



McGill

DOCTORAL THESIS

**Development of a dose-surface map workflow
for dose-outcome studies of the rectum
during radiotherapy for prostate cancer**

Haley M. Patrick

Medical Physics Unit,
Biological and Biomedical Engineering,
McGill University, Montréal, Québec, Canada

June 2023

A thesis submitted to McGill University in partial fulfillment of
the requirements of the degree of *Doctor of Philosophy*

©2023 Haley Patrick

Abstract

Prostate cancer patients undergoing radiotherapy (RT) frequently experience rectal toxicities as an unintended consequence of their treatments. While guidelines to minimize toxicity risk are prevalent, limited understanding of the relationship between rectal radiation dose and toxicity hinders the establishment of improved guidelines. This may be attributable in part to the underutilization of spatially inclusive delivered dose information in dose-outcome studies.

The dose-surface map (DSM) is an emerging dosimetric tool that can quantify the spatial dose distribution to the surface of an organ. It is well suited to be used to study dose deposition patterns in hollow organs such as the rectum. Recently, the use of DSMs for dose accumulation during RT has been proposed but not yet validated experimentally. The objective of this thesis was therefore to develop and validate a means of performing dose accumulation for the rectum during prostate cancer RT using DSMs. Three studies were devised to meet the objective.

As programs to calculate DSMs are not freely available to the scientific public, it was necessary to first create a code base capable of calculating DSMs from treatment plan data. A Python package was developed to reproduce the diverse selection of DSM calculation methods present in the literature. Options included different ways to space out the slicing of structures used to create DSMs and different ways of orientating these slices (planar or non-coplanar alignments).

An investigation into the equivalency of DSMs produced by different calculation methods and their suitability for dose accumulation purposes formed the basis of the first study. Rectum and bladder DSMs were calculated with different slice spacing and orientation methods and compared to each other. DSMs produced using different slice

orientation methods were found to be non-equivalent to each other to the degree that they could impact the results of a cohort comparison, which could have consequences for research reproducibility. Overall, non-coplanar slice alignments were found to produce the most appropriate representation of the rectal surface in DSM form and were thus recommended to be adopted as the standard calculation method for producing DSMs of the rectum and for rectal dose accumulation. Recommendations for other structures were also proposed.

The second study was an experimental validation of the accuracy at which DSM accumulation can recreate the dose delivered to a rectum. Multi-fraction RT treatments were delivered to a rectum phantom for a variety of inter-fraction motion scenarios and delivered surface doses were experimentally quantified using radiochromic film. Accumulated DSMs were calculated for each scenario and compared to the film measurements using gamma analysis. The best agreement between DSMs and films was observed when the non-coplanar slicing method was used (gamma pass rate $\geq 94.5\%$ for all scenarios, 3%/2mm criteria).

The final study used DSMs to compare planned and accumulated rectal doses for prostate cancer patients treated with two different RT regimens. Findings indicated that patients treated with the longer duration (and lower dose per fraction) regimen exhibited more dose deviations from planned than the shorter (and higher dose per fraction) regimen. Using the spatial information contained within DSMs, it was determined that these deviations were attributable to localized changes in rectal wall position over the course of treatment.

This work demonstrates that DSMs show great promise as a tool to calculate and visualize rectum accumulated dose in dosimetric studies. Future work should focus on further validation of the spatial representations presented in single and accumulated DSMs, and aim at promoting improved calculation standardization within the field. Ultimately, DSMs can be used to determine new, spatially-inclusive dosimetric constraints and evaluate RT treatments for the rectum and other hollow organs at risk.

Résumé

Les patients atteints d'un cancer de la prostate et soumis à un traitement de radiothérapie (RT) présentent fréquemment de la toxicité rectale causée involontairement par le traitement. Malgré la prévalence des règles visant à minimiser le risque de toxicité, la compréhension de la relation entre la dose de rayonnement rectal et la toxicité est limitée, ce qui prévient l'amélioration des règles. Cela peut s'expliquer en partie par la sous-utilisation des informations sur la dose spatiale délivrée dans les études dose-effet.

La carte de dose-surface (CDS) est un nouvel outil dosimétrique permettant de quantifier la distribution spatiale de la dose à la surface d'un organe. Elle est bien adaptée à l'étude des schémas de dose administrée dans les organes creux tels que le rectum. Récemment, l'utilisation des (CDSs) pour visualiser l'accumulation de la dose pendant la RT a été proposée, mais n'a pas encore été validée expérimentalement. L'objectif de cette thèse est de développer et de valider un moyen d'effectuer l'accumulation de dose au rectum pendant la RT du cancer de la prostate en utilisant des CDSs. Pour atteindre cet objectif, trois études ont été conçues.

Comme les programmes de calcul des CDSs ne sont pas librement accessibles au public scientifique, il a été nécessaire de créer d'abord un code de base capable de calculer les CDSs à partir des données du plan de traitement. Un module Python a été développé pour reproduire les diverses méthodes de calcul des CDSs présentes dans la littérature. Les options comprenaient différentes façons d'espacer les coupes de structures utilisées pour créer les CDSs ainsi que différentes façons d'orienter ces coupes (alignements planaires ou non planaires).

La première étude constitue d'une investigation sur l'équivalence des CDSs produites par différentes méthodes de calcul et leur adéquation à des fins d'accumulation de doses. Les

CDSs du rectum et de la vessie ont été calculées à l'aide de différentes méthodes d'espacement et d'orientation des coupes et comparées entre elles. Les CDSs produites se sont révélées non équivalentes au point d'avoir un impact sur les résultats d'une comparaison entre cohortes, ce qui pourrait avoir des conséquences sur la reproductibilité de la recherche. Dans l'ensemble, les alignements de coupes non coplanaires représentent la surface rectale la plus appropriée sous forme de CDS. Il est donc recommandé de les adopter comme méthode standard de calcul pour la production de CDSs du rectum et pour l'accumulation de la dose rectale. Des recommandations pour d'autres structures ont également été proposées.

La deuxième étude était une validation expérimentale de la précision avec laquelle l'accumulation de CDS peut recréer la dose délivrée au rectum. Des traitements de RT à fractions multiples ont été administrés à un fantôme de rectum pour une variété de scénarios de mouvement entre les fractions et les doses de surface délivrées ont été quantifiées expérimentalement à l'aide des films radiochromiques. Les CDSs accumulées ont été calculées pour chaque scénario et comparées aux mesures des films par une analyse gamma. La meilleure concordance entre les CDSs et les films a été observée lorsque la méthode de découpage non coplanaire a été utilisée (taux de succès gamma $\geq 94.5\%$ pour tous les scénarios, critère $3\%/2$ mm).

La dernière étude a utilisé les CDSs pour comparer les doses rectales planifiées et accumulées chez des patients atteints de cancer de la prostate et traités avec deux différents schémas de RT. Les résultats ont indiqué que les patients traités avec le schéma de plus longue durée (et de plus faible dose par fraction) présentaient plus d'écarts de dose par rapport à la dose planifiée que le schéma de plus courte durée (et de plus forte dose par fraction). En utilisant les informations spatiales contenues dans les CDSs, les écarts étaient déterminés d'être attribuables à des changements localisés dans la position de la paroi rectale au cours du traitement.

Ce travail démontre que les CDSs sont un outil très effectif pour calculer et visualiser la dose accumulée dans le rectum dans les études dosimétriques. Les travaux futurs sont de concentrer sur une validation profonde des représentations spatiales présentées dans les CDSs uniques et accumulées, et promouvoir une meilleure normalisation des calculs. En conclusion, les CDSs peuvent être utilisées pour déterminer de nouvelles contraintes

dosimétriques spatiales et évaluer les traitements de RT pour les organes à risques selon le rectum et autres organes creux.

Acknowledgements

When I moved to Montreal seven years ago, I had no idea what was in store for me for my graduate studies. Now, as I write this, I am so grateful to be able to say I was incredibly fortunate to be a part of a department filled with thoughtful, supportive people who have helped me grow in more ways than I will ever be able to count.

I would be remiss if I did not first acknowledge my sincere gratitude to my supervisor John Kildea. His guidance, insight, and sharp eye for detail has been instrumental in my development as a scientist and I always left our discussions feeling refreshed and reinvigorated. John has also been instrumental in the many clinical collaborations I have had the fortune to contribute to and I will be forever grateful for the experience those opportunities provided. Sincere thanks must also be extended to my committee members Amine Kamen and Cynthia Ménard their input and assistance across our many meetings.

I would also like to extend my appreciation to the clinical team members at the McGill University Health Centre for their many contributions to this work and my journey. Thank you to Dr. Souhami for being such a supportive collaborator for all these years and always having the time to share your clinical expertise with me. Thank you to all the medical physicists, especially Emily Poon, Russel Ruo, Tanner Connel, Saad Aldelaijan, and Horacio Patrocinio for the (many) questions you fielded from me and the time you dedicated to teaching me just a fraction of your clinical knowledge. Thank you to Joe Larkin for his help with constructing measurement equipment I could never hope to create myself, and to Margery Knewstubb and Tatjana Nisic for all your help with administrative questions and for keeping the department running so smoothly. Lastly, I would like to wish Michael Evans a happy retirement and thank him for being such a supportive presence throughout my studies; from spontaneous radiation safety teaching moments to life advice I've appreciated

so much of what you've taught me.

To the many members of the NICE-ROKS team, thank you so much for your support and friendship. To those who graduated before me, Logan and Hossein, thank you for all your help in my programming endeavors and for your calming presence over these many years. To Felix, Kayla, and James, thank you for always helping me where you could, be it in reviewing manuscripts or being a sympathetic ear to my bumps along the way. And to Luc Galarneau, thank you for all your assistance with various protocol documents and my French writing endeavors, I don't think I could have succeeded without you.

There have been countless friends and colleagues who have been such amazing sources of laughter and companionship here in Montreal. If we have ever shared a meal or heartfelt conversation, please know I include you here as someone who makes the MPU such a special place. I would like to extend special thanks to Liam Carroll, Chris Lund, and Morgan Maher, as well as the many members of the Whiskey O'Clock crew, both virtual and in-person, for your friendship all these years.

Finally, thank you to my parents, Dave and Pat, and sister, Sarah, for your unwavering support and unsolicited puns throughout this entire process. You're shrimply the best.

Contents

1	Introduction	4
1.1	Prostate Cancer	4
1.1.1	Treatment Options	5
1.2	Prostate Radiotherapy	7
1.2.1	Ionizing Radiation	7
1.2.2	Radiation Toxicities	7
1.2.3	Therapeutic Ratio	9
1.2.4	Fractionation Schemes	11
1.2.5	Treatment Planning Workflow	13
1.3	Delivery Uncertainties	15
1.3.1	Prostate Motion	15
1.3.2	Image Guided Radiotherapy	17
1.4	Thesis Goal and Objectives	19
1.5	Overview of Thesis	20
1.5.1	Peripheral Publications	21
2	Background	33
2.1	A Brief History of Dose-Outcome Research	33
2.1.1	Early Era	33
2.1.2	Megavoltage Era	34
2.1.3	Conformal Era	34
2.1.4	Modern Era	36
2.2	Limitations of Existing Dose-Outcome Studies	37

2.2.1	Planned versus Delivered Dose	38
2.2.2	Spatial Dose Effects	38
2.3	Dose Accumulation	40
2.3.1	Anatomical Information	40
2.3.2	Dose Calculation Method	41
2.3.3	Dose Summation Strategies	43
2.4	Dose-Surface Maps	46
2.4.1	Dose-Surface Map Accumulation	48
3	Technical Note: <i>rtdsm</i> — An open-source software for radiotherapy dose-surface map generation and analysis	65
3.1	Preface	65
3.2	Abstract	66
3.3	Introduction	67
3.4	Methods	68
3.4.1	Planning Phase	68
3.4.2	Stage 1: Mesh Creation	70
3.4.3	Stage 2: Mesh Slicing	71
3.4.4	Stages 3 and 4: Dose Sampling and Unwrapping	74
3.4.5	Post-processing Functionalities	74
3.5	Results	75
3.5.1	Dose-Surface Maps	75
3.5.2	DSM Conversions and Combinations	76
3.5.3	Spatial Features	76
3.6	Discussion	78
3.7	Conclusions	81
3.8	Acknowledgements	81
3.9	Supplemental Materials	88
4	More than one way to skin a dose volume: the impact of dose-surface map calculation approach on study reproducibility	93

4.1	Preface	93
4.2	Abstract	94
4.3	Introduction	95
4.4	Methods	97
4.4.1	Patient Cohort	97
4.4.2	DSM Calculation Workflow	99
4.4.3	DSM Analysis Technique	99
4.5	Results	101
4.5.1	Influence of Calculation Approach on Pixel-wise Comparisons	101
4.5.2	Influence of Calculation Approach on DSM Features (Aims 1 and 2)	105
4.5.3	Influence of Calculation Approach on the Conclusions of a Cohort Comparison (Aim 3)	106
4.6	Discussion	110
4.6.1	Aim 1: Equivalence of planar and non-coplanar DSMs	110
4.6.2	Aim 2: Equivalence of scaled and fixed slice spacing DSMs	112
4.6.3	Aim 3: DSM Feature robustness against DSM calculation method	114
4.6.4	Relation to current state of reproducibility in the literature	114
4.7	Conclusion	116
4.8	Acknowledgements	117
4.9	Supplemental Materials	121
4.9.1	Supplement A: Multiple Comparison Permutation Testing	121
4.9.2	Supplement B: Additional Figures	127
5	Experimental validation of dose accumulation for the rectum using dose-surface maps	130
5.1	Preface	130
5.2	Abstract	131
5.3	Background	132
5.4	Material and Methods	134
5.4.1	Phantom Design	134

5.4.2	Treatment Planning	134
5.4.3	Experimental Setups	135
5.4.4	Film Dosimetry	137
5.4.5	Dose-Surface Map Calculation and Accumulation	137
5.4.6	Analysis	138
5.5	Results	139
5.5.1	(1) Dose plane in solid water	139
5.5.2	(2) Dose to a stationary rectum	140
5.5.3	(3) Dose to a rectum with inter-fraction motion	142
5.5.4	(4) Dose to a rectum with inter-fraction volume changes	142
5.5.5	Accumulated dose as calculated by DIR	142
5.6	Discussion	145
5.7	Acknowledgements	148
6	A spatial investigation of real-world rectal dose delivery accuracy during radiotherapy for prostate cancer	154
6.1	Preface	154
6.2	Abstract	155
6.3	Introduction	156
6.4	Methods	158
6.4.1	Patient Cohorts	158
6.4.2	DSM Calculation	159
6.4.3	Comparison of Planned and Delivered Doses	159
6.4.4	Influence of Rectal Shape Metrics	160
6.5	Results	160
6.5.1	Comparison of Planned and Delivered Doses	160
6.5.2	Influence of Rectal Shape Metrics	162
6.6	Discussion	164
6.7	Conclusions	166
6.8	Acknowledgements	167

6.9	Supplemental Materials	173
7	Summary and Outlook	176
7.1	Summary	176
7.1.1	Objective 1: Development of software for DSM calculation	177
7.1.2	Objective 2: An evaluation of DSM stability with calculation approach	177
7.1.3	Objective 3: Experimental validation of DSM accumulation	178
7.1.4	Objective 4: Practical demonstration of DSM accumulation	179
7.2	Future Directions	180
7.2.1	Further investigations of DSM-based dose accumulation	180
7.2.2	Further clinical studies	181
7.2.3	Improvements to DSM calculation software	182
7.2.4	Broader clinical adoption of DSMs	183

List of Figures

1.1	Examples of TCP and NTCP curves for (a) an ideal scenario where high tumour control probability and low normal tissue complication probability can be achieved, and (b) a non-ideal scenario where one cannot be achieved without sacrificing the other.	10
1.2	Visual depictions of the dose deposition distributions of single beam, 3D-CRT, IMRT, and VMAT treatments. The prostate target is represented by a blue circle while the rectum and bladder and represented by black and yellow circles, respectively.	10
1.3	(a) Single-fraction linear quadratic survival curves for typical cancerous and normal tissues. (b) Effective survival curve for healthy tissue created by fractionation, with a single fraction healthy tissue survival curve for reference.	11
1.4	Sagittal view of the key components of the male pelvic anatomy relevant to prostate radiotherapy.	16
2.1	Visual explanation of the DVH concept. A diamond-shaped structure contains a distribution of dose voxels (left), which can be used to obtain a DVH (right) by computing a cumulative histogram of the voxel values contained within it. Examples of the V_{doseGy} and $D_{vol\%}$ are also shown.	35
2.2	Dose-volume limits proposed in the literature and used as a foundation for QUANTEC rectum dosimetric constraints. All values have been converted to be for standard fractionation (2 Gy/fraction). Thicker lines indicate that higher toxicity rates were observed. Figure reproduced from [13].	37

2.3	Despite rectums A and B receiving very different dose distributions, their DVHs are identical after the spatial information is removed.	39
2.4	Beam and detector configurations for (a) fan beam and (b) cone beam CT imaging.	42
2.5	Overview of the process by which a deformable image registration algorithm deforms a moving image into the reference image.	45
2.6	Flowchart of the basic DSM calculation process. Using a contour from a treatment plan, slices of the structure are defined along which dose will be sampled from the plan's dose distribution. The cylindrical set of dose points will then be cut open and unfurled to produce a DSM.	47
2.7	DSM features are derived from masks of isodose regions within a DSM. Several popular features in the literature are shown.	49
2.8	DSM-based dose accumulation is hypothetically feasible provided the dose-surface map calculation process can accurately map the dose to a given point in the rectum to the same location within the DSM for each fraction.	49
3.1	The four key stages of a typical dose-surface map generation workflow for an organ at risk, as identified in our literature review and implemented in <i>rtdsm</i>	71
3.2	Visual explanations of specific operations performed during planar and non-coplanar mesh slicing. (a) Two-step ray-casting approach to acquire dose sampling points used by both slicing methods. (b) Calculation of CAP tangent vectors to define planes for non-coplanar slicing. (c) Correction method to resolve overlapping non-coplanar slices.	73
3.3	Planar DSMs (left), non-coplanar DSMs (centre), and the slices of the ROI used to construct the DSMs (right) of four example patients (units: Gy). Individual vertical axes are used for each DSM in order to illustrate how choice of slicing method influences sampling path length and DSM shape. In the right column, planar slices are shown in pale red and non-coplanar slices in dark blue. For clarity, only every second slice is shown.	77

3.4	Average dose-surface maps for the IMRT and SBRT cohorts (10 patients each), before and after EQD2 Gy conversion, in units of Gy. Difference maps are also shown with contours indicating specific dose thresholds.	78
3.5	Boxplots of the dose features of the IMRT and SBRT cohorts, converted to EQD2 Gy: (a) percent volume (from DVHs), (b) percent area (from DSMs), (c) lateral projection (from DSMs), and (d) longitudinal projection (from DSMs).	79
4.1	Visualization of all DSM calculation approaches and comparisons used in this work for (a) rectum and (b) bladder structures. Individual comparisons are indicated with capital letters and colors that are consistent across all figures in this article. Examples of typical bladder and rectum DSMs and how to interpret them are shown in (c).	98
4.2	Visualization of the DSM features used in this work, calculated as originally defined in the literature [6, 14]. As is convention, each feature is reported as a percentage of either the total area or total dimensional extent (lateral or longitudinal) of the DSM used to calculate it.	101
4.3	Quantitative comparisons of DSM calculation approaches. Each comparison is indicated as a row or column with its letter label and arrow from Figure 1, and a correspondingly colored box. Each comparison box contains: the cohort-mean DSMs of the two approaches being compared, their DDM, and the map of significantly different pixels between the two as determined by MCP testing. All maps are shown in units of Gy except for the MCP test result maps, which show p-values.	103
4.4	Comparisons of DSM calculation approaches for our representative patient. Each comparison is indicated with its respective label, arrow, and colored box and includes the individual DSMs calculated by each approach along with their DDM. All maps are in units of Gy.	104

