are changes in breathing motion and breathing deformation between planning (4D CT) and treatment (4D CBCT). The $D_{acc}$ (Figure 3.2B) was compared with a modified accumulation (Figure 3.2C) that deforms exhale 4D CT to exhale 4D CBCT (identical to $D_{acc}$) to account for all inter-fraction changes but subsequently applies the predicted 4D CT deformation map (not the 4D CBCT deformation map). Both scenarios identically model inter-fraction errors (e.g. setup, baseline shifts, and deformation), allowing the effect of breathing changes on dose accumulation to be measured.
Residual setup: These are positioning errors, or the rigid liver COM differences between exhale 4D CT and exhale 4D CBCT. The $D_{\text{pred}}$ (Figure 3.2A) was compared with a modified accumulation (Figure 3.2D) that accounts for daily setup errors using rigid liver registration between exhale 4D CT and exhale 4D CBCT (not deformable registration as in $D_{\text{acc}}$) and subsequently applies the predicted 4D CT deformation map (identical to $D_{\text{pred}}$). Each scenario modeled breathing identically, whereas the latter accounts for residual rigid setup effects in the dose accumulation. Other inter-fraction changes (e.g. baseline liver shifts, deformation) are not modeled in this comparison.

Deformations: This is abdominal deformation. It includes baseline differences in tissue position (relative to liver) and shape changes of all tissues between exhale 4D CT and exhale 4D CBCT. Two modified accumulations scenarios were compared. The first (Figure 3.2C) accounts for inter-fraction changes between exhale 4D CT and exhale 4D CBCT using DIR (setup errors, baseline shifts, and deformation), whereas the second (Figure 3.2D) uses rigid liver COM registration (ignoring baseline shifts and deformation). Both scenarios subsequently apply the predicted 4D CT deformation map to identically model breathing motion. Therefore, accounting for inter-fraction changes with either deformable or rigid registration is measured.

3.3 Results

Dose was accumulated for 30 patients at planning ($D_{\text{pred}}$) and over the entire treatment course ($D_{\text{acc}}$), using DIR of 60 4D CT and 360 4D CBCT images respectively.

3.3.1 Geometric uncertainties

Group geometric uncertainties evaluated after DIR are summarized in Table 3.2. Between exhale 4D CT and exhale 4D CBCT, 9 patients (30%) had individual systematic (mean) residual GTV COM errors $>3$ mm in any direction. These occurred in the LR, AP, and SI directions in 2, 6, and 5 patients, respectively, up to a maximum of 11 mm. Sixteen patients (53%) had 3 or more fractions with residual GTV COM errors $>3$ mm. Between 4D CBCT and 4D CT, 16 patients (53%) had mean changes in the GTV COM breathing magnitude of at least 3 mm in any direction. These changes of -11 to 8 mm occurred in LR, AP, and SI directions in 2, 6, and 12 patients, respectively. Sixteen patients (53%) had at least 3 fractions with changes in GTV COM breathing motion $>3$ mm.
3.3.2 Accumulated dose changes

Accumulated dose is compared with the planned and predicted dose in Table 3.3. Relative to the planned dose ($D_{\text{acc}} - D_{\text{plan}}$), 21 patients (70%) had large dose deviations ($|\Delta| \geq 5\%$) to any tissue or GTV. Thirty-nine tissues in these 21 patients had deviations, and 34 (87%) of these were dose decreases compared with $D_{\text{plan}}$. Relative to the predicted dose ($D_{\text{acc}} - D_{\text{pred}}$), 16 patients (53%) had large dose changes to any tissue or GTV. Thirty-two tissues in these 16 patients had significant changes, with 55% decreasing and 45% increasing in dose compared with $D_{\text{pred}}$. Mean accumulated dose ($D_{\text{acc}}$) differed significantly ($p<0.05$) from the planned dose ($D_{\text{plan}}$) for the majority of organs. Mean accumulated deviations relative to the predicted dose ($D_{\text{acc}} - D_{\text{pred}}$) differed significantly, and were often smaller in magnitude, from mean deviations relative to the static plan ($D_{\text{acc}} - D_{\text{plan}}$). This suggests that $D_{\text{pred}}$ may be better overall at predicting the accumulated dose, particularly for the normal gastrointestinal organs. However, a large range of deviations were still observed. An example of dose deviations relative to both the planned and predicted dose is shown in Figure 3.3.

The contribution of each geometric error on deviations from the predicted dose ($D_{\text{acc}} - D_{\text{pred}}$) is shown in Figure 3.4 for selected critical organs. For the 32 tissues that did have large deviations, residual setup errors were the largest cause in 17 tissues (effect size: -15% to 18%), deformation in 8 tissues (effect size: -22% to 11%), and breathing variations in 7 tissues (effect size: -9% to 7%). Although many tissues had small overall dose deviations <5%, geometric errors were individually observed to have large effects that could offset each other (Figure 3.4).

Three patients had significant GTV dose deviations ($D_{\text{acc}} - D_{\text{pred}}$). One patient with three GTVs had an increase of 5% to a portal-vein thrombus being treated to a lower dose than the other GTVs. This thrombus had residual systematic errors of 5 mm left and 10 mm posterior, shifting it toward the other higher dose GTVs. Another patient had a GTV dose increase of 13% relative to $D_{\text{pred}}$ (Figure 3.3). The third patient with a decrease in minimum GTV dose of 14% due to deformation is shown in Figure 3.5.

Liver NTCP changes after accumulation were small because the initial planned NTCP was often low (<2%), and mean liver dose deviations ($D_{\text{acc}} - D_{\text{pred}}$) were small. Two patients had NTCP increases relative to $D_{\text{pred}}$ of 8.8% and 9.4%, caused by 3 mm residual systematic setup errors
and 8 mm less superior-inferior mean breathing motion, respectively. These errors moved more normal liver into the high-dose region.

Figure 3.3. Deviations from the accumulated dose (D_{acc}) are shown in a dose difference image (top) and dose-volume histogram (bottom). PTV coverage was compromised inferiorly on the static exhale 4D CT plan (D_{plan}) to spare the large bowel, whereas 4D CT (D_{pred}) predicted decreased dose as these tissues moved inferior away from the high-dose region. Geometric errors seen on 4D CBCT moved these tissues back toward the high-dose region.
Table 3.2. Population geometric errors (in mm) at planning and during 6-fraction SBRT evaluated with deformable image registration.

<table>
<thead>
<tr>
<th>Region</th>
<th>Direction</th>
<th>Exhale 4D CT to Inhale 4D CT</th>
<th>Exhale 4D CT to Exhale 4D CBCT</th>
<th>Exhale 4D CBCT to Inhale 4D CBCT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Rigid (mean, SD)</td>
<td>Deformation (SD*, RMS†)</td>
<td>Rigid (Mean, ΣCOM, σCOM)</td>
</tr>
<tr>
<td>GTV</td>
<td>LR</td>
<td>-1.3, 2.1</td>
<td>1.0, 1.0</td>
<td>0.7, 1.8, 1.7</td>
</tr>
<tr>
<td></td>
<td>AP</td>
<td>3.7, 3.0</td>
<td>1.7, 0.8</td>
<td>-0.8, 2.2, 1.9</td>
</tr>
<tr>
<td></td>
<td>SI</td>
<td>8.8, 5.0</td>
<td>1.9, 1.2</td>
<td>-0.5, 2.6, 2.4</td>
</tr>
<tr>
<td>Liver</td>
<td>LR</td>
<td>-1.1, 2.0</td>
<td>0.5, 1.5</td>
<td>0.5, 1.7, 1.9</td>
</tr>
<tr>
<td></td>
<td>AP</td>
<td>3.9, 2.3</td>
<td>0.4, 1.8</td>
<td>-0.8, 2.2, 1.9</td>
</tr>
<tr>
<td></td>
<td>SI</td>
<td>9.0, 5.2</td>
<td>0.5, 2.0</td>
<td>-0.4, 2.1, 2.4</td>
</tr>
</tbody>
</table>

Abbreviations: CT=computed tomography; CBCT=cone-beam CT; COM=centre-of-mass; DEF=deformation; SD=standard deviation; RMS=root mean square; Σ=group systematic error; σ=group random error; GTV=gross tumor volumes. Notes: positive mean values are motion in the left, anterior, inferior directions. *SD of each patient's mean elements' motion (minus liver COM). †RMS of the SD of each patient's elements' motion (minus liver COM).
Table 3.3. Accumulated treatment dose ($D_{acc}$) deviations relative to the planned static dose ($D_{plan}$) and the 4D CT-predicted breathing dose ($D_{pred}$).

<table>
<thead>
<tr>
<th>Tissue (dose parameter)</th>
<th>Accumulated ($D_{acc}$) vs. planned ($D_{plan}$)</th>
<th>Accumulated ($D_{acc}$) vs. predicted ($D_{pred}$)</th>
<th>$D_{acc}$-$D_{plan}$ vs. $D_{acc}$-$D_{pred}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean±SD in Gy</td>
<td>Range in Gy (% of Rx)</td>
<td>% patients with a change in dose ≥5%</td>
</tr>
<tr>
<td>GTV (min), $n=54$</td>
<td>-0.2±1.0</td>
<td>-4.4, 2.3 (-15, 5)</td>
<td>10</td>
</tr>
<tr>
<td>Liver (mean), $n=30$</td>
<td>-0.2±0.5*</td>
<td>-1.7, 0.9 (-6, 2)</td>
<td>3</td>
</tr>
<tr>
<td>Liver (NTCP)$^\dagger$, $n=30$</td>
<td>-0.5±2.5$^\dagger$</td>
<td>-8.3, 8.0$^\dagger$</td>
<td>10</td>
</tr>
<tr>
<td>Large bowel (max), $n=30$</td>
<td>-1.1±1.5*</td>
<td>-5.3, 1.3 (-15, 3)</td>
<td>33</td>
</tr>
<tr>
<td>Small bowel (max), $n=15$</td>
<td>-1.3±2.2*</td>
<td>-7.8, 0 (-26, 0)</td>
<td>20</td>
</tr>
<tr>
<td>Duodenum (max), $n=30$</td>
<td>-1.5±2.6*</td>
<td>-12.6, 0.7 (-42, 3)</td>
<td>33</td>
</tr>
<tr>
<td>Esophagus (max), $n=29$</td>
<td>0.3±0.8*</td>
<td>-0.8, 2.4 (-3, 8)</td>
<td>7</td>
</tr>
<tr>
<td>Stomach (max), $n=30$</td>
<td>-0.4±1.5</td>
<td>-4.3, 4.6 (-14, 8)</td>
<td>17</td>
</tr>
<tr>
<td>Right kidney (mean), $n=30$</td>
<td>-0.4±0.7*</td>
<td>-2.0, 0.6 (-5, 2)</td>
<td>10</td>
</tr>
<tr>
<td>Left kidney (mean), $n=30$</td>
<td>-0.1±0.3</td>
<td>-1.2, 0.4 (-3, 1)</td>
<td>0</td>
</tr>
<tr>
<td>Heart (max), $n=25$</td>
<td>-0.5±1.0*</td>
<td>-4.0, 0.8 (-13, 2)</td>
<td>8</td>
</tr>
</tbody>
</table>

Abbreviations: NTCP=normal tissue complication probability; Rx=prescription dose. Other abbreviations as in Table 1. Notes: * $p < 0.05$ on paired t-test. $^\dagger$ Values are % NTCP. Min and Max are to 0.5 cm$^3$. 

48
Figure 3.4. Relative contribution of geometric errors to changes between the accumulated ($D_{\text{acc}}$) and predicted ($D_{\text{pred}}$) dose for all patients’ stomachs and bowels. Data is sorted by decreasing magnitude and direction of dose deviation.
Figure 3.5. An example of a large dose deviation caused primarily by deformation. Left: patient was planned on exhale 4D CT for 30 Gy with abdominal compression (liver=purple; GTV=red). Middle: deformation in the abdomen shown on CBCT was possibly due to a misaligned compression plate and increased stomach contents. Right: dose difference between the accumulated and predicted doses ($D_{\text{acc}} - D_{\text{pred}}$) showing delivered decreases (dark blue region) to the minimum GTV of 4.3 Gy, and the maximum duodenum by 11.5 Gy.

3.4 Discussion

Accumulated dose was investigated for 30 liver SBRT patients. This is the first study using DIR of daily 4D CBCT to reconstruct the delivered dose over the entire course of 6-fraction liver SBRT, simultaneously accounting for and evaluating the relative importance of different geometric uncertainties. Significant accumulated dose deviations relative to either the planned distribution ($D_{\text{acc}} - D_{\text{plan}}$) or the breathing dose distribution modeled with the planning 4D CT ($D_{\text{acc}} - D_{\text{pred}}$) were observed in the majority of patients. Residual setup errors, observed after DIR of 4D CBCT, followed by deformation and breathing variations, were the most common source of deviations.

This study compared $D_{\text{acc}}$ with the static plan on exhale 4D CT ($D_{\text{plan}}$) and with the planning 4D CT-predicted breathing dose ($D_{\text{pred}}$), to help determine which distribution is more representative of the delivered dose. Overall, the mean changes between $D_{\text{acc}}$ and $D_{\text{pred}}$ significantly differed, and were often smaller in magnitude, compared with changes between $D_{\text{acc}}$ and $D_{\text{plan}}$ (Table 3.3). Four-dimensional breathing dose distributions ($D_{\text{pred}}$) may be more robust than static distributions ($D_{\text{plan}}$) particularly for the duodenum and bowels, often dose-limiting at planning[109].

In the clinical trials that these patients were treated on the prescribed dose was limited by the liver NTCP or other normal tissue doses on $D_{\text{plan}}$, which were more often lower after accumulation ($D_{\text{acc}} - D_{\text{plan}}$). Large dose deviations were also observed relative to $D_{\text{pred}}$, suggesting
that full dose accumulation ($D_{acc}$) may be beneficial for future SBRT trials. One patient’s duodenum exceeded the planned tolerance when the maximum dose increased by 0.7 Gy, or 4% ($D_{acc} - D_{pred}$), not thought to be clinically significant. Research is ongoing to correlate accumulated dose with clinical outcomes. Exploiting anatomic changes with adaptive SBRT could possibly safely allow further dose escalation. Deviations caused by uncertainties not modeled in this study, including intra-fraction position changes, stomach filling, and changes to the dose grids caused by external contour variations, also need to be quantified.

Many tissues had large geometric errors exceeding the IGRT tolerance of 3 mm yet had negligible dose changes. Gross tumor volumes are buffered by the PTV and degree of surrounding dose conformity. It is also expected that sharper dose gradients are more sensitive to geometric errors causing larger dose deviations[139]. For 4D CT-predicted dose distributions in the liver, larger tumor changes have been observed for IMRT versus conventional plans[115]. Deformable image registration and dose accumulation can aid in evaluating the robustness of planning solutions on predicting the accumulated dose. Highly conformal techniques (i.e. IMRT) can improve target coverage, spare normal tissue, and possibly allow dose escalation on the static distribution[60], or predicted breathing distribution. Stricter IGRT tolerances are required to minimize accumulated dose deviations for more conformal plans.

Residual setup caused the majority of dose deviations. Many patients had half their SBRT course with tumor displacements (exhale 4D CT vs. exhale 4D CBCT) exceeding the IGRT tolerance. These residual setup errors are likely due to the inherent uncertainties in using the liver as a surrogate, breathing artifacts on the CBCT, and the rigid registration used. Baseline shifts in liver position relative to bony anatomy were largely accounted for through the use of CBCT and soft-tissue targeting, and thus not a significant contributor to dose deviations. Neither DIR nor 4D CBCT for online localization were available clinically during the study period.

Four-dimensional CBCT has the potential to capture both baseline shifts and breathing motion even with simple rigid liver registration[94,95]. Combining online DIR with a priori motion models to improve the quality of the 4D CBCT[140], or target the tumor with limited two-dimensional imaging may also improve IGRT[141]. Many patients had breathing variations (4D CBCT vs. 4D CT), though they were relatively stable, as has been observed in other 4D CBCT studies of the lung and liver[94,142]. Breathing dose distributions ($D_{pred}$) should be implemented only if the planning 4D CT is representative of the treatment breathing motion.
Case et al. reported that mean 4D CBCT motion correlated well with 4D CT, though this was evaluated using rigid liver registration[94]. Substantial geometric differences have been observed between rigid and DIR modeling of breathing motion[109].

Although deformation had less impact on $D_{acc}$ than residual setup error, this should not be interpreted as meaning that DIR can be replaced by rigid registration. Deformable image registration facilitated dose accumulation and was used for all breathing models. In cases in which large anatomic deformations are observed and cannot be corrected for with IGRT alone (Figure 3.5), re-simulation and re-planning is strongly recommended. Deformable image registration and dose accumulation may also facilitate this, though rigorous geometric and dosimetric validation is necessary before clinical implementation.

3.5 Conclusions

Deformable registration of 4D CBCT and dose accumulation improved estimates of the delivered dose to targets and normal tissues in free-breathing liver SBRT patients. The majority of patients had accumulated dose deviations >5% relative to the static clinical plan. Breathing dose distributions predicted with the planning 4D CT can help reduce the overall uncertainty at planning in certain normal tissues, but large deviations still occurred in over half the patients. Breathing dose distributions may need to be coupled with improvements in IGRT before clinical implementation because residual setup uncertainties commonly caused dose deviations. Full dose accumulation during SBRT can account for residual anatomic deformations and may facilitate the development of adaptive therapy and the pursuit of further safe dose escalation. Accumulated dose may help interpret clinical outcomes of SBRT response and toxicity.
Chapter 4
Deformable Dose Reconstruction to Evaluate IGRT Strategies in Free-Breathing Liver SBRT
4 Deformable Dose Reconstruction to Evaluate IGRT Strategies in Free-Breathing Liver SBRT

4.1 Introduction

Technological advances have allowed for potent radiation doses to be prescribed to liver tumors while sparing adjacent normal tissues and toxicity. A hallmark of all SBRT plans is the use of steep dose gradients to conform high doses around PTVs, ensuring tumor dose and normal tissue sparing is maximized. Highly conformal liver SBRT plans though desirable, are more sensitive to geometric treatment uncertainties than less conformal plans[115]. Image-guidance is critical in liver SBRT to minimize geometric uncertainties to ensure PTVs are not prohibitively large, and that the delivered doses do not substantially deviate from the planning.

Image-guidance for liver SBRT may use cone-beam CT (CBCT) to target the liver, a surrogate for the tumor, as baseline shifts in liver position relative to bony anatomy exceed 3 mm in 89% of fractions[95]. Inter-fraction liver and abdomen deformation can occur requiring clinicians to rigidly align the planning liver contour to the CBCT with consideration of tumor location and how surrounding normal tissues fall with respect to the high dose regions. The geometric accuracy of this approach results in residual random and systematic population errors, based on liver measurements, of <3 mm[96,100,101]. Imaging free-breathing patients with a CBCT acquired over several minutes results in blurring artifacts that make interpretation difficult. Respiratory-correlated, or ‘4D’ CBCT, can resolve this breathing motion[93], so that dose distribution optimized on a single static image (i.e. exhale 4D CT) can be registered to the corresponding breathing phase at each fraction (i.e. exhale 4D CBCT). Unless implanted fiducials are present, the liver or diaphragm is often used as a tumor surrogate.

