

# **Patient-specific planning target volume margins for liver stereotactic robotic radiosurgery**

By

**Ming Liu**

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs  
in partial fulfillment of the requirements for the degree of

**Doctor of Philosophy**

in

**Physics**

**Specialization in Medical Physics**

Ottawa-Carleton Institute for Physics  
Department of Physics  
Carleton University  
Ottawa, Ontario, Canada  
© 2020  
Ming Liu

## **Abstract**

**Purpose:** To create analytical tools for a proposed clinical workflow, which implements patient-specific planning target volume (PTV) margins for liver radiation therapy treatments.

**Methods and Materials:** Treatment log files are analyzed for liver treatments to assess tumour motion-tracking accuracy. A uniform PTV margin is estimated that considers motion-tracking errors and deformations, provided that the impact of uncorrected rotations is minimized. A supervised machine learning algorithm employing retrospective data, which emulates a dry-run session prior to treatment planning, is used to investigate if motion-tracking errors are less than 2 mm, and consequently, the standard PTV margins can be reduced by 2 mm. For a safer implementation in the clinic, we employ a warning system that quantifies the probability of a geographic miss if the PTV margin is reduced for every subsequent fraction. A dosimetric analytical tool is proposed to retrospectively assess the dose to a target against different types of delivery errors. The tool is validated by radiochromic film measurements for two very different types of treatments, liver and trigeminal neuralgia (a cranial nerve disorder). Case studies are conducted to assess the suitability of a selected PTV margin based on the geometrical and dosimetric coverage of targets. The range of rotations that can be safely allowed for trigeminal neuralgia treatments is quantified since rotational corrections cannot be applied by the system for this specific disease site.

**Results:** Isotropic 4 mm PTV margins are sufficient to account for tracking errors and deformations for 95% of patients. For patient-specific PTV margins, the accuracy of predicting if motion-tracking errors are less than 2 mm is  $0.84 \pm 0.06$  using 5-fold cross-validation. Using the warning system, 11 out of 64 cases predicted to be treated with 2 mm reduced PTV margins might

## *Abstract*

---

require replanning, but for each fraction they have more than 96% of target(s) encompassed by the reduced PTV. For experiments with different types of geometrical errors the dose measurements with radiochromic film agree well (2%/2 mm level) with the dose distributions estimated using the dosimetric analytical tool. Dose to targets considering delivery errors can be significantly improved if treatments are planned following certain guidelines. For trigeminal neuralgia treatments, target rotations of up to  $1^\circ$  can be safe for some patients.

**Conclusions:** For treatment adaptation it is feasible to implement patient-specific PTV margins in the clinic, assisted with an early-warning system and dosimetric analytical tool to warn of a potential geographic miss and underdosing of target(s).

## **Acknowledgements**

To my supervisors and mentors, Drs. Joanna Cygler and Eric Vandervoort, I would like to thank you all for your incredible support in my study. Your patience of guidance, especially in helping me improve my programming, presentation and scientific writing skills are always very appreciated. To Dr. Cygler and her family, your kindness and love have always made me feel at home.

Thanks to my parents, siblings, and friends for their warm support and understanding throughout my Ph.D. study. Your everlasting support have always motivated me to work hard and enjoy my life in Canada. Special thanks to Christopher Dydula, Lingyue Sun, Ericka Venturina, and Keren Mayorov, who have always given me great support these years.

Thanks to Dr. Dal Granville, the medical physicist at the Ottawa Hospital Cancer Centre (TOHCC) for supporting the initial stage of the machine learning project. Thanks to the radiation oncologists Drs. Graham Cook and David Tiberi, the radiation oncologist fellow Dr. Kadir Mullings, the clinical physicist Dr. Byron Wilson, the radiation therapist Wendy Yue, and the CyberKnife® team at our clinic for the technical support in my project. Thanks to Lisa Goggin and Fabienne Hirigoyenberry from Accuray Inc, who provided some technical specifications of the CyberKnife Synchrony® Respiratory Tracking System for the Synchrony log file study.

This work has been funded by the Ontario Consortium for Adaptive Interventions in Radiation Oncology (# RE-04-026), a research grant from Accuray Inc., CA, the Department of Medical Physics at TOHCC, the Physics Department at Carleton University, Ontario Graduate Scholarship, John Lyndhurst Kingston Scholarship, and Carleton University 75<sup>th</sup> Anniversary Scholarship, Robert L. Clarke Graduate Scholarship, and Bader Student Travel Scholarships.

## *Acknowledgements*

---

Thanks to the examination board: Issam El Naqa (external examiner; H. Lee Moffit Cancer Centre), Matthew Holden (Computer Science, Carleton), Delphine Gourdon (Physics, University of Ottawa), Rowan Thomson (Physics, Carleton University), Joanna Cygler (supervisor, Physics, TOHCC), Eric Vandervoort (supervisor, Physics, TOHCC), and Sean Barry (chair of defence, Chemistry, Carleton University). I appreciate their valuable and constructive comments and discussions on my thesis work.

## **Statement of Originality**

The author confirms that she is the only author of the written work here enclosed and that it consists of the author's research throughout her doctoral program at Carleton University. Parts excepted are corrections of form and content by the thesis supervisors, Drs. Joanna Cygler and Eric Vandervoort, who are clinical physicists working at the Ottawa Hospital Cancer Centre, where the entire thesis work has been conducted.

Drs. Joanna Cygler and Eric Vandervoort were involved in all aspects of the thesis work, including but not limited to revising the author's written work (the thesis, manuscripts, and conference abstracts), and conducting the experiments. All the in-house Python and MATLAB scripts for data processing used in the thesis work were produced by the author, except that the MATLAB scripts used for analyzing the measured dose using radiochromic films and for gamma analyses were written by Dr. Eric Vandervoort. Using the CyberKnife Precision treatment planning system, Eric created the planned dose distributions for all experimental validations (Chapter 4) and case studies of liver treatments (Chapter 5). Drs. David Tiberi, Kadir Mullings, and Eric Vandervoort altogether completed the delineation of structures of interest for the trigeminal neuralgia project (Chapter 5) using the treatment planning system.

All patient data were collected as part of a local clinical trial approved by the Ottawa Hospital Research Ethics Board. The Carleton University Research Ethics Board-B (CUREB-B) has granted ethics clearance for the entire project. The author has completed the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans Course on Research Ethics (TCPS 2: CORE).

As outlined below, most of the thesis work has been presented at conferences and published as conference abstracts. Most of the content in Chapters 2 and 3 was published in peer-reviewed

journals to which the author made major contributions. The author's contributions to the peer-reviewed papers and the conference abstract published in *Gastrointestinal Endoscopy* (as the second author) are elaborated below. Two papers whose contents make up Chapters 4 and 5 are in preparation.

### **Peer-reviewed papers**

- I. **M. Liu**, J. E. Cygler, E. Vandervoort, "Patient-specific PTV margins for liver stereotactic body radiation therapy determined using support vector classification with an early-warning system for margin adaptation." *Med Phys.* **47** (10), 5172-5182, 2020. Most of the published work has been reused in Chapter 3. The author's contributions included but were not limited to data extraction, data processing, building a pipeline for the machine learning model, testing the model, designing the clinical workflow assisted by the model, and statistical analyses.
- II. **M. Liu**, J. E. Cygler, E. Vandervoort, "Geometrical tracking accuracy and appropriate PTV margins for robotic radiosurgery of liver lesions by SBRT." *Acta Oncologica.* **58** (6), 906-915, 2019. Most of the published work has been reused in Chapter 2. The author's contributions included but were not limited to data collection, data processing, and statistical analyses.

### **Conference abstracts (\*indicates presenting author)**

#### **Oral presentations**

- I. \***M. Liu**, A. Ross, J. E. Cygler, and E. Vandervoort. "TH-A-SAN2-10: Adaptive Margins with An Early Warning System for Motion-Tracking Errors in Liver SBRT." *Med. Phys.* **46**(6), 499-500, 2019. Presented at the 61<sup>st</sup> American Association of Physicists in Medicine Annual Meeting, Jul. 2019, San Antonio, TX.

## Statement of Originality

---

- II. \***M. Liu**, D. Granville, J. E. Cygler, and E. Vandervoort. “SU-F-209-3: Predicting Motion Compensation Errors for CyberKnife SBRT Liver Patients Using a Machine Learning Algorithm.” *Med Phys.* **45** (6), 142, 2018. Presented at the 60<sup>th</sup> American Association of Physicists in Medicine Annual Meeting, Jul. 2018, Nashville, TN.
- III. \***M. Liu**, E. Vandervoort, and J. E. Cygler. “Abstract #169: Tracking accuracy of robotic radiosurgery for liver lesions.” **T19-057**, Pages 863-864. The Book of Abstract accessible at <https://guarant.topinfo.cz/iupesm2018/en/book-of-abstracts>. Presented at the World Congress on Medical Physics and Biomedical Engineering, June 2018, Prague, Czech Republic.
- IV. \***M. Liu**, E. Vandervoort, and J. E. Cygler. “TH-CD-205-10: Use of CyberKnife Log-files to Aid the Derivation of PTV Margins for Liver Treatments Using Synchrony Respiratory Compensation.” *Med. Phys.*, **44** (6), 3295, 2017. Presented at the 59<sup>th</sup> American Association of Physicists in Medicine Annual Meeting, Jul. 2017, Denver, CO.

### Poster presentations

- I. \***M. Liu**, J. E. Cygler, D. Tiberi, J. Doody, S. Malone, J. Pantarotto, E. Vandervoort, “BReP-SNAP-T-50: Dosimetric Sensitivity to Sources of Uncertainties for Liver Cancer and Trigeminal Neuralgia Treated with Ablative Radiosurgery”. *Med Phys.* **47** (6), e457, 2020. Accepted for a Blue Ribbon ePoster, Virtual AAPM COMP Joint Meeting, Jul. 2020, Vancouver, BC.
- II. F. Jowhari, **M. Liu**, E. Vandervoort, *et al.* “How Good Is EUS Guided Fiducial Placement for Targeted Cyberknife Radiotherapy?” *Gastrointestinal Endoscopy*, **87** (6), AB438, 2018. Accepted for a poster presentation at Digestive Disease Week (DDW), June 2018, Washington, DC. The author’s contributions included the data extraction from

gastrointestinal (GI) treatment log files and analyses of treatment motion-tracking related errors for 90 patients, for whom the fiducial tracking markers were implanted via Endoscopic ultrasound (EUS) or percutaneously.

- III. **\*M. Liu**, E. Vandervoort and J. E. Cygler. “Effect of translational and rotational errors on the accuracy of CyberKnife liver treatments.” *Med Phys.* **44** (8), 4383, 2017. Accepted for a poster presentation at the 63<sup>rd</sup> Canadian Organization of Medical Physicists Annual Scientific Meeting, June 2017, Ottawa, ON.
- IV. **\*M. Liu**, J. E. Cygler and E. Vandervoort. “Assessment of Synchrony respiratory compensation error for CyberKnife liver treatment.” *Med. Phys.* **43** (8), 4948, 2016. Accepted for a poster presentation at the 62<sup>nd</sup> Canadian Organization of Medical Physicists Annual Scientific Meeting, June 2016, St. John’s, NL.

### **Seminar presentations**

- I. **\*M Liu**, J. E. Cygler and E. Vandervoort, “Patient-specific PTV margins for liver SBRT with an early-warning system for margin adaptation.” Presented at the Ottawa Medical Physics Institute (OMPI) seminar, Nov. 2019.
- II. **\*M Liu**, J. E. Cygler and E. Vandervoort, “Accuracy of the CyberKnife Synchrony respiratory tracking system for liver cancer.” Presented at the Ottawa Medical Physics Institute (OMPI) seminar, Feb. 2017.

---

## Table of Contents

<b>Abstract</b> .....	<b>i</b>
<b>Acknowledgements</b> .....	<b>iii</b>
<b>Statement of Originality</b> .....	<b>v</b>
<b>Table of Contents</b> .....	<b>ix</b>
<b>List of Tables</b> .....	<b>xiii</b>
<b>List of Figures</b> .....	<b>xiv</b>
<b>List of Abbreviations and Symbols</b> .....	<b>xvii</b>
<b>Chapter 1. Introduction</b> .....	<b>1</b>
1.1. Overview of Radiation Therapy.....	1
1.1.1. Liver Cancer .....	1
1.1.2. Radiobiological Mechanism.....	2
1.1.3. Physics of Photon Beam Radiation Therapy .....	4
1.1.4. Definition of Volumes.....	6
1.1.5. Patient Specific Quality Assurance and Dose Evaluation Metric .....	10
1.2. Target Motion Management .....	13
1.3. Robotic Stereotactic Body Radiation Therapy .....	27
1.3.1. CyberKnife Stereotactic Radiosurgery System .....	29
1.3.2. Synchrony Tracking System .....	33
1.3.3. Typical Clinical Workflow.....	34
1.3.4. Challenges of Accurate Treatment Delivery .....	37
1.4. Thesis Overview .....	43
<b>Chapter 2. Uniform PTV Margins</b> .....	<b>45</b>
2.1. Introduction.....	45
2.2. Method and Materials .....	46
2.2.1. Patient Data .....	46
2.2.2. Synchrony Errors.....	47
2.2.3. Rotational Corrections.....	48
2.2.4. Rigid Body Error .....	51
2.2.5. Model Linearity.....	51

## Table of Contents

---

2.2.6.	Hypothesis Testing.....	52
2.3.	Results.....	52
2.3.1.	Synchrony Error.....	52
2.3.2.	Rotational Corrections.....	56
2.3.3.	Rigid Body Error.....	61
2.3.4.	Proposed Uniform PTV Margin.....	63
2.3.5.	Model Linearity.....	63
2.3.6.	Hypothesis Testing.....	64
2.4.	Limitations.....	65
2.5.	Discussions.....	66
2.5.1.	Synchrony Error.....	66
2.5.2.	Rotational Corrections.....	68
2.5.3.	Rigid Body Error.....	70
2.5.4.	Model Linearity.....	71
2.6.	Conclusions.....	72
	<b>Chapter 3. Patient-Specific PTV Margins .....</b>	<b>73</b>
3.1.	Introduction.....	73
3.2.	Methods and Materials.....	75
3.2.1.	Data Collection and Analyses.....	75
3.2.2.	Synchrony Respiratory Tracking System.....	76
3.2.3.	Margin Reduction Methods.....	77
3.2.4.	Proposed Clinical Workflow.....	77
3.2.5.	Features and Binary Labels.....	80
3.2.6.	Support Vector Classification (SVC).....	81
3.2.7.	Parameter Tuning and Feature Selection.....	82
3.2.8.	Classifier Evaluation.....	83
3.3.	Results.....	84
3.3.1.	Planning Data Analyses.....	84
3.3.2.	Simple Two-Feature Analysis.....	85
3.3.3.	Parameter Tuning and Feature Selection.....	86
3.3.4.	Classifier Evaluation.....	87
3.3.5.	Regression Evaluation.....	88
3.3.6.	Early Warning System.....	89

## *Table of Contents*

---

3.4.	Discussion .....	93
3.4.1.	Planning Data Analysis .....	93
3.4.2.	Simple Two-Feature Analysis .....	93
3.4.3.	Classifier Evaluation Using More Features.....	94
3.4.4.	Early-Warning System .....	96
3.5.	Conclusions.....	98
<b>Chapter 4. Experimental Validations of the Dose Perturbation Tool .....</b>		<b>99</b>
4.1.	Introduction.....	99
4.2.	Methods and Materials.....	103
4.2.1.	Phantoms .....	103
4.2.2.	Respiratory Motion Platforms .....	107
4.2.3.	Film Dosimetry Protocol .....	108
4.2.4.	Dosimetric Perturbation Method.....	112
4.2.5.	Dose Penumbra Broadening Effects.....	120
4.2.6.	Experimental Validations .....	121
4.3.	Results.....	125
4.3.1.	Rotational Errors .....	125
4.3.2.	Tracking Residual Errors.....	134
4.4.	Discussion .....	146
4.5.	Conclusions.....	151
<b>Chapter 5. Dosimetric Assessments .....</b>		<b>153</b>
5.1.	Liver.....	153
5.1.1.	Introduction .....	153
5.1.2.	Method and Materials.....	156
5.1.3.	Results .....	157
5.1.4.	Discussion .....	174
5.2.	Trigeminal Neuralgia .....	178
5.2.1.	Introduction .....	178
5.2.2.	Method and Materials.....	181
5.2.3.	Results .....	185
5.2.4.	Discussions.....	192
5.3.	Conclusions.....	194

---

<b>Chapter 6. Final Overview and Future Work .....</b>	<b>195</b>
6.1. Summary .....	195
6.2. Future Work .....	198
<b>Appendix A. Different Coordinate Systems used with Cyberknife .....</b>	<b>202</b>
<b>Appendix B. Translation and Rotation Matrix .....</b>	<b>203</b>
<b>Appendix C. GTV Voxel Occupancy Probability .....</b>	<b>205</b>
<b>Appendix D. Phase Shift Between Internal and External Motions .....</b>	<b>208</b>
<b>Appendix E. X-ray Imaging Dose (Liver Insert).....</b>	<b>209</b>
<b>Appendix F. Screenshots of the Precision Treatment Planning Interface.....</b>	<b>211</b>
<b>References .....</b>	<b>213</b>

## **List of Tables**

Table 2-1. Fiducial placement characteristics for 72 liver treatments	60
Table 4-1. Characteristics of experiments using the rocket phantom to investigate impacts of uncorrected rotations on dose distributions	126
Table 4-2. Characteristics of experiments for validating the dose perturbation model	135
Table 5-1. Several potentially relevant quantities for underdosing of targets	164
Table 5-2. D99 of target(s) considering different delivery errors and for the original plan for three case studies	169
Table 5-3. Characteristics of the patient-specific dosimetric assessment prior to a trigeminal neuralgia treatment	190
Table A-1. Different coordinate systems used with CyberKnife	202
Table E-1. Mass attenuation coefficients, density, and path length for phantom materials	210

## **List of Figures**

Figure 1-1. A surface rendering of a patient CT with treatment beams	7
Figure 1-2. 2D illustration of target volumes commonly used in treatment planning for upper abdominal tumours	9
Figure 1-3. A CT slice for a liver lesion with treatment planning contours	10
Figure 1-4. Two fiducials identified in CT images	15
Figure 1-5. Immobilization devices at the Ottawa Hospital Cancer Centre	17
Figure 1-6. CyberKnife G4 Stereotactic Radiosurgery System at the Ottawa Hospital	29
Figure 1-7. The interface of the Synchrony Respiratory Tracking System	34
Figure 1-8. The prescription dose and PTV D99 for 148 treatment plans	41
Figure 2-1. Synchrony errors and inter-model standard deviations of radial errors	54
Figure 2-2. Inter- and intra- fractional variations of correlation and predictor errors	55
Figure 2-3. The offsets of correlation and predictor errors across eight respiratory phases	56
Figure 2-4. Rotational corrections grouped by treatment fraction for 23 patients	59
Figure 2-5. Inter- and intra- patient variations of rotational corrections for 23 patients	60
Figure 2-6. The mean, SD, and absolute values of signed rigid body errors	61
Figure 2-7. Inter- and intra- fractional variations of rigid body errors for 66 patients	62
Figure 2-8. The $R^2$ value per model between target and marker positions	64
Figure 3-1. The proposed clinical workflow for patient specific PTV margins	79
Figure 3-2. The demonstration of the marker displacement during treatment.	81
Figure 3-3. (a) The prescription dose and PTV D99 for compromised and non-compromised treatments. (b) D99 of $PTV_{std-2mm}$ and $PTV_{std}$ using the same dose distribution	85
Figure 3-4. The confusion matrix and the ROC metric	88
Figure 3-5. The relationship between predicted and true tracking errors	89
Figure 3-6. The minimum fraction of the perturbed GTV within the $PTV_{std-2mm}$ (a) and the $PTV_{std}$ (b), respectively, for predicted Class 1 and Class 2 for all fractions of treatment.	90
Figure 3-7. The fractional GTV voxel occupancy probability within the $PTV_{std-2mm}$ for eleven treatment courses.	91
Figure 3-8. GTV voxel occupancy probability for 148 plans using PTV margins predicted by a regression model	92

## *List of Figures*

---

Figure 4-1. Photos of phantoms used in the thesis work	106
Figure 4-2. The Synchrony platform used to create periodic motion for the phantom and markers moving in perpendicular directions	107
Figure 4-3. The experiment setup to create in-phase (a) and out-of-phase motion (b) for markers and a cylindrical insert using the Quasar phantom(s)	108
Figure 4-4. A flowchart for film dosimetry using EBT3 films	111
Figure 4-5. The distributions of correlation, predictor, and end-to-end errors in three anatomical directions observed in two experiments.	114
Figure 4-6. A flowchart of the dose perturbation tool accounting for intrafraction translational and rotational errors	117
Figure 4-7. A flowchart of the dose perturbation tool for targets taking into account rigid body errors.	119
Figure 4-8. Line profiles for the planned, perturbed, and measured dose	120
Figure 4-9. Demonstration of profile characteristics for dose measurements	121
Figure 4-10. The dose difference for the planned, perturbed, and measured dose for a trigeminal neuralgia plan delivered using 5 mm cone beams.	125
Figure 4-11. 2D gamma analyses for experiment R1	127
Figure 4-12. 2D dose difference for planned, perturbed, and measured dose of an “L”-shape target (experiment R2).	128
Figure 4-13. 2D gamma analyses for experiment R2	130
Figure 4-14. Spherical targets planned at the film plane and out-of-plane	131
Figure 4-15. 2D gamma analyses for experiment R3	133
Figure 4-16. 2D gamma analyses for an isocentric plan using a ball cube phantom	136
Figure 4-17. 2D gamma analyses for a non-isocentric plan using a lung insert	139
Figure 4-18. The relationship between target motion and marker motion for a linear model and a non-linear model	140
Figure 4-19. The patient waveform used for liver experiments	141
Figure 4-20. 2D gamma analyses for an experiment (liver insert, close-to-lung target)	144
Figure 4-21. 2D gamma analyses for an experiment (liver insert, far-to-lung target) with the large uncorrected rotational offset of 4.0 degrees.	145

## *List of Figures*

---

Figure 5-1 Target rotations, correlation errors, rigid body errors, and the planned and treated D99/Rx for 70 patients	160
Figure 5-2. D99/Rx for planning and treatments considering different delivery errors	163
Figure 5-3. D99/Rx for planning and treatments without considering rigid body errors	163
Figure 5-4. The difference between treated and planned D99/Rx	165
Figure 5-5. The relationship between $\Delta$ D99/Rx and the planned D99/Rx when a large rotational offset is present.	167
Figure 5-6. Geometric and dosimetric coverage for a GTV using 5 mm PTV margin	173
Figure 5-7. Geometric and dosimetric coverage for a GTV using 3 mm PTV margin	174
Figure 5-8. Transverse slices through the planned dose distributions and planning CT for representative liver and trigeminal neuralgia cases.	181
Figure 5-9. An illustration of good and poor alignments of the longer length of trigeminal nerve and the treatment target.	183
Figure 5-10. Workflow for patient-specific assessment prior to treatment to inform allowable rotations for prospective trigeminal neuralgia treatments	185
Figure 5-11. The planned and treated length of the nerve covered by Rx and the brainstem dose for 24 trigeminal neuralgia treatments.	187
Figure 5-12. (a) The treated and planned nerve length covered by Rx. (b) The planned and treated volume overlap between the nerve and Rx isodose.	187
Figure 5-13. Plan sensitivity to rotational offsets for the length of the left-sided and right sided nerve covered by prescription dose and toxicity to the brainstem	189
Figure E-1. A diagram to illustrate the axial plane of the phantom inserts	209
Figure F-1. A screenshot of the Precision treatment planning interface comparing the replanned and original dose distributions with one target	211
Figure F-2. A screenshot of the Precision treatment planning interface comparing the replanned and original dose distributions with two targets	212

## **List of Abbreviations and Symbols**

2D: two-dimensional

3D: three-dimensional

4DCT: four-dimensional computed tomography

ABC: active-breathing control

CBCT: cone-beam computed tomography

CI: confidence interval

CN5: trigeminal nerve

CT: computerized tomography

CTV: clinical target volume

D95: the minimum dose to 95% of a volume of interest

D99: the minimum dose to 99% of a volume of interest

DIBH: deep-inspiration breath hold

DICOM: Digital Imaging and Communications in Medicine

DIR: deformable image registration

DRR: digitally reconstructed radiograph

DTA: distance to agreement

DVH: dose volume histogram

E2E: end-to-end

GTV: gross target volume

HCC: Hepatocellular carcinoma

IGRT: image guided radiation therapy

IMRT: intensity modulated radiation therapy

ITV: internal target volume

kV: kilovoltage

kVp: x-ray tube voltage

linac: linear accelerator

## *List of Abbreviations and Symbols*

---

mAs: milliamperere-seconds  
MC: Monte Carlo  
MLC: multi leaf collimator  
MR: magnetic resonance  
MRI: magnetic resonance imaging  
MTS: motion tracking system  
MV: megavoltage  
NTCP: normal tissue complication probability  
OAR: organ at risk  
PC1: the first principal component  
PCA: principal component analysis  
PDF: probability distribution function  
PET: positron emission tomography  
PRV: planning risk volume  
PTV: planning target volume  
QC: quality control  
RMS: root-mean-square  
RMSE: root-mean-square error  
Rx: prescription dose  
SAD: source-to-axis distance  
SBRT: stereotactic body radiation therapy  
SD: standard deviation  
SRS: stereotactic radiosurgery  
SVC: Support vector classification  
SVM: Support vector machine  
TCP: tumour control probability  
TN: Trigeminal neuralgia  
V15Gy: volume of interest receiving 15 Gy  
XML: Extensible Markup Language

## **Chapter 1. Introduction**

### **1.1. Overview of Radiation Therapy**

#### **1.1.1. Liver Cancer**

Cancer ranks as the second leading cause of death worldwide, and liver cancer is one of the most common lethal cancers, together with lung, colorectal, stomach and breast cancers. About 3100 Canadians are predicted to be diagnosed with liver cancer in 2020, while 1450 will die as a result.<sup>1</sup> According to the US Cancer Statistics of 2009-2015, the 5-year survival rate for primary liver cancer is only 18%, which means less than one-fifth of patients can survive more than five years after the diagnosis of liver cancer.<sup>2</sup> Primary liver cancer consists of malignant tumours that originate from liver cells. Hepatocellular carcinoma (HCC) is the main liver malignancy that destroys hepatocytes, mostly arising from hepatitis B virus, hepatitis C virus, and autoimmune hepatitis.<sup>3</sup> However, liver metastases are more common than liver primaries.<sup>4</sup> Metastases consist of tumour cells that originate from one organ or tissue type and spread to other parts of the body.

Depending on the cancer stage and liver function, treatments for liver cancer include surgery, trans-arterial chemoembolization (TACE), radiofrequency ablation, and radiation therapy.<sup>5</sup> As a non-invasive treatment method, radiotherapy can be used for both curative and palliative treatments.<sup>6</sup> For curative care, it can be combined with surgery (often post-operative) or chemotherapy for early stage cancer or locally advanced metastases. For palliative treatments, a shorter treatment period is often preferred to relieve pain and prevent symptoms from getting worse. Radiotherapy is often used to treat liver cancer when the tumour cannot be removed with surgery, is too large for ablation or when the metastatic spread is at an advanced stage. The goal

of radiation therapy is to sterilize tumour cells while minimizing the damage to surrounding healthy tissues to reduce the chance of complications.

### **1.1.2. Radiobiological Mechanism**

**DNA Damage.** In external beam radiation therapy, high energy particles (such as photons or charged particles) are targeted at tumours, and some of the energy that the particles carry is deposited in other tissues along the beam path. A series of radiation chemistry processes result in the production of ionized atoms and molecules, along with solvated electrons and radicals which are set into motion.<sup>7</sup> Sufficient DNA damage will lead to cell death. Both the charged particles and radicals can, directly and indirectly, damage the DNA in the cell and lead to cell death, which is the principal target for biological effects.<sup>7</sup> Absorbed dose, a non-stochastic quantity, is defined as the energy deposited in the matter per unit mass (unit: J/kg = Gy).<sup>8</sup> For deep seated tumours, the beams pass through healthy tissues before reaching the tumour, and the damage to healthy tissues is a concern. It is sometimes necessary to reduce the prescribed dose and, therefore, compromise the tumour control probability (TCP) to reduce the chance of complications resulting from damage to healthy tissues.

**Side Effects.** Side effects that result from radiation therapy can be separated broadly into early and late occurring effects. Common early effects include skin burns, hair loss, and nausea that occur during or shortly after treatments. Due to the low cell renewal rate for some normal tissue types, late effects, such as demyelination of nerve cells, fibrosis, stenosis, and ulceration, can be very destructive and appear months or years after the start of treatment.<sup>9</sup>

**Hypofractionation.** When tumour cells undergo cell division, they are more radiosensitive in the late G2 and mitosis phases, compared to the late synthesis phase of the cell cycle.<sup>7</sup> The radiation damage can be repaired through biological mechanisms. Usually, a treatment course is fractionated,

which means the prescription dose of a treatment course is delivered over multiple fractions. The response of tissues to conventionally fractionated radiotherapy (repeated daily radiation doses of approximately 2 Gy per fraction) is largely determined by the five “R’s”: repair of DNA damage, repopulation of surviving cells, redistribution of proliferative cells to preferably stay in a radiosensitive phase of a cell cycle, reoxygenation of hypoxic tumour cells to be more radiosensitive, and radiosensitivity of normal tissues during treatments.<sup>10</sup> If a treatment course is split into many fractions, both the tumour and healthy cells have a higher chance of surviving. However, tumour cells are more sensitive to radiation than healthy tissues. This leads to a trade-off between repairing healthy tissues and effective tumour control. Hypofractionation accelerates the dose delivery using a small number of high-dose fractions. Advances in external beam radiation therapy technology have allowed a larger dose per fraction to be delivered safely to tumours. This helps achieve better tumour control and provides convenience for patients and caregivers.<sup>7,9</sup> Hypofractionation is also used in palliative treatments to relieve pain and reduce cancer-induced symptoms. The entire treatment course is relatively short (at most 2 to 3 weeks), and the cancer recurrence or the late effects would occur later than the statistically expected survival.<sup>6</sup> However, one should not neglect the late effects for the sake of improved local control.<sup>7,9</sup> Contemporary hypofractionated treatments provide more conformal dose delivery and steep dose gradients around the target, which spares high dose to organs at risk. However, there is an increased risk for hypofractionated treatments that the delivered dose accidentally misses the target due to patient motion and setup errors. Guidelines on dose constraints and appropriate margins (discussed later in Section 1.1.4) for a target ensure the safety of these treatments.<sup>9</sup>

### 1.1.3. Physics of Photon Beam Radiation Therapy

Radiation can be classified into two major categories: non-ionizing radiation and ionizing radiation, based on the ability of particles to ionize matter. For neutral particles like photons, ionizing radiation can indirectly ionize matter by creating charged particles that deposit energy in the medium along their trajectory of charged particles.<sup>11</sup> This is achieved by Coulomb interactions between secondary charged particles and orbital electrons of the medium.

**Linear Accelerator.** For radiation therapy, photon beams can be generated using a medical linear accelerator (linac). The linac accelerates electrons in a waveguide, using electromagnetic radiofrequency waves. The accelerating potentials are typically from 4 to 25 MV. High energy electrons collide with high  $Z$  materials, such as tungsten, and produce high energy photons. Collimators are used to restrict photon trajectories to produce a desired beam shape.<sup>12</sup> The angular distribution of bremsstrahlung x rays generated in an MV linac is highly forward peaked in the incident electron beam direction. Some conventional linacs are equipped with a specialized flattening filter to generate a more uniform radial dose profile at depth. Conventional and modern linacs using the flattening filter free (FFF) fields are commercially available and used in the clinic to achieve a higher dose rate and a shorter treatment.<sup>13</sup> The CyberKnife system, discussed in Section 1.3.1, is a modern linac that employs FFF beams using circular collimators (range of diameters: 5 mm to 60 mm).

**Interactions with Matter.** Photons interact with the medium in several modes. The relative probability of an interaction mode depends on the photon energy and the material they interact with, and the interaction mode affects how much energy gets deposited in the medium. For the photoelectric effect and Compton scattering, electrons are set in motion which deposit energy along their trajectories. For megavoltage (MV) beams, Compton scattering plays the most

important role compared to other modes of photon interactions. In contrast, for kilovoltage (kV) beams, the photoelectric effect is more prominent for high atomic number ( $Z$ ) materials such as metal, which means photons are more easily attenuated in high  $Z$  materials.

**Small Field Dosimetry.** For photon beams, the dose rises rapidly from the surface as the depth increases until it reaches the maximum dose at a specific depth and then it decreases exponentially (dominated by photon attenuation). The depth of the buildup region is proportional to beam energy because it originates from the range of secondary electrons (about 1.5 cm and 2.5 cm, respectively, for 6 MV and 10 MV beams). Beyond the maximum dose, the transient charged particle equilibrium is achieved, which means the collision kerma (an acronym for “kinetic energy released per unit mass”) is proportional to the absorbed dose, which is a measurable quantity.<sup>14</sup> The collision kerma, a calculable quantity, is the energy transferred from photons to charged particles per unit mass in collision interactions (such as photoelectric, Compton scattering, and pair production), and it contributes to the local absorbed dose.<sup>15</sup>

The accuracy of dose measurements for small fields (typically less than  $3 \times 3 \text{ cm}^2$ ) is subject to the type of dosimeters used.<sup>16</sup> If the field size is comparable with the size of a detector, the presence of a detector made of heterogeneous materials cannot be treated as nonperturbing. The measured dose profile can be blurred with the maximum dose underestimated because of the volume averaging effect.<sup>17,18</sup> For small field dosimetry, the advantages and drawbacks of different types of radiation dosimeters have been summarized by Wilcox *et al.*<sup>19</sup> They investigated the reproducibility, sensitivity, dose linear response, energy dependence, spatial resolution, water equivalence and ease of use for ionization chambers, thermoluminescent detectors, semiconductors, radiographic films, and radiochromic films. Except for the films and plastic scintillation detectors, correction factors are needed for the measured dose to account for the dose

perturbation due to the presence of the detector made of inhomogeneous materials and the lack of charged particle equilibrium. As stated in the American Association of Physicists in Medicine (AAPM) Task Group 55 Report,<sup>20</sup> radiochromic films can be a reliable radiation dosimeter that provides two-dimensional dose measurement with a high spatial resolution for small beam fields. They are made of tissue-equivalent materials, consisting of a single or double layer(s) of radiosensitive organic microcrystal monomers coated with a transparent polyester base. Unlike silver halide radiographic films, radiochromic films do not require chemical processing and they are also much less sensitive to room light.

#### **1.1.4. Definition of Volumes**

Figure 1-1 shows a graphical representation of a CyberKnife (discussed later in Section 1.3.1) radiation therapy treatment plan. The image shows a grayscale 3D surface rendering of a patient based on the cross-sectional anatomy determined using a computed tomography (CT) image set. Very briefly, a CT scan is a tomographic image volume reconstructed from a series of transmission x-ray projection images taken from different angles around the patient body. Also shown are blue cylinders which are a representation of circular collimated beams and their directions of incidence that have been selected by the treatment plan optimization algorithm. A treatment plan is generated by optimizing the delivered dose weightings on each beam (indicated by the length of each cylinder in the figure) using different planning objective functions which ensure the dose constraints for treatment targets and organs at risk are met following published or institutional treatment planning guidelines (see Section 1.3.3 for a more detailed discussion of treatment planning).

During treatment planning, margins can be added around targets and organs to account for different types of delivery errors. The International Commission on Radiation Units and

Measurements (ICRU) Report 62<sup>21</sup> defines the volumes commonly used in treatment planning, as discussed in the following two subsections.

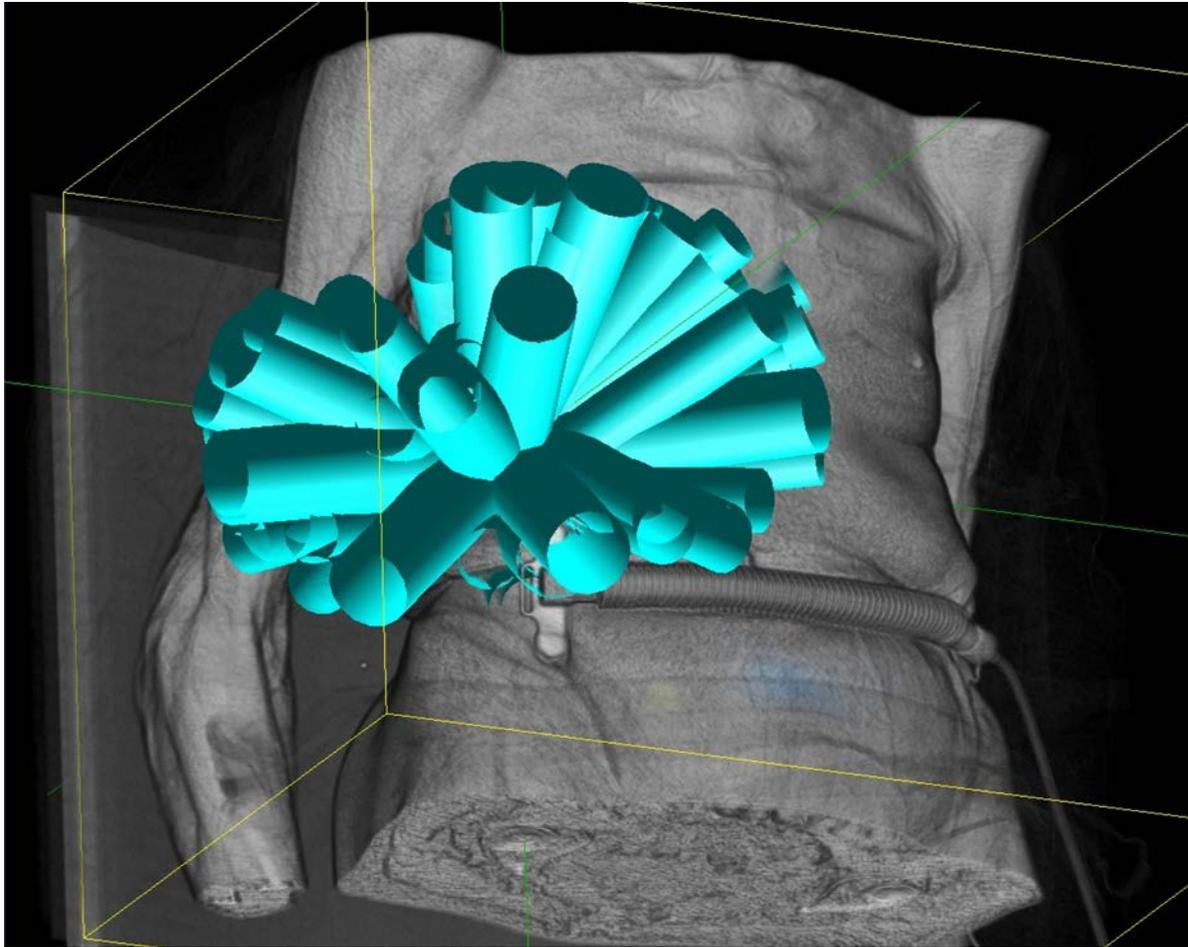


Figure 1-1. A grayscale 3D surface rendering of a patient computed tomography scan image set. The blue cylinders indicate simulated collimated beams in different directions. The screenshot was taken from the MultiPlan® treatment planning system.

**Definitions of Volumes (Targets).** The gross tumour volume (GTV) is the palpable or visible macroscopic extent of a tumour. The GTV is typically defined using CT scan and/or magnetic resonance (MR) scans. Functional imaging, including functional magnetic resonance imaging (fMRI)<sup>22</sup> and positron emission tomography (PET),<sup>23,24</sup> can also be used to help delineate the GTV. As illustrated in Figure 1-2, the clinical target volume (CTV) is often an expansion of the GTV to account for the presumed microscopic extension of the tumour taking into account

anatomical barriers to tumour spread. To consider target motion and deformation, an internal margin may be added to the CTV, to form an internal target volume (ITV). The four-dimensional CT (4DCT) image set (discussed further in Section 1.2) consists of CT images acquired over time and binned for different respiratory phases, and it can be used for generating ITV.<sup>25</sup>

For tumours that move with respiration, an ITV can be generated by adding an internal margin to the union of the CTV at different respiratory phases to encompass the target motion. To account for patient setup uncertainties, another margin, called the setup margin, is added to the ITV to form a planning target volume (PTV). The delineation of an ITV is optional in defining the PTV when the uncertainties of target motion dominate the setup uncertainty and/or when the uncertainties of the GTV location is independent of the setup uncertainty. The setup margin accounts for patient positioning and beam alignment uncertainties.<sup>26</sup>

According to ICRU Report 83,<sup>26</sup> the PTV is a geometrical concept meant to ensure that the prescribed dose is delivered within clinically acceptable probability levels to all parts of the CTV, despite geometrical uncertainties. In principle, uncertainty due to organ motion, setup variations and even the uncertainty in the delineation of CTV or GTV should be included in the PTV.<sup>27</sup> The dose prescribed to the PTV by the radiation oncologists should be sufficient to ensure adequate tumour control while controlling the risk of complications for OARs. The minimum dose to 99% of the volume received by the PTV (referred to as the near-minimum dose and denoted as “PTV D99”) was introduced in ICRU Report 83 as an important metric to compare different dose distributions and to correlate them to tumour control probabilities. Unlike the CTV and GTV where the extent of the volume is related to the tumour volume itself, the size of the PTV-GTV or PTV-CTV margin is a geometrical concept that may vary from one cancer center to another, depending on the technology used, the disease site, the immobilization and other patient setup strategies used.

This margin takes into account the uncertainties discussed above for the specific patient population under consideration.<sup>26</sup>

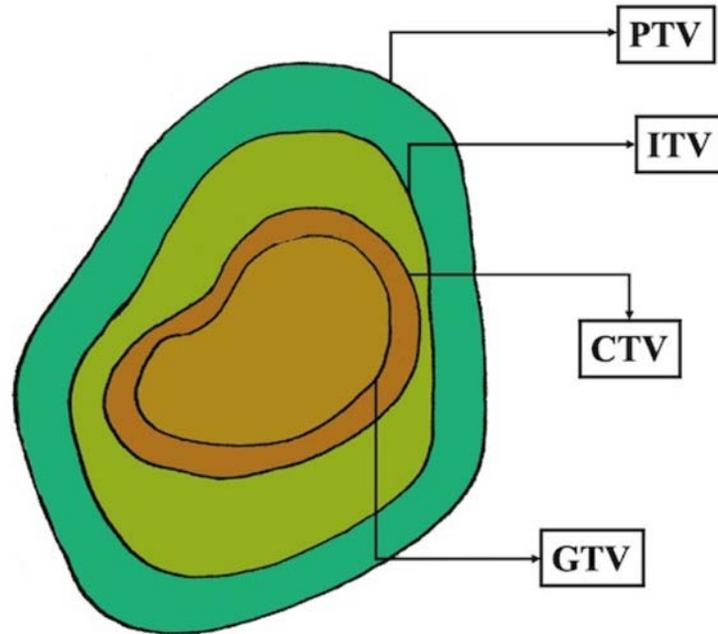


Figure 1-2. Two-dimensional illustration of target volumes commonly used in treatment planning for upper abdominal tumours: gross tumour volume (GTV), clinical target volume (CTV), internal target volume (ITV), and planning target volume (PTV). The dose is prescribed to the PTV for tumour control.

**Definitions of Volumes (Organs at Risk).** The dose to the healthy tissues should be as low as possible to minimize toxicity. Organs at risk (OARs) and critical normal tissues, usually close to the tumour, can be radiosensitive and the patient might suffer from severe morbidity when the dose exceeds a certain level.<sup>7,26</sup> OARs are either classified as serial or parallel depending on whether the functional units within the organ are independent of each other. For example, esophagus and small intestine are serial structures, and liver and kidney are parallel structures. To compensate for the motion of OARs, analogous to a PTV-GTV margin, a margin is added to OARs to generate the planning risk volume (PRV). A schematic diagram showing a possible arrangement

of PRV, OAR, PTV, and GTV structures is shown in Figure 1-3. The tolerance doses for OARs are usually much lower than the prescription dose typically used for ablative liver treatments, and the maximum dose for serial structures (digestive organs, spinal cord, etc.) is especially important. When serial OARs are close to the tumour, the possible complications may be too severe or even life threatening, in which case, we often lower the prescription dose or compromise the dose to the PTV to meet the dose constraints for OARs.

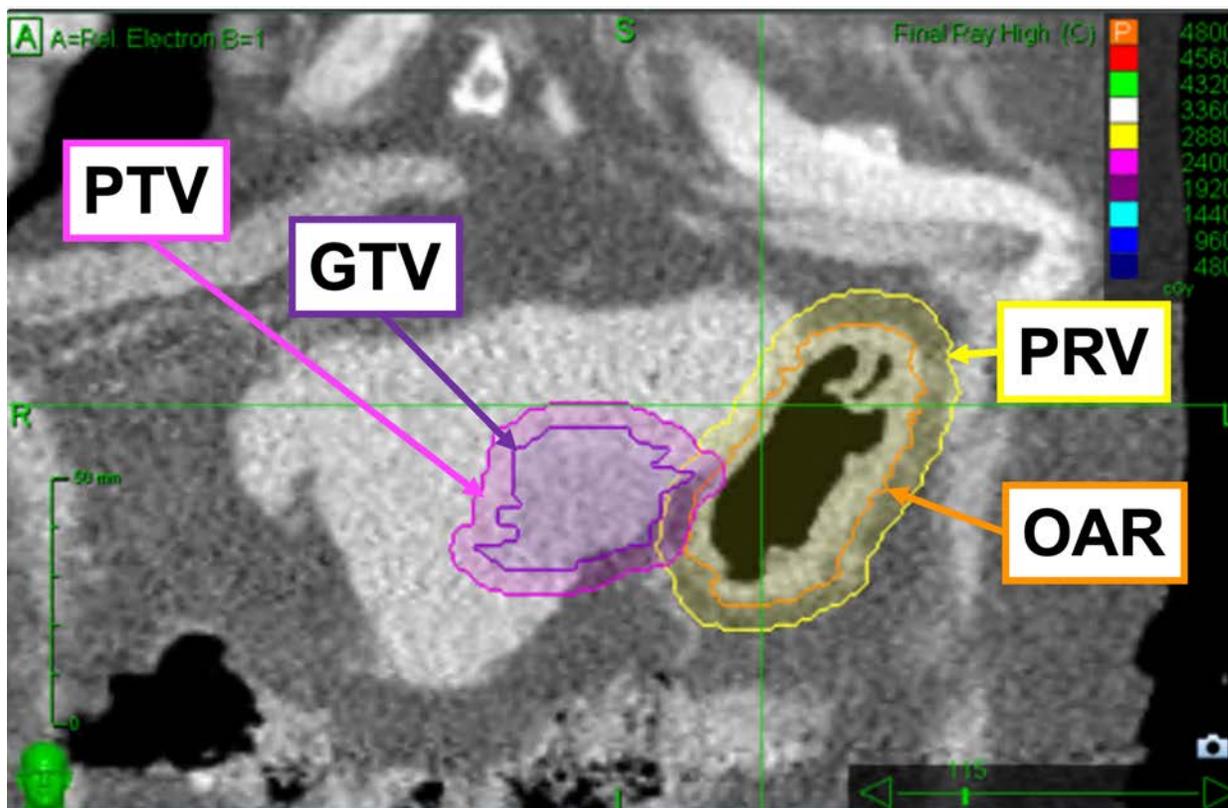


Figure 1-3. A CT slice for a liver lesion with contours of volumes commonly used for treatment planning: gross tumour volume (GTV), planning target volume (PTV), planning risk volume (PRV), and organs at risk (OAR). In this example, there is an area of overlap between the PTV and PRV. It may be necessary to either compromise the dose to the PTV or accept a higher risk of complication.

### 1.1.5. Patient Specific Quality Assurance and Dose Evaluation Metric

When commissioning a treatment planning system or performing patient-specific quality assurance, the planned dose should be verified experimentally to ensure the differences between

the planned and measured dose are clinically acceptable. Since human tissue is approximately water equivalent, the dose to tissue can be estimated by dose measurements using dosimeters in water or more conveniently in solid material, called “phantoms” (discussed later in Section 4.2.1). Phantoms are designed to have similar radiological properties as water. We can directly (calorimeter) and indirectly (ion chambers, thermoluminescent dosimeters, radiochromic film, etc.) measure the absorbed dose in these materials. In this study, we use radiochromic film to sample 2D dose distributions in near water equivalent plastic phantoms (discussed in detail in Section 4.2.1).

For quality assurance, the dose distribution calculated using the planning system is often compared with the measured dose using an evaluation metric called the gamma index. This metric combines a specific dose difference criterion ( $\Delta D$ ) and distance-to-agreement criterion ( $\Delta d$ ) in the same metric, which should correspond to the clinical level of difference between the reference and evaluated dose. For example, for many applications  $\Delta D = 2\%$  of the maximum dose and  $\Delta d = 2$  mm are considered clinically acceptable criteria for patient-specific quality assurance for contemporary linac treatments. The gamma index ( $\gamma$ ) at a given reference point ( $\vec{r}_r$ ) is given by the equation below,

$$\gamma_{\Delta D/\Delta d}(\vec{r}_r) = \min \left\{ \sqrt{\frac{(D_e(\vec{r}_e) - D_r(\vec{r}_r))^2}{(\Delta D)^2} + \frac{\|\vec{r}_e - \vec{r}_r\|^2}{(\Delta d)^2}} \right\} \forall \{\vec{r}_e\},$$

where  $D_e(\vec{r}_e)$  and  $D_r(\vec{r}_r)$  are the evaluated (usually as calculated by the planning system) and reference (usually the measured data) dose, respectively, at the reference point. For each reference point, we progressively calculate the gamma index for all pixels ( $\vec{r}_e$ ) of the evaluated dose distribution within a certain search radius (larger than  $\Delta d$ ) of the same reference location ( $\vec{r}_r$ ). At any reference position for which  $\gamma \leq 1$ , the evaluated dose is considered acceptable. The gamma

pass rate describes the percentage of pixels in the measured dose distribution that meet the gamma criteria ( $\gamma < 1$ ). In this thesis work, we use in-house MATLAB scripts based on the fast implementation of the gamma analysis as described by Wendling *et al.*<sup>28</sup> to complete the 2D and 3D gamma analyses.

In the AAPM TG 135 report<sup>29</sup> on quality assurance for robotic radiosurgery, it is recommended that for treatments using the Synchrony® tracking system (introduced in Section 1.3.2), the gamma pass rate using gamma criteria of 3%/3mm should be above 90% for a volume encompassed by 50% of the maximum dose. In the Canadian Partnership for Quality Radiotherapy (CPQR) Technical Quality Control Guidelines for CyberKnife, an annual test of the Synchrony motion tracking system with the internal fiducial and external LED marker motions out of phase for at least 20° is suggested.<sup>30</sup> An analysis of the penumbral spread is recommended but the tolerance levels are left at the discretion of the user. Tolerance and action levels of 2 and 3 mm change in penumbral widths, respectively, are proposed for comparing the static and out-of-phase phantom motion, which is tracked using the Synchrony system.<sup>30</sup> See Section 4.2.5 for calculating the penumbral width of a dose profile.

In this thesis work, the dose difference criterion is chosen to be a small fraction (usually 1% or 2%) of the maximum reference dose. In Chapter 4, we use relatively strict gamma criteria (2%/2mm or 1%/1mm) to compare the measured dose with the planned dose which is perturbed to account for delivery errors. We also look specifically at penumbral widths in the direction of phantom motion. The distance-to-agreement ( $\Delta d$ ) should not be smaller than the minimum dose resolution, which is usually 1 mm for our phantom studies.

## 1.2. Target Motion Management

Delivering a high dose accurately to tumours during treatments while minimizing the toxicity for surrounding healthy tissues can be especially challenging for targets that move due to respiration, such as tumours in the lung, liver, pancreas, and kidney.<sup>31,32</sup> Many studies have quantified the magnitudes of liver tumour motion using kV images,<sup>33</sup> cone beam computed tomography (CBCT),<sup>34-36</sup> 4DCT,<sup>37</sup> and MR image data.<sup>38,39</sup> In Liang *et al.*<sup>33</sup>'s study of 90 fractions of treatment for 14 liver patients, the median motion range estimated using kV image data recorded in treatment log files was 11.9, 1.3, and 3.8 mm in the superior-inferior, left-right, and anterior-posterior directions, respectively.

Brandner *et al.*<sup>40</sup> reviewed the imaging modalities used for respiratory motion monitoring for thoracic and upper abdominal tumours and compared the advantages and disadvantages of different imaging systems. A more recent review by Bertholet *et al.*<sup>41</sup> discussed technologies used for monitoring intra-fraction motion during external beam radiation treatment. For intra-fraction motion, the tumour can be tracked using either ionizing radiation, such as kV x-rays (fluoroscopy)<sup>42,43</sup> and MV images (beam's eye view),<sup>44,45</sup> or non-ionizing radiation (or technology), including MR images,<sup>46-48</sup> ultrasound,<sup>49-52</sup> high definition (HD) cameras (e.g. surface tracking for chest wall and breast irradiation),<sup>53,54</sup> implanted radiofrequency transponders,<sup>55</sup> and an electromagnetic positioning tracker (placed on the patient surface or inside the body cavity) that can also simultaneously measure the dose.<sup>56</sup>

KV images compared to MV images provide better image contrast for soft tissues. However, for liver lesions the contrast for soft tissues and tumours is still not sufficient without intravenous contrast (not routinely used during treatment). Therefore, fiducials are often used as a surrogate of tumour(s) for tracking.<sup>57</sup> Fiducials are small metal (platinum or gold) seeds that are easily

distinguished from soft tissue in kV x-ray images, as shown in Figure 1-4. They are biologically inert and are permanently implanted near the tumour as a surrogate of the tumour position. A linac or radioisotope therapy unit ( $^{60}\text{Co}$ ) can be equipped with an on-board kV imaging system for image-guided radiation therapy (IGRT). In this way, the target motion can be monitored and taken into account if possible. Refer to a review by Dawood *et al.*<sup>5</sup> for clinical IGRT modalities to treat liver metastasis. The imaging dose from kV x-ray images introduces an extra dose, and the associated risk of secondary cancer should be maintained as low as possible. For fiducial tracking, we can also employ an electronic portal imaging device (EPID) during treatment, so that the MV treatment beam could be used to localize the fiducials continuously during the respiration cycle.<sup>45</sup>

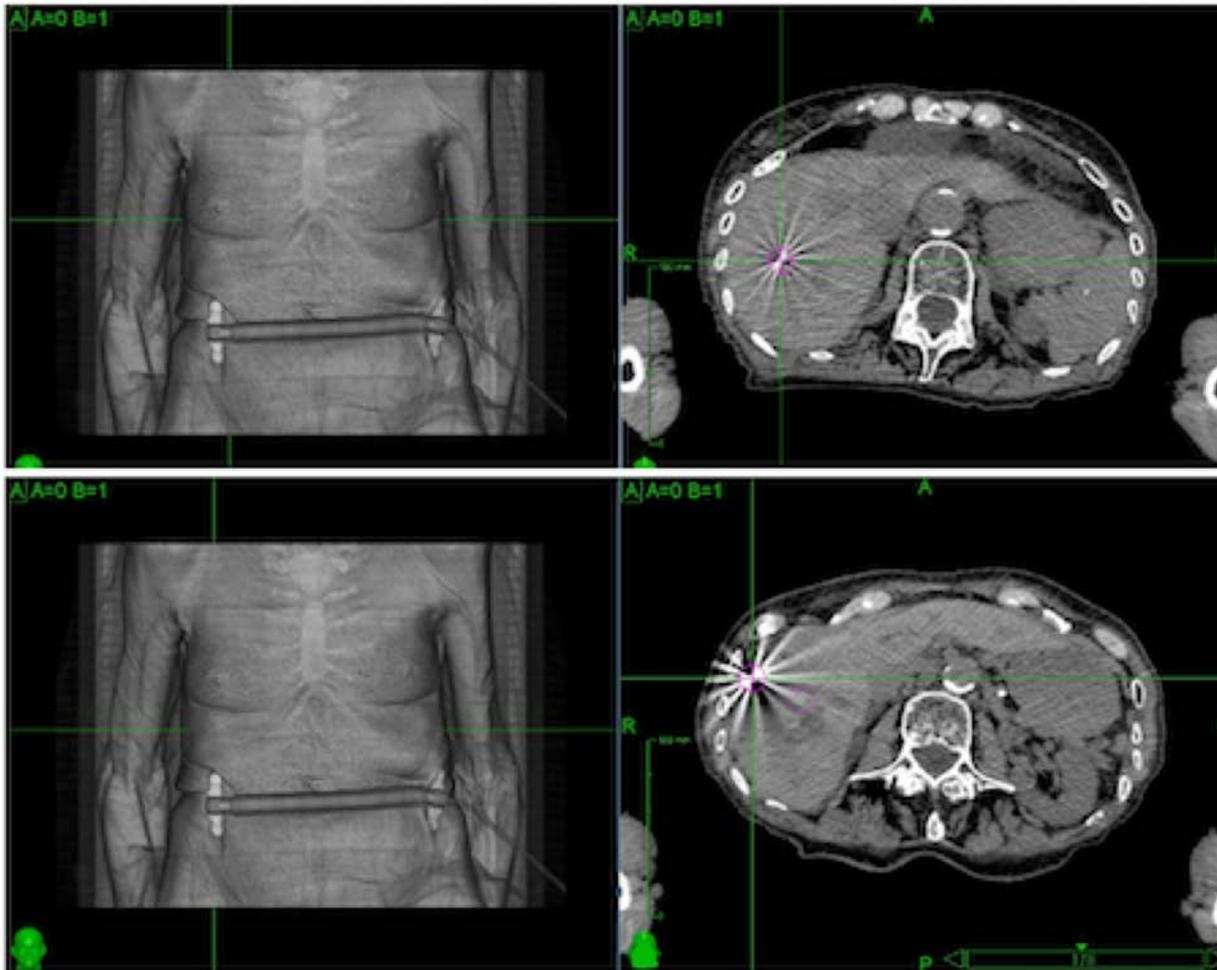


Figure 1-4. Two fiducials (indicated by the crosshair) are identified in CT images for a patient. The plots show the surface renderings of a patient anatomy.

Respiration introduces different types of artifacts in different imaging modalities which can lead to target and OAR delineation errors during treatment planning.<sup>27,58</sup> The errors due to target motion can often be thought of as a displacement of the target relative to the planned dose distribution. These errors may be random or systematic in nature. Most random errors are well approximated as a dose blurring due to the intra-fractional target motion. In some cases, systematic errors can be approximated as a shift of the target with respect to the planned dose distribution which persists for the entire treatment. Systematic errors can have a much more severe impact on target coverage than random errors and should be minimized with accurate target localization and

frequent monitoring during treatment.<sup>27</sup> Several commonly used strategies for motion management are described in the following sections, including immobilization, breath-hold, gating, and real-time tracking.

**Immobilization.** Immobilization is used from treatment planning step to daily treatment delivery. Figure 1-5 illustrates some commonly used immobilization tools for different sites and body regions used in the clinic. Some of them (subfigures b, d, and e) are customized to an individual patient's anatomy to provide patient-specific immobilization over their entire treatment course. As shown in Figure 1-5 (e), we can use an abdominal compression plate which is intended to force a patient to breathe shallowly, and thereby reduce the amplitude of breathing motion. Eccles *et al.*<sup>38</sup> studied the impacts of abdominal compression on 60 patients treated for liver lesions using magnetic resonance imaging data. They found that target motion reduced by more than 3 mm in the superior-inferior direction for 40% of the patient cohort. In comparison, less than 2% of patients had an increased motion amplitude in the same direction. For hypofractionated liver treatments on CyberKnife (discussed later in Section 1.3.1), the patient breathes freely as the system actively compensates for tumour motion. A reproducible setup is very important to reduce systematic setup uncertainties for this precisely delivered treatment. For these treatments, a Vac-Lok™ (Civco® Medical Solutions, Orange City, IA), shown in Figure 1-5 (e), is used for CT simulation and treatment at our center. This system creates a customized mold of the patient's body by evacuating a low-density bead filled cushion while the patient lies in it, hence "locking" the patient into a reproducible treatment position. Similar to a Vac-Lok, the BodyFix® (Elekta, Stockholm, Sweden) vacuum posterior cushion is also used in the clinic, as shown in Figure 1-5 (f). This system employs an anterior plastic cover sheet to stabilize a patient and an abdominal compression plate to facilitate shallower breathing.

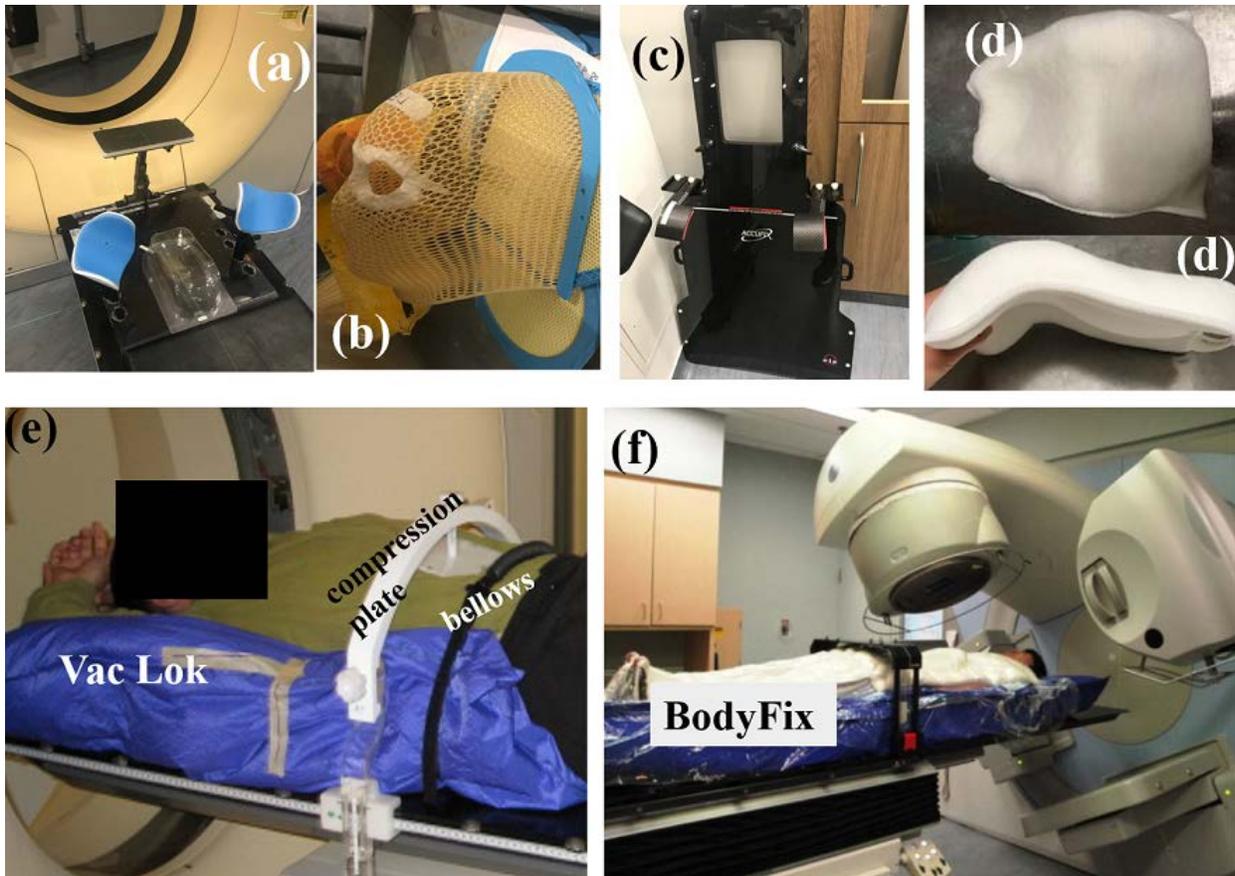


Figure 1-5. Immobilization devices at the Ottawa Hospital Cancer Centre. (a) The lung board used for lung and regular abdomen treatments. (b) The brain mask to restrict the head motion. (c) AccuFix board for shoulder compression used in head-and-neck treatments. (d) A customized Klarity™ cushion for the neck. (e) Vac-Lok posterior customized cushions (blue); an in-house abdominal compression plate (white) for shallow breathing; and the air bellows (black) for determining the respiratory phase during the CT simulation. (f) The BodyFix® patient positioning posterior cushion (in blue) and anterior plastic cover sheet that fit a patient's body contour, assisted by a dual vacuum pump. Figure obtained with permission from Tseng *et al.*<sup>59</sup>

**Breath Hold.** Breath-hold techniques including deep-inspiration breath-hold (commonly used for treating breast tumours),<sup>60</sup> active-breathing control (ABC),<sup>44,60-62</sup> and other methods require patients to voluntarily hold their breath close to the end of inspiration when the tumour and organs at risk are more spatially separated, or at a predefined phase of the respiratory cycle. The ABC device (Elekta Oncology Systems Ltd, Crawley, UK) relies on monitoring the patient breathing (volume and flow rate) and determines the respiratory phase using a mechanical spirometer.

Patients, with their nose clamped, are trained to hold their breath at predefined respiratory phases during the beam delivery.<sup>61</sup> This allows for a smaller margin to be added to the clinical target volume (CTV) to account for the target motion and the setup uncertainties. Dawson *et al.*<sup>44</sup> studied the impacts of implementing ABC by monitoring the air-diaphragm interface using megavoltage imaging for twenty liver stereotactic body radiation treatments (refer to Section 1.3 for stereotactic body radiation treatments). The absolute diaphragm motion, which was correlated with the target motion, was reported to be 1.9 mm on average based on the analyses of MV images acquired during treatments for 52 fractions. However, 13 out of 34 patients were found to be intolerant of holding their breath for more than ten seconds, and five had poor end-exhale breath hold reproducibility (i.e., target motion >2 mm) determined using fluoroscopy.<sup>44</sup> Since active breathing control, breath hold, and abdominal compression can introduce discomfort and may not be possible for all patients, methods of respiratory compensation with the patient breathing normally during treatment is a desirable alternative approach.

**Respiratory Gating Using External Markers.** For respiratory gating, the breathing cycle is monitored continuously during treatment so that beams can be delivered at a specific period (gating window) of the respiratory cycle.<sup>63</sup> The difference between respiratory gating and active breathing control in determining the respiratory phase (or amplitude) is that the former is achieved by direct imaging, while the latter relies on a spirometer.<sup>6</sup> The gating window can be based on either phase or amplitude. Some studies showed that amplitude-based gating gave a more consistent estimation that the tumour position was within the gating window for dose delivery.<sup>64,65</sup> To be able to use gating for treatment, we need to create a treatment plan based on a CT volume for the “gated” position (a specific phase or range of breathing motion displacement), at which the beams are meant to be delivered. This is acquired using 4D CT, which provides information on three-

dimensional target motion as a function of time (see Section 1.1.4 for 4D CT).<sup>25</sup> During the acquisition of 4D CT, the respiratory phase can be determined using an air bellows placed over the patient abdomen, as shown in Figure 1-5e. The air bellows contracts and expands due to the abdomen motion, which is converted into an electronic signal that indirectly indicates the respiratory phase.<sup>6</sup> Gating does not track the tumour motion directly. One approach commonly used for respiratory motion detection during the treatment of thoracic and upper abdominal lesions is the Real-time Position Management (RPM) System (Varian® Medical Systems, CA).<sup>66,67</sup> It is a displacement-based gating system that uses an infrared tracking camera to monitor the position of a plastic infrared marker placed on the patient's surface. The gating window selected involves a trade-off between the efficiency of beam delivery, the residual motion of the target, and the effectiveness of dose sparing to healthy tissues.<sup>27</sup> For treatments using the RPM gating system, beams are turned on only while the marker motion is within the gate. The in-room imaging system (usually three-dimensional kV volumetric imaging) is also used to confirm the target location and anatomy of the patient. The treatment beam is interrupted when the system detects irregular breathing or a large phase shift between the target and the marker. Since the gating system relies on monitoring an external marker to determine if the target is within the gating window, the treatment does not require an additional invasive procedure (such as implanting fiducials or transponders).

**Respiratory Gating Using Internal Markers.** Respiratory gating can also be achieved using internal fiducial markers for *direct* motion tracking, along with a real-time tumour tracking system (Mitsubishi Electronics Co., Ltd., Tokyo).<sup>42,68,69</sup> In Shirato *et al.*'s study,<sup>42</sup> thirty patients with tumours in the liver, lung, and prostate were carefully selected for implanting fiducials considering the tumour size and the patient health condition. The implantation of fiducials was an image guided

procedure using ultrasound, CT, or endoscopy, depending on the body site. Before treatments, the target motion was assessed using the 4DCT image set (breath hold, exhalation, inhalation, and intermediate phases). A 5 mm CTV-GTV margin was used to encompass the microscopic tumour extension. To account for other uncertainties, they employed a PTV-CTV margin that ranged from 1 mm to 10 mm, depending on the proximity to nearby OARs, the motion pattern (determined using the CT image set), and the setup uncertainty. During treatments, the system relied on fluoroscopy (kV x-ray images acquired 30 times per second) to monitor the target motion. To verify if fiducials were identified correctly during treatment, digitally reconstructed radiographs (DRRs) were created. DRR is a computer-generated two-dimensional projection of the planning CT volume, which is commonly used in radiotherapy to emulate planar x-ray image acquisition. The DRRs (with fiducials identified) were compared with fluoroscopic images for validating the fiducial locations. The beams were turned on only if the fiducials were within a pre-determined volume.<sup>42</sup> Issues of fiducial migration and the relative motion between fiducials and tumours made it challenging to determine the uncertainties of treatments.<sup>42</sup> When the image quality is not optimal for tracking or the patient's health condition does not tolerate fiducial implants, gating with fiducials may not be ideal for motion management.