-
- 4.5 Mean DSM feature values (and standard uncertainty of the mean) from the four rectum DSM calculation approaches. In order, they are (a) cluster area, (b) cluster lateral extent, (c) cluster longitudinal extent, (d) ellipse lateral extent, and (e) ellipse longitudinal extent. Significantly different pairs of features are indicated with a bracket and an asterisk. Figures of the other dose levels are provided in the supplemental material. 107
- 4.6 DDMs (left) and maps of significantly different pixels (right) between the benchmark and comparison cohorts for each of the DSM calculation approaches tested. “P” refers to planar slice orientations and “NP” to non-coplanar. DDMs are in units of Gy. 108
- 4.7 Mean differences in DSM feature values between the benchmark and comparison cohorts for each of the four rectum DSM calculation approaches. Error bars indicate 95% confidence intervals and statistically significant differences between the cohorts are indicated with asterisks. 109
- 4.8 Examples of how the choice of DSM calculation approach influences the slices of the rectum (and subsequent appearance of the DSM) of the representative patient. Subfigures cover each comparison made in this work. The 15, 35, and 55 Gy isodose regions are included as purple, pink, and yellow overlays, respectively, for dosimetric context. Additionally, for visual clarity, only every fourth slice is shown. Dashed lines are used to represent every 12th slice to make visual comparisons of relative slice positions between slicing styles easier. 111
- 4.9 (a) Comparison of scaled bladder DSMs calculated with 30 scaled spacing slices (median 2.5 mm) and fixed slice spacing DSMs. Mean DSMs are shown for both approaches along with their DDM and p-value map. (b) Visual demonstration of the slice desynching that occurs between scaled and planar bladder DSMs. Only every 4th slice is shown for visual clarity, with every 12th slice represented with dashed lines. 113

4.10	An example of a basic permutation test. We hypothesize the average male is taller than the average female, therefore under the null hypothesis we would expect the mean height difference (our test statistic) to be zero. To test this, we calculate the mean height difference many times, randomly grouping our data into "male" and "female" categories each time. Once we have a distribution we can determine where the test statistic of our observed sample lies within it and determine if the null hypothesis can be rejected.	122
4.11	Visual explanation of how accounting for variance in test statistic value impacts the choice of T_{max} value. If variance is not taken into consideration, a value from the 75th percentile of one voxel's distribution of values can be selected over a value in another voxel's 98th percentile, which is arguably a greater outlier.	123
4.12	Visual depictions of the unpaired (a) and paired (b) MCP testing process.	125
4.13	Mean DSM feature values (and standard uncertainty of the mean) from the four rectum DSM calculation approaches of the 35 and 45 Gy clusters. Significantly different pairs of features are indicated with a bracket and an asterisk.	128
4.14	DSM feature values from the four rectum DSM calculation approaches for the representative (median) patient. Features from the patients with the shortest (blue) and tallest (red) rectums are included as well.	129
5.1	(a) Set-up of the water tank with the smallest diameter rectum pipe in place. (b) Visualizations of the four measurement set-ups tested in this study. A three-dimensional visualization is provided for the solid water set-up and a sagittal view visualization is provided for the water tank experiments.	135
5.2	Total accumulated dose distributions for each measurement scenario, as measured by film and calculated DSMs, along with their gamma index maps for the 3%/2 mm criteria. Isodose lines are included to facilitate visual comparison of the film and DSM dose distributions.	141

5.3	Dose profiles for the (a) anterior, (b) posterior, (c) left, and (d) right walls of the rectum phantom in the volume changing scenario. Gamma index values for the 3%/2 mm criteria are included. Both DSM- and DIR-based results are shown.	144
5.4	Visual comparison of DSMs created from a cylinder angled 10° in a striped dose distribution aligned with the cylinder’s central axis to demonstrate the effect that the choice of DSM calculation style has on DSM appearance. A film measurement of this scenario would feature parallel stripes of high and low dose. For calculated DSMs, the stripes will appear straight and match the film measurement for a non-coplanar DSM (a), but will be distorted and non-equivalent to the film for a planar DSM (b).	147
6.1	Planned minus accumulated-delivered dose difference maps (DDMs) of the patients with the most different (a-b), median different (c-d), and most similar (e-f) planned and accumulated DSMs, in units of Gy. Patients in the IMRT (60 Gy in 20 fractions) cohort are in the left column and patients in the SBRT cohort (36.25 Gy in 5 fractions) are in the right column. Subregions with statistically significant dose differences are contoured in black. Patients are identified by the grey labels.	161
6.2	Cohort-average dose difference maps for the IMRT (a) and SBRT (b) cohorts.	162
6.3	Violin plots of the relative change (planned minus delivered) in daily rectal volume from planning value for the overall IMRT and SBRT cohorts (All), as well as individually numbered patients. Statistically significant results are indicated in red. Distribution medians are shown as rectangular markers.	163
6.4	Mean rectal wall and PTV positions at planning and during treatment for the IMRT (a) and SBRT (b) cohorts, as well as their mean positions over the first (c) and last (d) five fractions for the IMRT cohort. Shaded regions indicate standard uncertainty of the mean, with purple areas indicating regions of statistically significant anterior-posterior shifts.	164

List of Tables

3.1	Python library dependencies of <i>rtdsm</i>	70
3.2	Spatial dose features supported by <i>rtdsm</i>	75
5.1	Gamma pass rates for each combination of experimental set-up and test criterion.	139

Glossary

3D-CRT	Three-Dimensional Conformal Radiotherapy.
AAPM	American Association of Physicists in Medicine.
AMAB	Assigned Male At Birth.
AUC	Area under the receiver operating characteristic Curve.
BED	Biologically Effective Dose.
CAP	Central Axis Pat.
CBCT	Cone-Beam CT.
CT	Computed Tomography.
CTV	Clinical Target Volume.
DDM	Dose Difference Map.
DIR	Deformable Image Registration.
DSB	Double Strand Break.
DSC	Dice Score Coefficient.
DSH	Dose Surface Histogram.
DSM	Dose-Surface Map.
DVH	Dose-Volume Histogram.
EBRT	External Beam Radiotherapy.
ED	Electron Density.
EQD	Equivalent Dose.
ESTRO ACROP	European Society for Radiotherapy and Oncology Advisory Committee for Radiation Oncology Practice.
FBCT	Fan-Beam CT.

FSU	Functional Sub-Unit.
GTV	Gross Target Volume.
HU	Hounsfield Unit.
HyTEC	Hypofractionated Treatment Effects in the Clinic.
ICRU	International Commission on Radiation Units and measurements.
IGRT	Image Guided Radiotherapy.
IMRT	Intensity Modulated Radiotherapy.
MCP	Multiple Comparisons Permutation.
MLC	Multi-Leaf Collimator.
MRI	Magnetic Resonance Image (or Imaging).
MV	Megavoltage.
NTCP	Normal Tissue Complication Probability.
OAR	Organ At Risk.
PSA	Prostate Specific Antigen.
PTV	Planning Target Volume.
QUANTEC	QUantitative Analyses of Normal Tissue Effects in the Clinic.
ROI	Region Of Interest.
SBRT	Stereotactic Body Radiotherapy.
SSR	statistically Significant Subregion.
TCP	Tumour Control Probability.
TERMA	Total Energy Released per unit MAss.
TG	Task Group.
TPS	Treatment Planning Software.
VMAT	Volumetric Arc Therapy.

Preface and Contribution of Authors

This thesis contains four original manuscripts, one of which is published and three that have been submitted to journals. Each manuscript represents an original contribution to the fields of medical physics and radiotherapy dose-surface mapping and can be summarized as follows:

1. The creation of the first open-source code base for the calculation and analysis of a diverse range of dose-surface maps.
2. An investigation of the stability of dose-surface map information in response to different calculation approaches which found significant non-equivalencies and an associated set of recommendations for future best practices.
3. The first experimental investigation of the accuracy at which accumulated dose-surface maps can reproduce the true delivered dose to a rectum phantom for a multi-fraction treatment.
4. A comparison study of planned and delivered rectal doses during two dosimetrically different courses of hypofractionated radiotherapy to demonstrate a potential clinical use case of dose-surface map accumulation.

Author contributions

Chapters 1, 2, and 7 constitute the introduction, theory, and conclusion of this thesis for which I am the sole author. Editing and proof reading of these chapters was provided by John Kildea, James Manalad, and Kayla O’Sullivan-Steben. Douaa El Abiab, Mélodie Cyr,

and Luc Galarneau edited the French translation of the abstract. The author contributions to each of the manuscripts contained in Chapters 3 to 6 are as follows:

1. **Chapter 3:** *Haley M Patrick and John Kildea. Technical note: rtdsm-An open-source software for radiotherapy dose-surface map generation and analysis. Medical Physics. 2022 Nov; 49(11):7327-7335.*

I performed the literature review, developed and quality-tested the software package, and wrote the package documentation and manuscript. John Kildea advised on the package design and reviewed the manuscript. While not an author, Luc Galarneau is acknowledged for his assistance in preparing the package for public release.

2. **Chapter 4:** *Haley M Patrick and John Kildea. More than one way to skin a dose volume: the impact of dose-surface map calculation approach on study reproducibility. Submitted to Physics in Medicine & Biology and awaiting review.*

I designed the study, calculated the dose-surface maps, developed a paired implementation of a previously published pixel-wise multiple comparisons permutation test for analysis, analyzed and interpreted the results, and wrote the manuscript. John Kildea assisted in the design of the study and reviewed the manuscript.

3. **Chapter 5:** *Haley M Patrick, Emily Poon, and John Kildea. Experimental validation of a novel method of dose accumulation for the rectum during prostate radiotherapy. Acta Oncologica (accepted undergoing revisions).*

I designed the study, acquired the measurements, performed the analysis, and wrote the manuscript. Emily Poon provided guidance on film dosimetry, acquired the CT sim images required for treatment planning of the phantom, and assisted with the measurements and analysis. John Kildea provided guidance on all aspects of the study and training to use clinical equipment. All authors reviewed the manuscript.

4. **Chapter 6:** *Haley M Patrick and John Kildea. A spatial investigation of real-world rectal dose delivery accuracy during radiotherapy for prostate cancer. Submitted to Journal of Applied Clinical Medical Physics and awaiting review.*

I designed the study, prepared and processed the data, performed the analysis, and wrote the manuscript. John Kildea contributed to the study design, data interpretation, and reviewing the manuscript.

Chapter 1

Introduction

1.1 Prostate Cancer

Cancer is the leading cause of death in Canada, with 85,100 deaths from the disease projected for 2022 [1]. While approximately 1 in 2.3 Canadians are currently expected to be diagnosed with, and 1 in 4.3 to die from, cancer in their lifetimes [2], these rates are expected to climb with the aging of the population. Prostate cancer is the most common cancer diagnosis in individuals assigned male at birth (AMAB), representing 20.3% of new cases diagnosed in 2021 [2]. Incidence rates have fluctuated over the years in response to screening practices, such as with the advent of prostate-specific antigen (PSA) testing, but have recently stabilized in response to new recommendations to prevent over-diagnosis [3].

While the most prevalent cancer in AMAB people, prostate cancer only kills about 18% of individuals diagnosed with it [4]. Five- and ten-year survival rates for prostate cancer are among the highest of all cancers, being 91% and 88% for all individuals diagnosed and notably higher for those younger than 75 [2]. This can be attributed in part to prostate cancer's slow growth rate. *In vivo* tumour doubling rates are estimated to be around two years, nearly three times slower than estimates for breast cancer [5]. As such, individuals with prostate cancer are more likely to be diagnosed with early stage disease, increasing the diversity and effectiveness of treatment options available to them.

1.1.1 Treatment Options

Active Surveillance and Watchful Waiting

When prostate cancer is diagnosed, an oncologist may recommend active surveillance if the tumour is small, slow-growing, and low-risk. While on active surveillance, patients undergo routine PSA tests, biopsies, and physical exams every 3-6 months to track disease progression in order to avoid unnecessary interventional treatments and the risks of side effects they present. While many people remain on it indefinitely, active surveillance is not curative, meaning individuals with worsening disease will be recommended to switch to a curative treatment option.

Watchful waiting follows the same side effect sparing mentality as active surveillance, but omits regular testing, even in cases where disease has progressed. Instead, the patient monitors their own symptoms and informs their oncologist in the case of a change. It is typically only offered to older individuals with a life expectancy under five years, with the understanding that their quality of life is being prioritized over cancer treatment and that any future treatments offered will be palliative.

Hormone Therapy

Prostate cancer cell growth is aided by androgen hormones like testosterone and can therefore be slowed through androgen deprivation. Modern hormone therapy consists of a combination of compounds that block numerous steps in the testosterone producing pipeline and interfere with testosterone receptors [6, 7]. While this therapy significantly reduces prostate cancer aggressiveness and proliferation it is not curative. As such, it is usually combined with other curative treatments to supplement their effectiveness.

Chemotherapy and Targeted Therapies

Chemotherapy is the use of cytotoxic drugs to kill rapidly-proliferating cells, while targeted therapies use specific molecules or monoclonal antibodies to inhibit proteins that promote cancer growth. Although not commonly used to treat prostate cancer, chemotherapy and targeted therapy may be used to treat metastatic prostate cancer that is no longer be

responding to other therapies.

Surgery

If the cancer is localized to the prostate gland, an oncologist may opt to surgically remove the entire organ through a radical prostatectomy. Prostatectomies can be performed through either open or laparoscopic surgery and can be approached through either the perineum or retropubic abdominal wall [8]. In general, the retropubic approach is favoured as it allows for the biopsy or removal of nearby pelvic lymph nodes and better preservation of sexual function. Patients undergoing prostatectomies may also receive hormone therapy or radiotherapy to control any remaining microscopic disease.

Radiotherapy

Radiotherapy, or radiation therapy, is the use of high-energy electromagnetic waves or subatomic particles to kill cancer cells. Approximately 30% of all prostate cancer patients receive radiotherapy as part of their treatment course, with stage 3 (i.e. intermediate-risk) patients being the most frequent recipients ($\sim 53\%$) [4]. Radiotherapy can be broadly grouped into two categories: external-beam radiotherapy (EBRT) and brachytherapy. EBRT is a non-invasive technique that uses a specialized machine called a linear accelerator (or linac) to direct radiation beams at a patient's cancer from the outside, whereas brachytherapy is an invasive technique that inserts radiation emitting devices directly into the patient's body. Brachytherapy has experienced a steady decline in popularity since 2002 while EBRT usage has continued to increase due to new technological advancements [9]. As this thesis focuses exclusively on EBRT treatments of prostate cancer, all future mentions of radiotherapy will refer to EBRT unless otherwise stated.

1.2 Prostate Radiotherapy

1.2.1 Ionizing Radiation

Radiation is defined as the transmission of energy through space in the form of waves or particles and includes fixtures of everyday life like radio waves and visible light. Radiation is usually categorized as either ionizing or non-ionizing, in reference to its ability to ionize the material it encounters. Ionizing radiation has sufficient energy to eject orbital electrons from atoms, thereby ionizing them, whereas non-ionizing radiation may have sufficient energy to excite electrons, but not eject them.

Ionizing radiations are further categorized as directly or indirectly ionizing based on the mechanism by which they ionize. Directly ionizing radiation involves charged particles (electrons, protons, or heavy ions) that deposit their energy directly into a medium through Coulomb interactions. Indirectly ionizing radiation involves uncharged particles (photons or neutrons) that must first interact with the medium to produce charged particles that then go on to deposit their energies [10]. In all cases, the deposited energy is measured as absorbed dose, defined as the energy deposited per unit mass and uses the SI unit of gray (Gy), which is equivalent to 1 J/kg.

While energy can be deposited in tissue and can damage any cellular component, DNA's essential role for cellular division makes it a critical target for cell kill by ionizing radiation [11]. Cell death is commonly associated with double-strand breaks (DSBs), which are the most difficult type of DNA damage to repair. If a sufficient number of DSBs are created, the DNA damage can be fatal, and the cell will die by either apoptosis or necrosis. However, if the cell is able to sufficiently repair the damage it will be able to continue living and dividing, with any incorrectly repaired damage persisting as one or more genetic mutations.

1.2.2 Radiation Toxicities

Ionizing radiation is non-discriminating and damages cancerous and healthy cells alike. Although healthy cells are typically more radioresistant than cancer cells, healthy cell death is still common at sufficiently high doses of radiation. If just a small proportion of

the cells that make up a tissue die, the surviving cells may be able to compensate for the loss and retain the function of the tissue. However, if a critical threshold of cell death is exceeded, the tissue (or organ) may become compromised and the cell damage manifests itself as radiation toxicity [12].

Radiation toxicities can be either acute or late. Acute toxicities primarily occur due to inflammation of rapidly proliferating cells within hours to weeks of radiation exposure, but also tend to resolve over time. Late toxicities, on the other hand, occur months to years after exposure and are more likely to be irreversible, chronically reducing a person's quality of life.

The radiation tolerance and response of a given tissue or organ are dependent on the organization of clonogenic cells that replicate to replace those killed off. Clusters of these cells are referred to as functional subunits (FSUs) which can be either structurally distinct or structurally undefined and are able to transfer cells between each other [13]. When the FSUs of an organ are organized to operate in tandem with each other, the organ is considered a 'parallel' organ that, much like a parallel electronic circuit, it can retain its function even if several FSUs are damaged (e.g., the kidneys). In comparison, organs that have FSUs that rely on their neighbors to fulfill their functions (e.g., the spinal cord) are considered 'serial' organs and will lose function if even a single FSU is lost. In general, radiation toxicities caused by localized high doses are from serial responses, while toxicities caused by dose baths to entire organs are from parallel responses. The main organs at risk for radiation toxicities during prostate radiotherapy are the bladder and rectum, which have both been shown to act like serial organs for some toxicities and serial-parallel hybrid organs for others [14]. Interestingly, there is growing evidence to suggest that rectum and bladder toxicities are associated with specific spatial dose distribution patterns that go beyond localized high doses or whole-organ dose baths [15–17]. Such associations may suggest that simplified serial/parallel organ models may not sufficiently capture the nuances of organ radiosensitivities and that more sophisticated models may be required.

Due to its proximity to the prostate, the rectum frequently experiences radiation toxicities following prostate radiotherapy. A study by Olopade *et al.* found that 89% of patients who receive pelvic radiotherapy experience chronic changes in bowel habits, with 49% of these

patients significantly affected [18]. It has been estimated that 20-40% of prostate cancer radiotherapy patients will develop some form of bowel toxicity that causes a moderate to severe reduction in their overall quality of life [19]. This is particularly noteworthy in relation to prostate cancer's high survival rates, as patients may continue to suffer from toxicities many years after treatment.