Although liver tumors are not typically visible on CBCT, Brock et al. demonstrated that biomechanical-based DIR can model the liver tumor location on CBCT acquired under breath-hold and may play a role in liver IGRT[103]. In Chapter 3, it was shown that residual setup errors after rigid liver targeting on 3D CBCT, had the largest impact on accumulated dose deviations compared with the planning 4D CT-predicted dose. Combining 4D CBCT and DIR has the potential to target the tumor directly and eliminate these errors. The applications of biomechanical DIR have been previously demonstrated at planning (e.g. incorporating 4D CT motion into the planned dose distribution, target delineation[109,131]) and post-treatment (e.g.
deformable dose reconstruction[143]). Prior to translating DIR into an online imaged-guidance application, evaluation of its potential gains is required. Any improvement in GTV targeting accuracy should then be weighed against the increased complexities and uncertainties associated with DIR. To date the benefit of DIR-based IGRT for free-breathing patients has not been evaluated and a comparison of the dosimetric impact of different IGRT strategies using deformable dose reconstruction has not been explored.

The purpose of this study was to quantify the potential geometric and dosimetric gain in using advanced IGRT strategies for free-breathing liver SBRT patients. These were compared to the results of the clinical IGRT strategy, building on the results reported in Chapter 3. Using the actual clinical SBRT plans for the 30 previously studied patients, IGRT strategies were retrospectively simulated by either rigidly aligning the exhale liver, or the exhale liver tumor position modeled with biomechanical-DIR, between the planning 4D CT and the retrospectively-sorted 4D CBCT. The delivered doses for both new strategies were then accumulated over the course of liver SBRT to account for residual deformations and breathing variations during the course of SBRT. These two strategies were compared to the delivered doses reconstructed with DIR in Chapter 3, following the clinical IGRT protocol based on rigid liver registration and 3D CBCT.

It was hypothesized that coupling the best estimate of dose at planning, the breathing dose predicted with DIR of 4D CT, with direct tumor targeting via DIR with IGRT would reduce the dose deviations between the planned and delivered dose to targets and normal tissues. This analysis should provide insight as to where future efforts should be made at improving liver IGRT, and whether DIR needs to be translated to the treatment unit or if 4D imaging and rigid registration is adequate. Additionally, reducing geometric uncertainties in liver SBRT is warranted prior to consideration of PTV reduction which may allow for escalated doses to be safely delivered with potential improvements in clinical outcomes.

4.2 Methods

4.2.1 Image-guidance strategies

The 30 liver SBRT patients described in Chapter 3 were investigated further. Briefly, these patients were unsuitable for ABC-breath hold and treated free-breathing or with abdominal compression. All were clinically planned on the exhale 4D CT dataset, and this static dose
calculation did not incorporate any breathing motion using dose accumulation. The PTV was asymmetrically applied around the GTV on the exhale 4D CT to encompass the breathing motion measured on the pre-treatment imaging studies, plus a further expansion to account for other SBRT uncertainties (details in Chapter 3, Table 3.1). The minimum PTV allowed in any direction was 5 mm.

For daily IGRT, fluoroscopy initially verified that the diaphragm breathing amplitude was consistent to within 3 mm of that measured at planning and simulation. A 3D CBCT was then acquired in free-breathing and the planning liver contour was rigidly registered to the liver on CBCT. For patients with an amplitude >5 mm, the alignment was biased towards the superior part of the blurred diaphragm (i.e. breathing artifact) which is closer to the daily exhale liver position. Patients were re-positioned and re-imaged if the patient rotation was >5°. This strategy based on 3D CBCT was used to treat the patients clinically, hence is termed ‘clinical-IGRT’.

The delivered doses resulting from clinical-IGRT were retrospectively estimated using deformable dose reconstruction and were reported in Chapter 3. Briefly, each fraction’s final CBCT after clinical-IGRT was retrospectively sorted into 4D CBCT images[93]. Biomechanical-model based DIR, Morfeus, enabled dose accumulation. The planning exhale 4D CT was deformed to exhale and then inhale 4D CBCT, allowing dose accumulation over the entire SBRT course.

For the present investigation two additional IGRT strategies based on 4D CBCT were investigated. The dose reconstruction process above provides a 4D biomechanical model of the patient based on 4D CBCT at each fraction. By rigidly translating these 4D models, mimicking translations of the treatment couch with IGRT, it is possible to simulate alternative IGRT strategies (Figure 4.1).

The first 4D strategy simulated, termed ‘4D-IGRT’, involved aligning the liver between each exhale 4D CBCT to the planning exhale 4D CT. Each fraction’s 4D CBCT model from DIR was rigidly translated to resolve the liver centre-of-mass (COM) displacement calculated with DIR of exhale 4D CT to exhale 4D CBCT. This 4D-IGRT strategy simulates rigid registration of the liver as a surrogate for the GTV (as is the case with the clinical-IGRT) however the uncertainty in liver position caused by CBCT breathing artifacts would be resolved with the use of 4D CBCT.
The second 4D strategy simulated, termed ‘DIR-IGRT’, involved directly aligning the GTV between each exhale 4D CBCT to the planning exhale 4D CT. Each fraction’s 4D CBCT model from DIR was rigidly translated to resolve the GTV centre-of-mass (COM) displacement calculated with DIR of exhale 4D CT to exhale 4D CBCT. For patients with multifocal lesions (n=15), the average COM displacements across all GTVs were corrected for. This DIR-IGRT strategy simulates the online application of biomechanical-DIR to directly target GTVs, thus avoiding the potential limitations of the liver as a surrogate on the exhale 4D CBCT images. Compared to the clinical-IGRT and 4D-IGRT strategies, DIR-IGRT was hypothesized to represent a ‘best-case’ scenario.
4.2.2 Evaluation

The geometric result of each DIR strategy was evaluated by calculating the populations’ mean (M), systematic (Σ) and random (σ) geometric residual errors after the simulated IGRT strategy relative to the planning exhale 4D CT[66]. The error was calculated about each tissue’s COM displacement.

The dosimetric result was evaluated by comparing the different delivered doses following each DIR strategy to the breathing dose predicted at planning which was common to all strategies (Figure 4.2). The clinical plans (i.e. $D_{\text{plan}}$ in Chapter 3) did not model the planning 4D CT breathing motion into the dose distribution. Therefore DIR was performed between the planning exhale and inhale 4D CT, and dose was accumulated over the breathing cycle (i.e. 4D CT images) to generate a 4D CT-‘predicted’ breathing dose distribution. This was previously calculated (i.e. $D_{\text{pred}}$ in Chapter 3) and is used in the current analysis as the baseline dose to which the delivered doses from each IGRT strategy are compared to. The 4D CT-predicted dose facilitated a comparison to the delivered dose because they both incorporate breathing motion.

Deformable dose reconstruction was used to accumulate the delivered doses for each IGRT strategy using the calculations presented in Chapter 2. For clinical-IGRT, the delivered dose was accumulated based on the patient’s treatment position following 3D CBCT guidance. These delivered doses for the clinical-IGRT were previously reconstructed (i.e. $D_{\text{acc}}$ in Chapter 3), and are partially restated in the present Chapter to compare them to the advanced IGRT strategies that were simulated.

For 4D-IGRT, the delivered doses were accumulated after aligning the liver COM between the planning exhale 4D CT and each exhale 4D CBCT. For DIR-IGRT, the delivered doses were accumulated after aligning the GTV COM between the planning exhale 4D CT and each exhale 4D CBCT (where the GTV position on CBCT was determined with DIR). The delivered doses following these new simulated IGRT strategies were then compared to the same 4D CT-predicted dose that the clinical-IGRT was compared to. Deviations between the delivered dose and the 4D CT-predicted dose (e.g. $D_{\text{acc}}-D_{\text{pred}}$) thus accounted for any inter-fraction residual post-IGRT translational errors (for which the geometric magnitude depends on the IGRT strategy), and residual deformations and breathing variations (for which the geometric magnitude is equal for all strategies).
For an individual patient, delivered dose deviations to the relevant clinical dosimetric criteria (e.g. min GTV dose, max bowel dose) of greater than 5% in magnitude (or >1 Gy) were considered to be potentially clinically significant.

Applying the methods from Chapter 3, the dosimetric contribution of the following geometric uncertainties on the delivered doses were also quantified for each strategy:

- Breathing variations: change in breathing magnitude and deformation between planning (DIR of exhale-to-inhale 4D CT) and treatment (DIR of each exhale-to-inhale 4D CBCT).
- Abdominal deformations: change in organ configuration and shape between planning and treatment (DIR of exhale 4D CT to each exhale 4D CBCT).
- Residual rigid setup errors: the rigid liver COM displacement between planning and treatment (rigid liver registration of exhale 4D CT to each exhale 4D CBCT).

### 4.3 Results

#### 4.3.1 Geometric impact of each IGRT strategy

The residual population geometric errors in GTV and liver position following each IGRT strategy calculated between the exhale planning 4D CT and each treatment exhale 4D CBCT are shown in Table 4.1. The population random and systematic errors for the GTV ranged from 1.7–
2.6 mm with the Clinical-IGRT (as reported in Chapter 3), and were reduced to 0.8–1.4 mm with the 4D-IGRT, and ≤0.9 mm with DIR-IGRT. Assuming direct GTV targeting with DIR-IGRT, the residual GTV errors for patients with single tumors were eliminated (the implications of this are discussed later). For multifocal cases (n=15), residual systematic (random) errors were 0.8 (0.6), 1.3 (0.5) and 0.9 (0.6) mm in the LR, AP and SI directions respectively with DIR-IGRT. This was a result of liver inter-fraction deformations and rotations that prevented one translation from aligning all GTVs simultaneously.

Table 4.1. Residual mean (M), systematic (Σ), and random (σ) errors following each IGRT strategy, between the planning exhale 4D CT and treatment exhale 4D CBCTs (in mm).

<table>
<thead>
<tr>
<th></th>
<th>Clinical-IGRT</th>
<th>4D-IGRT</th>
<th>DIR-IGRT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LR</td>
<td>AP</td>
<td>SI</td>
</tr>
<tr>
<td>GTV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>0.7</td>
<td>-0.8</td>
<td>-0.5</td>
</tr>
<tr>
<td>Σ</td>
<td>1.8</td>
<td>2.2</td>
<td>2.6</td>
</tr>
<tr>
<td>σ</td>
<td>1.7</td>
<td>1.9</td>
<td>2.4</td>
</tr>
<tr>
<td>Liver</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>0.5</td>
<td>-0.8</td>
<td>-0.4</td>
</tr>
<tr>
<td>Σ</td>
<td>1.7</td>
<td>2.2</td>
<td>2.1</td>
</tr>
<tr>
<td>σ</td>
<td>1.9</td>
<td>1.9</td>
<td>2.4</td>
</tr>
</tbody>
</table>

The proportion of patients with individual systematic (mean) GTV errors ≥3 mm in any direction was 30% (n=9, four had multiple GTVs) with clinical-IGRT (max: 9.6 mm SI), to 13% (n=4, all had multiple GTVs) with 4D-IGRT (max: 4.9 mm, AP), to 10% (n=3, all had multiple GTVs) with DIR-IGRT (max: 3.3 mm, SI).

The population random and systematic errors for the liver ranged from 1.7–2.4 mm with the Clinical-IGRT, and these were reduced to ≤1 mm with the DIR-IGRT, and to 0 mm for the 4D-IGRT as the intention of this strategy was to directly align the exhale liver. Figure 4.3 depicts the residual population errors for all tissues following each IGRT strategy, as measured from their centre-of-mass displacement relative to the planning exhale 4D CT. Compared to the Clinical-IGRT strategy, each of the non-liver normal tissues errors were in general not reduced by use of either 4D-IGRT or DIR-IGRT. This is a result of the translational corrections applied to the 4D CBCTs models which were intended to align only the GTV or livers respectively. The geometric magnitude of deformations and breathing variations for all tissues (reported in
Chapter 3, Table 3.2) were equal amongst all three strategies because the tissues differed only in their translational position.

4.3.2 Dosimetric impact of each IGRT strategy overall

Differences in the accumulated dose relative to the planning 4D CT-predicted dose for all IGRT strategies are shown in Table 4.2. On average the percent (%) dose deviations between the delivered dose and the 4D CT-predicted doses were ≤2.5% for all tissues using any IGRT strategy, with a large range of dose deviations (Table 4.2). Sixteen of thirty (53%) of patients had large dose deviations (i.e. an absolute change >5%) to any tissue with Clinical-IGRT, with dose deviations ranging from -38% to 10%. Thirteen of thirty (43%) of patients had large dose deviations to any tissue with 4D-IGRT, with dose deviations ranging from -21% to 8%. Fifteen of thirty (50%) of patients had large dose deviations to any tissue with DIR-IGRT, with dose deviations ranging from -24% to 8%.

The duodenum was the only tissue that had an increased proportion of patients (n=6, or 20%) with large dose deviations with the DIR-IGRT, compared to either the Clinical-IGRT or 4D-IGRT (both n=2, or 7%). The other tissues had either equivalent or reduced rates of patients with large dose deviations using the advanced IGRT strategies, compared to the Clinical-IGRT (Table 4.2).

Improved correlation ($R^2$ value) was observed between relevant dose-volume metrics calculated between the planning 4D CT-predicted dose distribution and the delivered doses, with the advanced IGRT strategies versus the Clinical-IGRT. The $R^2$ for the minimum GTV dose was 0.974 with Clinical-IGRT, 0.991 with 4D-IGRT and 0.995 with DIR-IGRT. Compared to the Clinical-IGRT, improved correlation between the 4D CT-predicted and delivered dose was observed for all other tissues using either the 4D-IGRT or DIR-IGRT. An exception was observed for the left kidney where the $R^2$ was 0.986 for Clinical-IGRT, 0.979 for 4D-IGRT and 0.977 for DIR-IGRT. However, the 4D CT-predicted mean left kidney doses were low (average: 2.0 Gy) and the low $R^2$ was skewed by one outlier patient with a small delivered dose deviation <5% (Clinical-IGRT = 2%, 4D-IGRT = 3%, DIR-IGRT = 4%). The overall $R^2$ value (taken as the mean $R^2$ across all tissues in Table 4.2) is 0.978 for Clinical-IGRT, 0.989 for 4D-IGRT and 0.987 for DIR-IGRT. This suggests that the least amount of dosimetric variance occurs between
Figure 4.3 (partial figure, continued on next page)
Figure 4.3 (partial figure, continued from previous page). Population residual geometric errors (based on each tissue’s centre-of-mass displacement) following three image guidance strategies, relative to the planning exhale 4D CT position.
Figure 4.4. The mean (absolute) total dose deviations (%) between the delivered dose and the 4D CT-predicted dose (square markers), for each tissue and IGRT strategy. The stacked colored bars depict the mean (absolute) impact of the different geometric errors towards the total dose deviation.
Table 4.2. Deviations between the delivered dose and the planning 4D CT-predicted dose.

<table>
<thead>
<tr>
<th>Tissue (dose metric*)</th>
<th>Proportion of patients with</th>
<th>Magnitude of % dose deviations, mean-standard deviation (range)</th>
<th>Correlation between delivered and 4D CT-predicted dose ($R^2$)†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proportion of patients with</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$</td>
<td>\text{dose deviations}</td>
<td>&gt; 5%$</td>
</tr>
<tr>
<td></td>
<td>Clinical-IGRT</td>
<td>4D-IGRT</td>
<td>DIR-IGRT</td>
</tr>
<tr>
<td>GTV (min)</td>
<td>10</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>Liver (mean)</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Duodenum (max)</td>
<td>7</td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td>Large bowel (max)</td>
<td>20</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>Small bowel (max)</td>
<td>33</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Esophagus (max)</td>
<td>14</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Stomach (max)</td>
<td>20</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Right kidney (mean)</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Left kidney (mean)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Heart (max)</td>
<td>16</td>
<td>8</td>
<td>4</td>
</tr>
</tbody>
</table>

Notes: *max and min doses are to 0.5cm³; †calculated using absolute dose.
the delivered and 4D CT-predicted doses over all GTVs and normal tissues when the dose is delivered with the 4D-IGRT strategy. The correlations between the delivered and predicted doses (R value) were significantly higher in most tissues (all $p<0.05$) for either advanced IGRT strategy compared to the clinical-IGRT. The only exceptions were for the small bowel and left kidney where the correlations did not differ significantly for the advanced IGRT strategies versus the clinical-IGRT ($p>0.06$)

4.3.3 Delivered GTV doses

The proportion of patients with an absolute change in minimum delivered dose of $\geq5\%$ to any GTV was 10\% (n=3) for Clinical-IGRT, but in only 1 of the 3 cases was the deviation a decrease in dose. The proportion of patients with an absolute change in minimum delivered dose of at least 5\% to any GTV was 10\% (n=3) for 4D-IGRT and 6.7\% (n=2) for DIR-IGRT, and all these deviations were decreases in dose. The patients with decreases in minimum GTV dose are described in detail below.

With Clinical-IGRT one patient with two GTVs had deviations in the minimum delivered dose of -4\% (GTV1) and -14\% (GTV2) (previously shown to have substantial abdominal deformation in Chapter 3, Figure 3.5). The magnitudes of the GTVs’ dose deviations were reduced to -1\% and -7\% with 4D-IGRT, or to -2\% and -6\% with DIR-IGRT for the same respective GTVs. The residual mean GTV errors averaged over both GTVs (and max to any GTV) were 6.3 mm (10.5 mm for GTV2) with Clinical-IGRT, 1.5 mm (5.7 mm for GTV2) with the 4D-IGRT and 0 mm (4.2 mm for GTV2) with DIR-IGRT, all in the SI direction. For the clinical treatment (i.e. Clinical-IGRT) the image-guidance goal as noted in the patient’s chart was specifically biased towards the smaller, superior GTV (GTV1). In this patient the advanced IGRT strategies reduced the dose deviations but were unable to fully correct the deformation.

For two patients following the Clinical-IGRT, deviations in the minimum delivered GTV(s) ranged from -3\% to 13\% (i.e. there were no decreases of 5\% or more):

The first patient with three GTVs had dose deviations of -2\% (GTV1), -2\% (GTV2) and -3\% (GTV3) following the Clinical-IGRT. These dose deviations changed to -7\%, -2\% and 1\% with 4D-IGRT, and to -6\%, -2\% and 1\% with DIR-IGRT, for the same GTVs respectively. The residual mean GTV errors (averaged over all GTVs) were 4.0 mm with Clinical-IGRT, 1.8 mm
with the 4D-IGRT and 0.0 mm with DIR-IGRT, all in the AP direction. However because of liver deformation and rotation the change in residual error for individual GTVs was variable as it was not possible to simultaneously align all GTVs. For example GTV1, which had the largest dosimetric difference between Clinical-IGRT and DIR-IGRT, had a residual mean AP error of 0.7 mm with Clinical IGRT (dose deviation = -2%), which increased to 3.3 mm with DIR-IGRT (dose deviation = -6%). Whereas for GTV2 and 3 the residual mean AP errors were 3.9 and 7.4 mm with Clinical IGRT (dose deviations = -2%, -3%) which decreased to 0.1 and 3.4 mm with DIR-IGRT (dose deviations = -2%, -1%). Therefore, in this patient the advanced IGRT strategies increased the residual geometric error of GTV1 in an attempt to align all GTVs. However, for the clinical treatment (i.e. Clinical-IGRT) the image-guidance goal as noted in the patient’s chart was specifically biased to the largest GTV (GTV1). Had DIR-IGRT also solely aligned GTV1 as was done clinically, the resulting dose deviations would have been -2% (GTV1), -1% (GTV2) and 0% (GTV3).