**Tracking Based on Correlation Models.** Two commercial respiratory tracking systems using a kV imaging system, which rely on internal/external breathing motion correlation models instead of gating for respiratory motion compensation, are discussed as follows. They are the “robot-based” CyberKnife system (Accuray Inc., CA, USA) and “gimbal-based” Vero system (Brainlab AG, Feldkirchen, Germany).<sup>70</sup> For motion compensation, both systems require the real-time tracking of external markers to predict the internal tumour location (indicated by fiducial markers) using correlation models.

The CyberKnife system is a compact linac mounted on a robotic arm. The Synchrony® system (discussed in more detail in Section 1.3.2) is a real-time respiratory tracking and compensation system. It generates a correlation model between the internally implanted fiducials and external LED markers placed on the patient's abdomen or chest. The created model instantaneously predicts the tumour position. The system evaluates the discrepancy between the target and the predicted positions. If the difference is greater than 5 mm, the system interrupts the delivery, and a new model must be generated to continue the treatment.<sup>71</sup> Usually, it takes about one to two hours for treatments using the Synchrony system because the beam is turned on for 20% to 30% of the time. For the remaining (beam-off) time, it is used for repositioning the linac using the robotic arm and acquiring x-rays to update the respiratory correlation model.<sup>71</sup>

The CyberKnife also provides a tracking system called Xsight® Lung, which tracks lung tumours *without* using fiducials. It works in a similar manner to the Synchrony system but uses the image contrast between the solid tumour and lower density lung tissue to *directly* track the “visible” portion of the tumour (>15 mm in diameter) using kV images.<sup>72</sup> The skeletal structure of the spine is used for patient alignment prior to the start of treatment. Tumours that are too small or are at a location where bony anatomy interferes with the contrast between lung and tumour are not candidates to be treated using Xsight Lung.<sup>72,73</sup>

The Vero system also creates a model to correlate the internal target location as determined by orthogonal kV x-ray images but uses infrared markers to track external breathing motion. The linac and MLC are mounted together on orthogonal gimbals to enable panning and tilting of the therapeutic beam which is used for real-time MLC tracking of moving tumours.<sup>74</sup> The system can have dose delivery from any angle about a patient's superior-inferior axis because the gantry can rotate 360 degrees about the treatment couch and the ring itself  $\pm 60^\circ$  about the vertical axis, which

allows for non-coplanar beam delivery. In contrast, CyberKnife mostly delivers beams from nodes above the couch level. The Vero imaging system is equipped with collimators for a smaller x-ray field to reduce the imaging dose, and it rotates with the gantry to spare imaging dose to healthy tissues, whereas the x-ray beams of the CyberKnife imaging system are “static”.<sup>75</sup> Moreover, Vero’s system latency is 48 ms,<sup>74</sup> which can be taken into account by its prediction algorithm to forecast target motion. In contrast, there is a 115 ms latency<sup>76</sup> for CyberKnife but the Synchrony system employs a predictive model that corrects for this effect (see section 1.3.2). Unfortunately, the Vero system has been discontinued by the manufacturer.<sup>77</sup>

There are some key differences in the details of the correlation models used by the two systems. The correlation model of CyberKnife only keeps up to fifteen of the most recent data points acquired every 30 to 60 s, while a higher frequency (0.2 to 1 Hz) of x-ray acquisitions and more data points at a lower x-ray tube voltage are possible with the Vero motion tracking system.<sup>75</sup> A CyberKnife correlation model predicts the target location using the marker position only.<sup>78</sup> In contrast, a Vero correlation model is represented by a quadratic function in terms of the marker’s position and speed.<sup>75,78</sup> Poels *et al.*<sup>78</sup> compared the modelling algorithms for the correlation models in CyberKnife and Vero systems. The correlation models for both systems were “trained” using fluoroscopy images (11 Hz). No significant difference was found in determining the geometrical tracking accuracy, as measured by the 95<sup>th</sup> percentile of tracking errors and the probability of obtaining tracking errors larger than 5 mm.

A non-commercial real-time tumour monitoring system, the Combined Optical and Sparse Monoscopic Imaging with Kilovoltage x-rays (COSMIK), was implemented on a conventional gantry-based linac.<sup>79</sup> It also used a correlation model that accounted for the baseline drift between internal fiducial and external marker motions. The model was built using CBCT data prior to

treatment with an in-house algorithm for the auto-segmentation of fiducials.<sup>80</sup> During treatment, the model used the last three x-ray images (sampled at 0.33 Hz) to correct for the overall baseline shift of fiducials, but did not include this data to adapt the model during the treatment. The latency in the system was overcome by predicting the fiducial location using the marker location acquired at 20 Hz. The authors validated the system's accuracy of estimating the tumour location in phantom experiments using recorded liver motion data and external marker motion data for 15 patients (without abdominal compression) during stereotactic body radiation therapy (introduced later in Section 1.3). The root-mean-square error (RMSE) for the difference between the model-estimated and true target locations was 1.49 mm. The system was combined with a dose reconstruction algorithm using in-house scripts to assess the real-time delivered dose considering the target motion.<sup>81,82</sup>

**Dynamic MLC Tracking for Conventional Linac Systems.** For intensity modulated radiation therapy (IMRT), the multi-leaf collimator (MLC) consists of tungsten leaf pairs, which can be used to modulate the incident photon fluence and thereby create an intensity modulated field. For real time MLC tracking, the treatment beams are adapted continuously by the MLC based on the tracked tumour's position, shape, and proximity to OARs, allowing for a real-time adaptive treatment to make the dose delivery safer and more effective. Research on the dynamic MLC tracking system with a target tracked by an infrared camera system (Varian RPM system),<sup>83,84</sup> kV/MV images (Varian Trilogy™ system),<sup>85</sup> and using radiofrequency transponders (Calypso® 4D Localization system, Calypso Medical Technologies, Seattle, WA)<sup>86</sup> has been reported. Falk *et al.*<sup>84</sup> compared the measured and planned dose for two plans, with and without the dynamic MLC tracking, and the target motion was *indirectly* determined using the RPM system. The measurements and dosimetric analyses were performed using a Delta<sup>4</sup>® (ScandiDos AB,

Uppsala, Sweden) dosimetric phantom that was moving in the superior-inferior direction for the motion ranges of 5, 10, 15, 20 and 25 mm. They found the PTV margin could be reduced by about 2 mm and more than 3 mm for motion ranges of 15 mm and 25 mm, respectively, when dynamic MLC tracking was used, because the dose delivery was more conformal to the target. The gamma pass rates for the planned and measured dose distributions using 3%/3mm criteria were higher than 98% for both cases with dynamic MLC tracking and the range of target motion  $\geq 15$  mm. In contrast, they were  $\leq 75\%$  when the tracking was not employed. The system temporal latency (160 ms) for the system to respond to a new leaf position was not accounted for in their study. Cho *et al.*<sup>85</sup> integrated the kV and MV imaging system (images acquired at 6.7 Hz) that directly tracked the fiducial location with dynamic MLC tracking. The system latency from imaging acquisitions to adjusting the MLC positions was about 450 ms due to the image acquisition and fiducial extraction processes, which required correction for accurate delivery. With a prediction algorithm to account for this system latency, the average tracking error was less than 1 mm, and the dosimetric error was negligible (gamma pass rate = 99.8% using 3%/3mm criteria). The failure rate using the same criteria was 22.5% without the intra-fraction motion tracking. Keall *et al.*<sup>86</sup> studied the dosimetric accuracy of dynamic MLC tracking for prostate and lung treatments using implanted radiofrequency transponders of the Calypso 4D localization system. The treatment plans were delivered to a phantom with its motion programmed using measured patient traces, and the dose was measured using a two-dimensional ion chamber array. The gamma pass rates using 3%/3mm criteria were 98% and 99%, respectively, for lung and prostate cases with motion tracking. Without dynamic MLC tracking, they were 66% and 86%, respectively. The system latency for this system was 150 ms and was compensated for using a prediction algorithm.

For IMRT using dynamic MLC, the interplay effects that originate from the motion of the MLC becoming unintentionally synchronized to the tumour motion cannot be neglected, especially when only a few fractions are delivered.<sup>87,88</sup> The dose differences between the planned and delivered dose distributions are related to the motion amplitude, the number of fractions, the dose rate, and the dose inhomogeneity.

**Dynamic MLC Tracking for MR Linac Systems.** In contrast to the imaging systems with ionizing radiation discussed above, commercial MR-Linac systems have been developed that combine a MR scanner and linac to provide real-time direct tracking of tumours. Two commercial entities available for in-room MR-Linac with fast treatment planning and adaptation are the ViewRay® (ViewRay Inc., Oakwood Village, OH) MRIdian® and Elekta (Elekta AB, Stockholm, Sweden) Unity systems. The first ViewRay systems combined MR guided adaptive therapy with three <sup>60</sup>Co treatment sources but the manufacturer has since moved on to an MR-Linac design. Due to the excellent soft tissue contrast in MR images, invasive procedures for implanting fiducial markers are not essential for tracking liver lesions.<sup>89,90</sup> It also makes real time tracking for the target motion (including translation, rotation, and deformation) and organ motion, and the acquisition of functional information possible.<sup>91</sup>

MRI alone cannot provide electron density data required for dose calculation. These data can be approximated using deformable image registration (DIR) with the planning CT, since DIR allows for “tissue tracking” between images using a deformable vector field for accurate dose accumulation.<sup>92</sup> The inter-fractional plan optimization before each treatment has been clinically achieved using the MRIdian system to account for anatomical changes of targets and OARs.<sup>93,94</sup> For inter-fractional adaptation, Raaymakers *et al.*<sup>95</sup> reported that the Elekta system took about 5 min for online replanning to treat spine bone metastases (N = 4). Kontaxis *et al.*<sup>96,97</sup> demonstrated

the proof of principle for the online intra-fraction plan adaptation using simulated SBRT plans by inter-beam replanning, with an average replanning time of 25.6 seconds. The replanning pipeline proposed by their study not only considered the anatomical changes in dose reconstruction but also corrected for the baseline drifts of targets. With the dose delivery status recorded every 40 ms in treatment log files and online MR imaging data, the dosimetric impacts of anatomical motion for targets and OARs can be retrospectively assessed on a daily basis.<sup>98,99</sup> It can be possible in the future to enable real time adaptive replanning if the system is robust for reoptimizing the plan. The DIR dose accumulation, therefore, has to be very efficient (i.e., low latency) with desired automated contouring of the targets and OARs and ensured in vivo verification for patient delivery quality assurance.

MR guided radiation therapy has the benefits of improving tumour control and reducing irradiated volume and toxicity to the normal tissues for abdominal tumour treatments. Fast *et al.*<sup>90</sup> used “tumour trailing” that corrected the baseline shifts of time-averaged target positions treated on an Elekta MR-Linac. For seventeen liver cases, the median dose to 98% of the GTV was found to be increased by 0.5 - 1.9 Gy per fraction for all four scenarios (a linear drift, a single drift during the treatment, a periodic drift, and measured baseline drifts) with different target drifting patterns. For the experiment conducted using a phantom set in periodic motion with a peak-to-peak range of 15 mm in the inferior-superior direction, one of the precalculated plans was selected for adaptive replanning whenever the superior-inferior baseline motion incremented by 1 mm. They reported that the gamma pass rates using 3%/3mm increased from  $58\% \pm 4\%$  (reference) to  $99\% \pm 1\%$  using trailing, when a baseline shift of 9 mm was present and corrected during dose delivery. Currently only 2D images with high temporal resolution can be acquired for real time tracking, but it can be combined with dynamic MLC tracking or gating when the target is within a tolerance of the

reference position. However, with the flexibility in the orientation and location of the chosen slice, one can still view different anatomical planes. AI-Ward *et al.*<sup>91</sup> evaluated the radiobiological advantages of MLC tracking (GTV+5 mm PTV margin) using an Elekta MR-Linac, as compared to conventional hypo-fractionated treatment plans (ITV+5 mm PTV margin) for liver, pancreas, and kidney cases (N = 9 in total). They assessed the size of PTV with and without an ITV margin using the Monaco treatment planning system (Elekta AB, Stockholm, Sweden) and determined the normal tissue complication probabilities (NTCPs) for OARs using the Lyman Kutcher Burman model<sup>100-103</sup>. In their study, the average size of the PTV which did not employ an ITV was reduced by 31.1%, 26.3% and 26.9% for liver, pancreas, and kidney patients, respectively. They demonstrated the potential in reducing the NTCPs for kidney, bowel, and duodenum using MLC tracking on a MR-Linac.

Regarding the robustness of the system, Fast *et al.*<sup>48</sup> evaluated the geometrical accuracy of dynamic MLC tracking using an Elekta MR-Linac and assessed the system latency using MV portal images. The tracking accuracy and the system latency were dependent on the MV imaging frequency and the maximum target speed. They reported the system latency for three treatment scenarios: the leaf motion (56–61 ms), “jaw+leaf” motion (71–78 ms), and “dynamic leaf guides+leaf motion” (58–72 ms) when the target position was sampled at 30 Hz. For tracking accuracy, the root-mean-squared errors (RMSE) were less than 1.5 mm for the leaf motion and “jaw+leaf” motion scenarios.

### **1.3. Robotic Stereotactic Body Radiation Therapy**

The name “stereotactic” historically referred to surgical procedures with a high degree of three-dimensional spatial precision enabled by rigidly fixing a frame to the skull. In the 1950s, neurosurgeon Lars Leksell from Stockholm, Sweden, developed a frame-based surgery system

called *stereotaxy* to navigate the surgical instruments to specific points accurately.<sup>104</sup> The first frame-based stereotactic radiosurgery (SRS) system, the Gamma Knife unit (Elekta AB, Stockholm, Sweden), was proposed by Lars Leksell, and a radiation physicist, Borge Larsson, “to replace the needle by narrow beams of radiant energy directed at the target in the brain and thereby produce a local destruction of the tissue”.<sup>105</sup> Many contemporary systems for the radiotherapy for brain tumours and functional abnormalities can also be completed by a frameless linac-based SRS system using a personalized mask (Figure 1-5b) for patient immobilization.<sup>106</sup> While SRS typically refers to intracranial treatments with high dose delivered in a single fraction, stereotactic body radiation therapy (SBRT) usually indicates a high dose per fraction delivered in one to five fractions in a treatment course to a non-cranial target.<sup>107</sup> SBRT treatments can be delivered on a conventional gantry-based linac or using the CyberKnife robotic radiosurgery system.<sup>108</sup>

As addressed in Section 1.2, the motion range for liver tumours can be as large as 11.9 mm (median) in the superior-inferior direction.<sup>33</sup> For SBRT on a conventional linac, abdominal compression, a larger internal target volume (see Section 1.1.4), and the gating technique might be needed for motion management.<sup>109,110</sup> In comparison, the CyberKnife stereotactic radiosurgery (SRS) system allows for free breathing because the tumour motion is compensated for. The planning target volume can be much smaller than that for treatments on a conventional linac, because of a reduced PTV-CTV margin and because no ITV is necessary. Apart from reduced irradiated volume and improved patient comfort, the biological benefits and concerns of late effects due to hypofractionation can be found in Section 1.1.2.

In the following subsections, the CyberKnife SRS system and the Synchrony motion compensation system are introduced, followed by a description of the clinical workflow for SBRT treatments using the Synchrony system at our center. Finally, the main issues associated with liver

treatments are addressed. A thorough description of the CyberKnife technical aspects can be found in a review by Kilby *et al.*<sup>73</sup>

### 1.3.1. CyberKnife Stereotactic Radiosurgery System

Major components of the CyberKnife G4<sup>TM</sup> Stereotactic Radiosurgery System at the Ottawa Hospital Cancer Centre are shown in Figure 1-6. CyberKnife is an image-guided radiosurgery system with a compact 6 MV linac mounted to a robotic manipulator (a robot arm). The beams can be shaped using a set of fixed-diameter circular tungsten collimators (12 cones available with aperture diameters from 5 to 60 mm), the Iris<sup>TM</sup> variable aperture collimator, and the InCise<sup>TM</sup> multileaf collimator (note: only available for CyberKnife M6<sup>TM</sup> and S7<sup>TM</sup>).<sup>73,111</sup>

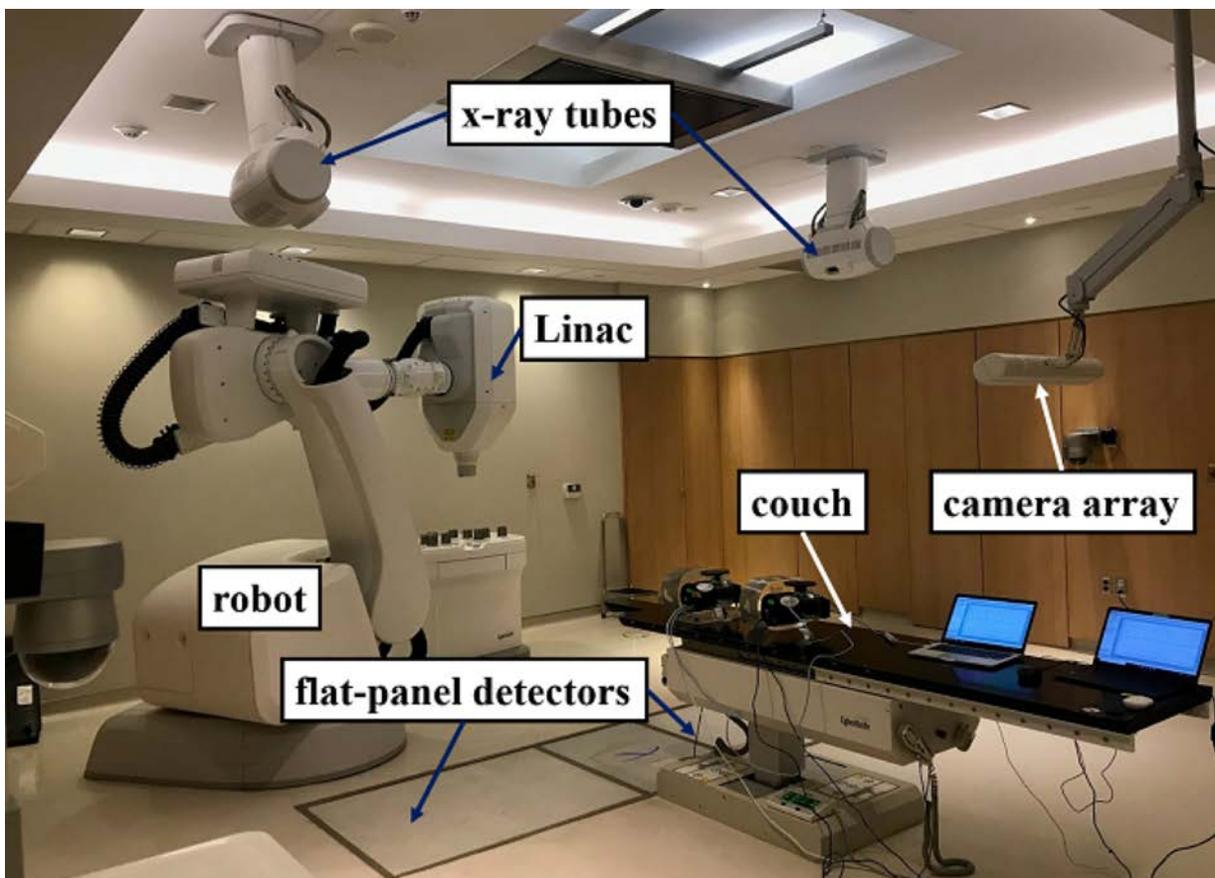


Figure 1-6. CyberKnife G4<sup>TM</sup> Stereotactic Radiosurgery System at the Ottawa Hospital.

The robot arm has six rotation axes to precisely maneuver the linac into different positions. The linac focal spot (the source point) can be positioned at a finite number of virtual locations, called *nodes*, in a three-dimensional workspace around the patient. The system is preconfigured for different *path sets*, consisting of a series of nodes, for intracranial and extracranial treatments. For example, the system employs different anatomy-specific paths for cranial and extra-cranial treatments with further specialized path sets for certain disease sites such as prostate cancer and trigeminal neuralgia (a nerve disorder discussed in greater detail in Section 5.2.1).<sup>73</sup> During treatments, the robotic arm is programmed to move the linac efficiently through the selected nodes along a *path* to safely deliver beams without colliding into the patient. The robot is programmed to keep the distance between the source and the center of the nodes (imaging isocentre) constant while the beam direction can be changed to treat lesions at different positions within the system's treatment workspace volume. This distance is different for the head path, body path, and trigeminal path, which is 80 cm, 100 cm, ~65 cm, respectively.<sup>112</sup> The robot arm can change the beam directions from each node to compensate for translations and rotations of the target during the treatment. The magnitude of rotation that the robot can correct depends on the path set and tracking method used for treatment.<sup>73</sup> Two patient positioning couches are available on this system. The standard couch, which is used at our clinic, adjusts the patient positioning with five degrees-of-freedom: three-dimensional translations, with roll and pitch angles up to  $\pm 5^\circ$  corrected automatically by the couch; but yaw angles must be corrected by manually repositioning the patient. The CyberKnife RoboCouch® Patient Positioning System has six degrees-of-freedom positioning capable of a larger range of translational motion than the standard couch and can correct  $\pm 5^\circ$  rotations about any axis.

There are two x-ray tubes (Figure 1-6) mounted on the ceiling that generate orthogonal kV x-ray image pairs. Two amorphous silicon flat-panel detectors (nominal resolution of 0.4 mm/pixel) are located on the floor, and image processing software is used to calculate the translational and rotational corrections by the geometric back projection based on a live x-ray image pair.<sup>73</sup> Before any treatment, the live images are registered to the digitally reconstructed radiographs (DRRs) generated using the planning CT volume. When fiducials are used, they are identified and aligned in both the DRRs and live x-ray images. The beam directions are repositioned by the robot to correct for translations (and rotations when available) of the target positions so that the beams are precisely aligned relative to the patient the same way as they are planned.<sup>71,73</sup> Note that a fiducial can be “disabled” for tracking if its location in the x-ray images has sufficiently deviated from its location in the DRRs due to potential migration.<sup>73</sup>

Three LED markers (not shown in Figure 1-6) can be placed on the patient surface, and the camera array registers the marker positions every 40 milliseconds. The system can predict the real-time tumour location using a correlation model (see Section 1.3.2) based on marker positions during beam delivery. During treatments, geometrical tracking errors (discussed later in Section 2.2.2) can be determined by calculating how far the model-assumed target location deviates from where it is planned. Due to the geometry of the dual orthogonal x-ray imaging system, the rotations in pitch (i.e., rotations about a patient’s left-right axis) and yaw (about the anterior-posterior axis) are “in-plane”, that is, they are directly observable in the projections. Rotations in roll (about the superior-inferior axis) are “out-of-plane” for both x-ray images.<sup>113</sup> Pitch and yaw angles can be directly determined from in-plane rotations by comparing x-ray images and DRRs. In contrast, a roll angle is determined by finding the “best match” of a pre-rotated DRR (from a library of rotated

reference DRRs) for a given x-ray image. For technical details of the algorithm used for estimating the target rotations during treatment, refer to Book Section 26.5 in Fu *et al.*<sup>72</sup>

The treatment planning systems used for this thesis work are the Accuray MultiPlan® for CyberKnife systems v.10.6 or earlier, and Accuray Precision® System for v.11.0 and v.11.5. The same two dose calculation algorithms (ray-tracing and Monte Carlo) are available in these two treatment planning systems. The simplest method with the shortest calculation time is the ray-tracing algorithm, which is usually used when treating brain or soft tissue that can be well approximated as radiologically similar to water.<sup>73,114</sup> The algorithm calculates the dose deposited in every voxel from each beam, based on measurements made in water at the time of commissioning. The electron density for a voxel is converted from the CT number, which corresponds to the image voxel intensity of a CT image for a specific CT scanner. To account for the heterogeneity of the medium, the algorithm calculates the radiological water equivalent depth along ray-lines from the source to a given position within the planning CT (target point) using the electron density relative to water. The dose for each voxel is calculated in terms of the effective depth, the off-centre radius, and the collimator size using the empirically measured commissioning data. This approach does not explicitly model photon scattering or the deposition of energy by electrons set in motion in the medium.

Alternatively, for more accurate dose calculations in heterogeneous tissues, one can use Monte Carlo (MC) dose calculation to consider all photon and electron processes. This method provides the dose calculation within an uncertainty of 2% for a reasonable calculation time (<5 min).<sup>115</sup> As described by Ma *et al.*,<sup>115</sup> with precise source modelling and beam commissioning, the MC model simulates the histories of primary and scattered photons. Random numbers are generated to determine the trajectories of particles until they are absorbed (energy below a predefined cut-off

value) or exit the patient geometry, the type of photon interactions, and energy deposition by electrons set in motion. To shorten the calculation time, precalculated electron histories simulated in water are used rather than explicitly modelling electron transport. The electron step length, scattering and stopping power ratios are modified based on the local material density.<sup>115</sup> The MC dose calculation is preferred for anatomical sites that contain heterogeneities, such as bone, lungs, and sinus cavities. For the smaller cone sizes (diameter < 15 mm) used with CyberKnife, the impacts of the lack of electron transport in low density regions were shown to be dosimetrically significant for lung tumours.<sup>114,115</sup> Wilcox *et al.*<sup>114</sup> compared the dose distributions generated by the two dose calculation algorithms provided by CyberKnife for 33 patients treated for lung tumours to demonstrate the necessity of MC calculation for lung tumours or tumours near the diaphragm. In their study, the ray-tracing method overestimated the maximum dose to the PTV by a factor as large as 1.63, and the discrepancy was worse when using smaller cones.

### **1.3.2. Synchrony Tracking System**

Before the treatment, fiducials (cylindrical platinum seeds, 0.8 mm in diameter and 3.0 mm long) are implanted, preferably near a tumour. The target location is determined based on fiducial locations used for tracking. The LED markers are placed on the patient's surface, and the locations are recorded by the camera array at 26 Hz. The Synchrony Respiratory Motion Tracking System (MTS) compensates for tumour motion using a predictive model (henceforth referred to as the 'model') which correlates the internal breathing motion derived from the fiducial positions with the external motion observed in near real-time using the LED and camera system.<sup>71</sup> At least one model is generated by the system for each treatment fraction and each model is updated throughout the treatment as new x-ray images are acquired.

A screenshot of part of the interface of the tracking system is shown in Figure 1-7. For each x-ray image acquisition, a new model point containing the locations of the target and markers is added to update the model. The model is based on, at most, the fifteen most recent model points.<sup>71,73</sup> The model can be linear or non-linear and is fit to the marker and fiducial locations.<sup>73,116</sup> It predicts the target position 115 ms in advance to overcome the latency between the hardware and software involved in beam delivery.<sup>73</sup>

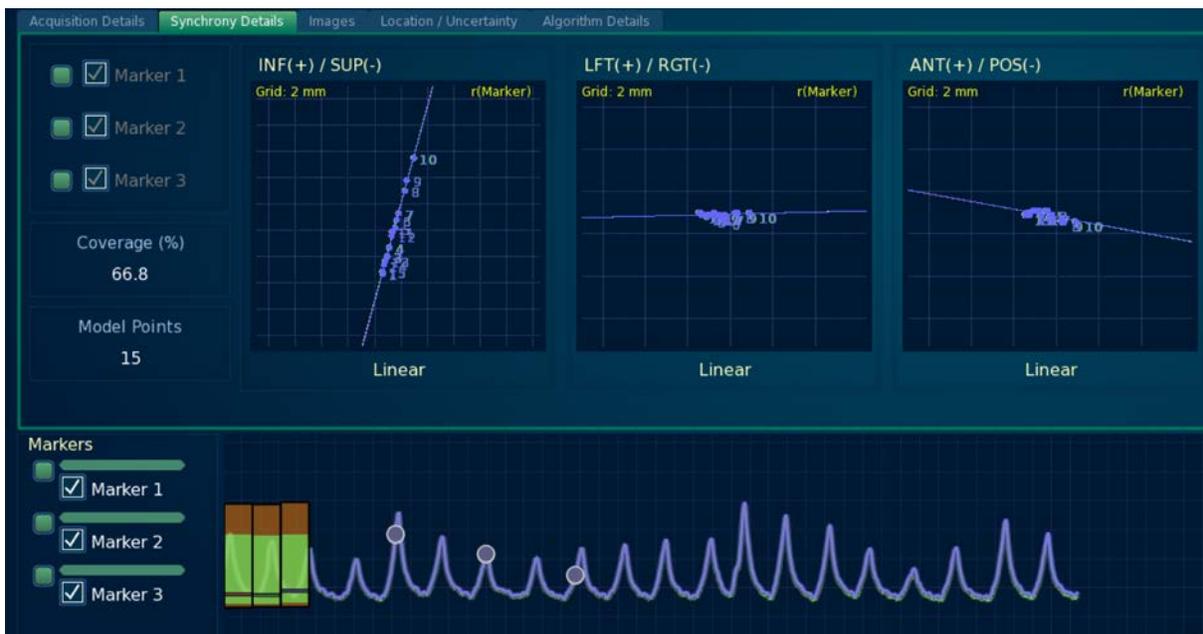


Figure 1-7. A screenshot of part of the interface of the Synchrony Respiratory Tracking System (v.11.5) during treatment. The three plots on the top indicate a linear correlation between the fiducial and marker locations in three anatomical directions. Data points are sampled at the acquisition of x-ray images. The bottom plot shows the marker motion tracked by the camera at 26 Hz.

### 1.3.3. Typical Clinical Workflow

The typical clinical workflow used for robotic liver SBRT treatments on CyberKnife at our clinic is discussed below.

**Patient Selection.** At the Ottawa Hospital Cancer Centre, approximately 70% of the patients treated on CyberKnife for liver cancer have metastatic tumours in origin (the remaining are from primary liver cancer). Metastatic colorectal cancer is the most common type of liver metastasis we treat on CyberKnife. Many of these patients are considered to have an oligometastatic disease, a proposed disease state<sup>117,118</sup> in which there are only a few sites of metastatic disease and localized cancer treatments are curative in a subset of patients.

**Fiducial Implantation.** A week before the CT simulation and magnetic resonance imaging (MRI), the fiducials are implanted trans-cutaneously in an ultrasound-guided (sometimes CT-guided) procedure, or through endoscopy. Generally, smaller tumours (diameter < 5 cm) are preferred and fiducials are implanted in the healthy liver tissue surrounding the lesion since the needles can spread the disease if they pass through a tumour during the implant procedure.

**CT Simulation.** Three image datasets (planning CT, contrast CT, and 4DCT) are acquired *without* abdominal compression during the CT simulation for treatment planning. The planning CT is a non-contrast end-of-expiration breath-hold CT image set. It is used for dose calculation, and the slice thickness is usually 1 mm. The voluntary breath hold helps to minimize the image artifact due to target motion and is essential so that the bony anatomy and fiducials are well defined in high resolution for the DRRs. The end-of-expiration breath-hold contrast CT is also used to visualize the tumour location and vasculature. The delineation of targets at our center is usually done on MR images only. However, on very rare occasions the contouring of targets is determined from the contrast CT if the patient cannot have an MRI or if the MRI quality or fusion is unsatisfactory.

A 4DCT scan is acquired at inhalation, expiration, and an arbitrary respiratory phase. The maximum-intensity-projection (MIP) is reconstructed for 4DCT images to evaluate the fiducial

motion shown as “streaks” at different respiratory phases. If a fiducial’s trajectory during the breathing cycle moves differently (in either the amplitude or the direction of the streak), the treating therapists at CyberKnife suite are notified and therapists might preferentially disable that fiducial for motion tracking during treatment. The inhale and exhale phases of 4DCT are fused to the planning CT by fiducial-to-fiducial alignment. If the exhale phase CT and the planning CT do not fuse well (especially with very large angles), it indicates a potentially significant difference between the patient’s normal breathing and the breath hold. In that case, a second CT simulation may be needed to reduce the difficulties for patient setup before treatments. The inhale phase CT is used to evaluate the motion of OARs with respiration and assess if the planning risk volumes (PRVs) are reasonable.

**Treatment Planning.** The radiation oncologists prescribe the dose and delineate the GTV based on the planning CT fused with secondary MR images. Using the Accuray treatment planning system (MultiPlan® or Precision®), they contour the GTV and review OAR and PRV contours generated by the radiation therapy staff and may contour some critical dose-limiting OARs. PTVs are expanded from the GTV with ‘no margin’ for a presumed microscopic extension of the tumour (i.e., GTV = CTV). This is based on the approach adopted for lung SBRT trials by the Radiation Therapy Oncology Group.<sup>119</sup> Isotropic 5 mm PTV margins expanded from GTVs are typically used for planning at our institution.<sup>119</sup> Occasionally, 7 mm PTV margins are used to account for poor image quality, difficulty achieving a good match between fused images, imperfect fiducial implantation (especially distance from target), very small lesion size, and other clinical considerations. If targets are adjacent to OARs, the prescription dose can be reduced to meet dose constraints for OARs. In our institutional planning guideline, the prescription dose is 60 or 54 Gy in 3 fractions, respectively, for the tumour size < 3 mm and between 3 to 5 cm in diameter without

abutting OARs. The prescription is reduced to 42 Gy to 48 Gy in six fractions if the tumour is larger than 5 cm in diameter or within 1 cm of the duodenum or stomach (serial structures, see Section 1.1.4). A plan is then generated in a semi-automated optimization process in which potential arrangements of candidate beams at different nodes using a specific path set and their dosimetric weights are adjusted to meet planning objectives for targets and OARs.<sup>73</sup> Dose is calculated using either the ray-tracing or Monte Carlo dose calculation algorithm. If the plan is satisfactory, it is approved by the physician and made ready for treatment following several quality assurance checks by therapists and a physicist.

**Treatment Delivery.** Using an in-room laser localization system the immobilization devices (headrest on a thick sponge, leg cushion, and Vac-Lok vacuum cushion), patients are set up for treatment in the same position as for CT simulation based on skin marks (tattoos) and straightening lines on the Vac-Lok cushion. When the treatment starts, kV x-ray images are acquired, and a Synchrony model is created. Treatments usually start within a week or two after the CT simulation. Treatment fractions are completed at least every other day and normally the entire treatment course takes less than 10 days.

**Follow-Up.** The standard follow-up is three months after treatment. It includes a physical exam, follow-up contrast CT scans to determine the tumour size and distal recurrence disease, and albumin blood tests for hepatic toxicity.<sup>120-122</sup>

### **1.3.4. Challenges of Accurate Treatment Delivery**

The work of this thesis is partially motivated by two issues that occur frequently for CyberKnife liver SBRT which may impact the efficacy of the treatment and accuracy of the delivered dose. These issues are the compromised prescription dose due to PRV-PTV overlap, and uncorrected and/or uncorrectable rotations.

**PRV-PTV Overlap and Adaptive Margins.** Compared to certain other trials for liver SBRT treatments conducted in other institutions,<sup>123,124</sup> the prescription at our clinic is relatively conservative, mostly due to the concerns of treatment toxicity and based on the evidence from Princess Margaret Hospital in Toronto.<sup>125,126</sup> Typically, a dose of 42-60 Gy over three to six fractions of a treatment course is prescribed by the radiation oncologist, depending on the tumour size, number of lesions, and proximity of tumours to OARs (see Section 1.3.3). The 95% confidence interval of liver volumes for females and males were reported as 1294-1502 and 1796-1956 cubic centimetres (cc), respectively, by Andersen *et al.*<sup>127</sup> To preserve adequate liver function, our institutional treatment planning guidelines require that at least 700 cc of the healthy liver (liver volume excluding the GTV and radiofrequency ablation volumes if present) receive a dose less than 15 Gy.<sup>128</sup> In the clinic, we observed a significant difference ( $P < 0.05$ ) in the prescription dose and PTV D99 (near-minimum dose to the PTV) when there is an overlap between the PTV and serial organ PRVs such as in case of the digestive organs (stomach, small bowel and duodenum), as illustrated in Figure 1-8. PTV D99 is an important metric that correlates to the tumour control probability.<sup>26</sup> When the PTV is adjacent to OARs, the prescription dose is often reduced to meet the dose constraints for OARs, resulting in a *compromised* treatment plan. This motivates us to investigate if patient-specific margins could be used in an adaptive clinical workflow for these treatments. If smaller PTV margins could be safely used for a subset of patients, the overlap volume could be reduced and potentially a higher dose can be prescribed to tumours without an increase in toxicity. If the margins being used for certain patients are too large and do not realistically reflect the delivery errors associated with their treatments, the near-minimum dose to the tumour will be underestimated and can be a poor predictor of tumour control.

A review by Green *et al.*<sup>129</sup> classified the adaptive process for a treatment course into three types based on timescales: offline (inter-fraction), online (immediately before a treatment fraction), and real-time (intra-fraction). The first step for adaptive treatment therapy is to identify the necessity (such as changes in internal anatomy and improper immobilization devices) and timing for change. The adaptation can happen during any step of a clinical treatment workflow, including dose prescription, CT simulation, planning, delivery, and inter-fraction assessment. In our study, the necessity of reducing the PTV margin is based on the magnitude of motion-tracking errors. In Chapter 3, we propose a clinical workflow to ensure the safety and efficiency of implementing patient-specific PTV margins for prospective liver treatments on CyberKnife. The process requires, prior to treatment planning, using a machine learning algorithm called support vector classification (SVC) to determine if a specific patient is a good candidate for margin reduction. An off-line assessment of motion tracking errors is also performed at the end of each fraction to assess the geographical target coverage for any patient treated with reduced margins.

Machine learning was initially introduced to model neuron cell interactions, first described by Donald Hebb.<sup>130</sup> As a powerful approach for data management and modelling, it has been applied to many applications within medicine, including the field of medical physics. In general, it involves building a mathematical model to make decisions based on input data, also known as “training data”. Machine learning has three categories that are commonly used in radiation oncology: supervised learning, unsupervised learning, and semi-supervised learning.<sup>131</sup> A supervised machine learning algorithm, such as linear discriminant analysis or support vector machine, uses labelled data to train a model, which requires that the label is known for training and testing data. In contrast, unsupervised machine learning algorithms draw inferences from unlabeled data, which is typically used in clustering analysis and dimension reduction (e.g., principal component

analysis). A semi-supervised learning combines supervised and unsupervised learning and requires labelled and unlabeled data, which is preferable when it is costly to get sufficient labelled data. A review by El Naqa *et al.*<sup>132</sup> discussed the applications in radiation oncology that employ machine learning models, and emphasized the importance of validating model robustness to reduce the risk of model overfitting. Support vector classification, the supervised machine learning algorithm used in this thesis work (see Chapter 3), has been applied to classification problems for many different applications in radiation oncology and medical physics. Some examples of these applications include local<sup>133</sup> and distant<sup>134</sup> tumor control, toxicity studies,<sup>135,136</sup> staging,<sup>137</sup> image segmentation,<sup>138</sup> and predicting quality control (QC) results.<sup>139</sup>

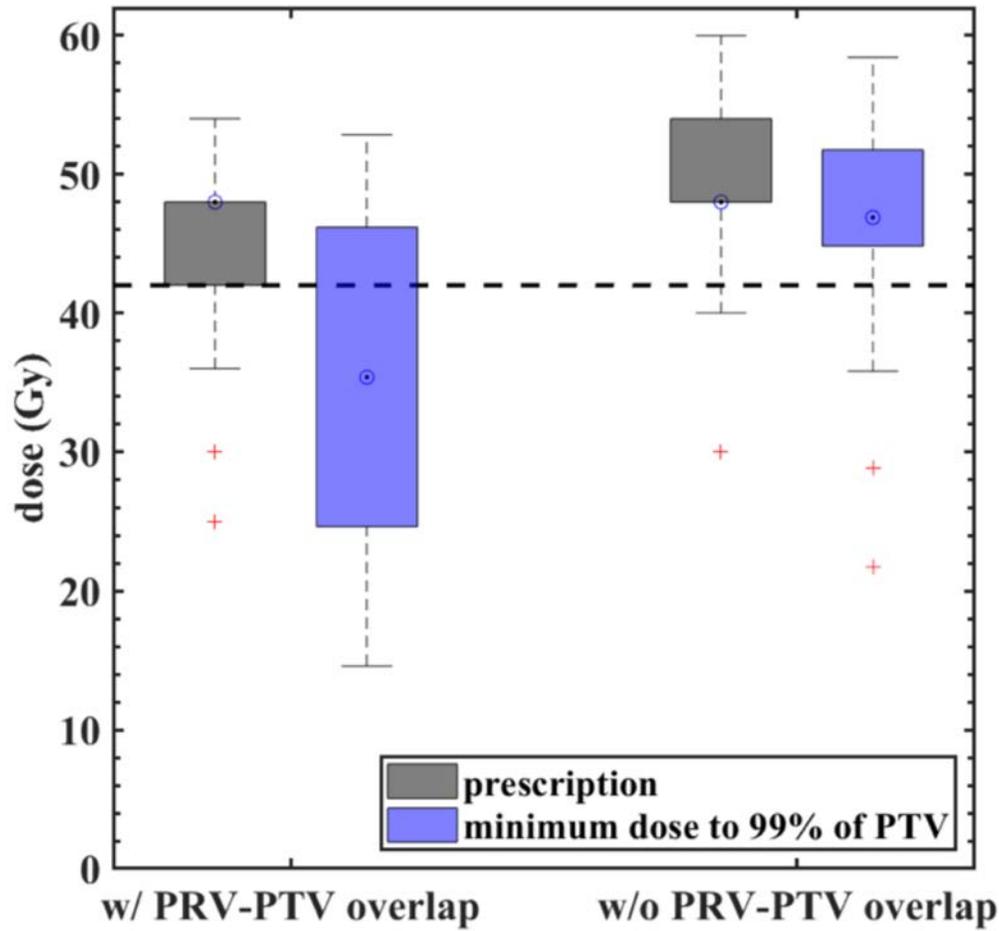


Figure 1-8. Comparison of the prescription dose and the near-minimum dose to PTV for 57 targets with non-zero volume overlap of PTV and PRV (compromised) and 121 targets without volume overlap of PRV and PTV (non-compromised). Each box spans from the first quartile to the third quartile. The circles and red crosses indicate medians and outliers, respectively. The dashed line indicates 42 Gy, the lower threshold of the prescription dose typically used for robotic liver SBRT treatments.

**Uncorrected Rotations (Liver).** Target rotations during treatment are estimated using the kV image data with respect to the DRRs (see Section 1.3.1 for the algorithm used to estimate rotations), but they are often uncorrected by the system. The treatment delivery system has somewhat limited capability to correct for target rotations during the treatments. The standard CyberKnife couch can correct for up to  $\pm 5^\circ$  in pitch and roll but yaw must be corrected by manually repositioning the patient. Using the recommended settings for extra-cranial treatments on CyberKnife with

Synchrony motion tracking, the robotic arm can only correct for up to  $\pm 1.5^\circ$  for roll and pitch, and  $\pm 3.0^\circ$  for yaw. In addition, due to the complexity of the fiducial implant procedure, it can be challenging to meet the spatial configuration of fiducials required for the system to accurately estimate the rotations. This requirement is that at least three fiducials form a triangle with the smallest angle larger than  $15^\circ$  and with the shortest side length larger than 20 mm (see Section 2.2.3).<sup>140</sup> In cases where these requirements are not met, it may be safer not to correct for rotations since the uncertainty in their values may be too high.<sup>141</sup> Using Monte Carlo generated random configurations of fiducials, Murphy investigated three computational solutions to determine the fiducial location and orientation.<sup>141</sup> Murphy found implantation of at least four fiducials, as compared to three, gave much improved accuracy for estimating the rotations, and addressed the importance of having a larger inter-fiducial distance to reduce the uncertainty in the estimation. If the system does not correct for rotations, it can lead to a large dose reduction to targets. Effects associated with uncorrected rotations should be quantified and acknowledged for the sake of treatment efficacy and patient safety.<sup>140,142</sup> A dose perturbation tool is needed to evaluate the near-minimum dose to targets affected by the uncorrected rotations. In Chapter 4, we present a new tool for evaluating the dosimetric impacts of delivery errors associated with the CyberKnife system. We apply the tool for liver treatments (which is the primary focus of this work) and to trigeminal neuralgia, a very specialized cranial treatment done on CyberKnife, which is discussed in more detail in the following section.

**Uncorrected Rotations (Trigeminal Neuralgia).** Rotations are very important not only for liver treatments, but also for trigeminal neuralgia (a type of cranial nerve disorder) treatments. This is because the target is very small (typically 6 mm for the nerve root), dose is very high (60 Gy in a single fraction), and OARs (such as the brainstem) are nearby. For this site, a specialized set of

beams (see Section 1.3.1 for “trigeminal path”) are used, and the linac is brought closer to the patient to deliver the required high dose. However, rotational corrections cannot be applied because of the risk of collision with the patient, and so that target rotations are only monitored and not corrected for during treatments. The volume and the length of the nerve treated with high dose are important for the treatment outcome.<sup>143</sup> The dose perturbation tool is also applied to this site to demonstrate its general applicability and address the appropriate tolerances on rotational corrections for these treatments.

#### **1.4. Thesis Overview**

Tumour motion is compensated for using the Synchrony system on CyberKnife, but different types of delivery errors cannot be neglected. The detailed post-treatment system log file data and extensible markup language (XML) data can provide information to determine PTV margins. Analysis of these very detailed log files may provide insight into what breathing pattern features are predictive of higher or lower model errors and could be used for margin reduction prior to treatment. The main goal of this thesis work is to validate a clinically feasible approach for assigning patient-specific PTV margins for liver treatments on CyberKnife. This approach is intended for use in a prospective clinical trial, with a proposed workflow to ensure the patient’s safety and the efficacy of the treatment. The thesis work includes several steps below.

- (1) Quantification of tracking errors, including tracking residual errors, target rotations and deformations, for 72 liver treatment plans. (Chapter 2)
- (2) Classification of the magnitude of tracking errors using features identified from the previous step via a supervised machine learning algorithm. (Chapter 3)
- (3) Proposal for a clinical workflow to treat patients with reduced PTV margins when they are predicted to have lower tracking errors. This includes the implementation of a warning

- system to monitor if the reduced PTVs are sufficient to enclose the GTV after taking into account the motion-tracking residual errors for every subsequent fraction. (Chapter 3)
- (4) Proposal for a general analytical tool to evaluate the dosimetric impacts of delivery errors for CyberKnife treatments, which might be included in the warning system for post-fraction dosimetric assessment of delivery errors for liver SBRT. (Chapter 4)
- (5) Experimental validations using radiochromic films for the perturbed dose distributions considering translational tracking residual errors and rotations. (Chapter 4)
- (6) Dosimetric assessments for two very different types of treatments, liver cancer (N = 148) and trigeminal neuralgia (N=24), to demonstrate the broad applicability of the dose perturbation tool. (Chapter 5)
- (7) Quantification of plan sensitivity to uncorrected rotations for liver and trigeminal neuralgia treatments and their clinical implications are explored. (Chapter 5)

## Chapter 2. Uniform PTV Margins

### 2.1. Introduction

For stereotactic body radiation therapy (SBRT) treatments which employ respiratory motion compensation, the planning target volume (PTV) margin should take into account the delivery errors for motion tracking, uncorrected rotations, deformations, and estimates of the overall delivery accuracy measured for static targets.<sup>68,144,145</sup> In this chapter, we estimate population-based PTV margins for a relatively large cohort of 72 patients treated between 2011 and 2017 at the Ottawa Hospital Cancer Centre for liver cancer with CyberKnife robotic radiosurgery using Synchrony respiratory motion compensation.\*

Researchers have used treatment log files to estimate treatment margins for thorax,<sup>145</sup> lung,<sup>146-148</sup> and liver<sup>142,149,150</sup> lesions treated with CyberKnife SBRT treatments. The previous studies<sup>76,142</sup> that quantified uncertainties for liver treatments on CyberKnife had a relatively small cohort size (N<30). Details of inter- and intra- patient variations in tracking errors, effects of uncorrected rotations, and target deformations were not evaluated in most of those earlier studies. Some research groups assessed the deformations<sup>151</sup> and the resulting dosimetric effects<sup>152-154</sup> using either 4DCT or CBCT data. Cao *et al.*<sup>155</sup> evaluated the systematic and random errors associated with rotational setup variations for SBRT liver treatments by manually matching CBCT images with the planning CT images and found that more than 14.7% of the patient population had more than 2° of rotational error in at least one rotational direction (roll/pitch/yaw). In one study examining CBCT data for 29 liver patients, Bertholet *et al.*<sup>156</sup> found a large day-to-day variation of rotations

---

\* The results in this Chapter have been published in a peer-reviewed journal

(up to 16.9° and 28.6° in roll and yaw, respectively), which may be important for liver treatments especially without abdominal compression as is typical for CyberKnife treatments.

As described in greater detail in Section 1.3.3, isotropic 5 mm PTV margins are typically used for planning at our institution.<sup>119</sup> At our center, the PTV is expanded directly from the GTV (with no CTV margin) for liver SBRT treatments following the approach adopted for lung SBRT trials by the Radiation Therapy Oncology Group (RTOG). Treatment delivery errors (including tracking errors, target rotations, and rigid body errors) and the linearity of tracking models, which potentially affect the geometric tracking accuracy for liver SBRT treatments, are assessed in this chapter. We also investigate if non-isotropic margins are warranted based on our data.

## **2.2. Method and Materials**

### **2.2.1. Patient Data**

Treatment log files have been assessed for 72 patients (67%/33%: liver metastases/liver primaries, 70 ± 11 years old, 46 Males/26 Females, typically 3 to 6 fractions) treated between 2011 and 2017. Special filtering was applied to exclude data that was not directly used for treatment (e.g., exclusion of data acquired during the patient setup, beam-off periods and manually rejected by radiation therapy technologists). For data used in this study, we received the informed consent of patients as a part of the Research Ethics Board (REB) approved local clinical trial.

For these treatments completed prior to 2018, different versions of the CyberKnife MultiPlan® planning system (from v.3.5 to v.4.6) and delivery system (from v.9.0 to v.9.6) were used. Depending on the size of the patient, the CT spatial resolution was 1.0×1.0×1.0 mm<sup>3</sup> or smaller. Recommended margins are rounded up to the nearest integer in mm due to partial volume effects associated with contour expansions in the planning system. In this study, data analyses are performed using MATLAB 2018b (MathWorks Inc., Natick, MA).

### 2.2.2. Synchrony Errors

The combined uncertainty in the CyberKnife delivery and Synchrony motion tracking system (MTS) should be accounted for when estimating PTV margins for these treatments. Synchrony error is calculated as  $\sqrt{e_{corr}^2 + e_{pred}^2 + e_{e2e}^2}$ ,<sup>76</sup> where  $e_{corr}$ ,  $e_{pred}$ , and  $e_{e2e}$  are correlation, predictor, and end-to-end error, respectively. The correlation error is a measure of the motion-tracking residual errors and is the difference between the predicted and the measured target positions determined for each X-ray image pair acquired during patient treatment. The predictor error is the error compensating for system latency in beam positioning. It is calculated as the difference between the target position predicted 115 ms earlier and the predicted target position inferred from the Synchrony model using the current LED positions that are sampled at a frequency of 26 Hz. The end-to-end error is a global beam positioning error associated with the CyberKnife system. It is estimated from measurements made with a static phantom as part of routine QC. According to quality assurance guidance documents,<sup>29,30</sup> the end-to-end error should be maintained at < 1 mm in any direction, and < 1 mm total radially.

Correlation errors that are larger than 5 mm in any direction can trigger a “soft stop” of a treatment. At that point, the radiation therapists can reassess the model validity and/or determine if the system has identified the fiducials correctly in the acquired images. It may be necessary to reset the Synchrony model a few times during each treatment fraction due to changes in breathing patterns, patient movement, or software issues. Therefore, the cause of > 5 mm correlation errors cannot be determined in the Synchrony system log file versions. In this chapter, the x-ray acquisitions with correlation errors that are larger than 5 mm have been filtered from our margin estimates.

The correlation error per model is calculated as the root-mean-square (RMS) of correlation errors reported within the model. The predictor error per model is the RMS of predictor errors during beam-on periods within a model, following the approach of Pepin *et al.*<sup>147</sup> To assess the intra-fractional variation of the predictor, the RMS predictor error over each beam is also calculated. The Synchrony error for each model is calculated as the square root of a sum of squares of RMS correlation error, RMS predictor error and end-to-end error (set to its maximum 1 mm in any direction and radially). The radial Synchrony error per patient is calculated as the mean value across models. We extract the marker position and split the external respiratory traces into eight equal-width time bins to investigate the phase dependence of the *signed* correlation and predictor error per patient (i.e., to see how the *direction* of these errors change during breathing).

One of Van Herk's most well-known margin recipes<sup>157</sup> is used for comparison with our estimation of the PTV margins. It is calculated as  $2.5 \Sigma + 0.7 \sigma$ , where  $\Sigma$  is the systematic error and  $\sigma$  is the random error.  $\Sigma$  is calculated as the standard deviation (SD) of the residual setup error (taken as the radial Synchrony error for this study) across the entire patient population.  $\sigma$  is the RMS across the entire patient population of the SD of the residual setup error computed across all the fractions for each patient.

### 2.2.3. Rotational Corrections

During the treatment, tracking on specific fiducials may be disabled when the system cannot identify them in the X-ray image data or when acquired image data cannot match the digitally reconstructed radiographs (DRRs). Sometimes the rotational corrections are manually disabled, or rotations are not tracked because of frequent soft stops initiated by larger rotations that the system cannot correct for. Having more than three fiducials used for tracking is recommended to reduce uncertainty in rotation corrections.<sup>32</sup> During the treatment, the rotation can be accurately calculated

when the spatial configuration of fiducials meets the following criteria: at least a set of three fiducials forming a triangle with the minimal angle  $>$  the collinearity angle and with the minimum side length  $>$  the spacing threshold.<sup>140</sup> The collinearity and spacing threshold, by default, are  $15^\circ$  and 20 mm, respectively, according to specifications stated in the MultiPlan Treatment Planning Manual. Only rotational corrections calculated using fiducial positions that meet all these criteria are assessed. For all treatments using the Synchrony system at our center, the patient is initially set up using the Xsight spine tracking setup plan. If the criteria for accurate rotational corrections are not met or the magnitude of rotational variation exceeds system limits, rotational corrections are turned off and the couch rotations are not changed from the spine setup values.

Note that the rotational matrix in three-dimensions is not commutative, so the order in which the rotations are applied is important. Appendix A describes the relevant coordinate systems and conventions that are used by different CyberKnife components and in log files. The translational and rotational corrections reported in this thesis are given in a patient-based coordinate system using Euler angles. The axes are defined such that inferior is the positive x-direction, the patient's left is positive in the y-direction, and anterior is positive in the z-direction. Roll, pitch, and yaw are defined as rotational corrections about the x, y, and z axes, respectively, with signs following the right-hand rule. For example, a positive rotational correction in roll means that the patient should be rotated about the superior/inferior axes toward the patient's right. The transformation matrix listed in Appendix B was verified by private correspondence with Accuray and confirmed in Chapter 4.

As addressed in Section 1.3.4, the system is limited in the magnitude of rotational corrections that it can apply during the treatment. For body treatments using recommended planning settings, the system can correct for up to  $\pm 1.5^\circ$  for roll and pitch, and up to  $\pm 3.0^\circ$  for yaw. According to

the MultiPlan treatment delivery manual, rotation corrections about the imaging system isocenter are applied (if they can be calculated, are correctable by the system and are enabled) based on the *last* x-ray image acquired before the beam delivery.

The contours and planned dose distributions are extracted from Digital Imaging and Communications in Medicine (DICOM) files for 72 patients, and the dose-volume histogram (DVH) for the planned dose distributions are calculated. To estimate the dosimetric impact of rotations, we rotate the planned dose distributions about the imaging isocenter using the inverse of the rotation correction matrix and recalculate the DVH. The changes in dose to the GTV are quantified using the minimum doses received by 95% (denoted as D95) and 99% (denoted as D99) of the target volume after the rotations are applied and expressed as a fraction of the prescribed dose ( $R_x$ ) for each patient.

To verify this approach, for three patients we rotated CTs by the average rotation over a treatment course using a rotation matrix (see Appendix B) and recalculated the dose using the planning system. Note that rotating the entire CT volume assumes that a patient is a rigid body. For these patients, rotation corrections at the time of treatment were larger than  $3^\circ$  in at least one direction on average across the entire treatment course. Three-dimensional gamma analyses<sup>28,158-160</sup> (see Section 1.1.5) were performed using 2%/2mm criteria for these patients to assess the impact of neglecting changes in tissue density and the physical path length. Since the method used to estimate the dose to target(s) ignored the change of material density along the beam path, the correlation between the dosimetric change to targets and the distance between the imaging isocenter and the center of fiducials was also investigated.

#### **2.2.4. Rigid Body Error**

According to specifications stated in the MultiPlan Treatment Delivery Manual, the rigid body error is calculated by the system when an X-ray image pair is acquired. It is defined as the *absolute* change in the inter-fiducial distance per active fiducial pair compared to that determined in the CT volume.<sup>145</sup> However, the user cannot distinguish between target contraction and expansion by these absolute quantities. Hence, we calculate *signed* rigid body errors where positive and negative values indicate target expansion or target contraction, respectively. The average of signed rigid body errors for all active fiducial pairs represents the *average* change of inter-fiducial distances. Since tracking on specific fiducials may be turned on or off, we only consider fiducials that are presently active in the Synchrony model. A fiducial may be excluded from a model if it is the cause of a very large rigid body error that could indicate fiducial migration. A fiducial may also be removed from a model due to poor live image contrast. In this case, no fiducial may be visible to the treating therapist where the software has identified it or the system itself may report that the uncertainty for fiducial identification is too high. Rigid body errors of 66 patients (with >1 fiducial enabled) are assessed. The number of fiducials used for tracking and implanted inside the GTV contours is investigated.

#### **2.2.5. Model Linearity**

Since the system models the relationship between external and internal motions as either linear or curvilinear,<sup>116,161</sup> the linearity between internal and external motions (hysteresis effect) is investigated. The target and marker positions are determined per X-ray image acquisition and reported in a log file. A linear fit is determined for the marker and target positions as a function of time across the entire model, and the drift component is corrected. We also quantify the correlations among Synchrony errors, the range of target SUP-INF motions, and the range of

marker radial motions for all models. The multiple correlation coefficient<sup>162</sup> is used to evaluate how well the internal and external motions can predict Synchrony errors. The  $R^2$  value per model is used as an indicator of model linearity between SUP-INF positions of the target and the average radial positions of external markers within a model. Inter-patient and intra-patient (inter-model) variations of the magnitudes of the target and marker motions are also assessed.

### 2.2.6. Hypothesis Testing

Two-tailed Student's and Welch's t-tests<sup>163,164</sup> are used to calculate statistical significance ( $P \leq 0.05$ ), and the 95% confidence interval for the mean difference between two samples is reported. Spearman correlation coefficients (Spearman r) are used if Lilliefors statistics test  $P > 0.05$ .<sup>165</sup> The impacts on the tumour tracking accuracy of age, gender, and lesion position (left or right side of the liver), and the distance along the superior/inferior direction to the maximum superior slice of the liver contour (specified as "dome-to-target distance"), are assessed. Refer to Appendix A for how to convert a point of interest (such as the fiducial location) from the patient coordinate to the DICOM coordinate. Also investigated is the relationship between the magnitude of rigid body errors and the number of fiducials used for tracking. Most statistical tests and figures generated in this chapter are performed using OriginPro 2018b (OriginLab Corp., Northampton, MA).

## 2.3. Results

### 2.3.1. Synchrony Error

No significant differences in Synchrony errors (95% CI -0.1 to 0.0 mm,  $P \leq 0.05$  for SUP-INF, LT-RT, and ANT-POST directions) are observed along the positive and negative anatomical directions (Figure 2-1a). The errors in the SUP-INF direction are significantly different from those in the other two directions (95% CI +0.4 to +0.5 mm,  $P \leq 0.0001$  for LT-RT direction; 95% CI +0.3 to +0.5 mm,  $p \leq 0.0001$  for ANT-POST direction). The 95<sup>th</sup> percentiles of the Synchrony errors in SUP-INF, LT-RT, and ANT-POST directions for our patient population are 2.3, 1.6, and 1.8 mm,

respectively. Figure 2-1b illustrates the inter-patient and intra-patient (inter-model) variations in Synchrony errors. The standard deviation (SD) of Synchrony errors for a treatment course is less than 1 mm for most patients. Figure 2-2 shows inter- and intra- fractional variations of correlation errors per X-ray image pair and RMS predictor errors per treatment beam. Using the Van Herk margin recipe for the entire patient population, 2 mm-isotropic margins will be needed to cover the CTV with 95% of the prescription dose for 90% of the patients.

Figure 2-3 shows for all patients the average and SD of signed correlation and predictor errors across the respiratory phases per patient. The systematic effect for both correlation and predictor errors across all respiratory phases approximately cancels out within  $\pm 0.5$  mm and  $\pm 0.05$  mm, respectively, for most patients. The average SD of correlation errors across respiratory phases in SUP-INF, LT-RT and ANT-POST directions are 0.4, 0.2 and 0.3 mm, respectively.

Synchrony errors and inter-model SD of radial errors for 72 patients

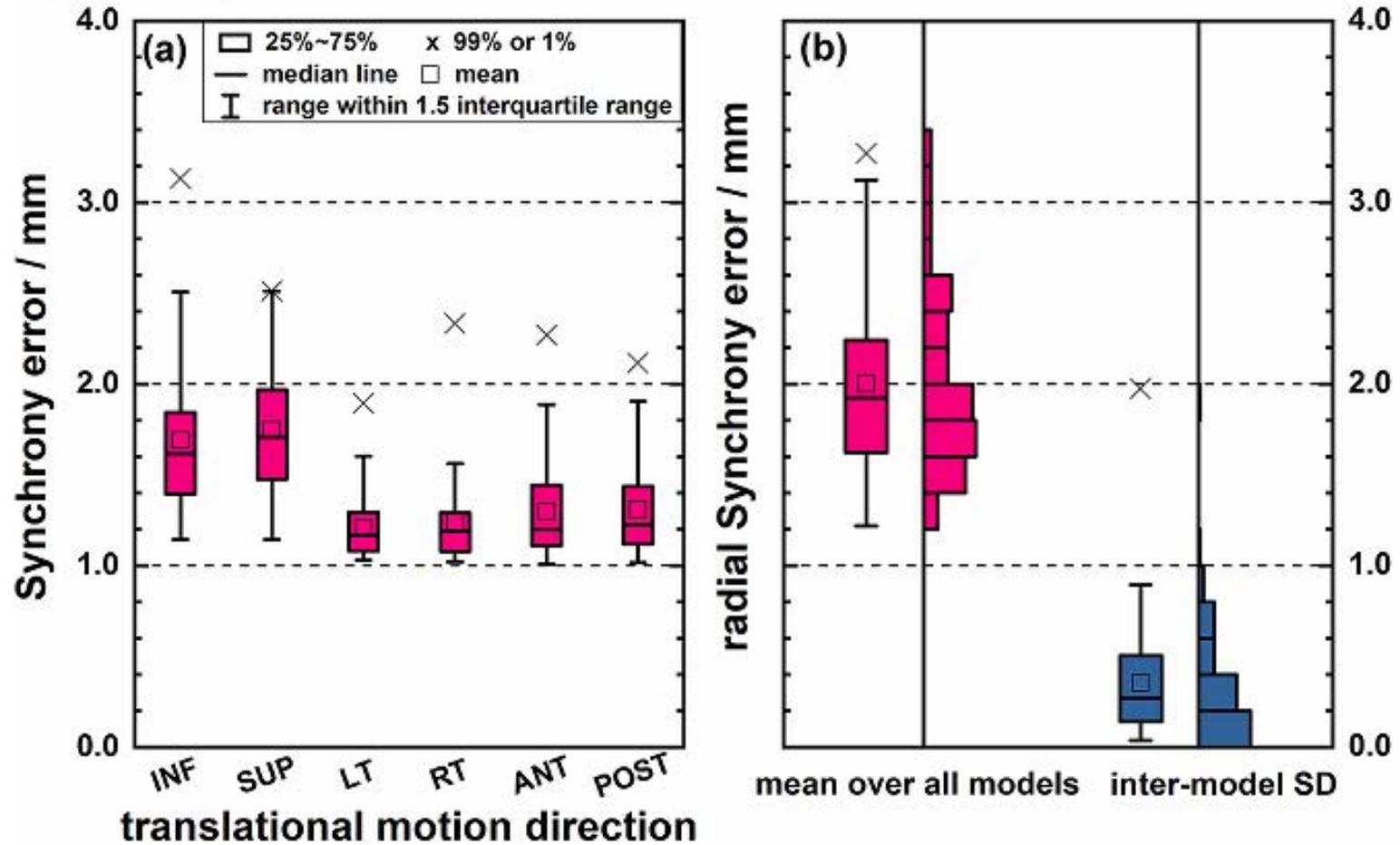


Figure 2-1. Synchrony errors and inter-model standard deviations (SD) of radial errors. Red boxplots (a, b) indicate the mean error for each patient across all their models, while the blue boxplot (b) indicates the SD across all models for each patient. (b) The length of the bar is proportional to the number of patients within each bin.

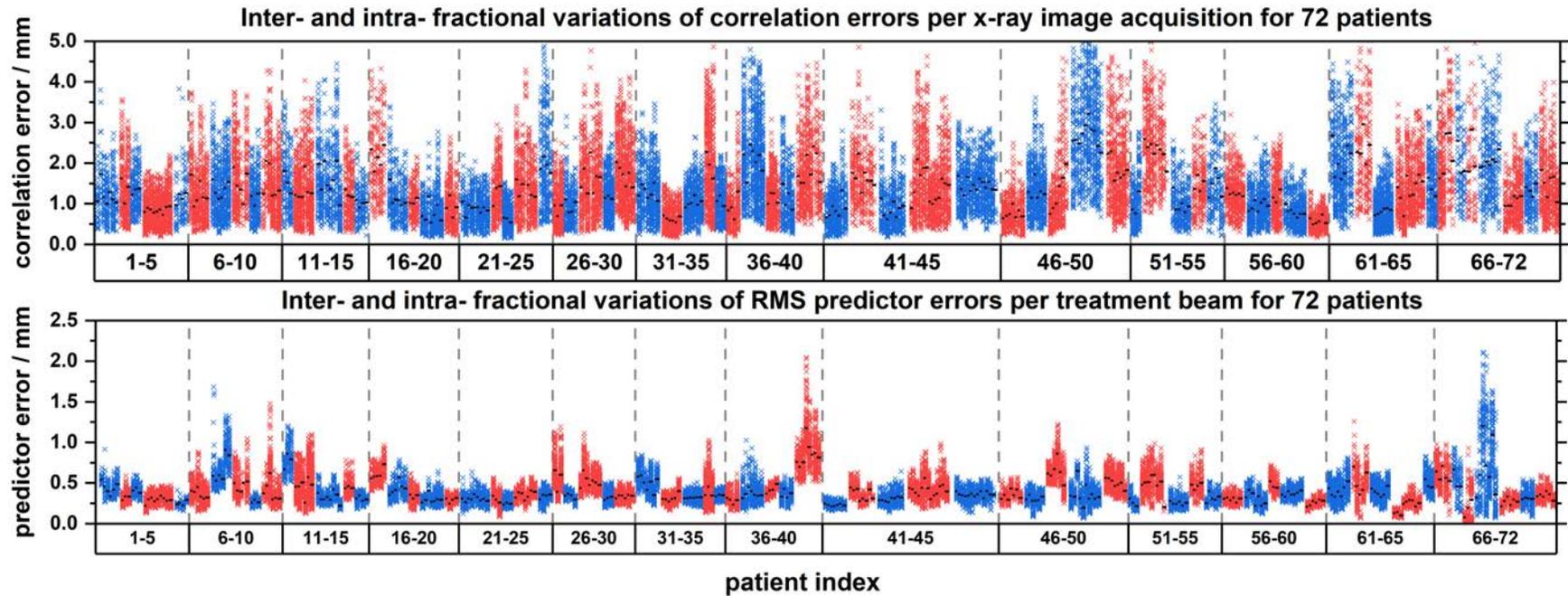


Figure 2-2. Inter- and intra- fractional variations of correlation and predictor errors for 72 patients. Radial correlation errors per X-ray image pair (top plot) and radial RMS predictor errors per treatment beam (bottom plot) are grouped by fraction for all patients. Each vertical data cluster represents the 5<sup>th</sup> to 95<sup>th</sup> percentile of the data for a treatment fraction. Clusters belonging to individuals alternate between red and blue.