1.2.3 Therapeutic Ratio

Since radiation toxicities may diminish patient quality of life, the primary goal of curative radiotherapy is to maximize the damage to cancerous cells while sparing healthy tissues as much as possible. The radiation response of cancerous and healthy cells can be respectively modeled using Tumour Control Probability (TCP) and Normal Tissue Complication Probability (NTCP) curves: sigmoid functions that model the likelihood of a given outcome as a function of absorbed dose (Fig 1.1) [20]. The challenge for radiotherapy is to therefore maximize the ratio of TCP to NTCP (referred to as the therapeutic ratio) for a given absorbed dose and cancer site. In ideal circumstances, the TCP function is shifted further towards lower doses than the NTCP function (Fig 1.1a), making optimization of the therapeutic ratio straightforward. In practice however, the curves may be much closer together, making optimization much more challenging (Fig 1.1b). As such, radiotherapy makes use of various strategies to separate the curves and increase the therapeutic ratio.

One way to improve the therapeutic ratio is through the use of multiple radiation beams from different angles, centering their area of overlap on the tumour so that it receives a higher dose than nearby healthy tissues. Radiation dose distributions can also be shaped to mimic the shape of the cancerous mass using multi-leaf collimators (MLCs) and intensity modulation, further focusing high radiation doses on the tumour. Different EBRT modalities can be differentiated by the type of MLC and linac gantry motion that occurs when the radiation beam is turned on, with 3D conformal and volumetric modulated arc radiotherapy (3D-CRT and VMAT) being the most and least conformal of the main modalities, respectively (Fig 1.2) [10].

Another way in which the therapeutic ratio can be improved is by breaking up the radiotherapy treatments from one large dose session into multiple, smaller dose sessions

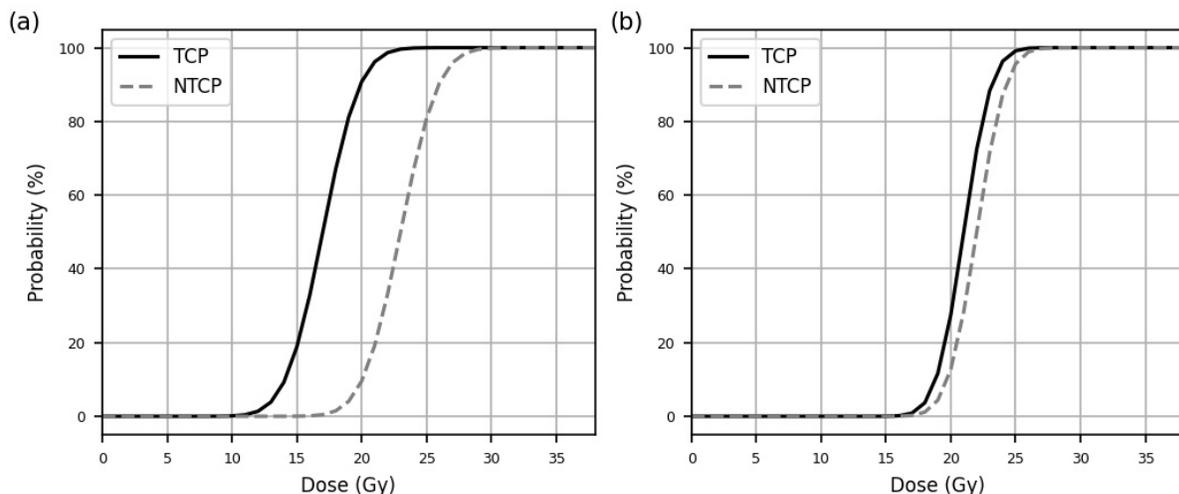


Figure 1.1: Examples of TCP and NTCP curves for (a) an ideal scenario where high tumour control probability and low normal tissue complication probability can be achieved, and (b) a non-ideal scenario where one cannot be achieved without sacrificing the other.

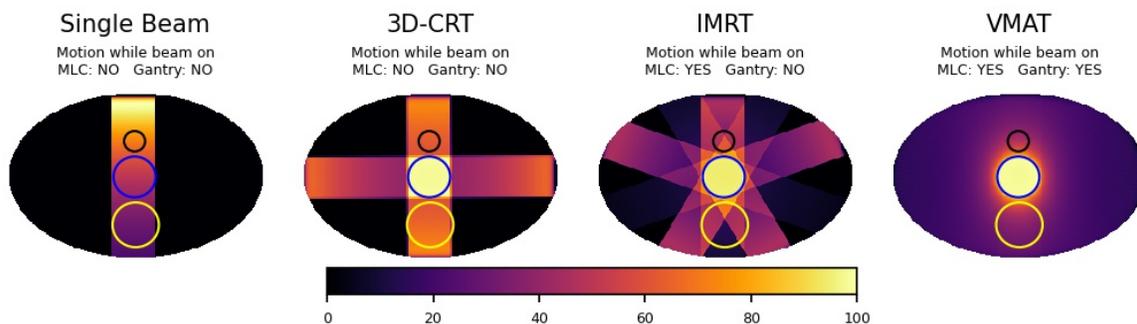


Figure 1.2: Visual depictions of the dose deposition distributions of single beam, 3D-CRT, IMRT, and VMAT treatments. The prostate target is represented by a blue circle while the rectum and bladder are represented by black and yellow circles, respectively.

called fractions. This practice, called fractionation, accounts for the differences in radiation responses of different tissue types to the same radiation dose. The theory behind the practice revolves around the linear-quadratic model of cell survival,

$$SF = e^{-\alpha D - \beta D^2} \quad (1.1)$$

where SF is the surviving fraction of cells after irradiation, D is radiation dose and the

coefficients α and β reflect the DSBs created by single and multiple radiation interaction events, respectively [11]. The radiosensitivity of a given tissue can be characterized by the ratio α/β . In general, fast replicating tissues like cancerous tumours have high α/β ratios whereas slower replicating tissues have lower ones (Fig 1.3a) [13]. By fractionating the dose delivery, the biologically effective dose (BED) of a treatment delivering D total dose (in Gy) is reduced to:

$$BED = D \left[1 + \frac{D/n}{\alpha/\beta} \right] \quad (1.2)$$

where n is the number of fractions. This reduction is because fractionation effectively restarts the cell survival curve at each fraction, transforming the survival curve from linear-quadratic to linear (Fig 1.3b). By carefully selecting the dose per fraction, differences between the α/β ratios of cancer and healthy tissues can be taken advantage of, thereby improving the therapeutic ratio.

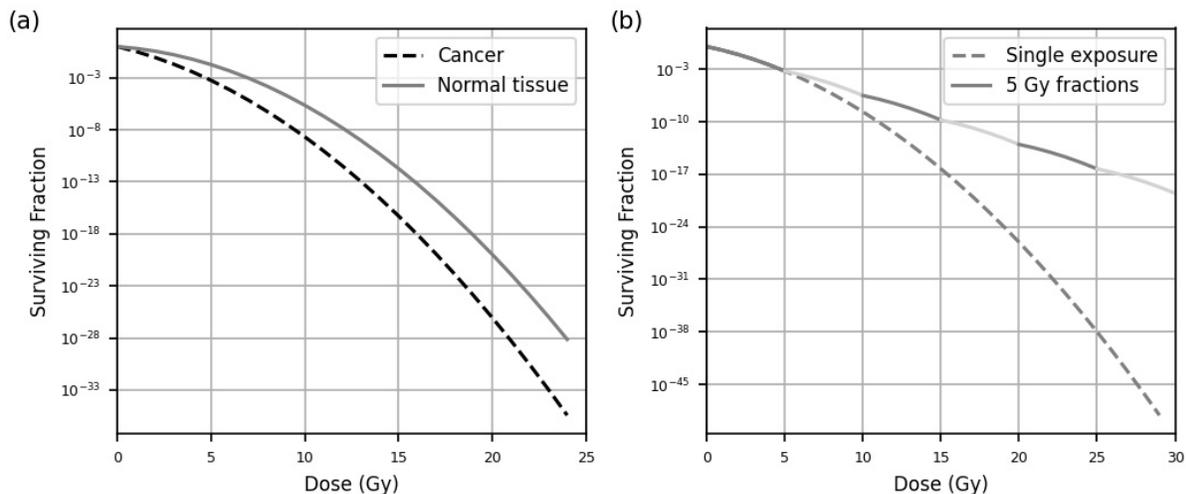


Figure 1.3: (a) Single-fraction linear quadratic survival curves for typical cancerous and normal tissues. (b) Effective survival curve for healthy tissue created by fractionation, with a single fraction healthy tissue survival curve for reference.

1.2.4 Fractionation Schemes

Most cancers have α/β ratios near 10 Gy and are treated with conventional fractionation, delivering 2 Gy per fraction to provide a good balance of cancer kill and normal tissue

sparing. However, prostate cancer is somewhat unique, possessing an α/β ratio closer to 1.5 Gy [21, 22], lower than that of nearby healthy tissues. Hypofractionation, which is the use of higher doses per fraction across fewer fractions, is therefore hypothesized to be more biologically effective at killing prostate cancer while still having sufficiently low NTCP [23]. The effectiveness of hypofractionated prostate radiotherapy compared to conventional fractionation has been the subject of several randomized controlled clinical trials. Most used 3D-CRT or IMRT and focused on the treatment of low or intermediate risk patients. Two of the larger trials, CHHiP and PROFIT, used prescriptions of 60 Gy in 20 fractions and were found to be non-inferior to their equivalent conventional treatment with no increase in late radiation toxicities [24, 25]. Other trials looked at more escalated hypofractionated doses, finding that while patients on the hypofractionated arm had non-inferior biochemical relapse free survival, they also had increased risk of late genitourinary and gastrointestinal toxicities [26–29].

More recently there has been an interest in “extreme hypofractionation”, or stereotactic body radiation therapy (SBRT) for prostate radiotherapy. Rather than standard hypofractionation, which uses doses per fraction around 2.5-4 Gy, SBRT delivers closer to 6-9 Gy per fraction, potentially triggering different physiological and immune responses that contribute to cell kill [30]. While long-term outcomes are still being collected, current evidence suggests SBRT is well suited to treat localized prostate cancer. 5-year failure free survival has been reported to be non-inferior by the HYPO-RT-PC trial for 42.7 Gy in 7 fractions compared to the conventional 78 Gy in 39 fractions along with similar patient quality of life [31, 32]. The PACE-B trial has similarly reported no significant differences in toxicity rates between 36.25 Gy in 5 fractions and 78 Gy in 39 fractions regimens [33], with oncological outcomes still to come. Other clinical trials are also ongoing to assess 36.25 Gy in 5 fraction treatments [34, 35].

In addition to enhanced prostate cancer cell kill, hypofractionation allows for more efficient radiotherapy treatments, reducing the time, clinical resources, and number of visits required to treat a single patient. For this reason, hypofractionation experienced a large swell in adoption during the COVID-19 pandemic in order to minimize patient visits to hospitals [36, 37]. Indeed, extreme hypofractionation was adopted by some centres as the

standard of care for several tumour sites including prostate cancer. While evidence to support this transition does exist, it will be important to evaluate how this rapid change in clinical practice translates to long-term patient outcomes [38–40].

1.2.5 Treatment Planning Workflow

In order to maximize the therapeutic ratio and best treat each radiotherapy patient, each patient’s treatment plan must be designed to fit their unique anatomy and needs. As such, the conventional treatment planning workflow for EBRT prostate treatments is divided into the following four steps:

(1) Computed Tomography Simulation Scan

The treatment planning process begins with the acquisition of a computed tomography (CT) scan to obtain a full 3D anatomical representation of a patient. CT images are preferred over other medical imaging modalities as they provide not only anatomical information, but electron density information that is crucial to accurately model how radiation will interact with the body. During the simulation appointment, the patient is positioned exactly as they will be during the treatment with the help of immobilization devices that ensure reproducible positioning and radio-opaque fiducial markers attached to the patient to define the coordinate system that will be used for treatment planning and delivery. Acquisition of a pelvic magnetic resonance image (MRI) is also recommended, when possible, to better identify prostate localization and disease burden.

(2) Contouring of Critical Structures

Contouring is an important step that defines the locations of critical structures that inform the treatment planning process. Simulation CT images are imported into treatment planning software (TPS) and registered to one another for trained medical professionals to review, identify, and delineate these structures.

Critical structures are divided into two groups: organs at risk (OARs) and targets. OARs are any organs or anatomical structures with specific radiation dose tolerances that

must be respected by the treatment plan, whereas targets are a nested set of volumes that encompass the tumour and various biological and physical uncertainties in the treatment planning process. For prostate radiotherapy there are three main types of target structures [10]:

- **Gross Tumour Volume (GTV):** encompasses the malignant growth visible on the simulation CT images.
- **Clinical Tumour Volume (CTV):** includes the GTV, plus an additional margin that accounts for the actual or expected microscopic spread of disease based on clinical experience. Conventionally, the CTV encompasses the entire prostate.
- **Planning Target Volume (PTV):** includes the CTV, plus additional margins to account for anticipated geometric uncertainties in prostate positioning due to patient set-up or organ motion. This is the structure that the radiation dose is prescribed to in order to ensure all cancerous cells receive the intended dose. In general, modern PTV margins for a prostate CTV range from 3-7 mm depending on the treatment modality and protocol used.

(3) Treatment Planning

Once the simulation CT image is contoured, it is used to determine the optimal way to deliver radiation to the PTV to meet the treatment planning goals. Many delivery factors like beam energy, the number, orientation, shape, and size of radiation fields, and motion of the MLCs and linac gantry can be adjusted during this process using a clinical TPS. The TPS typically contains advanced dose calculation algorithms that use the information from CT images to calculate how radiation beams will interact with a patient's body. The optimization of treatment delivery factors can be done through either forward or inverse planning. However, modern prostate radiotherapy planning is performed almost exclusively through inverse planning. Unlike forward planning, which requires the user to manually set all delivery factors, inverse planning works backwards from the treatment goals of each structure to determine what delivery factors are required to meet them. This approach is an iterative process that tweaks delivery parameters to minimize an objective function

that evaluates the satisfaction of various dose-volume constraints for targets and OARs. Constraints usually take the form of “X% of this structure must receive no more/less than Y Gy” and are described in more detail in Chapter 2. Once a satisfactory treatment plan is generated, it is reviewed by radiation oncologists and medical physicists before being approved for delivery.

(4) Treatment Delivery

For every fraction of the treatment plan, the patient is positioned on the linac couch as closely as possible to the simulation position. Patient alignment is obtained and verified using a combination of immobilization devices, laser guidance systems, and imaging. Once positioning is confirmed, the treatment fraction is delivered as planned.

1.3 Delivery Uncertainties

The accuracy of radiotherapy delivery is just as, if not more important than, the quality of the treatment plan to be delivered. Treatment machine output, treatment planning algorithms, and patient positioning all contribute uncertainties that affect delivery accuracy. The ICRU (International Commission on Radiation Units and measurements) report 24, the AAPM (American Association of Physicists in Medicine) Task Group (TG) 106, and other authoritative bodies recommend that cumulative dosimetric uncertainties be kept within 3-5% [41–43], which has shown to be reasonably achievable with modern dosimetry protocols. While many uncertainty factors can be controlled and accounted for before a patient begins treatment, one of the largest for prostate radiotherapy, organ motion, must be accounted for at every treatment fraction.

1.3.1 Prostate Motion

Historically, management of patient positioning uncertainties was restricted to external alignment. Radiation therapists would ensure surface markers on the patient’s skin were in the same position as during the CT simulation, and their internal organs were assumed to match the simulation positioning within a predictable uncertainty. While this assumption

did not need to hold for treatments during the pre-conformal radiotherapy era, the transition to 3D-CRT and increased usage of dose escalation in the late 1980s [44] necessitated a reevaluation and recognition of the existence of significant prostate motion.

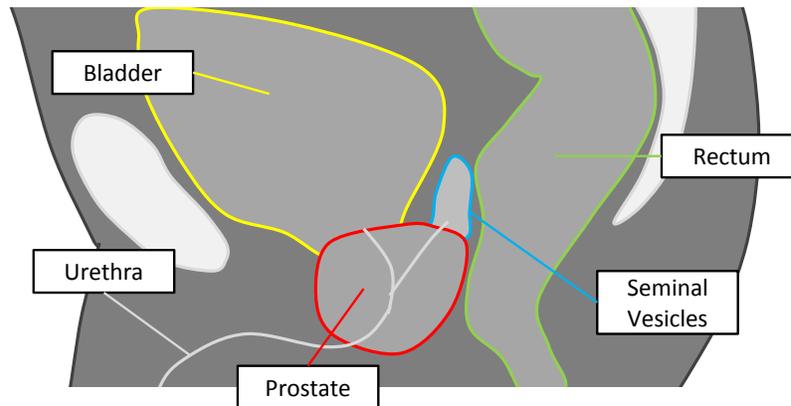


Figure 1.4: Sagittal view of the key components of the male pelvic anatomy relevant to prostate radiotherapy.

The male pelvic cavity contains the rectum, bladder, and reproductive organs along with major arteries, veins, and nerves. The organs are closely packed together, as shown in Figure 1.4, with the prostate snugly located posteriorly-inferiorly to the bladder and anteriorly to the rectum. Due to this proximity, changes in the volume or shape of a neighboring structure can influence the positioning of the prostate. These positional changes are typically attributed to changes in bladder and rectum filling and can occur over the duration of a single fraction (intra-fraction motion) or between fractions (inter-fraction motion). In general, intra-fraction motion is small, on the order of 1-3 mm [45, 46], and tends to occur predictably with gradual increasing of bladder filling over time [47]. Inter-fraction motion, on the other hand, is much larger and more random in nature, posing a significant challenge to treatment delivery accuracy.

The earliest reports of prostate inter-fraction motion were from controlled filling experiments that compared prostate positioning before and after injecting contrast fluid into the rectum and bladder. In these studies, the prostate was observed to shift by 5 mm or more in 62% of participating patients when the rectal volume was increased by 50 cm³ [48, 49]. In the most extreme cases, prostate displacement was nearly 20 mm, well

beyond the typical PTV margins of 10-15 mm used in the early 3D-CRT era [50]. More studies followed that aimed to characterize the level of prostate displacement that occurred over a full course of radiotherapy. It was quickly determined that the largest standard deviations in prostate positioning occurred in the anterior-posterior and superior-inferior directions, in the range of 2.6-3.7 mm and 1.7-3.6 mm, respectively, with maximum displacements on the order of 7-15 mm [51–55]. Lateral motion and rotations may also occur but are generally within ± 1 mm and 16° [56].

Dosimetrically, prostate inter-fraction motion poses a significant risk to adequate CTV coverage. If only external alignment is used, the average prostate dose can reduce by 0.4 Gy for 8 mm isotropic margins or 2.1 Gy for 7 mm margins [57, 58], while the average dose to 95% of the CTV can reduce by 5.1 ± 9.0 Gy [58]. Consequently, patients treated with conformal techniques are at increased risk of local and biochemical failure [59–62], especially if the rectal volume at planning is not representative of the volume during treatment [63]. Dose to the rectum and bladder are similarly, if not more, sensitive to inter-fraction motion [58, 64, 65], as sub-volumes receiving high doses can vary by as much as 26.0% or 62.3%, for these organs respectively, day to day [57]. For this reason, verification of daily prostate positioning through image guidance became increasingly common in the 2000s and is now considered standard practice [66].

1.3.2 Image Guided Radiotherapy

Image guided radiotherapy (IGRT) is broadly defined as the use of an imaging technology to locate the position of the prostate during radiotherapy treatment, typically in conjunction with a conformal treatment modality. While IGRT adoption rates vary by anatomical site, targets that exhibit large inter-fraction motion like the prostate generally have the highest rates of usage [67]. Nowadays, most linear accelerators have at least one type of on-board imaging system integrated into their design, making for straightforward IGRT positioning adjustments prior to treatment.