For the second patient the initial PTV dose-coverage for this solitary GTV was compromised at planning in order to spare the large bowel. The GTV had a residual mean error of 2.3 mm SI in the opposite direction off the under-dosed PTV region resulting in a dose deviation of 13% (i.e. an increase in dose) following the Clinical-IGRT (this patient was also shown in Chapter 3, Figure 3.3). The residual GTV error was reduced to 1.3 mm SI with 4D-IGRT (dose deviation = -6%) and 0.0 mm with the DIR-IGRT (dose deviation = -1%). In this patient the advanced IGRT strategies reduced the residual geometric GTV errors and their dosimetric impact, and reduced the absolute magnitude of the deviations between the delivered dose and the 4D CT-predicted dose. However, the residual dosimetric impact of breathing motion, and inter-fraction deformation resulted in a delivered dose decrease after either of the advanced IGRT strategies. A review of the treatment chart (i.e. for Clinical-IGRT) did not reveal any specific image-guidance instructions for this patient.

4.3.4 Delivered normal tissue doses

Dose changes to non-liver, dose-limiting organs were investigated in detail. ‘Dose-limiting’ was defined as cases where the prescribed dose was reduced from what was allocated by the liver-NTCP model, or where PTV dose coverage was compromised to spare normal tissue. Organs were identified as such if they were planned to receive doses within 2 Gy of the maximum criteria on the static plan exhale 4D CT dose distribution, or flagged as ‘dose-limiting’ in the
patients’ charts. With the Clinical-IGRT, 10 dose-limiting organs in 9 patients had delivered dose deviations of at least 5% (range: -38% to 10%), and 5/10 deviations were dose increases. With the 4D-IGRT, 9 dose-limiting organs in 7 patients had delivered dose deviations of at least 5% (range: -21% to 8%), and 3/9 deviations were doses increases. With DIR-IGRT, 11 dose-limiting organs in 9 patients had dose deviations of at least 5% (range: -21% to 8%), and 3/11 deviations were dose increases.

The absolute delivered normal tissue doses did not exceed any of the planning dose constraints by more than 1 Gy for any strategy, with one exception. Relative to the 4D CT-predicted dose, for one patient’s duodenum the delivered dose increased by 1.2 Gy (4%) with Clinical-IGRT. With 4D-IGRT the esophagus increased by 0.7 Gy (2%) relative to the 4D CT-predicted dose. With DIR-IGRT the esophagus decreased by -0.3 Gy (-1%) relative to the 4D CT-predicted dose.

4.3.5 Relative impact of geometric errors for each strategy

To better understand the causes of the deviations between the delivered and 4D CT-predicted doses for all 10 tissues (i.e. GTV, all normal tissues), the mean impact of the various geometric errors was evaluated (Figure 4.4). For Clinical-IGRT the dominant source of the dose deviations (depicted as the length of longest colored bar for each tissue in Figure 4.3) was the residual rigid liver setup errors (navy bars) for 7 tissues, and abdominal deformation (red bars) for 3 tissues. For both 4D-IGRT and DIR-IGRT, the dominant source of the dose deviations was abdominal deformation (red bars) for 9 tissues, and breathing variations for 1 tissue (orange bars).

For the Clinical-IGRT, breathing variations caused mean absolute dose deviations of 0.9%, averaged over all tissues (for each tissue, the mean impact ranged from 0.2–1.7%). For both the 4D-IGRT and DIR-IGRT, breathing variations caused mean absolute dose deviations of 0.8%, averaged over all tissues (for each tissue, the mean impact ranged from 0.2–1.6%). This indicates that breathing variations between the treatment 4D CBCT and the planning 4D CT impacted the delivered dose by the least magnitude, and the magnitude was fairly consistent across the IGRT strategies. This is visualized in Figure 4.4 by the orange bars having, in general, a consistent size for each IGRT strategy.

For the Clinical-IGRT, abdominal deformations caused mean absolute dose deviations of 1.2%, averaged over all tissues (for each tissue, the mean impact ranged from 0.3–2.5%). For both the
4D-IGRT and DIR-IGRT, abdominal deformations caused mean absolute dose deviations of 1.2%, averaged over all tissues (for each tissue, the mean impact ranged from 0.3–2.4% for 4D-IGRT, and from 0.4-2.4% for DIR-IGRT). This indicates that abdominal deformations have an equivalent dosimetric impact for each IGRT strategy when averaged over all tissues. However, when considering the duodenum alone, abdominal deformation caused mean absolute dose deviations of 2.5% for Clinical-IGRT, 2.1% for 4D-IGRT and 1.9% for DIR-IGRT. This indicates that the more advanced IGRT strategies reduced the dosimetric impact of the deformation on the duodenum compared to Clinical-IGRT, although the geometric magnitude was unchanged. Conversely, when considering the esophagus alone, abdominal deformation caused mean absolute dose deviations of 0.6% for Clinical-IGRT, 0.8% for 4D-IGRT and 0.9% for DIR-IGRT. This indicates that the more advanced IGRT strategies increased the dosimetric impact of the deformation on the esophagus compared to Clinical-IGRT, although the geometric magnitude was again unchanged.

For the Clinical-IGRT, residual rigid setup errors caused mean absolute dose deviations of 1.4%, averaged over all tissues (for each tissue, the impact ranged from 0.3–2.5%). This was the most dominant effect on the delivered doses for Clinical IGRT. For the 4D-IGRT, rigid setup errors caused mean absolute dose deviations of 0.2%, averaged over all tissues (for each tissue, the impact ranged from 0–0.4% for 4D-IGRT). For the 4D-IGRT, rigid setup errors caused mean absolute dose deviations of 0.7%, averaged over all tissues (for each tissue, the impact ranged from 0.2–1.7% for 4D-IGRT). Residual rigid setup errors had the smallest effect on the delivered doses for 4D-IGRT and DIR-IGRT, as was expected after directly targeting the liver or GTV(s) respectively with translational corrections. The impact of residual rigid setup errors was smallest for the 4D-IGRT strategy (visualized as having the shortest navy bars in Figure 4.4) because the setup error for this specific dosimetric analysis was defined about the liver’s position.

4.4 Discussion

This study compared image-guidance strategies based on rigid liver-to-liver registration of 3D CBCT (Clinical-IGRT), 4D CBCT (4D-IGRT), or direct tumor-targeting enabled through DIR of 4D CBCT (DIR-IGRT). Compared to the Clinical-IGRT strategy, 4D-IGRT based on liver alignment to the exhale 4D CBCT reduced the residual GTV errors by approximately half, resulting in population random and systematic GTV errors ≤1.4 mm. If biomechanical DIR and
4D CBCT were available for online guidance, as was simulated with the DIR-IGRT strategy, the GTV could be targeted directly reducing the residual population GTV errors to <1 mm. Compared to the Clinical-IGRT, the advanced 4D- and DIR-IGRT strategies both improved the dosimetric correlation between predicted 4D CT-dose distribution incorporating breathing motion, and the delivered dose distribution incorporating inter-fraction geometric errors. Compared to Clinical-IGRT, the biggest overall geometric and dosimetric IGRT improvement came from the use of 4D CBCT with only a modest further gain with the addition of DIR.

**Geometric impact of IGRT**

To simulate the 4D- and DIR-IGRT strategies, translations were applied to 4D CBCT model to align the exhale liver or GTV’s centre-of-mass for the two strategies respectively. Clinical factors such as normal-tissue sparing, or targeting a specific GTV (e.g. the largest tumor burden) were not considered for the 4D- and DIR-IGRT strategies. Deformations and rotations in the liver also prevented alignment of multiple GTVs simultaneously, even with DIR-IGRT. Robotic treatment couches with movement in 6 degrees of freedom are able to correct rotations[144]. These couches are not currently used for liver SBRT, however their application could potentially further reduce residual GTV translational errors.

The residual GTV errors observed after Clinical-IGRT could have resulted from the radiation therapists intentionally biasing the 3D CBCT registration to spare critical normal tissues planned to receive high doses. The assumption of this technique is that displaced GTVs would have still received the planned dose if they remained inside the PTV. Whether this occurred or not is often not documented in the patient’s chart. To help determine if the GTV errors were intentional, or a result of poor targeting (e.g. due to a lack of 4D CBCT and DIR) the nine patients with systematic GTV errors ≥3 mm in any direction between exhale 4D CT and exhale 4D CBCT were examined in detail:

- Three patients had single GTVs with mean errors ≥3 mm in any direction. One patient had a 3.5 mm LR error, one had a 4.1 mm SI error, and all three had errors of 3.5–5.2 mm in the AP direction. For all three patients the adjacent large bowel was dose-limiting at planning, and it had mean displacements of 1.1–1.9 mm LR, 0–0.6 mm AP and 0.6–2.4 mm SI. There were no specific IGRT guidelines to avoid the bowels. As the bowel
displacements were generally smaller in magnitude than the GTVs however, it is possible the GTV errors resulted from bowel avoidance in these three patients.

- The six other patients had mean GTV errors ≥3 mm in any direction (3 with single GTVs, 3 with multiple GTVs). The GTV errors were all ≤3 mm LR, three patients had AP errors of 3.4–5.2 mm, and four patients had SI errors of 3.2–9.6 mm. Two of the six patients had single GTVs with no adjacent dose-limiting normal tissues, therefore it is possible these GTV errors resulted from targeting issues. Four of the six patients had GTVs (one with a single GTV, three with multiple GTVs) that were adjacent dose-limiting gastrointestinal tissues. There was specific IGRT guidelines to bias the registration towards one specific GTV (2 patients), and/or to avoid the gastrointestinal tissues (3 patients). In all four patients the mean gastrointestinal tissue displacements were ≥3 mm, up to a maximum of 6.1 mm LR, 4.5 mm AP and 8.3 mm SI, and they all occurred in the same direction as the GTV errors. Therefore it is possible the GTV errors resulted from targeting issues in these four patients.

Compared to the Clinical-IGRT strategy, 4D-IGRT reduced the magnitude of residual population GTV errors by nearly a half. Respiratory-correlated (4D) CBCT is achieved by processing the raw CBCT data and it does not require equipment different from 3D CBCT. Artifacts and uncertainties in the sorted 4D phases, caused for example by sorting errors secondary to breathing irregularities[93], are visualized in the images themselves. Compared to complexity potentially introduced with online DIR-based guidance, using 4D CBCT guidance with rigid registration is relatively simple and greatly reduces residual GTV errors. Since 2012, 4D CBCT has been implemented clinically at PMH for liver SBRT for patients with a breathing amplitude >5 mm, and is based on rigid liver registration to the exhale 4D phase.

This study assumed perfect targeting about the GTV’s centre-of-mass with the DIR-IGRT strategy. Other uncertainties remain due to residual deformation, breathing variations, intra-fraction motion, and the accuracy of the DIR algorithm itself (~2 mm[124]). Therefore, even directly targeting the GTV would still require a PTV to account for these uncertainties. Clinical implementation of DIR-IGRT requires fast and accurate DIR. The current study used multi-organ DIR to accumulate the delivered dose. However using DIR on the liver alone to track the GTV on CBCT could be calculated in <1 minute excluding contouring time[103]. Nguyen et al. developed a technique that adapts a patient-specific motion model created with DIR of 4D CT to
2D treatment images such as fluoroscopy[141]. The accuracy in predicting the liver GTV position was similar to full biomechanical DIR, potentially eliminating the requirement for contouring for online biomechanical DIR-based targeting of liver tumors.

Non-DIR strategies relying on GTV surrogates are widely used for liver IGRT. Wunderink et al. demonstrated that when fiducials are implanted 37–103 mm from liver tumors and multiple GTVs are simultaneously targeted, the resulting individual GTV errors could be greater than guidance based on diaphragm imaging[102]. Directly targeting liver tumors without DIR may result from ongoing research aimed at integrating MRI systems with linear accelerators[145]. A prototypic system is currently being developed at PMH combining a conventional CBCT-equipped linac with an MRI-on-rails[146]. The planned specification of this project is to perform MRI-based IGRT, robotically re-position the patient between the MRI bore and the linac, and being SBRT delivery within 90 seconds.

**Dosimetric impact of IGRT**

For all image-guidance strategies, decreases in the delivered minimum tumor doses of more than 5% relative to the 4D CT-predicted doses occurred infrequently (<5% of patients). This was expected given that the minimum PTV expansion was 5 mm in these patients. The range of minimum GTV dose deviations with the Clinical-IGRT (-14% to 13%) was reduced with the 4D-IGRT (-7% to 4%) and DIR-IGRT (-6% and 4%). With smaller PTVs, the dosimetric improvement in GTV doses with the advanced strategies would likely be more pronounced. The range of delivered normal tissues dose deviations >5% with the Clinical-IGRT (-38% to 10%) was reduced with the 4D-IGRT (-21% to 8%) and DIR-IGRT (-24% and 8%). Overall, significantly improved correlations between the 4D CT-predicted doses and the delivered doses were observed with the advanced IGRT strategies.

The dosimetric impact of residual setup errors were reduced for all tissues as was hypothesized to occur with the advanced IGRT strategies compared to the Clinical-IGRT. Although the geometric magnitude of deformations and breathing variations did not change, the dosimetric impact changed as tissues were translated to different locations within the dose matrices and onto different dose gradients. For example, the advanced IGRT strategies increased the dosimetric impact of deformations in the esophagus and small bowel (Figure 4.4). As residual
setup errors are reduced with improved IGRT, the consideration of residual deformation becomes of greater importance.

A potential solution to the difficulty in targeting multiple GTVs simultaneously when liver deformation is present would be to treat each GTV individually, although this would substantially increase the workload. Also, the potential risk of intra-fraction overlap of the high-dose regions from each individual GTV plan resulting in accumulated over-doses in the overlap region requires further investigation. Alternatively, daily IMRT re-optimization of the beam orientation and fluences has been proposed to adaptively account for the abdominal deformation occurring in liver SBRT and improve delivered doses[147]. Again, the increased workload likely inhibits clinical implementation at present. Both strategies could potentially benefit from deformable dose reconstruction during the course of SBRT to evaluate whether the overall accumulated doses are clinically acceptable.

### 4.5 Conclusion

Compared to IGRT based on rigid liver targeting using 3D CBCT, the addition of 4D CBCT reduced the population residual GTV errors by half. The dosimetric correlation between the planning 4D CT-predicted breathing dose and the delivered dose modeled with deformable dose reconstruction was improved with the 4D imaging. The addition of DIR to the 4D strategy to directly target liver GTVs, and overcome the limitation of the liver as a surrogate when deformation exists, further improved the accuracy of SBRT delivery. These strategies may allow for PTV reduction and dose-escalation. With the advanced IGRT strategies however, approximately half of liver SBRT patients still have dose deviations >5% caused by residual deformations and breathing variations. In the minority of patients where tumors are being under-dosed or normal tissues over-dosed, adaptive radiotherapy interventions facilitated by prospective deformable dose reconstruction could be explored to improve the quality of SBRT.
Chapter 5
Preliminary Validation of Biomechanical-Based Deformable Dose Reconstruction
5 Preliminary Validation of Biomechanical-based Deformable Dose Reconstruction

5.1 Introduction

The overall goal of all DIR algorithms is to model the complex anatomical and physiological behaviour of patients occurring between imaging points. These models require validation prior to implementation to understand the scale and complexity of the uncertainties related to these models that will be introduced into the clinical environment. The majority of published DIR studies have focused on geometric comparisons of features between the DIR-predicted images and the actual target images. Common evaluation metrics include volume overlap measures (e.g. of organs, tumors) which do not quantify the structure’s internal accuracy, or point-based registration errors (e.g. of anatomical points, implanted fiducials) which only offer a small sample of measurements in sites such as the liver.

In the context of dose accumulation, the delivered doses modeled via DIR should ideally be compared to physical measurements directly related to dose. Deformable dosimeters that can track dose in 3D have been investigated for this application. Yeo et al. developed a tissue-equivalent 3D dosimeter containing a gel matrix with a radiochromic dye that polymerizes after radiation[148]. The 3D change in optical density representing absorbed dose can be reconstructed with a specialized scanner, called optical CT, which replaces the ionizing radiation used in standard CT with a visible light source. Eleven intensity-based DIR algorithms had varying degrees of accuracy in predicting the deformed delivered dose inside the gel with an overall trend of worsening DIR performance for larger deformation magnitudes[149].

Gel dosimeters were also investigated by Niu et al. with the dose distribution imaged using MRI, which was possible because the radiation-initiated polymers affect water molecule mobility in the gel[136]. The dosimeter was repeatedly deformed under a piston-like device mimicking 4D respiratory motion under irradiation. The delivered dose distribution predicted with Morfeus was compared to the measured dose on MRI, revealing >95% of the voxels were accurately predicted within the measurement uncertainty (4.7%) and within a 3 mm agreement range. Niu et al. used a simple centrally located spherical dose distribution, and applied a contact-surface model[150] between the piston and the gel. Further investigation is warranted using more intricate dose-distributions which may better relate to clinical IMRT plans, and to
evaluate biomechanical-DIR relying on the guided-surface projections which are used throughout this thesis.

The primary purpose of this study was to further evaluate the accuracy of biomechanical DIR-based dose reconstruction. This was done using acquired imaging data from a novel plastic (non-gel) dosimeter from Duke University. A secondary aim was to compare the dosimetric results between biomechanical DIR-models using either guided-surface projections, or contact-surfaces.

5.2 Methods

5.2.1 Dosimeter characteristics and experimental setup

A novel 3D deformable dosimeter was previously developed and characterized by Juang et al. and is summarized here[151]. It is a modification of the commercially-available Presage dosimeter (Heuris Inc., Skillman, USA). The elastic 60×47.5 mm cylinder consists of a transparent polyurethane matrix doped with a light-absorbing radiographic leuco dye. The dosimeter is water-equivalent and the leuco dye exhibits a linear change in optical density with absorbed dose. The mechanical properties are similar to biological tissue with a Young’s Modulus (E) of 13.5–887 kPa and a Poisson’s ratio (ν) of 0.475. Young’s Modulus represents the material stiffness and is defined by the ratio of stress over strain, along the direction of the applied stress. Poisson’s ratio represents the material compressibility and is defined by the ratio of the percent-extension normal to the applied stress, over the percent-contraction in the direction of the stress.