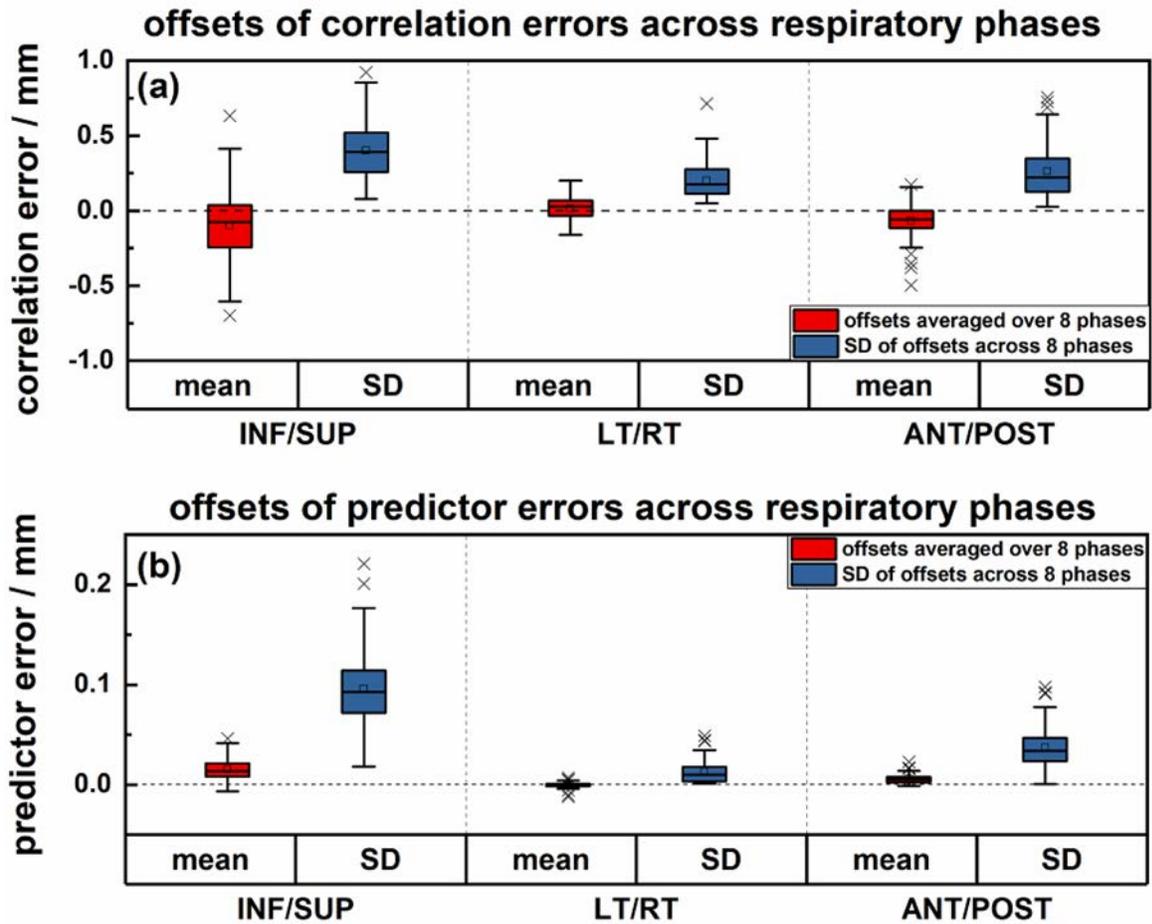


Figure 2-3. The offsets of correlation (a) and predictor errors (b) across eight respiratory phases for 72 patients.

### 2.3.2. Rotational Corrections

Geometric target rotations are an independent source of error<sup>166</sup> that should be considered in determining treatment margins, especially when rotational corrections cannot be applied. Fiducial placements for only 23 out of 72 patients (32%) meet all the criteria for accurate calculation of rotations (for a total of 84 fractions). In 18 out of these 23 patients (78%) more than three fiducials were used for tracking which is also recommended for lower uncertainty in rotation corrections.<sup>141</sup> Rotational corrections for every treatment fraction of 23 patients are shown to illustrate inter- and intra- fractional variations (Figure 2-4). Figure 2-5 shows the average and SD of rotational

corrections across all x-ray acquisitions of a treatment course, which illustrates the inter-patient and intra-patient variations of rotational corrections, respectively. The average rotational errors for the 23 patients are  $-0.1^\circ \pm 3.3^\circ$  in roll,  $-1.0^\circ \pm 3.7^\circ$  in pitch, and  $0.5^\circ \pm 3.8$  in yaw. The rotational corrections are found to be uncorrelated with Synchrony errors (Spearman  $r < .3$ ).

Only 6 (patient indices #2, 5, 12, 14, 17, and 20 in Figure 2-4) out of 23 with accurate rotational corrections have more than 70% of their observed rotational corrections within the system thresholds as shown in Figure 2-4, meaning that the system can correct rotations for only 8% of the entire population of our patients. For 26 out of 84 fractions (31%), the 5<sup>th</sup> to the 95<sup>th</sup> percentile of corrections span within  $3^\circ$ ,  $3^\circ$  and  $6^\circ$  in roll, pitch, and yaw, respectively, of their mean values (Figure 2-4). This means that if these systematic rotational corrections were avoided by setting up patients such that their rotational corrections averaged to zero across the respiratory cycle, then these rotations could be corrected by the system at least 90% of the time for these 26 fractions.

After applying the average angle of rotation of a treatment course to target(s), the changes to D95 and D99 as a fraction of the prescribed dose ( $R_x$ ) are shown in Figure 2-4 (d) for the 23 patients (31 targets) with accurate rotations. Rotations of any magnitude are assumed to be not corrected by the system and equal to the average rotation per x-ray image pair across the entire treatment course. Some targets were intentionally underdosed to spare a nearby OAR. We define compromised treatments as those with more than 10% difference between the D99 and D95 in the planned target DVH. For this type of compromised plans, the dose fall-off is fast so that the dose to target can be sensitive to different types of delivery errors. The arrows indicate three targets that were treated with compromised plans (3 out of 31). The percentage  $D95/R_x$  and  $D99/R_x$  decrease from  $105.1\% \pm 13.2\%$  to  $100.8\% \pm 17.6\%$  and from  $101.1\% \pm 14.5\%$  to  $96.7\% \pm 19.0\%$ ,

respectively, for all 31 targets. For uncompromised targets (28 out of 31), the average change to the percentage  $D_{99}/R_x$  is  $-2.7\% \pm 5.8\%$ . Excluding targets from patient 4, 18, and 21 with rotations exceeded  $\pm 5^\circ$  in at least one direction, the average change to  $D_{99}/R_x$  is  $-1.1\% \pm 2.3\%$ .

A 3D gamma analysis is used to assess changes to the planned dose distribution due to patient orientation using average observed rotations for three patients (# 1, 4 and 18 in Figure 2-4). The minimum and average 2%/2mm pass rates are 91% and 95% respectively, for all voxels with dose greater than 50% of the maximum planned dose. This suggests that the approximation of rotating the dose distribution and neglecting attenuation changes is reasonable even for large angles of rotations.

Fiducials are implanted mostly percutaneously using needles and are preferably not implanted inside the tumour to reduce the risk of the spread of tumour cells. Table 2-1 shows some statistics for fiducial placement. The Spearman correlation coefficient for the change in  $D_{99}/R_x$  and the distance between imaging isocenter to the fiducial center is -0.4. This suggests that a larger distance between the imaging isocentre and the centre of fiducials might lead to a larger dose reduction to target if rotations are not corrected.

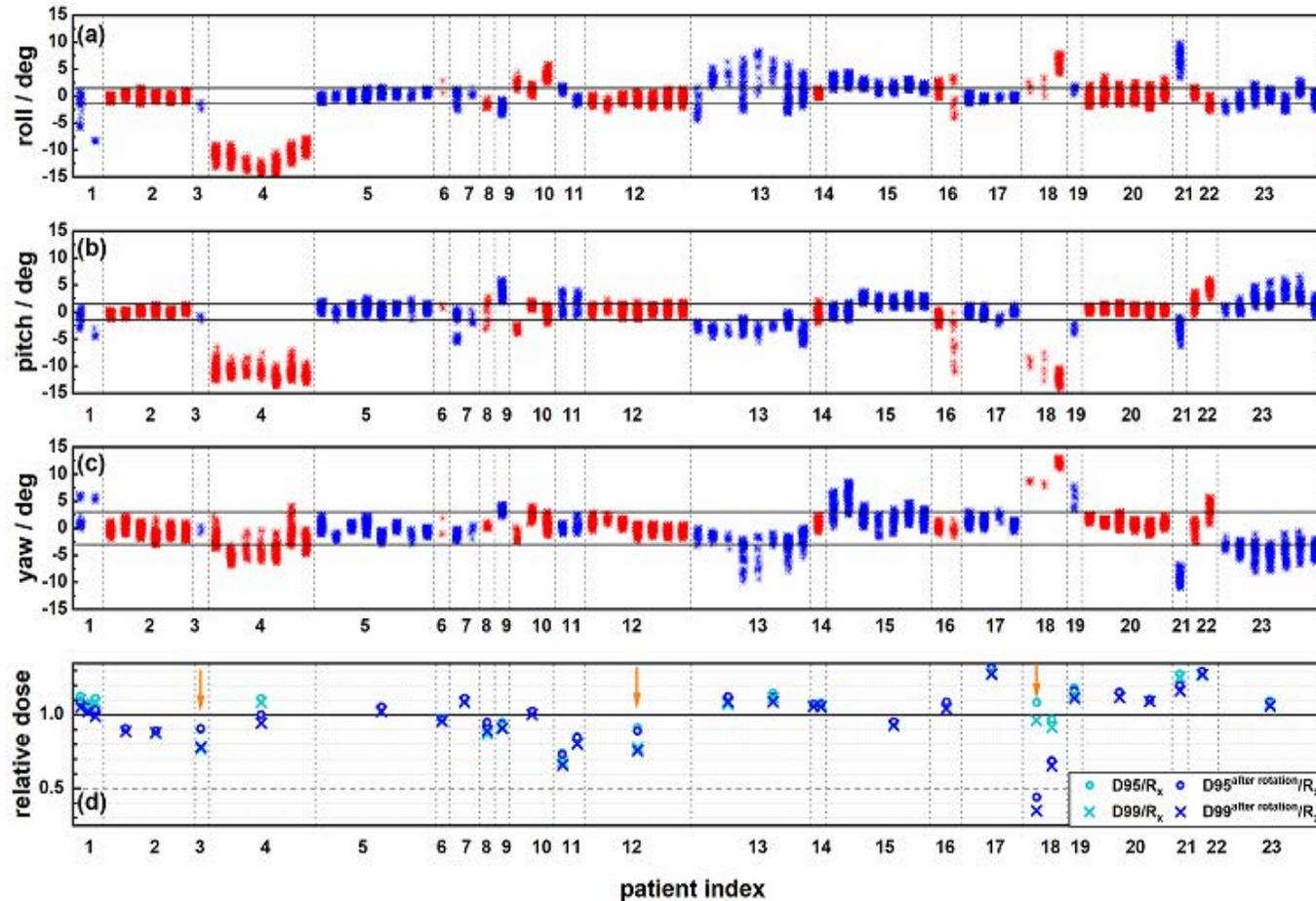


Figure 2-4. Rotational corrections grouped by treatment fraction for 23 patients. Each data cluster represents the 5<sup>th</sup> to the 95<sup>th</sup> percentile of corrections for a treatment fraction. Red and blue markers alternate for different patients. (a-c) The dashed lines indicate the boundaries of rotation for which the system can correct. Three patients with the largest rotations had two to three distinct large lesions with fiducials in the most inferior segment (VI) of the liver closest to the bowel structures, which are sensitive to radiation. (d) The changes in dose to 95% of the volume (D95) and 99% of the volume (D99), as a fraction of the prescribed dose (Rx) for 31 targets. The arrows point at targets treated with compromised plans.

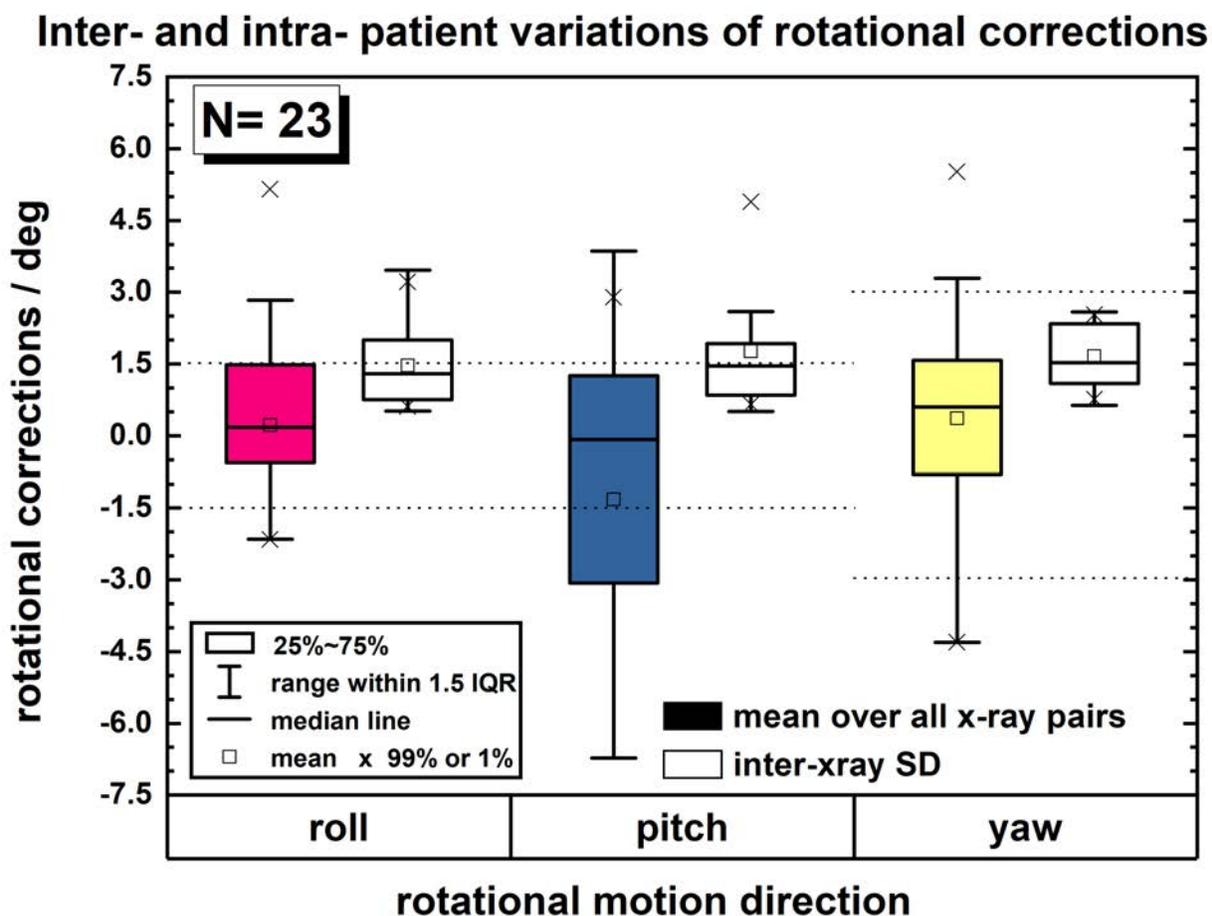


Figure 2-5. Inter- and intra- patient variations of rotational corrections for 23 patients. Mean rotational corrections across all X-rays for each patient (color boxplots) and inter-x-ray SD across all X-rays for each patient (white boxplots). The dashed lines indicate the range of rotational corrections which the system can correct.

Table 2-1. Characteristics of the distance between points of interest for fiducial placement for 72 liver treatments. The points of interest considered are the imaging isocentre, the centre of GTV, and the centre of fiducials used for tracking. The GTV-to-liver dome distance is calculated as the perpendicular distance from the centre of GTV to the most superior plane of the liver contour.

	distance between centroids of interest (cm)			
	isocenter-to-GTV	fiducial-to-GTV	isocenter-to-fiducial	GTV-to-liver dome
mean	4.6	4.1	3.5	6.3
SD	2.1	2.3	1.8	3.1

### 2.3.3. Rigid Body Error

We found that 20.4% of the 89 GTVs for the 66 patients (with rigid body reported) had at least one fiducial implanted in the GTV. The distance between the centre of a target and the center of fiducials is  $4.1 \pm 2.4$  cm for them. Figure 2-6 shows the average and SD of signed rigid body errors for all x-ray image acquisitions over a treatment course, as well as the average of the absolute value of these rigid body errors. The absolute value of the rigid body error is a quantity that can trigger treatment interrupts. It very frequently exceeds the default limit (2 mm) for our patients. Rigid body errors grouped by fractions for 66 patients are shown in Figure 2-7.

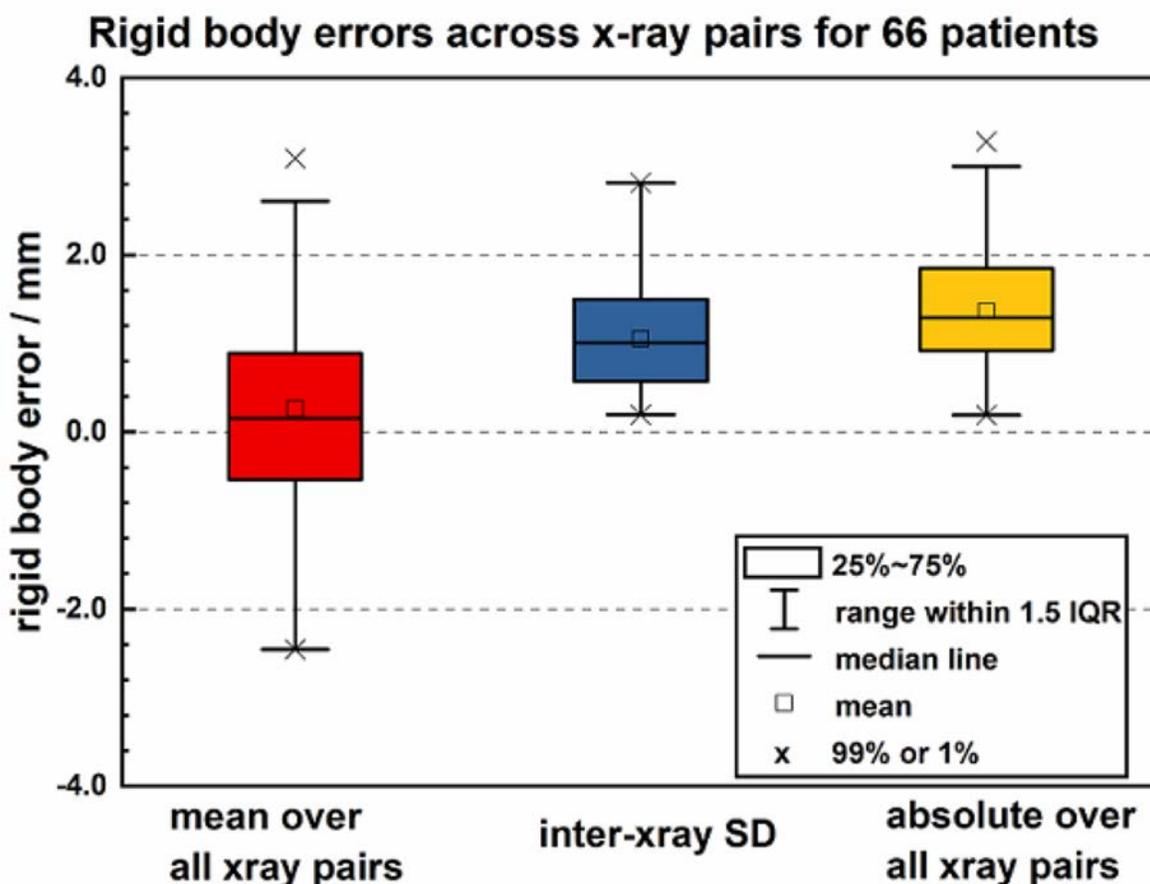


Figure 2-6. Rigid body errors for all x-ray acquisitions. The mean, SD, and absolute values of signed rigid body errors are shown as boxplots from left to right.

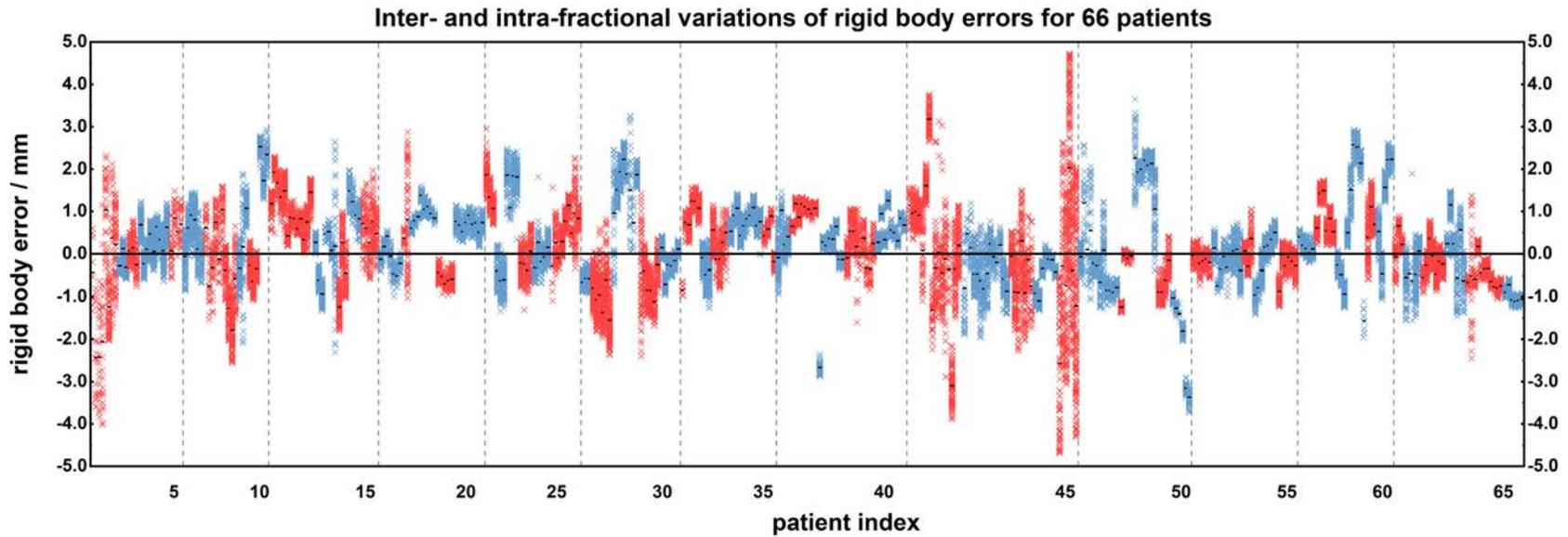


Figure 2-7. Inter- and intra- fractional variations of rigid body errors for 66 patients. Rigid body errors for every x-ray image pair are grouped by fraction for these patients. Each vertical data cluster represents the 5<sup>th</sup> to the 95<sup>th</sup> percentile of the rigid body errors for every fraction. Clusters belonging to individuals alternate between red and blue.

### 2.3.4. Proposed Uniform PTV Margin

The absolute rigid body errors are found to be uncorrelated with Synchrony errors (Spearman  $r < 0.2$ ). We propose PTV margins based on adding the absolute rigid body error and the Synchrony error quadratically, which results in 3.7, 3.3, and 3.4 mm in SUP-INF, LT-RT, and ANT-POST directions, respectively. However, considering the limited CT resolution, errors are rounded to the next integer. Therefore, 4 mm isotropic margins are sufficient to compensate for Synchrony and rigid body errors for 95% of the patient population, if the target rotations have a negligible impact on target coverage (e.g., spherical targets centred about the imaging isocenter). Rotational angles and rigid body errors, in our study, are found to be uncorrelated (Spearman  $r < 0.2$  for roll, pitch, and yaw rotational directions), which may imply that the large rotations observed for some patients are not associated with fiducial migration.

### 2.3.5. Model Linearity

A relatively strong linear correlation (Pearson  $r = .67$ ) is found between the range of target SUP-INF motions ( $17.4 \pm 7.4$  mm) and marker radial motions ( $10.0 \pm 4.7$  mm) for the 624 models but a negligible multiple correlation coefficient ( $r = .10$ )<sup>162</sup> is found among internal and external motion amplitudes, and radial Synchrony errors over all models ( $2.0 \pm 0.6$  mm). The  $R^2$  value for every model (representing model linearities or conversely the lack of hysteresis) are shown as a function of SUP-INF correlation errors in Figure 2-8. The dashed lines show the 95<sup>th</sup> percentile of SUP-INF correlation errors (2.2 mm) and 80<sup>th</sup> percentile of  $R^2$  values (0.80).

The average range of target motions in SUP-INF, LT-RT, ANT-POST directions across models for all patients are  $17.5 \pm 5.3$ ,  $4.0 \pm 1.9$ , and  $7.5 \pm 3.3$  mm, respectively, and the corresponding inter-model standard deviations are  $4.0 \pm 2.5$ ,  $1.3 \pm 0.9$ , and  $2.0 \pm 1.4$  mm,

respectively. The average range of marker radial motion across models for 72 patients is  $10.8 \pm 3.7$  mm. The corresponding inter-model standard deviation for all patients is  $2.7 \pm 1.5$  mm.

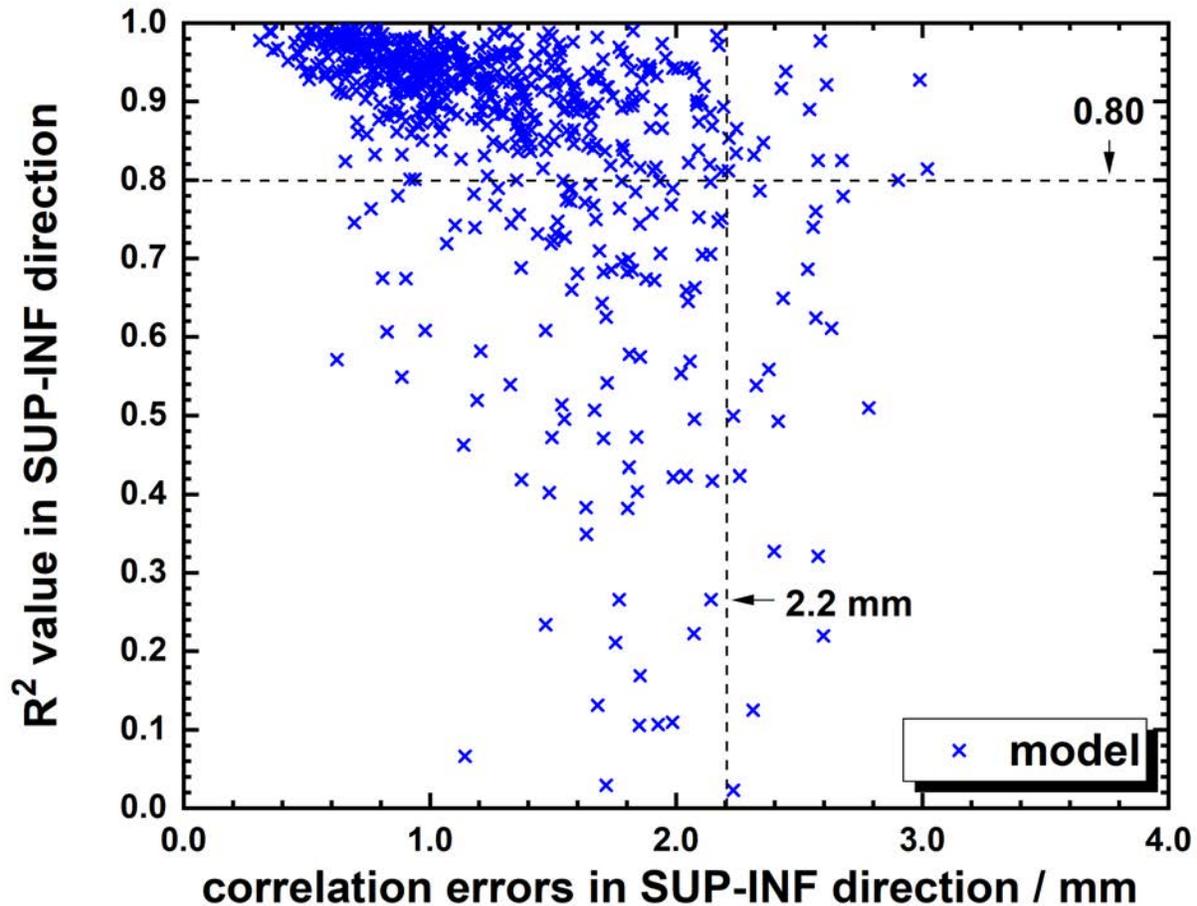


Figure 2-8. The  $R^2$  value per model between SUP-INF positions of the target and the average radial positions of external markers shown as a function of correlation errors in SUP-INF direction.

### 2.3.6. Hypothesis Testing

Following the organ contouring consensus guidelines,<sup>167</sup> a radiation oncologist identified the liver segment in which each patient had their fiducials implanted. Sixty-one patients were identified as having fiducials exclusively on either the left or right side of the liver. No significant difference in radial Synchrony errors were found for either gender (95% CI  $-0.1$  to  $+0.3$  mm,  $P = 0.3$ ) or left/right lesion location (95% CI  $-0.4$  to  $+0.1$  mm,  $P = 0.2$ ). Similarly, no significant

difference was found between the range of SUP-INF target motion (95% CI -0.7 to +4.7 mm,  $P = 0.1$ ) or rigid body errors (95% CI -0.3 to +0.5 mm,  $P = 0.7$ ) for left/right lesion location. The dome-to-target distance is  $6.3 \pm 3.1$  cm for 72 patients (98 GTVs). No correlation was found between the radial Synchrony errors and dome-to-target distance (Spearman  $r = -0.1$ ) or between rigid body errors and dome-to-target distance (Spearman  $r = -0.1$ ). The correlation between age and radial errors was negligible (Spearman  $r = 0.3$ ). We also found that the number of fiducials used for tracking did not make a significant difference in the amplitude of rigid body errors (95% CI -0.5 to +0.5 mm,  $P > 0.3$ ).

## 2.4. Limitations

Due to the partial volume effect, we propose 4 mm isotropic margins even though our data support the use of non-isotropic margins. We acknowledge that the proposal of the PTV margins based on root-mean-square averaged tracking errors that cover the 95<sup>th</sup> percentile of our patient population may be overly conservative as some geometric errors may cancel out when they are in opposite directions.<sup>142</sup>

This study does not investigate uncertainties in rotational corrections associated with using less than four fiducials,<sup>32</sup> the variation of motion in different liver segments,<sup>168</sup> larger spacing between fiducials and targets,<sup>68</sup> or changes in liver rotational orientation across the respiratory cycle.<sup>142</sup> Due to the small number of patients ( $N=23$ ) with implanted fiducials that meet the criteria for accurate calculation of rotation corrections, the results may be biased in some way. For example, patients who have fiducials implanted that meet these criteria may have other common characteristics, such as lesion location, the number and size of lesions, and health status. The impact of having an uncorrected large rotational offset during treatments is further studied in Section 5.1, in which more treatment data with rotations tracked by the system is available. Our

method to estimate the dose change to targets due to uncorrected rotations treats the patients as rigid bodies and neglects the change in beam attenuation. The effect of changing beam attenuation is only investigated for three cases. The PTV expansion used to assess uncorrected rotations was typically a 5 mm isotropic margin, slightly larger than the 4 mm margin that we propose here.

Rigid body errors may also be an overly simplified way to characterize the margin required for organ deformation, especially for complicated target volumes and when there are larger distances between the lesion and the fiducials. The dose increases to OARs (stomach, duodenum, the rest of the small bowel, and heart) are not discussed, which might be another important concern when there are uncorrected rotations or organ deformation.

## 2.5. Discussions

### 2.5.1. Synchrony Error

For 95% of patients, margins of 3, 2 and 2 mm in SUP-INF, LT-RT and ANT-POST directions, respectively, are sufficient to account for Synchrony errors (see Figure 2-1). Winter *et al.*<sup>76</sup> gave quantitative estimates of the total Synchrony error averaged over all treatment fractions for liver treatments on CyberKnife (N=27). The correlation errors were selected for data immediately before the beam delivery. Predictor errors were considered within a 20-second window before the corresponding x-ray image pair. They reported the 95<sup>th</sup> percentile of absolute correlation errors in each direction as 3.3, 2.1, and 1.8 mm for the SUP-INF, LT-RT, and ANT-POST directions, respectively, and an overall radial PTV margin of 4 mm. Although our work considers the RMS error for every model, instead of mean errors for every fraction, our results agree well with Winter *et al.*'s data.

Pepin *et al.*<sup>147</sup> extracted different percentiles of residual errors from 100 Synchrony models for 23 lung patients. The correlation, predictor, and targeting errors were added linearly to estimate

non-isotropic margins of 6.9 mm in the SUP-INF direction, 4.6 mm in the ANT-POST direction, and 3.5 mm in the LT-RT direction. However, they used a more stringent requirement on 95% of model points, rather than the 95% of patients used in our study, and they included all data from the entire treatment. In our study, the variations of radial Synchrony errors over treatment models for individuals are relatively small (mostly  $<1$  mm), as shown in Figure 2-1b. Figure 2-2 illustrates that distributions of correlation and predictor errors for each patient are similar from fraction to fraction, indicating that the system tracks the target motion in a relatively consistent way for an individual. These observations indicate that patients tend to have higher or lower Synchrony errors consistently throughout the treatment course. Hence, patient-specific margins are a possibility for these treatments, perhaps based on the data from the first treatment fraction or a dry-run session prior to treatment planning, similar to what proposed by Pepin *et al* for lung patients.<sup>147</sup>

Using Van Herk's margin recipe with the radial Synchrony error as residual setup error for these patients, a 2 mm PTV margin is needed for 90% of the patient population to ensure the CTV receives 95% of the prescription dose. However, that margin recipe does not take into account target deformation or rotation, makes specific assumptions about the dose penumbra, and it was proposed to be used for highly fractionated treatments and spherical targets. The dosimetric evaluation of a selected PTV margin is investigated using an in-house dose perturbation tool developed in our later study (see Section 4.2.4).

For the phase dependence of the correlation and predictor errors, the systematic effects across all respiratory phases approximately cancel out within  $\pm 0.5$  mm and  $\pm 0.05$  mm, respectively, for most patients in any direction, as shown in Figure 2-3. However, the SD across the respiration phases for individuals is not negligible. The dose blurring effect can get worse when the motion is out of phase, and it leads to less sharp penumbra.<sup>169</sup> This finding confirms our relatively

conservative approach of using unsigned errors to estimate PTV margins that take the dose blurring effects into consideration. Chan *et al.*,<sup>142</sup> using ten phases over the respiratory cycle, pointed out that the correlation errors were not statistically phase-dependent ( $P > .05$ ), which is consistent with our work.

### 2.5.2. Rotational Corrections

Due to the limitation in magnitudes of corrections the Synchrony system can apply, only six out of 23 patients could be treated with most of their rotations corrected. Ideally, fiducial placements should be improved so that rotations can be tracked more frequently. This could be done using a kV X-ray c-arm, CT, or MR imaging system, which would allow the fiducial positions to be validated during the implant procedure using oblique planar images. The Synchrony system should be modified to correct for larger rotations. For example, the system can correct for up to  $\pm 5^\circ$  for prostate treatments,<sup>170,171</sup> which is beneficial for the patients. Errors could also be reduced if the patients were set up such that rotational corrections average to zero across the respiratory cycle. The user's confidence in these rotations could be improved with on-board volumetric imaging planned to be included in a future release of the CyberKnife system.

The rotational corrections in all directions are non-negligible and may contribute to systematic beam delivery errors if their effects are not minimized. The relatively large inter-patient variations (Figure 2-5) indicate that those patients are set up with rotational errors systematically different from each other. The intra-fractional variations (Figure 2-4) indicate that non-negligible rotations also occur from x-ray to x-ray, over timescales comparable to the respiratory cycle. This can be even more problematic for the remaining 49 patients whose rotational corrections are unknown. The dosimetric change due to systematic rotations has been assessed for this small group of patients with accurate corrections (N=23). Since more than 4% of reduction in D95/Rx and D99/Rx

has been observed for their treatments, in any case, the imaging isocenter should be placed as close as possible to the center of fiducials to minimize dose change introduced by uncorrected rotations.

Chan *et al.*<sup>142</sup> approximated the cumulative dosimetric effects due to residual correlation, predictor, and rotational errors using data binned into ten breathing phases for 24 liver patients treated with CyberKnife. They reported relatively small systematic rotational corrections of  $-0.26^\circ \pm 0.51^\circ$ ,  $0.14^\circ \pm 0.71^\circ$ ,  $0.65^\circ \pm 0.61^\circ$  in roll, pitch, and yaw directions, respectively. The rotations observed for our patient population (N =23) deviate more from zero and have a larger SD than those reported by Chan *et al.*<sup>142</sup> However, some of the aggregate errors presented in that work were averaged over each bin of the respiratory cycle and over the total number of fractions, so our results cannot be compared directly. We present intra- and inter-fractional rotation corrections for each x-ray pair acquisition, so it is perhaps not surprising that we have larger SDs across the patient population. Chan *et al.*<sup>142</sup> appear having more success setting up patients with rotational corrections close to zero than in our experience. However, there are some similarities between our observations. Chan *et al.* also observed larger rotations for a subset of their patients, remarking that 23% of patients had rotations  $> 2^\circ$ . Also, their fiducial implants did not always meet the criteria for accurate rotation corrections, with only 24 of 46 patients meeting their criteria.

Chan *et al.*<sup>142</sup> estimated the dosimetric effect of correlation, predictor, and uncorrected rotation errors using a “pseudo-4D” dose calculation method. They considered changes to CTV and PTV coverage and investigated if a 3 mm isotropic CTV-to-PTV expansion was sufficient to cover all residual tracking errors. They rotated and translated the contours for the CTV and PTV based on the aggregate translational and rotational errors for several bins of the respiratory cycle, each equally weighted in the final dose accumulation. They found a median dose reduction of 1.1% and 0.0% for PTV and CTV, respectively, for their patients. Based on these results and other concerns,

including the increased dose to OARs for patients with small residual errors, they argued against larger and non-isotropic margins. However, changes to coverage based on rotating contours, both in Chan's work and in ours, may neglect the dose broadening effects and the associated increase in penumbra due to tracking errors.<sup>27,172</sup> These errors may affect different parts of the target differently due to the large numbers of small fields used by the CyberKnife system and when errors with opposite signs occur at different extremes of the respiratory cycle. Further study is warranted to assess more accurately the dosimetric changes due to Synchrony and rigid body errors as well as uncorrected rotations for these treatments.

Xu *et al.*<sup>140</sup> calculated translational and rotational corrections for 23 liver patients by comparing live x-ray images, and deformations relative to one reference x-ray image pair near the beginning of the treatment. They found that absolute rotational angles of roll, pitch, and yaw were  $1.2^\circ \pm 1.8^\circ$ ,  $1.8^\circ \pm 2.4^\circ$ , and  $1.7^\circ \pm 2.1^\circ$ , respectively, the SD of which were more consistent with our findings.

### **2.5.3. Rigid Body Error**

The intra-patient variations in rigid body errors, presumably associated with the respiratory cycle, are relatively small ( $0.6 \pm 0.4$  mm), as shown in Figure 2-6, suggesting that rigid body errors for most individuals are relatively constant over their treatment courses. This could be applied to adaptive liver treatments. For example, if a patient had positive rigid body errors (i.e., target expansion) during the first treatment fraction, a margin expansion could be considered for future treatments.

Xu *et al.*<sup>140</sup> concluded that intra-fractional fiducial migration was unlikely based on small 3D fiducial displacements ( $0.2 \pm 0.2$  mm) when the fiducials were aligned rigidly. Furthermore, only small systematic volumetric increases (1-2%) were observed in that study in subsequent fractions

relative to the first fraction. They believed it might be related to radiation-induced edema. In our study, rotational angles and rigid body errors are not found to be strongly corrected, implying that the large rotations observed are not associated with fiducial migration. However, organ deformation may contribute to biases in the rotational corrections included in our analysis, since the changes in fiducials' spacing may lead to the exclusion of certain data if the criteria for accurate rotation corrections are not met.

#### **2.5.4. Model Linearity**

Target motion amplitudes were found to be strongly correlated with predictor errors but were weakly correlated with correlation errors.<sup>76</sup> In our study, using the range of SUP-INF target motion and radial marker motion cannot predict the size of Synchrony errors when the coefficient is negligible (multiple correlation coefficient  $r = 0.10$ ). However, most models correlating internal and external marker motions are nearly linear, which can be applied to other types of treatment modalities for liver treatments.

The results and methodology concerning target motion amplitudes for liver lesions vary a lot among research groups (see review by Langen *et al.*<sup>173</sup>). Gierga *et al.*<sup>174</sup> found that the internal liver tumour motion and external marker motions are well correlated with the maximum range of SUP-INF motion of 25 mm. This result is consistent with our data although only four liver patients were assessed in their work. Liang *et al.*<sup>33</sup>'s log file study of SBRT liver patients (N=14) found the range of liver motions (5% cut off of the inspiration and expiration positions) for their patient population were 5.1-17.3, 0.4-4, and 0.9-7.7 mm in the SUP-INF, LT-RT and ANT-POST directions, respectively. Their results are similar to our findings of  $17.5 \pm 5.3$ ,  $4.0 \pm 1.9$  and  $7.5 \pm 3.3$  mm in the SUP-INF, LT-RT and ANT-POST directions, respectively. The large amplitudes of target motion suggest that motion management, such as abdominal compression, gating or active

breathing control should be used if these treatments are delivered using conventional LINAC systems.

## **2.6. Conclusions**

Our study proposes 4 mm isotropic PTV margins that encompass Synchrony errors and rigid body errors for liver treatments on CyberKnife. The small intra-patient variations in both Synchrony errors and rigid body errors suggest that customized margins for individuals may be a possibility, perhaps based on a dry-run session prior to planning. More patient data (with accurate system calculated rotational corrections) are required to better estimate the impact of uncorrected rotations. However, this error can be minimized with improved fiducial placement and by taking more care during CT simulation to ensure that the imaging isocenter placed as close as possible to the center of fiducials. Linear models correlating internal and external marker motions might be applied to other treatment modalities for liver radiotherapy.

## Chapter 3. Patient-Specific PTV Margins

### 3.1. Introduction

Target motion and deformation during treatments should be taken into consideration for treatment planning and delivery.<sup>175,176</sup> Statistical analyses,<sup>177-181</sup> machine learning,<sup>133-139,182-186</sup> and deep learning<sup>187</sup> using image and post-treatment delivery system log file data are increasingly being reported on for patient-specific and adaptive radiation treatments. However, there are few applications of machine learning/deep learning for patient-specific treatments using stereotactic body radiation therapy (SBRT) for liver lesions.<sup>187,188</sup>

Radiation-induced complications in bowel structures, such as ulceration, obstruction, and fistula, are especially of concern in hypofractionated treatments.<sup>9,123</sup> Reducing planning target volume (PTV) margins can be beneficial for SBRT treatments, since higher dose could be delivered while sparing high dose to surrounding healthy tissues. This in turn could improve local tumor control and reduce treatment toxicity. Velec *et al.*<sup>109</sup> created “*iso-toxic*” plans for SBRT liver treatments using dose-probability-based PTVs, which were smaller than the internal target volume (ITV) based PTVs. They found that plans using dose-probability-based PTV margins enabled a dose escalation of 4 Gy on average over six fractions. It has been postulated that patient-specific PTV margins are needed for SBRT liver treatments with motion tracking.<sup>142,181,189</sup> However, adjusting PTV margins requires an estimate of motion-tracking errors for the patient in advance. In Section 2.3.1, using data for 72 liver patients treated with robotic SBRT, we showed that inter-fractional standard deviations for motion-tracking errors were small (< 1 mm in any anatomical direction). We found that the standard 5 mm PTV margin greatly exceeded the observed motion-tracking errors consistently throughout the treatment courses for some patients. Inoue *et al.*<sup>190</sup> extracted the respiratory motion data from cine MR scans for lung patients (N = 91)

treated with SBRT and simulated the motion trajectories using a dynamic phantom. They calculated the motion-tracking errors and proposed patient-specific PTV margins based on features extracted from the MR images that were highly related to motion-tracking accuracy.

Some studies have investigated the use of adaptive margins for liver SBRT based on patient-specific response following treatment or specific toxicity risks evaluated using biomarkers,<sup>191</sup> the planned dose distributions,<sup>187</sup> and MRI imaging data.<sup>192,193</sup> Boda-Heggenmann *et al.*<sup>181</sup> proposed an *in vivo* method to assess the accuracy of dose delivery for gantry based linac (deep inspiration breath-hold gating with ultrasound image guidance) and robotic radiosurgery (Synchrony MTS) evaluated for 40 liver SBRT targets. However, they did not directly measure delivery accuracy; rather they correlated morphologic changes in post-treatment MRI images with planned isodose volumes using deformable image registration. Their data supported the use of patient-specific safety-margins for certain patients. They also concluded that a 5 mm PTV margin could compensate for the observed random error components, which was similar to our previous findings in Section 2.3.4.

To our knowledge, there has been no literature report on post-treatment delivery log file data to facilitate patient-specific PTV margins for SBRT liver treatments with motion tracking. This study aims to identify patients who could be assigned smaller PTV margins for robotic SBRT liver treatments, based on a small subset of CyberKnife® (Accuray Inc., Sunnyvale, CA) system log file data retrospectively emulating a dry-run session (mock) prior to planning. The support vector machine (SVM), a supervised machine learning algorithm<sup>194</sup> is employed in our study to predict whether the maximum root-mean-square (RMS) radial motion-tracking error for all treatment fractions will be above or below a predefined threshold to customize the PTV margin for a treatment course.

In Section 2.3.4, we found that 4 mm isotropic PTV margins were sufficient to encompass the motion-tracking residual errors (referred to as Synchrony error in Chapter 2) as well as target deformations for 95% of the patient population. However, the standard deviations across different motion tracking models for the same patient were relatively small and many patients always had RMS tracking errors per model below 2 mm. In this study, we explore using SVC to predict which patients will have maximum RMS motion tracking errors per model less than 2 mm for all subsequent treatments based on dry-run session data. In which case, we consider a 2 mm margin reduction (half of the 4 mm margin required for the entire population) to lower the normal tissue radiation exposure. We also compare this binary classification scheme with a regression model for predicting the maximum motion tracking error. Since some patients may be misclassified, we propose an early-warning system which performs a per fraction assessment of the tracking errors based on the volumetric coverage of the gross target volume (GTV) within the PTV. One hundred forty-eight liver treatment courses are studied to investigate the feasibility of implementing patient-specific PTV margins assisted by the early-warning system.<sup>†</sup>

### 3.2. Methods and Materials

#### 3.2.1. Data Collection and Analyses

Patients in this study were treated for liver lesions (metastases/primaries: 70%/30%) between 2010 and 2018 using the CyberKnife with the Synchrony™ Respiratory Motion Tracking System (MTS),<sup>71</sup> with the planning system v.3.5-v.5.2 and delivery system v.9.0-v.10.5. All patient data were collected as part of a local clinical trial approved by the Hospital Research Ethics Board.

As mentioned previously, either 5 mm or 7 mm isotropic PTV margins expanded from GTVs are used for treating liver SBRT on CyberKnife at our institution.<sup>119</sup> In this chapter, the standard

---

<sup>†</sup> The results in this Chapter have been published in a peer-reviewed journal

PTV margins (either 5 or 7 mm), which were used for planning, are denoted as  $PTV_{std}$ . Typically, a dose of 42-60 Gy given in three to six fractions per treatment course is prescribed by a radiation oncologist, depending on the size and location of the tumor(s) in the liver. To preserve adequate liver function, our institutional treatment planning guidelines require that at least 700 cc of the healthy liver (liver volume excluding GTV and radiofrequency ablation volumes if present) receive dose  $< 15$  Gy.<sup>128</sup> When the PTV is adjacent to the OARs, the prescription dose is often reduced to meet the dose constraints for OARs.

Dosimetric and volumetric analyses are performed for the PTV, GTV, the healthy liver, and serial OARs, including bowel structures, stomach, spinal cord, and inferior vena cava. We extract a binary three-dimensional (3D) mask from the contours (OARs, PTVs, and GTVs) and the planned 3D dose matrix from exported DICOM files. We use in-house MATLAB scripts (v.9.5.0, MathWorks Inc., Natick, MA) to isotropically expand or contract the contoured 3D masks. A 5 mm planning risk volume (PRV) is created by expanding each OAR, and an alternative PTV which is 2 mm smaller than  $PTV_{std}$  (denoted as  $PTV_{std-2mm}$ ) are generated. We calculate dose-volume histograms (DVHs) for  $PTV_{std}$  and the volume of overlap between PRV and PTV (denoted as  $PTV \cap PRV$ ).

### 3.2.2. Synchrony Respiratory Tracking System

Each fraction of treatment consists of one or more Synchrony models (henceforth referred to as the ‘models’) created by the MTS.<sup>71</sup> Refer to Section 1.3.2 for the details of the MTS. The Synchrony error the square root of a sum of the motion-tracking residual error (correlation error), the error compensating for system latency in beam positioning (predictor error), and the global beam positioning error (end-to-end error) with data filtering as described in Section 2.2.2. As in Section 2.2.1, the CT image resolution used for these treatments is typically  $\leq 1.0 \times 1.0 \times 1.0$  mm<sup>3</sup>

and margin expansion or contraction is rounded up to the nearest mm to consider partial volume effects.

### **3.2.3. Margin Reduction Methods**

For our binary classification scheme, we propose to reduce PTV margins by 2 mm, only when a patient's maximum per model RMS Synchrony errors are predicted to be less than 2 mm. The selection of this threshold allows us to have more balanced classes, an important consideration for machine learning applications with relatively small sample sizes. Given the numerical precision of the reported Synchrony errors, we could in principle use a continuous margin reduction based on a regression analysis. However, in the CyberKnife planning system the margin expansion is discretized by the dose matrix voxel size (typically  $1.0 \times 1.0 \times 1.0 \text{ mm}^3$  at our centre). In practice, this means that we would still have discrete possible categories for margin reduction, for example 1, 2, and 3 mm given a PTV margin of 5 mm, and there may not be enough samples from each class.<sup>133,136,185</sup> Therefore, we round the continuous tracking error prediction for the margin reduction. We considered two possible margin reduction formulae (see Figure 3-5) based on the regression analysis, either the predicted Synchrony error (rounded up to the nearest integer) subtracted from 5 mm (more aggressive, denoted as “margin = 5 mm - pred”) or from 4 mm (more conservative, denoted as “margin = 4 mm - pred”, where “pred” indicates the rounded predicted Synchrony error).

### **3.2.4. Proposed Clinical Workflow**

We propose that a patient is given a mock session without therapeutic dose delivery prior to planning and that a model, based on 15-20 x-ray acquisitions taken over approximately 15 minutes, is generated. In this retrospective study, the first model generated by the MTS of the first fraction of a treatment course emulates a mock session. We chose to estimate the maximum RMS error per

model, instead of the average error across all models, as reported in Chapter 2, to be more conservative and reduce the risk of underestimating the margins. The features derived from delivery system log files are used to train an SVM to predict if the maximum RMS radial Synchrony error is below 2 mm for all remaining models. Each treatment course is labeled as Class 1 if the maximum RMS radial Synchrony error is  $< 2$  mm for all remaining models and Class 2 otherwise. For predicted Class 1,  $PTV_{std-2mm}$  could be assigned for planning; otherwise,  $PTV_{std}$  will be maintained.

Since we cannot evaluate the maximum RMS radial Synchrony error until the end of a patient's treatment course, we propose to have an early warning system for patients predicted to belong to class 1 and therefore having plans created with the smaller margin ( $PTV_{std-2mm}$ ). To estimate if GTV coverage is appropriate for the PTV for each fraction, a MATLAB script is used to perturb the GTV by the 3D Synchrony errors for each x-ray image acquisition. We compute the fractional GTV voxel occupancy probability ( $P_i$ ) within the smaller and larger PTV averaged across all Synchrony errors for each treatment fraction  $i$ . See Appendix C for how to calculate  $P_i$  using delivery errors that are sampled with different frequencies.

We selected a conservative threshold of 0.99 for  $P_i$  to ensure adequate coverage of the target based on the analysis of the entire patient population using the standard PTV margins (discussed in Section 3.3.4). The early-warning system flags treatment courses if  $P_i$  is smaller than 0.99, indicating that replanning using  $PTV_{std}$  may be required for subsequent fractions of treatment, as illustrated in Figure 3-1.

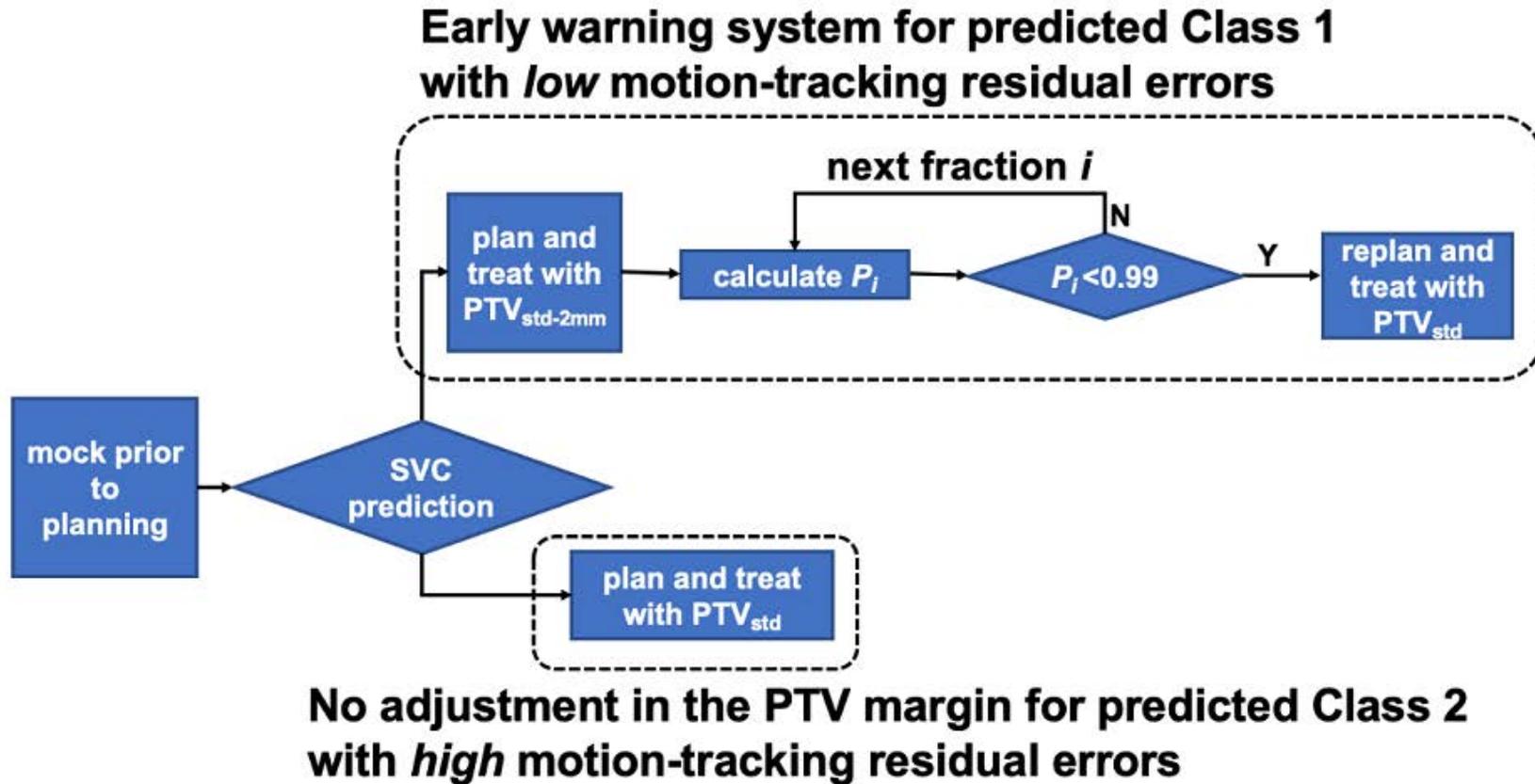


Figure 3-1. The proposed clinical workflow for patient-specific PTV margins: support vector classification (SVC) is used to predict if the patient should be planned with smaller ( $PTV_{std-2mm}$ ) or the standard ( $PTV_{std}$ ) margins, respectively. For patients planned with  $PTV_{std-2mm}$ , replanning is required using  $PTV_{std}$  when the GTV voxel occupancy probability ( $P_i$ ) is below 0.99 for any subsequent treatment fraction  $i$ . In these cases, the dose delivered to the periphery of PTV can be increased for subsequent fractions to compensate for the potential geographical miss.

### 3.2.5. Features and Binary Labels

Twenty features from the first model of each treatment course are calculated in this study. Eleven of these features are calculated based on quantities determined for each x-ray acquisition: (features 1-3)  $R^2$  values of the target positions in three anatomical directions (SUP/INF, LT/RT, and ANT/POST) versus average radial marker positions; (4) the RMS radial Synchrony error, and (5-7) the standard deviations of Synchrony errors in three anatomical directions; (8-10) target motion range in three anatomical directions and (11) the average radial marker motion range. The rest of the features are related to external marker trajectories in an arbitrary coordinate system (see Appendix A), representing the breathing patterns. The first two principal components (PC 1 and PC 2) of marker trajectories are calculated using singular value decomposition, and PC 1 and PC 2 represent marker motion and marker drifting (illustrated in Figure 3-2), respectively. The ranges (features 12-13) and standard deviations (14-15) of the first two principal components are also calculated to characterize the marker trajectories. The drift rate (16) is defined as the square root of the sum of the square of the slopes determined using linear regression for all three principal components versus time. The total drift (17) is the drift rate multiplied by the duration of the model. After applying a fast Fourier transform to PC 1, we calculate feature (18), the peak frequency ( $f_{\text{peak}}$ ) of marker motion within 0.1 ~ 1.0 Hz,<sup>195</sup> the signal strength (19) defined as the definite integral between  $0.75f_{\text{peak}}$  and  $1.25f_{\text{peak}}$ , and the signal-to-noise ratio (20) where the noise is defined as the integral outside the signal region (in the range 0.1 ~ 1.0 Hz).

Each treatment course is assigned to one of two classes based on whether the maximum RMS radial Synchrony error is above or below 2 mm. The sizes of Class 1 (< 2 mm) and Class 2 ( $\geq$  2 mm) are slightly imbalanced with 41% Class 1 and 59% Class 2.

To assess the need for employing a machine learning algorithm for this application, we investigate if we could use the average and standard deviation of radial Synchrony errors in the mock session as the only two quantities to determine the class label in a two-feature analysis.

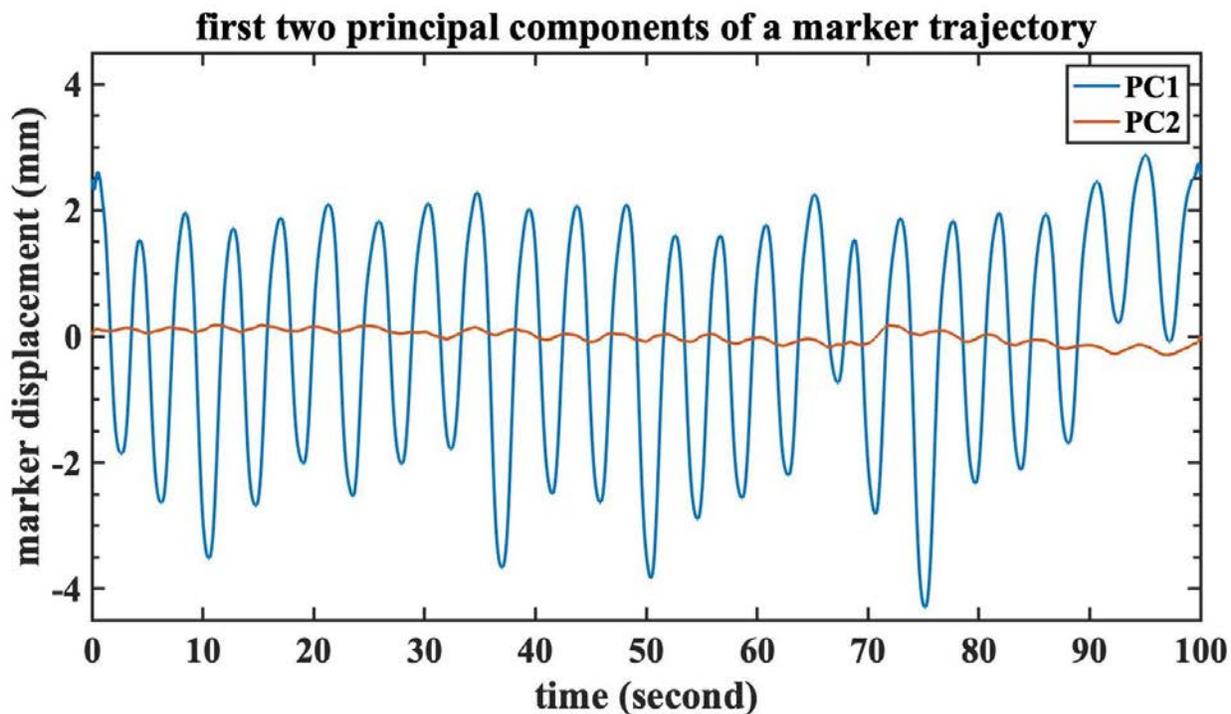


Figure 3-2. A demonstration of the marker displacement during treatment. The first and second principal components of the marker position are shown in blue and orange, respectively, after removing the high frequency components using fast Fourier transfer.

### 3.2.6. Support Vector Classification (SVC)

SVC is commonly used in binary classification with various kernels.<sup>183,185,194,196</sup> In-house python scripts (v.3.7.0) with scikit-learn python packages (v.0.20.3)<sup>196</sup> are used for SVC in this study. Ideally, an optimal linear hyperplane with boundary limits can entirely separate the feature space of two classes. In reality, a subset of data called support vectors, can lie between the boundary limits or even cross the hyperplane to the other class.

To determine the decision function defining the hyperplane, we need to maximize the width between the boundary limits and minimize the sum of the distances between the hyperplane and

the support vectors. A user-defined parameter,  $C$ , restricts the magnitude of penalization.<sup>185,197</sup> A smaller  $C$  value might lead to a smoother decision boundary but more misclassifications.<sup>196</sup> For imbalanced binary classes, the class with a smaller sample size can be given a higher relative class weight to penalize misclassifications more heavily.<sup>196</sup> The kernel functions can map the original features into a higher-dimensional feature space, and hence, non-linear decision boundaries can be obtained using non-linear kernels. The linear kernel and non-linear kernels, including the polynomial and exponential radial basis function (RBF) kernels, are used in this study.<sup>196</sup>

The entire dataset is split into the training and the testing sets. The decision function is determined using the training set, and the classification accuracy is validated using the testing set. All the features are rescaled using  $(x-\mu)/\sigma$ , where  $x$  is the feature,  $\mu$  and  $\sigma$  are the mean and standard deviation, respectively, of that feature across the entire training set. Principal component analysis (PCA) is then used to reduce dimensionality.<sup>138,184</sup> To avoid leaking statistics from the testing set into the training model, a pipeline composite estimator<sup>196</sup> is used, which combines the rescaling function, the PCA transformation, and the SVM classifier, and is applied independently to the training and testing sets.

### **3.2.7. Parameter Tuning and Feature Selection**

Five-fold cross-validation is used for parameter tuning and to reduce the number of features. The data is randomly split into five sets, and each is separately assigned as the testing set in different iterations.<sup>196</sup> The accuracy of the testing set indicates the probability of correct predictions for both classes. In this study, parameter tuning is performed using a grid search to exhaustively loop over all possible combinations of the number of PCs (1-20), the  $C$  values (0.01, 0.1, 1, 10, and 100), the relative class weights (Class 1: Class 2 = 1:0.8, 1:0.9 and 1:1), and the kernel type (linear, polynomial, and RBF). A fraction of the features are removed using recursive

feature elimination.<sup>198</sup> The optimal parameters for classification are those which result in the maximum accuracy using 5-fold cross-validation. We score the importance of features by the decrease in accuracy for each feature that is removed from classification.

### 3.2.8. Classifier Evaluation

After parameter tuning and feature selection, we considered four different methods to evaluate the classifier performance. The first method used to evaluate the classifier is leave-one-out cross-validation. In this process, one treatment course is assigned as the testing set, and the rest of the data is used as the training set. Leave-one-out cross-validation has the advantage that it allows for all the treatment courses to be uniquely sorted into two groups based on the predicted class labels.

The second evaluation is based on the confusion matrix and associated metrics,<sup>199</sup> using 500-repeated random sub-sampling cross-validation. For each of the 500 iterations, 80% of the data is selected randomly as the training set, and the remaining data is the testing set. The confusion matrix consists of the number of true positives (TPs), false positives (FPs), true negatives (TNs), and false negatives (FNs). Higher Synchrony respiratory motion-tracking error (Class 2) is designated as “positive” in this binary classification. The precision ( $\frac{TP}{TP+FP}$ ), sensitivity ( $\frac{TP}{TP+FN}$ ), specificity ( $\frac{TN}{TN+FP}$ ), and F1 score ( $\frac{2}{\text{precision}^{-1}+\text{sensitivity}^{-1}}$ ) metrics are also calculated.<sup>200</sup> High precision indicates less false alarms for Class 2, while a high sensitivity means we can identify as many Class 2 as possible. The F1 score gives equal weights to both precision and sensitivity.

The third evaluation considered uses receiver operating characteristic (ROC) metric with 5-fold cross-validation. The performance of the binary classification is quantified by plotting the true positive rate (sensitivity) as a function of the false positive rate (FPR = 1-specificity).<sup>201</sup> It

visualizes all possible thresholds, and it is not sensitive to imbalanced classes. Generally, a larger area under the curve (AUC) indicates better performance for classification.

The final evaluation uses a classifier trained using the liver dataset ( $N = 148$ ) but applied to an independent dataset ( $N = 48$ ) for other anatomical sites (kidney/adrenal glands: pancreas: others = 25:12:11). There were no specific selection criteria for this patient cohort, except that these patients were also treated using Synchrony respiratory motion tracking on CyberKnife between 2010 and 2018 using the body path (Synchrony system). The range of respiratory motion observed for the other sites ( $N = 48$ ) is comparable to the liver cancer patient cohort ( $N = 148$ ). The liver targets exhibited motion ranges in INF/SUP, LT/RT, and ANT/POST directions of  $16.5 \pm 6.5$ ,  $4.0 \pm 2.0$ ,  $7.4 \pm 3.8$  mm, respectively. For other sites, the observed ranges in INF/SUP, LT/RT, ANT/POST directions were  $12.1 \pm 5.6$ ,  $3.2 \pm 1.6$ ,  $5.6 \pm 2.6$  mm, respectively. This not only tests the robustness of SVC but also illustrates the potential benefit for smaller cohorts of patients where using a disease site specific training set would not be possible.

### 3.3.Results

#### 3.3.1. Planning Data Analyses

The Spearman correlation coefficient between volume of a healthy liver receiving  $< 15$  Gy and the size of  $PTV_{std}$  for 148 plans is  $-0.30$ , indicating a weak correlation. If  $PTV_{std-2mm}$  were used instead of  $PTV_{std}$  for planning, the average volume reduction in the PTV would be  $(30.6 \pm 8.4)\%$  for all treatment plans. For the 52 plans with a non-zero overlap of PTV and PRV volumes ( $PTV \cap PRV$ ) this volume could have been reduced by an average of  $1.8 \pm 1.9$  cc if 2 mm smaller margins were used. Statistically significant differences in the prescribed dose and PTV D99 ( $P < 0.05$ ) are found between targets with ( $N = 57$ ) and without ( $N = 121$ )  $PTV \cap PRV$ . Figure 3-3 (a) illustrates lower prescription dose and PTV D99 for targets with and without  $PTV \cap PRV$ . The

Spearman’s rank correlation coefficient between  $PTV \cap PRV$  and PTV D99 was -0.47 and was -0.40 between  $PTV \cap PRV$  and PTV D99/Rx, which indicated  $PTV \cap PRV$  was not strongly correlated with either PTV D99 or PTV D99/Rx. Figure 3-3 (b) shows the increase in the near-minimum dose (D99) the target receives if we reduce the PTV isotropically by 2 mm using the same planned dose.

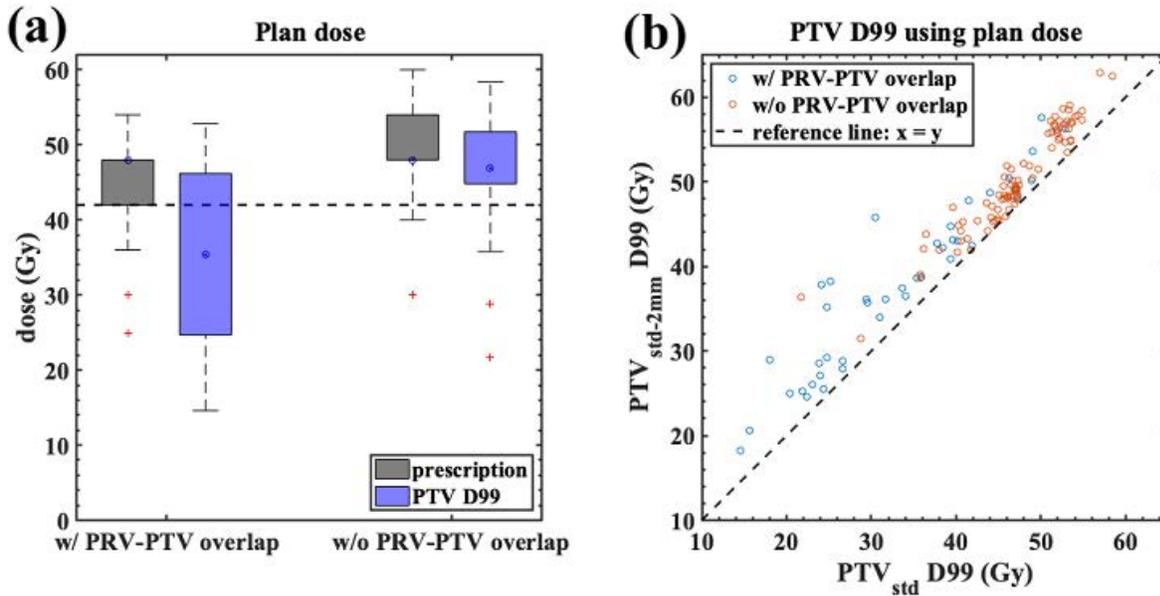


Figure 3-3. (a) Comparison of the prescription dose and PTV D99 for 57 targets with  $PRV \cap PTV_{std}$  (compromised) and 121 targets without  $PRV \cap PTV_{std}$  (non-compromised). Each box spans from the first quartile to the third quartile. The circles and red crosses indicate medians and outliers, respectively. The dashed line indicates 42 Gy, the lower threshold of the prescription dose typically used for robotic SBRT liver treatments at our institution. (b) The relationship between D99 of  $PTV_{std-2mm}$  and  $PTV_{std}$  for 148 plans using the same dose distribution, where  $PTV_{std}$  denotes the standard PTV margin, which was used for planning, and  $PTV_{std-2mm}$  denotes a PTV 2 mm smaller than  $PTV_{std}$ .