The European Society for Radiotherapy and Oncology Advisory Committee for Radiation Oncology Practice (ESTRO ACROP) recommends the use of one of the following IGRT techniques to verify daily prostate positioning [66]:

- **Computed Tomography (CT) Imaging** is the most common modern prostate IGRT technique as it provides 3D visualization of bone and soft tissues. Image quality will depend on the type of CT imager used and the imaging radiation dose.
- **2D x-ray radiographs** taken at two perpendicular angles. As soft tissue visualization is limited for this modality, radio-opaque markers must be implanted into the prostate to verify daily positioning.
- **Electromagnetic transponders** implanted in the prostate. No anatomical information is visualised with this modality.
- **Ultrasound** imaging performed either trans-abdominally or trans-perineally. Caution must be taken to ensure the pressure with which the ultrasound transducer is held against the skin does not induce additional prostate displacements [68].
- **Magnetic Resonance Imaging (MRI)** performed using new MR-linacs with fully integrated MR scanners.

The use of IGRT has significantly contributed to the advancement of dose escalation, hypofractionation, and conformal treatments. Inclusion of IGRT has been shown to counteract the effect of rectal distention on failure-free survival [69–71] and significantly reduce late toxicities for otherwise identical treatments [71, 72]. However, although prostate IGRT undeniably improves the accuracy and precision with which dose is delivered to the prostate, it is important to note that is inherently target-focused. This means that while daily-delivered prostate dose will be within a few percent of the intended planned dose, daily rectum and bladder doses can vary up to 30% from planned values [57, 58]. The rectum in particular has been shown to exhibit interfraction motion beyond what can be predicted based on prostate tracking [73], causing rectum delivered dose to differ from what was planned. Consequently, a patient’s rectum dose may exceed known radiation tolerances and cause radiation toxicity despite the original plan indicating otherwise.

Radiation-induced rectal toxicities continue to be common for prostate radiotherapy, even with the adoption of IGRT and conformal treatment modalities. Nearly a third of

patients experience late toxicities, whereas between 54-70% experience acute toxicities, all of which negatively impact quality of life [25, 32]. The situation likely stems from two key consequences of prostate inter-fraction motion:

1. A patient may receive a different rectum dose than what was planned;
2. The treatment planning constraints used to create prostate radiotherapy plans are based on dose-outcome relationships derived from planned, rather than delivered, doses.

This is further confounded by the exclusion of spatial information in most dose-outcome research possibly masking the existence of radiosensitive subregions in the rectum. For these reasons, treatment outcomes cannot be expected to be improved without due investigation and characterization of the relationships between delivered doses and radiation toxicities. The development of methods to calculate and visualize delivered rectal dose are therefore essential to any meaningful advancements in toxicity mitigation.

1.4 Thesis Goal and Objectives

This thesis presents research to establish and validate a methodology to calculate and visualize delivered dose to the rectum wall in a spatially-inclusive manner. Specifically, the use of two dimensional dose-surface maps for dose accumulation purposes was investigated. This was achieved through four objectives:

1. To develop software for the calculation and accumulation of rectal dose-surface maps.
2. To assess the stability of dose-surface map appearance with regards to different calculation methodologies.
3. To experimentally validate the accuracy of accumulated dose distributions calculated from daily dose-surface maps.
4. To demonstrate a use case of accumulated dose-surface maps.

1.5 Overview of Thesis

Chapter 1 introduced the treatment of prostate cancer with radiotherapy and the consequences of prostate inter-fraction motion on rectum delivered dose and toxicities, leading to the description of the objectives of this thesis.

Chapter 2 summarizes the history of dose-outcome research, limitations of the current standards, and alternatives. The concepts of dose accumulation and dose-surface maps are introduced, along with their relevance to dose-outcome studies of the rectum.

Chapter 3 outlines the development of a software package to calculate dose-surface maps in satisfaction of objective 1. The resulting package, *rtdsm*, represents the first comprehensive dose-surface map calculation code base to implement multiple calculation methodologies, allowing for replication of the breadth of existing methodologies in the literature. The open-source release of the software package also represents a potential major step towards improved accessibility of dose-surface map research, which has so far remained largely segregated. This work was published as a technical note in *Medical Physics* [74].

In **Chapter 4**, the first study investigating the sensitivity of dose-surface map appearance and analysis to calculation methodology is presented (objective 2). It was found that different calculation methodologies produced significantly different dose-surface maps, leading to the identification of different analytical findings. The possible implications on reproducibility within the field of dose-surface map research are discussed and several measures to improve the current situation are recommended. This manuscript for this chapter has been submitted to *Physics in Medicine & Biology*.

Chapter 5 compares accumulated rectum doses calculated using dose-surface maps to absorbed dose measurements performed with radiochromic film (objective 3). The accuracy of rectal dose-surface map accumulation was evaluated for different inter-fraction motion scenarios and different calculation methodologies. Results indicated that dose-surface map based dose accumulation generally agrees with measured dose distributions within TG-218 recommended tolerance limits, provided a specific calculation approach is used. This chapter represents the first attempt to experimentally validate dose-surface map based dose accumulation and is undergoing revisions for publication in *Acta Oncologica*.

Chapter 6 demonstrates how dose-surface maps can be used to evaluate the effects of inter-fraction motion on rectum dose during a course of prostate radiotherapy using the methodology identified in Chapter 5. It was found that dose-surface maps could identify deviations in rectal delivered dose that could not be identified using conventional dose-volume histogram data. This chapter has been submitted as a manuscript to the *Journal of Applied Clinical Medical Physics*.

The thesis concludes with **Chapter 7**, which summarizes the work and discusses future possible avenues of accumulated dose-surface map based research.

1.5.1 Peripheral Publications

The following articles were published in parallel to my doctoral studies and represent additional work contributing to, but not directly related to, this thesis or collaborations with other researchers from my institution:

1. Naseri H, Skamene S, Tolba M, Faye MD, Ramia P, Khriugian J, **Patrick H**, Andrade Hernandez AX, David M, Kildea J. ‘Radiomics-based machine learning models to distinguish between metastatic and healthy bone using lesion-center-based geometric regions of interest’. *Sci Rep*. 2022 Jun 14; 12(1): 9866.
2. Di Lalla V, **Patrick H**, Siriani-Ayoub N, Kildea J, Hijal T, Alfieri J. ‘Satisfaction among Cancer Patients Undergoing Radiotherapy during the COVID-19 Pandemic: An Institutional Experience’. *Curr Oncol*. 2021 Apr 10; 28(2): 1507-1517.
3. **Patrick HM**, Souhami L, Kildea J. ‘Reduction of inter-observer contouring variability in daily clinical practice through a retrospective, evidence-based intervention’. *Acta Oncol*. 2021 Feb; 60(2): 229-236.
4. Sepulveda E, **Patrick H**, Freeman CR, Kildea J. ‘Implementation of a DVH Registry to provide constraints and continuous quality monitoring for pediatric CSI treatment planning’. *J Appl Clin Med Phys*. 2021 Jan; 22(1): 191-202.
5. **Patrick HM**, Hijal T, Souhami L, Freeman C, Parker W, Joly L, Kildea J. ‘A Canadian Response to the Coronavirus Disease 2019 (COVID-19) Pandemic: Is

There a Silver Lining for Radiation Oncology Patients?’ *Adv Radiat Oncol.* 2020 Jun 29; 5(4): 774-776.

Bibliography

- [1] Statistics Canada, “Deaths, 2019,” 2020.
- [2] Canadian Cancer Statistics Advisory Committee in collaboration with the Canadian Cancer Society, Statistics Canada and the Public Health Agency of Canada, “Canadian Cancer Statistics 2021,” tech. rep., 2021.
- [3] N. Bell, S. Connor Gorber, A. Shane, M. Joffres, H. Singh, J. Dickinson, E. Shaw, L. Dunfield, and M. Tonelli, “Recommendations on screening for prostate cancer with the prostate-specific antigen test,” *CMAJ*, 2014.
- [4] Cancer Research UK, “Prostate cancer treatment statistics,” 2017.
- [5] D. V. Fournier, E. Weber, W. Hoeffken, M. Bauer, F. Kubli, and V. Barth, “Growth rate of 147 mammary carcinomas,” *Cancer*, 1980.
- [6] C. Huggins, R. E. Stevens, and C. V. Hodges, “The effects of castration on advanced carcinoma of the prostate gland,” *JAMA: Surgery*, 1941.
- [7] E. D. Crawford, M. A. Eisenberger, D. G. McLeod, J. T. Spaulding, R. Benson, F. A. Dorr, B. A. Blumenstein, M. A. Davis, and P. J. Goodman, “A Controlled Trial of Leuprolide with and without Flutamide in Prostatic Carcinoma,” *New England Journal of Medicine*, 1989.
- [8] S. R. Denmeade and J. T. Isaacs, “A history of prostate cancer treatment,” 2002.
- [9] D. G. Petereit, S. J. Frank, A. N. Viswanathan, B. Erickson, P. Eifel, P. L. Nguyen, and D. E. Wazer, “Brachytherapy: Where has it gone?,” 2015.
- [10] E. B. Podgorsak, *Radiation Oncology Physics: A Handbook for Teachers and Students*.
- [11] E. J. Hall and A. J. Giaccia, *Radiobiology for the radiologist*. 2012.

-
- [12] M. S. Peach, T. N. Showalter, and N. Ohri, “Systematic Review of the Relationship between Acute and Late Gastrointestinal Toxicity after Radiotherapy for Prostate Cancer.,” *Prostate cancer*, vol. 2015, p. 624736, 2015.
- [13] H. Rodney Withers, J. M. Taylor, and B. Maciejewski, “Treatment volume and tissue tolerance,” *International Journal of Radiation Oncology*Biophysics*, vol. 14, pp. 751–759, apr 1988.
- [14] C. Fiorino, T. Rancati, and R. Valdagni, “Predictive models of toxicity in external radiotherapy: Dosimetric issues,” in *Cancer*, 2009.
- [15] F. Buettner, S. L. Gulliford, S. Webb, M. R. Sydes, D. P. Dearnaley, and M. Partridge, “Assessing correlations between the spatial distribution of the dose to the rectal wall and late rectal toxicity after prostate radiotherapy: An analysis of data from the MRC RT01 trial (ISRCTN 47772397),” *Physics in Medicine and Biology*, 2009.
- [16] I. Improta, F. Palorini, C. Cozzarini, T. Rancati, B. Avuzzi, P. Franco, C. Degli Esposti, E. Del Mastro, G. Girelli, C. Iotti, V. Vavassori, R. Valdagni, and C. Fiorino, “Bladder spatial-dose descriptors correlate with acute urinary toxicity after radiation therapy for prostate cancer,” *Physica Medica*, 2016.
- [17] L. E. Shelley, M. P. Sutcliffe, S. J. Thomas, D. J. Noble, M. Romanchikova, K. Harrison, A. M. Bates, N. G. Burnet, and R. Jena, “Associations between voxel-level accumulated dose and rectal toxicity in prostate radiotherapy,” *Physics and Imaging in Radiation Oncology*, 2020.
- [18] F. A. Olopade, A. Norman, P. Blake, D. P. Dearnaley, K. J. Harrington, V. Khoo, D. Tait, C. Hackett, and H. J. Andreyev, “A modified Inflammatory Bowel Disease questionnaire and the Vaizey incontinence questionnaire are simple ways to identify patients with significant gastrointestinal symptoms after pelvic radiotherapy,” *British Journal of Cancer*, 2005.
- [19] J. Andreyev, “Gastrointestinal symptoms after pelvic radiotherapy: a new understanding to improve management of symptomatic patients,” 2007.

- [20] D. Begosh-Mayne, S. S. Kumar, S. Toffel, P. Okunieff, and W. O'Dell, "The dose-response characteristics of four NTCP models: using a novel CT-based radiomic method to quantify radiation-induced lung density changes," *Scientific Reports*, 2020.
- [21] D. J. Brenner and E. J. Hall, "Fractionation and protraction for radiotherapy of prostate carcinoma," *International Journal of Radiation Oncology Biology Physics*, 1999.
- [22] R. Miralbell, S. A. Roberts, E. Zubizarreta, and J. H. Hendry, "Dose-fractionation sensitivity of prostate cancer deduced from radiotherapy outcomes of 5,969 patients in seven international institutional datasets: $\alpha/\beta = 1.4$ (0.9-2.2) Gy," *International Journal of Radiation Oncology Biology Physics*, 2012.
- [23] T. Daly, "Evolution of definitive external beam radiation therapy in the treatment of prostate cancer," *World Journal of Urology*, 2020.
- [24] D. Dearnaley, I. Syndikus, H. Mossop, V. Khoo, A. Birtle, D. Bloomfield, J. Graham, P. Kirkbride, J. Logue, Z. Malik, J. Money-Kyrle, J. M. O'Sullivan, M. Panades, C. Parker, H. Patterson, C. Scrase, J. Staffurth, A. Stockdale, J. Tremlett, M. Bidmead, H. Mayles, O. Naismith, C. South, A. Gao, C. Cruickshank, S. Hassan, J. Pugh, C. Griffin, and E. Hall, "Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: 5-year outcomes of the randomised, non-inferiority, phase 3 CHHiP trial," *The Lancet Oncology*, vol. 17, pp. 1047–1060, aug 2016.
- [25] C. N. Catton, H. Lukka, C. S. Gu, J. M. Martin, S. Supiot, P. W. Chung, G. S. Bauman, J. P. Bahary, S. Ahmed, P. Cheung, K. H. Tai, J. S. Wu, M. B. Parliament, T. Tsakiridis, T. B. Corbett, C. Tang, I. S. Dayes, P. Warde, T. K. Craig, J. A. Julian, and M. N. Levine, "Randomized trial of a hypofractionated radiation regimen for the treatment of localized prostate cancer," *Journal of Clinical Oncology*, vol. 35, pp. 1884–1890, jun 2017.
- [26] S. Aluwini, F. Pos, E. Schimmel, E. van Lin, S. Krol, P. P. van der Toorn, H. de Jager, M. Dirksen, W. G. Alemaeyehu, B. Heijmen, and L. Incrocci, "Hypofractionated versus

- conventionally fractionated radiotherapy for patients with prostate cancer (HYPRO): Acute toxicity results from a randomised non-inferiority phase 3 trial,” *The Lancet Oncology*, vol. 16, pp. 274–283, mar 2015.
- [27] S. Aluwini, F. Pos, E. Schimmel, S. Krol, P. P. van der Toorn, H. de Jager, W. G. Alemayehu, W. Heemsbergen, B. Heijmen, and L. Incrocci, “Hypofractionated versus conventionally fractionated radiotherapy for patients with prostate cancer (HYPRO): late toxicity results from a randomised, non-inferiority, phase 3 trial,” *The Lancet Oncology*, 2016.
- [28] L. Incrocci, R. C. Wortel, W. G. Alemayehu, S. Aluwini, E. Schimmel, S. Krol, P. P. van der Toorn, H. de Jager, W. Heemsbergen, B. Heijmen, and F. Pos, “Hypofractionated versus conventionally fractionated radiotherapy for patients with localised prostate cancer (HYPRO): final efficacy results from a randomised, multicentre, open-label, phase 3 trial,” *The Lancet Oncology*, 2016.
- [29] W. R. Lee, J. J. Dignam, M. B. Amin, D. W. Bruner, D. Low, G. P. Swanson, A. B. Shah, D. P. D’Souza, J. M. Michalski, I. S. Dayes, S. A. Seaward, W. A. Hall, P. L. Nguyen, T. M. Pisansky, S. L. Faria, Y. Chen, B. F. Koontz, R. Paulus, and H. M. Sandler, “Randomized phase III noninferiority study comparing two radiotherapy fractionation schedules in patients with low-risk prostate cancer,” *Journal of Clinical Oncology*, 2016.
- [30] M. Garcia-Barros, F. Paris, C. Cordon-Cardo, D. Lyden, S. Rafii, A. Haimovitz-Friedman, Z. Fuks, and R. Kolesnick, “Tumor response to radiotherapy regulated by endothelial cell apoptosis,” *Science*, 2003.
- [31] A. Widmark, A. Gunnlaugsson, L. Beckman, C. Thellenberg-Karlsson, M. Hoyer, M. Lagerlund, J. Kindblom, C. Ginman, B. Johansson, K. Björnlinger, M. Seke, M. Agrup, P. Fransson, B. Tavelin, D. Norman, B. Zackrisson, H. Anderson, E. Kjellén, L. Franzén, and P. Nilsson, “Ultra-hypofractionated versus conventionally fractionated radiotherapy for prostate cancer: 5-year outcomes of the HYPO-RT-PC randomised, non-inferiority, phase 3 trial,” *The Lancet*, 2019.

- [32] P. Fransson, P. Nilsson, A. Gunnlaugsson, L. Beckman, B. Tavelin, D. Norman, C. Thellenberg-Karlsson, M. Hoyer, M. Lagerlund, J. Kindblom, C. Ginman, B. Johansson, K. Björnlinger, M. Seke, M. Agrup, B. Zackrisson, E. Kjellén, L. Franzén, and A. Widmark, “Ultra-hypofractionated versus conventionally fractionated radiotherapy for prostate cancer (HYPO-RT-PC): patient-reported quality-of-life outcomes of a randomised, controlled, non-inferiority, phase 3 trial,” *The Lancet Oncology*, 2021.
- [33] D. H. Brand, A. C. Tree, P. Ostler, H. van der Voet, A. Loblaw, W. Chu, D. Ford, S. Tolan, S. Jain, A. Martin, J. Staffurth, P. Camilleri, K. Kancherla, J. Frew, A. Chan, I. S. Dayes, D. Henderson, S. Brown, C. Cruickshank, S. Burnett, A. Duffton, C. Griffin, V. Hinder, K. Morrison, O. Naismith, E. Hall, N. van As, D. Dodds, E. Lartigau, S. Patton, A. Thompson, M. Winkler, P. Wells, T. Lymberiou, D. Saunders, M. Vilarino-Varela, P. Vavassis, T. Tsakiridis, R. Carlson, G. Rodrigues, J. Tanguay, S. Iqbal, S. Morgan, A. Mihai, A. Li, O. Din, M. Panades, R. Wade, Y. Rimmer, J. Armstrong, and N. Oommen, “Intensity-modulated fractionated radiotherapy versus stereotactic body radiotherapy for prostate cancer (PACE-B): acute toxicity findings from an international, randomised, open-label, phase 3, non-inferiority trial,” *The Lancet Oncology*, 2019.
- [34] M. Abramowitz, D. Kwon, D. Freeman, N. Dogan, T. Eade, S. Punnen, A. D. Pra, A. Kneebone, R. Stoyanova, and A. Pollack, “Early Toxicity and Patient Reported Outcomes from a Radiation Hypofractionation Randomized Trial of Extended vs Accelerated Therapy for Prostate Cancer (HEAT),” *International Journal of Radiation Oncology*Biography*Physics*, 2018.
- [35] S. Yu, J. Mathew, M. Dyer, B. Bloch, and A. Hirsch, “Stereotactic Prostate Radiation Therapy with Simultaneous Moderate Dose Escalation to the Dominant Intraprostatic Lesion: A Dosimetric Study,” *International Journal of Radiation Oncology*Biography*Physics*, 2019.