Previous CT images of the dosimeter were acquired in its original cylindrical state, and following a 16 mm (27%) lateral compression between two plates (Figures 5.1.A, 5.1.B). A checkerboard radiation dose pattern comprised of twenty nine 5×5 mm² fields was planned on the CT image (Figure 5.1.C) in the compressed geometry, and the dosimeter was then irradiated while remaining compressed. The planned dose was calculated with the Eclipse (Varian Medical Systems, Palo Alto, USA) treatment planning system at a 1×1×1.25 mm³ resolution. After removing the plates, the dosimeter returned to its un-compressed geometry and the measured delivered dose distribution, now deformed, was reconstructed using an optical CT system with a 1 mm isotropic resolution[152]. All dosimeter characterization, irradiation and imaging for the
experiment was carried out for a previous investigation by Titania Juang, Shiva Das and Mark Oldham at Duke University Medical Center, Durham, USA[151]

Figure 5.1. A. CT images of the dosimeter with and without the compression applied (with the plates digitally removed). B. Model of the experimental setup. C. The dose distribution planned on the CT in the compressed geometry.

5.2.2 DIR models

For this current investigation the CT images were first registered using Morfeus, a biomechanical model-based DIR algorithm. Morfeus establishes a correspondence between images using guided-surface projections between surfaces of contoured structures (e.g. organs). This describes the deformation and displacement at the surface which serves as boundary conditions for the finite element analysis. The analysis uses linear-elastic material properties and
determines the deformation and displacements of the internal nodes, comprising each element. For this study, surface projections were applied between the uncompressed and compressed dosimeter contours and material properties of $E=0.95$ kPa and $\nu=0.480$ were assigned. Varying the $E$ assigned had no effect the nodes’ displacements because the dosimeter is modeled as a homogenous material. Varying the $\nu$ from 0.17 to 0.499 changed the nodes’ average absolute displacements by $<0.5$ mm, indicating the model was insensitive to the exact $\nu$ applied. Morfeus has been validated in several anatomic sites (e.g. accuracy within the liver is 2 mm[124]). Hereafter this model is called MorfeusSurfProj.

A potential limitation of using guided-surface projections to register anatomy is they may not accurately reflect the true tissue biomechanics at the surfaces. This approximation is made for anatomical simulations as it is often difficult to model the forces that are acting on the organ. Due to the relatively simple experimental setup (i.e. compression by two plates normal to the dosimeter’s surface) it was hypothesized that accurate registration could also be achieved by modeling the interactions of the plate with the cylindrical gel. Therefore, a modified DIR model was investigated consisting of the uncompressed gel atop a baseplate, and between two rigid side plates spaced 5 cm away (as shown in Figure 5.1.B). The side plates were then moved inwards by 5 cm into their final position seen on the compressed CT. In this model the plates served as boundary conditions, and contact surfaces at the dosimeter-plate interfaces allowed small sliding. For this model Poisson’s ratio was optimized to maximize the agreement of the dosimeter’s external contour between the DIR-prediction and the actual CT image (uncompressed geometry), resulting in $\nu$ of 0.499. Hereafter, this model is called MorfeusPlates. For both DIR models the CT image of uncompressed dosimeter was deformed into the CT of the compressed dosimeter.

To investigate whether the results of using guided surface projections in DIR are affected by the magnitude of deformation, MorfeusSurfProj was compared to three known deformations generated by the MorfeusPlates model. These included modeling the original dosimeter compressed by the full amount with the plates (equal to the actual physical compression used in the experiment, Fig. 5.1.B), then by 2/3 of the full compression, and finally by 1/3 of the full compression. MorfeusSurfProj was then performed using the original surface of the uncompressed dosimeter and each deformed surface generated by MorfeusPlates (full, 2/3, or 1/3 compression). The node-by-
node displacements throughout the entire dosimeter between Morfeus$_{\text{SurfProj}}$ and Morfeus$_{\text{Plates}}$ were directly compared for each compression level.

5.2.3 Evaluation and dose warping

The deformed CT images from DIR were first compared to the actual CT image of the compressed dosimeter. The distance-to-agreement (DTA) of the external contour between the predicted and the actual image was calculated. The mean and overall maximum residual (i.e. after DIR) vector displacements between the surfaces were reported. The volume of overlap of the compressed dosimeter’s external contour between the deformed image and the actual image was also quantified using the Dice similarity coefficient (DSC):

$$DSC = \frac{2(\text{predicted dosimeter} \cap \text{actual dosimeter})}{\text{predicted dosimeter} + \text{actual dosimeter}}$$  \hspace{1cm} (1)

DIR-predicted deformed dose distributions (un-compressed geometry) were created as follows. DIR provides a deformation vector field throughout the dosimeter volume, which is a set of irregularly spaced vectors for each node in the model from the un-compressed to the compressed geometries. The vector field is first inverted to give displacements from the compressed (geometry of the Eclipse dose grid) to the un-compressed geometries. This inverted vector field is interpolated onto an isotropic 1 mm$^3$ grid, equal to the optical CT volume in the un-compressed geometry. The dose at each node from the compressed space (calculated in Eclipse) was mapped back to the new grid in the un-compressed space along the inverted vectors. Tri-linear interpolation was used to map the nodes onto the dose grids.

These DIR-predicted doses were compared to the measured distribution of optical density changes (a surrogate for dose) on optical CT (un-compressed geometry). All distributions were first rigidly registered together according to the best visual alignment of the deformed fields. These three volumes were subsequently all cropped by a volume equal to the un-compressed dosimeter contour on CT that was contracted by 4 mm. Contraction by 3–5 mm is common for this dosimetry method to avoid the lost data due to edge artifacts inherent with the optical CT[153,154].

The relative dose distribution differences were compared using the following measurements. Each deformed field was auto-segmented (Pinnacle$^3$ v9.0, Philips Medical Systems, Madison, USA) using a threshold equal to half of the maximum intensity change across the field (Figure
This was performed at three coronal cross-sections corresponding to depths of 12, 23 (depth at maximum compression) and 34 mm relative to the incident radiation beams. Field locations were compared between the DIR-prediction and optical CT using the centroid of each segmented field. Field shapes were compared by measuring the segmented field’s full width at half the maximum value (FWHM) in the horizontal and vertical directions (Figure 5.2). These metrics provide an analysis of the relative distributions independent of how the optical CT data is converted and normalized into dose. The measurement error is ~1 mm (the optical CT resolution). These metrics and measurement points are equivalent to Juang’s work[151], and were intentionally reproduced here to facilitate a comparison between the studies.

Figure 5.2. A. Each deformed field was auto-segmented at half the maximum value of the intensity profile across that field. B. The full widths at half the maximum (FWHM) values were measured in the horizontal and vertical directions to quantify the field shape (optical CT is shown).

The absolute dose distribution differences were compared as follows. Optical CT data was converted to dose using the linear response calibration previously reported[151]. The mean dose from the converted optical CT data was normalized to the mean dose for each DIR-predicted dose distribution. Normalization is required to account for slight inter-dosimeter response differences (e.g. due to formulation, temperature during irradiation etc.) between the calibration batch and subsequent dosimeters[153]. The optical CT and DIR-predicted distributions were compared using voxel-by-voxel dose differences. Additionally each voxel in the DIR-predicted distribution was tested against the optical CT distribution to fall within specified acceptance criteria, and the 3D passing rate over all voxels, termed the gamma (γ) index, was reported[155].
The commonly used criteria of 3% dose difference and 3 mm DTA was used ($\gamma_{3\%/3\text{mm}}$) and calculated in the Computational Environment for Radiotherapy Research (CERR) (Washington University, St. Louis, USA). The reference $\gamma_{3\%/3\text{mm}}$ for a non-deformed control dosimeter was previously reported as 96%[151].

5.3 Results

The DSC of the compressed dosimeter contours between the DIR-predicted image and the actual image was 0.988 for Morfeus\textsubscript{Plates} and 0.994 for Morfeus\textsubscript{SurfProj} (DSC=1.0 indicates perfect overlap). The mean (maximum) surface DTA was 0.8 (2.7) mm for Morfeus\textsubscript{Plates} and 0.3 (1.9) mm for Morfeus\textsubscript{SurfProj}. High DSC and low DTA values indicate excellent registrations of the dosimeter’s external boundary. This was expected for Morfeus\textsubscript{SurfProj}, indicating the surface alignment prescribed by the technique was executed correctly. For Morfeus\textsubscript{Plates} this indicates that modeling the plate-dosimeter interaction successfully generated an accurate deformed position. The displacements generated by Morfeus\textsubscript{Plates} and Morfeus\textsubscript{SurfProj} were directly compared at varying levels of dosimeter deformation. The mean±SD node-by-node differences were 1.5±0.8 mm for the full compression (the actual compression used in the physical experiment), 0.9±0.6 mm for 2/3 compression, and 0.5±0.4 mm for 1/3 compression. Figure 5.3 shows the distribution of displacement differences between the DIR models as a function of the deformation level.

Figure 5.3. Cumulative frequency distribution of the node-by-node displacement differences between Morfeus\textsubscript{SurfProj} and Morfeus\textsubscript{Plates}. 
5.3.1 Comparison with the relative optical CT distribution

The deformation map from DIR was then applied to the original dose grid (compressed geometry) to predict the deformed dose distribution (un-compressed geometry). This allowed for comparison to the measured distribution of optical density changes on optical CT (un-compressed geometry) in Figure 5.4. Differences in the deformed checkerboard fields’ centroid location and shape (i.e. FWHM) are reported in Table 5.1. It was not possible to quantify the centroid and FWHM measurements for 8 of the 29 fields due to their partial beam incidence and the necessary optical CT cropping, leaving 63 points (21 per cross-section) for evaluation.

Figure 5.4. Relative deformed dose distributions in the dosimeter’s uncompressed geometry measured with optical CT (represented as change in optical density), or predicted using either DIR model. Central cross-sections are shown in the coronal (top), axial (middle), and sagittal (bottom) planes.
Table 5.1. Field location and shape errors (in mm) between the dose distributions predicted using two DIR models, each compared to the measured distribution of optical density changes measured via optical CT (predicted–measured).

<table>
<thead>
<tr>
<th>Cross-section depth</th>
<th>Value</th>
<th>Δ Field location centroid</th>
<th>Δ Field shape</th>
<th>Δ Field location centroid</th>
<th>Δ Field shape</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Horizontal FWHM</td>
<td>Vertical FWHM</td>
<td>Horizontal FWHM</td>
<td>Vertical FWHM</td>
</tr>
<tr>
<td>12 mm</td>
<td>Mean±SD</td>
<td>0.9±0.5</td>
<td>-0.1±0.6</td>
<td>-0.4±0.7</td>
<td>1.0±0.5</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>0.0, 1.7</td>
<td>-1, 1</td>
<td>-2, 1</td>
<td>0.0, 2.3</td>
</tr>
<tr>
<td>23 mm</td>
<td>Mean±SD</td>
<td>1.0±0.5</td>
<td>0.1±0.6</td>
<td>-0.2±0.8</td>
<td>1.0±0.5</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>0.0, 1.6</td>
<td>-1, 1</td>
<td>-2, 1</td>
<td>0.3, 1.9</td>
</tr>
<tr>
<td>34 mm</td>
<td>Mean±SD</td>
<td>0.9±0.4</td>
<td>0.1±0.9</td>
<td>-0.4±0.6</td>
<td>1.0±0.6</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>0.0, 1.7</td>
<td>-2, 2</td>
<td>-1, 1</td>
<td>0.0, 2.4</td>
</tr>
<tr>
<td>Overall</td>
<td>Mean±SD</td>
<td>0.9±0.5</td>
<td>0.0±0.7</td>
<td>-0.4±0.7</td>
<td>1.0±0.5</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>0.0, 1.7</td>
<td>-2, 2</td>
<td>-2, 1</td>
<td>0.0, 2.4</td>
</tr>
</tbody>
</table>

Abbreviations: FWHM=full width at half the maximum value; SD=standard deviation. Notes: the FWHM measurement error is ~1 mm (the resolution of the optical CT data).

Errors of similar magnitude were observed between the cross-section corresponding to the plane of maximum deformation/compression (depth=23 mm) and the remaining cross-sections. For each DIR model the overall mean field centroid errors were ≤1 mm. For MorfeusPlates the maximum field centroid errors were 1.7 mm, occurring in 4 of 63 (6%) of the measurements. For MorfeusSurfProj the maximum field centroid error was 2.4 mm and errors ≥2 mm occurred in 4 of 63 (6%) of the measurements (1 of these 4 errors occurred in the same location as the largest error for MorfeusPlates). For both DIR models, the overall average absolute FWHM errors were ≤1 mm in both the horizontal and vertical directions. The horizontal FWHM errors varied from 2 mm narrower to 2 mm wider with MorfeusPlates, and from 3 mm narrower to 3 mm wider with MorfeusSurfProj.

5.3.2 Dosimetric evaluation

Optical CT data was then converted from optical density change to dose. Compared to the optical CT dose, the mean±SD voxel-by-voxel dose difference was 0.2±4.7 Gy (9±39%) for MorfeusPlates, and 0.1±4.2 Gy (8±36%) for MorfeusSurfProj. Figure 5.5 shows central dose profiles in the dosimeter, with good overall agreement between both DIR models to the optical CT. The
Figure 5.5. Dose profiles along a line bisecting the central row of deformed dose fields at varying coronal cross-sections.
largest voxel-by-voxel dose differences for both DIR models occurred throughout the whole dosimeter along the field edges where the dose gradients are steepest. The $\gamma_{3%/3\text{mm}}$ passing rate over the whole volume was 90% for Morfeus$^\text{Plates}$ and 91% for Morfeus$^\text{SurfProj}$. Without either DIR model, the $\gamma_{3%/3\text{mm}}$ for the original planning dose grid (compressed geometry) rigidly registered to the delivered optical CT distribution (uncompressed geometry) was only 58%.

5.4 Discussion

Progress towards validation of biomechanical DIR-based dose accumulation was achieved using data from a deformable dosimeter and optical CT system, complementing a previous study using MRI-based deformable dosimeters. By tracking the deformed 3D dose distribution direct dosimetric evaluation was possible at >60 measurement points. This is in contrast to anatomic or other imaging features commonly used for geometric validation, which are often sparse or non-existent on images of homogenous structures. Two biomechanical DIR algorithms were evaluated. The first, Morfeus$^\text{SurfProj}$, was identical to the standard implementation of Morfeus and used a guided-surface projection between the dosimeter’s surfaces on each image to serve as boundary-conditions (i.e. constraints) for the deformation model. The second, Morfeus$^\text{Plates}$, directly modeled the motion of two plates compressing the dosimeter, which more realistically reflects the mechanics of the experimental setup. The accuracy of both models to predict the deformed dose fields’ location and shape were $\leq 1$ mm on average compared to the measured distribution read from optical CT with a 1 mm resolution.

Juang et al. originally used the imaging data to evaluate a commercially available, intensity-based DIR platform called Velocity AI (Velocity Medical Solutions, Atlanta, USA) based on a B-spline model with mutual information [151]. For Velocity, the external dosimeter contour was registered with sub-millimeter accuracy. The internal mean (maximum) dose field centroid errors were 4.2 (9.0) mm, the mean (range) dose field shape errors (i.e. horizontal FWHM) were 3.8 (-6, 14) mm, and the $\gamma_{3%/3\text{mm}}$ was 60% [151]. Those DIR errors are substantially worse than either biomechanical-based DIR models evaluated presently. Both results are contrasted in Figure 6. Velocity’s geometric accuracy has been reported elsewhere to range from 1–3 mm on average, similar to the image slice thicknesses in those studies [156,157]. Notably they all used patient derived CT images containing high-contrast, heterogeneous image intensities with prominent anatomic features unlike the homogeneous CT of the dosimeter.
To study DIR inconsistencies with varying contrast levels in more detail, Kirby and colleagues developed a single-slice, patient-derived virtual CT phantom alongside a geometrically equivalent physical phantom containing high-resolution optically tracked markers[156]. This allowed simultaneous geometric accuracy evaluation at both high and low contrast regions. Velocity’s error notably increased with deformation magnitude particularly in low contrast areas[158], supplementing the poor results observed in the Juang paper (Figure 5.6.B). Interpreted together, these studies should serve as a caution when extrapolating the accuracy of intensity-based DIRs validated on feature-rich imaging to images or regions that lack prominent features. Unlike intensity-based DIR, the accuracy of biomechanical DIR is largely independent of the image intensities and modalities. This is relevant for liver tumors which are often indistinguishable from the normal liver tissue when intravenous contrast is not routinely used (e.g. during cone-beam CT guidance used in stereotactic radiotherapy).

The results of this work, combined with the previous study by Niu et al.[136] demonstrate that biomechanical DIR is potentially well suited for inter-fraction dose-accumulation in sites like the liver. For biomechanical DIR, the voxel-by-voxel dose differences in this study compared to the ground truth (8±36% for MorfeusSurfProj) are larger than reported in Niu’s study (3±14%)[136]. Whereas the previous work used a central spherical dose distribution, this study used a checkerboard field pattern resulting in a higher proportion of field edges throughout the dosimeter. Dose difference errors (e.g. due to registration error) are particularly sensitive at
these edges where dose gradients are steepest and likely contribute to the difference between the studies. The 3D voxel passing rate ($\gamma_{3\%/3mm} = 91\%$ for MorfeusSurfProj) is comparable to Niu’s study ($\gamma_{4.7\%/3mm} = 97\%$) which used a more relaxed acceptance criteria to reflect the higher level of uncertainty in the dosimetric measurement accuracy of the gel dosimeter.

The mean displacement difference between MorfeusSurfProj and MorfeusPlates was 1.5 mm when the dosimeter was compression laterally by 16 mm (27%), and <1 mm for smaller magnitudes of simulated deformations. Although there was a trend for greater disagreement at larger deformation levels (Fig. 5.3), some organs such as the liver rarely deform to that extreme[97]. Therefore, MorfeusSurfProj appears suitable for clinical application of biomechanical-DIR in the liver where it is often not practical to directly model the forces acting on liver’s surface.

Both DIR models predicted the measured optical CT distribution well overall as seen in Figures 5.4 and 5.5. Lateral shifts of the dose profile of 1–2 mm (e.g. voxels 10–30, Figure 5.5) are likely caused by DIR errors. Other mismatches could be caused by out-of-plane registration errors or uncertainties in the optical CT data. For example, the lone spike in optical CT dose >35 Gy for the field centered on voxel ~38 (Figure 5.5, 12 mm depth) is 6% higher than the maximum point dose on the entire original dose grid from the planning system. Such a difference in maximum dose could be conceivably caused by a substantial change in the dosimeter shape between CT imaging and irradiation (unlikely because the compression plates were unchanged between these time points), or inherent artifacts caused by the dosimeter material. Another mismatch occurs at the field centered on voxel ~56 (Figure 5.5, all depths) where the optical CT peak dose is 27% lower than predicted by DIR. This could be caused by standard optical CT artifacts at the edge of the dosimeter (although all analyses were done 4 mm from the dosimeter edge), or by a field that was only partially incident on the dosimeter. The latter would be indicative of a translational setup error during irradiation.