### 3.3.2. Simple Two-Feature Analysis

In Section 2.3.1, we observed relatively small (mostly <1 mm) inter-model standard deviations in patients’ Synchrony motion tracking errors, suggesting that patients tend to have higher or lower Synchrony errors consistently throughout their treatment courses. Before employing machine learning with additional features for this classification task, we wanted to first test if the maximum

RMS radial Synchrony error for all remaining models can be predicted using only the average and standard deviation of the smaller sample of radial Synchrony errors observed during the mock session. The multiple correlation coefficient<sup>162</sup> relating the maximum error and the average and standard deviation of errors observed during the mock was 0.70. If we use only these two features for SVC, an accuracy of  $0.73 \pm 0.05$  is achieved suggesting that the use of machine learning with additional features is warranted.

### **3.3.3. Parameter Tuning and Feature Selection**

In an initial assessment of different classification algorithms, we investigated random forest, decision tree, linear discriminant analysis and support vector classification (SVC) and found that SVC was the only classifier to provide accuracy greater than 0.8. In the 5-fold cross-validation (using 148 treatment courses), a maximum accuracy of  $0.84 \pm 0.06$  was achieved using the following parameters: fourteen of the initial 20 identified features (features numbered 1, 7, 9, 10, 14, and 18 in Section 3.2.5 were eliminated); ten principal components; a C value of 10; a relative weighting for Class 1: Class 2 of 1:0.9; and a linear kernel type. The four most important features are the standard deviations of Synchrony errors in the left-right and inferior-superior directions, model linearity in the anterior-posterior direction, and average radial marker motion range, since the accuracy would drop by 0.08, 0.03, 0.05, and 0.03, respectively, if those features were removed from classification.

A few different regression models were also investigated, including multiple regression, regression trees, and support vector regression (SVR). The same set of 20 features investigated for binary classification were used as inputs. The smallest root-mean-square error was 0.5 mm for SVR using a polynomial kernel and 10 principal components.

### 3.3.4. Classifier Evaluation

Leave-one-out cross-validation is accurate in determining the class labels 80% of the time. This technique is also used for the early warning system discussed in Section 3.3.4. The rationale being that during the implementation of the proposed clinical workflow we would have a relatively large patient population to train the data on and would be assessing the model on a patient-by-patient basis.

Figure 3-4 (a) shows the averages and standard deviations of TPs, TNs, FPs, and FNs for the testing set ( $N = 30$ ) over 500-repeated random sub-sampling validations. The sensitivity, precision, specificity, F1 score, and accuracy are  $0.81 \pm 0.09$ ,  $0.85 \pm 0.08$ ,  $0.80 \pm 0.11$ ,  $0.83 \pm 0.06$ , and  $0.80 \pm 0.07$ , respectively.

The ROC curve for the classifier using 5-fold cross-validation is shown in Figure 3-4 (b). An AUC of  $0.87 \pm 0.05$  was achieved with this classifier.

The independent cohort of treatment courses ( $N = 48$ ) for other anatomical sites was used as an additional cross-validation to test the robustness of the classification. The accuracy of classification is 0.77 (sensitivity = 0.60, specificity = 0.89, and precision = 0.80) if liver data is used as the training set and non-liver data is used as the testing set. If only non-liver data were used for training, the accuracy of the prediction would be  $0.71 \pm 0.16$  using 5-fold cross-validation.

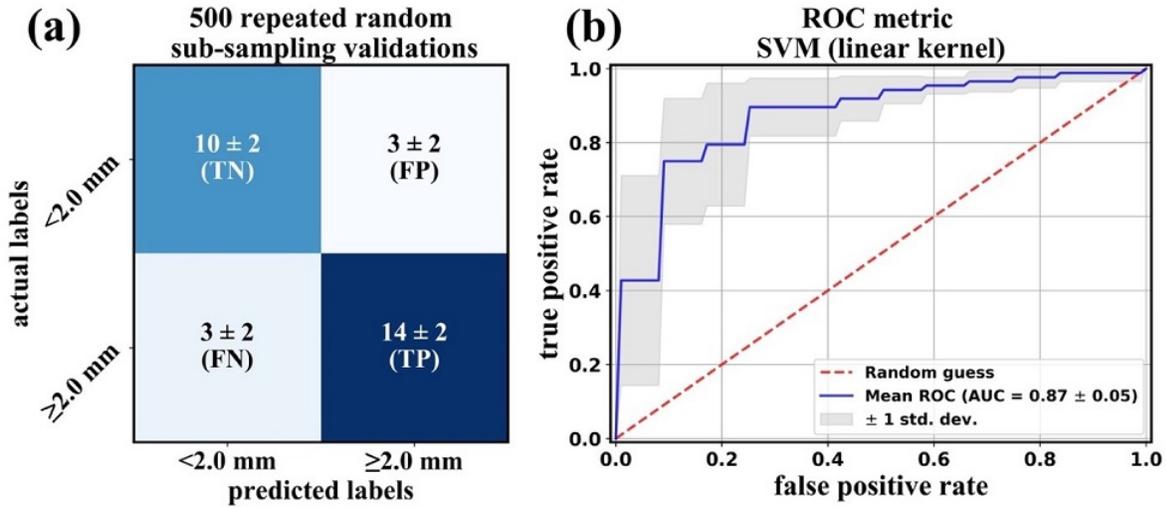


Figure 3-4. The confusion matrix evaluated using 500-repeated random sub-sampling cross-validation (a) and the ROC metric evaluated using 5-fold cross-validation (b).

### 3.3.5. Regression Evaluation

The SVR regression model was validated by 5-fold cross-validation. The relationship between predicted errors and true errors is shown in Figure 3-5 (a), where each error is determined using leave-one-out cross validation. When the predicted errors are rounded to the next integer, as described in Section 2.2.1, the rate of underestimation of tracking error is 17/61 (28%), 11/79 (14%), and 2/7 (29%) for predicted tracking errors of 2 mm, 3 mm, and 4 mm, respectively. In total, for 20% of cases the tracking errors would be underestimated, as shown in Figure 3-5 (b).

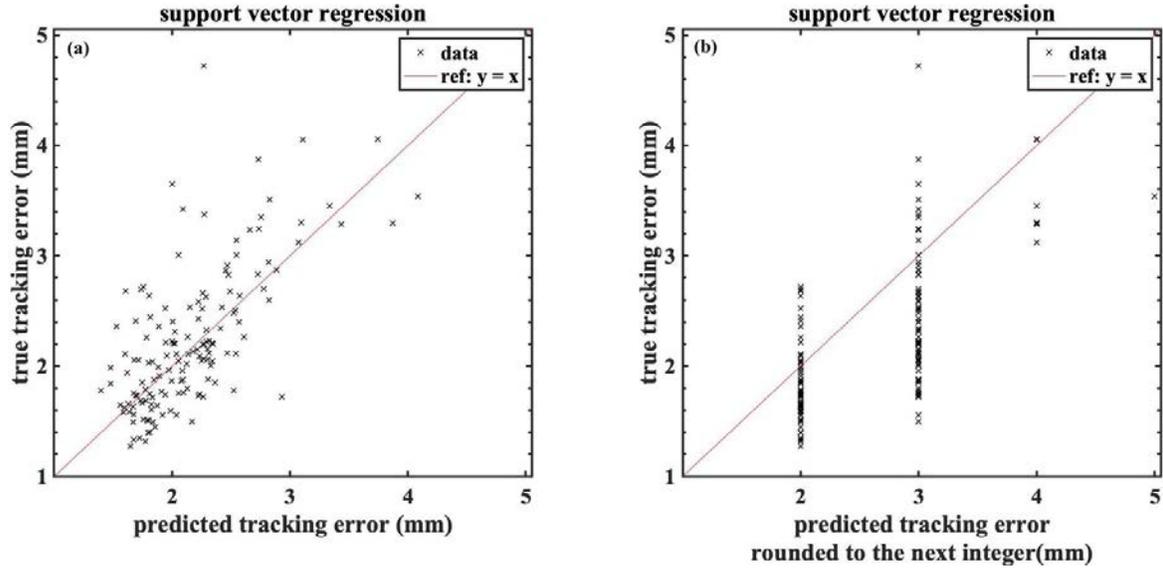


Figure 3-5. (a) The relationship between predicted tracking errors and true tracking errors using a nonlinear regression model. The predicted error for each plan is determined using leave-one-out cross validation. The points above the reference line are cases where the tracking errors are underestimated. (b) The relationship between the rounded predicted tracking errors and true tracking errors. The predicted errors are rounded to the next integer due to the limited CT resolution.

### 3.3.6. Early Warning System

Based on the class labels predicted by leave-one-out cross-validation, 64 and 84 treatment courses are predicted to be of Class 1 ( $<2$  mm error) and Class 2 ( $\geq 2$  mm error), respectively.  $P_i$  is a metric to assess the risk of a geographical miss for a GTV perturbed by residual tracking errors. Among the 64 predicted Class 1, 11 of them have  $P_i < 0.99$  for a target in at least one fraction of treatment and would require replanning using the larger  $PTV_{std}$  following our proposed clinical workflow. The dashed ellipse indicates those treatment courses in Figure 3-6 (a). Figure 3-6 (b) illustrates that using  $PTV_{std}$  ensures that  $P_i$  is larger than 0.99 for all treatment courses. This indicates that the  $PTV_{std}$  margins are sufficiently large to encompass these respiratory tracking errors for the entire population.<sup>27</sup> This also means that in our proposed clinical workflow, all patients will be treated with at most one fraction with  $P_i < 0.99$ , in order to minimize the potential

impact of underdosing the GTV. Figure 3-7 shows additional information available from the GTV voxel occupancy analysis at each fraction which would have been flagged by the early warning system for the eleven treatment courses. A treatment course typically consists of three to six fractions. Any remaining partially delivered treatment fractions which were not completed in the allotted two-hour session per fraction are completed in the makeup fraction. Treatment courses no. 4, 7, 8, and 11 had the seventh fraction as the “makeup fraction”. Many of these plans have more than one target and may not require larger margins for every target. For treatments with more than three fractions, trending across the fractions could also be useful to identify outliers for lower GTV voxel occupancy.

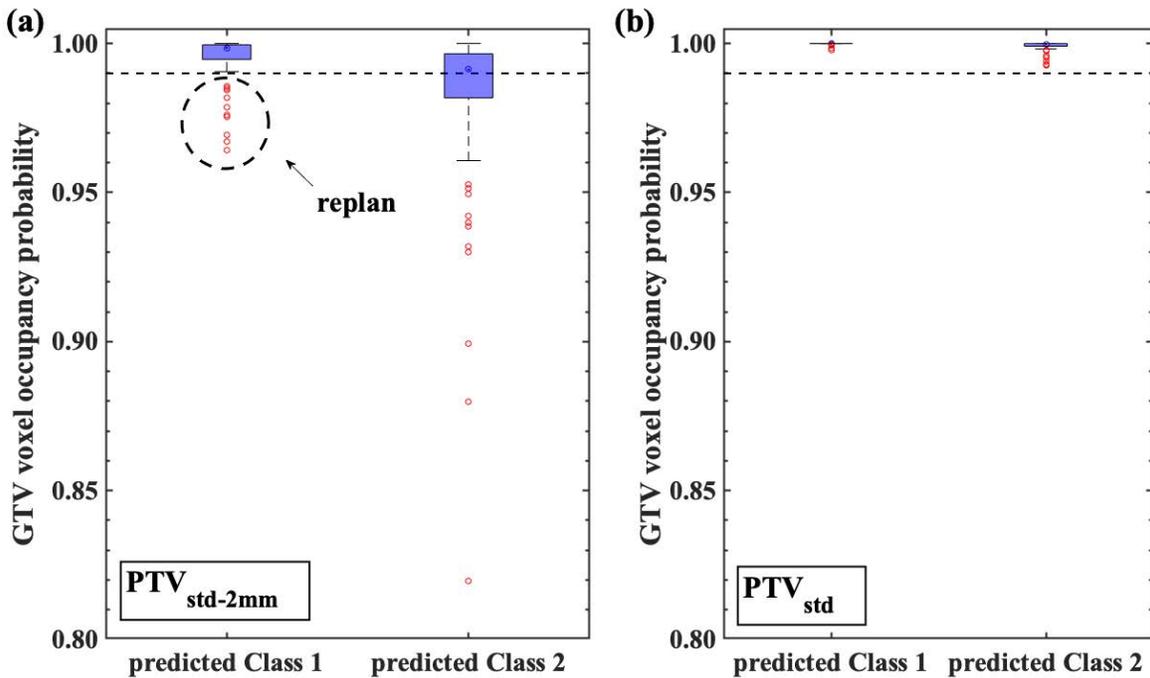


Figure 3-6. The minimum fraction of the perturbed GTV within the  $PTV_{std-2mm}$  (a) and the  $PTV_{std}$  (b), respectively, for predicted Class 1 and Class 2 for all fractions of treatment. The GTV voxel occupancy probability is a metric to assess the risk of geographical miss of a GTV perturbed by residual tracking errors. The dashed ellipse indicates treatment courses that require replanning using  $PTV_{std}$  for subsequent treatments.

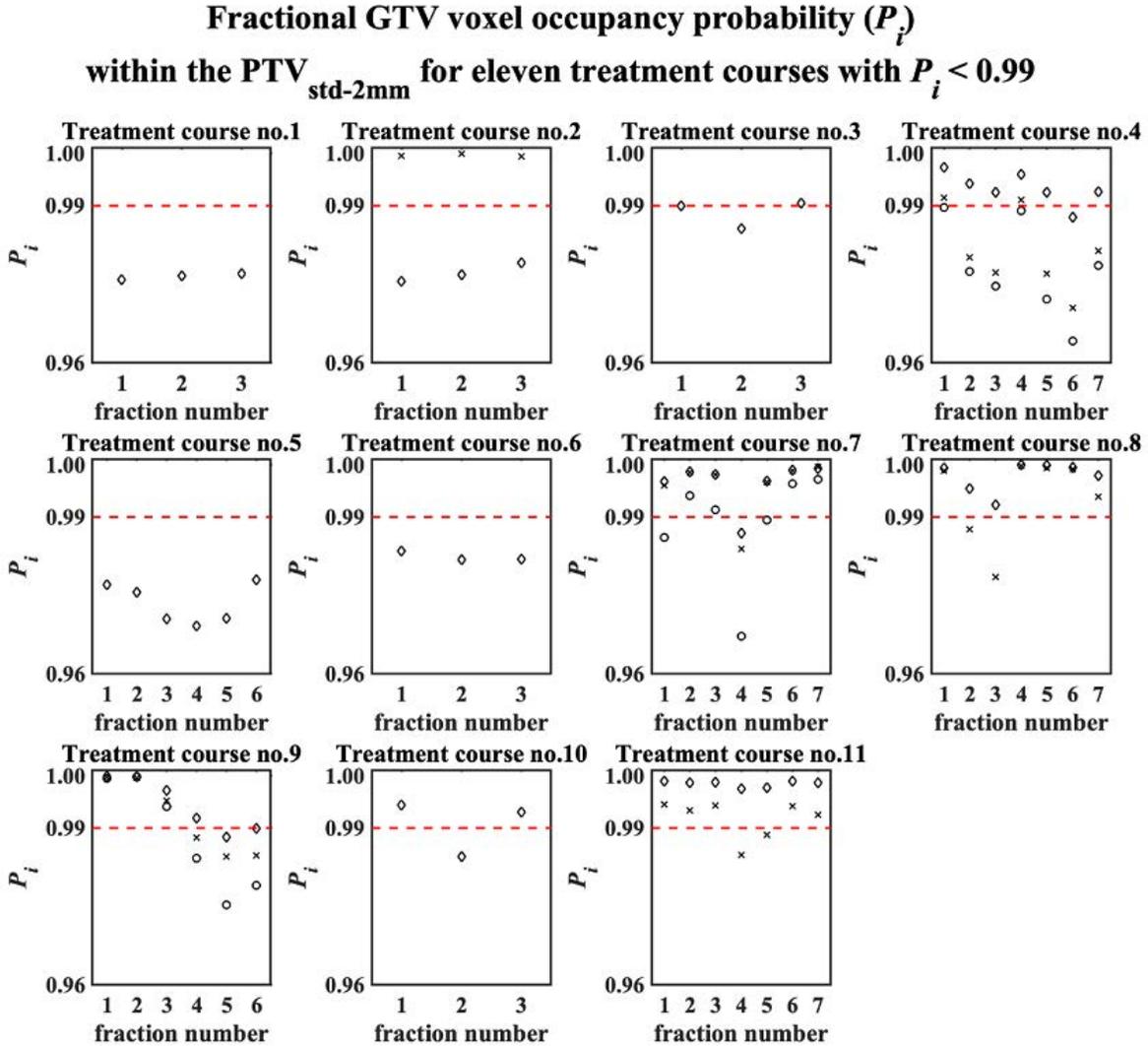


Figure 3-7. The fractional GTV voxel occupancy probability ( $P_i$ ) within the PTV<sub>std-2mm</sub> for eleven treatment courses. These plans have at least one target with  $P_i < 0.99$ , as indicated by the black points below the dashed line ( $P_i = 0.99$ ). If present, different targets treated in the same plan are shown as diamonds, crosses, and circles. There is more than one target for plans no. 2, 4, 7, 8, 9, and 11. The seventh fraction is the “makeup fraction” for treatment courses no. 4, 7, 8, and 11. Any remaining partially delivered treatment fractions which were not completed in the allotted two hour session per fraction are completed in the makeup fraction.

When the regression model discussed in Section 3.3.5 is used to predict the tracking error, we can reduce the PTV margin accordingly and calculate the GTV voxel occupancy probability ( $P_i$ ). As shown in the left boxplot of Figure 3-8, if we reduce the PTV margin 5 mm minus the rounded predicted maximum tracking error, we would have discrete margin reductions of 3, 2, and 1 mm,

respectively, for predicted tracking errors being 2, 3, and 4 mm, and the  $P_i$  per target falls below 0.99 for 52.7% of patients in at least one fraction. If we use the more conservative margin reduction formula, the PTV margins would be reduced by 2 and 1 mm for predicted tracking errors of 2 and 3 mm, respectively, and the  $P_i$  per target falls below 0.99 in 14.2% of patients.

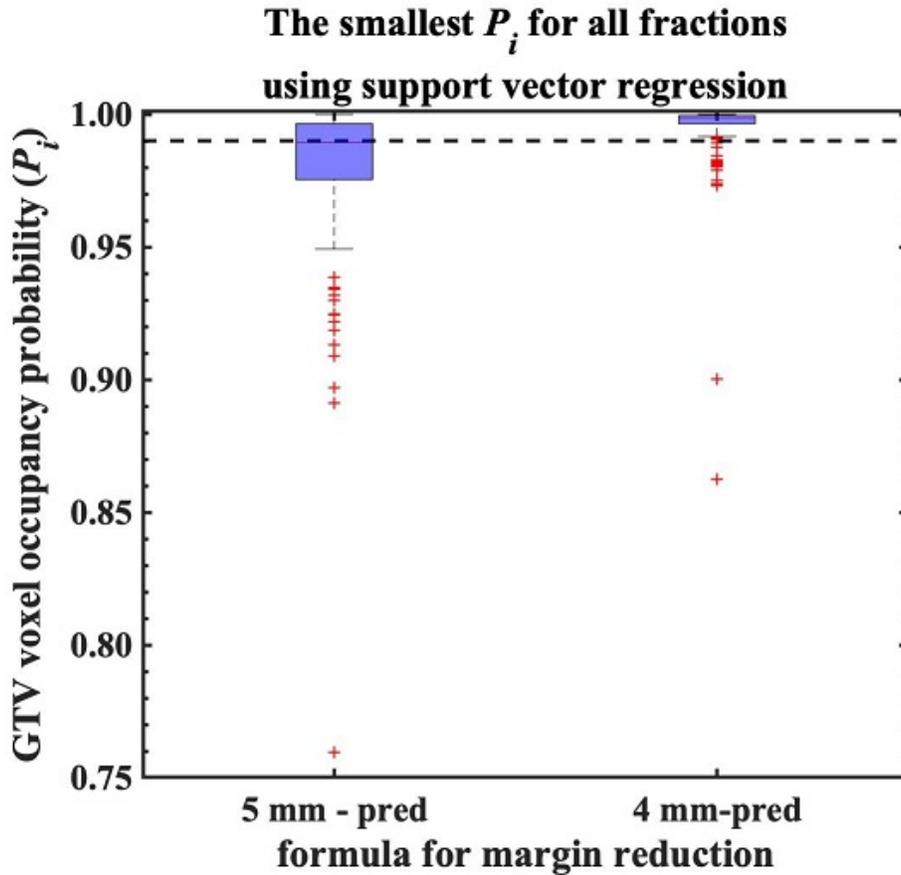


Figure 3-8. The distribution of GTV voxel occupancy probability ( $P_i$ ) for 148 plans shown as the left and right boxplots if the PTV margins are reduced by 5 mm minus rounded predicted errors, and reduced by 4 mm minus rounded predicted errors, respectively. For cases where  $P_i < 0.99$ , the replanning is needed since  $P_i$  quantifies the geographical miss of a GTV perturbed by residual tracking errors. The replanning rates for these two scenarios are 52.7% and 14.2%, respectively.

## 3.4. Discussion

### 3.4.1. Planning Data Analysis

Dosimetric and volumetric analyses illustrate the potential benefit of reducing PTV margins for liver SBRT. A weak correlation (Spearman  $r = -0.30$ ) between the volume of healthy liver receiving less than 15 Gy and the PTV was observed in our data. Whether this planning constraint can be met depends on the target size and location within the liver as well as total volume of the liver. In our institution a more dose-limiting constraint tends to be the proximity or overlap of targets with serial organs at risk. In our study, 52 out of a total of 148 plans had overlap between a 5 mm isotropic serial structure PRV and the PTV. As a result, dose to the target is often compromised, as indicated by the reduction in both the PTV D99 and the prescribed dose as shown in Figure 3-3 (a). Figure 3-3 (b) indicates that without replanning D99 of PTV can be increased by  $3.7 \pm 2.9$  Gy if the PTV is reduced by 2 mm isotropically. If smaller PTV margins are used for certain patients with smaller motion-tracking errors, the overlap volume is reduced and potentially a higher dose can be prescribed to tumors without an increase in toxicity. If the margins being used for planning are too large and do not realistically reflect the motion tracking errors for the patient, the near-minimum dose to the tumor will be underestimated and may be poor predictor of tumor control.

### 3.4.2. Simple Two-Feature Analysis

The two-feature analysis provides a simple method to predict the maximum Synchrony error for future fractions based only the average and standard deviation of the radial Synchrony errors observed during the mock session. However, relying only on these two features may not provide satisfactory predictive accuracy ( $0.73 \pm 0.05$ ) for safe implementation for patient-specific margins using only these two quantities.

### 3.4.3. Classifier Evaluation Using More Features

Using 14 features and algorithm parameters tuned to maximize classification accuracy, the 5-fold cross-validation accuracy is  $0.84 \pm 0.06$  and for 500-repeated random sub-sampling cross-validations is  $0.80 \pm 0.07$ . The sensitivity, precision, specificity, F1 score provide additional information about classifier performance. The relatively high precision ( $0.85 \pm 0.08$ ) implies that treatment courses of Class 2 ( $\geq 2$  mm error) get flagged with few false alarms. However, some treatment courses will be misclassified as Class 1 ( $< 2$  mm error) due to the relatively lower sensitivity ( $0.81 \pm 0.09$ ). The SVR regression analysis using a polynomial kernel achieves a relatively low root-mean-square error of 0.5 mm but presents some cost-benefit concerns for planning resources which are discussed more detail in the following section.

At our institution, we have a much smaller cohort of treatment courses for non-liver sites treated using the Synchrony MTS. The cross-site validation that trains the classifier using the liver data set ( $N = 148$ ) and predicts for other sites ( $N = 48$ ) validates the robustness of the model. If only non-liver data is used for training and testing, the accuracy of classification was only 0.71 with a relatively high standard deviation of  $\pm 0.16$  using 5-fold cross-validation. If liver data is used for training, the prediction accuracy for non-liver data increases to 0.77, however a relatively low sensitivity of 0.60 suggesting that there may be differences between the anatomical sites not accounted for in this model. To improve the accuracy and robustness of the predictions, data sharing between multiple different centers would increase the sample size and to make it possible to apply our SVC model to other institutions and disease sites. Features in this study are closely related to motion-tracking and can be calculated for any patient using only post-treatment log file data. Demographic or histological characteristics are not considered in features. In Section

2.3.1, we found that age, gender and location of the tumor within the liver were not correlated with the magnitude of Synchrony errors; therefore, these features are not included in this study.

One limitation of our study, compared to other studies using SVC for binary classification in radiotherapy applications, is our relatively small sample size. However, our choice of margin reduction threshold allows us to have well-balanced classes and relatively high prediction accuracy. Chen *et al.*<sup>185</sup> predicted whether patients (N = 219) had Grade 2+ radiation-induced pneumonitis after 3D conformal radiation therapy using SVC (RBF kernel) with both dose and non-dose related features and obtained an AUC of 0.76 (sensitivity/specificity = 0.74/0.75). Their study had imbalanced classes with negative/positive outcome = 185: 34, and the accuracy and precision with uncertainty were not reported. Pella *et al.*<sup>136</sup> used SVC (polynomial kernel) to predict if the prostate toxicity level was Grade > 2 for 321 patients and had an AUC of 0.7, validated using an independent test sample of 30 patients. Using SVC (RBF kernel), Klement *et al.*<sup>133</sup> found seven most important prognostic factors for predicting the local failure in non-small-cell-lung-cancer (NSCLC) patients (N=399). They used a feature forward selection algorithm (same as Chen *et al.*<sup>185</sup>) and achieved an accuracy of  $0.73 \pm 0.2$  using 100-repeated 10-fold cross-validation. Schiller *et al.*<sup>202</sup> studied the risk of radiation induced pneumonitis resulting from thoracic treatments and used an ensemble of SVMs with balanced samples from the dataset for training and testing, to reduce the impacts of unbalanced data. Granville *et al.*<sup>139</sup> found that ten features derived from VMAT plans and quality control metrics (N = 1620) could help classify hot (dose higher than planned), normal, and cold plans based on the median dose deviation of the measured to the planned dose distributions. They obtained a micro-averaged AUC of 0.93 and a macro-averaged AUC of 0.88 using linear SVC when 25% of the entire dataset was tested independently. In our study, data from other anatomical sites are used for validation independently.

### 3.4.4. Early-Warning System

For the proposed early-warning system, less than 8% of patients (11 of 148 total patients) would require replanning with larger margin expansions ( $PTV_{std}$ ) using our binary classification scheme. For these patients, we could boost the peripheral of PTV for the one or more fractions to compensate for a potential geographical miss. A total of 36% (53/148) of patients could be treated using smaller margins ( $PTV_{std-2mm}$ ) with fractional GTV voxel occupancy ( $P_i$ ) greater than 0.99 for all fractions. For those patients, dose escalation may be achievable. However, for the 8% of patients incorrectly assigned smaller margins, a clinical decision to deescalate the dose may be necessary. If we followed a slightly different clinical workflow and assessed  $P_i$  for every patient, 66% of the patient population could be treated with smaller margins, which would be clearly beneficial to an even larger population but would require more treatment planning resources. As illustrated in Figure 3-7, the early warning system per fraction  $P_i$  data, trending across fractions could help to identify outliers and identify targets less sensitive to tracking errors which would allow for more informed replanning. Since our proposed workflow uses a mock session prior to planning, patient-specific PTV margins can be determined before treatments offline with little delay in the patient's clinical schedules as each treatment course requires additional 15 min for a mock session prior to planning. It takes less than 10 minutes to process the post-mock log file data and determine whether to reduce the PTV margin for a new treatment course using the in-house scripts. It takes less than 5 min to calculate  $P_i$  for each fraction. These times are not excessive compared to the typical treatment fraction time being 1 to 2 hours.

One limitation of this workflow is that it does not include any replanning for treatment courses predicted as Class 2 ( $> 2$  mm) and will therefore overestimate the required margins for some patients as shown in Figure 3-6. It also does not explicitly model the dosimetric impact of tracking

errors, including the synchronization of the sign of the motion tracking error with respiratory phase<sup>142</sup> which can lead to decreased target dose due to penumbral broadening.<sup>172</sup> The dosimetric impacts of these types of delivery error are investigated in more detail in Chapter 4 and Chapter 5. In Chapter 5, we extend the GTV voxel occupancy estimation to also include the effect of uncorrected rotations and rigid body errors and show how it can be combined with dosimetric information in a modified version of the early warning system.

As shown in Figure 3-8, if we reduce the PTV margin semi-continuously using predicted errors determined by a regression model, at least 140 out of 148 cases could have PTV margins reduced. However, if we reduced PTV margins by 5 mm minus the rounded predicted tracking error, as shown in the left boxplot of Figure 3-8, the replanning rates would be 52.7%. This would be an aggressive approach that may not be suitable for implementation in a busy clinic with limited resources. However, if we reduced PTV margins by 4 mm minus the rounded predicted tracking error, as shown in the right boxplot of Figure 3-8, the replanning rate would be 14.2% which is higher than the 8% required using the binary classification model. Using the more conservative regression-based margin reduction formula, 79 plans would be treated with PTV margins reduced by only 1 mm. In this case, the balance between the benefit to the patient and increased planning resources may be difficult to justify.

One of the challenges for safe clinical implementation of offline adaptive treatments is the balance of risk, benefit and resources required to determine the appropriate time to intervene.<sup>129</sup> Since this a single institutional study, the SVC model might not be generally applicable to other institutions, however, our results for different anatomical sites show good predictive accuracy (0.77) indicating the generality of this method. The early-warning system flags treatment fractions for replanning only when the PTV margin is too small. This is a conservative and efficient

approach, which may make dose escalation and/or margin reduction more palatable to clinicians and administrators.

### **3.5. Conclusions**

This study presents a clinical workflow for determining patient-specific margins for liver SBRT using robotic radiosurgery. Data presented retrospectively emulate a dry-run session (mock session) prior to planning which is used to evaluate motion-tracking errors and other features to assess suitability for margin reduction. Supervised machine learning is used prior to planning to predict the maximum RMS motion-tracking errors in subsequent fractions with an accuracy of  $0.84 \pm 0.06$  using 5-fold cross-validation. An early-warning system is also proposed to evaluate target coverage at each fraction based on the observed motion tracking errors, which would flag a small subset of patients for replanning using larger PTV margins for subsequent treatments.

## Chapter 4. Experimental Validations of the Dose Perturbation Tool

### 4.1. Introduction

As described in Chapter 3, the GTV voxel occupancy probability is a geometric coverage metric which is useful for evaluating the suitability of patient-specific PTV margins in the proposed early warning system. Plans of many patients treated on CyberKnife have very heterogeneous dose distributions. Some have very high dose in the central regions of the PTV and others have steep dose gradients within the PTV to reduce dose to nearby OARs. The probability that a GTV receives the prescribed dose depends on the specific features of the dose distributions as well as geometric considerations. The dosimetric impact of uncorrected rotations, motion-tracking residual errors, and target deformations should also be evaluated, since it can lead to underdose of targets or excessive dose to OARs. In this chapter, we propose and validate a dose assessment tool that perturbs the dose to target(s). This tool can be applied to any treatment on CyberKnife.

Lujan *et al.*<sup>203</sup> proposed a method to estimate dose to targets affected by intrafraction motion assuming the target was a rigid body. They convolved the static dose distribution for treating a liver lesion with the probability distribution function (PDF) of simulated organ motion data. They also recalculated the dose distribution using direct simulations, which meant the isocentre and beams were shifted in the opposite directions from the target motion, and the time-weighted average dose was calculated for the organ position at each time bin. The convolved dose distribution agreed well with the recalculated dose with differences less than 2% in over 98% of the healthy liver volume. The dose convolution was easy to apply, in contrast to direct simulations which required multiple calculations using the planning system for targets at different respiratory phases. The authors also addressed the limitation of applying the convolution-based method to all

structures since some structures did not move with the target, such as the spinal cord, and suggested that organ-specific PDFs could be quantified and used for those structures. Karlsson *et al.*<sup>204</sup> also compared dose-shift (using convolution) and beam-shift (shift of isocentre) based methods in estimating dose delivered to lung targets accounting for rigid translational uncertainties (with setup errors up to 5 mm and intrafraction target motion amplitude up to 10 mm in any anatomical direction) for SBRT treatments. They found that the minimum dose to the target, quantified as D98, calculated using the dose-shift method was within about 2% difference to that calculated by the beam-shift method (ground truth), which was simulated using clinically relevant motion and setup errors. The agreement in the magnitude of the dose calculated using these two methods supports the fact that ignoring the difference of tissue densities between tumour and surrounding tissues is a reasonable assumption as long as the magnitude of the shift is small.

Another approach to consider geometric uncertainties in dose calculation is fluence convolution, in which the fluence incident on the patient is convolved with geometric uncertainties prior to dose calculation. It has been proved to be particularly useful in lung tumours in which tissue inhomogeneities are more prominent.<sup>205,206</sup> Chetty *et al.*<sup>205</sup> compared dose-convolution and fluence-convolution calculated dose for three lung treatments. They found the dose-convolution based calculation had a 5% lower estimation for dose close to the lung compared to the fluence-convolution based method, which compensated for tissue inhomogeneities in dose calculation. Beckham *et al.*<sup>206</sup> convolved the incident photon fluence with a PDF of the random setup errors for a virtual phantom, and compared it to dose calculated using the TPS (ground truth, MC calculation) and the dose-convolution method. They found that dose-convolution based calculation underestimated dose perturbation in boundaries of heterogeneous materials (lung/water and

bone/water). The dose profiles of fluence-convolution and dose-convolution based calculations had discrepancies of up to 5% when the 0.4 cm setup error was present.

Ravkilde *et al.*<sup>207</sup> proposed a fast online dose reconstruction method that accounted for target motion and compared the reconstructed dose with measured dose using a phantom containing two planar diode arrays (Delta4PT, Scandidos, Sweden). The dose for each detector plane was estimated using a pencil beam convolution algorithm prior to scaling the calculated dose to user-defined points of interest in other planes using the inverse square law and a percentage depth dose curve measured in water. Calculation points were shifted rigidly based on the average target position to emulate the translational motion of targets. The RMS gamma pass rate for all experiments was 97.4% using 3%/3mm criteria for comparing the reconstructed and measured dose. In another study, the same research group demonstrated the accuracy of the online dose reconstruction in a phantom study for five liver SBRT plans.<sup>82</sup> The mean  $\pm$  SD of the dose difference for the cumulative dose between reconstructed and measured dose (also using Delta4 phantom) was  $-1.6\% \pm 3.9\%$  for all experiments (with and without MLC motion tracking). Skouboe *et al.*<sup>81</sup> applied the same online reconstruction tool to fifteen simulated SBRT liver treatments (90 sessions). The dose was also calculated using the TPS, where motion was modelled as a series of isocentre shifts. They compared the change of D95 for the CTV ( $\Delta$ D95) due to motion estimated using these two methods (water density was assumed for TPS calculation) and found an RMS deviation of 0.80%-points. The limitation of these studies discussed above is that they all assumed the target was a rigid body and target rotation was not considered.

For small field dosimetry on the CyberKnife system, it is challenging to obtain accurate dose measurements due to high dose gradients, a lack of lateral charged particle equilibrium, and nonnegligible fluence perturbation by a dosimeter.<sup>208-210</sup> The advantages and drawbacks of

different types of radiation dosimeters have been summarized by Wilcox *et al.*<sup>19</sup> They investigated the reproducibility, sensitivity, dose linear response, energy independence, spatial resolution, water equivalence and ease of use for ionization chambers, thermoluminescent detectors, semiconductors, radiographic films, and radiochromic films. As stated in the American Association of Physicists in Medicine (AAPM) Task Group 55 Report,<sup>20</sup> radiochromic films can be a reliable radiation dosimeter that provides two-dimensional dose measurement with a high spatial resolution for small fields. They are made of nearly tissue-equivalent materials, consisting of a single or double layer(s) of radiosensitive organic microcrystal monomers coated with a transparent polyester base. The active layer becomes darker as the absorbed dose increases and the optical density (defined as the base ten logarithm of the ratio of unirradiated and irradiated film intensity of transmitted light<sup>211</sup>) is calibrated using a non-linear function to the absorbed dose. Unlike silver halide radiographic films, radiochromic films do not require chemical processing and can be used for accurate radiation dosimetry using only a desktop flatbed scanner.<sup>212-216</sup> They are also much less sensitive to room light.

The most common radiochromic films currently used for medical radiation dosimetry are Gafchromic™ EBT3 (Ashland Inc., Covington, KY) films. These films are highly suitable for small field dosimetry measurements on CyberKnife.<sup>217,218</sup> Moignier *et al.*<sup>217</sup> found that EBT3 films had the best agreement with Monte Carlo simulation in off-axis ratio measurements (measured dose profiles normalized at the central axis intensity), compared with three other detectors (PTW31014 ion chamber, and PTW60017 and SN EDGE diodes). The correction factors for those detectors were estimated using Monte Carlo simulations. Using a respiratory motion phantom, Marants *et al.*<sup>169</sup> compared measurements, made simultaneously with a micro-MOSFET in contact with EBT3 films, with the planned dose distributions for several different types of breathing traces.

They found the film measurements and RADPOS measurements agreed within 2% for all dynamic and static cases. The correction factors for the micro-MOSFET were provided by Ploquin *et al.*<sup>219</sup> using the Monte Carlo calculations.<sup>219</sup>

Using experiments with EBT3 films positioned in a phantom, we validate the dose perturbation tool by comparing the perturbed dose, which accounts for different types of delivery errors, with film measurements as a ground truth using gamma analyses.

## 4.2. Methods and Materials

### 4.2.1. Phantoms

To study the dose deposited in a patient body, we use “phantoms” consisting of materials in which photon and electron interactions are similar to those in specific tissue types or body regions. Acrylic, Solid Water™ (Gammex Middleton, Wisconsin, USA), and expanded polystyrene (Styrofoam™), on the decreasing order of density, are common materials used in phantom construction.

Four phantoms are used in this study, as illustrated in Figure 4-1. They will be referred to as the rocket phantom, the ball cube, and the lung and liver inserts (for the Quasar platform). Each phantom has a minimum of four fiducials embedded in the phantom material. They are used to position the phantom and allow translational and/or rotational motion tracking using either the CyberKnife system’s static fiducial tracking or Synchrony respiratory motion compensation. Precise spatial co-registration of the planned and delivered dose is achieved by fixing the films within each phantom using high density pins which can easily be identified in the CT images and holes in the films which can be automatically identified in the film scan images.

The rocket phantom is an in-house manufactured bullet-shaped heterogeneous phantom meant to emulate a head-like geometry. It is made primarily of Solid Water with a slot for a stack of four

films (film size:  $10.2 \times 12.7 \text{ cm}^2$ , six fiducials) which can be oriented in coronal or sagittal plane orientations. It was designed for patient-specific quality assurance on CyberKnife. In patient-specific quality assurance, a patient's plan is recalculated using a CT scan of the phantom, the planned dose is delivered to the phantom, and the measured and planning system doses are compared. This phantom emulates a situation where a tumour is surrounded by homogeneous water-equivalent tissues such as in the brain.

The end-to-end (E2E) ball cube (side length 6.4 cm) is a ball-shape plastic phantom provided by Accuray for quality assurance purposes with a denser spherical "hidden" target located in the centre of the cube. Two precisely cut films (size:  $6.3 \times 6.3 \text{ cm}^2$ ) can be placed at the orthogonal slots in transverse and sagittal planes. As shown in Figure 4-1, the cubic film insert fits inside an acrylic ball-shaped phantom. This phantom is referred to as the ball-cube phantom and is meant to be used to perform quality assurance for each of the CyberKnife system's fiducial-based tracking methods. It can also be placed on a motion platform (described in more detail in Section 4.2.2) to test the motion tracking capabilities of the Synchrony system. The geometric accuracy of the delivered dose is assessed using the same clinical processes, which are used for treating patients, beginning with CT simulation and ending with the dose delivery. For that reason, we refer to this type of quality assurance as end-to-end testing.

The cylindrical lung insert (diameter 7.9 cm, length 16.1 cm) was partially constructed in-house and is mostly made of low-density cedar (approximately  $0.26 \text{ g/cm}^3$ ) to represent lung tissue, with a split sphere (diameter 3 cm) made of Solid Water to emulate a lung tumour located in the centre of the cedar cylinder. The cylindrical liver insert was also constructed in-house and is meant to emulate the interface of the liver and lung tissue. The lower half of the insert is made of black Perspex acrylic provided by the vendor, and the upper half was made in-house and consists of

approximately 1/3 Styrofoam and 2/3 Solid Water. For lung and liver inserts, two EBT3 films ( $7.8 \times 15.6 \text{ cm}^2$  and  $7.4 \times 15.6 \text{ cm}^2$  for the lung and liver inserts, respectively) can be exposed for each experiment. These inserts were designed to be compatible with the Quasar respiratory motion platform discussed in Section 4.2.2.

For the studies in this thesis, the dose volume for heterogeneous phantoms such as the E2E and rocket phantoms is calculated using Ray-Tracing. Monte Carlo calculation is used for the lung and liver inserts to account more accurately for the heterogeneity of the medium.

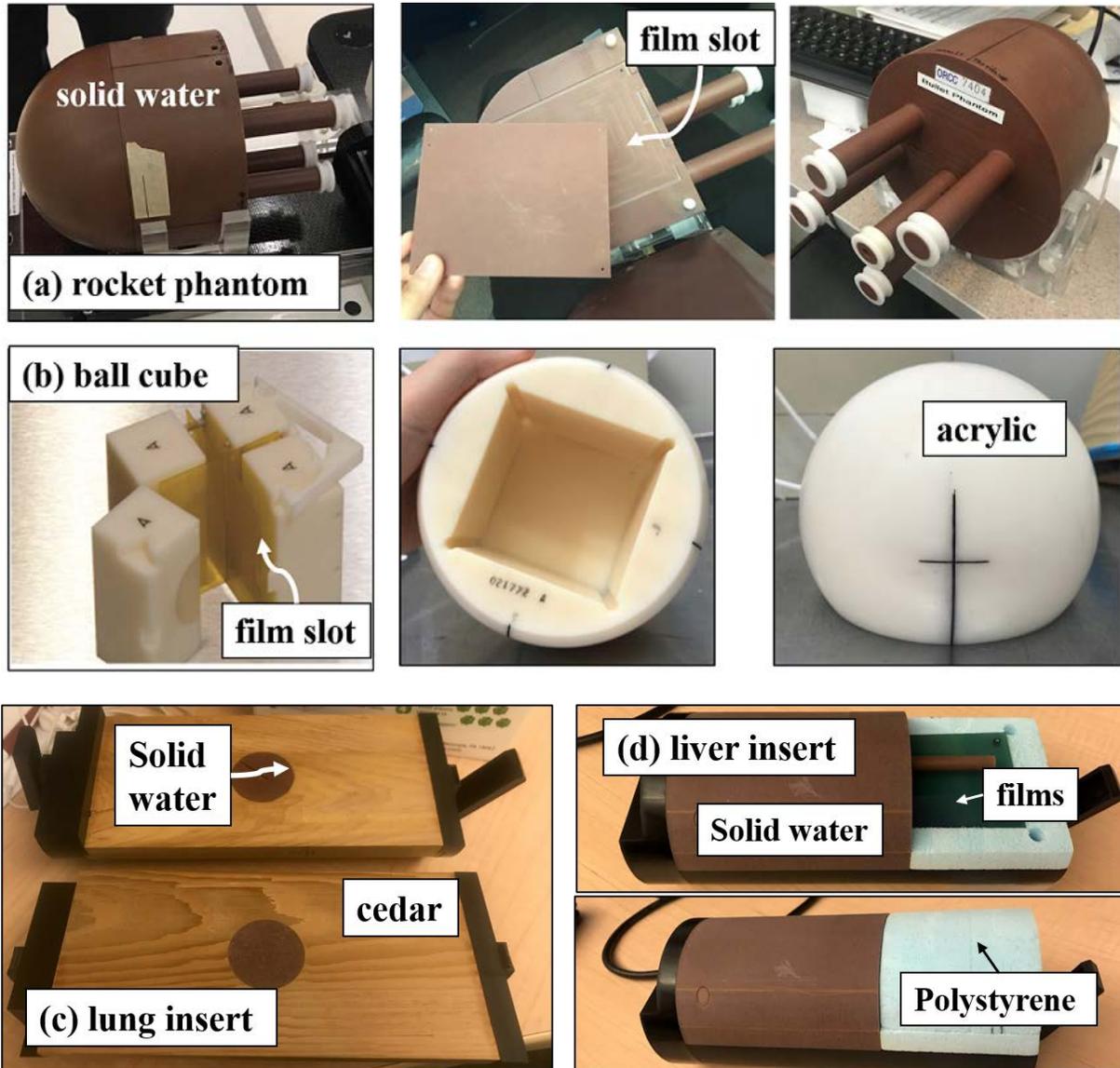


Figure 4-1. (a) The rocket phantom is made primarily of Solid Water with a slot for a stack of four films which can be oriented in coronal or sagittal plane orientations. (b) The ball cube ball-shape plastic phantom is provided by Accuray for quality assurance purposes with a denser spherical “hidden” target located in the centre of the cube. Two custom cut films can be placed at the orthogonal slots in coronal and sagittal planes. (c) The lung insert is made of low-density cedar to represent lung tissue and its center is a sphere made of Solid Water to emulate a lung tumour. (d) Half of the liver insert is made of black Perspex acrylic provided by the vendor, and the other half is customized (approximately 1/3 Styrofoam, 2/3 Solid water).

#### 4.2.2. Respiratory Motion Platforms

A simple motion platform (Figure 4-2) is provided by Accuray to perform end-to-end quality assurance of the Synchrony motion tracking system. It creates simple periodic motion with a constant frequency and amplitude on two separate platforms moving in orthogonal directions. The LED markers are placed on a platform moving in the anterior-posterior direction, and this motion can be adjusted to be in phase or out of phase with another platform on which the ball cube phantom moves in the superior-inferior direction.

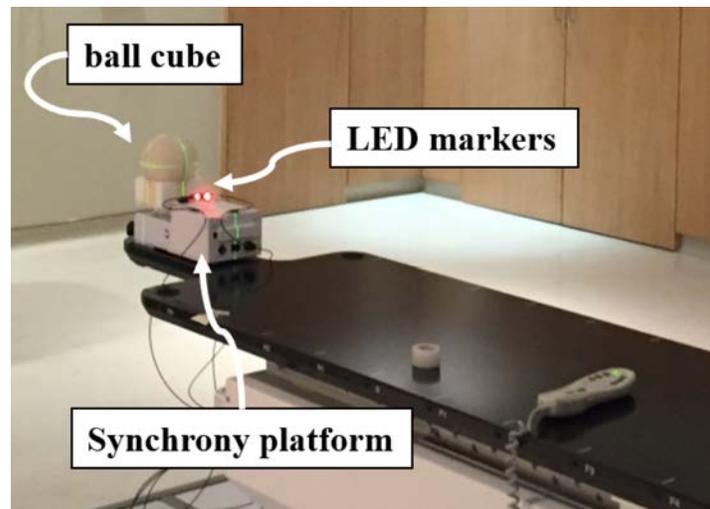


Figure 4-2. The Synchrony platform is used to create periodic motion for the phantom and LED markers moving in perpendicular directions.

For more complicated motion, we employ a Quasar™ (Modus QA Inc., London, ON, Canada) respiratory motion phantom, which has a motor that drives a cylindrical insert in one direction (usually superior/inferior) to emulate internal breathing motion in the thorax or abdomen. A platform simultaneously moves in the anterior-posterior direction to emulate the breathing motion of the chest wall. LED markers are placed on the platform to emulate the external breathing signal. The difference between this phantom and the motion platform provided by Accuray (shown in Figure 4-2) is that the Quasar can be programmed for different motion types. These include

periodic motion with variable amplitude and frequency, and the user can also input an arbitrary respiratory motion trace. The motion of the platform is in phase with that of a cylindrical insert (Figure 4-3a).

Out of phase motion can be emulated by running the same motion trace on two different Quasar systems with a time delay between setting the phantoms in motion as shown in Figure 4-3b. For out of phase experiments, we fit an ellipse to the internal (fiducial) and external (LED) locations for a running window of 15 data points (see Appendix D). An average phase shift between the phantom and marker motions is calculated for each experiment. EBT films (or other dosimeters) can be loaded in the cylindrical insert to measure dose in a 2D plane moving with the target, and therefore, test respiratory motion compensation techniques.

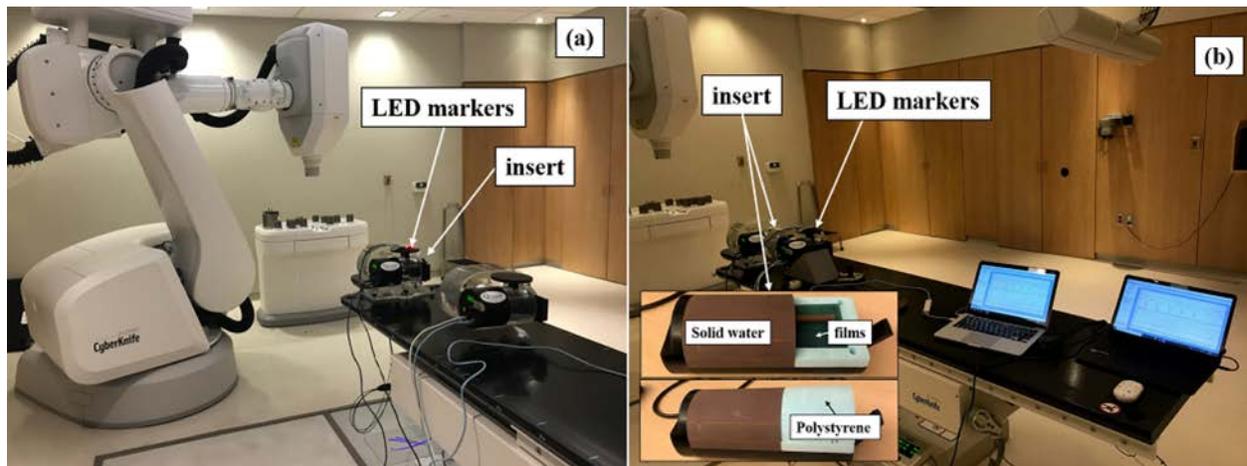


Figure 4-3. The experiment setup to create in-phase (a) and out-of-phase motion (b) for markers and a cylindrical insert using Quasar(s). Two laptops (b) used for controlling the motion of two Quasars.

### 4.2.3. Film Dosimetry Protocol

Gafchromic EBT3 films (thickness 0.28 mm) are suitable for measurements in the dose range from 200 cGy to 1000 cGy. They have a high spatial resolution ( $<25 \mu\text{m}$ ), and a nearly energy-independent dose response for MV photon beams.<sup>220-222</sup> EBT3 films are near tissue equivalent with an effective atomic number of films of 6.7, compared to water's effective atomic number of 7.4,

which minimizes the perturbations caused by the presence of the detector materials within the irradiated media.<sup>223</sup> At least two films are used in each experiment to reduce the uncertainty of dose measurements (overall uncertainty reported by other studies: 1.3% using two films<sup>216</sup> and 0.9% using five films<sup>224</sup>).

The workflow for film processing is illustrated in Figure 4-4. The films are scanned prior to irradiation and at approximately 24 hours after irradiation using an Epson 10000XL scanner (Epson, Nagano, Japan) with 48-bit colour, 150 dpi, and in transmission mode. For lower dose experiments (maximum planned dose <300 cGy), film calibration is performed using the red channel following the procedure and functions described in Devic *et al.*<sup>212</sup> For higher dose experiments, a specific function of the net optical density is used which linearizes the dose response curve for the green channel. The measured dose is proportional to  $(-1) \cdot \text{netOD}^{2/3} / \ln(\text{netOD})$ , where OD represents the optical density.<sup>225</sup> For more recent film batches, it is found that the exponent in the numerator should be adjusted from 2/3 to 4/5 to minimize the  $R^2$  of the fit function. For all experiments, a rescaling is performed following the approach used by Lewis *et al.*<sup>226</sup> The rescaling is based on two pieces of film cut from the same sheet and scanned at the same time as the film piece used for the experimental measurement. One piece of film is left unirradiated but exposed to the same environmental conditions as the experimental test films. This film accounts for any changes because of heat or light exposure and differences in the scanner lamp brightness during the pre- and post- irradiation scans. The other piece of film is irradiated to a known dose using the 60 mm collimator at 1.5 cm depth in Solid Water, 80 cm from the source. In-house MATLAB scripts are employed for film processing and gamma statistics.

The measured dose not only results from MV beam deliveries but also consists of kV x-ray imaging dose. No correction is applied to experiments where the estimated imaging dose is

negligible (<1% of the maximum planned dose), except for experiments using the liver insert (explained in Appendix E). This is determined by using an upper threshold on the surface imaging dose estimated using the measured data provided by the CyberKnife imaging whitepaper,<sup>227</sup> which is rescaled based on milliamperere-seconds (mAs) of an x-ray tube. A slightly higher x-ray tube voltage (kVp) is chosen for the estimation if the table does not provide the imaging dose for the exact kVp used in the experiment.

For experiments using the lung insert, the imaging dose for the film plane was measured by Marants *et al.*<sup>169</sup> For the experiment using the liver insert, it was estimated using the same measured data but modified based on the differences of phantom materials (described in Appendix E) since the x-ray imaging setting was the same for lung experiments. When comparing the perturbed dose with the measured dose using liver and lung inserts, the imaging dose is subtracted from the measured dose.

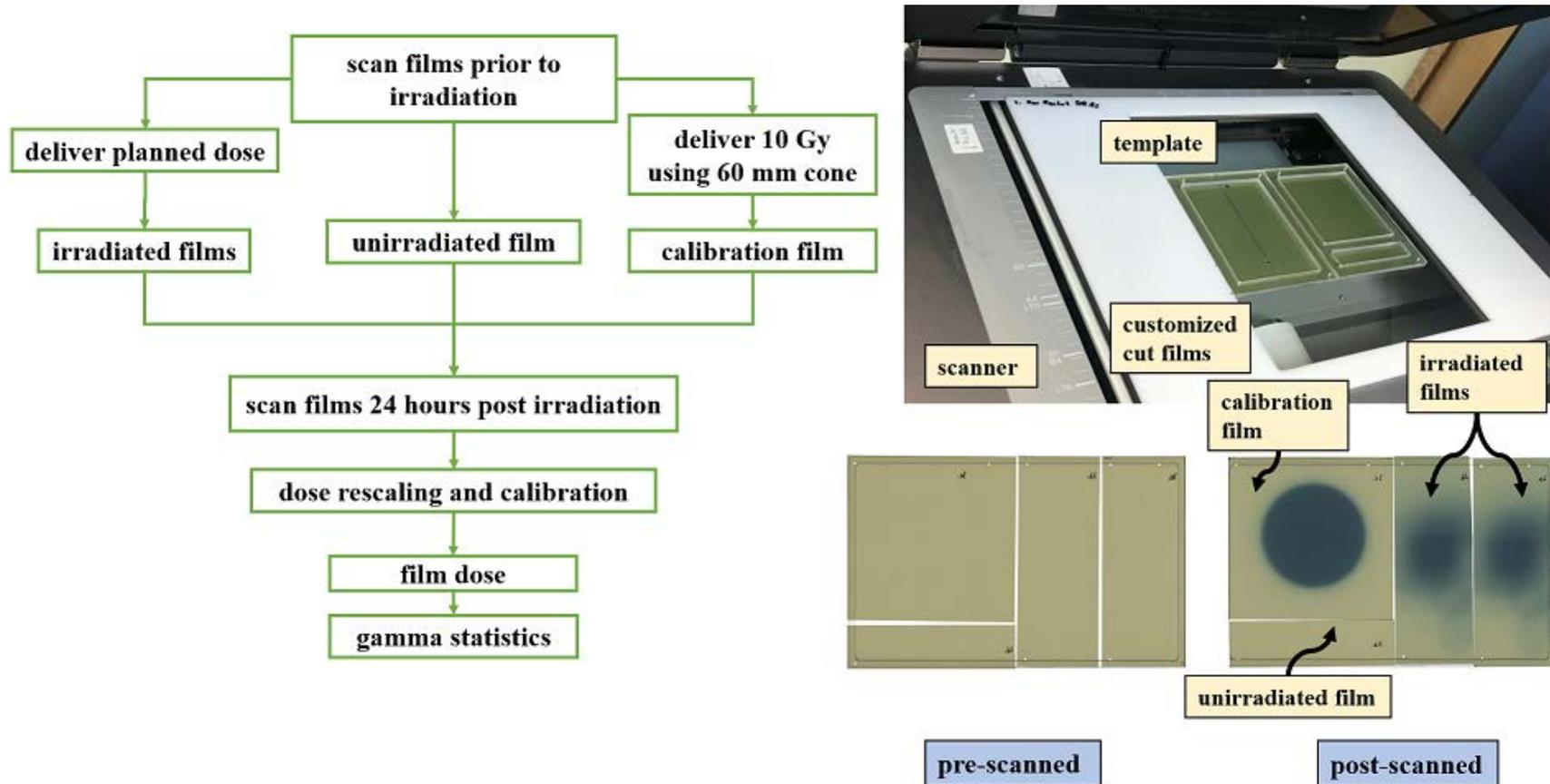


Figure 4-4. A flowchart of the film dosimetry using EBT3 films for CyberKnife treatments. Figures on the right illustrate the template (upper right corner) for stabilizing customized cut films before scanning, and scanned films (lower right corner) before and after irradiation.

#### **4.2.4. Dosimetric Perturbation Method**

For each treatment-course, the DICOM structure and dose files, the per-fraction treatment delivery XML file and other system log-files are extracted. In-house MATLAB scripts are employed to perturb the dose in a sub volume encompassing the films to account for different types of delivery errors. Most experiments in this chapter use a stack of two or more films which are averaged for analyses. The perturbed dose plane at each film position is calculated and averaged for the comparison. Each film plane position in the stack is assumed to be separated by the film thickness. Translational and rotational tracking residual errors derived from system delivery files are specified separately for treatments using Synchrony tracking (liver treatments) and 6D skull tracking (TN treatments).

##### **Translational Tracking Residual Errors**

For treatments using Synchrony tracking, correlation errors are determined for every x-ray acquisition. The predictor errors during beam-on and static phantom end-to-end errors from CyberKnife annual quality assurance measurements (2015 to 2019) are obtained. The average  $\pm$  SD of the end-to-end measurement using the body path are  $0.1 \pm 0.3$  mm,  $0.0 \pm 0.3$  mm, and  $-0.3 \pm 0.2$  mm, respectively, in the inferior-superior, left-right, and anterior-posterior directions. Using the trigeminal path, the average  $\pm$  SD are  $0.0 \pm 0.2$  mm,  $0.1 \pm 0.3$  mm, and  $-0.3 \pm 0.3$  mm, respectively (data acquired from 2013 to 2020). These three types of errors are considered as independent sources of translational tracking residual errors for liver treatments. The distributions of observed errors in two experiments are illustrated using histograms in Figure 4-5.

For trigeminal neuralgia treatments the CyberKnife system uses 6D skull tracking, in which the system compares the bone structures of live images with DRRs and calculates translational corrections (and rotational corrections which are not applied for the trigeminal path set) for every

x-ray acquisition.<sup>113</sup> The differences between the translational corrections from one image to the next approximate how much the patient moves in the time period between x-ray acquisitions. These data are used to generate a PDF of residual translational tracking errors for TN treatments. The magnitudes of these errors are mostly small (<1 mm) based on our observations for these patients.

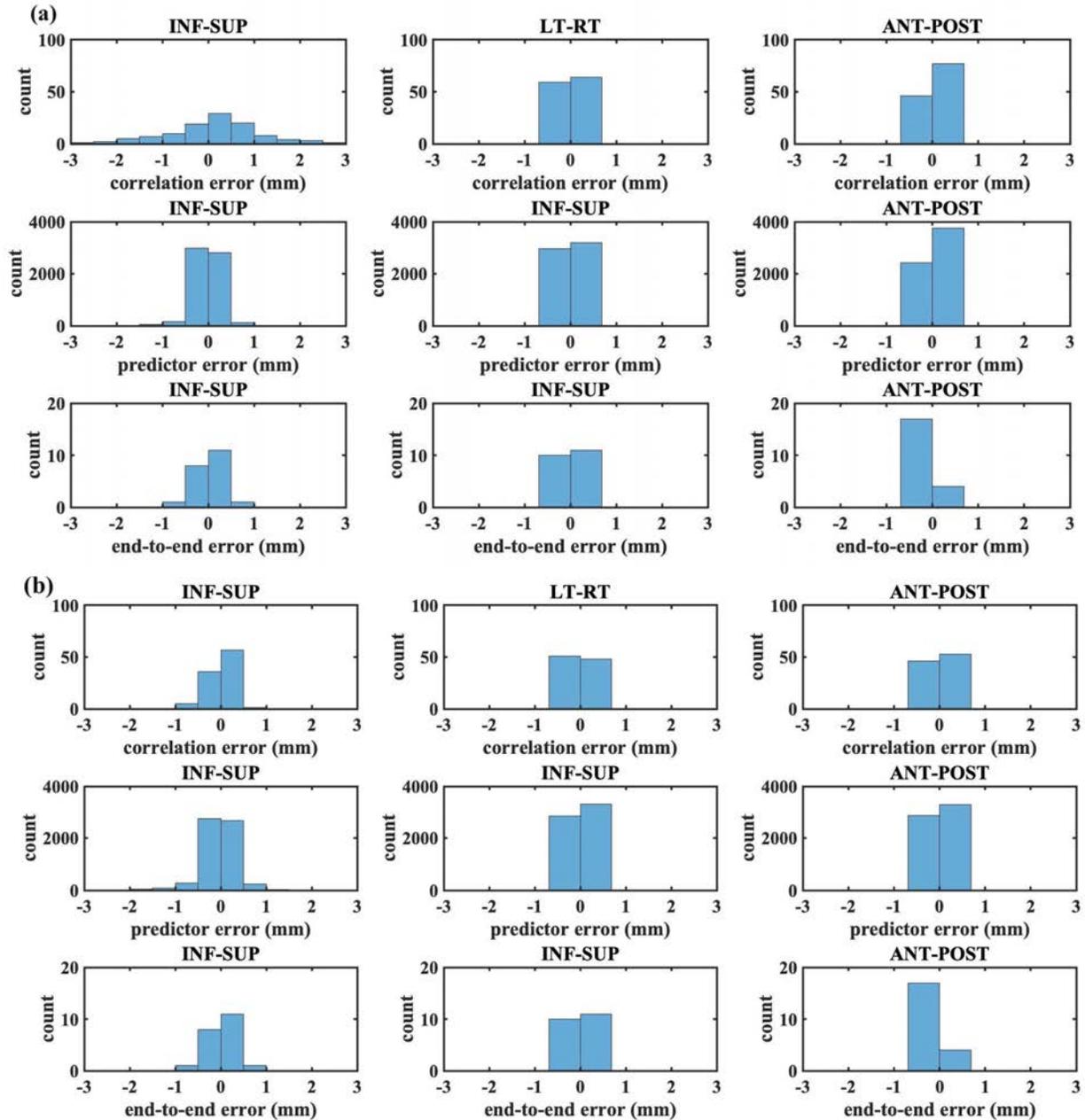


Figure 4-5. The distributions of correlation, predictor, and end-to-end errors in three anatomical directions observed in (a) the experiment “Liver-Far2Lung-1” and (b) the experiment “Liver-Far2Lung-2”, as labelled in Table 4-2. The bin size is chosen based on the resolution of the dose distribution. The probability distribution functions of three types of translational tracking residual errors for liver treatments are created using data.

### Rotational Tracking Residual Errors

The range of rotational corrections that can be applied during treatment using the body path and Synchrony tracking during liver treatments is  $\pm 1.5^\circ$  in roll and pitch and  $\pm 3^\circ$  in yaw. For treatments using the trigeminal path, rotations are not corrected for due to the potential for collision, as described in Sections 1.3.4. We consider a rotational error to be non-zero only if it is outside of the range of rotations that the system can correct. For example, if the system reports  $1^\circ$  in roll, pitch, and yaw for liver treatment (Synchrony tracking) and TN treatment (6D skull tracking using the trigeminal path), we consider the rotational error to be zero for the liver case, and non-zero for the TN case.

### Dose Perturbation – Neglecting Rigid Body Error

A flow chart demonstrating the algorithm that calculates the dose to target voxels is shown in Figure 4-6. Translational and rotational errors are modelled in two separate steps as described below.

(1) The first step is the spatial convolution of the planned dose using a probability distribution function (PDF) of translational residual errors. For liver treatments, correlation, predictor, and end-to-end errors are binned into three separate spatial PDF matrices using bin sizes that are the same size as the dose matrix voxel size. We convolve the planned dose with these three probability matrices to calculate the perturbed dose. Similarly, for TN treatments we convolve the planned dose with the PDF matrices of the tracking residual errors and end-to-end errors.

(2) Secondly, the new positions ( $\mathbf{X}_j, j = 1, 2, \dots, N$ ) of each target voxel are determined using the rotation matrix described in Appendix B, where  $\mathbf{X}_j$  is the vector position of the voxel due to the  $j^{\text{th}}$  observed rotational error. The accumulated dose for a voxel is calculated as  $\frac{\sum_{j=1}^N D_p(\mathbf{X}_j)}{N}$ , where  $D_p(\mathbf{X}_j)$  represents the perturbed dose for the voxel determined using a nearest neighbour

approximation, and  $N$  indicates the number of rotational errors assuming each perturbation is equally weighted. The angle of rotation determined by an x-ray image acquisition shortly before each beam delivery is the correction used by the CyberKnife robot manipulator, which means per-beam rotational errors are sampled slightly less frequently than the per x-ray translational tracking residual errors used in the first step.

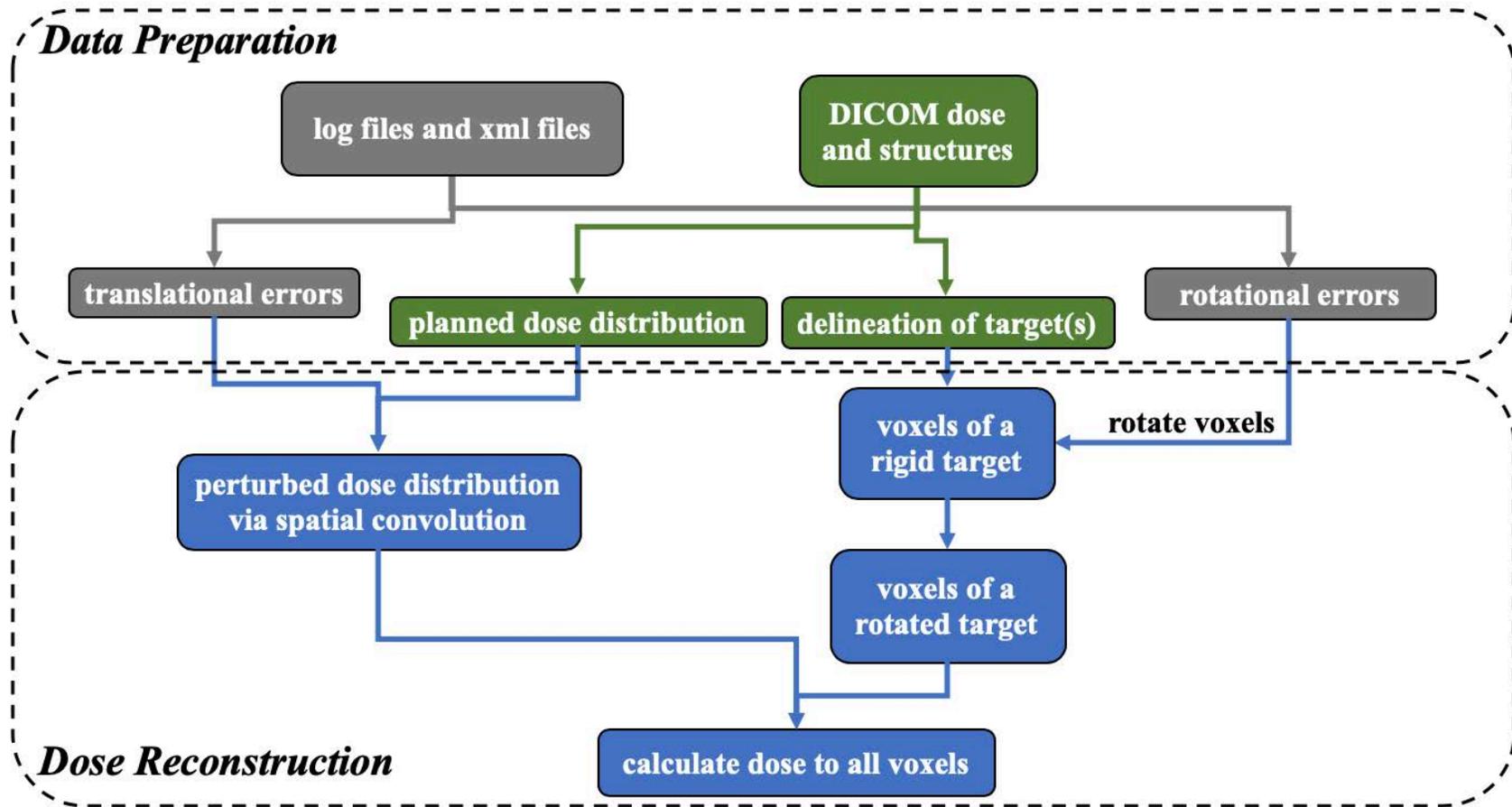


Figure 4-6. The flowchart illustrates the dose perturbation tool accounting for intrafraction translational and rotational errors.