- [36] K. Tamari, Y. Nagata, T. Mizowaki, T. Kodaira, H. Onishi, K. Ogawa, Y. Shioyama, N. Shigematsu, and T. Uno, "The impact of the COVID-19 pandemic on radiotherapy in Japan: nationwide surveys from May 2020 through June 2021," *Journal of Radiation Research*, vol. 55, 2022.
- [37] J. Nossiter, M. Morris, M. G. Parry, A. Sujenthiran, P. Cathcart, J. van der Meulen, A. Aggarwal, H. Payne, and N. W. Clarke, "Impact of the COVID-19 pandemic on the diagnosis and treatment of men with prostate cancer," *BJU International*, 2022.
- [38] H. Zhong, K. Men, J. Wang, J. van Soest, D. Rosenthal, A. Dekker, Z. Zhang, and Y. Xiao, "The impact of clinical trial quality assurance on outcome in head and neck radiotherapy treatment," *Frontiers in Oncology*, 2019.
- [39] L. J. Peters, B. O'Sullivan, J. Giralt, T. J. Fitzgerald, A. Trotti, J. Bernier, J. Bourhis, K. Yuen, R. Fisher, and D. Rischin, "Critical impact of radiotherapy protocol compliance and quality in the treatment of advanced head and neck cancer: Results from TROG 02.02," *Journal of Clinical Oncology*, 2010.
- [40] R. Kearvell, A. Haworth, M. A. Ebert, J. Murray, B. Hooton, S. Richardson, D. J. Joseph, D. Lamb, N. A. Spry, G. Duchesne, and J. W. Denham, "Quality improvements in prostate radiotherapy: Outcomes and impact of comprehensive quality assurance during the TROG 03.04 'RADAR' trial," *Journal of Medical Imaging and Radiation Oncology*, 2013.
- [41] R. J. Shalek, "Determination of Absorbed Dose in a Patient Irradiated by Beams of X or Gamma Rays in Radiotherapy Procedures," *Medical Physics*, 1977.
- [42] I. J. Das, C. W. Cheng, R. J. Watts, A. Ahnesjö, J. Gibbons, X. A. Li, J. Lowenstein, R. K. Mitra, W. E. Simon, and T. C. Zhu, "Accelerator beam data commissioning equipment and procedures: Report of the TG-106 of the Therapy Physics Committee of the AAPM," 2008.

- [43] D. Thwaites, "Accuracy required and achievable in radiotherapy dosimetry: Have modern technology and techniques changed our views?," in *Journal of Physics: Conference Series*, 2013.
- [44] B. A. Fraass, "The development of conformal radiation therapy," *Medical Physics*, 1995.
- [45] P. Juneja, A. Kneebone, J. T. Booth, D. I. Thwaites, R. Kaur, E. Colvill, J. A. Ng, P. J. Keall, and T. Eade, "Prostate motion during radiotherapy of prostate cancer patients with and without application of a hydrogel spacer: A comparative study," *Radiation Oncology*, 2015.
- [46] J. Boda-Heggemann, F. M. Köhler, H. Wertz, M. Ehmann, B. Hermann, N. Riesenacker, B. Küpper, F. Lohr, and F. Wenz, "Intrafraction motion of the prostate during an IMRT session: A fiducial-based 3D measurement with Cone-beam CT," *Radiation Oncology*, 2008.
- [47] D. M. de Muinck Keizer, L. G. Kerkmeijer, T. Willigenburg, A. L. van Lier, M. D. Hartogh, J. R. van der Voort van Zyp, E. N. de Groot-van Breugel, B. W. Raaymakers, J. J. Lagendijk, and J. C. de Boer, "Prostate intrafraction motion during the preparation and delivery of MR-guided radiotherapy sessions on a 1.5T MR-Linac," *Radiotherapy and Oncology*, 2020.
- [48] R. K. Ten Haken, J. D. Forman, D. K. Heimburger, A. Gerhardsson, D. L. McShan, C. Perez-Tamayo, S. L. Schoepfel, and A. S. Lichter, "Treatment planning issues related to prostate movement in response to differential filling of the rectum and bladder," *International Journal of Radiation Oncology, Biology, Physics*, 1991.
- [49] S. E. Schild, H. E. Casale, and L. P. Bellefontaine, "Movements of the prostate due to rectal and bladder distension: Implications for radiotherapy," *Medical Dosimetry*, 1993.
- [50] K. M. Langen and D. T. Jones, "Organ motion and its management," 2001.
- [51] M. van Herk, A. Bruce, A. P. Guus Kroes, T. Shouman, A. Touw, and J. V. Lebesque, "Quantification of organ motion during conformal radiotherapy of the prostate by three

- dimensional image registration,” *International Journal of Radiation Oncology, Biology, Physics*, vol. 33, pp. 1311–1320, dec 1995.
- [52] J. M. Crook, Y. Raymond, D. Salhani, H. Yang, and B. Esche, “Prostate motion during standard radiotherapy as assessed by fiducial markers,” *Radiotherapy and Oncology*, 1995.
- [53] J. Balter, H. M. Sandler, K. Lam, R. L. Bree, A. S. Lichter, and R. K. Ten Haken, “Measurement of prostate motion over the course of radiotherapy,” *International Journal of Radiation Oncology*Biology*Physics*, 1993.
- [54] J. C. Roeske, J. D. Forman, C. F. Mesina, T. He, C. A. Pelizzari, E. Fontenla, S. Vijayakumar, and G. T. Chen, “Evaluation of changes in the size and location of the prostate, seminal vesicles, bladder, and rectum during a course of external beam radiation therapy,” *International Journal of Radiation Oncology, Biology, Physics*, 1995.
- [55] E. Vigneault, J. Pouliot, J. Laverdière, R. Jean, and M. Dorion, “Electronic portal imaging device detection of radioopaque markers for the evaluation of prostate position during megavoltage irradiation: A clinical study,” *International Journal of Radiation Oncology Biology Physics*, 1997.
- [56] J. F. Aubry, L. Beaulieu, L. M. Girouard, S. Aubin, D. Tremblay, J. Laverdière, and E. Vigneault, “Measurements of intrafraction motion and interfraction and intrafraction rotation of prostate by three-dimensional analysis of daily portal imaging with radiopaque markers,” *International Journal of Radiation Oncology Biology Physics*, vol. 60, pp. 30–39, sep 2004.
- [57] P. M. van Haaren, A. Bel, P. Hofman, M. van Vulpen, A. N. Kotte, and U. A. van der Heide, “Influence of daily setup measurements and corrections on the estimated delivered dose during IMRT treatment of prostate cancer patients,” *Radiotherapy and Oncology*, 2009.

- [58] D. J. Fraser, Y. Chen, E. Poon, F. L. Cury, T. Falco, and F. Verhaegen, “Dosimetric consequences of misalignment and realignment in prostate 3DCRT using intramodality ultrasound image guidance,” *Medical Physics*, 2010.
- [59] B. Engels, G. Soete, D. Verellen, and G. Storme, “Conformal Arc Radiotherapy for Prostate Cancer: Increased Biochemical Failure in Patients With Distended Rectum on the Planning Computed Tomogram Despite Image Guidance by Implanted Markers,” *International Journal of Radiation Oncology Biology Physics*, 2009.
- [60] W. D. Heemsbergen, M. S. Hoogeman, M. G. Witte, S. T. Peeters, L. Incrocci, and J. V. Lebesque, “Increased Risk of Biochemical and Clinical Failure for Prostate Patients with a Large Rectum at Radiotherapy Planning: Results from the Dutch Trial of 68 Gy Versus 78 Gy,” *International Journal of Radiation Oncology Biology Physics*, 2007.
- [61] W. D. Heemsbergen, A. Al-Mamgani, M. G. Witte, M. Van Herk, and J. V. Lebesque, “Radiotherapy with rectangular fields is associated with fewer clinical failures than conformal fields in the high-risk prostate cancer subgroup: Results from a randomized trial,” *Radiotherapy and Oncology*, 2013.
- [62] M. Marcello, M. Ebert, A. Haworth, A. Steigler, A. Kennedy, D. Joseph, and J. Denham, “Association between treatment planning and delivery factors and disease progression in prostate cancer radiotherapy: Results from the TROG 03.04 RADAR trial,” *Radiotherapy and Oncology*, 2018.
- [63] R. De Crevoisier, S. L. Tucker, L. Dong, R. Mohan, R. Cheung, J. D. Cox, and D. A. Kuban, “Increased risk of biochemical and local failure in patients with distended rectum on the planning CT for prostate cancer radiotherapy,” *International Journal of Radiation Oncology Biology Physics*, vol. 62, pp. 965–973, jul 2005.
- [64] R. C. Wortel, L. Incrocci, F. J. Pos, J. V. Lebesque, M. G. Witte, U. A. Van Der Heide, M. Van Herk, and W. D. Heemsbergen, “Acute toxicity after image-guided intensity modulated radiation therapy compared to 3D conformal radiation therapy in prostate cancer patients,” *International Journal of Radiation Oncology Biology Physics*, 2015.

- [65] I. Gauthier, J. F. Carrier, D. Béliveau-Nadeau, B. Fortin, and D. Taussky, “Dosimetric Impact and Theoretical Clinical Benefits of Fiducial Markers for Dose Escalated Prostate Cancer Radiation Treatment,” *International Journal of Radiation Oncology Biology Physics*, 2009.
- [66] P. Ghadjar, C. Fiorino, P. Munck af Rosenschöld, M. Pinkawa, T. Zilli, and U. A. van der Heide, “ESTRO ACROP consensus guideline on the use of image guided radiation therapy for localized prostate cancer,” *Radiotherapy and Oncology*, 2019.
- [67] D. R. Simpson, J. D. Lawson, S. K. Nath, B. S. Rose, A. J. Mundt, and L. K. Mell, “A survey of image-guided radiation therapy use in the United States,” *Cancer*, 2010.
- [68] M. Li, N.-S. Hegemann, F. Manapov, A. Kolberg, P. D. Thum, U. Ganswindt, C. Belka, and H. Ballhausen, “Prefraction displacement and intrafraction drift of the prostate due to perineal ultrasound probe pressure,” *Strahlentherapie und Onkologie*, 2017.
- [69] P. A. Kupelian, K. M. Langen, O. A. Zeidan, S. L. Meeks, T. R. Willoughby, T. H. Wagner, S. Jeswani, K. J. Ruchala, J. Haimerl, and G. H. Olivera, “Daily variations in delivered doses in patients treated with radiotherapy for localized prostate cancer,” *International Journal of Radiation Oncology Biology Physics*, 2006.
- [70] S. S. Park, D. Yan, S. McGrath, J. T. Dilworth, J. Liang, H. Ye, D. J. Krauss, A. A. Martinez, and L. L. Kestin, “Adaptive image-guided radiotherapy (IGRT) eliminates the risk of biochemical failure caused by the bias of rectal distension in prostate cancer treatment planning: Clinical evidence,” *International Journal of Radiation Oncology Biology Physics*, vol. 83, pp. 947–952, jul 2012.
- [71] M. J. Zelefsky, M. Kollmeier, B. Cox, A. Fidaleo, D. Sperling, X. Pei, B. Carver, J. Coleman, M. Lovelock, and M. Hunt, “Improved clinical outcomes with high-dose image guided radiotherapy compared with non-IGRT for the treatment of clinically localized prostate cancer,” *International Journal of Radiation Oncology Biology Physics*, vol. 84, pp. 125–129, sep 2012.

- [72] M. Valeriani, S. Bracci, M. F. Osti, T. Falco, L. Agolli, V. De Sanctis, and R. M. Enrici, “Intermediate-risk prostate cancer patients treated with androgen deprivation therapy and a hypofractionated radiation regimen with or without image guided radiotherapy,” *Radiation Oncology*, 2013.
- [73] J. Scaife, K. Harrison, M. Romanchikova, A. Parker, M. Sutcliffe, S. Bond, S. Thomas, S. Freeman, R. Jena, A. Bates, and N. Burnet, “Random variation in rectal position during radiotherapy for prostate cancer is two to three times greater than that predicted from prostate motion.,” *The British journal of radiology*, vol. 87, p. 20140343, oct 2014.
- [74] H. M. Patrick and J. Kildea, “Technical note: rtdsm-An open-source software for radiotherapy dose-surface map generation and analysis.,” *Medical physics*, vol. 49, pp. 7327–7335, nov 2022.

Chapter 2

Background

2.1 A Brief History of Dose-Outcome Research

A radiotherapy dose-outcome relationship characterizes the probability of a specific treatment outcome as a function of radiation dose. Dose-outcome relationships are the foundation on which radiotherapy treatments are based, dictating the radiobiological constraints that radiation oncologists must work within to maximize TCP and minimize NTCP. These constraints are usually implemented in clinical practice by way of dosimetric objectives and constraints: specific minimum prescription doses to the target and maximum dose thresholds for the OARs below which the risk of serious toxicity is rare (e.g. 5% over 5 years).

2.1.1 Early Era

During the infancy of radiotherapy, treatments were limited to lower energy (50-200 kV) sources of radiation that deposit most of their dose in the skin. As such, the main toxicities of concern were skin irritations or burns, and the earliest radiation dose limits were defined based on the exposures required to produce a skin rash (erythema) [1]. Further radiation toxicities for other organs were determined throughout the 1910s and 1920s with the discovery and commercialization of radioactive substances like radium, but skin toxicity remained the primary concern of external beam x-ray treatments until the 1940s.

2.1.2 Megavoltage Era

With the 1940s came the development of high-energy megavoltage (MV) treatment beams, which are more penetrating, provide a skin-sparing effect, and deposit most dose a few centimeters into the body. The transition to MV EBRT treatments was all but secured with the invention of linacs and cobalt teletherapy units in the 1950s [2], allowing for the treatment of more deep-seated cancers with fewer skin reactions. However, the treatments of this era were characterized by generous treatment beam margins to account for treatment setup uncertainties as well as inter- and intra-fraction motion of the target and OARs, often resulting in whole or partial volumes of internal organs receiving doses at or near the prescription dose [3]. While internal organ toxicities had become commonplace, a comprehensive attempt at their characterization did not occur until 1968 with the publishing of Rubin and Casarett's text *Clinical Radiation Pathology* [4]. A radiotherapist and radiopathologist respectively, Rubin and Casarett collaborated to fully describe the pathology and radiobiology of radiation toxicities for 20 organ groups, recommending best treatment practices for each. They pioneered the concept of tolerance dose—the dose at which a given toxicity has a probability X to occur in time period Y —that is still used today and was the dominant method to express toxicity in the 1970s and 1980s. However, it is worth noting that Rubin and Casarett's recommendations were derived based on whole-organ exposures, which became invalid with the next evolution of radiotherapy.

2.1.3 Conformal Era

The invention of the CT scanner in 1971 [5] represented a landmark shift for the field of radiotherapy. Tumours and internal organs could now be properly visualized in CT images that could be used to guide treatment plan design in burgeoning early computerized treatment planning systems. The computer revolution itself was allowing for better quantification of dose to internal organs, and advancing computer power was leading to the development of the 3-dimensional conformal radiotherapy (3D-CRT) modality. 3D-CRT uses CT images to compute views of the target from various angles around the patient's body and to conform the radiation beams to the shape of the target in each view using

MLCs. A significant development that aided 3D-CRT planning was the invention of the dose-volume histogram (DVH) by treatment planning system developers Goitein and Verhey in 1979 [6]. DVHs are 2D representations of the 3D dose deposition within a given volume of interest. They are created by counting the number of voxels in a given contoured object receiving a radiation dose of value X or greater (Fig 2.1). This allows for simple, straightforward visualization of the dose coverage of an OAR or target as well as quantification of the dose a fractional volume of a structure is exposed to, making DVHs a mainstay for treatment plan evaluation and optimization. DVH information is commonly reported in either “ V_{doseGy} ” or “ $D_{vol\%}$ ” form, describing either:

- The percent volume of the structure receiving X Gy or more ($V_{dose Gy}$), or
- The minimum radiation dose (in Gy) delivered to most exposed $Y\%$ of the structure ($D_{vol \%}$).

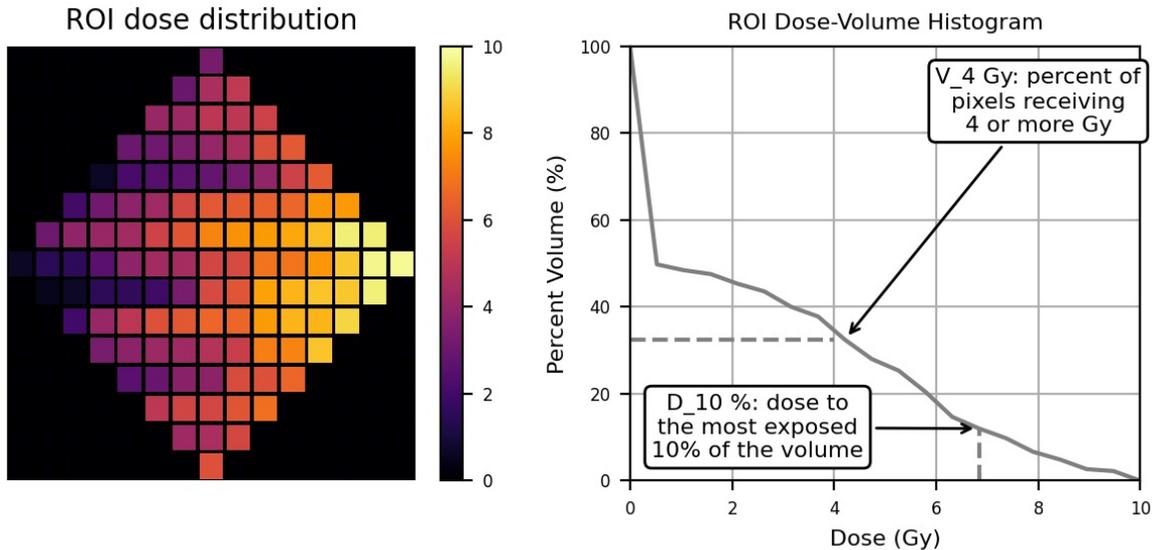


Figure 2.1: Visual explanation of the DVH concept. A diamond-shaped structure contains a distribution of dose voxels (left), which can be used to obtain a DVH (right) by computing a cumulative histogram of the voxel values contained within it. Examples of the V_{doseGy} and $D_{vol\%}$ are also shown.

With the advent 3D-CRT came a new issue for clinicians: now that treatment beams

were more conformal to target volumes, neighbouring OARs were being partially irradiated much more frequently, something for which they had no guiding tolerance doses. Paired with the shift from 2D to 3D planning, radiation oncologists were left to guess what constituted reasonable OAR doses based on their past experience. Recognizing the need for new constraints in this landscape, Emami assembled a team of experts and released his landmark paper *Tolerance of Normal Tissue to Therapeutic Irradiation* in 1991 [7]. Using their shared experience and what little quantitative data existed, the Emami paper established a set of recommended dosimetric constraints for partial volumes of all major OARs. In collaboration, Burman *et al* also fitted a Lyman NTCP model to the Emami constraints [8], allowing them to be extrapolated to any fractional volume of an organ. These guidelines, paired with the development of inverse planning algorithms for treatment planning systems, formed the basis of the next two decades of radiotherapy treatment planning.