The physical experiment could be considered an end-to-end test and therefore not only dependent on DIR accuracy. Localization of the dosimeter (e.g. with 3D cone-beam CT) was not done before irradiation, and is not physically possible within the optical CT system. Therefore inherent translational errors up to 2 mm can be reasonably expected at these time points. The dosimeter itself developed by Juang et al. is also novel and therefore requires further research to characterize its reproducibility, temporal stability, and dose-response linearity.
Geometric uncertainties in the linear accelerator, typically 1–2 mm, and dosimetric uncertainties, typically <2%, originating from the treatment planning system’s dose model would also cause small delivered dose deviations. Dose calculation engines used in clinical treatment planning systems need to balance accuracy and efficiency. Modeling fields <30×30 mm² are particularly challenging and are common sites of increased dose calculation uncertainties[159]. For example, Gagne et al. reported errors for Eclipse’s dose engine of 5% underneath 5 mm thick multi-leaf collimators and up to 12% at the leaf edges[160]. Ong et al. reported Eclipse overestimates the penumbra width, the distance between 90 and 50% of the peak dose, for 10×10 mm² fields by 0.2–0.4 mm but underestimates the peak dose by up to 2–6%[161]. For fields <10×10 mm², multi-leaf collimator positioning errors of 1 mm can result in field size differences of 30–40%[162]. These are potentially large sources of error for the current study considering field sizes of only 5×5 mm² were used.

Overall, this study accurately reconstructed the delivered dose from a single irradiation session via biomechanical-DIR of CT images. Further research is required to reconstruct the accumulated dose delivered over multiple time points, for example during respiratory-like motion and over multiple fractions. This dosimetry system could also be applied in those scenarios to evaluate uncertainties in the accumulated dose estimates, including those introduced by DIR. While this dosimeter behaves similar to a liver, more sophisticated designs are likely required to evaluate dose reconstruction in anatomy where density and mass are not conserved (e.g. lung, variable stomach filling etc.).

5.5 Conclusions

Biomechanical-based DIR was used to reconstruct the delivered dose distribution via registration between CT images of a homogenous, deformable dosimeter. Accuracy of the deformed distribution was compared to that measured throughout the dosimeter via optical CT. The mean predicted dose field location and shape errors were ≤1 mm, and this did not differ between two DIR models utilizing different boundary conditions. This validation supports the application of biomechanical-model based DIR to reconstruct the delivered radiotherapy dose, particularly for anatomical sites like the liver which lack prominent features in clinical images.
Chapter 6
Simplified Strategies to Determine the Mean Respiratory Position For Liver Radiation Therapy Planning

This chapter was published as: Velec, Joanne L. Moseley, Kristy K. Brock. Simplified strategies to determine the mean respiratory position for liver radiation therapy planning. Practical Radiation Oncology 2014, 4: 160-166.
6  Simplified Strategies to Determine the Mean Respiratory Position for Liver Radiation Therapy Planning

6.1  Introduction

Stereotactic body radiation therapy (SBRT) for liver cancer patients results in better clinical outcomes with higher doses[47,54]. SBRT doses up to 60 Gy in 6 fractions have been delivered without serious liver toxicity using individualized dose allocation[53]. However, this dose is often not achievable due to overlap of the planning target volume (PTV) with the neighboring luminal gastrointestinal structures. The volume of normal tissue irradiation, and thus dose allocation, is governed largely by the technical details in the planning of SBRT.

On average, the liver moves 11 to 25 mm during normal respiration[69]. A common way to incorporate the patient’s specific breathing motion into the PTV is to create an internal target volume (ITV), a composite volume of the tumor on each phase of the planning 4D CT. Unless automated intravenous contrast is synchronized with 4D CT acquisition liver tumors are not well visualized[73], requiring the use of surrogates (e.g. liver, diaphragm etc.) to measure motion. Because the tumor spends only a fraction of time at each position of the breathing cycle, the ITV can be considered unnecessarily large. If the planning CT represents the patient’s geometry in its time-averaged respiratory position, smaller dose-probability PTV margins can be used to achieve a specific clinical goal (e.g. 95% of prescribed dose is received by 90% of patients[163]). ITV-based PTV for lung radiation therapy has been shown to be 33% larger on average than dose-probability margins[76]. Similar gains may be possible in the liver provided the mean position can be established, currently a challenging task.

Wolthaus et al. developed the mid-ventilation CT for lung radiation therapy by reconstructing a single 4D CT phase around the time-percentage closest to the tumor’s time-averaged position[164]. This time-percentage is derived from the tumor or diaphragm motion, itself requiring an initial 10-phase 4D CT reconstruction, or directly using the external respiratory signal from image acquisition[165]. If hysteresis is present, the mid-ventilation CT will not represent the time-averaged position accurately. However, a mid-position CT accounting for hysteresis can be reconstructed, by means of deformable image registration (DIR) and an
average 4D deformation map[166]. A full 4D deformation map from 9 deformable registrations of 10 4D CT phases is needed, making this approach potentially computationally expensive.

These findings have not yet been confirmed for liver radiation therapy. Unlike the lung, liver tumors are not well visualized on 4D CT making the time-averaged position difficult to validate. The aim of this study was to evaluate simple and accurate methods to select a mid-ventilation CT, or create a mid-position CT for liver SBRT planning. The potential impact on PTV reduction was also quantified as this may facilitate substantial normal tissue sparing and dose-escalated liver SBRT.

6.2 Methods and Materials

6.2.1 Patient and imaging data

Data from 10 patients treated on clinical trials of liver SBRT were used (Table 6.1). 4D CT was acquired under normal breathing, using a multi-slice scanner coupled to a respiratory signal (abdominal bellows device or infrared chest marker). Ten phases were reconstructed using phase-based sorting, with each phase representing one tenth of the breathing period. Typically, 50% represents the end-exhale and 0% the end-inhale phases. Contouring and SBRT planning was generally done on the exhale 4D CT, with tumor delineation based on fused contrast-enhanced CT or magnetic resonance, both acquired under voluntary normal exhale breath hold.

Table 6.1. Patient and imaging details.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Descriptions (number of patients)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breathing motion management</td>
<td>Free-breathing (6), abdominal compression plate (4)</td>
</tr>
<tr>
<td>4D CT image resolution</td>
<td>1.0×1.0×2.0 mm (6), 1.0×1.0×2.5 mm (4)</td>
</tr>
<tr>
<td>Clinical planning image set</td>
<td>End-exhale 4D CT (7), voluntary exhale breath-hold CT (3)</td>
</tr>
<tr>
<td>Tumor delineation image set(s)</td>
<td>Contrast-enhanced CT (6), MRI (8)</td>
</tr>
<tr>
<td>Gross tumor volume (GTV)</td>
<td>Range 11-532 cm³</td>
</tr>
</tbody>
</table>

For this study, the liver was contoured on all 4D CT phases by a single observer. The gross target volume (GTV) contour on exhale 4D CT was unaltered from the clinical plan. In 3 cases where the planning data set was the contrast-enhanced breath-hold CT, the GTV contour was transferred to the non-contrast exhale 4D CT. The 4D CT was acquired immediately after the
breath-hold CT and therefore did not require any further registration. The accuracy of liver alignment was visually inspected on the fused images, and quantified using the Dice similarity coefficient (DSC) between the liver contours on the exhale 4D CT and breath-hold CT. The mean DSC was 0.96, thus differences in liver position were negligible. It was not possible to accurately contour the GTV on all 4D CT phases due to contrast washout during image acquisition.

6.2.2 Deformable image registration

The exhale 4D CT was registered to the remaining 4D CT phases using Morfeus, a biomechanical model-based DIR algorithm. Briefly, the primary liver contour on the exhale 4D CT is converted into a 3D surface mesh that establishes a correspondence with the secondary liver surfaces, created from liver contours on the other phases, by means of guided-surface projections (HyperMorph, Altair Engineering, Troy, USA). The internal liver and GTV deform according to assigned tissue biomechanics. Because this algorithm is independent of the image intensity, the accuracy of the GTV registration is independent of the contrast of the GTV in each image. The accuracy (absolute mean±SD) of Morfeus has been previously shown to be 1.2±0.7 mm left-right (LR), 1.7±1.4 mm anterior-posterior (AP) and 1.4±1.0 mm superior-inferior (SI)[124].

6.2.3 Strategies to estimate the mean respiratory position

The true time-weighted mean respiratory GTV position, hereafter referred to as the time-averaged position, was calculated as the average GTV position determined using DIR over all ten 4D CT phases. Two simplified strategies were investigated to select or reconstruct a CT that estimates the time-averaged position (concept shown in Figure 6.1). The accuracy of both is reported as the GTV’s center-of-mass position relative to the time-averaged position.

The first strategy selects one of the initial ten 4D CT phases (10% bins) as the mid-ventilation CT. Assuming a scenario where the 4D CT lacks GTV contrast and DIR is not available to track the GTV, the SI diaphragm motion was measured using a local rigid registration on each 4D CT (Figure 6.2). The mean diaphragm position was calculated over all 10 phases, and the phase with the smallest diaphragm displacement relative to the mean diaphragm position was chosen as the mid-ventilation CT. This assumes that the diaphragm and GTV motion correlate well with no phase shifts.
Figure 6.1. Left: Schema showing the 4D CT tumor motion (solid line, white circles); the possible mid-ventilation CT phase, the mid-position CT (gray circle) along the linear exhale-to-inhale trajectory (dashed line), versus the time-averaged position (black circle). Note that it is possible for multiple phases (e.g. 20% or 80%) to be selected as the mid-ventilation CT if they are equally close to the time-averaged position. Right: the SI diaphragm (white boxes) position on 4D CT. Mid-ventilation CT was selected as the phase with the diaphragm closest to the mean diaphragm position (gray box). The relative SI mean diaphragm position was also applied to the exhale-to-inhale 4D CT deformation map to generate the mid-position CT.

Figure 6.2. The local rigid registration volume is around the left diaphragm, and excludes the chest wall and heart. A mutual information algorithm was used, only considering translations.
The second strategy reconstructs the mid-position CT using a single DIR between the end-exhale and end-inhale 4D CT to capture the maximum GTV motion. Each patient’s SI diaphragm motion, from above, was normalized (exhale = 0, inhale = 1) and the mean position was calculated to estimate the mid-position of the GTV. This assumes the diaphragm and the GTV mean position occur at the same relative position between exhale and inhale. This percentage was then applied to the linear trajectory of the exhale-to-inhale deformation map to deform the exhale 4D CT into the mid-position CT assuming no hysteresis. Both the patient-specific diaphragm position and the population diaphragm position (averaged over all 10 patients) were tested. The proposed workflows are compared in Figure 6.3.

6.2.4 PTV margin calculations

The required GTV-to-PTV margins were quantified and compared for 2 planning strategies. The first uses the exhale position as the planning data set while expanding the ITV into the PTV. The second uses the mean position as the planning data set and applies a true dose-probability PTV that does not encompass the full motion extent. The margin recipe proposed by van Herk et al. was applied based on 95% of the prescribed dose being received by 90% of the population[67]. Both include the population inter-fraction and intra-fraction liver SBRT setup errors and the DIR accuracy, all previously quantified to be <2 mm individually. This results in a baseline margin of 3.7 mm LR, 4.5 mm AP, and 5.4 mm SI, excluding a component for respiration. For the ITV-based PTV, this margin was added linearly to the maximum GTV respiratory motion. For the dose-probability PTV, the population systematic errors (i.e. GTV error on the mid-position CT versus the time-averaged mean) and patient-specific random errors (i.e., one-third the GTV amplitude[163]) for respiration were added quadratically with the other population errors into the margin as shown in Table 6.2, with no further linear expansion.
Figure 6.3. Workflows to select a mid-ventilation CT or reconstruct a mid-position CT.
Table 6.2. Example calculation of the inferior PTV margin applied to the GTV on the exhale 4D CT where the breathing amplitude is fully encompassed (ITV-based PTV), or on the mid-position CT where the amplitude is incorporated into the PTV as a random error.

<table>
<thead>
<tr>
<th>Variable</th>
<th>ITV-based PTV used on exhale 4D CT</th>
<th>Dose-probability PTV used on the mid-position CT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inter-fraction setup errors ((\text{INTER}))</td>
<td>(\sum_{\text{INTER}}, \sigma_{\text{INTER}})</td>
<td>(\sum_{\text{INTER}}, \sigma_{\text{INTER}})</td>
</tr>
<tr>
<td>Intra-fraction setup errors ((\text{INTRA}))</td>
<td>(\sum_{\text{INTRA}}, \sigma_{\text{INTRA}})</td>
<td>(\sum_{\text{INTRA}}, \sigma_{\text{INTRA}})</td>
</tr>
<tr>
<td>Deformable image registration accuracy ((\text{DIR}))</td>
<td>(\sum_{\text{DIR}}, \sigma_{\text{DIR}})</td>
<td>(\sum_{\text{DIR}}, \sigma_{\text{DIR}})</td>
</tr>
<tr>
<td>Error in mean position ((\text{MID-POSITION CT ERROR}))</td>
<td>n/a</td>
<td>(\sum_{\text{MID-POSITION CT ERROR}})</td>
</tr>
<tr>
<td>Patient-specific GTV amplitude (A)*</td>
<td>A</td>
<td>(\sigma_{A} \approx A/3)</td>
</tr>
<tr>
<td>Required margin†</td>
<td>(2.5\sum_{\text{TOTAL}} + 0.7\sigma_{\text{TOTAL}} + A)</td>
<td>(2.5\sum_{\text{TOTAL}} + 0.7\sigma_{\text{TOTAL}})</td>
</tr>
</tbody>
</table>

Abbreviations/Symbols: A=amplitude; \(\Sigma\)=standard deviation of systematic errors; \(\sigma\)=standard deviation of random errors; Notes: *Modeled with deformable registration between exhale and inhale 4D CT; †All \(\Sigma\) and all \(\sigma\) from each column are added in quadrature to determine the \(\sum_{\text{TOTAL}}\) and \(\sigma_{\text{TOTAL}}\), respectively.

6.3 Results

6.3.1 Respiratory motion quantification

Accurate quantification of respiratory motion is required for both ITV-based and dose-probability PTV approaches. On average, the maximum SI diaphragm motion calculated with rigid registration was 2.1 mm larger \((p<0.02)\) than the GTV motion calculated with DIR (Table 6.3). Although the GTV and diaphragm motion correlated well \((R^2=0.77)\), the diaphragm had at least 3 mm greater motion in 30% of patients (range: 4.0–5.5 mm). Respiration-induced deformation caused multi-focal GTVs within the liver to have differences in amplitude up to 6 mm (Figure 6.4). The maximum motion for both structures occurred between the same 4D CT phase pairs in 9 of 10 patients. One patient had a slight phase shift where the maximum motion was between 50% and 0% for the GTV and 50% and 90% for the diaphragm; however, the difference in diaphragm motion between these phase pairs was only 0.7 mm.

Table 6.3. Maximum amplitude on 4D CT (in mm).

<table>
<thead>
<tr>
<th></th>
<th>LR</th>
<th>AP</th>
<th>SI</th>
<th>3D vector</th>
<th>Diaphragm motion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean ± SD</td>
<td>-1.9±1.5</td>
<td>2.9±1.9</td>
<td>10.3±3.8</td>
<td>11.1±3.9</td>
<td>12.4±4.3</td>
</tr>
<tr>
<td>Range</td>
<td>-4.1, 0.9</td>
<td>-0.6, 5.7</td>
<td>6.0, 17.3</td>
<td>6.4, 18.6</td>
<td>5.8, 19.4</td>
</tr>
</tbody>
</table>
6.3.2 4D GTV position versus mean position

The GTV was mapped from end-exhale to the remaining 4D CT phases using DIR. The GTV position at each of the 10 phases was then compared with its time-averaged position to quantify the inherent GTV error at each of the initial 4D CT phases reconstructed with the 10% bins. Using Figure 6.1 to illustrate, the distance between the time-averaged position (black circle) and each of the 4D CT phases (white circles) was calculated. Across all patients, the mean GTV errors were smallest at the 20% (exhalation) followed by the 80% (inhalation) phases. Three patients had GTV errors >2 mm on the 20% phase (range: 2.3–3.2 mm, all in the SI direction). Individually, however, the phase with the smallest GTV error was patient specific and was most commonly the 20% phase (5 patients), followed by the 80% phase (2 patients), then the 10%, 30%, and 70% phases (1 patient each).

![Diagram showing diaphragm and GTV motion over all 4D CT phases for 1 patient (GTV locations shown on the inset figure).](image)

Figure 6.4. Diaphragm and GTV motion over all 4D CT phases for 1 patient (GTV locations shown on the inset figure).
6.3.3 Mid-ventilation CT and mid-position CT accuracy

Because the 4D GTV position would not be known without DIR, mid-ventilation CT selection was based solely on the 4D diaphragm position. Using Figure 6.1 to illustrate, the phase (white box) with the diaphragm closest to the mean diaphragm position (gray box) would be selected. Ideally, this phase would also have the smallest GTV error relative to the time-averaged GTV position (white versus black circles). Selection of the mid-ventilation CT in this manner coincided with the phase having the smallest GTV error relative to the time-averaged position in 50% of patients, and the second smallest error in the other 50%. One patient had a GTV error of 2.1 mm SI while all others were ≤1.1 mm.

Table 6.4. Errors in liver GTV position for different planning CT datasets relative to the time-averaged position (in mm).

<table>
<thead>
<tr>
<th>Dataset</th>
<th>LR</th>
<th>AP</th>
<th>SI</th>
<th>Vector magnitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>20% 4D CT phase</td>
<td>0.1±0.5 (0.8)</td>
<td>0.2±0.8 (1.8)</td>
<td>0.7±1.6 (3.2)</td>
<td>1.7±1.6 (3.7)</td>
</tr>
<tr>
<td>80% 4D CT phase</td>
<td>−0.5±0.6 (1.9)</td>
<td>0.5±0.8 (2.9)</td>
<td>1.0±2.4 (4.8)</td>
<td>1.8±1.6 (6.0)</td>
</tr>
<tr>
<td>Mid-ventilation CT</td>
<td>−0.1±0.4 (0.8)</td>
<td>0.1±0.7 (1.1)</td>
<td>0.2±0.8 (2.0)</td>
<td>1.0±0.5 (2.0)</td>
</tr>
<tr>
<td>Mid-position CT (using patient-specific diaphragm position)</td>
<td>0.1±0.3 (0.4)</td>
<td>−0.1±0.4 (0.8)</td>
<td>0.0±0.6 (1.1)</td>
<td>0.6±0.3 (1.4)</td>
</tr>
<tr>
<td>Mid-position CT (using population diaphragm position)</td>
<td>0.1±0.3 (0.6)</td>
<td>−0.1±0.3 (0.6)</td>
<td>0.0±0.8 (1.4)</td>
<td>0.8±0.4 (1.5)</td>
</tr>
</tbody>
</table>

The mid-position CT was reconstructed by applying the patient’s mean normalized diaphragm position (average: 0.43±0.09, range: 0.30 to 0.61) as a percentage to the single exhale-to-inhale deformation map. This was within 0.01±0.06 (range: -0.05 to 0.13) of the normalized mean SI position of the GTV determined using DIR on all 4D phases. For the mid-position CT reconstructed with the patient’s mean normalized diaphragm position, 2 patients had SI GTV errors of 1.0 and 1.1 mm relative to the time-averaged position, while all other errors were <1 mm. In Figure 6.1, this error is the difference between the gray and black circles. Applying the population’s mean normalized diaphragm position of 43% to each patient’s deformation map resulted in similar average GTV errors on the mid-position CT created with the patient-specific
diaphragm position ($p=0.27$). Two patients had SI errors of 1.4 and 1.3 mm, while all other errors were <1 mm. Table 6.4 compares the mean GTV errors relative to the time-averaged position for the various possible planning data sets.