### **Dose Perturbation Including Rigid Body Errors**

For the experimental validation in this chapter, only rigid phantoms are used, therefore deformation is not modelled. For the patient studies in the next chapter, a simple model of deformation is used and described in this section. Target contours are either expanded or contracted isotropically based on the average value of the signed rigid body error (see Section 2.2.4) calculated for each x-ray image. The workflow to account for all three types of delivery errors in dose reconstruction is shown in Figure 4-7.

For a faster calculation, the signed rigid body errors are binned using the smallest dose matrix resolution in one dimension. Targets of different sizes based on binned rigid body errors are precalculated prior to dose perturbations considering translations and rotations as described in the previous section. We determine the dose to targets of different sizes and weigh the dose contributed from each deformation using the probability of the bin. For example, if the spatial bins for rigid body errors are -0.5, 0, 0.5 mm with the probability of 0.3, 0.2, and 0.5, respectively, the accumulated dose to a voxel is  $0.3D_{-0.5\text{mm}} + 0.2D_0 + 0.5D_{0.5\text{mm}}$ , where  $D_{\text{bin value}}$  represents the perturbed dose matrix values that account for rotational and translational errors to the voxel locations which have been isotropically shifted closer or further away from the center of the target. This simple model provides an approximate estimate of the dose to a non-rigid target which will be used to retrospectively estimate dose to targets for patient treatments in Chapter 5.

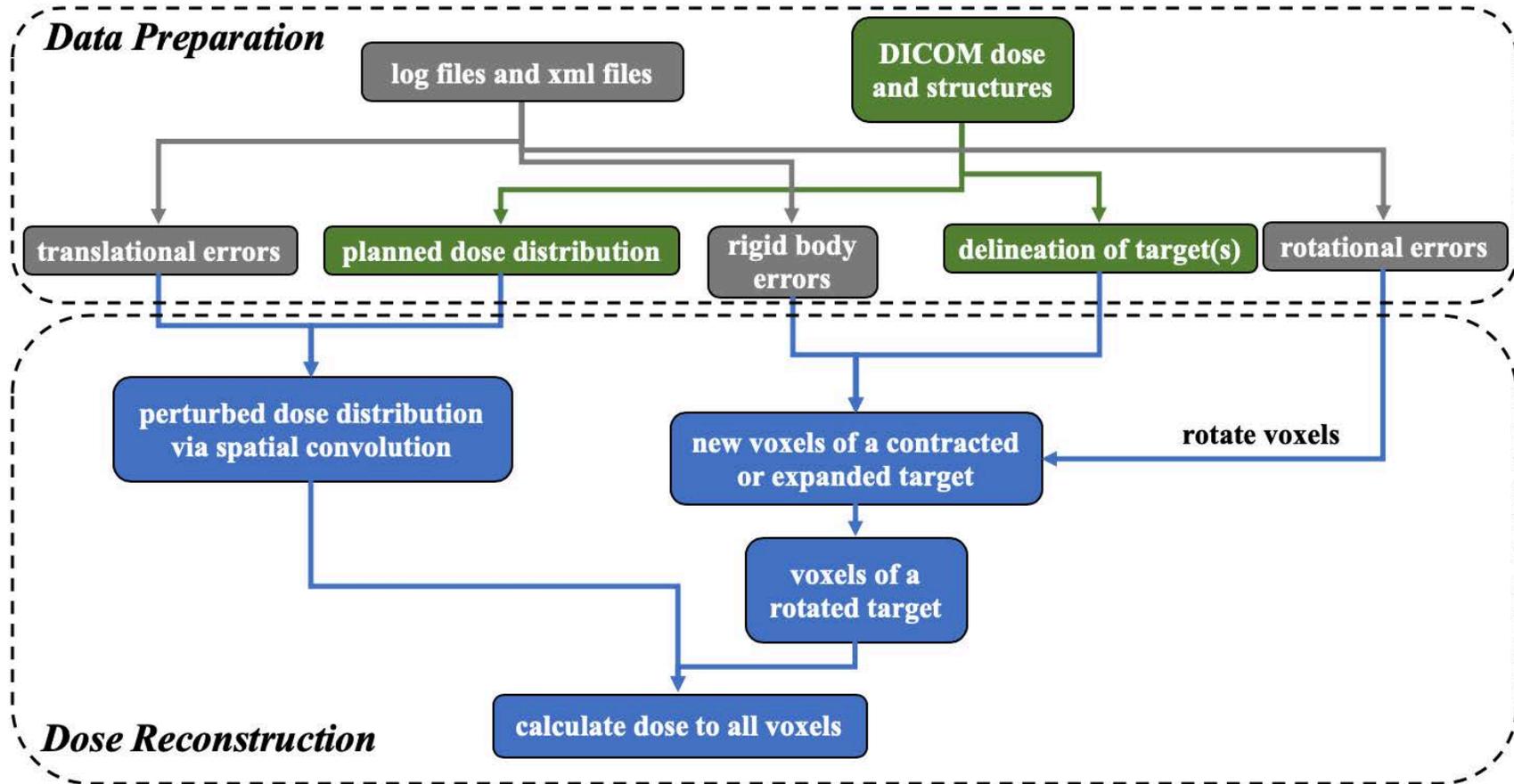


Figure 4-7. The flowchart illustrates the dose perturbation tool for targets taking into account rigid body errors.

#### 4.2.5. Dose Penumbra Broadening Effects

As recommended in the Canadian Partnership for Quality Radiotherapy (CPQR) report, the annual testing for Synchrony motion tracking should include an analysis of the penumbral width of the dose profile for out-of-phase internal and external motions.<sup>30</sup> To investigate how well the perturbed dose models the profile broadening effects associated with delivery errors, line profiles for measured and perturbed dose along the motion direction are compared. Figure 4-8 illustrates line profiles for the planned, perturbed, and measured dose for a phantom experiment described by Marants *et al.*<sup>169</sup> There is a measurable shift of dose profiles, and dose penumbra broadening is visible in the measured profile.

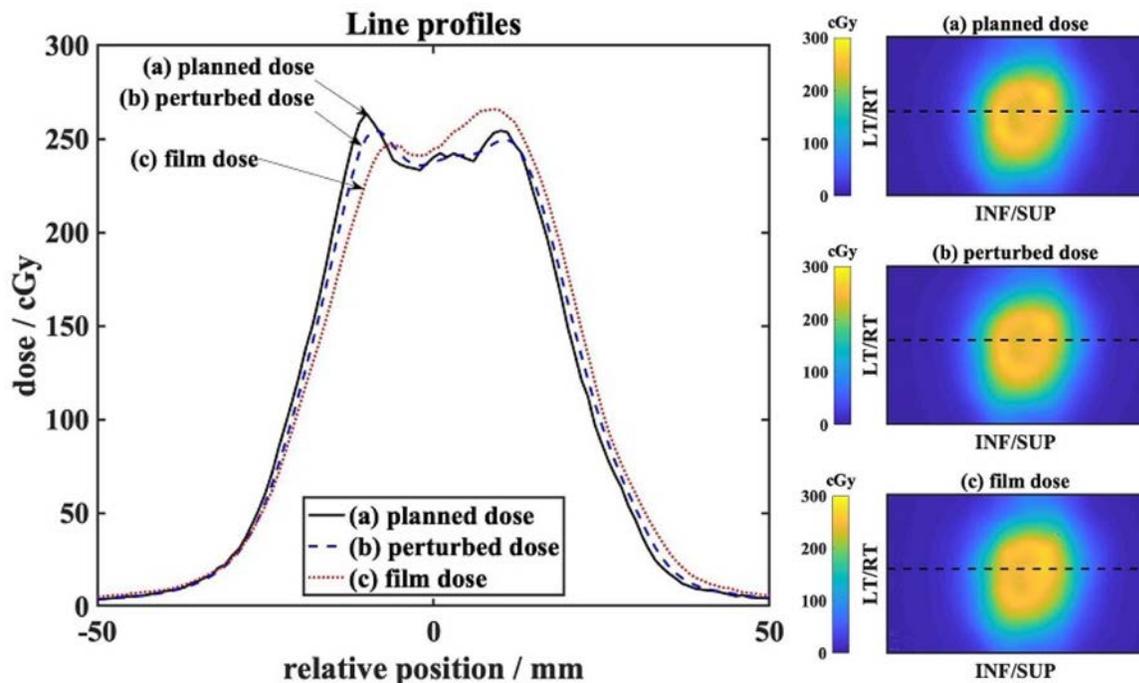


Figure 4-8. Line profiles for the planned, perturbed, and measured dose. The profiles are taken along the direction of the phantom motion, which is SUP/INF for this experiment.

The penumbral width of the line profile can be quantified as the lateral distance between 80% and 20% of the maximum dose.<sup>228,229</sup> As illustrated in Figure 4-9, we define the penumbral width as  $\frac{1}{2}(\text{FW}20\%-\text{FW}80\%)$ , half the difference between the full width at 80% (FW80%) and 20%

(FW20%) of the maximum dose. The line profile is selected that passes through the centre of the film. The penumbral width is rounded to the nearest millimetre because of the limited resolution in dose calculation.

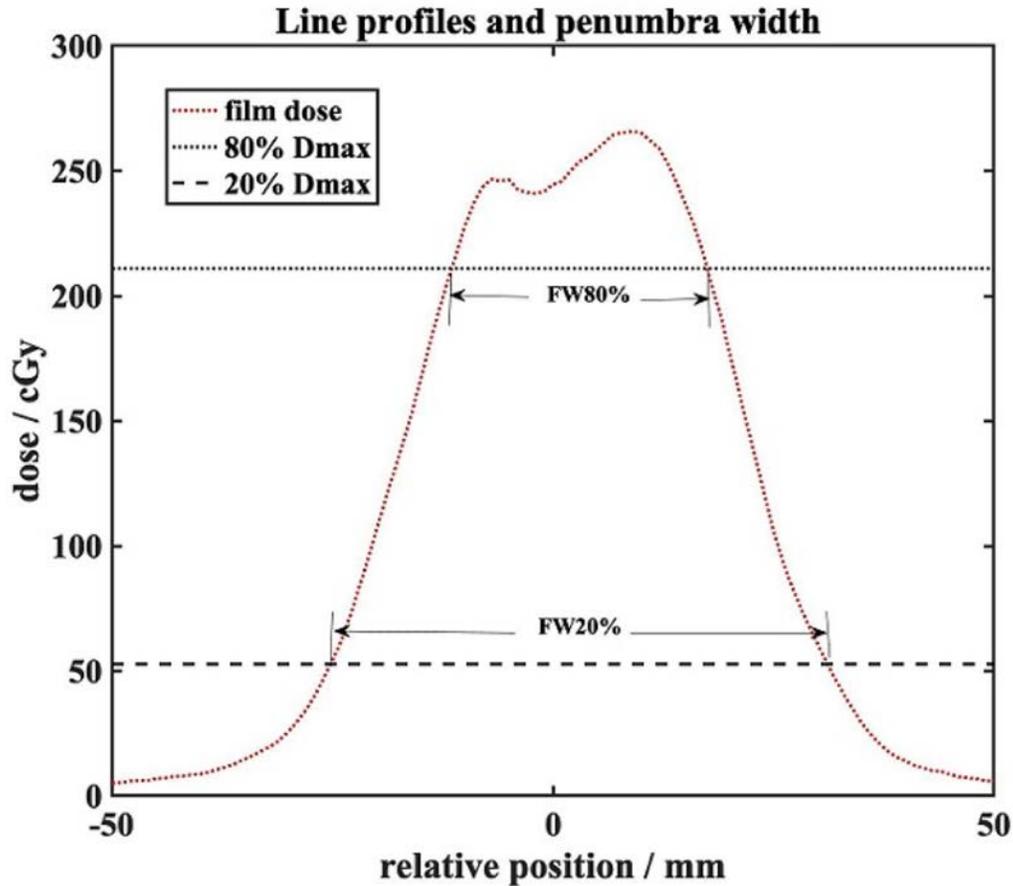


Figure 4-9. Demonstration of a line profile of a dose measurement. The FW80% and FW20% indicate the distances between the 80% isodose points, and that between 20% isodose points, respectively. The penumbral width for the dose broadening effect is defined as  $\frac{1}{2}$  (FW20% - FW80%).

#### 4.2.6. Experimental Validations

Our dose perturbation model was validated using film measurements in static phantoms and a programmable respiratory motion phantom. The agreement between the planned or the perturbed dose and the reference dose (measured) is quantified by the gamma pass rate. Three scenarios were considered to validate the model: (1) rotational offset only using the rocket phantom;

(2) translational errors only using the ball cube, lung insert, and liver insert; and (3) rotational offset and translational errors using the liver insert.

In Section 2.3.2, we demonstrated that uncorrected rotations could have a significant impact on the delivered dose in some clinical cases. In this chapter, we present experimental confirmation of our model of uncorrected rotations including the handedness of the coordinate system and the order of application for Euler angle rotation matrices (see Appendix B). Three experiments, in increasing order of complexity, were completed using the *static* rocket phantom for film measurements. The planned delivery is non-isocentric, with a maximum dose equal to 10 Gy. The phantom was manually positioned with an intentional rotational offset. The distance between the imaging isocenter (the origin about which rotational corrections are calculated) and the target must be reasonably large so that the effect of the uncorrected rotation has a measurable change to the delivered dose. At least four fiducials were used for tracking to estimate the phantom rotation accurately.<sup>142</sup> The rotational correction was either manually disabled when the rotation was small that could be corrected by the system or larger than what the system could correct.

For experiments that model residual tracking errors alone during treatments, the ball cube and the lung insert were used without a rotational offset. The systematic shift between the perturbed (or planned) and the measured dose was determined by comparing the dose-weighted centroids for those voxels in the film plane with  $> 80\%$  of the maximum dose in the planned dose distribution. The penumbral width (see Section 4.2.5) is also compared for planned, perturbed, and measured dose profiles.

Two experiments were conducted using the ball cube phantom with the Synchrony platform providing out-of-phase internal and external motion as shown in Figure 4-2. Both internal and external motions were periodic and approximately sinusoidal with the internal motion in the SUP-

INF direction. Two orthogonal EBT3 films (one in the transverse and one in the sagittal plane) were exposed during the E2E test, with the 35 mm diameter collimator using isocentric delivery. Only the sagittal plane film was considered for this analysis since it was oriented along the direction of motion. The maximum planned dose was 6 Gy for both experiments. The imaging dose was not accounted for in these experiments.

Five experiments were previously conducted using the lung insert, and the details of the experimental setup were described by Marants *et al.*<sup>169</sup> In their study, the maximum planned dose was 2.7 Gy to minimize the degradation of sensitivity associated with cumulative dose to the MOSFET dosimeter (part of the RADPOS system<sup>56</sup>). All measured data were corrected based on separate film measurements made in the same phantom and imaging settings (130 kVp, 200 mA, 80 ms) used in the experiment. The high dose was delivered to the Solid Water sphere which emulated a lung tumour surrounded by lower-density tissue. They compared the planned and measured dose using 2D gamma analyses (either 3%/1 mm or 3%/2 mm) and found the gamma failures were mostly around the edges of the high dose region for isocentric plans and were caused by dose broadening and a translational offset along the motion direction. It was hypothesized that the complexity of non-isocentric dose delivery would be more likely to lead to gamma failures in the high dose area. We reanalyzed their measured data and compared it with the perturbed dose. These experiments represented dose delivery to heterogeneous materials (cedar and Solid Water). Three of them used the same isocentric plan, and the phantom was driven using a sinusoidal waveform with different phase shifts. For the remaining two experiments, the dose was delivered using the same non-isocentric plan, and the phantom was set in motion using a waveform derived from the CyberKnife system log files for a lung patient.

For the lung insert experiment, the accuracy of the dose broadening generated by our perturbation model may be influenced by the presence of the low-density material surrounding the high dose region. For liver lesions, heterogeneity is usually less of a concern since the liver is a large organ with its density close to that of water, unless the tumour is at the dome of the liver and close to the diaphragm and lung tissue. Six experiments were conducted using the liver insert driven by a Quasar in the inferior-superior direction with the high dose region (> 50% of the maximum) either completely inside Solid Water or abutting the Styrofoam region of the liver insert. For in-phase motion, markers were placed on the same Quasar platform. The out-of-phase motion was completed using two Quasars as described in Section 4.2.2. The same patient waveform was used for all experiments (illustrated later in Figure 4-19, Section 4.3.2). The same imaging setting (130 kVp, 200 mA, 80 ms) for lung experiments was implemented, so that we could estimate the imaging dose for liver experiments using previously measured data.<sup>169</sup> The average imaging doses for the Solid Water region and the Styrofoam region for each x-ray image pair were estimated to be 0.028 and 0.078 cGy, respectively, and was subtracted from the measured dose distribution as described in Appendix E.

Two non-isotropic plans are used: high dose near the Solid Water and Styrofoam boundary (denoted as “close to lung”), and high dose in the Solid Water (denoted as “far to lung”). The center of high dose on the film plane is 1.7 cm and 3.0 cm away from the rotation axis, respectively, for the close-to-lung and far-to-lung plans. The liver phantom was moving in the superior-inferior direction for four experiments. For the remaining two experiments, we positioned the phantom with a rotational offset (2° and 4°) while the phantom was moving in one direction horizontally. The 2° rotational offset can be corrected by the system, which is equivalent to a zero rotational

offset in our dose perturbation model. In contrast, the  $4^\circ$  rotation is larger than what the system can correct.

### 4.3. Results

#### 4.3.1. Rotational Errors

##### Uncorrected Rotations for Trigeminal Neuralgia Treatment

For trigeminal neuralgia, the target is usually very small, and the planned delivery typically uses only the 5-mm collimator. As mentioned previously, rotational corrections are not applied by the system when the trigeminal path is used. The phantom was positioned with a rotational offset of  $2.8^\circ$  in yaw, and the rotational axis was 10 cm away from the centre of the high dose region, as illustrated in Figure 4-10. Two films were used for measurements leaving a small air gap (0.6 mm thick) in the film slot. The experiment was designed to have a very noticeable offset from the planned dose due to uncorrected rotations.

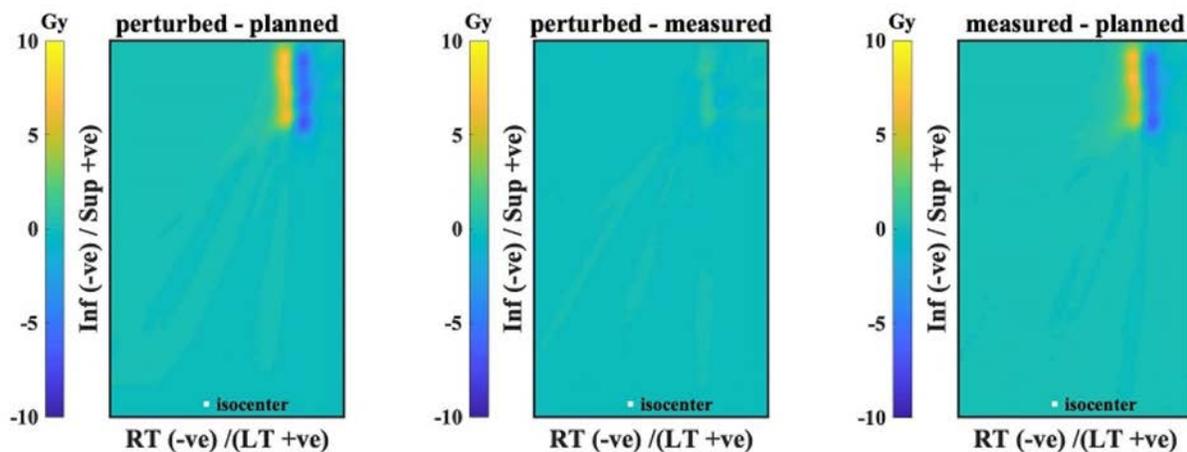


Figure 4-10. The dose difference for the planned, perturbed, and measured dose for a trigeminal neuralgia plan delivered using 5 mm cone beams. The white dot is the center about which the system calculates rotational corrections. A  $2.8^\circ$  offset about the rotation centre was present in the phantom setup, and the rotation was not corrected.

## Chapter 4. Experimental Validations of the Dose Perturbation Tool

As summarized in Table 4-1, the 2%/2mm criteria gamma pass rate increases from  $\gamma_0 = 90.0\%$  when the planned dose is compared with the measured dose to  $\gamma_1 = 99.5\%$  when the perturbed dose is compared with the measured dose. For a small target away from the origin within a high dose gradient, a relatively small uncorrected rotation can make a significant difference in the spatial distribution of absorbed dose. Figure 4-11 illustrates how well the perturbed dose models the impact of an uncorrected rotational offset. Most gamma failures are within the high dose area (very small area). The 2%/2mm gamma pass rates for perturbed and measured dose within the 80% Rx, 70% Rx, 10-70% Rx isodose lines are 75.4%, 87.5%, and 100.0%, respectively.

The imaging dose was negligible compared to the planned dose, since only 21 x-ray images were taken (x-ray parameters: 100 kVp, 40 mAs). If we consider 0.063 cGy per x-ray image pair (100 kVp, 40 mAs) surface dose provided by the CyberKnife imaging whitepaper,<sup>227</sup> the total imaging dose is 0.1% of the maximum planned dose (10 Gy).

Table 4-1. Characteristics of experiments using the rocket phantom to investigate impacts of uncorrected rotations on dose distribution. The phantom was static and positioned with a rotational offset for each experiment during beam delivery. The path set used was either the trigeminal or the body path. The agreement between the planned or the perturbed dose and the reference dose (measured) is quantified by the gamma pass rate.

label	film orientation	tracking method	path set	# of films	# of x-ray images	rotational offset (degree)			$\gamma_0$	$\gamma_1$
						roll	pitch	yaw		
R1	coronal	fiducial	trigeminal	2	21	0.0	0.0	2.8	90.0	99.2
R2	coronal	Synchrony	body	4	157	-0.5	-0.6	-3.7	82.4	99.6
R3	sagittal	Synchrony	body	4	36	3.1	-2.7	3.1	90.1	100.0

$\gamma_0$ : gamma pass rate (%) comparing planned and measured dose using 2%/2mm criteria

$\gamma_1$ : gamma pass rate (%) comparing perturbed and measured dose using 2%/2mm criteria

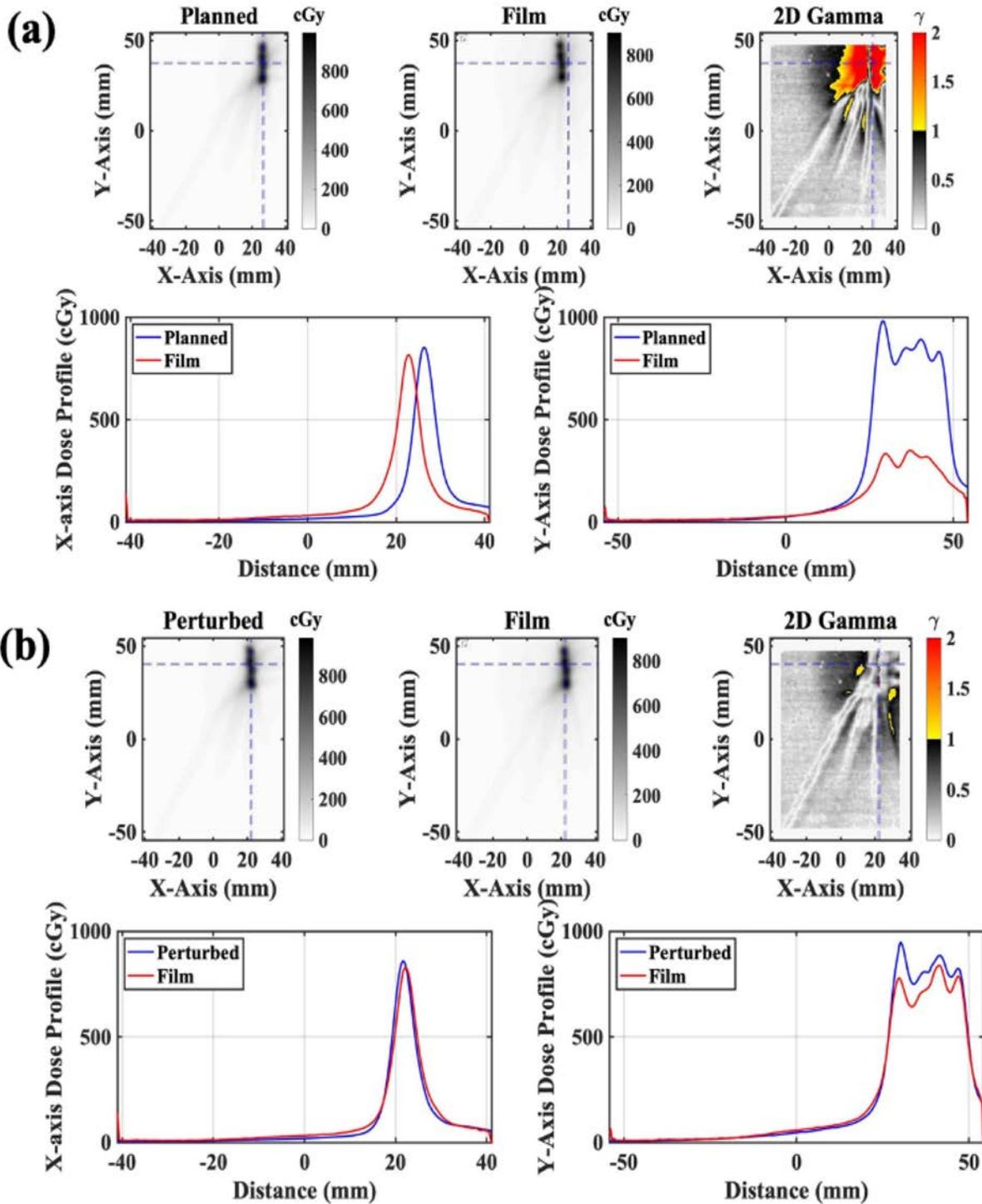


Figure 4-11. Two-dimensional gamma analyses (2%/2mm) on (a) planned and measured dose distribution for experiment R1, and (b) perturbed and measured dose distribution. The gamma pass rates for all non-zero pixels are 90.0% (a) and 99.2% (b).

### Uncorrected Rotations for Treatment Using Synchrony Tracking

A similar experiment was performed using Synchrony tracking and the body path to ensure that the system handles uncorrected rotations consistently for different tracking methods and path sets. The rocket phantom was placed with a  $3.7^\circ$  offset in yaw, and four (two dummy and two unirradiated) films pieces were irradiated without an air gap. Although Synchrony tracking was used for the delivery, the phantom together with its fiducials and the films all remained stationary, and an external breathing pattern was provided by LEDs placed on the Synchrony motion platform. This resulted in a “flat” correlation model with a negligible amplitude for internal motion. Since the tracking system corrects for a maximum angle of rotation of  $\pm 3^\circ$  in the yaw direction for the body path, this rotational offset will therefore be non-zero in the perturbation model. The dose was optimized to an “L”-shape target located about 10 cm away from the rotation centre using a non-isocentric beam set and the 7.5 mm and 12.5 mm collimators. The dose difference between the planned, perturbed, and measured dose is shown in Figure 4-12.

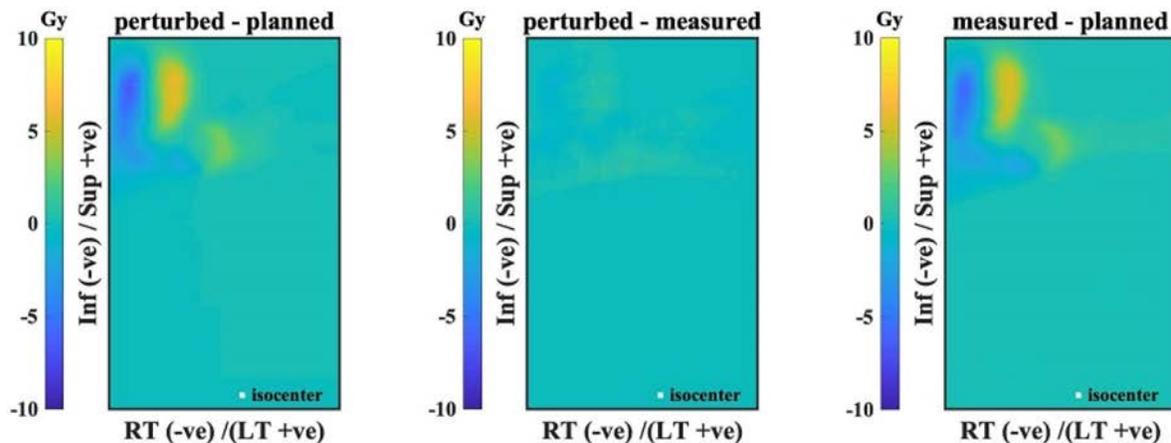


Figure 4-12. Two-dimensional dose difference for planned, perturbed, and measured dose of an “L”-shape target (experiment R2). The white dot indicates the point about which the system calculates rotation.

As summarized in Table 4-1 (experiment R2),  $\gamma_0$  and  $\gamma_1$  are 82.4% and 99.6%, for the comparison with planned and perturbed dose, respectively. The gamma pass rates for the comparison of perturbed and measured dose within the 80% Rx, 70% Rx, 10-70% Rx isodose lines are 99.9%, 100.0%, and 99.8%, respectively.

In this experiment, 157 x-ray images were taken (90 kVp, 40 mAs). If we consider 0.063 cGy per x-ray image pair (100 kVp, 40 mAs) surface dose provided by the CyberKnife imaging whitepaper,<sup>227</sup> the total imaging dose is just less than 1.0% of the maximum planned dose (10 Gy). The imaging dose is therefore considered negligible compared to the planned dose for this experiment.

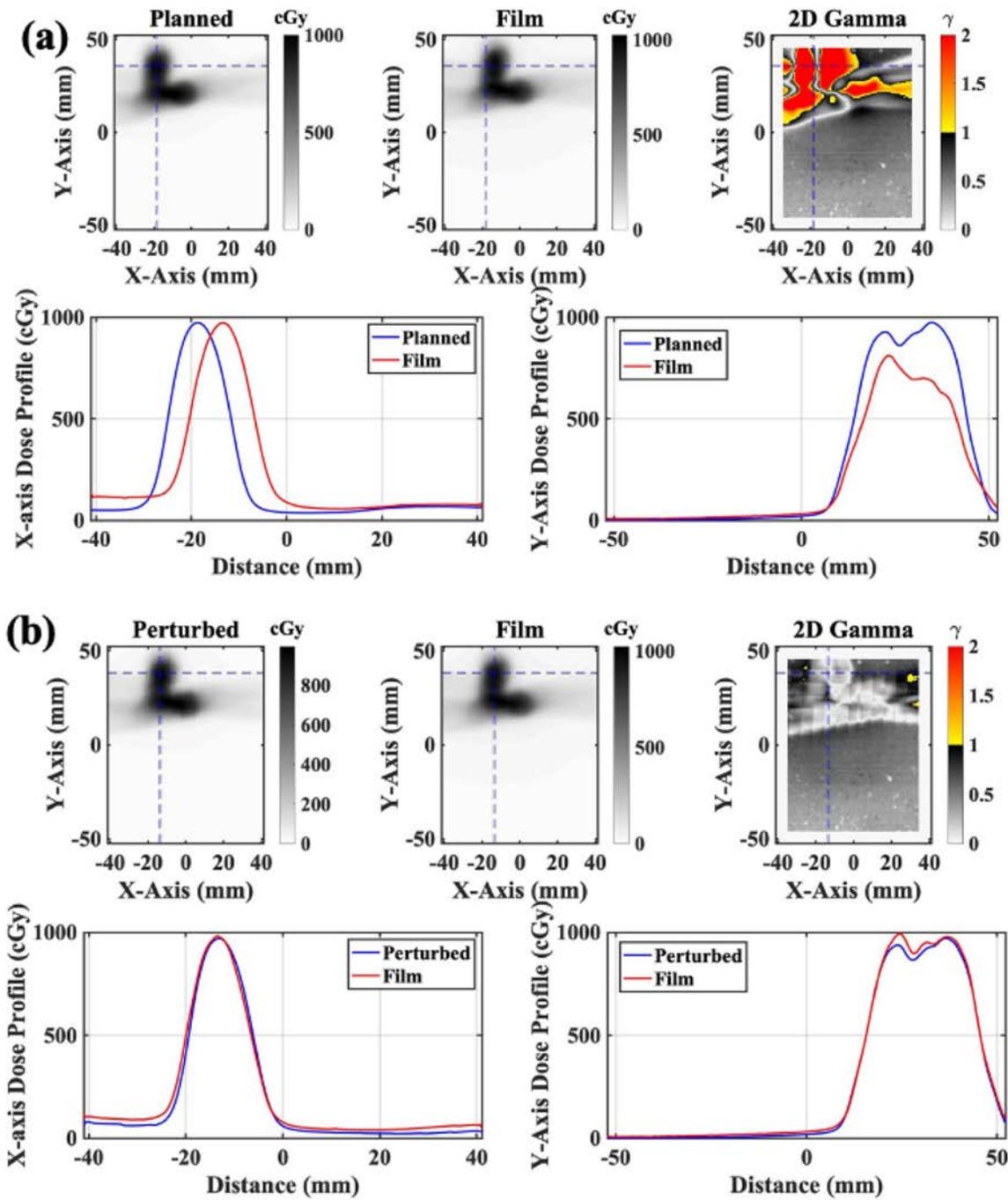


Figure 4-13. (a) Two-dimensional gamma analyses (2%/2mm) on planned and measured dose distribution for experiment R2, and (b) perturbed and measured dose distribution. The gamma pass rates (2%/2mm) for all non-zero pixels are 82.4% (a) and 99.6% (b).

### Rotations About Multiple Axes for Treatment Using Synchrony Tracking

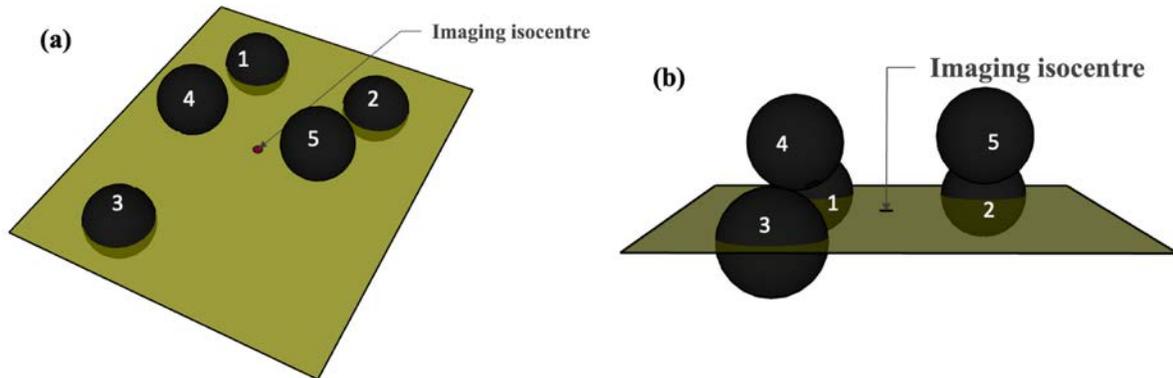


Figure 4-14. (a, b) Three (no. 1, 2 and 3) and two (no. 4 and 5) spherical targets, respectively, planned at the film plane and out-of-plane. The red dot indicates the imaging isocenter. (b) If the film plane is rotated counterclockwise and the system does not correct for the rotational offset, high dose of sphere no.5 will be delivered on the film plane. Diagrams completed using Sketchup 3D modeling software.

To confirm the assumed order of operations for rotations about different axes, a plan was generated for five targets using the 10 mm collimator with isocentric delivery. The plan was designed to be intentionally sensitive to the order of rotation about the different axes. For three of the targets, the treatment centres were placed in the film plane, and for the other two, were placed out of plane away from the image isocenter, as shown in Figure 4-14. The phantom was positioned with rotational offsets about all three axes (roll =  $3.1^\circ$ , pitch =  $-2.7^\circ$ , yaw =  $3.1^\circ$ ) relative to how it was positioned in the CT scanner. These rotations were greater than what could be corrected by the system. The difference would be detectable using a 2%/2mm  $\gamma$  criteria if the rotations were applied in different orders or with different signs.

As shown in Figure 4-15(a), the percentage difference between the planned and measured dose can be as significant as  $\pm 30\%$ . The dose difference between the perturbed and measured dose is significantly reduced shown in Figure 4-15 (b). The gamma pass rate increases from  $\gamma_0 = 90.1\%$  to  $\gamma_1 = 100.0\%$ . The gamma pass rates for perturbed and measured dose within the 80% Rx,

70% Rx, 10-70% Rx isodose lines are 100.0%, 99.7%, and 100.0%, respectively. The  $\gamma$ -pass rates for the perturbed dose are reduced for any other conventions we considered for applying the rotational corrections.

The imaging dose for this experiment is also negligible compared to the planned dose. In this case, 36 x-ray images were taken (120 kVp, 10 mAs). If we consider 0.024 cGy per x-ray image pair (120 kVp, 10 mAs) surface dose provided by the CyberKnife imaging whitepaper,<sup>227</sup> the total imaging dose is 0.1% of the maximum planned dose (10 Gy).

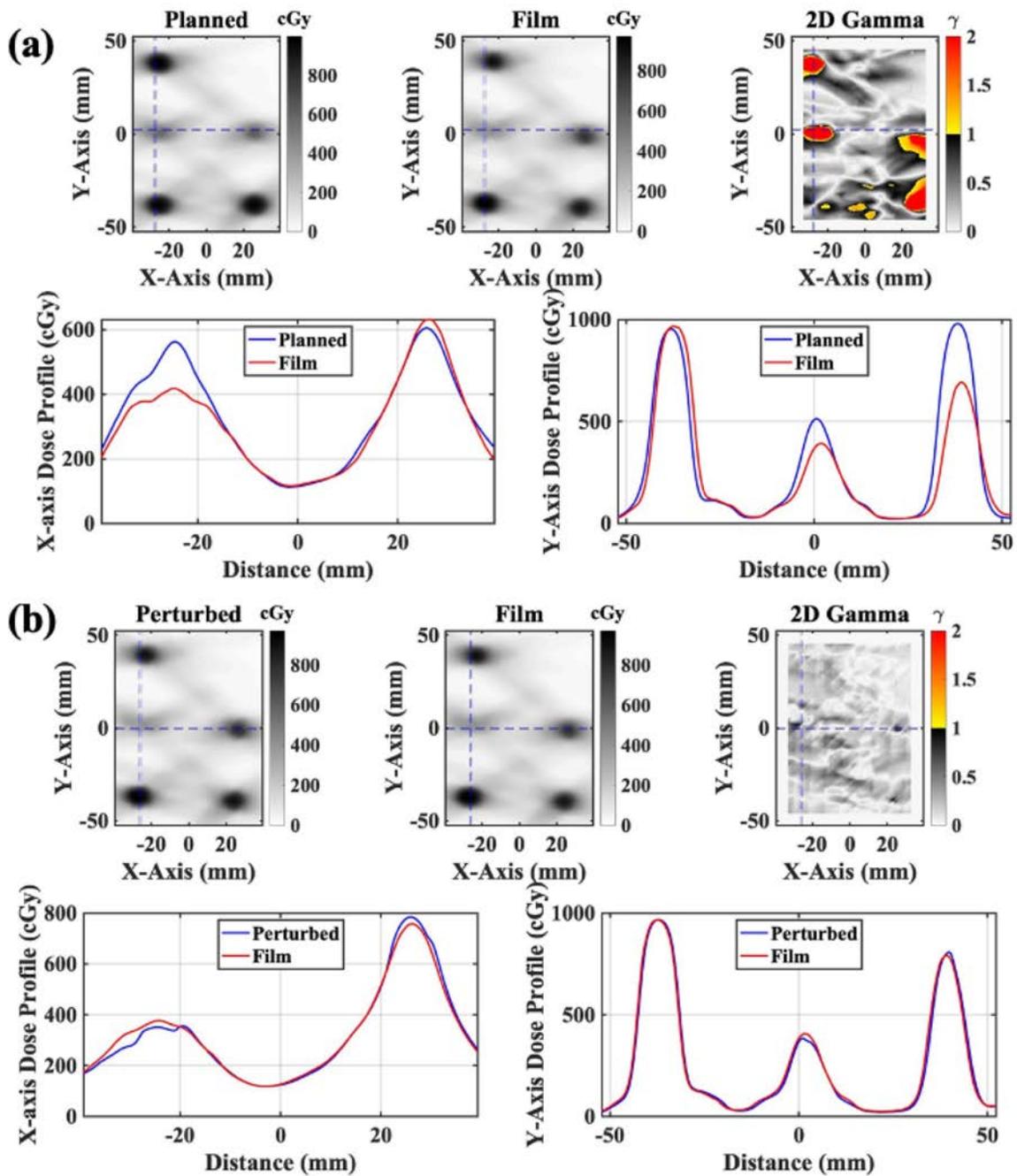


Figure 4-15. Two-dimensional gamma analyses (2%/2mm) on (a) planned and measured dose distribution, and (b) planned and measured dose distribution for experiment R3. The gamma pass rates (2%/2mm) for all non-zero pixels are 90.1% (a) and 100.0% (b).

### 4.3.2. Tracking Residual Errors

#### Using the Ball Cube

Some characteristics, including the nominal or calculated phase shift for internal and external motion, the tracking errors, the penumbral width, and the gamma pass rates are listed in Table 4-2. The table also includes a label to refer to each experiment in the following discussion. For the Ball cube experiments, the tracking error is small due to the simplicity of the phantom motion and small phase shifts (20° or 30°). The dose broadening effect was not apparent because the measured penumbral width increased by only about 1 mm for both experiments, compared to that of the planned dose. The penumbral widths of the perturbed dose were consistent with the measurements. Stricter gamma criteria (1%/1 mm) were selected because of the high gamma pass rate for comparing the planned and measured dose. For both experiments,  $\gamma_1$  are increased by about 1%-points on average compared to  $\gamma_0$ .

Figure 4-16 (a) shows a comparison of the planned dose (no perturbation for delivery errors) and measured dose for one of these experiments (denoted as Ball1 in Table 4-2). We found most gamma failures were around the dose falloff region along the motion direction (Y-axis in the figures). The gamma failures were reduced in the perturbed dose (fewer failures near the edges). A slight difference in the systematic dose shift (-0.5 mm) was observed for Ball1 as shown in Figure 4-16 (b).

The imaging dose is negligible compared to the planned dose for this experiment, with only 36 x-ray images were taken (100 kVp, 8 mAs). If we consider 0.015 cGy per x-ray image pair (100 kVp, 10 mAs) surface dose provided by the CyberKnife imaging whitepaper,<sup>227</sup> the total imaging dose is 0.1% of the maximum planned dose (6 Gy).

Chapter 4. Experimental Validations of the Dose Perturbation Tool

Table 4-2. Characteristics of experiments used to validate the dose perturbation model. For all these experiments, a phantom was moving on a platform while the beam was delivered. For two of these experiments, the phantom was also positioned with a rotational offset. The path set used was the body path for all experiments. The agreement between the planned or the perturbed dose and the reference dose (measured) is quantified by the gamma pass rate.

label	phantom	number of films	plan complexity	phase shift (degrees)	# of x-ray images	correlation error (mm)	SD of predictor errors (mm)	rotational offset (degrees)	gamma criteria	$\gamma_0$	$\gamma_1$	$s_0$ (mm)	$s_1$ (mm)	penumbral width (mm)						
														planned	perturbed	measured				
Ball1	ball cube	1	isocentric	20	45	$-0.0 \pm 0.8$	0.2	N/A	1%/1mm	97.4	98.1	-0.4	-0.5	5	6	6				
Ball2		1		30	34	$-0.2 \pm 0.7$	0.2			95.0	96.7	0.3	0.2	6	7	7				
Lung1	lung insert	2	isocentric	0	35	$-0.0 \pm 0.6$	0.2	N/A	2%/2mm	99.5	99.7	-0.3	-0.3	8	8	9				
Lung2		1		43 <sup>†</sup>	38	$0.0 \pm 1.7$	0.8			84.1	89.5	-0.3	-0.2	8	9	12				
Lung3		2		20 <sup>†</sup>	34	$0.0 \pm 0.7$	0.2			98.4	98.8	-0.7	-0.6	8	8	9				
Lung4	lung insert	2	non-isocentric	25 <sup>†</sup>	96	$0.4 \pm 2.2$	0.5	N/A	2%/2mm	95.5	96.8	0.6	0.3	13	14	14				
Lung5		2		18 <sup>†</sup>	70	$0.3 \pm 1.8$	0.5			86.7	96.3	-2.1	-1.7	14	14	14				
Liver-Far2Lung-1	liver insert	2	non-isocentric	0	123	$0.0 \pm 0.3$	0.4	N/A	2%/2mm	95.4	96.0	0.2	0.4	11	12	12				
Liver-Close2Lung-1		2		0	139	$-0.0 \pm 0.2$	0.4		N/A	2%/1mm	92.1						93.5	2%/2mm	97.1	98.2
Liver-Far2lung-2	liver insert	2	non-isocentric	40 <sup>†</sup>	155	$-0.5 \pm 2.1$	0.4	N/A	2%/2mm	93.2	98.8	2.0	1.6	12	13	14				
Liver-Close2Lung-2		2		17 <sup>†</sup>	112	$0.1 \pm 0.7$	0.3		N/A	2%/2mm	95.3						97.6	2%/1mm	85.0	94.0
Liver-Far2Lung-3		2		51 <sup>†</sup>	140	$-0.6 \pm 2.2$	0.6		4.0	2%/2mm	76.9						99.9	2%/1mm	63.0	96.8
Liver-Far2Lung-4	liver insert	2	non-isocentric	47 <sup>†</sup>	146	$-0.3 \pm 2.1$	0.6	2.0 <sup>‡</sup>	2%/2mm	99.1	99.6	0.3	0.1	12	13	12				
		2								2%/1mm	95.6						98.1			

$\gamma_0$ : gamma pass rate (%) comparing planned and measured dose (all non-zero voxels)

$\gamma_1$ : gamma pass rate (%) comparing perturbed and measured dose (all non-zero voxels)

$s_0$ : systematic shift of dose along phantom motion direction comparing planned and measured dose

$s_1$ : systematic shift of dose along phantom motion direction comparing perturbed and measured dose

<sup>†</sup>: phase shift calculation as described in Appendix D (nominal phase shift otherwise)

<sup>‡</sup>: a rotational offset that can be corrected by the system (equivalent to zero degree for the resulting rotational offset)

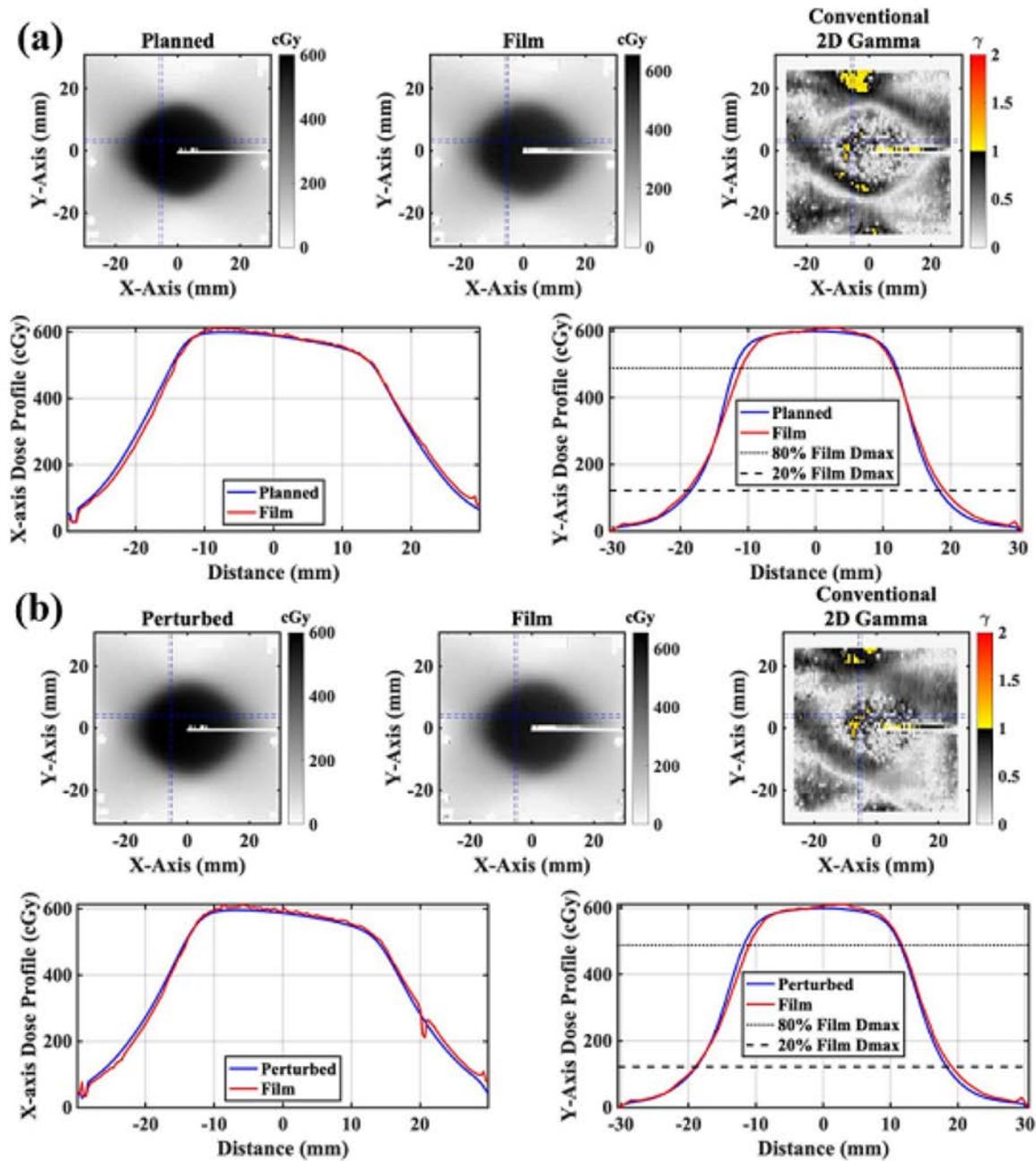


Figure 4-16. 2D gamma analyses (1%/1mm) on (a) planned and measured dose distribution, and (b) perturbed and measured dose distribution, for an isocentric plan using a ball cube phantom (experiment Ball1). The phantom was set in motion along Y axis and a cut film was loaded at the sagittal plane. The colored pixels indicate gamma failures. The gamma pass rate using 1%/1mm for all non-zero pixels are 97.4% (a) and 98.1% (b). The systematic shifts of dose profiles in Y axis are (a) -0.4 mm between the planned and measured dose, and (b) -0.5 mm between the perturbed and measured dose. (b) The gamma pass rate (1%/1mm) inside 80% isodose line = 96.6% (Percent of total voxels = 15.4%).

### **Using the Lung Insert**

The characteristics of the five experiments using the lung insert are summarized in Table 4-2. For the isocentric plans (labelled as Lung1, Lung2, and Lung3 in Table 4-2), the central axes of all beams intersect at a common point. The two experiments (Lung2 and Lung3) with a large phase shift between internal and external motion exhibit larger standard deviations for the correlation and predictor tracking errors. For experiments Lung1 and Lung 3,  $\gamma_0 > 98\%$  (using a 2%/2mm criteria), and  $\gamma_1$  are increased by 1%-points compared to  $\gamma_0$ . Most gamma failures are near the RADPOS detector (which has high density electronic components) for both the planned and perturbed dose. The difference in planned and measured penumbral widths for those two experiments is 1 mm.

For experiment Lung2 with the largest phase shift of  $43^\circ$ , the standard deviations of the correlation and predictor errors are 1.7 mm and 0.8 mm, respectively, much larger than those in Lung1 and Lung3. The  $\gamma$ -criteria failures for Lung2 for both perturbed and planned dose are mostly near the high dose gradient region along the motion direction. This suggests that dose broadening effects associated with motion tracking errors are more significant for this experiment and may be underestimated by the perturbation model. The 2%/2 mm gamma pass rate increases from  $\gamma_0 = 84.1\%$  to  $\gamma_1 = 89.5\%$  for perturbed relative to the planned dose. The measured penumbral width is 4 mm wider than the planned dose for this case.

For the experiments using non-isocentric delivery (labelled as Lung4 and Lung5 in Table 4-2), 2%/2 mm gamma pass rates are  $>96\%$  when comparing measured to perturbed dose. For Lung4, where the mean  $\pm$  SD of correlation errors is  $0.4 \pm 2.2$  and the phase shift between internal/external motion is  $25^\circ$ , the gamma pass rates increases from  $\gamma_0 = 95.5\%$  to  $\gamma_1 = 96.8\%$  by 1.3%-points. The systematic shifts for this experiment along the motion direction (Y axis) for planned vs. measured

and perturbed vs. measured are 0.6 mm and 0.3 mm, respectively. For Lung5, where the mean  $\pm$  SD of correlation errors is  $0.3 \pm 1.8$  mm and the phase shift is  $18^\circ$ , an increase of 9.6%-points from  $\gamma_0 = 86.7\%$  and  $\gamma_1 = 96.3\%$  is observed. For Lung5, the low gamma pass rate between the planned and measured dose (86.7%) is mainly caused by a systematic dose shift (-2.1 mm) in the motion direction, as shown in Figure 4-17 (a). This is partially accounted for in the perturbed dose where the systematic shift is reduced to -1.7 mm as shown in Figure 4-17 (b). This results in a much higher gamma pass rate (96.3%) since the residual shift is less than the distance to agreement criteria of 2 mm. The pass rate inside the 80% isodose line is about 82% for both Lung4 and Lung5.

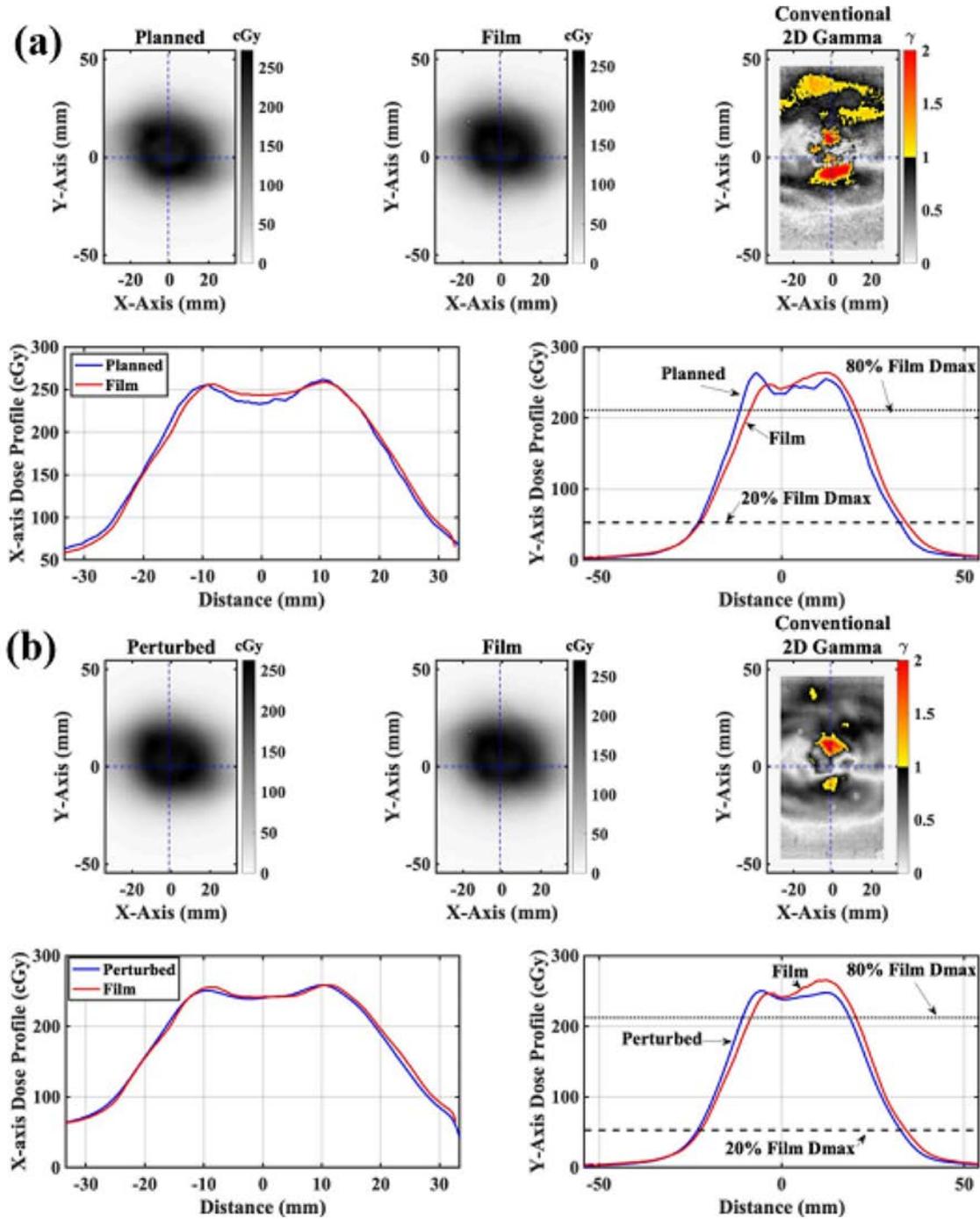


Figure 4-17. 2D gamma analyses (2%/2mm) on (a) planned and measured dose distribution, and (b) perturbed and measured dose distribution, for a non-isocentric plan using a lung insert (experiment Lung5). The gamma pass rates for all non-zero pixels are 86.7% (a) and 96.3% (b). The systematic shifts of dose profiles in Y axis are (a) -2.1 mm between the planned and measured dose, and (b) -1.7 mm between the perturbed and measured dose. (b) The gamma pass rate (2%/2mm) inside 80% isodose line = 82.3% (Percent of total voxels= 9.0%).

### Using the Liver Insert

The characteristics of the six experiments using the liver insert are summarized in Table 4-2. Refer to Table 4-2 for the labels for these experiments. Figure 4-18 shows the data points for two Synchrony models from two of the liver insert experiments, one with in-phase and the other one with out-of-phase internal/external motions (see Section 1.3.2 for model points). The same patient waveform is used for all experiments, and approximately 20 minutes of the motion trajectory for an LED marker is shown in Figure 4-19. The patient waveform changes from cycle to cycle, and there are several sudden changes in the motion trace. Larger tracking errors are expected in the case when there is a phase shift between the internal and external motion for this waveform, since the assumption that internal motion can be predicted from external motion is no longer valid.

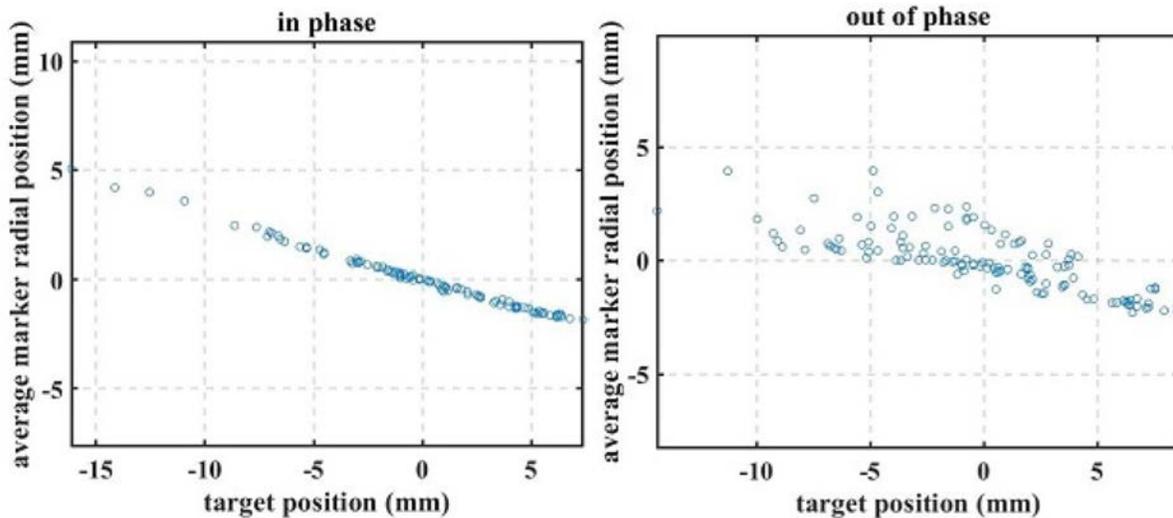


Figure 4-18. The relationship between internal target motion and average external marker radial motion (three markers) for an in-phase model (experiment Liver-Far2Lung-1) and an out-of-phase model (experiment Liver-Far2Lung-3). The data points are used to generate the predictive model used for Synchrony respiratory motion compensation.

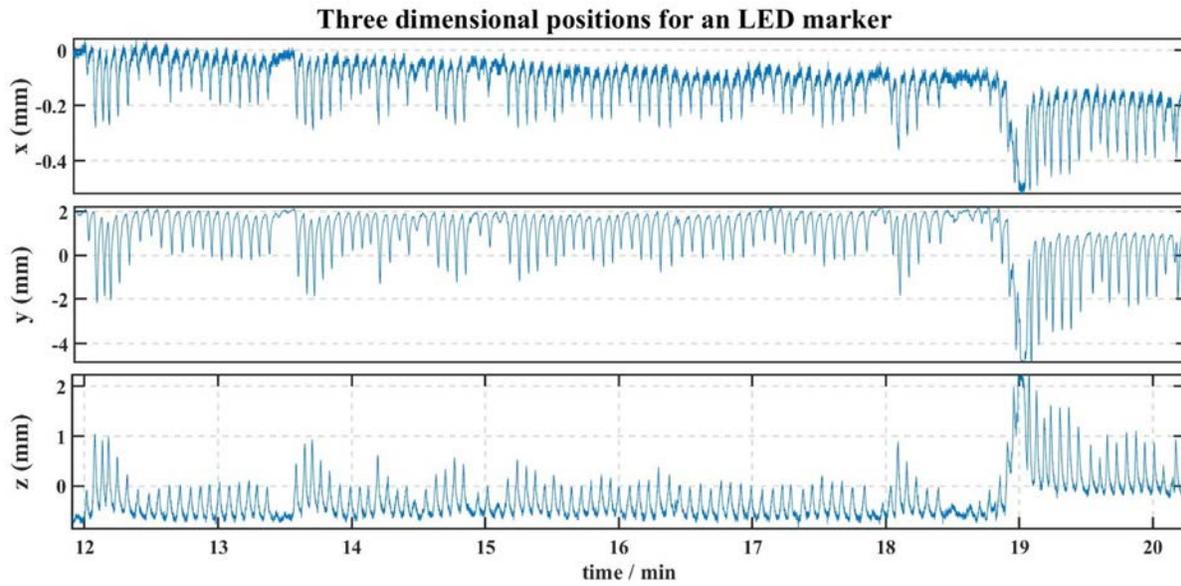


Figure 4-19. Part of the patient waveform implemented for all experiments using the liver insert. The data is extracted from the treatment log file (experiment Liver-Far2Lung-3, as labelled in Table 4-2).

For all experiments using the liver insert, except for Liver-Far2Lung-3 where there was a large uncorrected rotational offset, more than 90% of voxels passed the 2%/2mm criteria for the comparison of measured and planned dose when all non-zero dose voxels were considered. For these experiments, the difference in the penumbral width and systematic shift comparing the planned and measured dose in most cases is less than 2 mm. The gamma pass rates comparing the perturbed and measured dose ( $\gamma_1$ ) improve in all cases and are above 96% using 2%/2mm criteria. The  $\gamma_1$  are above 93% and above 87% using 2%/1mm criteria for all in-phase and out-of-phase experiments, respectively.

For experiments using the close-to-lung plans, gamma pass rates (using 2%/1mm criteria) inside the 80% isodose line when comparing the perturbed with the measured dose are above 98%. Most failures appear in the low density region of the insert and close to the edges of the film away from the high dose region. In contrast, for experiments using far-to-lung plans, most gamma

failures occur in the high dose regions. The mean  $\pm$  SD of gamma pass rates inside the 80% isodose volume is  $(73 \pm 15)$  % for these four far-to-lung experiments.

The experiments Liver-Close2Lung-1 and Liver-Close2Lung-2 emulate a liver target close to the diaphragm where the surrounding tissue densities are relatively more heterogeneous than those for a target away from the diaphragm. They have larger penumbral widths for the planned dose (16 mm) compared to the other experiments (about 12 mm) due to the presence of the low-density material near the high dose region. The penumbral widths increase by 1 to 2 mm in measurements due to dose perturbations by tracking residual errors. For these two experiments, the standard deviations of correlation errors and predictor errors were also relatively small, consistent with previous results with lower phase shifts, and  $\gamma_1$  are higher than 91% using 2%/1mm criteria. The two-dimensional gamma analyses using 2%/1mm criteria are shown in Figure 4-20 for experiment Liver-Close2Lung-2. The gamma failures are mostly close to the penumbra of measured dose (lower dose region) and near the media boundary.

The impact of having both large uncorrected rotations and tracking errors is illustrated in Figure 4-21 for the experiment Liver-Far2Lung-3. The  $\gamma_0$  using 2%/1mm criteria for experiments Liver-Far2Lung-3 and Liver-Far2Lung-4 are 63.0% and 95.6%, respectively. Their correlation errors (mean  $\pm$  SD =  $-0.6 \pm 2.2$  mm and  $-0.3 \pm 2.1$  mm) are not significantly different ( $P < 0.05$ ). The major difference between these two experiments is whether an uncorrected rotational offset ( $4^\circ$ ) is present, which is the major reason why  $\gamma_0$  is very low in Liver-Far2Lung-3. Using 2%/1mm criteria  $\gamma_1$  is above 96% for both experiments.

On average, 135 x-ray images were taken for an individual experiment (130 kVp, 16 mAs). If we consider the upper threshold of 0.062 cGy per x-ray image pair (150 kVp, 16 mAs) provided by the CyberKnife imaging whitepaper,<sup>227</sup> the total imaging dose is about 0.8% of the maximum

planned dose (10 Gy), which is a small component in the measured dose. However, the measured dose is corrected for the imaging dose for these experiments using the liver insert, and the estimated imaging dose is described in Appendix E using the measured data in the lung-insert experiment discussed in the previous section.