Inspired by Emami, the 1990s and 2000s saw the publication of numerous clinical studies on dose-outcome effects for many treatment sites. Updated dosimetric constraints based on DVH metrics were being suggested for organs based on data analysis emerging from NTCP models, rather than experiential consensus. IGRT was also being utilized more and more in the early 2000s to reduce setup uncertainty and to better accommodate the potential effects of inter-fraction motion. By 2007 the field recognized that an update to the Emami constraints was necessary, resulting in the forming of the QUANTEC (QUantitative Analyses of Normal Tissue Effects in the Clinic) initiative. QUANTEC summarized the dose-outcome knowledge base into a set of practical guiding dosimetric constraints in a special 2010 edition of the *Red Journal*, serving as the foundation for the last decade of treatment planning efforts [3].

2.1.4 Modern Era

Dose-outcome research has continued to remain popular since the publishing of QUANTEC. Recently the AAPM has released HyTEC (Hypofractionated Treatment Effects in the Clinic) [9], a sister report to QUANTEC covering constraints for the ultra-hypofractionated treatment modality, Stereotactic Body Radiation Therapy (SBRT) due to growing evidence of different biological response to hypofractionated doses.

2.2 Limitations of Existing Dose-Outcome Studies

It is important to note that while individual studies and larger initiatives like QUANTEC and HyTEC provide valuable guidance on OAR sparing practices, the dosimetric constraints they recommend are limited by the data used to derive them. For example, the literature base used to determine the QUANTEC rectum constraints featured many contrasting claims on what fractional volume of the rectum was able to tolerate doses between 30 and 60 Gy (Fig 2.2). Factors like sample size, treatment protocol, and outcomes metrics are among the many sources of uncertainty that can complicate the characterization of the true dose-outcome response of OARs and were the subjects of several “Vision Papers” in the QUANTEC report [10–12]. However, the focus of this thesis is on two particular limitations regarding the use of treatment planning DVH data in conventional dose-outcome studies: (1) the planned dose does not always represent the dose that is actually delivered, and (2) DVHs do not provide spatial information.

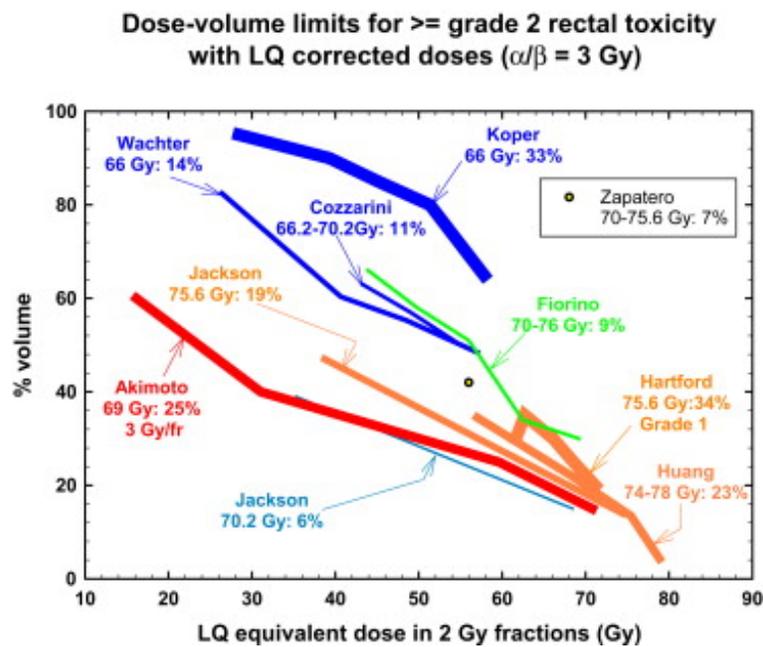


Figure 2.2: Dose-volume limits proposed in the literature and used as a foundation for QUANTEC rectum dosimetric constraints. All values have been converted to be for standard fractionation (2 Gy/fraction). Thicker lines indicate that higher toxicity rates were observed. Figure reproduced from [13].

2.2.1 Planned versus Delivered Dose

Patient delivered doses can differ from planned doses due to many factors including inter- or intra-fraction motion, tumour shrinkage over the course of treatment, weight fluctuations during treatment, and daily set-up uncertainties. However, nearly all dose-outcome studies use planned doses, rather than delivered doses, to evaluate the dose-outcome effect of OARs. Delivered dose calculations are complex and require a robust workflow that was not widely accessible until the 2010s, and even with a workflow in place, the extra effort required can disincentivize researchers from using delivered doses over readily-available planned doses. However, the dose differences involved are often not negligible and many studies of prostate radiotherapy have demonstrated significant disagreement between planned and delivered rectum doses. For example, Pearson *et al*, Godley *et al*, and Scaife *et al* all reported substantial increases in the proportion of the rectum receiving 70 Gy ($V_{70\text{ Gy}}$ increase 15-60%) during treatment delivery [14–16]. Statistically significant differences between planned and delivered mean rectum doses have also been reported by several authors [15, 17, 18], even in cases where endorectal balloons were used to limit the effects of rectal filling variations [19]. Studies conducted by Lebesque *et al* and Wen *et al* have suggested that these dosimetric differences can increase the rectal NTCP by up to 14% in the pre-IGRT era and 3% since [20, 21]. While one of the QUANTEC vision papers heavily advocated for a shift towards the use of delivered doses and outlined the efforts required [10], adoption has been slow. A small collection of studies using delivered rectum doses have been published recently, and together provide evidence that delivered doses may offer a non-negligible improvement in dose-outcome prediction over planned doses [22–24].

2.2.2 Spatial Dose Effects

While DVHs are used as a routine tool in radiotherapy practice, they are not without criticism. Most commentary on the limitations of DVHs revolve around the lack of spatial information and the convention of reporting DVH metrics in terms of relative volume, rather than absolute. Excluding spatial (or even absolute volume) information makes it

impossible to determine which region or what size region a $V_{X Gy}(\%)$ metric is referring to for a given organ. This can be especially troublesome for organs that exhibit large volume fluctuations like the rectum, which motivated the decision to provide HyTEC-recommended constraints for the rectum and bladder in units of absolute volume [25]. The collapsing of 3D spatial dose distributions into a 1D function also gives rise to another important limitation: the same DVH can be derived from hundreds of unique 3D dose distributions, requiring the assumption that all dose distributions that yield the same DVH yield the same dose response (Fig 2.3). This ties in with another key assumption: by discarding spatial information we assume that all organs are homogeneously radiosensitive. This assumption has arguably been utilized since the work of Rubin and Casarett [4], but numerous studies have demonstrated otherwise. Salivary gland, bladder, lung, and heart toxicities [26–29] have all been associated with dose to critical substructures or more actively functioning areas, whereas a dependence on spatial dose gradients has been reported for spinal cord toxicities [30]. Interest in spatial dose effects has increased in recent years, with discussions on available metrics and best practices ongoing [31].

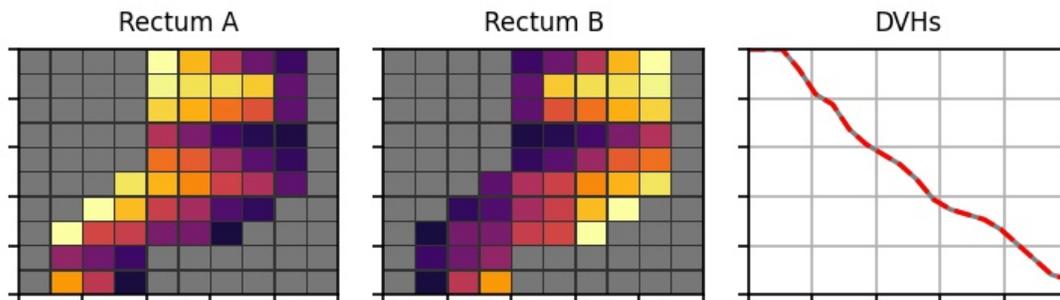


Figure 2.3: Despite rectums A and B receiving very different dose distributions, their DVHs are identical after the spatial information is removed.

Several alternatives to traditional DVH metrics have been suggested since early criticisms began at the beginning of the millennia [32]. The earliest – and simplest – alternative is the creation of DVHs for sub-volumes of a structure. Common sub-volume definitions for the rectum include separation into several subregions [33, 34], as well as slice-wise sub-volumes (commonly referred to as dose-line histograms) [35–37], both of

which have been demonstrated to produce improved dose-outcome models over those that use whole structure DVHs. More recently, full preservation and utilization of 3D dose information for outcomes research has been performed using deformable registration to align dose distributions to a single reference anatomy. Using this approach, Acosta *et al* identified a subregion in the anterior rectal wall which received 6 Gy more in patients who experienced rectal bleeding compared to than those who did not [38]. Similarly, Drean *et al* conducted 3D sub-volume analysis and identified that dose to the inferior-anterior quarter of the rectum was highly predictive of bleeding [33]. However, the most common alternative spatial dose representation in use for the rectum is 2D dose-surface mapping, which has been shown to yield toxicity associations not detectable with DVHs [39] and increased AUC (Area Under the receiver operating characteristic Curve) values relative to models that use only DVH information [40–42].

2.3 Dose Accumulation

Dose accumulation refers to the process of calculating the total dose delivered to an organ or body across multiple time points and requires three key elements [43]:

1. The *anatomical information* of the patient for each radiotherapy fraction;
2. A *dose calculation method* to determine the dose delivered at each fraction;
3. A *dose summation strategy* that accounts for the anatomical variations between fractions.

2.3.1 Anatomical Information

In order to accurately calculate and accumulate daily organ doses, the localization of a patient’s organs during each treatment fraction must be known. This information can be obtained from volumetric images obtained during radiation delivery, such as those acquired for daily IGRT. The imaging modality used is largely dependent on the imagers that are integrated into treatment suites, with cone-beam CT (CBCT) imaging being the most prevalent.

CT images are computational reconstructions of a 3D object based on many 2D projection images of the object acquired from many angles. Due to the projection images being transmission x-ray radiographs, the reconstructed CT voxels provide information on the radiation attenuation coefficients, μ , of the object for the x-ray beam used to acquire it [44]. This information can be used by dose calculation engines by relating the voxel intensities, given in Hounsfield Units (HU)

$$HU = 1000 \times \frac{\mu_{\text{voxel}} - \mu_{\text{water}}}{\mu_{\text{water}} - \mu_{\text{air}}} \quad (2.1)$$

to dosimetrically relevant material properties. However, it is important to note that HU values are dependent on both x-ray beam energy and beam shape, meaning different imagers may produce different readings for the same material.

CBCT imagers are primarily differentiated from traditional diagnostic, fan-beam CT (FBCT) imagers based on the shape of their radiation beam. FBCT uses a thin, collimated fan beam to acquire individual slices of a CT image, whereas CBCT imagers use a conical radiation beam and large detector array to acquire multiple CT slices in a single rotation of the gantry (Fig 2.4). While image resolution is slightly better with CBCT systems due to the use of flat panel detectors, overall image quality for CBCT is reduced. This is because the wider beam leads to significantly higher levels of scattered radiation reaching the flat-panel detector, decreasing the signal-to-noise ratio and image quality [45]. Additionally, the reconstruction algorithm used by most clinical CBCT scanners is less reliable at larger angles from the central plane, meaning the most superior and inferior slices of CBCT images tend to have less accurate reconstructions than the central-most [46].

2.3.2 Dose Calculation Method

To provide accurate dose distributions within 1-2% [47], a dose calculation method must be able to accurately model radiation interaction and energy deposition events within the many tissues of a patient. Clinically, this must also be done in a timely manner, meaning a balance must be struck between accuracy and speed. For example, the gold-standard, Monte Carlo simulation, explicitly models the interactions of particles as they travel through the patient

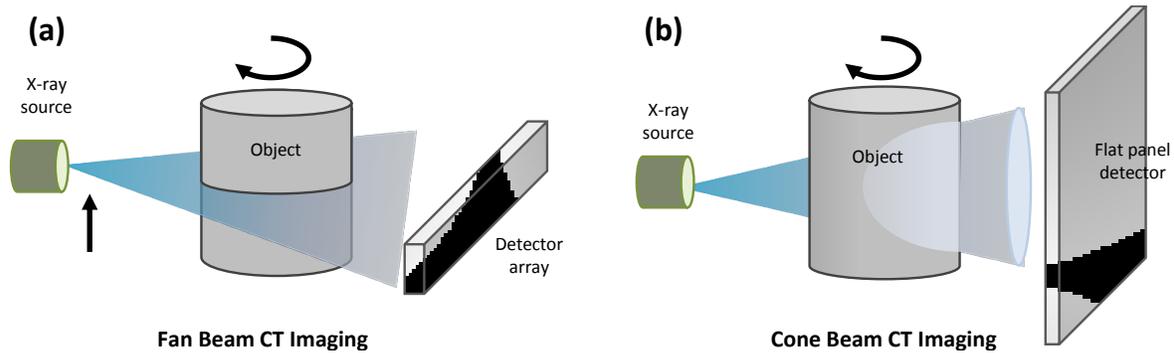


Figure 2.4: Beam and detector configurations for (a) fan beam and (b) cone beam CT imaging.

but it requires either substantial time or computing power to simulate the billions of particles required [48]. For this reason, most commercial dose calculation software applications use a Convolution-Superposition algorithm to break down the dose deposition process into two components to speed up the computation time while still preserving calculation accuracy [49]. The algorithm replicates the two-step process in which indirectly ionizing radiation deposits dose, along with a heterogeneity modifier as follows:

1. **TERMA component:** To start, the algorithm computes the distribution of the photons from the primary radiation beam as it travels through the patient, up until they interact with the tissues. This can be easily calculated by multiplying the energy fluence map of the primary photon beam by the mass attenuation coefficients of each point in the patient to yield primary TERMA (Total Energy Released per unit MA_{ss}).
2. **Kernel component:** From these interaction points, the algorithm then determines the distributions of the secondary photons and charged particles released at these points and how they deposit energy around them. Kernels are complex and depend on many factors and are therefore usually pre-calculated through Monte Carlo modeling and stored within the algorithm for easy lookup. These kernels are convolved with the TERMA component to produce the dose distribution.
3. **Heterogeneity modifier:** Because the transport of secondary photons and charged particles depends on the material composition of the irradiated object, the spatial

composition of the object needs to be accounted for. Rather than pre-calculate millions of unique kernels for every possible situation, both TERMA and kernel components receive a modifier that scales their distributions based on the radiologic densities within the heterogeneous patient before they are convolved together. This is known as the “superposition” aspect of the algorithm, providing suitable heterogeneity corrections while keeping the calculation manageable.

The heterogeneity modifier term is an essential component of accurate dose calculations, as its exclusion can lead to errors of up to 30% of the prescribed treatment dose [50]. These modifiers are based on tissue electron densities (EDs) [51], which can be linked to CT HU values through the use of a HU-ED curve determined by imaging calibration phantoms containing material inserts representing the many tissue types present in the human body [52]. However, it is important to note that the default HU-ED curves saved and maintained in dose calculation engines are for FBCT imagers, meaning either additional CBCT-specific HU-ED must be obtained [53, 54] or image post-processing workflows must be established [55, 56] in order to ensure accurate dose calculations when using CBCT images. For the work described in this thesis, we tested the performance of both approaches before settling on using CBCT-specific HU-ED curves for our dose accumulation workflow.

2.3.3 Dose Summation Strategies

Perhaps the most integral component of the dose accumulation process is the dose summation strategy. Simple addition of daily dose distributions neglects inter-fraction positional changes of organs and can yield different accumulated dose distributions than what was delivered in practice. While addition or averaging of daily DVHs or DVH metrics can be used to estimate total delivered dose [18, 57–59], it similarly fails to properly account for inter-fractional positional changes. For this reason, a reliable means to spatially align daily dose distributions to the same reference anatomy is a critical element of a quality dose accumulation strategy.

Deformable image registration (DIR) is the most commonly employed method to align daily delivered doses for dose accumulation. Using a *reference* anatomical image, a *moving* image of a different anatomical arrangement is iteratively deformed, evaluated, and updated

by a *deformable registration algorithm* to match the anatomy of the reference (Fig 2.5). These deformations are then applied to the moving image’s dose distribution, allowing it to be accumulated with other dose distributions aligned to the reference anatomy. While different DIR algorithms may be constructed in different ways, they all require three key elements:

1. A **Transformation Model** that defines the types of deformations that can be applied to the moving image. These can be either *parametric models*, which only calculate explicit deformations for a grid of control points and use splines to interpolate deformations for the remaining voxels [60, 61], or *non-parametric models* that explicitly calculate deformations for each voxel according to a physics force model (such as Demons force, elastic force, or fluid flow) [62–64].
2. A **Similarity Metric** that measures the similarity between the reference and deformed version of the moving image. These metrics can be either *intensity-based*, which compare voxel intensities, *feature-based*, which compare anatomic features like tissue boundaries or landmark points, or a *hybrid* of the two [65].
3. An **Optimization Strategy** to determine the best way to update the transformation model after each iteration to maximize the similarity metric. Most commercial algorithms use gradient-based optimization approaches, though some also employ more stochastic methods [66].

As the accuracy of DIR-based dose accumulation is heavily dependent on deformation quality, it is strongly recommended that all DIR algorithms be validated using quantitative performance metrics for each treatment site of interest. One of the simplest approaches is to quantify the accuracy at which an algorithm can deform organ contours from one image to another using Dice Similarity Coefficients (DSCs). DSCs provide a measure of the volumetric overlap of two contours (i.e. reference and deformed) between 0 and 1, with scores of >0.8 recognized as acceptable performance [66].

The performance of various deformable image registration algorithms for prostate radiotherapy applications has been the subject of numerous studies, covering CT-CT [15, 67–71], CT-CBCT [72–76], and MR-CT registrations [77, 78]. Over the years,

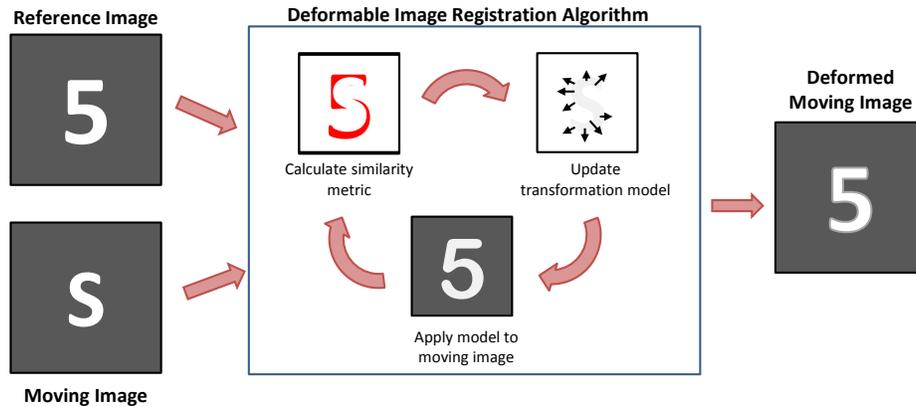


Figure 2.5: Overview of the process by which a deformable image registration algorithm deforms a moving image into the reference image.

most publications have focused on the performance of intensity-based algorithms, particularly those that use spline or Demons transformation models. Mean rectum DSC scores achieved by the algorithms are typically within the range of 0.60-0.77 [67, 68, 70, 74], though mean DSC scores as low as 0.37 have been reported by some authors for lower performing Demons algorithms [69, 73]. As noted by Foskey *et al*, many intensity-based algorithms struggle to handle the appearance and disappearance of rectal gas between treatment fractions, reducing registration accuracy [79]. While some authors have demonstrated that pre-processing images to remove gaseous pockets or assigning bulk HU values to entire organs can improve DSC scores (Godley *et al*: 0.74 to 0.93 and Gao *et al*: 0.51 to 0.71 [15, 71]), these approaches are generally undesirable as they often require additional labor from clinical staff to fully delineate the rectum on each daily image. More recently, hybrid algorithms that include geometric similarity metrics have begun to show promise, outperforming intensity-based algorithms in direct comparisons [73, 77, 80]. However, it is important to note that high DSC scores do not necessarily mean registrations are accurate. Both Thornqvist *et al* and Thor *et al* noted a significant proportion of their rectum registrations would be deemed unacceptable from a manual contouring standpoint, despite achieving DSCs between 0.71 and 0.78 [68, 72].