### 6.3.4 Impact of mean position on the required margin

Compared with planning at exhale with ITV-based PTV, planning on the mid-position CT with dose-probability PTV resulted in a GTV-to-PTV volume reduction (mean±SD) of 34±7% up to a maximum of 43%. This translates into 66±38 cm$^3$ (maximum 126 cm$^3$) of surrounding normal tissue that could be potentially spared full dose.

### 6.4 Discussion

At the time of this investigation, no other studies considered the accuracy of the mid-ventilation or mid-position CT for liver radiation therapy planning. Selecting a single 4D CT phase as the mid-ventilation CT using only rigid registration to quantify diaphragm motion in a commercial treatment planning system resulted in tumor errors relative to the time-averaged position of approximately 2 mm in all patients. Reconstructing a mid-position CT using a single deformation map (exhale-to-inhale) resulted in tumor errors of approximately 1 mm in all patients. Liver SBRT planning on these datasets with dose-probability PTV would encompass an average of 34% less normal tissue volume inside the PTV compared with planning at exhale with ITV-based PTV.

The mid-ventilation CT was simply selected from the initial 10 4D CT datasets (10% bins), avoiding the reconstruction of an extra CT around the exact time-percentage where the tumor is closest to the mean position. The latter requires either the raw data from the CT scanner, or the external respiratory signal which may be prone to errors caused by phase shifts[165]. The mean error on the mid-ventilation CT was small (1.0±0.5 mm) and likely a combination of the 4D CT binning, and tumor hysteresis including only considering SI diaphragm motion for mid-ventilation CT selection. GTV position errors versus the time-averaged position were smallest at the 20% and 80% phases, similar to the exact time-percentages reported previously in the lung (20.7±2.2% for exhalation and 78.8±2.9% for inhalation[164]). Hysteresis has been previously shown to be greater than 1 mm in only 20% of liver patients[70]. Because the tumor is moving relatively fast near the mid-ventilation phase it may be prone to imaging artifacts. One patient with a mid-ventilation CT error of 2 mm had 4D sorting artifacts near the liver on the
intermediate phases but not on exhale or inhale phases, likely contributing to this error. This patient’s GTV also had 14.9 mm of motion, though no strong correlations were seen between patients’ residual GTV errors and the maximum GTV amplitude (data not shown).

The mid-position CT previously proposed for lung requires DIR of each 4D CT phase to establish the time-averaged position[166]. Because all phases are used, delineation accuracy can be potentially improved due to the averaging out of phases with artifacts and reduced image noise[166]. Often however, liver GTVs are delineated directly on the planning exhale 4D CT but are based on the fused contrast-enhanced tri-phasic CT or MRI after performing a rigid liver-to-liver registration. In instances where substantial liver deformation is observed, the registration is focused to the region of the liver surrounding the GTV.

Although the mid-ventilation CT selection requires a smaller workload, the simplified mid-position CT reconstruction in this study uses only one DIR on the extreme phases and it avoids planning on the artifact-prone intermediate phases. In the absence of intravenous contrast synchronized with 4D CT acquisition, the use of DIR may be advantageous for quantifying GTV motion for PTV determination because of the diaphragms’ overestimation and by accounting for differences in GTV motion due to respiration-induced liver deformation (Figure 6.4). The deformation map was combined with either the patient’s or the populations mean normalized diaphragm position resulting in mean vector errors of 0.6±0.3 and 0.8±0.4 mm, respectively. The latter strategy is further simplified by avoiding the diaphragm motion analysis for each patient without sacrificing accuracy. Note that this study deformed only the liver for simplicity. To reconstruct the entire mid-position CT in clinical practice, the liver, spleen, and external body would require contouring on the exhale and inhale 4D CT to obtain a multi-organ deformation map.

Validation of the mean respiratory position would ideally compare the GTV position on mid-ventilation or mid-position CT with the average of the actual GTV position identified on all 4D CT phases. This was impossible to do on the majority of patients due to the poor tumor contrast. The errors reported above are therefore relative to the GTV’s time-averaged position predicted by DIR. Morfeus was previously shown to be accurate inside the liver to within 2 mm. For only 3 patients in this study additional validation was possible by identifying an anatomic landmark visible on all 4D CT phases. For 2 patients, residual intravenous contrast allowed the exhale GTV contour to be copied and rigidly fit to subsequent phases. In the third, residual Lipiodol
(Guerbet, Villepinte, France) from previous therapy was contoured using an auto-threshold segmentation tool on each 4D CT phase image. Comparing the true image-based time-averaged position of the landmarks’ centroid with that predicted by Morfeus resulted in an average registration error of 1.0±1.4 mm. Comparing the true image-based time-averaged position of the landmarks’ centroid to that predicted via the mid-ventilation or mid-position CT workflows resulted in errors of 1.1±0.3 and 0.8±0.7 mm, respectively, on the order of deformable registration accuracy and within the image resolution.

The importance of this work is supported by the increasing adoption of liver SBRT and the added commercially available cone-beam CT capability for daily 4D image-guidance to the mean position. Planning and delivery at the mean position with dose-probability based margins requires 34% less volume of normal tissue irradiation on average, compared with the commonly used ITV-based margins. These strategies may facilitate liver SBRT dose escalation, which observed dose-response relationships suggest could improve local control.

### 6.5 Conclusions

Simplified methods to select or reconstruct the mid-ventilation CT and mid-position CT were validated for liver radiation therapy planning with respect to the time-weighted mean respiratory position. It was shown that establishing the mean position at planning could be presently implemented with a reasonable 2 mm accuracy using simple, clinically and widely available tools (i.e. even without DIR).
Chapter 7
Dose-Escalated Liver SBRT at the Mean Respiratory Position

This chapter was published as: Michael Velec, Joanne L. Moseley, Laura A. Dawson, Kristy K. Brock. Dose-escalated liver stereotactic body radiotherapy (SBRT) at the mean respiratory position. *International Journal of Radiation Oncology, Biology, Physics* 89(5): 1121-1128, 2014.
7 Dose-Escalated Liver SBRT at the Mean Respiratory Position

7.1 Introduction

Evidence of a dose-response relationship for local tumor control is emerging for liver cancer patients treated with stereotactic-body radiotherapy (SBRT)[52,56,167,168]. Low rates of serious toxicities follow SBRT, providing a rationale to investigate iso-toxic dose-escalation techniques. Individualized dose-allocation schemes modeling liver normal tissue complication probabilities (NTCP) have been implemented clinically, however normal tissues still limit dose in the majority of patients[60].

Reducing the component of the planning target volume (PTV) accounting for breathing motion can spare normal liver tissue. Ten Haken et al. [79] modeled a 7% increase in tumor control probability after eliminating the breathing margin. This may be possible using active breath hold SBRT delivery, however approximately 40% of patients are unsuitable[100]. Abdominal compression passively reduces liver motion by only 2.3 mm on average[87]. Gating allows for reduced PTVs and an iso-NTCP median dose-escalation of 21%[82]. However, gating requires continuous real-time monitoring[169] and larger online workload versus other strategies.

For free-breathing radiotherapy, a simple and widely used PTV design creates an internal target volume (ITV), a union of the tumor’s positions on respiratory-correlated (4D) CT. A further linear expansion for setup uncertainties defines the PTV. This strategy aims for 100% tumor dose coverage over the entire breathing cycle, but effectively overcompensates as the tumor cannot be simultaneously at all phases. Margin ‘recipes’ have also been derived under simplified conditions, ensuring specific dosage and confidence levels (e.g. 90% of patients receive 95% dose)[67]. These dose-probability PTVs combine patient-specific breathing motion with population treatment uncertainties. They are 34% smaller on average than ITV-based PTV in liver SBRT[170]. Their dosimetric impact and robustness to liver SBRT uncertainties requires evaluation as there may be a reluctance to reduce margins clinically.

Delivered doses were previously accumulated in Chapter 3 with deformable registration of the treatment cone-beam CTs, for free-breathing liver SBRT plans using ITV-based PTVs[143]. Residual targeting errors exceeded the setup margin of 3 mm in 30% of patients however the delivered tumor doses were lower than prescribed in only 1 patient (3.3%) with substantial inter-
fraction liver deformation. The large ITV component of the PTV, intended to only compensate for breathing, may have nullified the dosimetric impact of these setup errors.

The aims of this study were to investigate liver SBRT planning at the mean breathing position with dose-probability PTVs enabling normal tissue sparing and dose-escalation. SBRT delivery was simulated with a guidance strategy based on 4D cone-beam CT and rigid registration. Deformable registration was used to reconstruct the delivered tumor and normal tissues doses to evaluate the robustness of the plans. Iso-toxic dose-escalation via margin reduction may be safely explored to improve local control provided the actual risk of target under-dosing does not increase.

7.2 Methods

7.2.1 Patient data

Twenty patients with primary or metastatic liver cancer previously treated on dose-escalation trials of liver SBRT for 27–49.8 Gy in 6 fractions were retrospectively investigated. The median total gross tumor volume (GTV) was 174 cm$^3$ (range: 26–2402 cm$^3$) over a median of 2 GTVs per patient (range: 1–3). Patients were treated free-breathing or with abdominal compression. The median GTV breathing amplitude was 8 mm (range: 1–21 mm).

Clinical planning was done on the exhale 4D CT (Pinnacle$^3$ v9.2; Philips Medical Systems, Madison WI). The GTV contour was based on fused contrast-enhanced voluntary breath-hold CT and magnetic resonance images (MRI). The PTV (clinical PTV) was applied asymmetrically around each GTV at exhale, encompassing the patient-specific breathing motion plus a 5 mm expansion to account for other uncertainties (e.g. setup errors). In practice, the patient-specific motion was assessed as 90-100% of the motion measured on 4D CT, fluoroscopy and cine-MRI. The lack of liver-GTV contrast on 4D CT prevented direct GTV contouring on all 4D phases and the creation of a traditional ITV.

Daily image-guidance involved rigidly registering the planning exhale 4D CT liver contour to a free-breathing kilovoltage 3D CBCT, biased towards the superior part of the blurred liver diaphragm (i.e. the exhale position). Uncertainty in targeting the GTV results from using non-4D and non-contrast CBCT, the liver as a surrogate and residual liver deformation[143]. The residual population systematic (random) GTV errors were quantified in Chapter 3 with
retrospective DIR of 4D CBCT to be 1.8 (1.7) mm left-right (LR), 2.2 (1.9) mm anterior-posterior (AP) and 2.6 (2.4) mm superior-inferior (SI)[143].

4D CBCT was not available clinically at the time these patients were treated, so the CBCTs were retrospectively 4D-sorted for this study[93].

7.2.2 SBRT planning

For this study, three plans were created with the same techniques. The beam arrangement was initially preserved from the clinical plan and intensity-modulated radiotherapy (IMRT) was re-optimized for all plans. Individual prescribed doses were maximized based on a previously implemented liver iso-NTCP scheme after minimizing the normal liver dose[53]. The patient’s risk level assigned clinically (5–20%) was maintained across the three plans. The primary IMRT objectives were to maximize the dose to 95% of the PTV (PTVD95) while constraining the maximum luminal gastrointestinal organ doses to 31–36 Gy to 0.5 cm³ (details in [56]). If the last criteria could not be achieved, a lower dose than permitted by the iso-NTCP schema was prescribed, and the plan was re-optimized until all criteria were met. Plans were normalized so the prescribed dose covered at least 95% of the PTV.

7.2.3 PTV strategies and planning datasets

For this study, three plans were compared overall, two on the exhale 4D CT and one on the mid-position CT (Figure 7.1).

The first exhale 4D CT plan used the clinical PTV described above, termed the ‘clinical PTV plan’. Because this encompassed 90-100% (based on physician preference) of the breathing motion plus a setup margin it is similar in concept to ITV approaches.

The second exhale 4D CT plan’s PTV was standardized across all patients by using an ITV based on the exhale and inhale GTV contours as follows, termed the ‘ITV-based PTV plan’. Each exhale 4D CT was deformed to the inhale 4D CT using Morfeus, a biomechanical model-based deformable image registration (DIR) algorithm. This DIR establishes correspondence between images using guided-surface projections between contours of the liver, spleen and body on each image. These serve as boundary conditions for the model while tissues without secondary contoured surfaces (e.g. GTV) deform via the assigned biomechanical properties and are solved with finite element analysis. The mean and standard deviation of the absolute errors
in the DIR-predicted positions of vessel bifurcations within the liver are ≤1.7 mm in any direction[124]. The exhale and deformed-inhale GTV contours were combined into an ITV and symmetrically expanded by 5 mm to create the ITV-based PTV, ensuring consistently constructed ITVs between all patients.

The third plan was created on a mid-position CT representing the patient anatomy in its time-averaged respiratory position allowing for dose-probability PTVs, termed the ‘dose-probability PTV plan’. Intermediate 4D CT phases were not archived for these patients therefore the mid-position CT was retrospectively reconstructed using DIR of each patient’s exhale and inhale 4D CT[170]. The exhale 4D CT and planning contours were deformed along the exhale-to-inhale linear trajectory into the mid-position CT geometry using the method developed in Chapter 6, by applying the populations’ mean diaphragm position (43% of the exhale-to-inhale distance). This accuracy of this simplification is <2 mm versus the true time-averaged position[170].

This mid-position CT plan used the van Herk[67] PTV derivation to ensure 90% of patients receive the nominal dose prescribed to the 90% isodose (typical for liver SBRT):

\[
\text{Margin} = 2.5\Sigma + 1.28\sqrt{(\sigma^2 + \sigma_{\text{penumbra}}^2)} - 1.28\sigma_{\text{penumbra}} (1)
\]

where \( \Sigma \) and \( \sigma \) are the population standard deviations (SD) of all systematic and random errors respectively, and \( \sigma_{\text{penumbra}} \) is the SD of the penumbra in water (3.2 mm[67]). The \( \Sigma \) and \( \sigma \) are the quadratic addition of the published uncertainties in GTV position due to DIR inaccuracy [124,125], intra-fraction errors specific to free-breathing liver SBRT[95], and residual inter-fraction GTV errors. It was assumed that after rigid liver alignment with 4D CBCT, deformation would cause residual errors in GTV position, therefore the inter-fraction errors were calculated as the centre-of-mass difference between the GTV and liver from the patients studied in Chapter 3[143]. The SD of the patient-specific 4D CT breathing motion was added quadratically to the population random errors (\( \sigma \)). Because the SD can be approximated as 1/3 the amplitude, this margin is smaller than the ITV-based PTV encompassing the full amplitude[80]. Rit et al. showed dose-probability PTVs are valid even with significant breath-to-breath variability[171]. The dose-probability PTV was applied symmetrically around the GTV on the mid-position CT. The baseline margin for all patients prior to the quadratic addition of the patient-specific breathing amplitude into \( \sigma \) was 3.1 mm LR, 3.9 mm AP and 4.7 mm SI.
Figure 7.1. Schema of the different PTV strategies (grey volumes). Clinically, the GTV was delineated on exhale 4D CT (1). DIR of 4D CT allowed for propagation of the exhale contours into the inhale (2) or time-averaged position (3). Figure adapted from [76].

7.2.4 Delivered dose reconstruction to evaluate PTV robustness

The delivered dose was modeled as previously described in Chapter 2 and 3, and is summarized in Figure 7.2[143]. DIR of all six 4D CBCT allowed for residual inter-fraction targeting errors, deformations, and daily breathing motion to be accumulated into the delivered dose distributions for each plan. For the two exhale 4D CT plans, the residual position from the clinical treatment was used, including any post-CBCT corrections. For the mid-position CT plan, the liver was rigidly corrected at mean breathing position on 4D CBCT retrospectively, prior to dose reconstruction. This correction simulates the current scenario where online deformable registration is not available for direct GTV targeting. Delivered dose changes >1 Gy relative to the planned dose are assumed to be potentially clinically significant. Maximum and minimum doses are reported to 0.5 cm$^3$. 

\[ \text{GTV position on exhale 4DCT} \]
\[ \text{GTV position on exhale 4DCT} \]
\[ \text{GTV position on mid-position CT (deformed from exhale to inhale 4DCT)} \]

\[ \text{Respiratory position:} \]
\[ \text{Exhale} \]
\[ \text{Inhale} \]
\[ \text{Time-averaged} \]

\[ \text{Expansion from planned position:} \]
\[ \text{Asymmetric; motion from imaging studies + 5 mm} \]
\[ \text{Asymmetric; exhale + inhale GTV 4DCT contours + 5 mm} \]
\[ \text{Symmetric; 4DCT GTV motion incorporated into Equation 1} \]
Figure 7.2. Workflows to model the delivered dose following CBCT-guidance based on the liver’s exhale or mean breathing position.

7.3 Results

7.3.1 Planning and dose-escalation

All PTVs are compared to breathing motion in Figure 7.3. Because the ITV-based PTV is a linear addition of breathing motion and setup margin, is it larger than the dose-probability PTV which adds the errors in quadrature (Equation 1). Clinical PTVs were on average 5% smaller in volume than ITV-based PTVs ($p=0.08$).

The average prescription isodose for the dose-probability PTV plans after optimization was 90% (range: 85-97%), equal to the isodose the margin was tailored for. Dosimetry for the dose-probability PTV plans compared to the clinical and ITV-based PTV plans is shown in Table 7.1. All plans met the normal tissue dose constraints with similar PTV coverage. For all plans, the
planned minimum GTV doses were 2-4% higher on average than the PTVD95, owing to the large dose heterogeneity allowed for SBRT distributions.

Figure 7.3. PTV size (total SI margin) versus the breathing amplitude. Above 6 mm of motion the difference in total SI expansion exceeds 5 mm between the dose-probability and ITV-based PTVs.

Compared to the clinical PTV plans, the PTVD95 for the dose-probability PTV plans were more than 5 Gy higher in 55% of patients for least one GTV (53% of all GTVs). Compared to ITV-based PTV plans, the PTVD95 for the dose-probability PTV plans were more than 5 Gy higher in 50% of patients for at least one GTV (50% of all GTVs). In 15% of patients at least one GTV’s PTVD95 for the dose-probability PTV plans was within ±1 Gy (the minimum ‘clinical relevance criteria’) of the clinical or ITV-based PTV plans’ PTVD95, but none were more than 1 Gy lower.