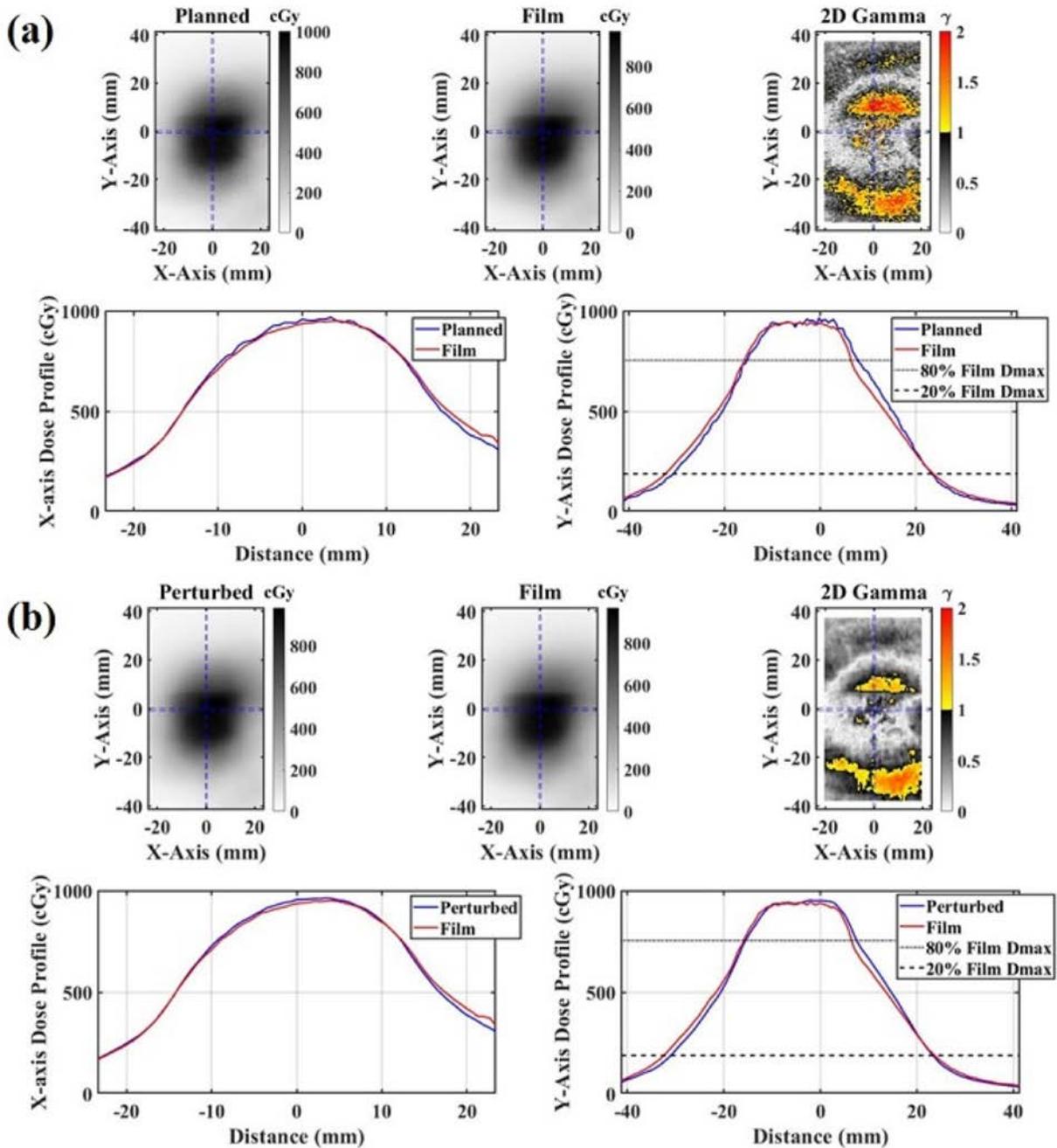


Figure 4-20. Two-dimensional gamma analyses (2%/1mm) on (a) planned and measured dose distribution, and (b) planned and measured dose distribution for experiment Liver-Close2Lung-2. The gamma pass rates (2%/1mm) for all non-zero pixels are 80.0% (a) and 91.5% (b). The systematic shifts of dose profiles in Y axis are 0.8 mm between the planned and measured dose (a), and 0.5 mm (b) between the perturbed and measured dose. (b) The gamma pass rate (2%/1mm) inside 80% isodose line = 100.0% (Percent of total voxels= 9.3%).

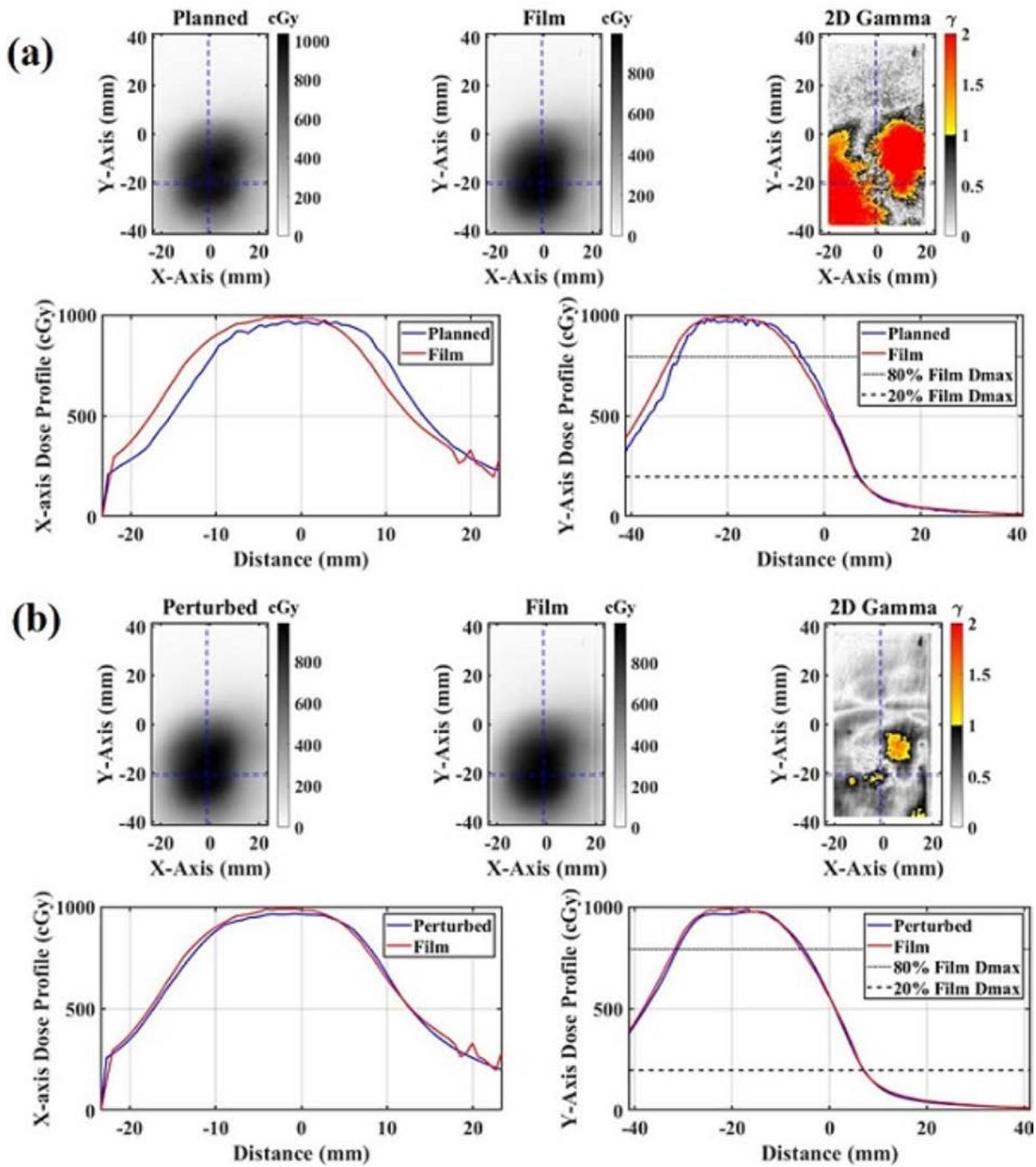


Figure 4-21. Two-dimensional gamma analyses (2%/1mm) on (a) planned and measured dose distribution, and (b) planned and measured dose distribution for experiment Liver-Far2Lung-3 with the large uncorrected rotational offset of 4.0 degrees. The gamma pass rates (2%/1mm) for all non-zero pixels are 63.0% (a) and 96.8% (b). The systematic shifts of dose profiles in Y axis are 2.2 mm between the planned and measured dose (a), and 0.9 mm (b) between the perturbed and measured dose. (b) The gamma pass rate (2%/1mm) inside 80% isodose line = 84.1% (Percent of total voxels= 8.4%).

#### **4.4. Discussion**

For experiments using the static rocket phantom positioned with a rotational offset, we validate the dose perturbation tool considering uncorrected rotations alone. A somewhat subtle artifact is visible in the 2D gamma plot in Figure 4-13b, mostly within the acceptable  $\gamma$ -criteria  $< 1$  grayscale of the colour bar appearing as a grid-like alteration of dark and light bands. We hypothesize that this is due to the rotated voxel locations being discretized and rounded to the nearest neighbouring voxel in the perturbed dose. This effect is more apparent when the rotation is an approximately constant value with a relatively large magnitude. It is worth noting that for a patient, the observed rotation angles vary throughout the treatment and from fraction to fraction (Figure 2-4 and Figure 2-5) and the artifacts in the perturbed dose would be less noticeable and more blurred out.

Using a 2%/2mm criteria for the perturbed dose,  $\gamma_1$  is above 99% for all three experiments (labelled as R1, R2, and R3 in Table 4-1) with uncorrected rotations and static phantoms. For the trigeminal plan, the measured dose is slightly lower than the perturbed dose in high dose regions. There are a few possible reasons this may occur. One possibility was the setup uncertainty, since the dose was delivered using the 5 mm collimator (a very small field) and there was a very steep dose gradient around the target. Any displacement from the high dose region due to setup uncertainty will generally lead to a lower dose. Another possible reason is that two films were used for the experiment when the film slot could fit four films, leaving a small air gap (with a thickness equal to that of two films which is approximately 0.6 mm) in the film slot. Many studies have shown that air-filled ion chambers exhibit a lower signal for small fields due to the perturbing presence of the air cavity.<sup>230,231</sup> The experiment illustrates that a relatively small uncorrected rotation can lead to a significant target underdose for small cylindrical targets such as the

trigeminal nerve. Our dose perturbation tool can be used to evaluate the effective length of the nerve covered by the prescription dose that takes into account these uncorrected angles of rotation (discussed later in Section 5.2).

The “L”-shape target using relatively larger collimators and the body path used for liver treatments is also very sensitive to uncorrected rotations. For the experiment with the phantom positioned with roll =  $3.1^\circ$ , pitch =  $-2.7^\circ$ , and yaw =  $3.1^\circ$ , we also confirm the signs and order of rotational corrections by the system, as described in Appendix B. Both experiments allow us to empirically confirm the conventions used for the application of translational and rotational corrections, as provided by Accuray in a private correspondence.

For experiments using Synchrony motion tracking, the agreement between the measured and perturbed doses improves (penumbral widths within 1 mm and 2%/2mm gamma pass rates > 95% for all experiments except Lung2) relative to the planned dose. The ball cube phantom is made of relatively homogeneous material close to tissue density. In isocentric delivery, the central axes of all beams intersect at a common point and the field edges are aligned. In these cases, the effects of penumbral broadening and systematic residual errors are simpler to interpret. Although the differences are somewhat subtle, the ball cube experiments provide evidence that our analytical model to estimate the dose due to translational perturbations is valid for simple types of motion and for isocentric delivery. There is a small systematic shift (-0.5 mm) between the perturbed dose and measured dose for the experiment Ball1. This might be introduced by the undersampling of x-ray images (only 36 acquisitions in the model), which could mean the tracking residual errors reported were not a good statistical representation of true errors during beam delivery.

For experiments using the lung insert, the tissue heterogeneity is a probable cause for the discrepancy observed between the perturbed and measured doses. The lung insert is composed of

heterogeneous materials where Solid Water (the target) is denser than the surrounding material (low-density cedar). The atomic numbers of Solid Water and the cedar (mostly carbon) are not very different, but secondary electrons travel further in low-density materials. Hence, electrons would deposit energy further from the site of photon interactions. Our perturbation model does not include this effect which may explain the gamma failures at the peripheral of the high dose regions for isocentric plans when the tracking error is large.

Liver tumours are not typically surrounded by low density media but may be adjacent to or near the interface of the lung and diaphragm. The liver insert models this situation and we consider two plans: one with the target far from and the other one close to the low-density phantom material (Styrofoam). The gamma pass rates ( $\gamma_1$ ) for the two close-to-lung experiments (Liver-Close2Lung-1 and Liver-Close2Lung-2) are greater than 97% using 2%/2mm criteria and 91% using 2%/1mm criteria. This indicates that our perturbation model can estimate dose to a target close to heterogeneous materials for translational tracking errors with a good agreement to the measured dose. For the close-to-lung experiments, the gamma failures are mostly in the penumbra of the measured dose (lower dose region) and near the media boundary. In contrast, for the far-to-lung experiments (LiverFar2Lung-1,2,3 and 4) most gamma failures are in the high dose area. This may be due to the scanning process which can cause some gamma failures near the edge of the films, because the template used (shown in Figure 4-4) cannot always keep the film corner flat during scanning (tilted corner can introduce incorrect optical density). In Marants *et al.*,<sup>169</sup> it was hypothesized that failures in the high dose regions of these types of non-isocentric plans may also be due to the complexity of the dose distribution when the beam central axes are no longer constrained to intersect at a common point. The blurring of entire dose distribution in our perturbation model does not include the effects of small individual beams intersecting and

overlapping in the high dose regions around the periphery of the target. Errors may be of a different magnitude and sign at different times during the delivery of individual beams, which could lead to changes in high dose regions not predictable by this model.

For the perturbed dose for the liver insert experiments, the systematic shift in the motion direction is within 2 mm and is reduced relative to the planned dose for all cases considered. The magnitude of the shift may be slightly underestimated in the perturbed dose, especially when the SD of the correlation and predictor errors are on the order of 2 and 0.5 mm, respectively. One possible contributing factor is that the systematic error in the model may not be well sampled in the correlation error distributions used in the perturbation model. The dose perturbation model assumes that x-ray images are taken uniformly throughout the respiratory cycle, so that the dose is perturbed using these errors with equal weights. Many of these experiments have highly out-of-phase internal/external motion, for which non-linear Synchrony models are used by the system. It is possible that respiratory motion amplitudes can be underestimated when the internal motion is changing rapidly but the external motion is in a more slowly varying part of the time trace. This may be a less significant problem for patient studies, since the majority of motion tracking models are linear (see section 2.2.5) and more than 200 x-ray pairs are often acquired during a typical liver treatment. The experiments discussed in this chapter all have less than 160 x-ray acquisitions for an individual experiment, so that the tracking residual errors reported might not a good statistical representation of true errors during beam delivery.

For most patients (see Section 2.3.1), correlation and predictor errors cancel out across the respiratory phases, so the dose broadening effect is probably the most important feature (besides uncorrected rotations) to match the experimental data to the perturbed dose. The differences between penumbral widths for perturbed and measured doses are within 1 mm for all experiments,

except for Lung2. The measured penumbra for experiment Lung2 is larger than the perturbed dose by 3 mm. However, for this version of the Synchrony motion tracking system, larger than 5 mm correlation errors that trigger the system “soft stop” of treatment are filtered from the system log files, and therefore cannot be applied to targets for dose perturbation, as described in Section 2.2.2. Larger than 5 mm correlation errors could occur during a sudden change in the breathing trace or misidentification of fiducials (for example, when the wires or part of the robot blocks the fiducials in the live images). If large correlation errors due to the sudden change in the breathing trace are frequent (e.g., when large phase shifts between internal and external motion are present) and removed from the analysis, we may underestimate the magnitude of tracking residual errors used for dose perturbation. The experiments employing the liver insert used the delivery system v.11.5. In this version, we can filter x-ray data when the corresponding rigid body errors exceed a user-defined threshold (caused by fiducial misidentification). In this way, large errors due to the change in the breathing trace can still be included in the dose perturbation. This was not possible for experiments (such as the lung experiments) conducted using the old delivery system (v.9.5 and v.10.5), where one could not link different types of errors using a unique x-ray image time stamp. Therefore, we might have underestimated the correlation errors for experiment Lung2, resulting in a narrower penumbral width. This may account for the improved accuracy of the dose perturbation model for the liver insert experiments for those experiments when the phase shifts were large.

The dosimetric impact of having a large uncorrected rotation is further emphasized by the low  $\gamma_0$  value (63.0%) for experiment Liver-Far2Lung-3, where a 4° rotation is uncorrected by the system. As shown in Figure 2-4 and Figure 2-5, a rotational offset of 4° is not unusual for liver treatments. This suggests that our dose perturbation tool could be used in the clinic to

retrospectively or prospectively, assess the dose to a target that considers both translational and rotational tracking errors for liver treatments. The measured penumbral widths are generally smaller (about 6 to 9 mm except for experiment Lung2) for experiments using isocentric plans, compared to those (about 12 to 17 mm) using non-isocentric plans. Non-isocentric plans where tissue heterogeneity is pronounced have larger penumbral widths, as seen in experiments using a liver insert (17 mm) and a lung insert (14 mm). For most of these experiments, Synchrony errors introduce a small change (about 1 to 2 mm) to the penumbral width. Details about how the target treatment is planned (using isocentric or non-isocentric) and tissue density heterogeneity seem to have more impact on the planned and measured penumbral widths. When the penumbra is broad, a plan is less sensitive to perturbations of translational errors. This suggests that the impact of tracking errors may be plan- and patient-specific. Therefore, this dose perturbation tool could be used to evaluate the dose coverage for a target at the end of each fraction, together with the early warning system (see Section 3.2.4), or prior to treatment to assess sensitivity to different types of error for a patient-specific PTV margin.

#### **4.5. Conclusions**

Our dose perturbation model was validated using film measurements in static phantoms and a programmable respiratory motion phantom for periodic motion and patient-derived breathing traces. The gamma pass rates using 2%/2mm that compare the perturbed and measured doses are above 95% for all experiments except for the Lung2 experiment. The gamma analyses and the comparison of penumbral widths have validated our analytical model's performance at the 2%/2mm level. This accounts for motion-tracking residual errors and uncorrected rotations for homogenous and heterogeneous phantom materials and in-phase and out-of-phase motion across a wide range of phase shifts. The dose perturbation tool can be used to retrospectively evaluate the

suitability of a specific PTV margin for a patient population or as a per-patient sensitivity to error assessment prior to treatment. It therefore enables adaptive planning processes and patient-specific margins for liver and more informed tolerances on allowable uncorrected rotations for TN.

## Chapter 5. Dosimetric Assessments

Using the dose perturbation tool validated in Chapter 4, we assessed the dose that was delivered to targets for liver and trigeminal neuralgia treatments. The dose assessments taking into account different delivery errors for these two body sites are discussed in Sections 5.1 and 5.2, respectively.

### 5.1.Liver

#### 5.1.1. Introduction

As discussed in Section 2.3.2, for liver treatments on CyberKnife, rotational corrections are often much larger than what the system can correct for. The dosimetric impact of uncorrected rotations, motion-tracking residual errors, and target deformations might compromise treatment outcomes.

To our knowledge, few studies have quantified the dosimetric effects of target rotations for SBRT liver treatments.<sup>142,155</sup> As discussed in Section 2.3.2, Chan *et al.*<sup>142</sup> rotated and translated the contours for the CTV and PTV based on the aggregate translational and rotational errors for ten phases of the respiratory cycle, each equally weighted in the final dose accumulation for CyberKnife liver treatments. They analyzed data for 24 liver patients and found a relatively small median dose reduction of 1.1% and 0.0% for PTV and CTV, respectively, when the aggregate rotations were of  $-0.26^\circ \pm 0.51^\circ$ ,  $0.14^\circ \pm 0.71^\circ$ ,  $0.65^\circ \pm 0.61^\circ$  in roll, pitch, and yaw directions, respectively. Cao *et al.*<sup>155</sup> estimated rotational setup errors by aligning to soft tissue using the liver as a surrogate organ in CBCT and planning CT images for 20 SBRT liver treatments using conventional gantry-based linac delivery. They found the systematic setup errors were  $-0.29^\circ \pm 0.62^\circ$ ,  $-0.06^\circ \pm 0.68^\circ$ , and  $-0.24^\circ \pm 0.61^\circ$  for roll, pitch, and yaw, with random errors of  $1.05^\circ$ ,  $0.80^\circ$ , and  $0.61^\circ$ , respectively. These values were of similar magnitude to the findings of Chan *et al.*<sup>142</sup>

However, they observed larger rotational errors with a more significant dosimetric impact on target coverage for certain fractions. After analyses of 75 CBCT images, they reported that 56.0%, 14.7%, and 1.3% of treatment fractions had at least one rotational error  $>1^\circ$ ,  $>2^\circ$ , and  $>3^\circ$ , respectively.<sup>155</sup> They evaluated the dosimetric consequences for three patients with  $>2.5^\circ$  in roll, by recalculating the dose on a CT image set that was rotated by the systematic rotational setup error, and found that the fraction of PTV receiving more than 80% of the prescription dose (specified as the PTV coverage) was reduced by  $7.5\% \pm 1.3\%$ .

A related area of research is evaluation of plan robustness against different types of delivery errors.<sup>155,232-236</sup> Liu *et al.*<sup>232</sup> proposed a fast offline approach for conventional linac treatments for evaluating the dosimetric impacts on OARs for  $\pm 2$  mm and  $\pm 1^\circ$  level systematic uncertainties (e.g., errors in the collimator, gantry, couch table, isocenter shifts, and leaf bank position). They perturbed the planned dose using rigid transformations considering these uncertainties for six brain and six spine plans. They found that  $98.9 \pm 1.3\%$  of voxels were within a 2% dose difference for dose values greater than 30% of the maximum value and at least 6-mm from the patient's surface. The previously mentioned study of Cao *et al.*<sup>155</sup> also looked at plan robustness. They rotated the planning CT images by  $1^\circ$ ,  $2^\circ$ ,  $3^\circ$ ,  $5^\circ$ , and  $7^\circ$  along roll axis, about the body centre, and recalculated the dose for seven liver patients treated on a conventional linac. They found that the PTV coverage dropped to 87.3% when  $3^\circ$  in roll was applied. Also, the plans showed the maximum dose increase of 13% to the spinal cord without translational corrections. However, the rigid-body assumption is not valid for many OARs, especially for bowel structures (contours were not changed in the replanning process). For conventional linac SBRT lung treatments, Yang *et al.*<sup>234</sup> quantified the dosimetric effects of rotational offsets (roll =  $\pm 1^\circ$ , pitch =  $\pm 3^\circ$ , and yaw =  $\pm 5^\circ$ ), and recalculated dose distributions based on the rotated CT images using the planning system. They found that 95.6%

of dose deviations were  $<1\%$  or  $<1$  Gy in PTV coverage and PTV  $D_{\max}$  for eleven patients. For the patients in their study,<sup>234</sup> the GTV was close to the isocentre so that the impacts of large rotational offsets were minimized. They also found that the nearly spherical targets and dose distributions for lung SBRT were less sensitive to rotational offsets. Liu *et al.*<sup>235</sup> evaluated the dosimetric effect of uncorrected rotations for  $1^\circ$ ,  $2^\circ$ ,  $3^\circ$ , and  $5^\circ$  in any direction for conventional linac SBRT lung treatments ( $N = 22$ ), and recalculated the dose distribution on rotated CT image sets. When the isocenter was  $> 6$  cm away from the GTV center, they found rotations as low as  $2^\circ$  might cause a significant dose reduction ( $>3\%$ ) for PTV D95. Similar to the results of Yang *et al.*,<sup>234</sup> when the isocenter was closer to the GTV center and rotations were  $< 2.5^\circ$ , negligible changes to PTV D95 were observed.

However, these studies were conducted using a relatively small cohort ( $N < 50$ ) of patients. For evaluating the plan robustness to angles of rotation, only a limited set of rotational offset observations was applied to a planned dose distribution in their studies. In Section 2.3.2, we observed that angles of rotation had a broad distribution (i.e., a large standard deviation for angles of rotation), and the use of an *average* target position could underestimate the impact of having a broad distribution of target rotations on the dose distribution.

To consider delivery errors, we validated a dose perturbation model using film measurements, as described in Chapter 4. In this chapter, we retrospectively apply the dose perturbation tool to evaluate the dose to targets considering delivery errors for a larger patient cohort. The tool can be used to inform the PTV margins selected for treatments. We can also apply the dose perturbation tool prior to a patient's treatment to assess plan robustness against a wide range of simulated angles of rotation. The tool could also provide additional information for the early warning system introduced in Chapter 3, since the planned dose distributions tend to have much higher dose values

(greater than the prescription) in the central region of the PTV, where the moving target spends more time.

### **5.1.2. Method and Materials**

The 3D contours and planned dose are extracted for 148 liver plans. As described in Section 4.2.4, the tracking errors are translational errors (including correlation, predictor, and end-to-end errors), large rotations that the system cannot correct for, and target deformations, and they are used to perturb the planned dose for the dose perturbation tool. Errors are all derived from system logfiles and XML files. For each plan, we calculate treated D99/Rx for GTV(s) where Rx is the prescription dose and compare it with the planned D99/Rx. We also retrospectively evaluate plan robustness based on the treated D99/Rx as we vary the simulated angles of rotation (roll and pitch:  $\pm 1.5^\circ$ ,  $\pm 3.0^\circ$ ,  $\pm 4.5^\circ$ , yaw:  $\pm 3.0^\circ$ ,  $\pm 4.0^\circ$ ,  $\pm 5.0^\circ$ ).

Case studies for three patients are also provided to illustrate how the treated dose coverage can be improved during the planning process and at the patient setup. The dose perturbation tool is also used to assess the suitability of the selected PTV margins for these patients. We replanned the three case studies using the following criteria to achieve robust planning goals and settings. First, the imaging isocenter was placed in the centre of the target. Secondly, isocentric and non-isocentric beams were combined in the same plan for the same target to increase the dose in the central part of the target. In addition to optimizing for PTV D95  $> 0.95$  Rx, which is required by the current planning guideline, the dose to the GTV was also optimized to achieve GTV D99 in the range from 105-110% of Rx.

Additionally, dosimetric assessment is combined with the early warning system (introduced in Section 3.2.4). Dose coverage and the potential for geographical miss could both be considerations when assessing margin suitability at the end of each fraction. We considered two different ways

this could be done. First, the GTV voxel occupancy tool is modified to include translational, rotational, and rigid body errors. This means, in addition to considering Synchrony errors in calculating the three-dimensional GTV voxel occupancy matrix (see Section 3.2.4), the method described in Section 4.2.4 that perturbs the dose matrix using the angles of target rotation and rigid body errors is applied to the GTV voxel occupancy matrix (see Appendix C). The distribution of voxel occupancy is then compared to the planned dose (not perturbed). Alternatively, dose coverage metrics such as D99 and D95 for the perturbed dose distribution can be assessed for the GTV itself. Two case studies, one predicted to be in Class 1 (smaller margin) and another in Class 2 (standard margin), are considered to illustrate the process. Each of their treatments is also replanned to evaluate robustness to delivery error as described above.

### 5.1.3. Results

**Patients with Rotation Tracking.** For the patient cohort assessed in this chapter, 47% of the patient treatments (70 out of 148 plans) had rotational corrections which could be accurately calculated and tracked by the system. Figure 5-1(a) and (b) show the inputs to the perturbation model for this group of treatments. Because we are illustrating the inputs to the perturbation model, these figures differ from those in Chapter 2. The *residual* uncorrected rotational errors for these treatment courses are shown in Figure 5-1 (a). If the rotation correction in *any* direction is greater than what the system can correct, rotational corrections cannot be applied by the system, and the rotational residual errors in all directions are non-zero in the dose perturbation model and in the figure. The rotational corrections are shown as zero if the system can correct for them. The correlation errors and rigid body errors are illustrated in Figure 5-1 (b). Some box plots for the patients with relatively small rigid body errors are too narrow to be visible because they are binned to have the smallest dimension of the dose matrix voxel size as described in Section 4.2.4. Twenty-

nine out of 70 patients (41%) had an average rotation offset larger than  $3^\circ$  in at least one direction for their treatments. Some patients (such as #2, 18, 32, and 53) had very large uncorrected rotations. Note that in Chapter 2 rotation data include both correctable and non-correctable angles. The data shown in Figure 2-4 are also for a smaller patient cohort ( $N = 23$ ) and grouped by treatment fraction.

The impacts on  $D99/Rx$  considering different types of delivery errors are shown in Figure 5-1 (c) and Figure 5-2. Eight out of 97 targets are treated with compromised dose, which we define as having  $(D95-D99)/D99 \cdot 100\% > 10\%$  during planning (defined in Section 2.3.2). After accounting for treatment delivery errors, approximately one-third of targets are underdosed, i.e.,  $D99/Rx < 1$ . If rigid body errors are neglected, about one-fifth of targets are underdosed in the perturbed dose volumes. All compromised targets (8% of the total number of targets considered) have treated  $D99/Rx < 1$ , whether the deformation is considered or not.

Two-sample t-tests for different quantities were performed for underdosed targets ( $n = 22$ ) and non-underdosed targets ( $n = 75$ ) based on the treated  $D99/Rx < 1$  definition and neglecting rigid body errors. The results are shown in Table 5-1. The prescription dose ( $Rx$ ), planned  $D99/Rx$ , target-to-isocenter distances, standard deviations of correlation errors, and translational offsets in pitch angles are significantly different for the two groups ( $P < 0.05$ ). The target volume, translational correlation errors, and rotational offsets of roll and yaw angles are not significantly different for the two groups. Figure 5-3 shows the comparison of planned and treated  $D99/Rx$  (without considering rigid body errors) for treatments with different prescriptions.

We also performed a two-sample t-test for treatments with both a large rotational offset ( $>3^\circ$ ) and target centroid  $> 2$  cm away from the imaging isocenter ( $n = 41$ ) compared to the remaining

patients ( $n = 56$ ). For these two groups a statistically significant difference ( $P < 0.01$ ) was found in the treated D99/Rx and the  $\Delta$ D99/Rx (the difference between the planned and treated D99/Rx).

Figure 5-4 shows a plot of  $\Delta$ D99/Rx with and without rigid body errors taken into consideration. In our perturbation model (see Section 4.2.4), targets are either expanded or contracted depending on the sign of the rigid body error. Targets that are on average expanded or contracted in the perturbation model are plotted in different colours in the figure. If the deformation is taken into consideration, we observe a change of more than 5% in the treated D99/Rx for 28 out of 97 targets. Most targets have a reduced treated D99/Rx when rigid body error is considered, which is indicated by more points under the reference line ( $x = y$ ). Furthermore, target expansion ( $n = 45$ ) introduces significantly larger reduction ( $P = 0.003$ ) in D99/Rx than target contraction ( $n = 52$ ).

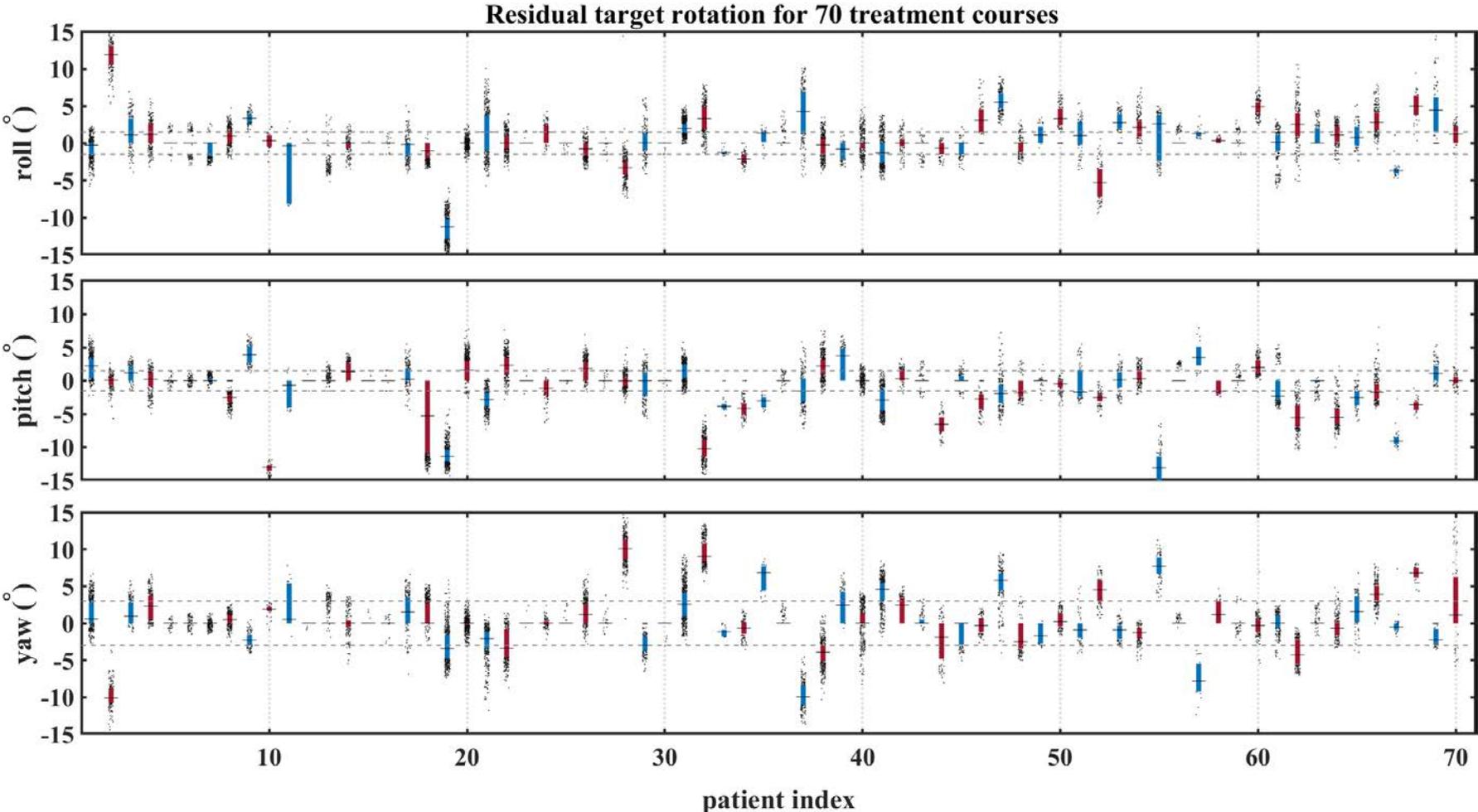


Figure 5-1 (a). Target rotations in roll, pitch, and yaw are illustrated for 70 patients (from left to right: treatment time in chronological order). Red and blue symbols alternate for different patients. For these patients, target rotations were tracked during treatment and met the criteria for accurate calculation of rotations. The horizontal dashed lines indicate the largest magnitude of corrections the system can apply. Each box spans from the first to the third quartile, and the horizontal solid line indicates the median. Data beyond the box are shown as black dots.

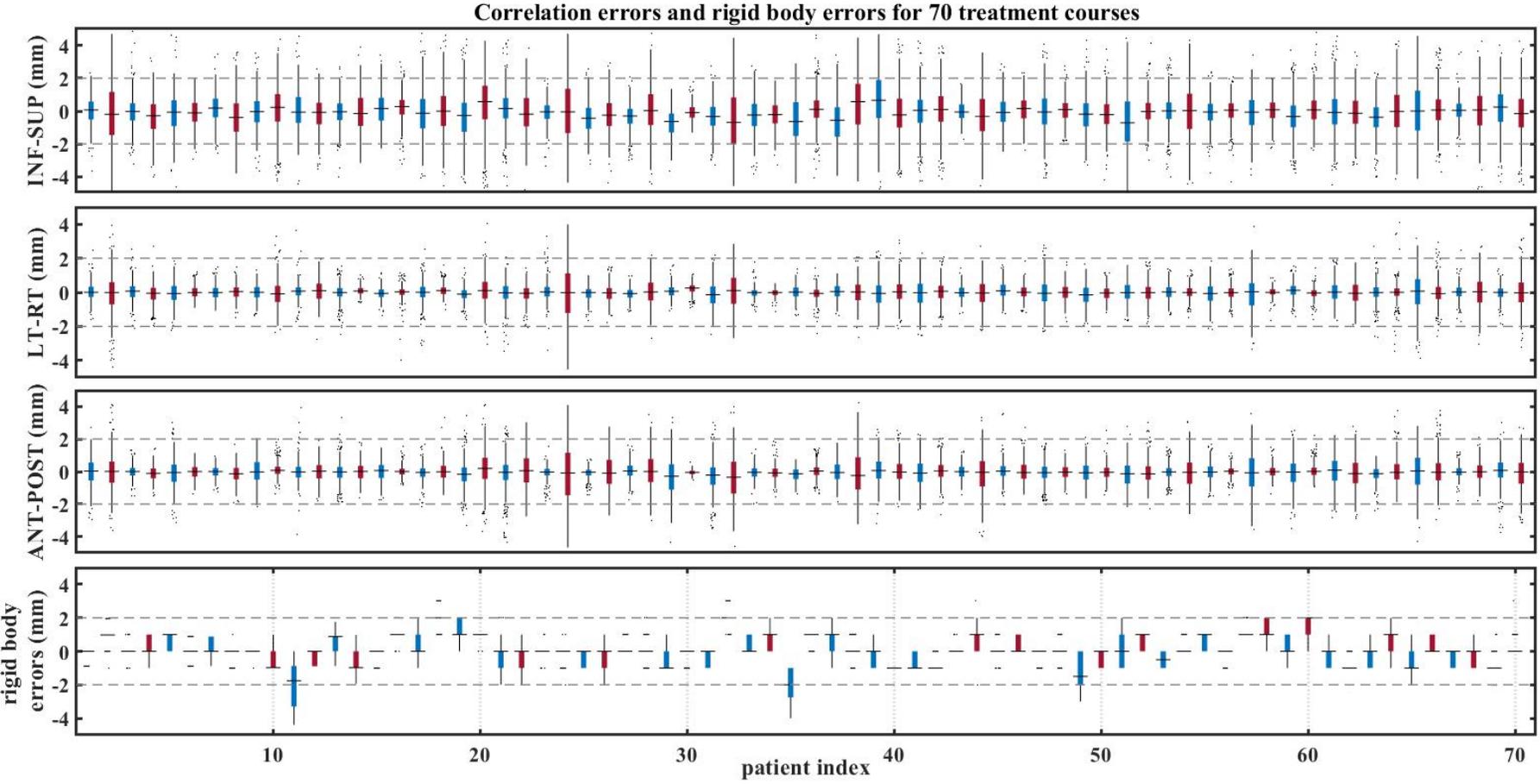


Figure 5-1 (b). Correlation errors and rigid body errors are illustrated for the same patients (from left to right: treatment time in chronological order). Red and blue symbols alternate for different patients. The horizontal reference lines indicate errors =  $\pm 2$  mm. Each box spans from the first to the third quartile, the vertical line represents the first to the 99<sup>th</sup> percentile of data, and the horizontal solid line inside the box indicates the median.

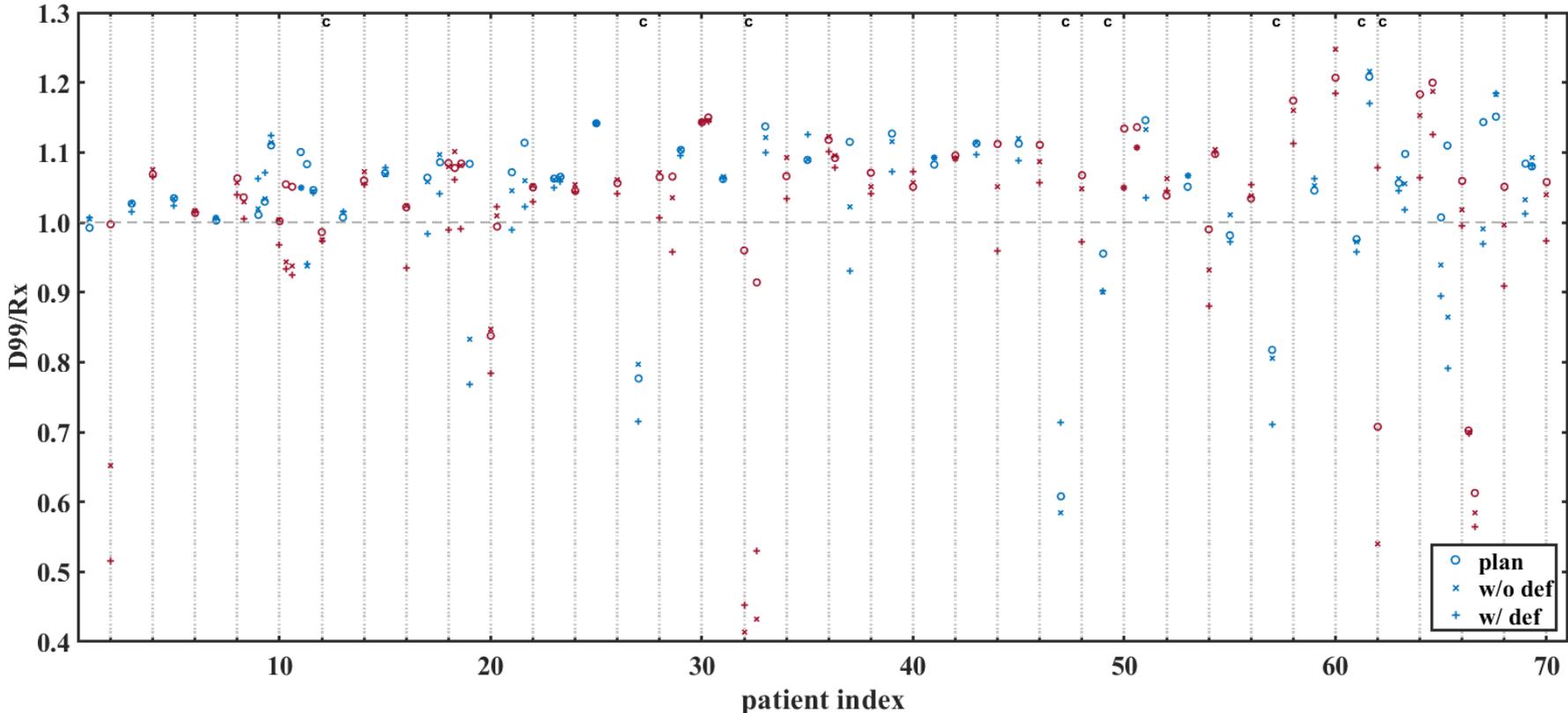


Figure 5-1. (c) The planned and treated (with and without rigid body errors) D99/Rx are illustrated for the same patients (97 targets). Red and blue symbols alternate for different patients. The letters “c” on the top indicates that data below belong to a compromised treatment plan. Each box spans from the first to the third quartile, the vertical line represents the first to the 99<sup>th</sup> percentile of data, and the horizontal solid line inside the box indicates the median.

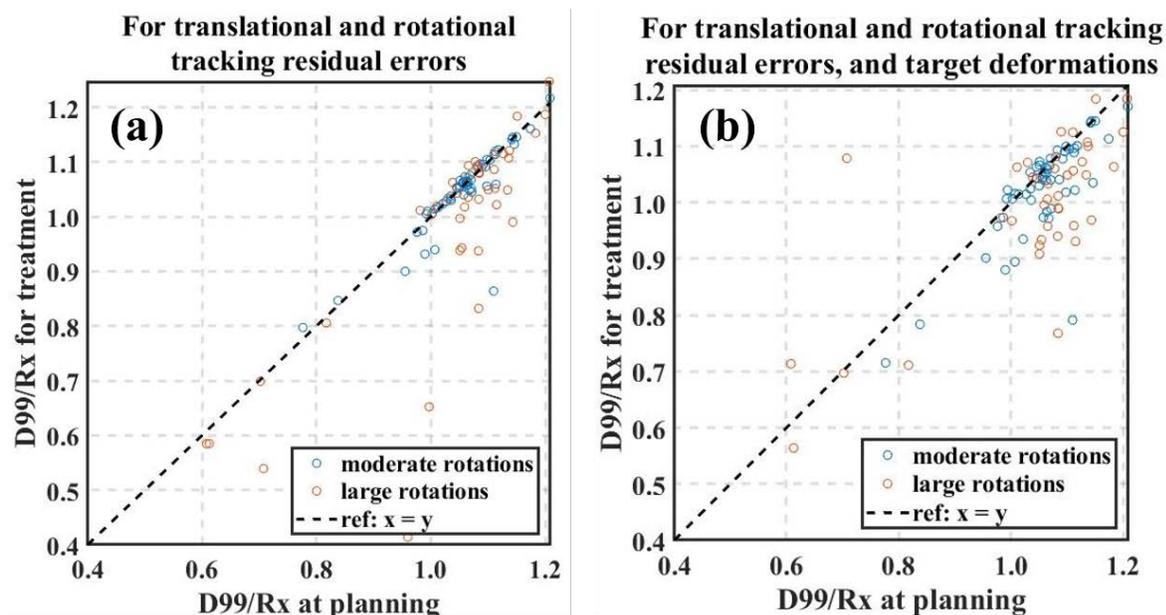


Figure 5-2. (a, b) D99/Rx for planning and treatments that considers different types of delivery errors. “Moderate rotations” indicate the average rotation during treatment is within  $3^\circ$  in roll, pitch, and yaw, while the “large rotations” means the average rotation is larger than  $3^\circ$ .

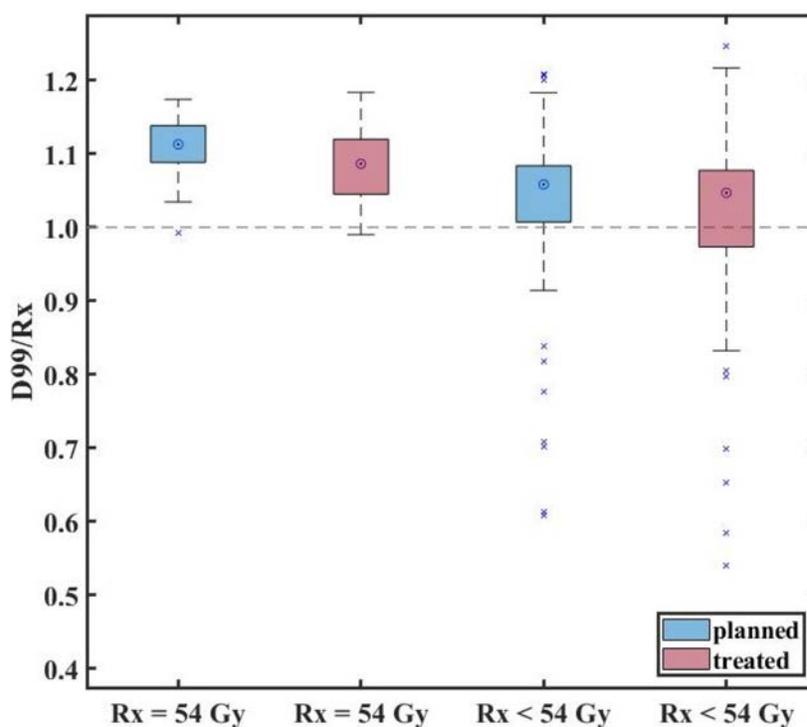


Figure 5-3. The comparison of planned and treated D99/Rx (not considering rigid body errors) for treatments with different prescriptions (Rx).

Table 5-1. Results of two-sample two-tailed t-tests at 5% significance level for several potentially relevant quantities for underdosed<sup>a</sup> and non-underdosed<sup>b</sup> targets considering translational and rotational tracking residual errors (but neglecting rigid body errors) for 70 patients (97 targets) with rotations tracked by the system.

Quantity Compared	Underdosed Targets <sup>a</sup> (N = 22)	Non-underdosed Targets <sup>b</sup> (N = 75)	P Value <sup>c</sup>
Rx (Gy)	45.8 ± 4.0	48.3 ± 4.2	<b>0.03</b>
planned D99/Rx	0.93 ± 0.16	1.08 ± 0.05	<b>&lt;0.001</b>
target volume (cc)	34.5 ± 70.3	33.2 ± 51.8	0.9
target centroid-to-isocentre distance (cm)	5.6 ± 2.7	4.4 ± 1.9	<b>0.03</b>
average roll angle (deg)	1.2 ± 4.2	0.4 ± 1.9	0.2
average pitch angle (deg)	-3.7 ± 5.0	-1.0 ± 3.1	<b>0.003</b>
average yaw angle (deg)	1.0 ± 4.7	0.5 ± 2.8	0.5
SD of roll angles (deg)	1.6 ± 0.9	1.4 ± 0.9	0.5
SD of pitch angles (deg)	1.5 ± 0.7	1.4 ± 1.0	0.9
SD of yaw angles (deg)	1.5 ± 0.7	1.4 ± 0.8	0.5
average INF/SUP correlation error (mm)	-0.1 ± 0.2	-0.1 ± 0.2	0.8
average LT/RT correlation error (mm)	0.0 ± 0.1	0.0 ± 0.1	0.3
average ANT/POST correlation error (mm)	0.0 ± 0.1	0.0 ± 0.1	0.7
SD of INF/SUP correlation errors (mm)	1.4 ± 0.4	1.1 ± 0.3	<b>0.007</b>
SD of LT/RT correlation errors (mm)	0.7 ± 0.3	0.5 ± 0.2	<b>&lt;0.001</b>
SD of ANT/POST correlation errors (mm)	0.8 ± 0.3	0.7 ± 0.3	<b>0.02</b>

<sup>a</sup> “underdosed” defined as treated D99/Rx <1

<sup>b</sup> “non-underdosed” defined as treated D99/Rx ≥1

<sup>c</sup> bold values denote statistical significance at  $P < 0.05$  level

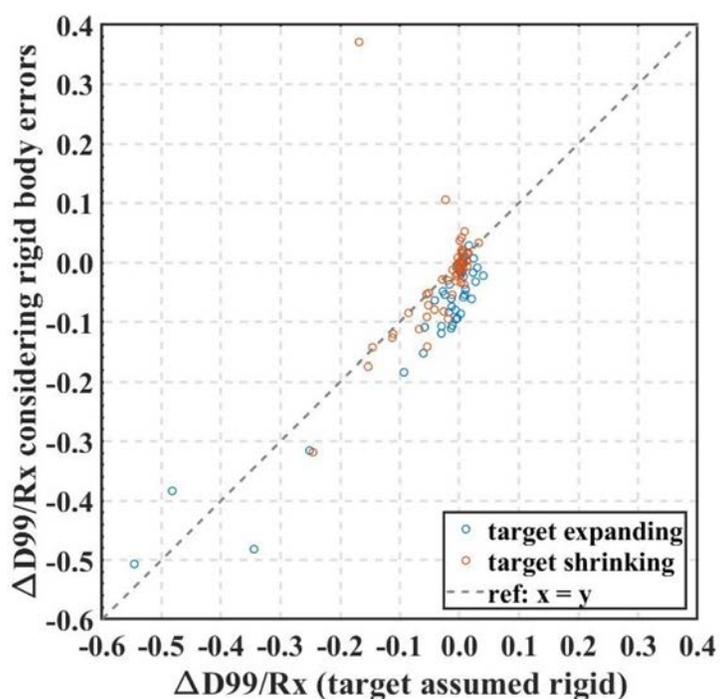


Figure 5-4. The difference between treated and planned D99/Rx is denoted as  $\Delta D99/Rx$ . This figure compares the impact of including or neglecting rigid body error. “Target expanding” and “target shrinking” indicate the average signed rigid body error being positive or negative, respectively.

**Plan Sensitivity.** The distributions of planned D99/Rx have a mean  $\pm$  SD =  $1.05 \pm 0.11$ , 5<sup>th</sup> – 95<sup>th</sup> percentile: 0.82 – 1.17 for 192 targets. For 32 of these targets (17% of the total), the dosimetric target coverage was compromised (because of OARs proximity) in the planned dose distributions, indicated by the planned D99/Rx < 1. This distribution is broad, indicating that target coverage is highly patient dependent and often compromised for organs at risk. As stated earlier, many patients did not have rotational corrections because they could not be calculated by the system. To consider the dosimetric impacts of having an uncorrected rotational offset during treatment, we considered all 148 patient plans and evaluated plan sensitivity to rotations. This is completed by calculating the treated D99/Rx using the other observed delivery errors (translational and rigid body errors) while separately varying the simulated angles of rotation about each axis. The angles assessed were  $\pm 1.5^\circ$ ,  $\pm 3.0^\circ$ , and  $\pm 4.5^\circ$  for roll and pitch, and  $\pm 3.0^\circ$ ,  $\pm 4.0^\circ$ , and  $\pm 5.0^\circ$  for yaw. In this analysis,

if the average rotations are less than  $\leq 3^\circ$ , 77% of targets have  $D99/Rx > 1$  if rigid body errors neglected. Little changes if rigid body errors are also taken into consideration (72% of targets have  $D99/Rx > 1$  for average rotations  $\leq 3^\circ$ ). The reduction in  $D99/Rx$  ( $\Delta D99/Rx$ ) due to a large rotational offset as a function of planned  $D99/Rx$  is shown in Figure 5-5. For the distance between the target centroid and imaging isocentre, the mean and SD is  $4.4 \pm 2.2$  cm for all targets. The red clusters indicate qualitatively that plans can be made more robust against large rotational offsets by having a smaller target-to-isocentre distance ( $< 3$  cm).

The hypothesis testing described in this paragraph was done for plan sensitivity to rotations considering the treated  $D99/Rx$  while neglecting rigid body errors. The compromised targets ( $n = 19$ ) have significantly lower treated  $D99/Rx$  ( $P < 0.0001$ ) for any simulated angle of rotation compared to the remaining targets. The average target volumes for compromised and non-compromised targets are  $59.1 \pm 87.5$  and  $30.7 \pm 47.9$  cc, respectively, which is also significantly different ( $P = 0.03$ ). The average and SD of treated  $D99/Rx$  are  $0.98 \pm 0.14$  and  $1.03 \pm 0.10$ , respectively, for large ( $>15$  cc,  $n = 88$ ) and small ( $\leq 15$  cc,  $n = 104$ ) targets. This implies that targets larger than 15.0 cc are more sensitive to rotations with significantly lower treated  $D99/Rx$  against all tested rotations ( $P < 0.05$ ).

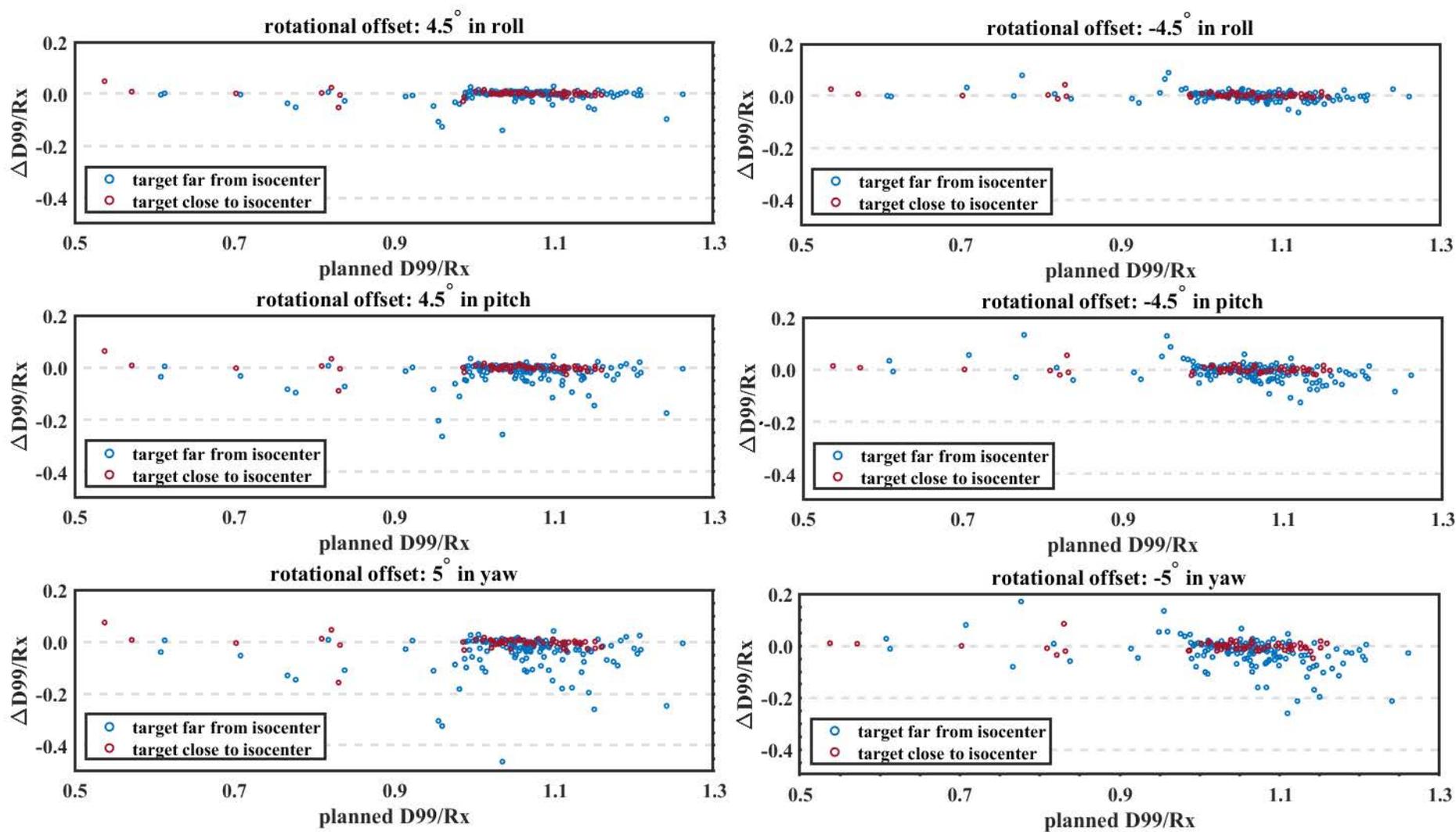


Figure 5-5. The relationship between  $\Delta D99/Rx$  when a large rotational offset is present and the planned D99/Rx. “Targets far from isocenter” ( $n = 137$ ) and “target close to isocenter” ( $n = 55$ ) indicate target-to-isocentre distance  $> 3$  cm and  $\leq 3$  cm, respectively.

**Case Studies (Retrospective Plan Robustness Evaluation).** As discussed above, treatments can be made less sensitive to tracking errors by having a higher dose in the central part of the PTV, reducing target-to-imaging isocenter distance at planning, and repositioning the patient to reduce the magnitude of rotational corrections at treatment. Three case studies for patients #19, 32, and 65 (as labelled in Figure 5-1) are used to illustrate different ways a plan can be made more robust against different types of delivery errors. The treated D99/Rx that considers different types of delivery error for the case studies is listed in Table 5-2.

Table 5-2. The near-minimum dose (D99) to target(s) considering different delivery errors and for the original plan are compared for three case studies. For each case except for the original plan, a change was made for one of the following: placing the imaging isocentre close to the centre of the target using the planned dose distribution, reducing the rotational offset using the planned dose distribution, and re-creating a dose distribution with higher dose delivered to the target centre. For each, the parameter(s) which illustrate a change are shown in bold font.

Case index *	Rx (Gy)	target-to-imaging isocentre distance (cm)	residual rotational offset (°)			average rigid body error (mm)	planned D99/Rx	treated D99/Rx (w/o rigid body errors)	treated D99/Rx (w/ rigid body errors)
			roll	pitch	yaw				
19 <sup>o</sup>	48	4.5	-11.4	-11.1	-3.0	1.4 ± 0.6	1.08	0.83	0.77
19 <sup>s1</sup>		<b>0.0</b>	-11.4	-11.1	-3.0			1.06	0.99
19 <sup>s2</sup>		4.5	<b>-6.4</b>	<b>-6.1</b>	<b>0.0</b>			0.98	0.88
19 <sup>r</sup>		<b>0.1</b>	-11.4	-11.1	-3.0			<b>1.11</b>	1.09
32 <sup>o(t1)</sup>	48	9.5	3.4	-10.1	9.6	2.0 ± 0.2	0.96	0.41	0.45
32 <sup>o(t2)</sup>		4.7	3.4	-10.1	9.6		0.91	0.43	0.53
32 <sup>s1(t1)</sup>		<b>0.0</b>	3.4	-10.1	9.6		0.96	0.71	0.78
32 <sup>s1(t2)</sup>		<b>0.0</b>	3.4	-10.1	9.6		0.91	0.86	0.76
32 <sup>s2(t1)</sup>		9.5	<b>0</b>	<b>-5.1</b>	<b>0</b>		0.96	0.78	0.70
32 <sup>s2(t2)</sup>		4.7	<b>0</b>	<b>-5.1</b>	<b>0</b>		0.91	0.99	0.87
65 <sup>o(t1)</sup>	48	10.6	0.9	-2.5	1.7	0.6 ± 0.6	1.01	0.94	0.89
65 <sup>o(t2)</sup>		12.1	0.9	-2.5	1.7		1.11	0.86	0.79
65 <sup>s1(t1)</sup>		<b>0.0</b>	0.9	-2.5	1.7		1.01	1.01	1.01
65 <sup>s2(t2)</sup>		<b>0.0</b>	0.9	-2.5	1.7		1.11	1.10	1.08
65 <sup>r(t1)</sup>		<b>1.2</b>	0.9	-2.5	1.7		<b>1.05</b>	1.05	1.04
65 <sup>r(t2)</sup>		<b>3.0</b>	0.9	-2.5	1.7		<b>1.19</b>	1.18	1.15

\*Subscript Notation:

“o” stands for the original plan as it was treated.

“s” stands for simulation, where the impact of changing a treatment or planning parameter without changing the planned dose is evaluated.

“r” stands for replanned for cases which were replanned to make the delivery more robust to delivery errors.

“t<sub>1</sub>” and “t<sub>2</sub>” represent two separated distinct targets that were treated in the same plan for patient index 32 and 65; these targets have different sensitivities to delivery error.

Patient 19 had a single target (28 cc) treated with Rx = 48 Gy in 6 fractions. A large reduction in D99/Rx was observed mostly due to a large rotational offset (see Table 5-2). As described in Section 1.3.1, the standard CyberKnife couch can only rotate the patient by a maximum angle of  $\pm 5^\circ$  in roll and pitch, and the yaw angle must be adjusted manually. We simulated the centre of rotation (imaging isocentre) being placed at the target centroid (“19<sup>s1</sup>” in Table 5-2) and found treated D99/Rx increased by more than 22% with and without rigid body errors considered. We replanned the case with the imaging isocentre placed near the target centroid (“19<sup>r</sup>” in Table 5-2). The planned D99 for the GTV were increased by just 3% of the prescription (D99/Rx increased from 1.08 to 1.11). The coverage was maintained with a treated D99/Rx  $\geq 1.0$  with and without rigid body errors considered. See Figure F-1 in Appendix F for the comparison of the dose distributions for the original and replanned treatments. If it were possible to manually reposition this patient to reduce the required rotational correction by  $5^\circ$  in roll and pitch positive directions and zero the yaw rotational offset (“19<sup>s2</sup>” in Table 5-2), the treated D99/Rx would improve  $>10\%$  no matter whether or not rigid body errors were considered.

Patient 32 had two anterior targets (8 cc and 17 cc) with Rx = 48 Gy in 6 fractions. The planned dose was slightly compromised to both targets in order to meet dose constraints for OARs such as bowel, and stomach, which has the dose limit Dmax  $< 30$  Gy. The original plan was for both targets to be treated at the same time. It was necessary to place the imaging isocenter posterior to the two targets due to limitations in how low the couch could be positioned in the anterior/posterior direction. At that time, all liver SBRT patients were set up with Styrofoam pads under the Vac-Lok cushion to lower their arm positions. This avoided treating through the arms for lateral beams directed at posterior targets. However, this was not necessary for this patient because of the anterior location of the lesions. A large rotational offset (about  $10^\circ$  in pitch and yaw) was observed that led

to low dose coverage for both targets (treated D99/Rx of 0.45 and 0.53, including rigid body errors). Ideally, when two targets are relatively distant from each other, errors can be minimized by creating separate treatment plans for each target with the imaging isocentre for each plan positioned at the target centroid. This can be simulated separately for each target in our dose perturbation model by computing the rotational corrections as though the imaging isocenter (i.e., the point about which rotations are calculated) for each target is located at its centroid. When the imaging isocentre was simulated to be at each target centre (“32<sup>s1(t1)</sup>” and “32<sup>s1(t2)</sup>” for each target in Table 5-2), treated D99/Rx was improved by about 30%. For this patient we also considered the impact if we had repositioned the patient closer to their rotational offset. As shown in Table 5-2, the treated D99/Rx could have been maintained at 0.70 and 0.87 for the two targets (labelled as “32<sup>s2(t1)</sup>” and “32<sup>s2(t2)</sup>” in Table 5-2). We did not consider creating a hotter plan which was more robust to error for this patient, since the plan was so constrained by the OAR doses.

Patient 65 had two small targets (1 cc and 3 cc) also treated with Rx = 48 Gy in 6 fractions. They were originally planned with both targets treated in the same plan with each target more than 10 cm away from the imaging isocentre, close to the average position of the combined lesions. The observed average rotational corrections (0.9°, -2.5°, 1.7° in roll, pitch, yaw, respectively) for the entire treatment course were relatively small. We considered a simulation with the isocentre placed at each target centroid (65<sup>s1(t1)</sup> and 65<sup>s1(t2)</sup>), and we replanned this case with separate plans for each target (65<sup>r(t1)</sup> and 65<sup>r(t2)</sup>). In the replanned case, it was necessary to move the imaging isocenter slightly away from each target centroid in order to ensure five fiducials are visible in the DRRs. This would also increase the likelihood that rotations could be tracked by the system. See Figure F-2 in Appendix F for the comparison of the dose distributions for the original and

replanned treatments. Treated D99/Rx is above 1.0 with and without rigid body errors considered for both simulation and replanning case studies for this patient.

**Case Studies (Prospective Dosimetric Assessment)** Following the workflow described in Chapter 3, patient #19 would be predicted as Class 2 and would have standard 5 mm PTV margins (PTV5mm) for their entire treatment course. Orthogonal planes through the GTV centroid are shown in Figure 5-6 for patient 19. The figure shows the PTV, the planned Rx isodose line (replanned for robustness but *not* perturbed for delivery error), and the GTV voxel occupancy, which includes translational and rotational delivery errors with (b) and without (a) rigid body errors. GTV voxels with occupancy greater than 0.5 are entirely inside the Rx isodose (Rx = 48 Gy in 6 fractions) for the planes shown. Dose coverage metrics are also evaluated for the perturbed dose distribution for this patient. When rigid body errors are neglected, D95 = 54.1 Gy, D99 = 52.3 Gy, and the voxel occupancy probability is  $P_i = 1.00$ . With rigid body errors (mean  $\pm$  SD =  $1.4 \pm 0.6$  mm) considered, D95 = 50.8 Gy, D99 = 48.0 Gy, even though  $P_i = 0.98$ . This means the PTV margin is suitable for treatment with this plan, as the target receives sufficient dose coverage.

Patient #1 is planned with a PTV<sub>3mm</sub> (Rx = 54 Gy in 3 fractions). The patient has three fractions of treatment, and  $P_i$  is 1.00 for every fraction when the rigid body errors are neglected. In the early warning system proposed in Chapter 3, replanning using a larger PTV margin is *not* required when  $P_i > 0.99$ . The average observed angles of rotation were  $0.2^\circ \pm 0.9^\circ$ ,  $1.7^\circ \pm 1.8^\circ$  and  $0.1^\circ \pm 1.1^\circ$  in roll, pitch and yaw, respectively, as shown in Figure 5-1. As shown in Figure 5-7, the GTV is well covered by Rx isodose for the planes shown but sometimes in the range of 0.001 to 0.5 outside of PTV<sub>3mm</sub>. For the perturbed dose distribution, after all fractions of treatment, D95 = 54.0 Gy and D99 = 51.2 Gy considering translational and rotational tracking errors. If rigid body errors (mean  $\pm$  SD =  $-0.1 \pm 0.3$  mm) are also considered, D95 and D99 increase to 58.9 Gy and

56.6 Gy, respectively, since the rigid body errors include contraction of the target. In this case, the near-minimum treated D99 dose well exceeds the prescribed dose. This means the PTV margin chosen is sufficient for treatments with sufficient dose coverage for both dosimetric assessments.

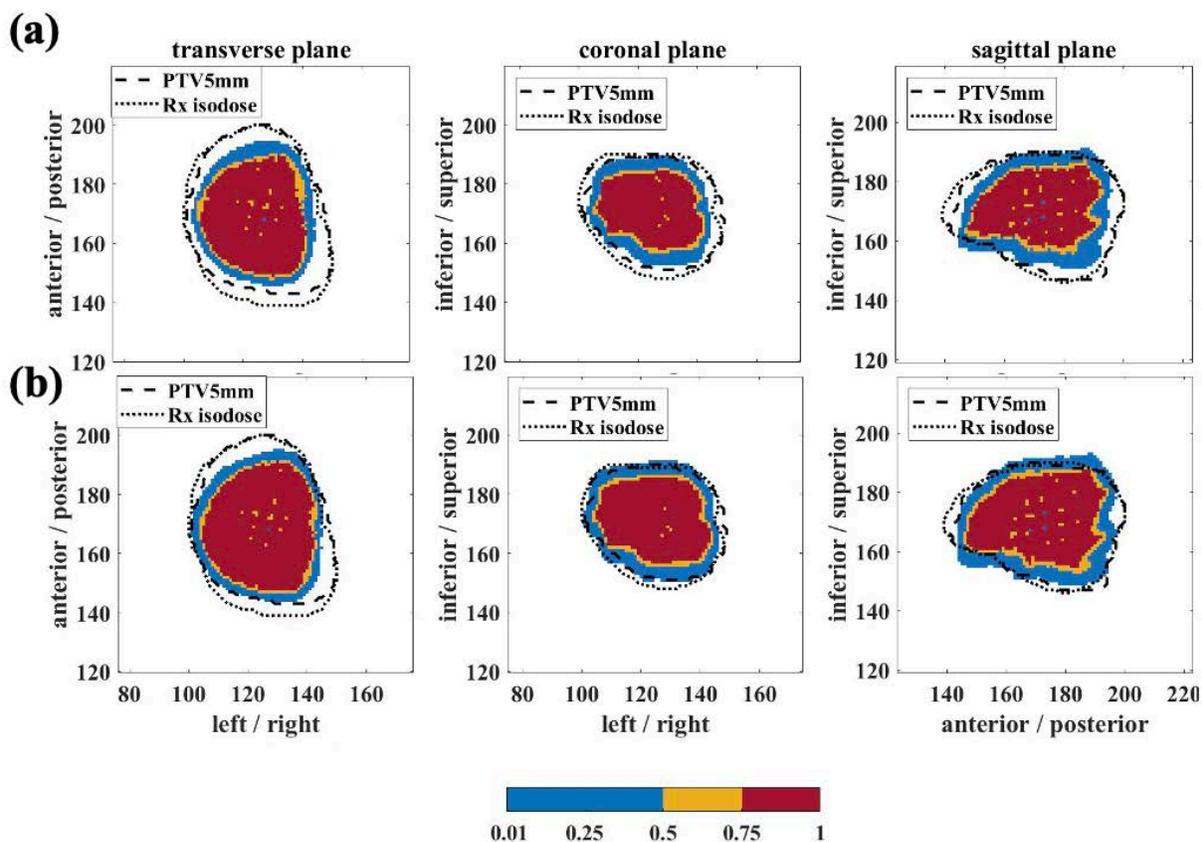


Figure 5-6. Geometric and dosimetric coverage for GTV of patient #19. The colour bar indicates the voxel occupancy probability for the GTV which includes the impact of translational and rotational delivery errors neglecting (a) and including (b) rigid body errors.