Beyond *in silico* performance metrics, which tend to be more global in nature, dose deformation and accumulation strategies can also be evaluated with computational and

physical phantoms. Computational phantoms typically take the form of before and after image pairs where the transformation model between them is explicitly known and can be useful for evaluating baseline algorithm performance. Physical phantoms on the other hand, can allow for a full end-to-end evaluation of the dose accumulation process that encompass imager noise, dose calculation uncertainties, and deformation accuracy. For example, deformable gel dosimeters are excellent end-to-end dose accumulation phantoms, as they allow for direct comparisons of delivered and calculated total doses [81–83]. More recently, fully 3D, anthropomorphic pelvis phantoms have been developed to validate MR-linac dose accumulation using embedded point detectors, radiochromic films, or gel dosimeters [84–86]. Testing with such phantoms can reveal inaccuracies in dose accumulations that are not readily apparent through visual inspection of image registrations alone [81].

2.4 Dose-Surface Maps

Dose-surface maps (DSMs) provide 2D representations of 3D surface doses that preserve the relative spatial information of the dose distribution on an organ’s surface. For this reason, they are particularly useful and popular for examining dose to hollow organs like the rectum and bladder [22, 29, 41, 87, 88], although studies have been conducted for the vagina, heart, duodenum, and esophagus [28, 89–91].

First described by Meijer *et al* [92], the DSMs reported in the literature have been calculated using contour and dose information from treatment planning systems using custom, in-house software. The general calculation process begins by defining multiple slices along the length of a structure’s surface (Fig 2.6), then sampling dose at many points around the circumference of each slice. A decision is then made on where the (approximately) cylindrical surface should be cut open, and the mesh of dose points unfurled and flattened into a 2D grid. In some cases, authors may post-process. For example, authors who do not normalize the number of dose points sampled per slice during DSM calculation may opt to apply this normalization after the fact [93]. Other authors may also chose to interpolate DSMs to a common $M \times N$ matrix size for dataset

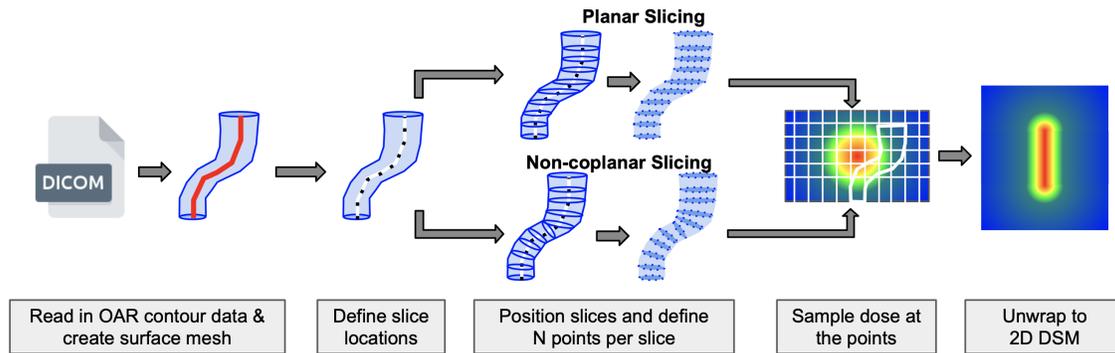


Figure 2.6: Flowchart of the basic DSM calculation process. Using a contour from a treatment plan, slices of the structure are defined along which dose will be sampled from the plan’s dose distribution. The cylindrical set of dose points will then be cut open and unfurled to produce a DSM.

consistency [41].

It is important to note that while the general DSM calculation approach largely remains consistent between research groups, implementations can and do vary significantly between research groups. The most noteworthy way by which methods diverge is in how DSM slices are oriented in 3D space. The earliest reports of rectum DSMs used the rectum’s central axis path to define slices orthogonal to it to account for its curvature in space (non-coplanar slicing) [39, 92, 94]. However, this approach was quite complex, requiring corrections or special path trajectories to handle overlapping slices [90, 95]. This eventually led many authors to simplify the calculation by using the parallel slices of the CT on which the rectum contour was drawn [41, 96, 97], and quickly became the dominant approach (planar slicing). DSM calculation approaches can also vary in several other ways. For instance, some authors use space DSM slices using a set distance (e.g. 1.5 or 3 mm) while others allow the distance to vary on a patient-by-patient basis such that each DSM comprises the same total number of slices. Sampling resolutions are highly variable as well, with some authors using coarse resolutions of 21×21 pixels [98] compared to others who use finer resolutions of 200×200 pixels or more [40, 93]. Even the location in which the DSM is cut open for unfurling can differ: some groups merely use the posterior-most point of each slice [41, 98, 99] while others use the points directly posterior to the centroids of each slice [22, 39, 100]. This diversity

of methods can be largely attributed to the in-house nature of most DSM calculation code bases, requiring each new research group to develop their own implementation.

While calculation methods can and do vary, the primary objective of most DSM studies is the same: identify spatial factors that are predictive of OAR toxicities. These factors are determined by comparing DSMs between patients with and without a given toxicity and characterizing how their spatial dose distributions differ. Two main categories of spatial factors have emerged over the years and enjoy similar levels of popularity in rectal DSM studies:

1. *DSM features* describe the size and shape of isodose clusters within a DSM, such as their area, center of mass, and lateral or longitudinal extent (Fig 2.7) Features that describe the size, shape, and angle of ellipses fitted to these clusters have also been commonly investigated after initial popularization by Buetter *et al* [41, 101]. Several features have been associated with the development of various rectal toxicities after standard fractionated radiotherapy. For instance, rectal bleeding has been associated with lateral extents of 40-60 Gy isodose clusters [22, 41, 93] and the area of 51 Gy clusters [41, 100], and diarrhea has been associated with area, lateral and longitudinal extents, and eccentricity of 23 Gy clusters [41, 100].
2. *DSM subregions* are areas within a given organ's DSM that receive significantly different dose between patients with and without a radiation toxicity, which may indicate the region is important for normal organ function. They are identified by comparing pixel values between DSMs from patient cohorts with and without the toxicity, usually with some kind of multiple comparisons correction to account for false positives [40, 87, 93, 102]. With a few exceptions, most identified subregions are located in the posterior rectal wall and suggest increased posterior wall dose may be associated with rectal bleeding [23, 93, 100, 103], proctitis [23, 100], and loose frequent stools [100, 103].

2.4.1 Dose-Surface Map Accumulation

As described above, dose-surface maps convert dose distributions on a 3D surface to a regularized, 2D map grid. Much like deformable image registration, this process takes one

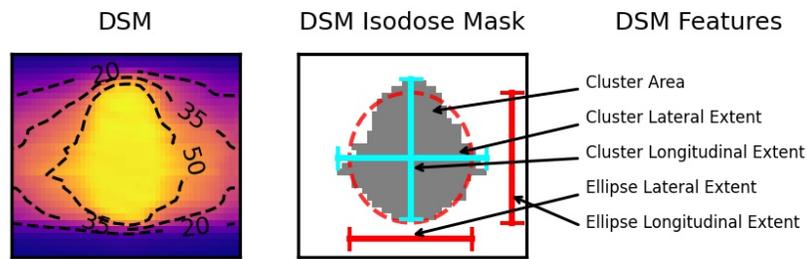


Figure 2.7: DSM features are derived from masks of isodose regions within a DSM. Several popular features in the literature are shown.

spatial representation of dose, applies a transformation function to calculate the DSM, and outputs a different spatial representation. Because of this, if one can safely assume that the transformation function can reliably deform the dose distribution of any rectum to the same standard 2D grid, dose accumulation should be possible by summing DSMs from each treatment fraction (Fig 2.8.)

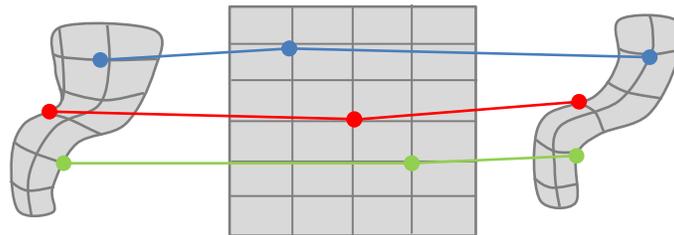


Figure 2.8: DSM-based dose accumulation is hypothetically feasible provided the dose-surface map calculation process can accurately map the dose to a given point in the rectum to the same location within the DSM for each fraction.

The first published proposal of this approach was made by Murray *et al* in a 2014 conference abstract [98]. Using the DSM calculation approach developed by their collaborators Buettner *et al* [41], Murray *et al* calculated and summed 21 x 21 pixel planar DSMs for each fraction of a prostate radiotherapy course to calculate accumulated DSMs for three patients. Unfortunately, Murray did not publish further on this idea. Their idea was however, continued by another student from the Royal Marsden Hospital, d’Aquino,

who used Murray’s DSM accumulation methodology to evaluate planned and delivered dose-outcome relationships for the rectum, ultimately finding them to be similar when strict IGRT and enema regimens are employed [104].

Separately, an interdisciplinary team from Cambridge led by Burnet established the “VoxTox Initiative” to examine rectal dose-outcome relationships using delivered doses calculated by way of DSM accumulation. Initially, the group used a similar DSM calculation and summation approach to Murray, calculating and summing $21 \times N$ pixel planar DSMs (N being the number of slices the rectum spanned) by aligning them at the rectum inferior border [16, 22]. Eventually they moved to a biomechanical modeling-based calculation strategy. Using the biomechanical properties of rectal tissue, the team would deform a cylindrical mesh to the shape of a given rectum for dose sampling. The mesh would then be returned to its rest state and unwrapped to produce a DSM. Shelley *et al* would go on to publish two papers on dose-outcome relationships using both planar and biomechanical DSMs, ultimately reaching a similar conclusion to d’Aquino that delivered doses were slightly more predictive of rectal toxicities compared to planned doses [22, 23].

Another DSM accumulation study comparing matched patients with and without grade ≥ 2 proctitis was conducted by Casares-Magaz *et al* [24]. Unlike Shelley *et al* and Murray *et al*, Casares-Magaz *et al* only used a subset of daily CBCTs representing 30% of all fractions to calculate accumulated planar DSMs. They noted patients with proctitis received higher dose to the inferior anterior rectum than matched control patients, who themselves received lower accumulated doses than planned. They also reported dose metrics from accumulated DSMs were more predictive of proctitis than planned DSMs, similar to previous reports.

While DSM accumulation has been explored in some detail already, it should be noted that some unexplored territories remain. Most notably, previous authors did not experimentally validate their DSM accumulation methods with physical measurements, meaning the potential limitations of DSM accumulation are not known. Studies have also focused thus far on standardly-fractionated treatments (2 Gy/fraction) which are delivered over the course of 30 or more fractions, rather than the short-course ultra-hypofractionated treatments that have become commonplace since the 2020 pandemic. As speculated by Shelley [105], longer courses of treatment may allow for random variations in daily doses to

average out over the full treatment course, yielding similar planned and delivered doses, whereas shorter courses of treatment may not. Investigation into this phenomenon is worthwhile to determine if treatment course length plays a role in the magnitude of delivered dose variation from the plan, especially given the recent rise of prostate SBRT [106]. Finally, the different DSM calculation methods used by these authors gives rise to the question of the equivalence of accumulated DSMs calculated in different manners. Investigation of this matter is important to determine the universal applicability of DSM-derived findings and whether or not a universal standard needs to be adopted.

These limitations served as the underlying motivations for the objectives of this thesis and are addressed in the following chapters. The equivalence of DSMs calculated in different manners is investigated in [Chapter 4](#) (objective 2), the validation of DSM accumulation in [Chapter 5](#) (objective 3), and equivalence of planned and delivered doses during short-course prostate SBRT in [Chapter 6](#) in parallel with objective 4.

Bibliography

- [1] E. H. Quimby and G. T. Pack, “The Skin Erythema for Combinations of Gamma and Roentgen Rays,” *Radiology*, 1929.
- [2] D. I. Thwaites and J. B. Tuohy, “Back to the future: The history and development of the clinical linear accelerator,” 2006.
- [3] S. M. Bentzen, L. S. Constine, J. O. Deasy, A. Eisbruch, A. Jackson, L. B. Marks, R. K. Ten Haken, and E. D. Yorke, “Quantitative Analyses of Normal Tissue Effects in the Clinic (QUANTEC): An Introduction to the Scientific Issues,” *International Journal of Radiation Oncology Biology Physics*, 2010.
- [4] J. R. Andrews, P. Rubin, and G. W. Casarett, “Clinical Radiation Pathology,” *Radiation Research*, 1968.
- [5] G.N. Hounsfield, “Computerised Transverse Axial Scanning (tomography) Part 1. Description of system,” *British journal of Radiology*, 1973.

- [6] B. A. Fraass, "The development of conformal radiation therapy," *Medical Physics*, 1995.
- [7] B. Emami, J. Lyman, A. Brown, L. Cola, M. Goitein, J. E. Munzenrider, B. Shank, L. J. Solin, and M. Wesson, "Tolerance of normal tissue to therapeutic irradiation," *International Journal of Radiation Oncology, Biology, Physics*, vol. 21, pp. 109–122, may 1991.
- [8] C. Burman, G. J. Kutcher, B. Emami, and M. Goitein, "Fitting of normal tissue tolerance data to an analytic function," *International Journal of Radiation Oncology, Biology, Physics*, 1991.
- [9] J. Grimm, L. B. Marks, A. Jackson, B. D. Kavanagh, J. Xue, and E. Yorke, "High Dose per Fraction, Hypofractionated Treatment Effects in the Clinic (HyTEC): An Overview," *International Journal of Radiation Oncology Biology Physics*, 2021.
- [10] D. A. Jaffray, P. E. Lindsay, K. K. Brock, J. O. Deasy, and W. A. Tomé, "Accurate Accumulation of Dose for Improved Understanding of Radiation Effects in Normal Tissue," *International Journal of Radiation Oncology Biology Physics*, vol. 76, pp. S135–9, mar 2010.
- [11] J. O. Deasy, S. M. Bentzen, A. Jackson, R. K. Ten Haken, E. D. Yorke, L. S. Constine, A. Sharma, and L. B. Marks, "Improving Normal Tissue Complication Probability Models: The Need to Adopt a "Data-Pooling" Culture," *International Journal of Radiation Oncology Biology Physics*, 2010.
- [12] A. Jackson, L. B. Marks, S. M. Bentzen, A. Eisbruch, E. D. Yorke, R. K. Ten Haken, L. S. Constine, and J. O. Deasy, "The Lessons of QUANTEC: Recommendations for Reporting and Gathering Data on Dose-Volume Dependencies of Treatment Outcome," *International Journal of Radiation Oncology Biology Physics*, 2010.
- [13] J. M. Michalski, H. Gay, A. Jackson, S. L. Tucker, and J. O. Deasy, "Radiation dose-volume effects in radiation-induced rectal injury.," *International journal of radiation oncology, biology, physics*, vol. 76, pp. S123–9, mar 2010.

- [14] D. Pearson, S. K. Gill, N. Campbell, and K. Reddy, “Dosimetric and volumetric changes in the rectum and bladder in patients receiving CBCT-guided prostate IMRT: Analysis based on daily CBCT dose calculation,” *Journal of Applied Clinical Medical Physics*, vol. 17, no. 6, pp. 107–117, 2016.
- [15] A. Godley, E. Ahunbay, C. Peng, and X. Allen Li, “Accumulating daily-varied dose distributions of prostate radiation therapy with soft-tissue-based KV CT guidance,” *Journal of Applied Clinical Medical Physics*, 2012.
- [16] J. E. Scaife, S. J. Thomas, K. Harrison, M. Romanchikova, M. P. F Sutcliffe, J. R. Forman, A. M. Bates, R. Jena, M. A. Parker, and N. G. Burnet, “Accumulated dose to the rectum, measured using dose-volume histograms and dose-surface maps, is different from planned dose in all patients treated with radiotherapy for prostate cancer,” *British Journal of Radiology*, 2015.
- [17] W. Y. Song, E. Wong, G. S. Bauman, J. J. Battista, and J. Van Dyk, “Dosimetric evaluation of daily rigid and nonrigid geometric correction strategies during on-line image-guided radiation therapy (IGRT) of prostate cancer,” *Medical Physics*, 2007.
- [18] J. A. Hatton, P. B. Greer, C. Tang, P. Wright, A. Capp, S. Gupta, J. Parker, C. Wratten, and J. W. Denham, “Does the planning dose-volume histogram represent treatment doses in image-guided prostate radiation therapy? Assessment with cone-beam computerised tomography scans,” *Radiotherapy and Oncology*, 2011.
- [19] J. Yu, N. Hardcastle, K. Jeong, E. T. Bender, M. A. Ritter, and W. A. Tomé, “On voxel-by-voxel accumulated dose for prostate radiation therapy using deformable image registration,” *Technology in Cancer Research and Treatment*, 2014.
- [20] J. V. Lebesque, A. M. Bruce, A. P. Guus Kroes, A. Touw, T. Shouman, and M. van Herk, “Variation in volumes, dose-volume histograms, and estimated normal tissue complication probabilities of rectum and bladder during conformal radiotherapy of T3 prostate cancer,” *International Journal of Radiation Oncology, Biology, Physics*, vol. 33, pp. 1109–1119, dec 1995.