Factors affecting the dose-escalation between dose-probability PTV and ITV-based PTV plans were investigated. Patients with GTV breathing motion <10 mm had an average±SD dose-
escalation using the dose-probability PTV plans of 5.2±3.8 Gy versus 3.8±1.7 Gy for >10 mm (p=0.28). Smaller GTV size (p=0.01) and PTVs that did not overlap normal luminal gastrointestinal tissues (p=0.04) had a higher magnitude of dose-escalation.

Table 7.1. Planning dosimetry for the dose-probability PTV plans, and its change from the clinical and ITV-based PTV plans.

<table>
<thead>
<tr>
<th>Dose-probability PTV plans on mid-position CT, mean±SD</th>
<th>Δ in dose-probability PTV plan parameters on mid-position CT compared to:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Clinical PTV plan on exhale CT, mean±SD</td>
</tr>
<tr>
<td><strong>Target dosimetry</strong></td>
<td></td>
</tr>
<tr>
<td>PTV–GTV volume</td>
<td>128±69 cm³</td>
</tr>
<tr>
<td>Nominal prescription</td>
<td>40.3±7.6 Gy</td>
</tr>
<tr>
<td>PTVD95</td>
<td>42.6±8.4 Gy</td>
</tr>
<tr>
<td><strong>Normal liver dosimetry</strong></td>
<td></td>
</tr>
<tr>
<td>Mean dose (excluding GTV)</td>
<td>16.1±3.2 Gy</td>
</tr>
<tr>
<td>Effective volume[^1]</td>
<td>0.39±0.12</td>
</tr>
<tr>
<td>Liver NTCP[^2]</td>
<td>2.4±4.0 %</td>
</tr>
<tr>
<td><strong>PTV coverage quality</strong></td>
<td></td>
</tr>
<tr>
<td>Volume receiving ≥ prescribed dose</td>
<td>99±1%</td>
</tr>
<tr>
<td>RTOG conformity index[^†]</td>
<td>1.29±0.15</td>
</tr>
<tr>
<td>RTOG coverage quality index[^† ‡]</td>
<td>0.98±0.05</td>
</tr>
<tr>
<td>RTOG homogeneity index[^#]</td>
<td>1.17±0.06</td>
</tr>
</tbody>
</table>

Abbreviations: ITV=internal target volume; NTCP=normal tissue complication probability; PTV=planning target volume; PTVD95=dose to 95% of the PTV; RTOG=Radiation Therapy Oncology Group; SD=standard deviation. Notes: ^* p<0.01 two-tailed paired Student’s T-test; ^[^1]calculated as per Kutcher et al.[37]; ^[^2]calculated with Lyman-linear quadratic model corrected for the dose/fraction (α/β=2.5 Gy)[53];[^†]ratio of nominal prescription isodose volume/PTV;[^† ‡]ratio of min PTV dose to 0.5cm³/nominal prescription dose;[^#]ratio of max PTV dose to 0.5cm³/nominal prescription dose.
7.3.2 Residual GTV treatment uncertainties

Following rigid alignment of each 4D CBCT’s mean liver position to the mid-position CT, the residual systematic (random) GTV errors, defined about the centre-of-mass, were estimated with DIR to be 0.9 (1.0) mm LR, 1.4 (1.1) mm AP and 1.0 (0.9) mm SI. These are reduced by 1.6 (1.5) mm SI versus errors reported following exhale CT to 3D CBCT liver alignment[143].

Mean changes in GTV breathing amplitude (4D CBCT-4D CT) were ≤1.5 mm. In 2 patients the amplitude increased by 5 and 8 mm, and these variations were present for all three PTV strategies.

7.3.3 Delivered versus planned doses

The delivered min GTV doses were first compared to their initial planned nominal prescribed doses. For all three strategies the delivered doses were on average 2.3–3.1 Gy (6–8%) higher and not more than 1 Gy lower for any patient. Using stricter criteria, the delivered min GTV doses were next compared to their planned PTVD95 (Figure 7.4). For all three strategies the delivered doses were on average 0.5–0.9 Gy (1–3%) higher than the PTVD95, but lower by at least 1 Gy in 10% of patients. The average dose deviation (delivered min GTV–planned PTVD95) was narrower for the dose-probability PTV plan (0.5 Gy, or 1%) versus the ITV-based PTV plan (0.9 Gy, or 3%), p=0.046.

In all three strategies, 2 different patients (10%) had at least one GTV with a delivered min dose more than 1 Gy lower than its PTVD95 (Figure 7.4.B). The largest delivered min GTV dose decrease was -2.2 Gy (-5%), -4.6 (-10%) and -4.7 Gy (-10%) for the clinical, ITV-based, and dose-probability PTV plans respectively. In the worst example, a patient with 3 GTVs (patient 3 on Figure 7.4.B) had liver deformation causing residual mean 3D GTV displacements of 5–7 mm, and a mean 8 mm larger breathing motion on 4D CBCT versus 4D CT. For this patient, all three plans had a delivered min GTV dose decrease of at least 1 Gy relative to their PTVD95. The delivered min doses however with the dose-probability PTV plan were >5 Gy higher than delivered with the clinical or ITV-based PTV plans, due to the planned dose-escalation that was possible with the dose-probability PTV.
For all three strategies the delivered normal luminal gastrointestinal tissues doses did not exceed planning constraints by more than 1 Gy with only one exception. The delivered max dose for one patient’s esophagus increased by 1.9 Gy (6%) relative to the planned dose only for the clinical PTV plan, and exceeded the planning constraint by 1.4 Gy.

7.3.4 Delivered doses between all strategies

The delivered min GTV doses are compared amongst the three strategies in Table 7.2. The delivered min GTV doses with the dose-probability PTV plans were not lower than either clinical or ITV-based PTV plans by more than 1 Gy in any case.

Table 7.2. Increase in the delivered GTV doses with the dose-probability PTV plan compared to the delivered doses with the clinical or ITV-based PTV plans.

<table>
<thead>
<tr>
<th>Delivered minimum GTV dose to 0.5cm³</th>
<th>Δ delivered dose for the dose-probability PTV plan compared to:</th>
<th>Delivered dose for the clinical PTV plan</th>
<th>Delivered dose for the ITV-based PTV plan</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVG±SD (max)</td>
<td></td>
<td>4.1±3.1 Gy⁺ (12.8 Gy) or 9±6 % (21 %)</td>
<td>3.5±3.0 Gy⁺ (11.4 Gy) or 8±6 % (19 %)</td>
</tr>
<tr>
<td>No. (%) GTVs with an increase ≥5 Gy</td>
<td></td>
<td>15 GTVs (44%)</td>
<td>14 GTVs (41%)</td>
</tr>
<tr>
<td>No. (%) patients with an increase ≥5 Gy to any GTV</td>
<td></td>
<td>8 patients (40%)</td>
<td>9 patients (45%)</td>
</tr>
</tbody>
</table>

Notes: *p<0.01 two-tailed paired Student’s T-test.
Figure 7.4. Delivered min GTV dose deviations relative to the planning PTVD95 for (A) all patients combined, and (B) separately for three patients each with a decrease ≥1 Gy to any GTV (Number 2 and 3 each have three GTVs).
7.4 Discussion

To the authors’ knowledge this is the first investigation of dose-probability PTV for liver SBRT including the quantification of dose-escalation. Dose-probability PTVs allowed for iso-toxic plans 4.0 Gy (9%) higher on average than ITV-based PTV plans. Of greater concern with margin reduction is whether the planned dose distributions conforming to reduced PTVs are equally robust to treatment uncertainties. To evaluate this, the delivered doses were modeled with deformable dose accumulation, assuming image-guidance was based on rigid CBCT liver alignment as per current clinical protocol. An equivalent rate of delivered tumor coverage was observed between PTV strategies, including the significantly smaller dose-probability PTVs. In summary, a complete liver SBRT planning and delivery strategy based on the mean breathing position was evaluated which can be implemented clinically with respiratory-correlated imaging, and simple rigid liver registration.

Linearly adding breathing motion to other margin expansions resulted in similar clinical and ITV-based PTVs (Table 1). Dose-probability PTVs however treat breathing motion as a random error only (≈1/3 the amplitude), assuming planning on the mid-position CT systematically accounts for the motion. The 8 mm median amplitude in this study is lower than the overall population because larger breathers received SBRT under active breath-hold. Dose-probability PTVs were also tailored to the 90% iso-dose, although iso-doses of 65% have been used for liver SBRT[68] requiring even smaller margins with the potential for further dose-escalation.

The dose-probability PTV was designed for free-breathing liver SBRT, and the exact margin may invalid for other techniques (e.g. active breath hold) or institutions. The original margin recipe included GTV delineation uncertainties[67]. These were not explicitly incorporated into any margin evaluated in this study, but are under investigation using DIR-enabled correlation of in vivo imaging to ex vivo histopathology[172]. In clinical practice these potential systematic errors are thought to be minimized using fused contrast-enhanced MRI, ‘over-contouring’ at ambiguous GTV borders, and contour peer-review. Additional margin validation is warranted using an independent cohort or modeling studies (e.g. Monte Carlo[173]), due to the limited number of patients investigated.

Dose-probability PTV adoption has been slow versus ITV methods, even for the more commonly treated lung which similarly benefits [76]. The barrier to widespread implementation
is likely the presumed-difficulty in establishing the mean breathing position. For liver SBRT planning, two groups recently validated the mid-position CT accuracy using DIR or simple rigid registration[129,170]. For delivery, this study modeled rigid liver localization at the mean 4D CBCT position to account for baseline shifts, which is a commercially-available capability. The population residual GTV errors following this strategy (caused by residual liver deformation) were likely reduced compared with exhale 4D CT to blurred 3D CBCT alignment, because the former strategy aligns analogous respiratory positions. PTVs were evaluated with DIR-based dose accumulation however image-guidance using rigid liver registration was simulated. Hugo et al.[174] also demonstrated the mean lung tumor position can be accurately localized by registering an average (i.e. blurred) planning 4D CT to free-breathing 3D CBCTs. This requires further investigation as liver tumors are poorly visualized on both 3D and 4D CBCT.

Leinders et al. [147] recently piloted daily adaptive IMRT re-optimization for liver SBRT. In 2 of 12 GTVs the intervention permitted better PTV dose coverage, but dose-escalation beyond the initial prescription was not investigated. A similar impediment to other strategies (e.g. gating), the authors acknowledge the current online workloads are prohibitive to clinical implementation. In comparison, the current study’s results suggest dose-escalation is feasible with PTV reduction, CBCT guidance and rigid liver registration.

Each PTV strategy had two patients (10%) where the delivered min GTV dose to any GTV was 5–10% lower than the PTVD95. These resulted from liver deformations, thus even margins fully encompassing the breathing motion (i.e. clinical and ITV-based PTV plans) also experienced tumor under-dosage. Despite the dose-escalation, the delivered normal tissue doses did not exceed the planning dose constraints for the dose-probability PTV plans. Dosimetric validation is progressing[136], but implementation of deformable dose reconstruction for SBRT quality assurance may allow for identification of delivered dose deviations and inform whether re-planning is warranted.

7.4.1 Validity of dose-probability PTV for SBRT

The van Herk formulas for the dose-probability PTV were originally derived with several assumptions and simplifications, and were demonstrated for use in conventionally fractionated radiotherapy. Although their application to liver SBRT has not been shown outside of the
current study, subsequent studies have shown the margin is valid in other scenarios. The original van Herk paper were assumed to be valid under the following major conditions[67]:

- A large fraction number: To account for the 6 fractions used in liver SBRT, all of the residual inter-fraction population systematic and random errors were similarly measured over 6 fractions. Over many fractions errors tend to average out to zero[163]. Therefore errors measured over fewer fractions are larger and the PTV incorporating these errors inherently accounts for the 6 fractions. The assumption was previously validated for a 3 fraction lung SBRT using Monte Carlo simulations[173]. In 297 patients who received 3 fraction lung SBRT using appropriately designed dose-probability PTV, 2-year local control of 98% was recently reported [175], providing further validation of the recipe.

- A spherical target, ignoring rotations and deformations: Most individual liver GTVs are approximately spherical, although half of the studied patients had multiple GTVs. There is no consensus on how to incorporate rotations and deformations into PTVs. The approach used in this study was to incorporate the residual GTV errors caused by liver deformation into the PTV margin. These were estimated as centre-of-mass differences between the GTV and liver using retrospective DIR of CBCT.

- Normally distributed errors: The only exception to this in liver SBRT is breathing motion. However, a follow up study validated the margin even for the bimodal breathing PDFs when the amplitudes are ≤10 mm[163]. Motion >10 mm requires slightly asymmetric but not necessarily larger margins. The median breathing amplitude in the current study was 8 mm. Rit et al. recently demonstrated that the impact of breathing asymmetry and large breath-to-breath variations within a fraction is limited on the required dose-probability PTV, and symmetric expansion is adequate in most cases[171].

7.5 Conclusion

Iso-toxic liver SBRT planning at the mean breathing position coupled with dose-probability PTVs allows for an average planned dose-escalation of 4 Gy (9%) in 6 fractions compared to ITV-based margins. When the delivered dose was evaluated with deformable dose reconstruction, the risk of under-dosing the tumor did not increase despite the reduced margins. This strategy has the potential to improve local control and can be readily implemented with respiratory-correlated imaging at planning and SBRT delivery and widely-available rigid registration.
Chapter 8
General Discussion
8 General Discussion

Deformable image registration and dose reconstruction has the potential to dramatically impact the field of radiation therapy. By closing the loop between treatment planning, where dose is estimated on a single CT image, and CBCT-guided treatment delivery, where the dynamic nature of patients is visualized, the estimates of the delivered doses can be substantially improved. This significant advancement is anticipated to impact both our current understanding of SBRT dose-response relationships for toxicity and tumor control, and how SBRT is designed and applied to future patients. In this thesis it was shown that despite the use of individualized treatment planning and daily image-guidance, the delivered dose significantly differed from the planned dose in the majority of liver cancer patients treated in free breathing. By accounting for the physical dose uncertainties we may improve our understanding of dose-effect relationships and the underlying biological differences in SBRT responses. Preliminary research shows that higher delivered liver SBRT doses correlate better with local tumor control than the doses estimated at planning which is the current standard in radiation oncology. Using novel treatment planning concepts, a liver SBRT strategy based on the mean breathing position and dose-probability margins was developed and evaluated. The smaller margins significantly spared more normal tissues and allowed for escalated delivered tumor doses without increasing the risks of liver toxicity or under-dosing the tumors during treatment delivery. This strategy, which can be readily implemented with existing technologies has the potential to improve local tumor control for future SBRT patients.

8.1 Summary of Chapters

Reducing uncertainties in the delivered SBRT dose estimates

Deviations in the delivered SBRT doses from the intended plans occur despite the use of IGRT. These can result from interpreting images with artifacts (e.g. blurry CBCT caused by respiration), the limitations of tumor surrogates, and the inability to completely correct for all anatomical and physiological changes using only one translation of the treatment couch. In Chapter 3, retrospective reconstruction of the delivered SBRT dose was performed using DIR of 4D CBCT. Compared to the planned static dose distributions calculated on exhale 4D CT, which remains as the current standard of care at PMH, 70% of patients had deviations in the delivered dose of >5%, ranging from -15% to 5% in tumors and -42% to 8% in normal tissues. When the
planning dose distributions incorporated the 4D CT breathing motion with DIR, deviations in the delivered doses >5% relative to those 4D CT-predicted doses still occurred in 53% of patients.

DIR modeled the motion and deformation of anatomy seen on 4D CBCT. Dose accumulation was based on the tracked motion of tissues within the dose grids calculated on the planning 4D CT, assuming the planning dose distribution is invariant to inter-fraction translations and deformation. Delivered dose accumulation studies for other anatomic sites suggest this approach is valid however future work should confirm this simplification for liver SBRT. It would be feasible to deform the exhale 4D CT image intensities into the geometry of the CBCT and then re-calculate the dose grids at each fraction. Correction factors have also been developed to calibrate CBCT intensities with those from CT, enabling direct dose grid calculations on CBCT to be accurate within 2–3% of the CT calculations[176].

Commissioning and evaluation of a commercial DIR and dose reconstruction platform is currently underway at PMH. Some logistical barriers observed in the current work need to be addressed for dose reconstruction to be used routinely in practice. Identifying and transferring the required data (e.g. the correct CBCTs, planning contours, dose grids etc.) between the IGRT software, treatment planning system and DIR/dose reconstruction platform is currently cumbersome. Efforts to automate these processes are not insurmountable. Contouring of the treatment CBCTs was also done manually in this thesis. Auto-segmentation tools developed for Pinnacle have reported mean DSC for CT images of liver of 0.94±0.01 compared to clinician contours[177], indicating quality contours could be produced automatically for biomechanical DIR. Therefore, efforts to streamline the workflow should facilitate clinical implementation of deformable dose reconstruction.

Dosimetric impact of advanced IGRT

Residual setup errors following image-guidance based on 3D CBCT were most often the largest cause of the delivered dose deviations relative to the 4D CT-predicted dose (reported in Chapter 3). Therefore the impact of advanced IGRT methods was explored in Chapter 4. Liver SBRT with 4D CBCT guidance based on rigid liver registration reduced the residual population GTV targeting errors by approximately half, compared to the 3D CBCT guidance strategy used to treat the patients. The planned dose distribution incorporating 4D CT breathing motion better
correlated with the delivered doses when the 4D CBCT strategy was simulated versus the 3D CBCT strategy used.

This research retrospectively sorted the 3D CBCT into 4D phases and simulated image-guidance by aligning the liver centre-of-mass on the exhale 4D CBCT to the planning exhale 4D CT. In 2012, 4D CBCT guidance based on rigid liver registration to the exhale phase was implemented clinically at PMH for liver SBRT for patients using commercially available IGRT software. Prospective 4D CBCT acquisition takes approximately 1–2 min longer than 3D acquisition in order to better sample the breathing motion and minimize residual artifacts[93]. Automated intensity-based rigid registration at exhale is then performed on the liver, mimicking the IGRT strategy simulated in Chapter 4. In practice the radiation therapists have the additional ability to refine the registration to the tumor-bearing liver region to improve targeting in the presence of deformations or rotations, or avoid normal tissues. The results of the current study have quantified the geometric and dosimetric gains in the use of 4D CBCT for liver SBRT, and the limitations in using the liver as a surrogate for the GTV position.

Directly targeting the GTV on each exhale 4D CBCT was also simulated assuming biomechanical DIR would be available prospectively for IGRT. This strategy further reduced GTV targeting errors compared to 4D CBCT and 3D CBCT based targeting of the liver, and further improved the correlation between the delivered and 4D CT-predicted breathing dose distributions. The gains observed with the addition of DIR to 4D imaging however were not as substantial as gains observed with the use of 4D over 3D CBCT. It is worth noting that decreases in the delivered minimum GTV dose were rare for all three IGRT strategies, likely because of the adequate PTV margin. Even with the current margins DIR could be reserved for situations where the residual position of multiple GTVs on CBCT is difficult to interpret because of liver deformation. At present, online biomechanical DIR would require approximately 5–10 min to manually contour the liver on CBCT and <1 min to run Morfeus.