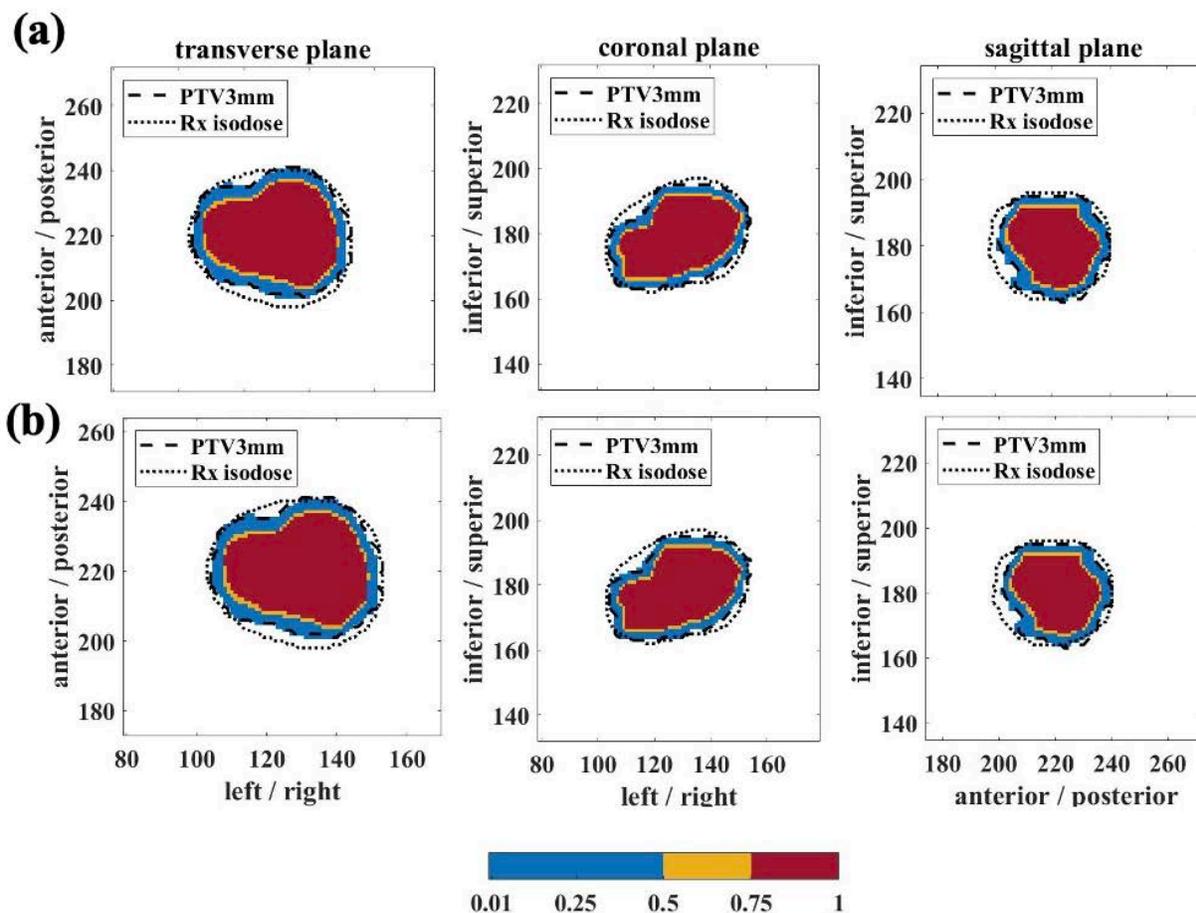


Figure 5-7. Geometric and dosimetric coverage for GTV of patient #1. The colour bar indicates the voxel occupancy probability for the GTV which includes the impact of translational and rotational delivery errors neglecting (a) and including (b) rigid body errors.

#### 5.1.4. Discussion

We investigated the dosimetric impacts for motion-tracking residual errors, uncorrectable large rotations, and target deformations for 148 liver patient plans treated on CyberKnife. The tracking residual errors, rigid body errors and target rotation are assumed to be independent sources of uncertainty and are therefore applied as separation convolutions in the dose perturbation, which is statistically reasonable for most patients based on the data presented in Chapter 2. However, if they are correlated for some patients, the dose broadening effects may cancel out over the respiratory cycle. Our method can model these synchronization effects for the most recent

CyberKnife treatment delivery systems (v.11.5 or later), since the three types of delivery errors can be identified for each x-ray image acquisition.

Similar to the results of Liu *et al.*'s study,<sup>235</sup> we found that a large target-to-isocenter distance could lead to underdosed targets. Unlike Cao *et al.*<sup>155</sup>'s study, we do not include dosimetric assessments for OARs since it is less realistic for highly deformable and mobile organs, such as bowel structures. We acknowledge that deforming the GTV contour isotropically using average rigid body errors is a simplified approach to model target deformation. Hence, the dosimetric assessments for liver treatments are reported with and without rigid body errors.

Different types of errors and their dosimetric impacts are shown for the 70 patient plans (of 148 in total) in which rotations could be accurately calculated by the system in Figure 5-1 and Figure 5-2. These figures illustrate the inputs and output of the dose perturbation model. There is a great deal of patient-to-patient variation suggesting a role for patient-specific adaption in planning and treatments. We divided this group into those with sufficient and insufficient GTV coverage (treated  $D99/Rx \geq 1$  and  $D99/Rx < 1$ ) and performed hypothesis testing. Several parameters were statistically different between the two groups. These parameters include the magnitudes of all types of tracking errors, the distance between the target centroid and imaging isocentre and the planned  $D99/Rx$ . When the planned  $D99/Rx$  is less than one it is often an indication that the target coverage is compromised because of proximity to OARs. As shown in Figure 5-3, plans with higher prescription dose ( $Rx \geq 54$  Gy) are more likely to have higher  $D99/Rx$  (hotter plan) for the target whether or not the tracking errors are taken into consideration. Plans with lower prescription doses are often for larger targets and/or targets closer to OARs (see Section 1.3.3). Hotter plans are also less sensitive to uncorrected rotations and how far the target

is away from the imaging isocentre. The standard deviation of the correlation errors is related to dose blurring and therefore plays an important role in the magnitude of treated D99/Rx.

When the system does not track rotations, the robustness of the plan against simulated large rotational offsets (quantified by D99/Rx) can be performed prior to treatment to inform the potential geographical miss or insufficient dose coverage. We examined the sensitivity of all 148 patient plans (192 targets) by simulating rotational errors about different axes. We found that 77% and 72% of targets had  $D99/Rx > 1$  for rotations  $\leq 3^\circ$ , if rigid body errors are neglected and considered, respectively. Many of these plans already had coverage of the target compromised for OARs in the planned dose distributions (17% had planned  $D99/Rx < 1$ ), suggesting a patient-specific assessment may be helpful for these treatments.

Two case studies (patient #1 and 19) also illustrate that we can proactively generate a plan more robust to delivery errors. To proactively reduce the impact of delivery error, one should place the imaging isocentre close to the target centroid during planning, increase the planned GTV D99/Rx (hotter plan), and adjust the patient setup to reduce the rotational offset. Sometimes it can be challenging to place the imaging isocentre in the center of a target because of couch height limitations and the potential for a collision between the patient, robotic arm, linac, and other components in the room. The CyberKnife system has the capability to run the plan in a demo mode, in which no radiation is delivered but the robot moves through all the beam positions in a plan while a user is in the room. The plan deliverability and potential for collision can be assessed prior to treatment and replanned if necessary. Three case studies (patients #19, 32, and 65) illustrated the potential for improving treated D99/Rx by changing different parameters during treatment and delivery and provided solution(s) to improve the plan robustness against different types of delivery errors.

Using two case studies shown in Figure 5-6 and Figure 5-7, we illustrate how a dosimetric assessment can provide valuable information for evaluating the suitability of a selected PTV margin. We demonstrated how these tools could be adapted for the early warning system presented in Chapter 3 that assesses the potential for geometrical miss of a target. The off-line computation generally takes less than ten minutes for every patient. It provides an independent dosimetric assessment prior to each new fraction of treatment which could be used for off-line dose adaptation.

We show two ways to illustrate geometric and dosimetric coverage at the end of each fraction. One method keeps the planned dose static but translates target voxels taking into account delivery errors. This is similar to the GTV voxel occupancy probability described in Chapter 3 that only included translation error, but with the addition of rotational and rigid body errors. This method of analysis is shown for two different patients in Figure 5-6 and Figure 5-7. The other method perturbs the planned dose and assesses dose coverage metrics for the target such as D99 and D95. Both types of dosimetric assessments could be used prior to treatment to evaluate margin suitability and provide an estimate of an individual plan's sensitivity to error.

The dose perturbation tools are developed to support a future prospective trial that will employ patient-specific PTV margins for liver SBRT treatments on this system.

## 5.2. Trigeminal Neuralgia

### 5.2.1. Introduction

Stereotactic radiosurgery (SRS) is used to treat some non-malignant disorders, such as intracranial arteriovenous malformation, meningioma (slow-growing lesions), acoustic neuroma, pituitary adenoma, trigeminal neuralgia, and epilepsy.<sup>237</sup> The trigeminal nerve (cranial nerve 5, denoted as CN5) has three major branches for chewing, biting, and facial sensation. Trigeminal neuralgia (TN) is a rare nerve disorder that causes severe facial pain. One of the treatments for this disorder is to ablate a segment of CN5 using radiosurgery. At the Ottawa Hospital Cancer Centre, 123 patients were treated on CyberKnife for this disease from 2010 to 2018. These treatments require a high degree of spatial accuracy and delivering a small ellipsoidal shaped isodose volume using non-isocentric delivery with the 5 and 7.5 mm collimator. Using this technique, one can safely deliver 50 to 60 Gy in a single fraction to a 4 to 10 mm length of nerve typically about three millimetres wide.<sup>238</sup>

Studies by Romanelli *et al.*<sup>239,240</sup> and Villavicencio *et al.*<sup>241</sup> have reported on the efficacy of CyberKnife SRS for the treatment of trigeminal neuralgia. Villavicencio *et al.*<sup>241</sup> reported a preliminary multi-center study in which 67% of patients (N = 95) experienced initial pain relief, and 50% continued pain relief in a two-year follow-up. Those patients were treated with a median maximum dose of 78 Gy. The overall complication rate in their study was 18% (e.g., decreased corneal reflexes, anesthesia dolorosa, lockjaw, and masticator weakness). Romanelli *et al.*<sup>239</sup> reported on a large single-center patient cohort, with 560 treatments for 527 patients with more than three years follow-up for 387 patients. In their study, a six-millimetre segment of the nerve was treated with a single-fraction dose of 60-65 Gy and 40-45 Gy, for the first treatment and for retreatment, respectively. They reported 12.8% of cases required a retreatment with a pain relief

rate of 76% for 343 patients within the follow-up period. In their study, technical details about the delivery were also included. It was stated that the average treatment duration was 50 minutes and that the head path was used so that rotations could be corrected.

As mentioned in Section 1.3.1, CyberKnife also has a specialized path set intended for trigeminal neuralgia. This path set brings the linac closer to the patient to more efficiently deliver the high dose required to ablate the nerve. The distance between the nodes and imaging isocenter is approximately 65 cm for the trigeminal path compared to 80 cm for the head path. The beam directions for this path are also optimized to confine the high dose to the cerebrospinal fluid (CSF) space around the brainstem (see Figure 5-8b). We have observed that for the same patient we can achieve better coverage of the nerve and lower planned dose to the brainstem and temporal lobe using the trigeminal path compared to the head-path. We, therefore, use the trigeminal path set for all TN treatments at our centre. However, rotational corrections cannot be applied without risk of collisions for this path.

Using univariate and multivariate regression analyses, Conti *et al.*<sup>143</sup> investigated in a multicentre study the factors for TN recurrence and radiation-induced sensory disturbance for 296 TN cases treated on CyberKnife. They found multiple sclerosis, a higher integrated dose over nerve target volume (mean dose times target nerve volume), and mean dose to target were related to pain relief during follow-up (median: 38 months). Their study implied that dose and target volume could be very important in treatment outcomes. Du *et al.*<sup>238</sup> studied the impacts of anatomical variables on the dosimetry of the brainstem for forty patients treated on a conventional linac cone-based radiosurgery. They found that the features important to treatment outcomes were spatial nerve orientation relative to the brainstem, the distance between the isocenter and brainstem, and the utilized cone size. Borchers *et al.*<sup>242</sup> examined the relationship between the success of TN

treatments and different anatomic and morphological features for 46 patients treated on CyberKnife. They found that having an angle of nerve takeoff from the brainstem to Meckel's cave of  $150^\circ$  to  $170^\circ$  relative to the sagittal plane was correlated with improved pain relief ( $P = 0.03$ ). Flickinger *et al.*<sup>243</sup> studied treatment outcomes for 87 patients (group 1: group 2 = 44: 43) treated using Gamma Knife, a  $^{60}\text{Co}$  multibeam frame-based radiosurgery system. They found that increasing the length of the nerve included in the treatment volume did not significantly raise pain relief for 87 patients (group 1: group 2 = 44: 43). They compared pain relief for two groups of patients with different average nerve lengths within the 50% isodose volume: one group with  $5.4 \pm 0.4$  mm (N=44) and the other with  $8.7 \pm 1.1$  mm (N=43). With a median follow-up of 26 months, they found younger age and fewer surgical procedures, and not the target length, were related to pain relief.

To give the reader a sense of size and the differences between these treatments, the dose distributions on the planning CT for liver and trigeminal nerve are compared in Figure 5-8. Uncorrected rotations can introduce significant dose reduction for TN treatments and can lead to high dose to nearby organs at risk such as the brainstem. The experiment with a small target delivered using a 5-mm collimator using the trigeminal path discussed in Section 4.3.1 also illustrated the importance of minimizing uncorrected rotations during treatment. Since 2013 an internal policy has been in place at our clinic to maintain the reported rotations to be less than  $0.5^\circ$  in any direction for TN treatments. The recommendation is based on simple geometric considerations and a rough estimate of the typical distance between the imaging isocenter and the trigeminal nerve. This decision resulted in longer treatment times for TN (typically about two hours) than for a typical brain tumour (approximately one hour).

We have applied the dose perturbation tool validated in Chapter 4 to demonstrate how this tool could be used to answer three clinical questions about these treatments delivered under our clinical current policy. First, what is the length and volume of CN5 receiving the prescription dose if one considers uncorrected rotations? Secondly, is the treated dose to the brainstem safe once we take into account delivery error? Thirdly, could we use different tolerances on uncorrected rotations for TN treatments to either make the delivery more efficient or to more accurately target CN5?

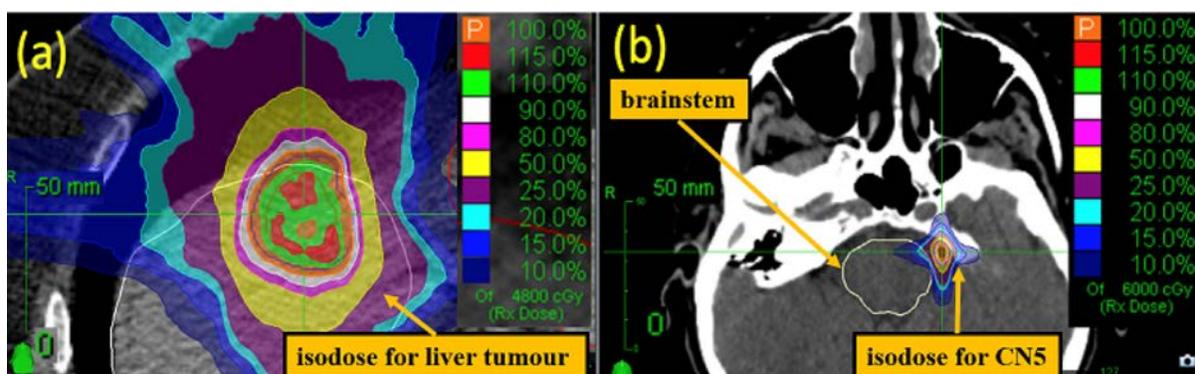


Figure 5-8. Transverse slices through the planned dose distributions and planning CT for representative liver and trigeminal neuralgia (TN) cases. (a) The treatment of a 23 cm<sup>3</sup> PTV at the dome of the liver using 15, 25, and 35 mm diameter collimators. (b) The TN target of 0.04 cm<sup>3</sup> located on the left trigeminal nerve (CN5). Treatment uses a 5 mm collimator and CyberKnife’s specialized TN path set for which rotational corrections are not possible. The brainstem is also shown to illustrate how close it is to the CN5 target.

### 5.2.2. Method and Materials

The contours, planned dose, and treatment delivery XML and log files are extracted for 27 treatments between 2011 and 2018. The typical prescription dose is 60 Gy in a single fraction, delivered to approximately 6 mm length of CN5 near the sensory root, excluding part of nerve 2 to 3 mm from the brainstem. For patients with recurrent pain or poor response, retreatment with 50 Gy in one fraction may be prescribed at least one year after the first treatment. For the brainstem,

our institutional treatment planning guideline requires  $V_{12Gy} \leq 0.3$  cc and  $V_{15Gy} \leq 0.03$  cc as a planning goal. Timmerman recommends  $V_{10Gy} \leq 1.00$  cc and a maximum dose of 15Gy.<sup>9</sup>

The contours are created using MR images. When MRI cannot be performed on a patient, a CT cisternogram is needed, which requires the contrast agent injection into the cerebrospinal fluid (CSF) via a spinal tap.<sup>238,244</sup> The radiation oncologists contour the OARs, including the brainstem, cochlea, ipsilateral facial nerve (CN7), auditory vestibular nerve (CN8), ipsilateral temporal lobe, as well as the target. The length of CN5 selected as the target is somewhat arbitrary, so once delivery errors are taken into consideration, portions of the prescribed isodose volume could fall outside of the target but still overlap with CN5. To evaluate this possibility, the entire length of CN5 from the sensory root entering the brainstem to the gasserian ganglion is therefore contoured by a second radiation oncologist for those 27 treatments (see Figure 5-9).

Since the nerve axis is at an angle with respect to the patient's anatomical axes, principal component analysis (PCA) is applied to voxels of the full nerve contour to determine the nerve orientation. Principal component 1 (PC1) explains 84.6% ( $\pm 4.7\%$ ) of data variability for all voxels so that PC1 can be considered the longitudinal axis of the nerve. Since the original target and the longer length of CN5 were delineated by different physicians, the CN5 contour and target can be slightly misaligned, as illustrated in Figure 5-9 (b). Because of this misalignment issue, the union of target and CN5 contours is considered as the full nerve contour. We exclude three treatments because the explained variance of PC1 is below 0.75 for them, indicating the nerve and target (high dose) are not aligned well (see Figure 5-9b). The average and standard deviation of the fraction of the full nerve within the target is  $0.78 \pm 0.26$  for the remaining 24 treatments. Three of those were retreatments. The length of the nerve is calculated as the full width at tenth maximum (FWTM) of the distribution of PC1 (histogram in Figure 5-9c). We calculate the effective length (denoted as

$L_{\text{eff}}$ ) of the full nerve receiving the prescription dose (Rx) by accumulating the nerve cross-sections in PC 2 and PC 3, if over half of the voxels for each cross-section receives more than Rx dose.

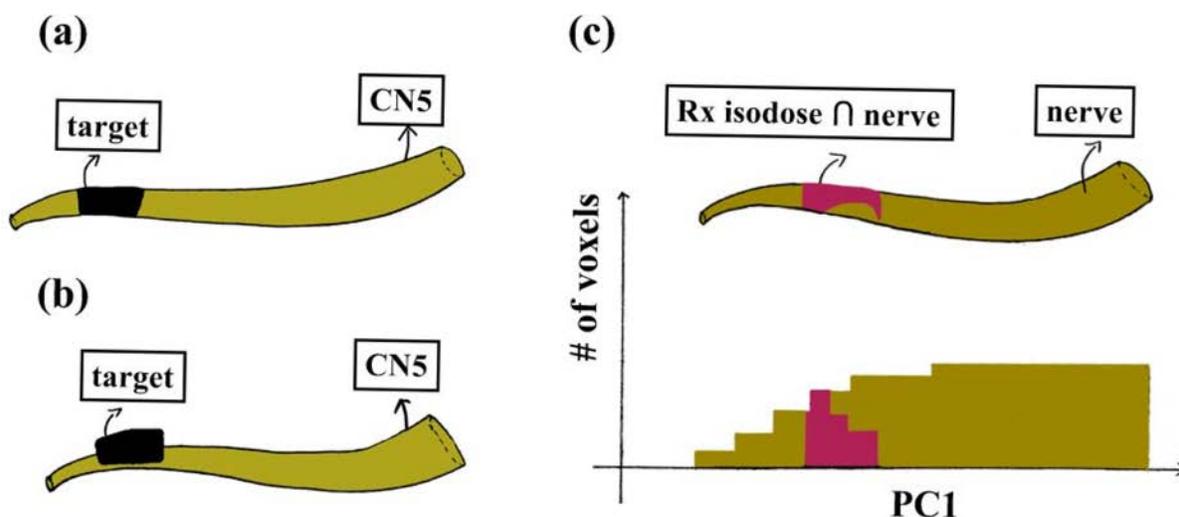


Figure 5-9. (a, b) Illustration of the good (a) and poor (b) alignment of the longer length of trigeminal nerve (denoted as “CN5”) and the treatment target. The original target and the longer length of CN5 are delineated by different physicians. In three cases, the CN5 contour and target are slightly misaligned as illustrated in subplot (b). Demonstration of the union of the CN5 and target voxels (denoted as “nerve”) projected onto their principal component 1 (PC1) axis using the principal component analysis. The bins in red indicate that part of these cross-sections (PC2 and PC3) receive the prescription dose (denoted as “Rx”).

Using the 6D skull tracking, the system calculates the translational and rotational corrections after the acquisition of each x-ray image pair (approximately one pair every 30 seconds). In the perturbation model, the translational tracking residual errors include beam positioning errors based on end-to-end QA measurements for this path set and the differences between translational corrections from one image to the next. The latter type of error approximates the maximum distance the patient moves within their immobilization in the time between two x-ray acquisitions. All rotational correction residual errors are non-zero in the perturbation model since they are not corrected for using the trigeminal path. We evaluate: the *volume* of the full nerve covered by Rx isodose denoted as “nerve  $\cap$  Rx; the *length* of nerve covered by Rx  $L_{\text{eff}}$  as defined above; and three

dose-volume metrics relevant to the brainstem (V10Gy, V12Gy, V15Gy). We do this assessment retrospectively for 27 previous treatments where a radiation oncologist has contoured a longer segment of CN5. We also demonstrate a workflow, shown in Figure 5-10, which could be used to assess the sensitivity of a particular plan for simulated angles of rotations prior to treatment. The results could be used to define patient-specific rotational tolerances that meet treatment plan objectives. We show how the plan sensitivity against uncorrected rotations and translational tracking errors can also be assessed prior to treatment using the same dosimetric parameters. For a prospective TN treatment, the image-to-image residual translational errors would not be available for this single fraction treatment. The pre-treatment assessment uses translational errors based on only end-to-end errors. The assessed simulated angles were between  $-1.5^{\circ}$  and  $1.5^{\circ}$  for roll and pitch in steps of  $0.25^{\circ}$ , and  $-2^{\circ}$  and  $2^{\circ}$  for yaw in steps of  $0.25^{\circ}$ . One could also complete a dosimetric assessment after every treatment course for quality control purposes, as also shown in Figure 5-10. This assessment could be used to retrospectively correlate treated dose accuracy with pain relief efficacy and neuro-toxicity.

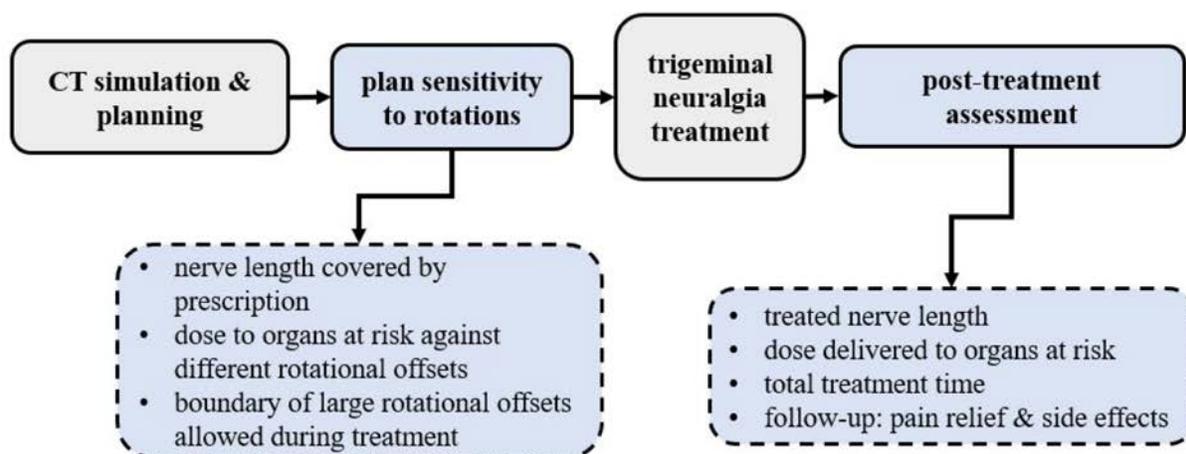


Figure 5-10. The workflow to demonstrate how to employ patient-specific dosimetric assessments prior to treatment to inform allowable rotations for prospective trigeminal neuralgia treatments on CyberKnife. Following, treatment, a dosimetric assessment could be done that considers the observed motion tracking related uncertainties. The results could be correlated in a retrospective study with treatment outcomes and normal tissue complications.

### 5.2.3. Results

Except for two cases treated before 2013, all patients had rotations  $< 0.5^\circ$  during treatment. The volumes of the target and full nerve are  $48 \pm 22 \text{ mm}^3$  and  $117 \pm 64 \text{ mm}^3$ , respectively, for 24 treatments. After accounting for observed rotations and translational errors to the target,  $L_{\text{eff}}$  and the toxicity to the brainstem are illustrated in Figure 5-11. By maintaining target rotations  $< 0.5^\circ$  during treatment,  $L_{\text{eff}} = 5 \pm 2 \text{ mm}$  if the delivery errors are considered compared to  $L_{\text{eff}} = 6 \pm 2 \text{ mm}$  for planning. As an approximate check of the nerve length calculation, we also estimated  $L_{\text{eff}}$  from nerve  $\cap$  Rx volumes assuming the nerve is a cylinder with a 3 mm diameter. From this rough estimate,  $L_{\text{eff}}$  would be  $7 \pm 4 \text{ mm}$  and  $6 \pm 4 \text{ mm}$ , respectively, for planned and treated lengths just slightly higher than the results estimated using principal component analysis.

The dose constraints for the treated dose to the brainstem are met for 23 out of 24 patients (see Figure 5-11b). Only one case had  $V_{15\text{Gy}} > 0.03 \text{ cc}$  when delivery errors were considered. However, for this case it already exceeded tolerance with  $V_{15\text{Gy}} = 0.04 \text{ cc}$  for the planned dose.

## Chapter 5. Dosimetric Assessments

---

The treated  $L_{\text{eff}}$  is reduced by  $1 \pm 1$  mm (or  $7\% \pm 21\%$ ) on average, compared to the planned  $L_{\text{eff}}$ , as illustrated in Figure 5-12 (a). However, Figure 5-12 (b) shows that the overlapped volume between the nerve and Rx isodose ( $\text{nerve} \cap \text{Rx}$ ) is not very sensitive to delivery errors. The volume overlap ( $\text{nerve} \cap \text{Rx}$ ) is reduced by  $(9.1 \pm 7.2)\%$  from  $47 \pm 26$  mm<sup>3</sup> to  $43 \pm 26$  mm<sup>3</sup> on average for 24 treatments considering the translational and rotational errors.

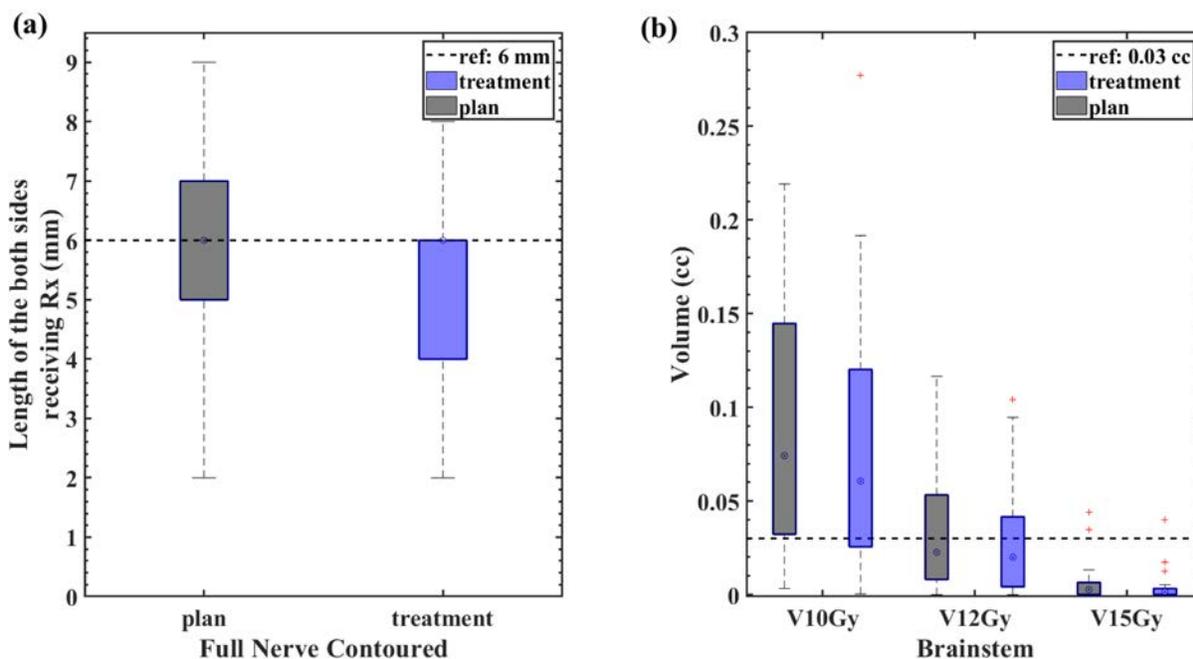


Figure 5-11. The planned and treated length of the nerve covered by Rx (a) and the brainstem dose (b) for 24 trigeminal neuralgia treatments.

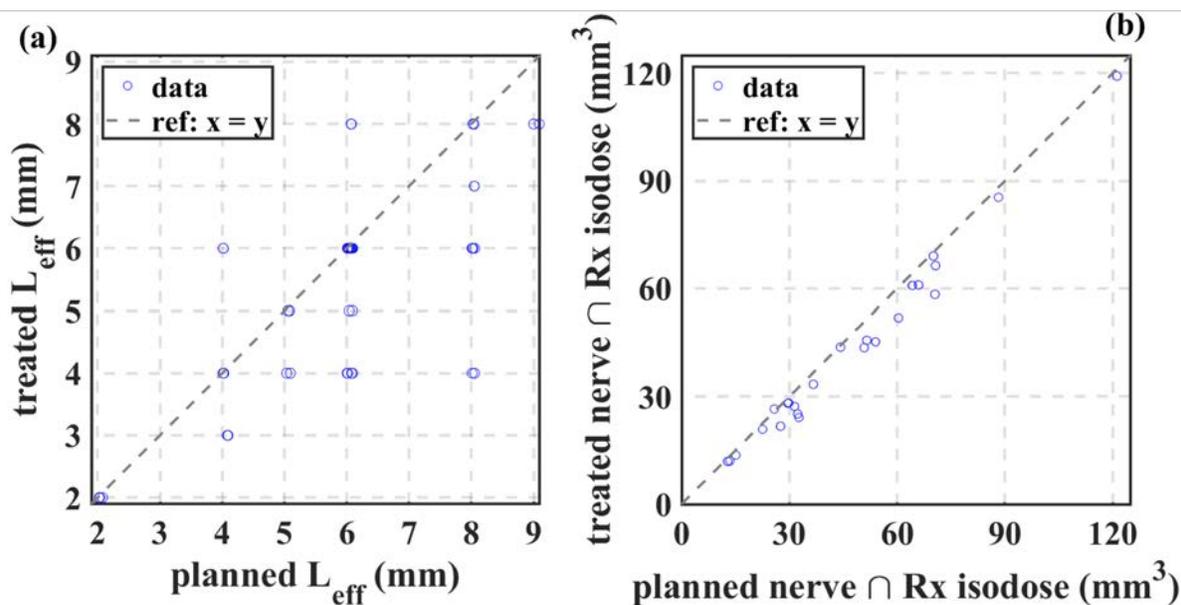


Figure 5-12. (a) The treated and planned nerve length covered by prescription dose (Rx). (b) The planned and treated volume overlap between the nerve and Rx isodose.

Since the target is either left-sided or right-sided, the direction of rotation is also important when considering the dosimetric impact of uncorrected rotations. Figure 5-13 suggests that larger

uncorrected rotations (up to  $\pm 1^\circ$ ) can be acceptable during treatments for some patients if a 4 mm length treated nerve (i.e.,  $L_{\text{eff}} > 4$  mm) and the brainstem  $V15\text{Gy} > 0.03$  cc is acceptable. As shown in Figure 5-13, for left-sided targets, the brainstem is more sensitive to rotations in roll to the patient right and yaw in the counterclockwise (toward the patient's right-hand side) direction. These signs are reversed for roll and yaw for right-sided nerves. For both right- and left- sided nerves brainstem dose increases for pitch in the head-down direction. Table 5-3 shows a patient-specific assessment of each treatment plan included in this study. For each plan we assess the maximum angles permissible in order to meet the planning goals of  $L_{\text{eff}} \geq 4$  mm and  $V15\text{Gy} < 0.03$  cc.

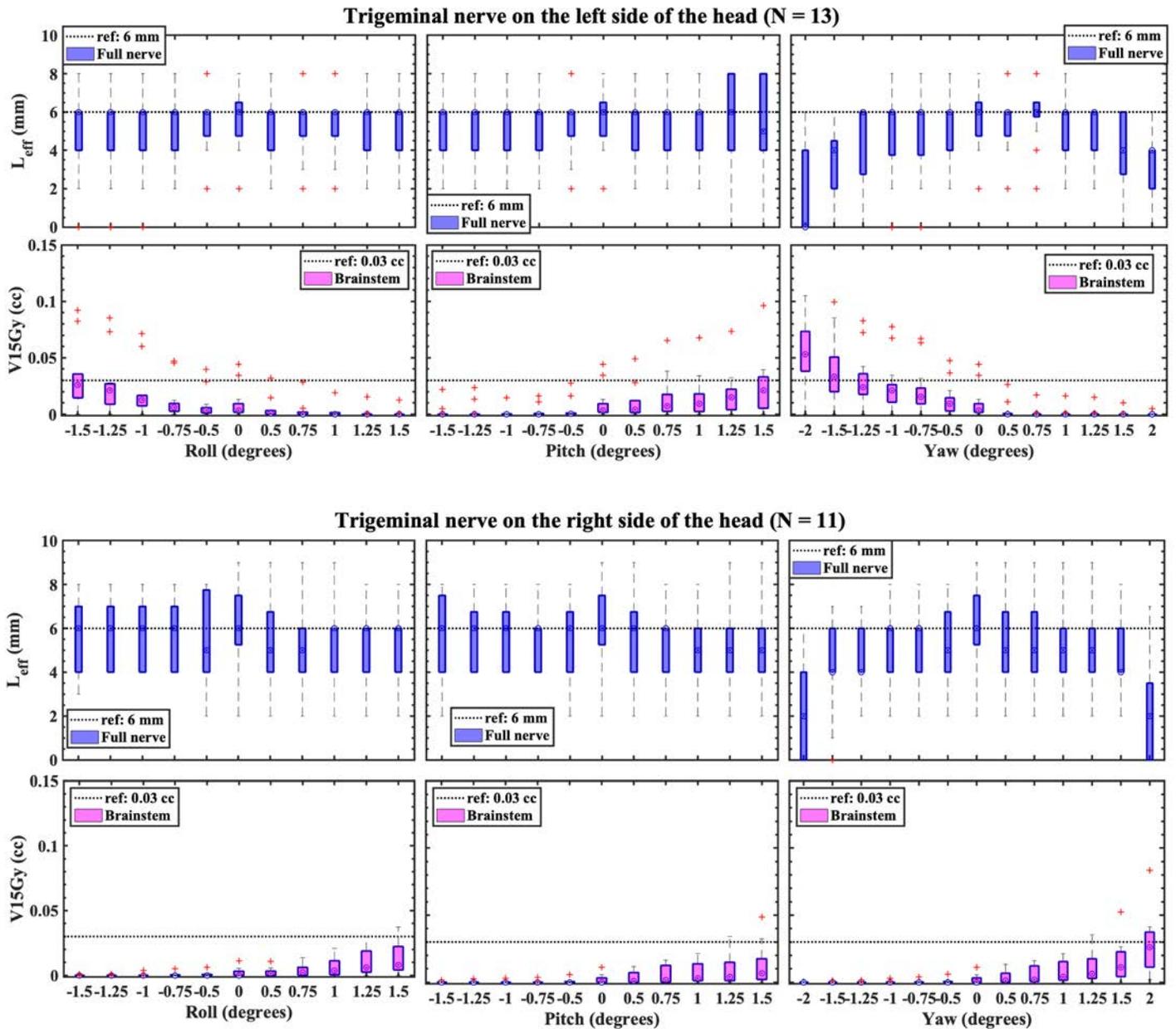


Figure 5-13. Plan sensitivity to rotational offsets for the length of the left-sided (N = 13) and right-sided (N = 11) nerve covered by prescription dose (Rx) and toxicity to the brainstem. The positive direction of roll, pitch, and yaw offsets are roll to the patient left, head down, and clockwise viewed from the ceiling, respectively. Each box spans from the first to the third quartile, and the circle in the box indicates the median.

Chapter 5. Dosimetric Assessments

Table 5-3. Characteristics of the patient-specific dosimetric assessment prior to a trigeminal neuralgia treatment. The length of the nerve covered by the prescription (Rx) and the dose to the brainstem (V15Gy and V12Gy) are quantified using the planned dose distribution. The range of target rotations along a specified direction of rotation (roll, pitch, or yaw) that allows at least 4 mm of the nerve covered by Rx and V15Gy of the brainstem less than 0.03 cc is investigated. “N/A” indicates that no angle is found that meets these two criteria within the search range of (-1.5°, 1.5°). The planned dose distribution is perturbed to account for translational tracking residual errors and the uncorrected target rotation.

unilaterality	patient index	prescription (Gy)	trigeminal nerve				brainstem		range of rotations (in degrees) allowed for $L_{eff} \geq 4$ mm and V15Gy < 0.03 cc		
			linearity of the nerve	Rx-covered nerve length (mm)	fraction of nerve-target overlap within target	Rx-covered nerve volume (mm <sup>3</sup> )	V15Gy (cc)	V12Gy (cc)	roll (min, max)	pitch (min, max)	yaw (min, max)
left sided	1	60	0.90	6	0.87	33	0.04	0.1	(0.75, 1.50)	(-1.50, -0.50)	(0.50, 1.50)
	2	60	0.84	6	0.65	44	0.00	0.0	(-1.50, 1.50)	(-1.50, 1.50)	(-1.50, 1.50)
	3	60	0.82	8	0.37	30	0.00	0.0	(-1.50, 1.50)	(-1.50, 1.50)	(-1.50, 1.50)
	4	60	0.87	6	0.97	37	0.01	0.1	(-1.25, 1.50)	(-1.50, 1.50)	(-1.00, 1.25)
	5	60	0.88	6	0.98	31	0.00	0.0	(-1.25, 1.50)	(-1.50, 1.50)	(-1.25, 1.50)
	6	60	0.79	6	0.30	26	0.00	0.0	(-1.50, 1.50)	(-1.50, 1.50)	(-1.25, 1.50)
	7	60	0.91	6	1.00	60	0.03	0.1	(-0.50, 1.50)	(-1.50, 0.50)	(0.50, 1.50)
	8	60	0.92	8	0.99	70	0.01	0.0	(0.50, 1.50)	(-1.50, 1.50)	(-0.50, 1.25)
	9	60	0.89	2	1.00	32	0.00	0.0	N/A	N/A	N/A
	10	60	0.81	4	0.98	51	0.01	0.1	(-1.50, 1.50)	(-1.50, 1.00)	(0.50, 1.00)
	11	60	0.89	8	0.94	88	0.00	0.0	(-1.50, 1.50)	(-1.50, 1.50)	(-1.00, 1.50)
	12	60	0.82	5	0.99	71	0.01	0.1	(-1.25, 1.50)	(-1.50, 1.25)	(-0.50, 1.50)
	13	50	0.76	4	0.90	70	0.00	0.0	(-1.50, 1.50)	(-1.50, 1.50)	(-1.50, 1.50)

Chapter 5. Dosimetric Assessments

<b>right sided</b>	14	60	0.43	4	0.88	15	0.01	0.0	(-1.50, 1.50)	(-1.50, 1.50)	(-1.50, 1.50)
	15	50	0.43	6	0.87	23	0.00	0.0	(-1.50, 1.50)	(-1.50, 1.50)	(-1.50, 1.50)
	16	50	0.51	8	0.85	30	0.00	0.0	(-1.50, 1.50)	(-1.50, 1.50)	(-1.50, 1.50)
	17	60	0.38	5	0.81	13	0.01	0.0	(-1.25, 1.25)	(-1.25, 1.25)	(-1.50, 1.50)
	18	60	0.38	6	0.82	13	0.00	0.0	(-1.50, 1.50)	(-1.50, 1.50)	(-1.50, 1.50)
	19	60	0.95	6	0.76	66	0.00	0.0	(-1.50, 1.50)	(-1.50, 1.50)	(-1.50, 1.50)
	20	60	0.84	6	0.81	51	0.00	0.0	(-1.50, 1.50)	(-1.50, 1.50)	(-1.50, 1.50)
	21	60	0.98	6	0.88	54	0.00	0.0	(-1.50, 1.25)	(-1.50, 0.75)	(-1.50, 1.25)
	22	60	0.98	8	0.79	64	0.00	0.0	(-1.50, 1.50)	(-1.50, 1.50)	(-1.25, 1.50)
	23	60	0.86	9	0.83	121	0.00	0.1	(-1.50, 1.50)	(-1.50, 1.00)	(-1.50, 1.00)
	24	60	0.98	2	0.89	28	0.00	0.0	N/A	N/A	N/A

#### 5.2.4. Discussions

The goal of TN treatments is to damage a sufficient length of CN5 to reduce or eliminate facial pain. We have assessed the volume and length of nerve covered by the prescription dose for treatments using the trigeminal path. We evaluate the length of nerve,  $L_{\text{eff}}$ , that receives the prescription for both planned and treated dose. The treated  $L_{\text{eff}} = 5 \pm 2$  mm if the delivery errors are considered, and are on average  $1 \pm 1$  mm smaller than the length of nerve covered in the planned dose.

To our knowledge, this is the first study of the impact of delivery errors on radiosurgical TN treatments. The preliminary analysis for this small sample size ( $N = 24$ ) demonstrates the feasibility of the approach. A larger dataset with the full nerve contoured would be required for statistical hypothesis testing. Some areas which could be explored include the question if the reduction in treated nerve volume or  $L_{\text{eff}}$  is significantly correlated with tracking errors or with treatment outcomes (e.g., pain relief, and the time between treatment and retreatment).

Romanelli *et al.*<sup>239</sup> reported on dosimetric parameters and outcomes for 387 TN treatment plans. As mentioned previously, for their plans they used the full head-path, where the CyberKnife robot could correct for rotations. The reported treatment durations were on average 50 minutes (no range or standard deviation reported), which is shorter than ours which can often take up to two hours (on average  $75 \pm 19$  min for our patient cohort) using the trigeminal path while maintaining rotational corrections under  $0.5^\circ$ . The planning goals seem to be similar to ours, but they allowed a slightly higher brainstem dose: near-maximum dose to  $0.035 \text{ cm}^3$  of the brainstem at 30% of the prescription dose (18 Gy for a 60 Gy Rx). They reported that the length of nerve covered by the Rx isodose line was assessed using the planning system for each treatment, with the goal of covering a 6 mm segment of nerve. However, 49% of patients (169/343) had only 4 to 5 mm of

nerve covered due to a shorter cisternal segment of the trigeminal nerve. The median prescription dose and target volume were  $60 \pm 5$  Gy and  $27 \pm 17$  mm<sup>3</sup>, respectively. In contrast, in our study (N = 24) the average prescription dose and target volume are  $59 \pm 3$  Gy and  $48 \pm 22$  mm<sup>3</sup> (median = 45 mm), respectively. The target volumes evaluated in our study are larger than those in Romanelli *et al.*'s study, suggesting that our planning goals may be more difficult to meet using the head-path.

From the dosimetric sensitivity data shown in Figure 5-13 and Table 5-3, allowing uncorrected rotations up to 1° appears to be safe for most patients, if the treated  $L_{\text{eff}} \geq 4$  mm for the trigeminal nerve is acceptable and a brainstem  $V_{15\text{Gy}} \leq 0.03$  cc is achieved in planning. The ranges of rotations that meet these criteria are patient-specific. Nine out of 24 of these treatments still meet the criteria up until 1.5° corrections in all directions. However, in this study the rotational error was only applied to the target along one anatomical axis at a time. We estimate that a more thorough calculation for each plan including all possible permutations of roll, pitch, and yaw angles from -1.0° to 1.0° for a step of 0.5° (i.e., 125 possible permutations) would require approximately 75 minutes.

The reduction in treated  $L_{\text{eff}}$  is affected by many parameters, including the magnitude of uncorrected rotations, the isocenter to nerve distance, the dose gradient, and the target to brainstem distance. The rigid nature of structures in the brain allows us to evaluate OAR dose more confidently for TN than for liver treatments. We consider cases for left- and right-sided nerves separately, since the high dose volume may rotate away or toward OARs depending on target laterality. The sensitivity study to rotational errors could be performed as a pre-treatment quality assurance test for each plan. This could allow for more informed tolerances on rotations for the trigeminal path set. Treatments with the smallest targets (e.g., #9 and 24) tend to have a smaller

region of CSF between the brainstem and gasserian ganglion. As shown in Table 5-3, the length of CN5 covered by the planned dose distribution is 2 mm for both treatment #9 and #24. It might be safer to use the head path for these patients in order to allow for rotations to be corrected.

### **5.3. Conclusions**

This study retrospectively assessed the dosimetric impact of delivery errors for liver and trigeminal neuralgia (TN) treatments on CyberKnife. For liver treatments with a relatively large sample size ( $N = 148$ ), the impacts of motion tracking errors, target rotations, and rigid body error on dose to targets have been quantified by using the treated D99/Rx for GTV(s). As a feasibility study, we also applied the dose perturbation tool to TN treatments, a disease site with very different characteristics and considerations. For a small sample size ( $N = 24$ ) of TN plans, we assessed the treated length/volume of the nerve and dose to the brainstem considering translational and rotational errors. For both liver and TN sites, we tested the plan robustness as measured by the dosimetric parameters against simulated angles of rotation and translational errors.

The dose perturbation tool, which can be used to retrospectively evaluate the suitability of a specific margin for a patient population or as per-patient sensitivity of dosimetric parameters to error assessment prior to treatment, enables adaptive planning processes and patient-specific margins for liver, and provides more informed tolerances on allowable uncorrected rotations for trigeminal neuralgia.

## **Chapter 6. Final Overview and Future Work**

### **6.1. Summary**

This study has retrospectively quantified each of the different types of delivery errors for liver treatments using the CyberKnife robotic radiosurgery system. Based on an analysis of the system log files for 72 patients, uniform 4 mm isotropic planning target volume (PTV) margins are required to account for the residual errors associated with motion tracking, beam positioning errors, and target deformations (rigid body errors) for 95% of this population. These margins are only sufficient if the impact of uncorrected rotational errors can be minimized. This can be achieved by moving the imaging isocenter as close as possible to the center of the target and implanting at least four fiducials around each target. For this patient cohort, the intra-patient variations in both respiratory motion tracking (Synchrony<sup>TM</sup>) errors and rigid body errors were quite small. This suggested that patient-specific margins might be a possibility based on a dry-run session in which the system's motion tracking accuracy for a particular patient is assessed prior to planning.

Based on these results, we propose a clinical workflow for these treatments using individualized PTV margins with a volumetric tracking error assessment after each fraction. A supervised machine learning algorithm is applied to retrospective data for 148 treatment courses. Data from the first fraction of a treatment course is used to emulate the dry-run session. We classify individual treatments into binary categories: either high or low Synchrony error using a support vector machine. Based on fourteen features related to motion-tracking accuracy, the maximum Synchrony error is predicted for all remaining respiratory motion models in a patient's treatment course. For patients predicted to have low maximum Synchrony errors, a reduced PTV margin can be used. Several methods for testing the robustness and accuracy of the classification model are employed including five-fold cross-validation, leave-one-out cross-validation, 500 repeated

random subsampling cross-validation and the receiver operating characteristic (ROC) metric. The classification is also independently cross-validated for a cohort of 48 treatment plans for other anatomical sites which also use Synchrony respiratory motion compensation. To mitigate the risk of geographical miss associated with margin reduction, we propose an “early warning system” that evaluates PTV suitability by perturbing GTV voxels by the observed motion tracking errors recorded in system log files. In this retrospective analysis, we find a total of 36% (53/148) of patients could have been treated using smaller margins with fractional GTV voxel occupancy greater than 0.99 for all fractions. One limitation identified in the early warning system is that it does not explicitly model the dosimetric impact of tracking errors or incorporate uncorrected rotations. Both effects are important considerations, which should be taken into account before prospectively reducing treatment margins for patients.

Therefore, we develop a dosimetric perturbation model which includes uncorrected rotations, motion-tracking residual errors, and target deformation. In principle, the model could be applied to any treatment on the CyberKnife system. To demonstrate its general applicability, we apply the dose perturbation tool for liver and trigeminal neuralgia treatments. They are two very different disease sites with unique delivery characteristics. The model is validated using film measurements in sixteen different experimental conditions including: in-phase and out-of-phase internal and external breathing motion patterns; with and without uncorrected rotations; for homogenous and heterogeneous phantoms emulating lung and liver patient anatomies. The measured and perturbed dose profile widths were only 1 to 2 mm different than the planned dose for all but one experiment. This suggests that, in most cases, the dose blurring associated with motion tracking errors is a relatively small effect. Large uncorrected rotations cause more significant changes to the delivered

dose. The perturbed and measured doses agree well with 2%/2mm criteria gamma pass rates above 95% for fifteen out of sixteen experimental runs.

Using this validated perturbation model, we retrospectively assess the dose which was delivered to targets for 70 liver treatment courses which had rotational corrections that could be accurately calculated and tracked by the system. In the perturbed dose distribution, we find that approximately one-third of targets are under-dosed, which we define as having a treated  $D99/Rx < 1$  for the GTV. If rigid body errors are neglected, about one-fifth of targets are under-dosed. Accurate rotational corrections are often not available if too few fiducials are visible in the live images or if the spatial arrangements of the fiducials are inadequate. For 148 treatment courses, we evaluated the treated  $D99/Rx$  while applying different constant rotational residual errors about each axis. For patients with average rotations  $\leq 3^\circ$ , 72% targets have treated  $D99/Rx > 1$ . However, many of these patients already had  $D99/Rx < 1$  for the planned dose distributions indicating that coverage was compromised to spare nearby OARs. The dosimetric sensitivity to uncorrected rotations depends on many factors including: the planned dose in the central region of the PTV; compromised coverage due to proximity of targets to OARs; the size of rotational corrections; and the distance between the target and imaging isocenter. For three representative cases, we demonstrate different ways the treatment can be made less sensitive to tracking errors using the dose perturbation tool. We also show how this tool can also be used to evaluate the suitability of a selected PTV margin as part of the early warning system proposed for the prospective margin reduction study.

For a relatively small sample size ( $N = 24$ ) of trigeminal neuralgia (TN) treatments, we demonstrated how the dose perturbation model could be used for other sites. We assessed the change in length and volume of nerve treated to the prescribed dose and the dose to the brainstem

in the perturbed dose distributions taking into account uncorrected rotations. For these treatments, target rotations of up to  $1^\circ$  can be safe for some patients if at least 4 mm of the nerve receives the prescription, and less than 0.03 cc of the brainstem receives more than 15 Gy. This work can provide more informed tolerances on allowable uncorrected rotations for trigeminal neuralgia treatment using the CyberKnife system.

The significance of this study is to prepare analytical tools (including the classification model, the early warning system, and the dose perturbation model) for a proposed margin reduction clinical trial, which implements patient-specific PTV margins for liver treatments safely and efficiently in the clinic. Some of the results have changed clinical treatment planning practices at the Ottawa Hospital Cancer Centre, including placing the imaging isocenter as close as possible to the treatment target center to reduce the impacts of uncorrected rotations.

## **6.2. Future Work**

The focus of this thesis was to develop the tools to safely implement the patient-specific margin reduction trial proposed in Section 3.2.4, assisted with the dose perturbation tool proposed in Section 4.2.4. Some patients can be treated with a reduced PTV margin for reduced chance of complication, and for the potential of an increased prescription dose to achieve better local control. The increase in PTV D99 which could be realized with a 2 mm smaller margin was indirectly demonstrated (making no change to the planned dose) in Figure 3-3. This analysis does not show the reduction in OAR dose which could also be an advantage of margin reduction. A decision must be made if this trial will involve dose escalation or only margin reduction while keeping the prescription dose constant. A retrospective planning study to better demonstrate the changes in the planned dose that could result from reducing the margins could help resolve this issue. It would also be interesting to retrospectively correlate local control with the treated D99 for GTVs in

Section 5.1. For a prospective clinical trial, the margin reduction is only for treatments using the Synchrony system. One area of future work would be determining the best way include new treatment course data for the trial. For example, should the tuning parameters or SVM be retrained each time a new patient is treated? We might also expand the dataset using treatments for other anatomical sites (such as pancreas and kidney) which also use the Synchrony tracking system or apply our model to data from another centre.

In the discussion in Section 4.4, we noted that the magnitude of the systematic shifts observed in the perturbed dose was slightly underestimated relative to the measured data. One limitation of using the correlation errors to estimate delivery errors is that they are associated with x-ray acquisitions which occur between the delivery of beams. These data may not fully capture the distribution of error while the beam is on. This might be especially true for highly out-of-phase internal/external motion, where the internal motion may change rapidly, while the external motion is in a more slowly varying part of its time trace. A six-dimensional dynamic motion system that employs a robot programmed to translate and rotate a phantom during beam delivery is available on the market.<sup>245</sup> Three-dimensional phantom motion and simultaneous target rotations can be programmed for more complicated phantom tests to validate our dose perturbation tool. An accurate positioning system like that could also be used to determine the location of the phantom during the beam-on time and, as an independent system, to verify how well the correlation errors represent the actual motion tracking errors during beam delivery.

A simple model of deformation involving uniform contraction or shrinking of the target was employed in Section 5.1. The validity of this approximation was not explored in this thesis, since the phantoms used in this study (Chapter 4) were rigid. One could conceive of a deformable phantom which could empirically validate this deformation model, but it would not answer the

question: is this a good deformation model for liver patients? To answer that question one could look at the 4DCTs and MR images with fiducials used as part of the planning process, and build a three-dimensional model to study the vector-based liver deformation.<sup>246</sup> This requires liver contours for CT images at different breathing phases for many patients for the results to be conclusive. Thus, auto-contouring of the liver using deep learning is desired. However, local deformations may be more indicative of target coverage than changes to the surface of the entire organ derived from contours. Target registration error<sup>247</sup> for points inside the liver close to the target should be evaluated for any deformation model since the liver is a large and highly deformable organ. The points used to evaluate the deformation should be independent of the fiducials, since that is the data available in log files and would therefore be the input of the deformation model. Some candidates include vessel bifurcations if contrast CT is available or surgical clips. When the target deformation is described using a more sophisticated and validated model the perturbed D99 of a target including deformation may be more accurate.

The dose perturbation tool can also be applied to other sites, tracking modalities, and even different radiation delivery systems. Under a research ethic board approved retrospective study at our institution we can investigate treatment outcomes, tumour control probability, toxicities, technical details of the delivery and plan for any patient treated on CyberKnife between 2009 and 2018. During that time 123 treatment courses were for trigeminal neuralgia. The feasibility study for trigeminal neuralgia described in Section 5.2 was for a relatively small cohort of 24 patients limited only by the availability of “full nerve” contours provided by the radiation oncologist. Extending this study to the full cohort of 123 patients could definitively answer the questions posed in that section. A simple planning study comparing plans generated using the trigeminal path set with those using full head path (where rotations can be corrected) would also be useful for this site.

In that study, the OAR doses and planning objectives should be kept constant for each patient and plan quality could be compared using the metrics described in Section 5.2. The impact of delivery errors could be assessed for both plans for each patient to evaluate the trade-off between plan quality and the increased uncertainty associated with not being able to apply rotations for the trigeminal path set.

The uncertainty in the experiment using the trigeminal path described in 4.3.1 could also be reduced. In that experiment, only two films were used when a maximum of four films could fit in the film slot. This left a small air gap above the films possibly leading to electron disequilibrium at the film plane contributing to measurement uncertainty. This experiment should be repeated without the air gap. For TN treatments where the dose gradient is steep and a small angle of uncorrected target rotation can lead to large dose reduction to a target, another set of experiments can be conducted to further validate the dose perturbation tool. The same phantom can be positioned with different magnitudes of rotational offsets (e.g. between  $0.5^\circ$  and  $1.5^\circ$  for roll, pitch, or yaw in steps of  $0.5^\circ$ ), and four EBT3 films will be used. The estimation of the length of the nerve could be empirically validated based on this measured data using a fictional 2D nerve contour in the film plane.

## Appendix A. Different Coordinate Systems used with Cyberknife

Different coordinate systems used on CyberKnife is shown in Table A-1. Note that the coordinate system (not listed in this table) for LED markers is arbitrary, since it depends on how the camera array is positioned. In this study, the principal component analysis is used for the marker trajectories that are recorded in the system log files (see Section 3.2.5).

Given the patient positioning offset  $(x_0^{AP}, y_0^{LR}, z_0^{IS})$  in the DICOM coordinate, the dose resolution  $(d_x^{AP}, d_y^{LR}, d_z^{IS})$ , the imaging isocenter  $(x_r^{LR}, y_r^{AP}, z_r^{IS})$  of the CT coordinate, and the fiducial position  $(x_f^{IS}, y_f^{LR}, z_f^{AP})$ , the voxel index  $(i^{AP}, j^{LR}, k^{IS})$  for a fiducial in the DICOM coordinate\* is expressed as follows,

$$i^{AP} = (-z_f^{AP} + y_r^{AP} - x_0^{AP}) / d_x^{AP},$$

$$j^{LR} = (y_f^{LR} + x_r^{LR} - y_0^{LR}) / d_y^{LR},$$

$$k^{IS} = (-x_f^{IS} + z_r^{IS} - z_0^{IS}) / d_z^{IS}.$$

Table A-1. The comparison of different coordinate systems used with CyberKnife.

	<b>x (+)</b>	<b>y (+)</b>	<b>z (+)</b>	<b>handedness</b>
<b>Patient coordinate (fiducial coordinate)</b>	inferior	patient left	anterior	right-handed
<b>DICOM coordinate</b>	posterior	patient left	inferior	right-handed
<b>DICOM coordinate*</b>	posterior	patient left	superior	left-handed
<b>CT Frame coordinate</b>	patient left	posterior	superior	right-handed

\* The DICOM coordinate\* is used in MATLAB scripts for the early warning system, for which the GTV voxel occupancy probability is calculated, and the dose perturbation model that is introduced in Section 4.2.4.

## Appendix B. Translation and Rotation Matrix

The translation and rotation matrices are defined as follows.

$$R_{yaw}(\alpha) = R_{AP}(\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) & 0 \\ 0 & \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$R_{pitch}(\beta) = R_{LR}(\beta) = \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$R_{roll}(\gamma) = R_{IS}(\gamma) = \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) & 0 & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_{APLRIS}(d_{AP}, d_{LR}, d_{IS}) = \begin{bmatrix} 1 & 0 & 0 & -d_{AP} \\ 0 & 1 & 0 & -d_{LR} \\ 0 & 0 & 1 & -d_{SI} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_{res}(e_{AP}, e_{LR}, e_{IS}) = T_{APLRIS}(e_{AP}, e_{LR}, e_{IS})$$

$$T_C(r_{AP}, r_{LR}, r_{IS}) = T_{APLRIS}(r_{AP}, r_{LR}, r_{IS})$$

$$T_{UC}(r_{AP}, r_{LR}, r_{IS}) = T_{APLRIS}(-r_{AP}, -r_{LR}, -r_{IS})$$

Where

$$X_i = T_{APLRIS}(d_{AP}, d_{LR}, d_{IS})T_{UC}R_{AP}(\alpha_{yaw})R_{LR}(\beta_{pitch})R_{IS}(\gamma_{roll})T_C(r_{AP}, r_{LR}, r_{IS})X_f \quad \text{Eq. (1)}$$

$X_f$ : the “live” target position in a DICOM image.

$X_i$  : the target position at planning.

$R_{AP}R_{LR}R_{IS}$ : a transformation matrix that applies rotation in the order of roll, pitch, and yaw.

$T_C$  : a transformation matrix that translates the center of rotation to the imaging isocenter

$(r_{AP}, r_{LR}, r_{IS})$

$T_{UC}$ : an inverse matrix of  $T_C$ , i.e.,  $T_C^{-1}(r_{AP}, r_{LR}, r_{IS}) = T_{UC}(r_{AP}, r_{LR}, r_{IS})$

Equation (1) can be rearranged to calculate  $X_f$  as follows,

$$X_f = T_C^{-1}(r_{AP}, r_{LR}, r_{IS})R_{IS}^{-1}(\gamma)R_{LR}^{-1}(\beta)R_{AP}^{-1}(\alpha)T_{UC}^{-1}(r_{AP}, r_{LR}, r_{IS})T_{APLRIS}^{-1}(d_{AP}, d_{LR}, d_{IS})X_i. \quad \text{Eq. (2)}$$

Note that  $(\theta_{yaw}, \theta_{pitch}, \theta_{roll})$  and  $(d_{AP}, d_{LR}, d_{IS})$  are corrections in the patient coordinate by the system. If there is a residual error  $(e_{AP}, e_{LR}, e_{IS})$  to be applied to the *target* in the DICOM coordinate\* (defined in Table A-1),

$$X_f = T_{UC}R_{IS}^{-1}(\theta_{roll})R_{LR}^{-1}(-\theta_{pitch})R_{AP}^{-1}(\theta_{yaw})T_C T_{res}(e_{AP}, -e_{LR}, e_{IS}) \cdot X_i. \quad \text{Eq. (3)}$$

Since  $(AB)^T = B^T A^T$  for matrix transformation, the equation above is equivalent to Eq. (4) shown below, which is the translation and rotation matrix applied to the binary mask of a target or the planned dose distribution in this thesis work.

$$X_f^T = X_i^T \cdot T_{res}^T(e_{AP}, -e_{LR}, e_{IS})T_C^T R_{AP}^T(-\theta_{yaw})R_{LR}^T(\theta_{pitch})R_{IS}^T(-\theta_{roll})T_{UC}^T. \quad \text{Eq. (4)}$$

## **Appendix C. GTV Voxel Occupancy Probability**

In this thesis, the GTV voxel occupancy probability defines the percentage of the gross tumour volume (GTV) enclosed by the planning target volume (PTV). For each x-ray acquisition, we perturb the binary mask of a GTV using a translational residual error and, if applicable, an uncorrected target rotation. For the early warning system proposed in Section 3.2.4, only translational errors are applied to perturb the GTV.

Note that three types of errors for treatments using the Synchrony tracking system are considered as independent sources of translational tracking residual errors. They are correlation, predictor, and end-to-end errors. The correlation errors are determined for every x-ray acquisition, which for liver treatments is acquired every 30 to 60 seconds. The predictor errors (sampled at 26 Hz) for only the beam-on intervals are considered. The end-to-end errors (SD = 0.3, 0.3 and 0.2 mm in the SUP-INF, LT-RT, and ANT-POST directions, respectively) are obtained from the CyberKnife annual quality assurance measurements (2015-2019).

After a number of perturbations applied to the GTV, the GTV travels to a volume of interest (VOI) that is larger than the size of GTV with different occupancy probabilities for voxels of VOI. The VOI can be larger than the size of PTV if the delivery errors are large. We denote  $p_j$  as the occupancy probability for voxel  $j$  of VOI. Using this definition,  $p_j$  is zero when the translated mask of GTV never “overlaps” with voxel  $j$ , while  $p_j$  is 100% when voxel  $j$  always overlaps with the perturbed mask of GTV for every x-ray acquisition. To obtain a treatment with high dose coverage for the GTV,  $p_j$  should be close to 100% when voxel  $j$  is near the centre of the PTV, but  $p_j$  can get lower when the voxel is near the edge of the PTV.

## Appendix C. GTV Voxel Occupancy Probability

---

For each x-ray acquisition, the perturbation is given the same weight. For treatment fraction  $i$ , the GTV voxel occupancy probability ( $P_i$ ) can be calculated as  $\frac{\sum^{within\ PTV} p_j}{\sum^{all} p_j}$ . Therefore,  $P_i$  measures the percentage of the GTV enclosed by the PTV. For example,  $P_i$  is 100% if for every perturbation the perturbed GTV is always enclosed by the PTV (indicating a sufficient PTV margin).

In Chapter 3, for each x-ray image acquisition, we shift the GTV voxels by a translational residual error. This error is calculated as the sum of a correlation error (determined at the x-ray image acquisition), a predictor error (randomly selected from the beam-on predictor errors for the same fraction), and an end-to-end error (randomly chosen from the annual QA measurements). Alternatively, this is achieved using the spatial convolution of the binary mask of the GTV with the probability distributions functions (PDFs) of correlation, predictor, and end-to-end errors, which has been used in two experimental studies discussed in Section 5.1.3 (see Figure 5-6 and Figure 5-7).

In addition to translational errors, a target rotation (if available) can be applied to perturb the GTV using the rotation matrix described in Appendix B. When the rigid body errors are also taken into consideration, we use a simplified method to calculate the GTV voxel occupancy probability matrix consisting of  $p_j$  as follows: firstly, the signed rigid body errors are binned using the smallest dose matrix resolution in one dimension; secondly, the GTV contours are programmed to either expand or contract isotropically, based on a binned rigid body error. In this way, binary masks of different sizes are precalculated; lastly, we determine the voxel occupancy matrices for every deformed mask by perturbing it using translational and/or rotational errors. For each binned rigid body error, the GTV voxel occupancy probability matrix is weighed using the probability of the bin.

There is one limitation in the nearest neighbor interpolation used in this study, which determines the translated voxel index. The discretization effect can be apparent when the rotational offset is large with a narrow distribution of rotation (i.e., a small standard deviation of angles of rotation), since a rotated voxel location is discretized and rounded to the nearest neighbour. For patient #19 as labelled in Figure 5-1, the distribution of angles of rotation has a large offset with a relatively small standard deviation. The average rotation over the treatment course was as large as  $-11.4^\circ \pm 1.8^\circ$ ,  $-11.1^\circ \pm 1.5^\circ$  and  $-3.0^\circ \pm 2.4^\circ$  in roll, pitch, and yaw, respectively. As shown in Figure 5-6, isolated pixels with low voxel occupancy probability ( $p_j$ ) at the centre of the PTV reflect the voxel discretization effects due to the nearest neighbour interpolation. In comparison, the effect is *not* noticeable for patient #1 (as labelled in Figure 5-1) because no isolated pixels with a  $p_j$  can be found in the centre of the PTV in Figure 5-7. We hypothesize that this is because the rotational offset is small ( $0.2^\circ \pm 0.9^\circ$ ,  $1.7^\circ \pm 1.8^\circ$  and  $0.1^\circ \pm 1.1^\circ$  in roll, pitch, and yaw, respectively) for this patient.

Similarly, in Chapters 4 and 5, the binary mask of GTV is perturbed using different types of delivery errors to determine the dose to a target. As discussed in Section 4.4, we hypothesize that the subtle artifact in the 2D gamma plot in Figure 4-13b is caused by the rotated voxel locations being discretized and rounded to the nearest neighbouring voxel in the perturbed dose. In our study, we do not use the trilinear interpolation to avoid the averaging effect for calculating  $p_j$ , particularly for treatments in which an accurate estimate of the maximum  $p_j$  for serial structures is important. For example, we found that the maximum dose to the brainstem was underestimated for trigeminal neuralgia treatments if the trilinear interpolation was used in determining the shifted or rotated voxel index.

## Appendix D. Phase Shift Between Internal and External Motions

Assuming the internal and external motion functions of the center of tracking fiducials and the average marker position are  $f_{int}$  and  $f_{out}$ , respectively, the simplified wave functions can be written as,

$$f_{int} = A_1 \sin(\omega t) + B_1 \quad \text{Eq. (1)}$$

$$f_{ext} = A_2 \sin(\omega t + \varphi) + B_2, \quad \text{Eq. (2)}$$

where  $A_1, A_2, B_1, B_2$  are constants. This assumes that these two motion functions do not have a time dependent drifting component (i.e.,  $B_1$  and  $B_2$  are constant). We can rewrite  $f_{ext} = A_2 \sin(\omega t) \cos(\varphi) + A_2 \cos(\omega t) \sin(\varphi) + B_2$ , Eq. (3). Since  $\sin^2(\omega t) + \cos^2(\omega t) = 1$ , by factoring out the  $\omega t$  term using equations 1, 2 and 3, the relationship between  $f_{int}$  and  $f_{ext}$  are as follows,

$$\frac{1}{A_1^2 \sin^2(\varphi)} (f_{int} - B_1)^2 - \frac{2 \cos(\varphi)}{A_1 A_2 \sin^2 \varphi} (f_{int} - B_1)(f_{out} - B_2) + \frac{1}{A_2^2 \sin^2(\varphi)} (f_{out} - B_2)^2 = 1. \quad \text{Eq. (4)}$$

In our study, we can fit the data points of the internal and external locations at the acquisitions of x-ray images using an elliptic function. The general form of an elliptic function that is used for curve fitting is written as  $a f_{int}^2 + b f_{int} f_{out} + c f_{out}^2 + d f_{int} + e f_{out} = 1$ , Eq. (5), where  $a, b, c, d$ , and  $e$  are coefficients.<sup>248</sup> By comparing Eq. (4) and (5), we can rewrite  $a, b$  and  $c$  using  $A_1, A_2$ , and  $\varphi$ , shown below.

$$a = \frac{1}{A_1^2 \sin^2(\varphi)}, b = -\frac{2 \cos(\varphi)}{A_1 A_2 \sin^2(\varphi)}, c = \frac{1}{A_2^2 \sin^2(\varphi)}. \quad \text{Eq. (6)}$$

The phase shift  $\varphi$ , in degree, can be calculated using coefficients  $a, b$  and  $c$  of the curve fit.

$$\varphi = \text{acos} \left( \sqrt{\frac{b^2}{4ac}} \right) \cdot \frac{180^\circ}{\pi}. \quad \text{Eq. (7)}$$

## Appendix E. X-ray Imaging Dose (Liver Insert)

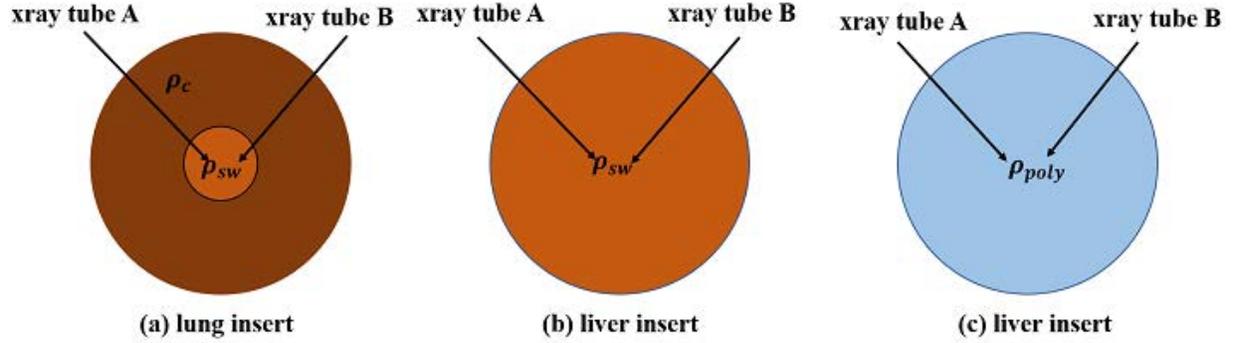


Figure E-1. A diagram to illustrate the axial plane of the phantom inserts, with two orthogonal x-ray beams intersect at the centre of the plane. Plane (a) represents the transverse plane of the lung insert. Planes (b) and (c) represent the axial planes at two different locations inside the liver insert. The subscripts “c”, “sw”, and “poly” represent the cedar, Solid Water, and Styrofoam, respectively.

For 130 kVp x-rays images, the average energy of the photons is about 1/3 of the peak photon energy, which is about 40 keV. The mass attenuation coefficient ( $\mu/\rho$ ) and density ( $\rho$ ) for phantom materials (<https://www.nist.gov/pml/x-ray-mass-attenuation-coefficients>) are listed in Table E-1. The mass attenuation coefficient of the Styrofoam is estimated below because the Styrofoam is made of approximately 95% of air and 5% of polystyrene.

$$\frac{\mu^{40\text{keV}}}{\rho_{\text{foam}}} = 5\% \cdot \frac{\mu^{40\text{keV}}}{\rho_{\text{poly}}} + 95\% \cdot \frac{\mu^{40\text{keV}}}{\rho_{\text{air,dry}}} = 0.05 \times 0.218 + 0.95 \times 0.249 = 0.247 \text{ cm}^2/\text{g}.$$

The diagram in Figure E-1 illustrates the axial plane of the two phantom inserts. In our experiments, we used the same x-ray settings for experiments employing the lung insert and liver insert. If the surface dose of two x-ray images is known as  $D_A$  and  $D_B$ , the dose at the centre of the circular plane for these three configurations of phantom materials can be calculated using equations 1-3 as follows.

$$D_{\text{lung,centre}} = (D_A + D_B) e^{-\left[\left(\frac{\mu_c}{\rho_c}\right)\rho_c x_c + \left(\frac{\mu_{\text{sw}}}{\rho_{\text{sw}}}\right)\rho_{\text{sw}} x_{\text{sw}}\right]} \quad \text{Eq. (1)}$$

$$D_{\text{sw,centre}} = (D_A + D_B) e^{-\left(\frac{\mu_{\text{sw}}}{\rho_{\text{sw}}}\right)\rho_{\text{sw}} x_{\text{sw}}}, \quad \text{Eq. (2)}$$

*Appendix E. X-ray Imaging Dose (Liver Insert)*

---

$$D_{\text{poly,centre}} = (D_A + D_B)e^{-\left(\frac{\mu_{\text{poly}}}{\rho_{\text{poly}}}\right)\rho_{\text{poly}}x_{\text{poly}}}, \quad \text{Eq. (3)}$$

where  $x$  represents the path length (radius of the cylinder). Since the dose at the centre of the lung insert ( $D_{\text{lung,sw}}$ ) was measured as 0.047 cGy per x-ray image pair,<sup>169</sup> we can factor out the  $(D_A+D_B)$  term in Equations (2) and (3) using Equation (1), and calculate the dose at the centre of the liver insert. Therefore, the estimated  $D_{\text{sw, centre}}$  is 0.028 cGy and  $D_{\text{poly, centre}}$  is 0.078 cGy on average for each x-ray image pair. On average, 135 x-ray images were taken for an individual experiment using the liver insert.

If we consider the imaging dose of 0.078 cGy and 0.028 cGy per x-ray image pair for the Styrofoam and Solid Water, respectively, the total imaging dose is about 1.1% and 0.4% of the maximum planned dose (10 Gy). The mean  $\pm$  SD of the imaging dose (measured using the lung insert) averaged over all film pixels was  $0.053 \pm 0.012$  cGy. Since the dose is not a constant across the axial plane and is the lowest at the centre, we subtract the 0.078 cGy and 0.028 cGy times the number of acquired x-ray images from the measured dose for the Styrofoam and Solid Water, respectively, to simplify the calculation.

Table E-1. Mass attenuation coefficients, density, and path length for phantom materials

<b>materials</b>	<b><math>\mu/\rho</math> (cm<sup>2</sup>/g)</b>	<b><math>\rho</math> (g/cm<sup>3</sup>)</b>	<b>path length (cm)</b>
Cedar (lung insert)	0.236	0.26	2.5
Solid Water (lung insert)	0.268	1.00	1.5
Solid Water (liver insert)	0.268	1.00	4.0
Styrofoam (liver insert)	0.247	0.05	4.0

## Appendix F. Screenshots of the Precision Treatment Planning Interface

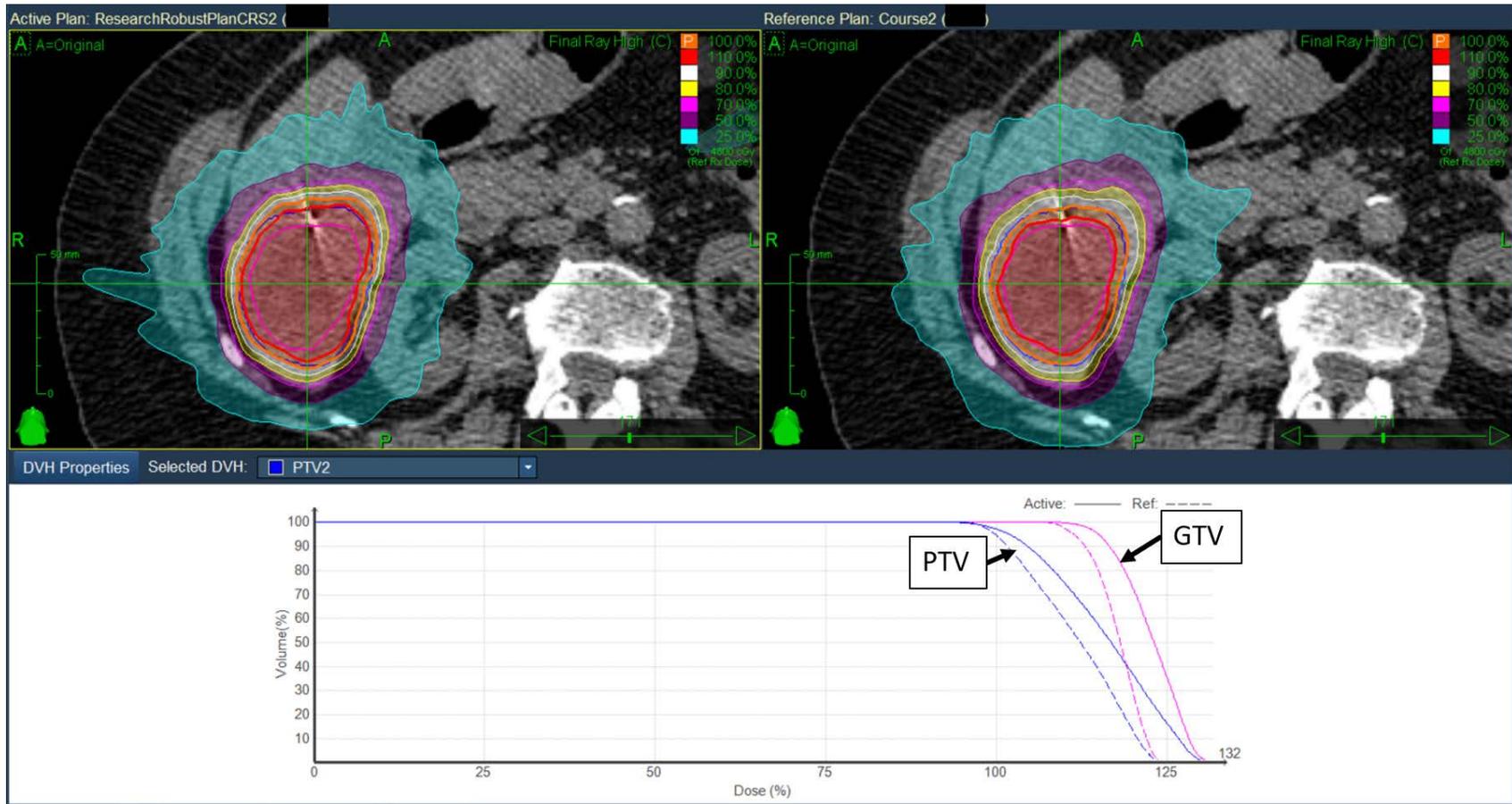


Figure F-1. A screenshot of the Precision treatment planning interface that compares the dose distributions (isodose lines in colour wash mode) for the replanned (upper left) and original (upper right) cases for patient #19 as labelled in Table 5-2. The bottom figure shows the dose-volume histogram for the PTV and GTV. For both plans, four fiducials can be used for tumour tracking. The green crosshair indicates the imaging isocentre is placed close to the target centroid for the replanned case. From the original to the replanned case, D99/Rx for the target increases from 1.08 to 1.11 by 3%.

## Appendix F. Screenshots of the Precision Treatment Planning Interface

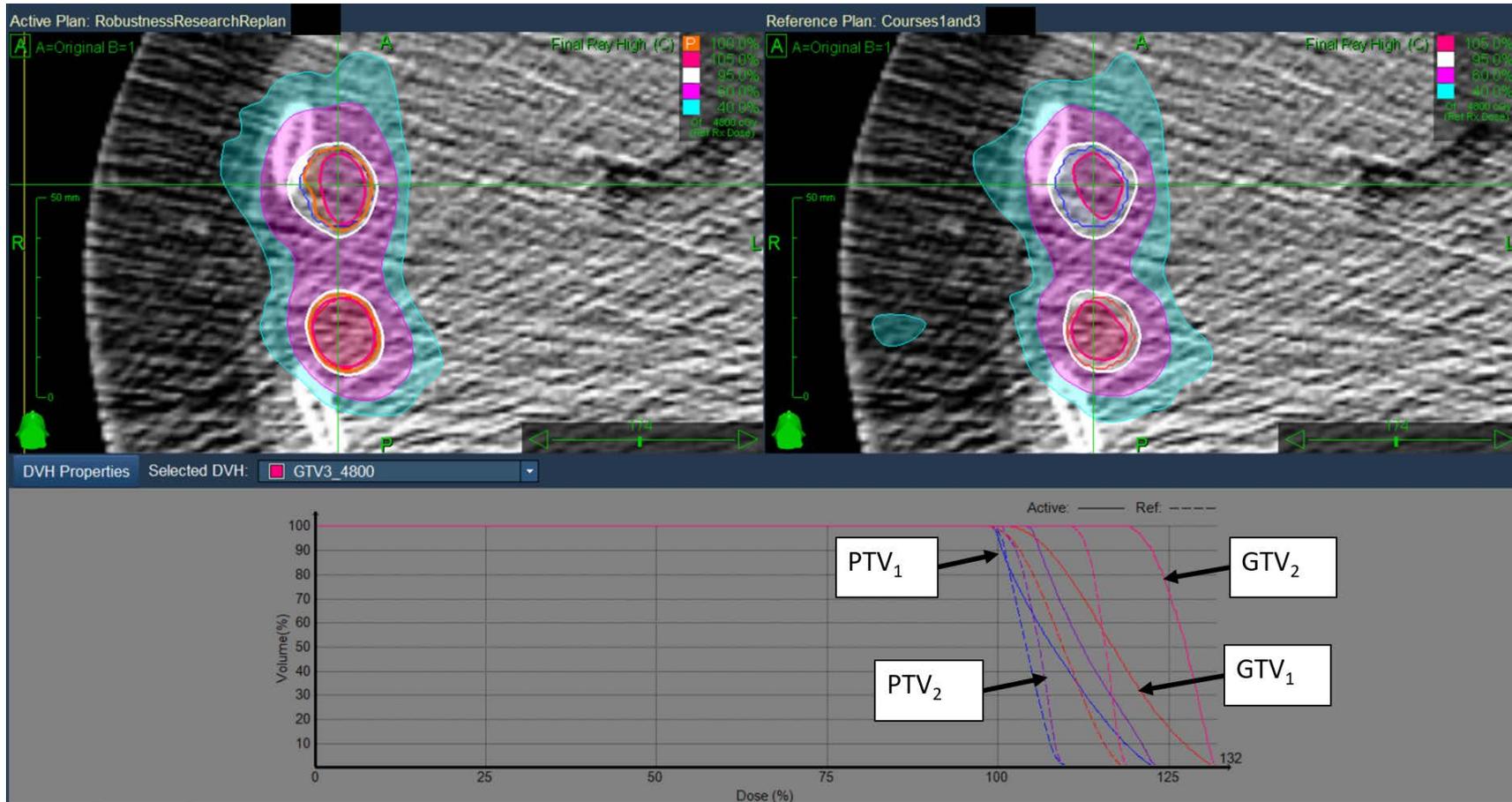


Figure F-2. A screenshot of the Precision treatment planning interface that compares the dose distributions (isodose lines in colour wash mode) for the replanned (upper left) and original (upper right) cases for patient #65 as labelled in Table 5-2. The bottom figure shows the dose-volume histogram for the PTVs and GTVs of two targets. Five and seven fiducials can be used for tumour tracking for the replanned and original cases, respectively. The green crosshair indicates the location of the imaging isocentre, which is placed close to target # 1 (as labelled in Table 5-2) for the replanned case. From the original to the replanned case, D<sub>99</sub>/Rx for target #1 increases from 1.01 to 1.05 by 4%, and for target #2 increases from 1.11 to 1.19 by 8%.

## References

1. Brenner DR, Weir HK, Demers AA, et al. Projected estimates of cancer in Canada in 2020. *Cmaj*. 2020;192(9):E199-E205.
2. Cancer Facts & Figures 2020. *American Cancer Society: Atlanta, GA, USA*. 2020.
3. Bruix J, Sherman M. Management of hepatocellular carcinoma: an update. *Hepatology*. 2011;53(3):1020-1022.
4. Griscom JT, Wolf PS. Cancer, Liver Metastasis. In: *StatPearls [Internet]*. StatPearls Publishing; 2020.
5. Dawood O, Mahadevan A, Goodman KA. Stereotactic body radiation therapy for liver metastases. *European Journal of Cancer*. 2009;45(17):2947-2959.
6. Mukherji A. *Basics of Planning and Management of Patients during Radiation Therapy*. Springer; 2018.
7. Hall EJ, Giaccia AJ. *Radiobiology for the radiologist*. 6th ed: Lippincott Williams & Wilkins; 2016.
8. ICRU. Report 85: Fundamental Quantities and Units for Ionizing Radiation. *ICRU Report 85, Journal of the ICRU*. 2011;11:1-33.
9. Timmerman RD. An overview of hypofractionation and introduction to this issue of seminars in radiation oncology. *Seminars in radiation oncology*. 2008;18:215-222.
10. Steel GG, McMillan TJ, Peacock J. The 5Rs of radiobiology. *Int. J. Radiat. Biol*. 1989;56(6):1045-1048.
11. Podgoršak EB. *Radiation Physics for Medical Physicists*. 2 ed. Heidelberg: Springer; 2010.
12. Cunningham JR, Johns HE. *The Physics of Radiology*. Springfield: Charles C. Thomas. 1983.
13. Sixel KE, Faddegon BA. Calculation of x-ray spectra for radiosurgical beams. *Med. Phys*. 1995;22(10):1657-1661.
14. Attix FH. *Introduction to radiological physics and radiation dosimetry*. John Wiley & Sons; 1986.
15. Attix F. Energy imparted, energy transferred and net energy transferred. *Phys. Med. Biol*. 1983;28(12):1385.
16. Das IJ, Ding GX, Ahnesjö A. Small fields: nonequilibrium radiation dosimetry. *Med. Phys*. 2008;35(1):206-215.
17. Bjärngård B, Tsai JS, Rice R. Doses on the central axes of narrow 6-MV x-ray beams. *Medical physics*. 1990;17(5):794-799.
18. Palmans H, Andreo P, Huq MS, Seuntjens J, Christaki KE, Meghzifene A. Dosimetry of small static fields used in external photon beam radiotherapy: Summary of TRS-483, the IAEA–AAPM international Code of Practice for reference and relative dose determination. *Med. Phys*. 2018;45(11):e1123-e1145.
19. Wilcox EE, Daskalov GM. Evaluation of GAFCHROMIC® EBT film for CyberKnife® dosimetry. *Medical physics*. 2007;34(6Part1):1967-1974.
20. Niroomand - Rad A, Blackwell CR, Coursey BM, et al. Radiochromic film dosimetry: recommendations of AAPM radiation therapy committee task group 55. *Medical physics*. 1998;25(11):2093-2115.
21. ICRU. Report 62: Prescribing, Recording and Reporting Photon Beam Therapy (Supplement to ICRU Report 50). *ICRU Report 62, Journal of the ICRU*. 1999.11-12.

## References

---

22. Ling CC, Humm J, Larson S, et al. Towards multidimensional radiotherapy (MD-CRT): biological imaging and biological conformality. *Int. J. Radiat. Oncol. Biol. Phys.* 2000;47(3):551-560.
23. Bassi MC, Turri L, Sacchetti G, et al. FDG-PET/CT imaging for staging and target volume delineation in preoperative conformal radiotherapy of rectal cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 2008;70(5):1423-1426.
24. Vanuytsel LJ, Vansteenkiste JF, Stroobants SG, et al. The impact of 18F-fluoro-2-deoxy-D-glucose positron emission tomography (FDG-PET) lymph node staging on the radiation treatment volumes in patients with non-small cell lung cancer. *Radiotherapy and Oncology.* 2000;55(3):317-324.
25. Keall P. 4-dimensional computed tomography imaging and treatment planning. *Seminars in radiation oncology.* 2004;14(1):81-90.
26. ICRU. Report 83: Prescribing, recording and reporting photon-beam intensity-modulated radiation therapy (IMRT). *ICRU Report 83, Journal of the ICRU.* 2012;10(1):97-99.
27. Keall PJ, Mageras GS, Balter JM, et al. The management of respiratory motion in radiation oncology report of AAPM Task Group 76. *Med. Phys.* 2006;33(10):3874-3900.
28. Wendling M, Zijp LJ, McDermott LN, et al. A fast algorithm for gamma evaluation in 3D. *Medical physics.* 2007;34(5):1647-1654.
29. Dieterich S, Cavedon C, Chuang CF, et al. Report of AAPM TG 135: quality assurance for robotic radiosurgery. *Medical physics.* 2011;38(6):2914-2936.
30. Vandervoort E, Patrocinio H, Chow T, Soisson E, Nadeau DB. COMP Report: CPQR technical quality control guidelines for CyberKnife® Technology. *Journal of applied clinical medical physics.* 2018;19(2):29-34.
31. Webb S. Motion effects in (intensity modulated) radiation therapy: a review. *Phys. Med. Biol.* 2006;51(13):403-425.
32. Ozhasoglu C, Murphy MJ. Issues in respiratory motion compensation during external-beam radiotherapy. *Int. J. Radiat. Oncol., Biol., Phys.* 2002;52(5):1389-1399.
33. Liang Z, Liu H, Xue J, et al. Evaluation of the intra- and interfractional tumor motion and variability by fiducial-based real-time tracking in liver stereotactic body radiation therapy. *Journal of applied clinical medical physics.* 2018;19(3):94-100.
34. Case RB, Moseley DJ, Sonke JJ, et al. Interfraction and intrafraction changes in amplitude of breathing motion in stereotactic liver radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 2010;77(3):918-925.
35. Case RB, Sonke J-J, Moseley DJ, Kim J, Brock KK, Dawson LA. Inter- and intrafraction variability in liver position in non-breath-hold stereotactic body radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 2009;75(1):302-308.
36. Park JC, Park SH, Kim JH, et al. Liver motion during cone beam computed tomography guided stereotactic body radiation therapy. *Medical physics.* 2012;39(10):6431-6442.
37. Worm ES, Høyer M, Fledelius W, Hansen AT, Poulsen PR. Variations in magnitude and directionality of respiratory target motion throughout full treatment courses of stereotactic body radiotherapy for tumors in the liver. *Acta Oncologica.* 2013;52(7):1437-1444.
38. Eccles CL, Patel R, Simeonov AK, Lockwood G, Haider M, Dawson LA. Comparison of liver tumor motion with and without abdominal compression using cine-magnetic resonance imaging. *Int. J. Radiat. Oncol. Biol. Phys.* 2011;79(2):602-608.
39. Shimizu S, Shirato H, Xo B, et al. Three-dimensional movement of a liver tumor detected by high-speed magnetic resonance imaging. *Radiotherapy and oncology.* 1999;50(3):367-370.

## References

---

40. Brandner ED, Chetty IJ, Giaddui TG, Xiao Y, Huq MS. Motion management strategies and technical issues associated with stereotactic body radiotherapy of thoracic and upper abdominal tumors: a review from NRG oncology. *Medical physics*. 2017;44(6):2595-2612.
41. Bertholet J, Knopf A, Eiben B, et al. Real-time intrafraction motion monitoring in external beam radiotherapy. *Physics in Medicine & Biology*. 2019;64(15):15TR01.
42. Shirato H, Shimizu S, Kunieda T, et al. Physical aspects of a real-time tumor-tracking system for gated radiotherapy. *Int. J. Radiat. Oncol., Biol., Phys.* 2000;48(4):1187-1195.
43. Kitamura K, Shirato H, Shimizu S, et al. Registration accuracy and possible migration of internal fiducial gold marker implanted in prostate and liver treated with real-time tumor-tracking radiation therapy (RTRT). *Radiotherapy and oncology*. 2002;62(3):275-281.
44. Dawson LA, Eccles C, Bissonnette J-P, Brock KK. Accuracy of daily image guidance for hypofractionated liver radiotherapy with active breathing control. *Int. J. Radiat. Oncol. Biol. Phys.* 2005;62(4):1247-1252.
45. Keall P, Todor A, Vedam S, et al. On the use of EPID-based implanted marker tracking for 4D radiotherapy: EPID-based marker tracking. *Medical physics*. 2004;31(12):3492-3499.
46. Buhl SK, Duun-Christensen AK, Kristensen BH, Behrens CF. Clinical evaluation of 3D/3D MRI-CBCT automatching on brain tumors for online patient setup verification—A step towards MRI-based treatment planning. *Acta Oncol.* 2010;49(7):1085-1091.
47. Crijns S, Raaymakers B, Lagendijk J. Proof of concept of MRI-guided tracked radiation delivery: tracking one-dimensional motion. *Phys. Med. Biol.* 2012;57(23):7863.
48. Fast MF, Nill S, Bedford JL, Oelfke U. Dynamic tumor tracking using the Elekta Agility MLC. *Medical physics*. 2014;41(11):111719.
49. Ozkan E, Tanner C, Kastelic M, Mattausch O, Makhinya M, Goksel O. Robust motion tracking in liver from 2D ultrasound images using supporters. *International journal of computer assisted radiology and surgery*. 2017;12(6):941-950.
50. De Luca V, Banerjee J, Hallack A, et al. Evaluation of 2D and 3D ultrasound tracking algorithms and impact on ultrasound-guided liver radiotherapy margins. *Medical Physics*. 2018;45(11):4986-5003.
51. Williamson T, Cheung W, Roberts SK, Chauhan S. Ultrasound-based liver tracking utilizing a hybrid template/optical flow approach. *International journal of computer assisted radiology and surgery*. 2018;13(10):1605-1615.
52. De Luca V, Székely G, Tanner C. Estimation of large-scale organ motion in B-mode ultrasound image sequences: A survey. *Ultrasound Med. Biol.* 2015;41(12):3044-3062.
53. Li S, Liu D, Yin G, Zhuang P, Geng J. Real-time 3D-surface-guided head re-fixation useful for fractionated stereotactic radiotherapy. *Med. Phys.* 2006;33(2):492-503.
54. Bert C, Metheany KG, Doppke K, Chen GT. A phantom evaluation of a stereo-vision surface imaging system for radiotherapy patient setup. *Med. Phys.* 2005;32(9):2753-2762.
55. Kupelian P, Willoughby T, Mahadevan A, et al. Multi-institutional clinical experience with the Calypso System in localization and continuous, real-time monitoring of the prostate gland during external radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 2007;67(4):1088-1098.
56. Cherpak A, Ding W, Hallil A, Cygler J. Evaluation of a novel 4D in vivo dosimetry system. *Medical physics*. 2009;36(5):1672-1679.
57. Shirato H, Seppenwoolde Y, Kitamura K, Onimura R, Shimizu S. Intrafractional tumor motion: lung and liver. *Seminars in radiation oncology*. 2004;14:10-18

## References

---

58. Keall P, Kini V, Vedam S, Mohan R. Potential radiotherapy improvements with respiratory gating. *Australas. Phys. Eng. Sci. Med.* 2002;25(1):1.
59. Tseng C-L, Eppinga W, Charest-Morin R, et al. Spine stereotactic body radiotherapy: indications, outcomes, and points of caution. *Global Spine J.* 2017;7(2):179-197.
60. Remouchamps VM, Letts N, Vicini FA, et al. Initial clinical experience with moderate deep-inspiration breath hold using an active breathing control device in the treatment of patients with left-sided breast cancer using external beam radiation therapy. *Int. J. Radiat. Oncol., Biol., Phys.* 2003;56(3):704-715.
61. Wong JW, Sharpe MB, Jaffray DA, et al. The use of active breathing control (ABC) to reduce margin for breathing motion. *Int. J. Radiat. Oncol., Biol., Phys.* 1999;44(4):911-919.
62. Eccles C, Brock KK, Bissonnette J-P, Hawkins M, Dawson LA. Reproducibility of liver position using active breathing coordinator for liver cancer radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 2006;64(3):751-759.
63. Kubo HD, Hill BC. Respiration gated radiotherapy treatment: a technical study. *Physics in Medicine & Biology.* 1996;41(1):83.
64. Vedam S, Keall P, Kini V, Mohan R. Determining parameters for respiration - gated radiotherapy. *Medical physics.* 2001;28(10):2139-2146.
65. Falk M, Pommer T, Keall P, et al. Motion management during IMAT treatment of mobile lung tumors—a comparison of MLC tracking and gated delivery. *Medical physics.* 2014;41(10):101707.
66. Keall P, Starkschall G, Shukla H, et al. Acquiring 4D thoracic CT scans using a multislice helical method. *Physics in Medicine & Biology.* 2004;49(10):2053.
67. Wagman R, Yorke E, Ford E, et al. Respiratory gating for liver tumors: use in dose escalation. *Int. J. Radiat. Oncol. Biol. Phys.* 2003;55(3):659-668.
68. Seppenwoolde Y, Shirato H, Kitamura K, et al. Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *Int. J. Radiat. Oncol., Biol., Phys.* 2002;53(4):822-834.
69. Shirato H, Harada T, Harabayashi T, et al. Feasibility of insertion/implantation of 2.0-mm-diameter gold internal fiducial markers for precise setup and real-time tumor tracking in radiotherapy. *Int. J. Radiat. Oncol., Biol., Phys.* 2003;56(1):240-247.
70. Sothmann T, Blanck O, Poels K, Werner R, Gauer T. Real time tracking in liver SBRT: comparison of CyberKnife and Vero by planning structure-based  $\gamma$ -evaluation and dose-area-histograms. *Physics in medicine and biology.* 2016;61(4):1677-1691.
71. Ozhasoglu C, Saw CB, Chen H, et al. Synchrony—CyberKnife respiratory compensation technology. *Med. Dosim.* 2008;33(2):117-123.
72. Fu D, Kahn R, Wang B, et al. Xsight Lung Tracking System: A Fiducial-Less Method for Respiratory Motion Tracking. In: Urschel HC, Kresl JJ, Luketich JD, Papiez L, Timmerman RD, Schulz RA, eds. *Treating Tumors that Move with Respiration.* doi: 10.1007/978-3-540-69886-9\_26 Berlin, Heidelberg: Springer Berlin Heidelberg; 2007:265-282.
73. Kilby W, Dooley JR, Kuduvalli G, Sayeh S, Maurer Jr CR. The CyberKnife® robotic radiosurgery system in 2010. *Technol. Cancer Res.* 2010;9(5):433-452.
74. Depuydt T, Verellen D, Haas O, et al. Geometric accuracy of a novel gimbals based radiation therapy tumor tracking system. *Radiotherapy and Oncology.* 2011;98(3):365-372.
75. Depuydt T, Poels K, Verellen D, et al. Initial assessment of tumor tracking with a gimbaled linac system in clinical circumstances: a patient simulation study. *Radiotherapy and Oncology.* 2013;106(2):236-240.

## References

---

76. Winter JD, Wong R, Swaminath A, Chow T. Accuracy of robotic radiosurgical liver treatment throughout the respiratory cycle. *Int. J. Radiat. Oncol., Biol., Phys.* 2015;93(4):916-924.
77. Ziegler M, Brandt T, Lettmaier S, Fietkau R, Bert C. Performance of gimbal-based dynamic tumor tracking for treating liver carcinoma. *Radiation Oncology.* 2018;13(1):1-12.
78. Poels K, Dhont J, Verellen D, et al. A comparison of two clinical correlation models used for real-time tumor tracking of semi-periodic motion: A focus on geometrical accuracy in lung and liver cancer patients. *Radiotherapy and Oncology.* 2015;115(3):419-424.
79. Bertholet J, Toftegaard J, Hansen R, et al. Automatic online and real-time tumour motion monitoring during stereotactic liver treatments on a conventional linac by combined optical and sparse monoscopic imaging with kilovoltage x-rays (COSMIK). *Physics in Medicine & Biology.* 2018;63(5):055012.
80. Bertholet J, Wan H, Toftegaard J, et al. Fully automatic segmentation of arbitrarily shaped fiducial markers in cone-beam CT projections. *Physics in Medicine & Biology.* 2017;62(4):1327.
81. Skouboe S, Ravkilde T, Bertholet J, et al. First clinical real-time motion-including tumor dose reconstruction during radiotherapy delivery. *Radiotherapy and Oncology.* 2019;139:66-71.
82. Ravkilde T, Skouboe S, Hansen R, Worm E, Poulsen PR. First online real-time evaluation of motion - induced 4D dose errors during radiotherapy delivery. *Medical physics.* 2018;45(8):3893-3903.
83. Keall PJ, Cattell H, Pokhrel D, et al. Geometric accuracy of a real-time target tracking system with dynamic multileaf collimator tracking system. *Int. J. Radiat. Oncol. Biol. Phys.* 2006;65(5):1579-1584.
84. Falk M, af Rosenschöld PM, Keall P, et al. Real-time dynamic MLC tracking for inversely optimized arc radiotherapy. *Radiotherapy and oncology.* 2010;94(2):218-223.
85. Cho B, Poulsen PR, Sloutsky A, Sawant A, Keall PJ. First demonstration of combined kV/MV image-guided real-time dynamic multileaf-collimator target tracking. *Int. J. Radiat. Oncol. Biol. Phys.* 2009;74(3):859-867.
86. Keall PJ, Sawant A, Cho B, et al. Electromagnetic-guided dynamic multileaf collimator tracking enables motion management for intensity-modulated arc therapy. *Int. J. Radiat. Oncol. Biol. Phys.* 2011;79(1):312-320.
87. Chen H, Wu A, Brandner ED, et al. Dosimetric evaluations of the interplay effect in respiratory-gated intensity-modulated radiation therapy. *Med. Phys.* 2009;36(3):893-903.
88. Court LE, Wagar M, Ionascu D, Berbeco R, Chin L. Management of the interplay effect when using dynamic MLC sequences to treat moving targets. *Med. Phys.* 2008;35(5):1926-1931.
89. Paganelli C, Seregni M, Fattori G, et al. Magnetic resonance imaging-guided versus surrogate-based motion tracking in liver radiation therapy: A prospective comparative study. *Int. J. Radiat. Oncol., Biol., Phys.* 2015;91(4):840-848.
90. Fast M, van de Schoot A, van de Lindt T, Carbaat C, van der Heide U, Sonke J-J. Tumor trailing for liver SBRT on the MR-linac. *Int. J. Radiat. Oncol. Biol. Phys.* 2019;103(2):468-478.
91. Al-Ward S, Wronski M, Ahmad SB, et al. The radiobiological impact of motion tracking of liver, pancreas and kidney SBRT tumors in a MR-linac. *Physics in Medicine & Biology.* 2018;63(21):215022.
92. Velec M, Moseley JL, Eccles CL, et al. Effect of breathing motion on radiotherapy dose accumulation in the abdomen using deformable registration. *Int. J. Radiat. Oncol., Biol., Phys.* 2011;80(1):265-272.

## References

---

93. Bohoudi O, Bruynzeel A, Senan S, et al. Fast and robust online adaptive planning in stereotactic MR-guided adaptive radiation therapy (SMART) for pancreatic cancer. *Radiother. Oncol.* 2017;125(3):439-444.
94. Acharya S, Fischer-Valuck BW, Kashani R, et al. Online magnetic resonance image guided adaptive radiation therapy: first clinical applications. *Int. J. Radiat. Oncol. Biol. Phys.* 2016;94(2):394-403.
95. Raaymakers B, Jürgenliemk-Schulz I, Bol G, et al. First patients treated with a 1.5 T MRI-Linac: clinical proof of concept of a high-precision, high-field MRI guided radiotherapy treatment. *Physics in Medicine & Biology.* 2017;62(23):L41.
96. Kontaxis C, Bol G, Lagendijk J, Raaymakers B. A new methodology for inter-and intrafraction plan adaptation for the MR-linac. *Phys. Med. Biol.* 2015;60(19):7485.
97. Kontaxis C, Bol G, Stemkens B, et al. Towards fast online intrafraction replanning for free-breathing stereotactic body radiation therapy with the MR-linac. *Phys. Med. Biol.* 2017;62(18):7233.
98. Menten MJ, Mohajer JK, Nilawar R, et al. Automatic reconstruction of the delivered dose of the day using MR-linac treatment log files and online MR imaging. *Radiother. Oncol.* 2020;145:88-94.
99. Kontaxis C, de Muinck Keizer DM, Kerkmeijer LG, et al. Delivered dose quantification in prostate radiotherapy using online 3D cine imaging and treatment log files on a combined 1.5 T magnetic resonance imaging and linear accelerator system. *Phys. Imag. Radiat. Oncol.* 2020;15:23-29.
100. Burman C, Kutcher G, Emami B, Goitein M. Fitting of normal tissue tolerance data to an analytic function. *International Journal of Radiation Oncology\* Biology\* Physics.* 1991;21(1):123-135.
101. Emami B, Lyman J, Brown A, et al. Tolerance of normal tissue to therapeutic irradiation. *International Journal of Radiation Oncology\* Biology\* Physics.* 1991;21(1):109-122.
102. Kutcher G, Burman C, Brewster L, Goitein M, Mohan R. Histogram reduction method for calculating complication probabilities for three-dimensional treatment planning evaluations. *International Journal of Radiation Oncology\* Biology\* Physics.* 1991;21(1):137-146.
103. Lyman JT. Complication probability as assessed from dose-volume histograms. *Radiat. Res.* 1985;104(2s):S13-S19.
104. Leksell L. A stereotaxic apparatus for intracerebral surgery. *Acta Chirurgica Scandinavica.* 1950;99(3):229-233.
105. Leksell L. The stereotaxic method and radiosurgery of the brain. *Acta Chirurg Scand.* 1951;102:316-319.
106. Minniti G, Scaringi C, Clarke E, Valeriani M, Osti M, Enrici RM. Frameless linac-based stereotactic radiosurgery (SRS) for brain metastases: analysis of patient repositioning using a mask fixation system and clinical outcomes. *Radiation oncology.* 2011;6(1):158.
107. Brown JM, Carlson DJ, Brenner DJ. The tumor radiobiology of SRS and SBRT: are more than the 5 Rs involved? *Int. J. Radiat. Oncol. Biol. Phys.* 2014;88(2):254-262.
108. Kuo JS, Yu C, Petrovich Z, Apuzzo MLJ. The CyberKnife stereotactic radiosurgery system: description, installation, and an initial evaluation of use and functionality. *Neurosurgery.* 2003;53(5):1235-1239.
109. Velec M, Moseley JL, Dawson LA, Brock KK. Dose escalated liver stereotactic body radiation therapy at the mean respiratory position. *Int. J. Radiat. Oncol., Biol., Phys.* 2014;89(5):1121-1128.

## References

---

110. Rubio C, Hernando-Requejo O, Aparicio DZ, et al. Image guided SBRT for multiple liver metastases with ExacTrac® Adaptive Gating. *Rep. Pract. Oncol. Radiother.* 2017;22(2):150-157.
111. Fürweger C, Prins P, Coskan H, Heijmen BJ. Characteristics and performance of the first commercial multileaf collimator for a robotic radiosurgery system. *Medical Physics.* 2016;43(5):2063-2071.
112. Dieterich S, Pawlicki T. Cyberknife image-guided delivery and quality assurance. *Int. J. Radiat. Oncol. Biol. Phys.* 2008;71(1):S126-S130.
113. Fu D, Kuduvali G. A fast, accurate, and automatic 2D–3D image registration for image-guided cranial radiosurgery. *Medical physics.* 2008;35(5):2180-2194.
114. Wilcox EE, Daskalov GM, Lincoln H, Shumway RC, Kaplan BM, Colasanto JM. Comparison of planned dose distributions calculated by Monte Carlo and Ray-Trace algorithms for the treatment of lung tumors with cyberknife: a preliminary study in 33 patients. *Int. J. Radiat. Oncol. Biol. Phys.* 2010;77(1):277-284.
115. Ma C, Li J, Deng J, Fan J. Implementation of Monte Carlo dose calculation for CyberKnife treatment planning. Paper presented at: J Phys Conf Ser2008.
116. Sayeh S, Wang J, Main WT, Kilby W, Maurer CR. Respiratory Motion Tracking for Robotic Radiosurgery. In: Urschel HC, Kresl JJ, Luketich JD, Papiez L, Timmerman RD, Schulz RA, eds. *Treating Tumors that Move with Respiration.* Berlin, Heidelberg: Springer Berlin Heidelberg; 2007:15-29.
117. Weichselbaum RR, Hellman S. Oligometastases revisited. *Nat. Rev. Clin. Oncol.* 2011;8(6):378.
118. Tree AC, Khoo VS, Eeles RA, et al. Stereotactic body radiotherapy for oligometastases. *Lancet. Oncol.* 2013;14(1):e28-e37.
119. Timmerman RD, Pass H, Galvin J, et al. *RTOG 0618: A Phase II Trial of Stereotactic Body Radiation Therapy (SBRT) in the Treatment of Patients with Operable Stage I/II Non-Small Cell Lung Cancer.* <https://www.rtog.org/ClinicalTrials/ProtocolTable/StudyDetails.aspx?action=openFile&FileID=4650> (accessed on July 08, 2020).
120. Liang P, Huang C, Liang S-X, et al. Effect of CyberKnife stereotactic body radiation therapy for hepatocellular carcinoma on hepatic toxicity. *OncoTargets and therapy.* 2016;9:7169.
121. Goyal K, Einstein D, Yao M, et al. Cyberknife stereotactic body radiation therapy for nonresectable tumors of the liver: preliminary results. *HPB Surgery.* 2010;2010.
122. Jarraya H, Mirabel X, Taieb S, et al. Image-based response assessment of liver metastases following stereotactic body radiotherapy with respiratory tracking. *Radiation Oncology.* 2013;8(1):24.
123. Pollom EL, Chin AL, Diehn M, Loo BW, Chang DT. Normal tissue constraints for abdominal and thoracic stereotactic body radiotherapy. *Seminars in radiation oncology.* 2017;27(3):197-208.
124. Høyer M, Swaminath A, Bydder S, et al. Radiotherapy for liver metastases: a review of evidence. *Int. J. Radiat. Oncol. Biol. Phys.* 2012;82(3):1047-1057.
125. Lee MT, Kim JJ, Dinniwell R, et al. Phase I study of individualized stereotactic body radiotherapy of liver metastases. *J. Clin. Oncol.* 2009;27(10):1585-1591.
126. Bujold A, Massey CA, Kim JJ, et al. Sequential phase I and II trials of stereotactic body radiotherapy for locally advanced hepatocellular carcinoma. *J. Clin. Oncol.* 2013;31(13):1631-1639.

## References

---

127. Andersen V, Sonne J, Sletting S, Prip A. The volume of the liver in patients correlates to body weight and alcohol consumption. *Alcohol Alcohol*. 2000;35(5):531-532.
128. Schefter TE, Kavanagh BD, Timmerman RD, Cardenas HR, Baron A, Gaspar LE. A phase I trial of stereotactic body radiation therapy (SBRT) for liver metastases. *Int. J. Radiat. Oncol., Biol., Phys.* 2005;62(5).
129. Green OL, Henke LE, Hugo GD. Practical Clinical Workflows for Online and Offline Adaptive Radiation Therapy. *Seminars in radiation oncology*. 2019;29:219-227.
130. Hebb DO. *The organization of behavior: a neuropsychological theory*. J. Wiley; Chapman & Hall; 1949.
131. Dey A. Machine learning algorithms: a review. *International Journal of Computer Science and Information Technologies*. 2016;7(3):1174-1179.
132. El Naqa I, Ruan D, Valdes G, et al. Machine learning and modeling: Data, validation, communication challenges. *Medical physics*. 2018;45(10):e834-e840.
133. Klement RJ, Allgäuer M, Appold S, et al. Support vector machine-based prediction of local tumor control after stereotactic body radiation therapy for early-stage non-small cell lung cancer. *Int. J. Radiat. Oncol., Biol., Phys.* 2014;88(3):732-738.
134. Zhou Z, Folkert M, Cannon N, et al. Predicting distant failure in early stage NSCLC treated with SBRT using clinical parameters. *Radiotherapy and Oncology*. 2016;119(3):501-504.
135. Valdes G, Solberg TD, Heskell M, Ungar L, Simone CB. Using machine learning to predict radiation pneumonitis in patients with stage I non-small cell lung cancer treated with stereotactic body radiation therapy. *Physics in Medicine & Biology*. 2016;61(16):6105-6120.
136. Pella A, Cambria R, Riboldi M, et al. Use of machine learning methods for prediction of acute toxicity in organs at risk following prostate radiotherapy. *Medical physics*. 2011;38(6):2859-2867.
137. Garapati SS, Hadjiiski L, Cha KH, et al. Urinary bladder cancer staging in CT urography using machine learning. *Medical physics*. 2017;44(11):5814-5823.
138. Way TW, Hadjiiski LM, Sahiner B, et al. Computer-aided diagnosis of pulmonary nodules on CT scans: Segmentation and classification using 3D active contours. *Medical physics*. 2006;33(7):2323-2337.
139. Granville DA, Sutherland JG, Belec JG, La Russa DJ. Predicting VMAT patient-specific QA results using a support vector classifier trained on treatment plan characteristics and linac QC metrics. *Physics in Medicine & Biology*. 2019;64(9):095017.
140. Xu Q, Hanna G, Grimm J, et al. Quantifying rigid and nonrigid motion of liver tumors during stereotactic body radiation therapy. *Int. J. Radiat. Oncol., Biol., Phys.* 2014;90(1):94-101.
141. Murphy MJ. Fiducial-based targeting accuracy for external-beam radiotherapy. *Medical physics*. 2002;29(3):334-344.
142. Chan M, Grehn M, Cremers F, et al. Dosimetric Implications of Residual Tracking Errors During Robotic SBRT of Liver Metastases. *Int. J. Radiat. Oncol., Biol., Phys.* 2017;97(4):839-848.
143. Conti A, Acker G, Pontoriero A, et al. Factors affecting outcome in frameless non-isocentric stereotactic radiosurgery for trigeminal neuralgia: a multicentric cohort study. *Radiation Oncology*. 2020;15(1):1-10.
144. Antypas C, Pantelis E. Performance evaluation of a CyberKnife® G4 image-guided robotic stereotactic radiosurgery system. *Physics in Medicine & Biology*. 2008;53(17):4697.

## References

---

145. Floriano A, García R, Moreno R, Sánchez-Reyes A. Retrospective evaluation of CTV to PTV margins using CyberKnife in patients with thoracic tumors. *Journal of applied clinical medical physics*. 2014;15(6):59-72.
146. Malinowski K, McAvoy TJ, George R, Dietrich S, D'Souza WD. Incidence of changes in respiration-induced tumor motion and its relationship with respiratory surrogates during individual treatment fractions. *International Journal of Radiation Oncology\* Biology\* Physics*. 2012;82(5):1665-1673.
147. Pepin EW, Wu H, Zhang Y, Lord B. Correlation and prediction uncertainties in the CyberKnife Synchrony respiratory tracking system. *Medical physics*. 2011;38(7):4036-4044.
148. Hoogeman M, Prévost J-B, Nuyttens J, Pöll J, Levendag P, Heijmen B. Clinical accuracy of the respiratory tumor tracking system of the CyberKnife: assessment by analysis of log files. *International Journal of Radiation Oncology\* Biology\* Physics*. 2009;74(1):297-303.
149. Winter JD, Wong R, Swaminath A, Chow T. Accuracy of robotic radiosurgical liver treatment throughout the respiratory cycle. *International Journal of Radiation Oncology\* Biology\* Physics*. 2015;93(4):916-924.
150. Xu Q, Hanna G, Grimm J, et al. Quantifying rigid and nonrigid motion of liver tumors during stereotactic body radiation therapy. *International Journal of Radiation Oncology\* Biology\* Physics*. 2014;90(1):94-101.
151. Paulsson AK, Yom SS, Anwar M, et al. Respiration-Induced Intraorgan Deformation of the Liver: Implications for Treatment Planning in Patients Treated With Fiducial Tracking. *Technology in cancer research & treatment*. 2017;16(6):776-782.
152. Lu X-Q, Shanmugham LN, Mahadevan A, et al. Organ deformation and dose coverage in robotic respiratory-tracking radiotherapy. *International Journal of Radiation Oncology\* Biology\* Physics*. 2008;71(1):281-289.
153. Velec M, Moseley JL, Craig T, Dawson LA, Brock KK. Accumulated dose in liver stereotactic body radiotherapy: positioning, breathing, and deformation effects. *International Journal of Radiation Oncology\* Biology\* Physics*. 2012;83(4):1132-1140.
154. Lu X-Q, Shanmugham LN, Mahadevan A, et al. Organ Deformation and Dose Coverage in Robotic Respiratory-Tracking Radiotherapy. *International Journal of Radiation Oncology\* Biology\* Physics*. 2008;71(1):281-289.
155. Cao M, Lasley FD, Das IJ, DesRosiers CM, Slessinger ED, Cardenes HR. Evaluation of rotational errors in treatment setup of stereotactic body radiation therapy of liver cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 2012;84(3):e435-e440.
156. Bertholet J, Worm E, Høyer M, Poulsen P. Cone beam CT-based set-up strategies with and without rotational correction for stereotactic body radiation therapy in the liver. *Acta Oncologica*. 2017;56(6):860-866.
157. Van Herk M, Remeijer P, Rasch C, Lebesque JV. The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. *Int J Radiat Oncol Biol Phys*. 2000;47(4):1121-1135.
158. Low DA, Dempsey JF. Evaluation of the gamma dose distribution comparison method. *Medical physics*. 2003;30(9):2455-2464.
159. Low DA. Gamma dose distribution evaluation tool. Paper presented at: Journal of Physics-Conference Series 2010.
160. Wendling M, Zijp LJ, McDermott LN, et al. A fast algorithm for gamma evaluation in 3D. *Medical physics*. 2007;34(5):1647-1654.

## References

---

161. Kilby W, Dooley J, Kuduvalli G, Sayeh S, Maurer Jr C. The CyberKnife® robotic radiosurgery system in 2010. *Technology in cancer research & treatment*. 2010;9(5):433-452.
162. Neter J, Kutner MH, Nachtsheim CJ, Wasserman W. *Applied linear statistical models*. Vol 4: Irwin Chicago; 1996.
163. Fagerland MW. t-tests, non-parametric tests, and large studies—a paradox of statistical practice? *BMC Medical Research Methodology*. 2012;12(1):78-84.
164. Ruxton GD. The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. *Behavioral Ecology*. 2006;17(4):688-690.
165. Mukaka MM. A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal*. 2012;24(3):69-71.
166. Langer MP, Papiez L, Spirydovich S, Thai V. The need for rotational margins in intensity-modulated radiotherapy and a new method for planning target volume design. *Int. J. Radiat. Oncol. Biol. Phys.* 2005;63(5):1592-1603.
167. Jabbour SK, Hashem SA, Bosch W, et al. Upper abdominal normal organ contouring guidelines and atlas: a Radiation Therapy Oncology Group consensus. *Practical radiation oncology*. 2014;4(2):82-89.
168. Kitamura K, Shirato H, Seppenwoolde Y, et al. Tumor location, cirrhosis, and surgical history contribute to tumor movement in the liver, as measured during stereotactic irradiation using a real-time tumor-tracking radiotherapy system. *Int. J. Radiat. Oncol. Biol. Phys.* 2003;56(1):221-228.
169. Marants R, Vandervoort E, Cygler JE. Evaluation of the 4D RADPOS dosimetry system for dose and position quality assurance of CyberKnife. *Medical physics*. 2018;45(9):4030-4044.
170. Holmes OE, Gratton J, Szanto J, et al. Reducing errors in prostate tracking with an improved fiducial implantation protocol for CyberKnife based stereotactic body radiotherapy (SBRT). *Journal of Radiosurgery and SBRT*. 2018;5(3):217-227.
171. Holmes OE, Gratton J, Szanto J, et al. Reducing errors in prostate tracking with an improved fiducial implantation protocol for CyberKnife based stereotactic body radiotherapy (SBRT). *Journal of radiosurgery and SBRT*. 2018;5(3):217.
172. Marants R, Vandervoort E, Cygler JE. Evaluation of the 4D RADPOS dosimetry system for dose and position quality assurance of CyberKnife. *Med. Phys.* 2018;45(9):4030-4044.
173. Langen K, Jones D. Organ motion and its management. *Int. J. Radiat. Oncol., Biol., Phys.* 2001;50(1):265-278.
174. Gierga DP, Brewer J, Sharp GC, Betke M, Willett CG, Chen GTY. The correlation between internal and external markers for abdominal tumors: implications for respiratory gating. *Int. J. Radiat. Oncol., Biol., Phys.* 2005;61(5):1551-1558.
175. Keall P, Poulsen P, Booth JT. See, Think, and Act: Real-Time Adaptive Radiotherapy. *Seminars in Radiation Oncology*. 2019;29:228-235.
176. Zou W, Dong L, Kevin Teo BK. Current State of Image Guidance in Radiation Oncology: Implications for PTV Margin Expansion and Adaptive Therapy. *Semin Radiat Oncol*. 2018;28(3):238-247.
177. Tolan S, Kong V, Rosewall T, et al. Patient-specific PTV margins in radiotherapy for bladder cancer—a feasibility study using cone beam CT. *Radiotherapy and Oncology*. 2011;99(2):131-136.
178. van den Bosch M, Öllers M, Reymen B, van Elmpt W. Automatic selection of lung cancer patients for adaptive radiotherapy using cone-beam CT imaging. *Physics and Imaging in Radiation Oncology*. 2017;1:21-27.

## References

---

179. Haasbeek CJA, Lagerwaard FJ, Cuijpers JP, Slotman BJ, Senan S. Is adaptive treatment planning required for stereotactic radiotherapy of stage I non-small-cell lung cancer? *Int. J. Radiat. Oncol., Biol., Phys.* 2007;67(5):1370-1374.
180. Woodford C, Yartsev S, Dar AR, Bauman G, Van Dyk J. Adaptive radiotherapy planning on decreasing gross tumor volumes as seen on megavoltage computed tomography images. *Int. J. Radiat. Oncol., Biol., Phys.* 2007;69(4):1316-1322.
181. Boda-Heggemann J, Jahnke A, Chan MKH, et al. In-vivo treatment accuracy analysis of active motion-compensated liver SBRT through registration of plan dose to post-therapeutic MRI-morphologic alterations. *Radiotherapy and Oncology.* 2019;134:158-165.
182. Kang JMDP, Schwartz RP, Flickinger JMD, Beriwal SMD. Machine Learning Approaches for Predicting Radiation Therapy Outcomes: A Clinician's Perspective. *Int. J. Radiat. Oncol., Biol., Phys.* 2015;93(5):1127-1135.
183. Bibault J-E, Giraud P, Burgun A. Big data and machine learning in radiation oncology: state of the art and future prospects. *Cancer letters.* 2016;382(1):110-117.
184. Zhu X, Ge Y, Li T, Thongphiew D, Yin FF, Wu QJ. A planning quality evaluation tool for prostate adaptive IMRT based on machine learning. *Medical physics.* 2011;38(2):719-726.
185. Chen S, Zhou S, Yin FF, Marks LB, Das SK. Investigation of the support vector machine algorithm to predict lung radiation-induced pneumonitis. *Medical Physics.* 2007;34(10):3808-3814.
186. Guidi G, Maffei N, Vecchi C, et al. A support vector machine tool for adaptive tomotherapy treatments: prediction of head and neck patients criticalities. *Physica Medica.* 2015;31(5):442-451.
187. Ibragimov B, Toesca D, Chang D, Yuan Y, Koong A, Xing L. Development of deep neural network for individualized hepatobiliary toxicity prediction after liver SBRT. *Medical physics.* 2018;45(10):4763-4774.
188. Toesca DAS, Ibragimov B, Koong AJ, Xing L, Koong AC, Chang DT. Strategies for prediction and mitigation of radiation-induced liver toxicity. *Journal of radiation research.* 2018;59(suppl. 1):40-49.
189. Vogel L, Sihono DSK, Weiss C, et al. Intra-breath-hold residual motion of image-guided DIBH liver-SBRT: an estimation by ultrasound-based monitoring correlated with diaphragm position in CBCT. *Radiotherapy and Oncology.* 2018;129(3):441-448.
190. Inoue M, Okawa K, Taguchi J, et al. Factors affecting the accuracy of respiratory tracking of the image-guided robotic radiosurgery system. *Japanese journal of radiology.* 2019;37(10):727-734.
191. Feng M, Suresh K, Schipper MJ, et al. Individualized adaptive stereotactic body radiotherapy for liver tumors in patients at high risk for liver damage: a phase 2 clinical trial. *JAMA oncology.* 2018;4(1):40-47.
192. Sanuki N, Takeda A, Oku Y, et al. Threshold doses for focal liver reaction after stereotactic ablative body radiation therapy for small hepatocellular carcinoma depend on liver function: evaluation on magnetic resonance imaging with Gd-EOB-DTPA. *Int. J. Radiat. Oncol., Biol., Phys.* 2014;88(2):306-311.
193. Cao Y, Wang H, Johnson TD, et al. Prediction of liver function by using magnetic resonance-based portal venous perfusion imaging. *Int. J. Radiat. Oncol., Biol., Phys.* 2013;85(1):258-263.
194. Vapnik V. *The nature of statistical learning theory.* Springer science & business media; 2013.

## References

---

195. Wang S, Bowen SR, Chaovalitwongse WA, Sandison GA, Grabowski TJ, Kinahan PE. Respiratory trace feature analysis for the prediction of respiratory-gated PET quantification. *Physics in Medicine & Biology*. 2014;59(4):1027.
196. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine learning in Python. *Journal of machine learning research*. 2011;12:2825-2830.
197. Chang C-C, Lin C-J. LIBSVM: A library for support vector machines. *ACM transactions on intelligent systems and technology*. 2011;2(3):1-27.
198. Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. *Machine learning*. 2002;46(1-3):389-422.
199. Stehman SV. Selecting and interpreting measures of thematic classification accuracy. *Remote sensing of Environment*. 1997;62(1):77-89.
200. Powers DM. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies*. 2011;2(1):37-63.
201. Fawcett T. An introduction to ROC analysis. *Pattern recognition letters*. 2006;27(8):861-874.
202. Schiller TW, Chen Y, El Naqa I, Deasy JO. Modeling radiation-induced lung injury risk with an ensemble of support vector machines. *Neurocomputing*. 2010;73(10-12):1861-1867.
203. Lujan AE, Larsen EW, Balter JM, Ten Haken RK. A method for incorporating organ motion due to breathing into 3D dose calculations. *Medical physics*. 1999;26(5):715-720.
204. Karlsson K, Lax I, Lindbäck E, Poludniowski G. Accuracy of the dose-shift approximation in estimating the delivered dose in SBRT of lung tumors considering setup errors and breathing motions. *Acta Oncologica*. 2017;56(9):1189-1196.
205. Chetty IJ, Rosu M, McShan DL, Fraass BA, Balter JM, Ten Haken RK. Accounting for center - of - mass target motion using convolution methods in Monte Carlo - based dose calculations of the lung. *Medical physics*. 2004;31(4):925-932.
206. Beckham W, Keall P, Siebers J. A fluence-convolution method to calculate radiation therapy dose distributions that incorporate random set-up error. *Physics in Medicine & Biology*. 2002;47(19):3465.
207. Ravkilde T, Keall PJ, Grau C, Høyer M, Poulsen PR. Fast motion-including dose error reconstruction for VMAT with and without MLC tracking. *Phys. Med. Biol.* 2014;59(23):7279.
208. Heydarian M, Hoban P, Beddoe AH. A comparison of dosimetry techniques in stereotactic radiosurgery. *Phys. Med. Biol.* 1996;41(1):93.
209. Rice R, Hansen J, Svensson G, Siddon R. Measurements of dose distributions in small beams of 6 MV x-rays. *Phys. Med. Biol.* 1987;32(9):1087.
210. McKerracher C, Thwaites D. Assessment of new small-field detectors against standard-field detectors for practical stereotactic beam data acquisition. *Phys. Med. Biol.* 1999;44(9):2143.
211. Meredith WJ, Massey JB. CHAPTER XVI - THE PROPERTIES OF THE X-RAY FILM. In: Meredith WJ, Massey JB, eds. *Fundamental Physics of Radiology (Third Edition)*. doi: <https://doi.org/10.1016/B978-0-7236-0778-6.50020-3>; Butterworth-Heinemann; 1977:175-190.
212. Devic S, Seuntjens J, Sham E, et al. Precise radiochromic film dosimetry using a flat-bed document scanner. *Med. Phys.* 2005;32(7Part1):2245-2253.
213. Todorovic M, Fischer M, Cremers F, Thom E, Schmidt R. Evaluation of GafChromic EBT prototype B for external beam dose verification. *Med. Phys.* 2006;33(5):1321-1328.
214. Lynch BD, Kozelka J, Ranade MK, Li JG, Simon WE, Dempsey JF. Important considerations for radiochromic film dosimetry with flatbed CCD scanners and EBT film. *Med. Phys.* 2006;33(12):4551-4556.

## References

---

215. Paelinck L, De Neve W, De Wagter C. Precautions and strategies in using a commercial flatbed scanner for radiochromic film dosimetry. *Phys. Med. Biol.* 2006;52(1):231.
216. Van Battum L, Hoffmans D, Piersma H, Heukelom S. Accurate dosimetry with GafChromic™ EBT film of a photon beam in water: What level is achievable? *Med. Phys.* 2008;35(2):704-716.
217. Moignier C, Huet C, Barraux V, et al. Determination of small field output factors and correction factors using a Monte Carlo method for a 1000 MU/min CyberKnife® system equipped with fixed collimators. *Radiation measurements.* 2014;71:287-292.
218. Bellec J, Delaby N, Jouyaux F, et al. Plan delivery quality assurance for CyberKnife: Statistical process control analysis of 350 film-based patient-specific QAs. *Physica Medica.* 2017;39:50-58.
219. Ploquin N, Kertzsch G, Vandervoort E, Cygler J, Andersen CE, Francescon P. Use of novel fibre-coupled radioluminescence and RADPOS dosimetry systems for total scatter factor measurements in small fields. *Phys. Med. Biol.* 2014;60(1):1.
220. Butson MJ, Cheung T, Peter K. Weak energy dependence of EBT gafchromic film dose response in the 50 kVp–10 MVp X-ray range. *Appl. Radiat. Isot.* 2006;64(1):60-62.
221. Fuss M, Sturtewagen E, De Wagter C, Georg D. Dosimetric characterization of GafChromic EBT film and its implication on film dosimetry quality assurance. *Phys. Med. Biol.* 2007;52(14):4211.
222. Sutherland J, Rogers D. Monte Carlo calculated absorbed-dose energy dependence of EBT and EBT2 film. *Med. Phys.* 2010;37(3):1110-1116.
223. Crijns W, Maes F, Van der Heide U, Van den Heuvel F. Calibrating page sized Gafchromic EBT3 films. *Medical physics.* 2013;40(1):012102.
224. Bouchard H, Lacroix F, Beaudoin G, Carrier JF, Kawrakow I. On the characterization and uncertainty analysis of radiochromic film dosimetry. *Med. Phys.* 2009;36(6Part1):1931-1946.
225. Devic S, Tomic N, Aldelaijan S, et al. Linearization of dose–response curve of the radiochromic film dosimetry system. *Medical physics.* 2012;39(8):4850-4857.
226. Lewis D, Micke A, Yu X, Chan MF. An efficient protocol for radiochromic film dosimetry combining calibration and measurement in a single scan. *Medical physics.* 2012;39(10):6339-6350.
227. Estimation of the imaging dose for the Cyberknife robotic radiosurgery system: white paper. In. Sunnyvale, CA: Accuray, Incorporated; 2010.
228. Cheung T, Butson MJ, Yu PK. Measurement of high energy x-ray beam penumbra with Gafchromic™ EBT radiochromic film. *Medical physics.* 2006;33(8):2912-2914.
229. Metcalfe P, Kron T, Elliott A, Wong T, Hoban P. Dosimetry of 6-MV x-ray beam penumbra. *Medical physics.* 1993;20(5):1439-1445.
230. Li XA, Yu C, Holmes T. A systematic evaluation of air cavity dose perturbation in megavoltage x-ray beams. *Med. Phys.* 2000;27(5):1011-1017.
231. Martens C, Reynaert N, De Wagter C, et al. Underdosage of the upper-airway mucosa for small fields as used in intensity - modulated radiation therapy: A comparison between radiochromic film measurements, Monte Carlo simulations, and collapsed cone convolution calculations. *Med. Phys.* 2002;29(7):1528-1535.
232. Liu S, Mazur TR, Fu Y, Li HH, Mutic S, Yang D. Technical Note: A method to evaluate dosimetric effects on organs-at-risk for treatment delivery systematic uncertainties. *Medical physics.* 2017;44(4):1552-1557.

## References

---

233. Korevaar EW, Habraken SJ, Scandurra D, et al. Practical robustness evaluation in radiotherapy—A photon and proton-proof alternative to PTV-based plan evaluation. *Radiotherapy and Oncology*. 2019;141:267-274.
234. Yang Y, Catalano S, Kelsey CR, Yoo DS, Yin F-F, Cai J. Dosimetric effects of rotational offsets in stereotactic body radiation therapy (SBRT) for lung cancer. *Medical Dosimetry*. 2014;39(1):117-121.
235. Liu H, Andrews M, Markovich A, Zhuang T. Dosimetric effect of uncorrected rotations in lung SBRT with stereotactic imaging guidance. *Physica Medica*. 2017;42:197-202.
236. Tilly D, Ahnesjö A. Fast dose algorithm for generation of dose coverage probability for robustness analysis of fractionated radiotherapy. *Physics in Medicine & Biology*. 2015;60(14):5439.
237. Seegenschmiedt MH, Makoski H-B, Trott K-R, Brady LW. CNS and Brain. In: *Radiotherapy for non-malignant disorders*. Springer; 2008.
238. Du Q, Zhang C, Zhu X, et al. Application of statistical and computational methodology to predict brainstem dosimetry for trigeminal neuralgia stereotactic radiosurgery. *Medical physics*. 2018;45(5):1822-1831.
239. Romanelli P, Conti A, Redaelli I, et al. Cyberknife radiosurgery for trigeminal neuralgia. *Cureus*. 2019;11(10).
240. Romanelli P, Conti A, Bianchi L, Bergantin A, Martinotti A, Beltramo G. Image-guided robotic radiosurgery for trigeminal neuralgia. *Neurosurgery*. 2018;83(5):1023-1030.
241. Villavicencio AT, Lim M, Burneikiene S, et al. Cyberknife radiosurgery for trigeminal neuralgia treatment: a preliminary multicenter experience. *Neurosurgery*. 2008;62(3):647-655.
242. Borchers III JD, Yang H-J, Sakamoto GT, et al. Cyberknife stereotactic radiosurgical rhizotomy for trigeminal neuralgia: anatomic and morphological considerations. *Neurosurgery*. 2009;64(suppl\_2):A91-A95.
243. Flickinger JC, Pollock BE, Kondziolka D, et al. Does increased nerve length within the treatment volume improve trigeminal neuralgia radiosurgery? A prospective double-blind, randomized study. *Int. J. Radiat. Oncol. Biol. Phys.* 2001;51(2):449-454.
244. Gospodarev V, Chakravarthy V, Harms C, et al. Computed Tomography Cisternography for Evaluation of Trigeminal Neuralgia When Magnetic Resonance Imaging Is Contraindicated: Case Report and Review of the Literature. *World Neurosurgery*. 2018;113:180-183.
245. Shiinoki T, Fujii F, Fujimoto K, Yuasa Y, Sera T. A novel dynamic robotic moving phantom system for patient-specific quality assurance in real-time tumor-tracking radiotherapy. *Journal of Applied Clinical Medical Physics*. 2020.
246. Preiswerk F, Arnold P, Fasel B, Cattin PC. A Bayesian Framework for Estimating Respiratory Liver Motion from Sparse Measurements. 2012; Berlin, Heidelberg.
247. Brock KK, Mutic S, McNutt TR, Li H, Kessler ML. Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132. *Med. Phys.* 2017;44(7):e43-e76.
248. Fitzgibbon A, Pilu M, Fisher RB. Direct least square fitting of ellipses. *IEEE Transactions on pattern analysis and machine intelligence*. 1999;21(5):476-480.