Despite the use of the advanced IGRT strategies to reduce targeting errors, approximately half the patients still had dose deviations >5% relative to the 4D CT-predicted dose. This was expected given the dosimetric interplay between setup errors, deformation and breathing variations observed in Chapters 3 and 4. These results bring light to the fact that improvements in IGRT may be ultimately limited by the degree to which residual deformations and breathing...
variations can be corrected or their impact minimized. This should also apply to other IGRT initiatives including MRI-guided radiotherapy facilities under development.

Validation of deformable dose reconstruction

The geometric accuracy of Morfeus biomechanical DIR has been previously evaluated to be within 2 mm based on clinical images from liver SBRT patients. Those studies relied on a sparsely identified vessel bifurcations spread throughout the liver. Validation on a larger sample of points is not possible with that approach, and the impact of those uncertainties on dose reconstruction is unclear. In Chapter 5, biomechanical DIR registered images of a deformable dosimeter, and the deformed dose predicted with DIR was compared to the delivered dose measure with optical CT. Morfeus predicted the deformed fields’ centroids with a mean (range) accuracy of 1.0±0.5 (0.0, 2.4) mm. The shape of the deformed fields’ was predicted with an accuracy of 0.2±1.3 (-3, 3) mm and -0.3±0.7 (-2, 1) mm parallel and perpendicular to the applied deformation respectively. The deformed dose distribution was accurately predicted in greater than 90% of the voxels throughout the dosimeter (γ3%/3mm=91%). This study supports deforming dose via biomechanical DIR, and the results complement the previous study by Niu et al. which validated dose accumulation under breathing-like 4D conditions[136].

An unresolved issue is how to relate the geometric accuracy measured in patient images to the dosimetric accuracy measured in deformable phantoms. For example, translating the results from Chapter 5 back to the clinical data is difficult. However, somewhat conservative dose metrics were reported when comparing deviations between the delivered and the planned (or 4D CT-predicted) SBRT doses. It was assumed that deviations >5% of the prescribed dose would be potentially larger than the uncertainties caused by DIR. Maximum and minimum doses were also reported to 0.5 cm³. Given that the dosimeter data was evaluated with fairly strict criteria (e.g. γ3%/3mm over voxels of 0.1 cm³) the clinical results reported throughout this thesis appear to be within the dose accumulation uncertainties due to Morfeus.

Studies such as Chapter 5 however can be important during DIR commissioning prior to clinical implementation, even if they may not be feasible to perform on a patient-specific basis. If the geometric and dosimetric uncertainties in a DIR algorithm are well characterized upfront, the geometric accuracy may only need to be assessed on a patient-specific basis (e.g. spot checking the vessel bifurcations). For this to be valid the conditions between the two scenarios should be
similar. Therefore, it would be beneficial for future work to construct anatomically correct abdominal phantoms. It should be possible for deformable dosimeter to be fabricated into the shape of a patient’s liver, and irradiated with an actual clinical SBRT plan over multiple fractions. This would improve our understanding of how DIR-based dose reconstruction uncertainties manifest in a more clinically plausible scenario.

*Optimizing the therapeutic ratio in liver SBRT*

The study in Chapter 3 revealed that despite residual inter-fraction geometric errors following CBCT-guided free-breathing liver SBRT, only a minority of patients had substantial decreases in the accumulated, delivered tumor doses. It was theorized that the clinical PTV used, incorporating the full breathing motion, nullified many of these errors and potentially smaller PTVs could be adequate. Furthermore, the delivered doses correlated better to the planning dose distribution when it incorporated breathing motion (i.e. 4 D CT-predicted doses) compared to the static dose calculation. These static dose calculations on the exhale 4D CT, still widely performed clinically, result from current treatment planning systems’ inability to incorporate full 4D CT datasets into dose calculations and the lack of clinically-available DIR algorithms. These points motivated the investigation into a planning technique that could potentially enable smaller PTVs while also better representing the patients’ anatomy during respiration, potentially minimizing the need for 4D CT-predicted dose distributions.

The purpose of Chapter 6 was to validate a technique to determine the mean respiratory position of liver tumors and to quantify the potential for planning target volume reduction using the dose-probability PTV proposed by van Herk[67]. Previous methods suggested to establish the mean lung tumor position are either complex or require DIR on full 4D CT datasets[164-166]. Using a single DIR between the extreme 4D CT phases, mid-position CT datasets were reconstructed with mean (max) errors of 0.6±0.3 (1.4) mm versus the GTV’s time-averaged position determined using DIR on the full 4D CT dataset. Additionally, a well-selected phase (i.e. the mid-ventilation CT) from the initial 4D CT dataset could also accurately represent the mean liver GTV position because hysteresis on 4D CT appeared small in the patients studied. This mid-ventilation CT, selected based on simple rigid registration of the diaphragm, had errors of 1.0±0.5 (2.1) mm versus the GTV’s time-averaged position. Compared to ITV-based PTV on exhale 4D CT, dose-probability PTV on CTs representing the mean breathing position reduced the irradiated volume by 34±7% on average. The accuracy of the mid-ventilation CT, which is
more representative of the overall breathing motion than end-exhale 4D CT, is within the image resolution and can be readily applied in current treatment planning systems without DIR. A limitation of this study was using DIR to model the GTV’s time-averaged position as the ground truth, as opposed to direct image-based measurements which was not possible on 4D CT in the majority of patients. Future work should validate these strategies using liver SBRT patients with larger breathing motion, as most of these patients were triaged to be treated with ABC.

Subsequent to the work performed in this thesis, researchers from The Netherlands Cancer Institute recently validated the mid-position CT for liver SBRT using optical flow DIR. Kruis et al. evaluated 11 patients each with 1 or 2 fiducial markers implanted in the liver which were digitally masked out prior to intensity-based DIR between the exhale 4D CT and the 9 remaining 4D CT phases[129]. Each phase was deformed into the mid-position CT using the 4D deformation map. DIR accuracy, evaluated by comparing the DIR-predicted position of the masked-fiducial images to their actual displacement, was 1.8±0.6 mm. By averaging the voxel intensities over all individual phases, their mid-position CT had 3 times less noise than individual phases potentially permitting direct GTV delineation. Contrast-enhanced exhale breath-hold CTs were also deformed to the mid-position CT with an accuracy of 3.0±0.9 mm.

The research in this thesis and that by Kruis et al. is complementary, and the choice between the mid-position or mid-ventilation CT likely depends on the institution’s method of GTV delineation. When the GTV is primarily defined by contrast-enhanced breath-hold CT and MRI, the mid-position CT requiring the use of DIR may not be necessary. Even with the much simpler mid-ventilation CT strategy however, DIR may be useful when multiple GTVs visualized on MRI or tri-phasic CT (both at exhale) need to be mapped to the mid-ventilation CT position. Respiration-induced liver deformation causing differential GTV motion up to 6 mm SI would likely limit the accuracy of rigid registration to map the GTVs in such cases to the mid-ventilation CT.

By planning at the mean respiratory position instead of at the extreme end-exhale 4D CT phase, the systematic geometric error related to respiration is eliminated. For 20 patients investigated in Chapter 7, the residual patient-specific breathing random motion (i.e. the SD of motion around the mean position) was then incorporated into a dose-probability PTV designed for liver SBRT, and applied to the GTV on the mid-position CTs. Based on the results obtained in Chapter 4, it was assumed that a delivery strategy using 4D CBCT and rigid liver registration could also be
used to target liver tumors within a reasonable 2 mm accuracy. Errors in GTV position following rigid 4D CBCT liver alignment, calculated on a population of patients receiving liver SBRT in Chapter 4, were additionally incorporated into the dose-probability PTV to account for residual uncertainty caused by liver deformation. The dose-probability PTVs applied on mid-position CT were on average 38% smaller than the ITV-based PTV applied on the exhale 4D CT, enabling an mean±SD increase in the planned dose to 95% of the PTV of 4.0±2.8 Gy (9±5%) on the mid-position CT. Dose probability PTV can be applied to either the mid-ventilation CT or mid-position CT because they both represent the mean breathing position. No other changes at planning are anticipated to be required to implement this technique. Dose escalation planned for an equivalent risk of RILD could potentially improve local control, provided that the reduced PTVs do not increase the risk of under-dosing the GTVs during SBRT delivery.

To evaluate this, the delivered doses were modeled with deformable dose reconstruction of the daily CBCTs based on liver targeting at the exhale or mean breathing position, for the exhale 4D CT and mid-position CT plans respectively. For both plans, the delivered minimum GTV doses were greater than the planned nominal prescribed dose in all 20 patients, and greater than the planned dose to 95% of the PTV in 18 (90%) patients indicating equivalent rates of tumor dosage. The dose-probability PTV achieved the goal of delivering the intended GTV dose to at least 90% of patients as per its design. Notably, the much larger ITV-based PTV plans did not guarantee 100% of the patients’ GTVs would receive the planned dose. These ITV-based PTVs are far more commonly used in liver SBRT than dose-probability PTV. The importance of these findings is strengthened by the fact that the delivered doses were modeled with 4D CBCT and rigid registration of the liver. The clinical IGRT systems currently have the capability to perform automated image-guidance to the mean 4D CBCT liver position, and this process does not take any addition time versus guidance based on the exhale 4D CBCT position. When combined with the mid-ventilation CT strategy for planning, this IGRT strategy has the potential to be readily implemented without the necessity of DIR. Unlike strategies such as gating or tracking which have been investigated to reduce PTVs, the GTV’s position within the breathing range does not need be continuously tracked or modeled during irradiation.

Nine patients (45%) had one or more GTVs with a delivered minimum dose more than 5 Gy higher with the mid-position CT plan using dose-probability PTV, compared to the delivered
dose with the exhale CT plan using ITV-based PTV. It is predicted that the delivered doses estimated with DIR of 4D CBCT more accurately model the true delivered dose and should in theory better relate to patient outcomes.

8.2 Future Directions

The research in this current thesis used DIR and dose accumulation to evaluate the delivered dose from different SBRT planning and delivery strategies. With 4D CBCT image guidance based on either rigid liver registration or DIR, the inability to correct for residual liver deformation online resulted in deviations in the delivered dose to tumors and normal tissues. However the ability to adapt the treatment plan based on the accumulated doses at each fraction has not yet been explored. Leinders et al. recently reported on the feasibility of daily re-optimization of IMRT plans for liver SBRT to account for inter-fraction deformation[147]. This approach ignores the previously delivered doses by optimizing the treatment plan for that specific fraction, and hence does not require DIR and dose accumulation.

It would be interesting to fully exploit the information provided by deformable dose reconstruction by re-optimizing the IMRT plan at each fraction with consideration of the doses delivered prior to that fraction. This strategy, termed ‘adaptive radiation therapy’, could allow for further PTV reduction in addition to the use of dose-probability PTV investigated in Chapter 7. If the treatment plan is continuously adapted to account for residual deformations, the expansion incorporated into the dose-probability PTV to account for deformation could be eliminated, resulting in greater normal tissue sparing and the potential for dose-escalation.

Abandoning dose-probability PTV margins and replacing them with robust 4D optimization has also been proposed as a means of sparing more normal tissues[178]. This strategy relies on a probabilistic optimization whereby the breathing motion is directly considered during IMRT optimization. PDFs for the breathing motion can be constructed by means of DIR and a 4D CT dataset. This strategy has not yet been explored for liver SBRT. In relation to this thesis, it would be additionally possible in theory to incorporate a population PDF of inter-fraction liver deformation measured using DIR of 4D CBCT to further spare normal tissues. Unlike PTVs, which are robust to motion of the GTV only, this strategy would account for motion of both the GTVs and normal tissues.
8.3 Clinical Impact on Liver Radiotherapy

In Chapter 7 the impact of dose-probability PTV was quantified at planning by escalating the prescription dose beyond what could be allocated with the larger ITV-based PTVs. This was based on clinical data suggesting a dose-response relationship exists for liver tumors which is strongest for colorectal liver metastases. For HCC, clinicians may not prescribed doses as high as the dose-allocation schema allows particularly for patients with poor pre-SBRT liver function as they are at an increased risk of non-RILD toxicities[179]. In such scenarios the application of reduced dose-probability PTV would spare >38% more normal tissue volume on average than the clinical PTVs, potentially decreasing the risk of toxicity for an equivalent prescribed dose. Sparing normal tissues via PTV reduction could also potentially increase eligibility for SBRT in patients that may not meet the technical planning criteria otherwise.

Deformable dose reconstruction was used specifically throughout this thesis to quantify the delivered doses for previously treated patients and to evaluate several proposed SBRT planning and delivery techniques. A larger hypothesis, not tested in this thesis, is that a better understanding of the delivered doses by means of deformable dose accumulation will improve our understanding dose-effect relationships. The optimal doses to treat liver tumors are not known, and the tolerance of normal tissues to hypo-fractionated regimes is not well understood. At PMH dose is based on a predefined risk of RILD. The patients studied in Chapter 3 for example had a planned mean (range) NTCP of 1.8% (0–15.2%) calculated using the normal liver dose on the static distributions. Despite hundreds of patients having been treated with this regime no RILD has been observed, suggesting it is a safe approach. On the other hand the NTCP model is likely overly conservative, and inaccurate possibly due to its parameters being derived from a different radiation dose/fractionation scheme. The delivered normal liver doses reconstructed from the CBCTs used for SBRT delivery changed the NTCP post-SBRT by a mean±SD (range) of -0.5±2.5% (-8.3 to 8.0%). This suggests that deformable dose reconstruction can improve our understanding of radiobiological responses by reducing the physical uncertainties in SBRT delivery.

It is envisioned that deformable dose reconstruction will eventually play an important role in routine quality assurance for liver SBRT as well as other disease sites. Clinicians may be faced with new challenges in deciding how to interpret delivered dose deviations. For simplicity the total delivered doses over 6-fraction SBRT were reported in this thesis, however the data can
also be viewed in time-series demonstrating the dose deviation up to each fraction as well as the
projected final delivered doses. Specific guidelines on how to adapt to potential dose deviations
during the course of SBRT do not yet exist, and strategies need to be evaluated. However, the
dose deviations reported in this thesis demonstrate the baseline variation in delivered doses for
free-breathing liver SBRT patients.

The results of this thesis hopefully illustrate that the technical details of liver SBRT play an
important role in its design and application. There is also evidence suggesting that the technical
quality of SBRT affects clinical outcomes. Bujold et al. analyzed SBRT for HCC using data
combined from the Phase I and II studies conducted at PMH[56]. Being on the latter trial was
significantly associated with better overall survival (21% for Phase I vs. 51% for Phase II, at 2
years). This may have resulted from less advanced HCCs being treated in Phase II, but it also
suggests that clinical experience in liver SBRT planning and IGRT can improve outcomes. Poor
treatment plan quality has been linked to statistically inferior clinical outcomes in retrospective
reviews of Phase III clinical trials for non-liver cancers. In a randomized trial of chemoradiation
for head and neck cancers for example, the impact of quality radiotherapy (i.e. evaluated as a
function of protocol compliance) was significantly greater than the potential incremental gains
from novel drugs[180]. In a randomized trial of chemoradiation for pancreatic cancer, protocol
deviations similarly resulted in poorer local control and increased toxicity in the experimental
arm[181]. These results, based primarily on the quality of the initial clinical treatment plan (i.e.
the static dose distribution), helped explain the outcomes in a subset of patients. As the
delivered SBRT doses deviated from the static planned dose distribution in 70% of patients
(Chapter 3), dose reconstruction could be applied to explain the outcomes for a greater
proportion of patients in trials for liver SBRT. As an example, a recently opened multi-
institution Phase III trial of Sorafenib plus SBRT for HCC permits a range of IGRT
technologies. Deformable dose reconstruction could also be applied to account for the
anticipated uncertainties in the delivered doses and aid in interpretation of outcomes.

This is in part currently being investigated at PMH. Swaminath et al. is investigating the dose-
response for patients with liver metastases treated with SBRT[182]. One aim is to determine
how local control relates to dose metrics either from the planned static dose calculation (e.g. the
PTV dose), or the delivered doses reconstructed with Morfeus-DIR and CBCT images (e.g.
minimum accumulated GTV dose). On a multivariate analysis of 172 GTVs in 93 patients the
primary diagnosis of breast cancer and higher SBRT doses had significantly better local control irrespective of GTV size. Delivered min GTV doses of <35, 35–45 and >45 Gy resulted in 18 month local control rates of 33%, 55%, and 83% respectively. Notably each 5 Gy increase in the delivered min GTV dose resulted in a greater reduction in the risk of local progression compared to an equivalent 5 Gy increase in min PTV dose calculated on the static dose distribution ($p=0.003$). This suggests the dose response relationship for local control is steeper with the delivered doses versus static planned doses. Even in this single institution analysis, reconstructing the delivered doses helped reduce noise caused by deviations in the delivered doses. This potentially allows for other clinical factors (e.g. primary diagnosis) to be identified.

8.4 Summary

In this thesis, deformable image registration was used as a tool to characterize the anatomical and physiological changes occurring in free-breathing liver SBRT patients. Applying deformable registration at SBRT planning to more accurately model the breathing doses or to plan on more optimal datasets, and to better target liver tumors with CBCT was demonstrated. Strategies to optimize liver SBRT planning and image-guidance with 4D imaging and widely available rigid registration were also evaluated with promising results. However, delivered normal tissue doses were observed to deviate substantially from the planned doses irrespective of the specific SBRT technique used. Deformable dose reconstruction can therefore model daily anatomic variations into the distribution and improve calculations of the delivered doses in liver SBRT. There is evidence to suggest that the delivered doses can help interpret clinical outcomes by accounting for residual uncertainties in SBRT delivery.
References


Copyright Acknowledgements

Permission was obtained through the Copyright Clearance Centre on March 7, 2014 from the publisher Elsevier to reproduce several published works for inclusion in this thesis.

Chapter 3 was reproduced from a full article published as: Michael Velec, Joanne L. Moseley, Tim Craig, Laura A. Dawson, Kristy K. Brock. Accumulated dose in liver stereotactic-body radiotherapy: positioning, breathing and deformation effects. *International Journal of Radiation Oncology, Biology, Physics* 2012;83(4):1132-40. License number: 3343840329197

Figure 5.6.B was reproduced from one figure previously published in: Titania Juang, Shiva Das, John Adamovics, Ron Benning, Mark Oldham. On the need for comprehensive validation of deformable image registration, investigated with a novel 3-dimensional deformable dosimeter. *International Journal of Radiation Oncology, Biology, Physics* 2013;87:414-421. License number: 3343841037952

Chapter 6 was reproduced from a full article published as: Michael Velec, Joanne L. Moseley, Kristy K. Brock. Simplified strategies to determine the mean respiratory position for liver radiation therapy planning. *Practical Radiation Oncology* 2013, 4: 160-166. License number: 3343840645937

Permission was also obtained through the Copyright Clearance Centre on July 11, 2014 from the publisher Elsevier to reproduce Chapter 7 from a full article published as: Michael Velec, Joanne L. Moseley, Laura A. Dawson, Kristy K. Brock. Dose-escalated liver stereotactic body radiotherapy (SBRT) at the mean respiratory position. *International Journal of Radiation Oncology, Biology, Physics* 89(5): 1121-1128, 2014. License number: 3425980357493