- [21] N. Wen, A. Kumarasiri, T. Nurushev, J. Burmeister, L. Xing, D. Liu, C. Glide-Hurst, J. Kim, H. Zhong, B. Movsas, and I. J. Chetty, “An assessment of PTV margin based on actual accumulated dose for prostate cancer radiotherapy,” *Physics in Medicine and Biology*, vol. 58, pp. 7733–7744, nov 2013.
- [22] L. E. A. Shelley, J. E. Scaife, M. Romanchikova, K. Harrison, J. R. Forman, A. M. Bates, D. J. Noble, R. Jena, M. A. Parker, M. P. F. Sutcliffe, S. J. Thomas, and N. G. Burnet, “Delivered dose can be a better predictor of rectal toxicity than planned dose in prostate radiotherapy,” *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*, vol. 123, no. 3, pp. 466–471, 2017.
- [23] L. E. Shelley, M. P. Sutcliffe, S. J. Thomas, D. J. Noble, M. Romanchikova, K. Harrison, A. M. Bates, N. G. Burnet, and R. Jena, “Associations between voxel-level accumulated dose and rectal toxicity in prostate radiotherapy,” *Physics and Imaging in Radiation Oncology*, 2020.
- [24] O. Casares-Magaz, S. Bülow, N. J. Pettersson, V. Moiseenko, J. Pedersen, M. Thor, J. Einck, A. Hopper, R. Knopp, and L. P. Muren, “High accumulated doses to the inferior rectum are associated with late gastro-intestinal toxicity in a case-control study of prostate cancer patients treated with radiotherapy,” 2019.
- [25] B. Wang, D. Q. Wang, M. S. Lin, S. P. Lu, J. Zhang, L. Chen, Q. W. Li, Z. K. Cheng, F. J. Liu, J. Y. Guo, H. Liu, and B. Qiu, “Accumulation of the delivered dose based on cone-beam CT and deformable image registration for non-small cell lung cancer treated with hypofractionated radiotherapy,” *BMC Cancer*, 2020.
- [26] W. Jiang, P. Lakshminarayanan, X. Hui, P. Han, Z. Cheng, M. Bowers, I. Shpitser, S. Siddiqui, R. H. Taylor, H. Quon, and T. McNutt, “Machine Learning Methods Uncover Radiomorphologic Dose Patterns in Salivary Glands that Predict Xerostomia in Patients with Head and Neck Cancer,” *Advances in Radiation Oncology*, 2019.
- [27] P. Van Luijk, S. Pringle, J. O. Deasy, V. V. Moiseenko, H. Faber, A. Hovan, M. Baanstra, H. P. Van Der Laan, R. G. Kierkels, A. Van Der Schaaf, M. J. Witjes,

- J. M. Schippers, S. Brandenburg, J. A. Langendijk, J. Wu, and R. P. Coppes, "Sparing the region of the salivary gland containing stem cells preserves saliva production after radiotherapy for head and neck cancer," *Science Translational Medicine*, 2015.
- [28] A. McWilliam, C. Dootson, L. Graham, K. Banfill, A. Abravan, and M. van Herk, "Dose surface maps of the heart can identify regions associated with worse survival for lung cancer patients treated with radiotherapy," *Physics and Imaging in Radiation Oncology*, 2020.
- [29] F. Palorini, C. Cozzarini, S. Gianolini, A. Botti, V. Carillo, C. Iotti, T. Rancati, R. Valdagni, and C. Fiorino, "First application of a pixel-wise analysis on bladder dose-surface maps in prostate cancer radiotherapy," *Radiotherapy and Oncology*, 2016.
- [30] H. P. Bijl, P. Van Luijk, R. P. Coppes, J. M. Schippers, A. W. Konings, and A. J. Van Der Kogel, "Unexpected changes of rat cervical spinal cord tolerance caused by inhomogeneous dose distributions," *International Journal of Radiation Oncology Biology Physics*, 2003.
- [31] M. A. Ebert, S. Gulliford, O. Acosta, R. De Crevoisier, T. McNutt, W. D. Heemsbergen, M. Witte, G. Palma, T. Rancati, and C. Fiorino, "Spatial descriptions of radiotherapy dose: Normal tissue complication models and statistical associations," 2021.
- [32] V. Moiseenko, J. Van Dyk, J. Battista, and E. Travis, "Limitations in using dose-volume histograms for radiotherapy dose optimization," in *The Use of Computers in Radiation Therapy*, 2000.
- [33] G. Dréan, O. Acosta, J. D. Ospina, A. Fargeas, C. Lafond, G. Corrége, J. L. Lagrange, G. Créhange, A. Simon, P. Haignon, and R. de Crevoisier, "Identification of a rectal subregion highly predictive of rectal bleeding in prostate cancer IMRTRectal subregion involved in bleeding," *Radiotherapy and Oncology*, 2016.
- [34] S. T. Peeters, J. V. Lebesque, W. D. Heemsbergen, W. L. Van Putten, A. Slot, M. F. Dielwart, and P. C. Koper, "Localized volume effects for late rectal and anal toxicity

- after radiotherapy for prostate cancer,” *International Journal of Radiation Oncology Biology Physics*, 2006.
- [35] C. W. Cheng and I. J. Das, “Treatment plan evaluation using dose-volume histogram (DVH) and spatial dose-volume histogram (zDVH),” *International Journal of Radiation Oncology Biology Physics*, 1999.
- [36] E. M. Donovan, N. J. Bleackley, P. M. Evans, S. F. Reise, and J. R. Yarnold, “Dose-position and dose-volume histogram analysis of standard wedged and intensity modulated treatments in breast radiotherapy,” *British Journal of Radiology*, 2002.
- [37] L. J. Hamlett, A. J. McPartlin, E. J. Maile, G. Webster, R. Swindell, C. G. Rowbottom, A. Choudhury, and A. H. Aitkenhead, “Parametrized rectal dose and associations with late toxicity in prostate cancer radiotherapy,” *British Journal of Radiology*, 2015.
- [38] O. Acosta, G. Drean, J. D. Ospina, A. Simon, P. Haigron, C. Lafond, and R. De Crevoisier, “Voxel-based population analysis for correlating local dose and rectal toxicity in prostate cancer radiotherapy,” *Physics in Medicine and Biology*, 2013.
- [39] W. D. Heemsbergen, M. S. Hoogeman, G. A. Hart, J. V. Lebesque, and P. C. Koper, “Gastrointestinal toxicity and its relation to dose distributions in the anorectal region of prostate cancer patients treated with radiotherapy,” *International Journal of Radiation Oncology Biology Physics*, 2005.
- [40] O. Casares-Magaz, L. P. Muren, V. Moiseenko, S. E. Petersen, N. J. Pettersson, M. Høyer, J. O. Deasy, and M. Thor, “Spatial rectal dose/volume metrics predict patient-reported gastro-intestinal symptoms after radiotherapy for prostate cancer,” *Acta Oncologica*, 2017.
- [41] F. Buettner, S. L. Gulliford, S. Webb, M. R. Sydes, D. P. Dearnaley, and M. Partridge, “Assessing correlations between the spatial distribution of the dose to the rectal wall and late rectal toxicity after prostate radiotherapy: An analysis of data from the MRC RT01 trial (ISRCTN 47772397),” *Physics in Medicine and Biology*, 2009.

- [42] X. Zhen, J. Chen, Z. Zhong, B. Hrycushko, L. Zhou, S. Jiang, K. Albuquerque, and X. Gu, “Deep convolutional neural network with transfer learning for rectum toxicity prediction in cervical cancer radiotherapy: A feasibility study,” *Physics in Medicine and Biology*, 2017.
- [43] I. J. Chetty and M. Rosu-Bubulac, “Deformable Registration for Dose Accumulation,” 2019.
- [44] S. T. Ratliff, *Webb’s Physics of Medical Imaging, Second Edition*. 2013.
- [45] J. H. Siewerdsen and D. A. Jaffray, “Cone-beam computed tomography with a flat-panel imager: Magnitude and effects of x-ray scatter,” *Medical Physics*, vol. 28, no. 2, pp. 220–231, 2001.
- [46] W. A. Kalender and Y. Kyriakou, “Flat-detector computed tomography (FD-CT),” 2007.
- [47] N. Papanikolaou, J. J. Battista, A. L. Boyer, C. Kappas, E. Klein, T. R. Mackie, M. Sharpe, and J. V. Dyk, “Tissue Inhomogeneity Corrections for Megavoltage Photon Beams,” tech. rep., AAPM TG #65, 2004.
- [48] E. B. Podgorsak, *Radiation Oncology Physics: A Handbook for Teachers and Students*.
- [49] A. Sethi, “Khan’s Treatment Planning in Radiation Oncology, 4 th Edition. Editor: Faiz M. Khan, John P. Gibbons, Paul W. Sperduto. Lippincott Williams & Wilkins (Wolters Kluwer), Philadelphia, PA, 2016. 648 pp. Price: \$215.99. ISBN: 9781469889979. (Hardcover) (*),” *Medical Physics*, 2018.
- [50] T. R. Mackie, E. El khatib, J. Battista, J. Scrimger, J. Van Dyk, and J. R. Cunningham, “Lung dose corrections for 6 and 15 MV x rays,” *Medical Physics*, 1985.
- [51] C. Constantinou, J. C. Harrington, and L. A. DeWerd, “An electron density calibration phantom for CT-based treatment planning computers,” *Medical Physics*, 1992.
- [52] J. P. Bissonnette, P. A. Balter, L. Dong, K. M. Langen, D. M. Lovelock, M. Miften, D. J. Moseley, J. Pouliot, J. J. Sonke, and S. Yoo, “Quality assurance for image-guided

- radiation therapy utilizing CT-based technologies: A report of the AAPM TG-179,” 2012.
- [53] H. Guan and H. Dong, “Dose calculation accuracy using cone-beam CT (CBCT) for pelvic adaptive radiotherapy,” *Physics in Medicine and Biology*, 2009.
- [54] J. Hatton, B. McCurdy, and P. B. Greer, “Cone beam computerized tomography: The effect of calibration of the Hounsfield unit number to electron density on dose calculation accuracy for adaptive radiation therapy,” *Physics in Medicine and Biology*, vol. 54, pp. N329–N346, aug 2009.
- [55] Y. Akino, Y. Yoshioka, S. Fukuda, S. Maruoka, Y. Takahashi, M. Yagi, H. Mizuno, F. Isohashi, and K. Ogawa, “Estimation of rectal dose using daily megavoltage cone-beam computed tomography and deformable image registration,” *International Journal of Radiation Oncology Biology Physics*, 2013.
- [56] R. Varadhan, S. K. Hui, S. Way, and K. Nisi, “Assessing prostate, bladder and rectal doses during image guided radiation therapy - Need for plan adaptation?,” *Journal of Applied Clinical Medical Physics*, 2009.
- [57] M. Thor, L. Bentzen, L. B. Hysing, C. Ekanger, S. I. Helle, Á. Karlsdóttir, and L. P. Muren, “Prediction of rectum and bladder morbidity following radiotherapy of prostate cancer based on motion-inclusive dose distributions,” *Radiotherapy and Oncology*, 2013.
- [58] P. A. Kupelian, K. M. Langen, O. A. Zeidan, S. L. Meeks, T. R. Willoughby, T. H. Wagner, S. Jeswani, K. J. Ruchala, J. Haimerl, and G. H. Olivera, “Daily variations in delivered doses in patients treated with radiotherapy for localized prostate cancer,” *International Journal of Radiation Oncology Biology Physics*, 2006.
- [59] L. Chen, K. Paskalev, X. Xu, J. Zhu, L. Wang, R. A. Price, W. Hu, S. J. Feigenberg, E. M. Horwitz, A. Pollack, and C. M. Ma, “Rectal dose variation during the course of image-guided radiation therapy of prostate cancer,” *Radiotherapy and Oncology*, vol. 95, pp. 198–202, may 2010.

- [60] J. Li, X. Yang, and J. Yu, "Compact support Thin Plate Spline algorithm," *Journal of Electronics*, 2007.
- [61] J. A. Schnabel, D. Rueckert, M. Quist, J. M. Blackall, A. D. Castellano-Smith, T. Hartkens, G. P. Penney, W. A. Hall, H. Liu, C. L. Truwit, F. A. Gerritsen, D. L. Hill, and D. J. Hawkes, "A generic framework for non-rigid registration based on non-uniform multi-level free-form deformations," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2001.
- [62] H. Wang, L. Dong, J. O'Daniel, R. Mohan, A. S. Garden, K. Kian Ang, D. A. Kuban, M. Bonnen, J. Y. Chang, and R. Cheung, "Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy," *Physics in Medicine and Biology*, 2005.
- [63] H. Zhong, T. Peters, and J. V. Siebers, "FEM-based evaluation of deformable image registration for radiation therapy," *Physics in Medicine and Biology*, 2007.
- [64] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "Deformable templates using large deformation kinematics," *IEEE Transactions on Image Processing*, 1996.
- [65] S. Oh and S. Kim, "Deformable image registration in radiation therapy," 2017.
- [66] K. K. Brock, S. Mutic, T. R. McNutt, H. Li, and M. L. Kessler, "Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132: Report," *Medical Physics*, 2017.
- [67] V. Zambrano, H. Furtado, D. Fabri, C. Lütgendorf-Caucig, J. Góra, M. Stock, R. Mayer, W. Birkfellner, and D. Georg, "Performance validation of deformable image registration in the pelvic region," *Journal of Radiation Research*, 2013.
- [68] S. Thörnqvist, J. B. Petersen, M. Høyer, L. N. Bentzen, and L. P. Muren, "Propagation of target and organ at risk contours in radiotherapy of prostate cancer using deformable image registration," *Acta Oncologica*, 2010.

- [69] N. Kadoya, Y. Y. Miyasaka, T. Yamamoto, Y. Kuroda, K. Ito, M. Chiba, Y. Nakajima, N. Takahashi, M. Kubozono, R. Umezawa, S. Dobashi, K. Takeda, and K. Jingu, “Evaluation of rectum and bladder dose accumulation from external beam radiotherapy and brachytherapy for cervical cancer using two different deformable image registration techniques,” *Journal of Radiation Research*, 2017.
- [70] Z. Saleh, M. Thor, A. P. Apte, G. Sharp, X. Tang, H. Veeraraghavan, L. Muren, and J. Deasy, “A multiple-image-based method to evaluate the performance of deformable image registration in the pelvis,” *Physics in Medicine and Biology*, 2016.
- [71] S. Gao, L. Zhang, H. Wang, R. De Crevoisier, D. D. Kuban, R. Mohan, and L. Dong, “A deformable image registration method to handle distended rectums in prostate cancer radiotherapy,” *Medical Physics*, 2006.
- [72] M. Thor, J. B. Petersen, L. Bentzen, M. Høyer, and L. P. Muren, “Deformable image registration for contour propagation from CT to cone-beam CT scans in radiotherapy of prostate cancer,” in *Acta Oncologica*, 2011.
- [73] K. Motegi, H. Tachibana, A. Motegi, K. Hotta, H. Baba, and T. Akimoto, “Usefulness of hybrid deformable image registration algorithms in prostate radiation therapy,” *Journal of Applied Clinical Medical Physics*, 2019.
- [74] Y. Takayama, N. Kadoya, T. Yamamoto, K. Ito, M. Chiba, K. Fujiwara, Y. Miyasaka, S. Dobashi, K. Sato, K. Takeda, and K. Jingu, “Evaluation of the performance of deformable image registration between planning CT and CBCT images for the pelvic region: Comparison between hybrid and intensity-based DIR,” *Journal of Radiation Research*, 2017.
- [75] M. Thor, E. S. Andersen, J. B. Petersen, T. S. Sorensen, K. O. Noe, K. Tanderup, L. Bentzen, U. V. Elstrom, M. Hoyer, and L. P. Muren, “Evaluation of an application for intensity-based deformable image registration and dose accumulation in radiotherapy,” *Acta Oncologica*, 2014.

- [76] N. Wen, C. Glide-Hurst, T. Nurushev, L. Xing, J. Kim, H. Zhong, D. Liu, M. Liu, J. Burmeister, B. Movsas, and I. J. Chetty, “Evaluation of the deformation and corresponding dosimetric implications in prostate cancer treatment,” *Physics in Medicine and Biology*, 2012.
- [77] M. Velec, J. L. Moseley, S. Svensson, B. Hårdemark, D. A. Jaffray, and K. K. Brock, “Validation of biomechanical deformable image registration in the abdomen, thorax, and pelvis in a commercial radiotherapy treatment planning system:,” *Medical Physics*, 2017.
- [78] T. Ishida, N. Kadoya, S. Tanabe, H. Ohashi, H. Nemoto, S. Dobashi, K. Takeda, and K. Jingu, “Evaluation of performance of pelvic CT-MR deformable image registration using two software programs,” *Journal of Radiation Research*, 2021.
- [79] M. Foskey, B. Davis, L. Goyal, S. Chang, E. Chaney, N. Strehl, S. Tomei, J. Rosenman, and S. Joshi, “Large deformation three-dimensional image registration in image-guided radiation therapy,” *Physics in Medicine and Biology*, 2005.
- [80] N. Kadoya, Y. Fujita, Y. Katsuta, S. Dobashi, K. Takeda, K. Kishi, M. Kubozono, R. Umezawa, T. Sugawara, H. Matsushita, and K. Jingu, “Evaluation of various deformable image registration algorithms for thoracic images,” *Journal of Radiation Research*, 2014.
- [81] U. J. Yeo, M. L. Taylor, J. R. Supple, R. L. Smith, L. Dunn, T. Kron, and R. D. Franich, “Is it sensible to ”deform” dose 3D experimental validation of dose-warping,” *Medical Physics*, vol. 39, pp. 5065–5072, jul 2012.
- [82] C. J. Niu, W. D. Foltz, M. Velec, J. L. Moseley, A. Al-Mayah, and K. K. Brock, “A novel technique to enable experimental validation of deformable dose accumulation,” *Medical Physics*, 2012.
- [83] E. Maynard, E. Heath, M. Hilts, and A. Jirasek, “Evaluation of an x-ray CT polymer gel dosimetry system in the measurement of deformed dose,” *Biomedical Physics and Engineering Express*, 2020.

- [84] D. Hoffmans, N. Niebuhr, O. Bohoudi, A. Pfaffenberger, and M. Palacios, “An end-to-end test for MR-guided online adaptive radiotherapy,” *Physics in Medicine and Biology*, 2020.
- [85] O. Bohoudi, F. J. Lagerwaard, A. M. Bruynzeel, N. I. Niebuhr, W. Johnen, S. Senan, B. J. Slotman, A. Pfaffenberger, and M. A. Palacios, “End-to-end empirical validation of dose accumulation in MRI-guided adaptive radiotherapy for prostate cancer using an anthropomorphic deformable pelvis phantom,” *Radiotherapy and Oncology*, 2019.
- [86] A. Elter, S. Dorsch, P. Mann, A. Runz, W. Johnen, C. K. Spindeldreier, S. Klüter, and C. P. Karger, “End-to-end test of an online adaptive treatment procedure in MR-guided radiotherapy using a phantom with anthropomorphic structures,” *Physics in Medicine and Biology*, 2019.
- [87] R. C. Wortel, M. G. Witte, U. A. van der Heide, F. J. Pos, J. V. Lebesque, M. van Herk, L. Incrocci, and W. D. Heemsbergen, “Dose–surface maps identifying local dose–effects for acute gastrointestinal toxicity after radiotherapy for prostate cancer,” *Radiotherapy and Oncology*, 2015.
- [88] N. Yahya, M. A. Ebert, M. J. House, A. Kennedy, J. Matthews, D. J. Joseph, and J. W. Denham, “Modeling Urinary Dysfunction After External Beam Radiation Therapy of the Prostate Using Bladder Dose-Surface Maps: Evidence of Spatially Variable Response of the Bladder Surface,” *International Journal of Radiation Oncology Biology Physics*, 2017.
- [89] M. Serban, A. A. de Leeuw, K. Tanderup, and I. M. Jürgenliemk-Schulz, “Vaginal dose-surface maps in cervical cancer brachytherapy: Methodology and preliminary results on correlation with morbidity,” *Brachytherapy*, 2021.
- [90] A. Witztum, B. George, S. Warren, M. Partridge, and M. A. Hawkins, “Unwrapping 3D complex hollow organs for spatial dose surface analysis,” *Medical Physics*, 2016.
- [91] F. Dankers, R. Wijsman, E. G. Troost, R. Monshouwer, J. Bussink, and A. L. Hoffmann, “Esophageal wall dose-surface maps do not improve the predictive

- performance of a multivariable NTCP model for acute esophageal toxicity in advanced stage NSCLC patients treated with intensity-modulated (chemo-)radiotherapy,” *Physics in Medicine and Biology*, 2017.
- [92] G. J. Meijer, M. Van Den Brink, M. S. Hoogeman, J. Meinders, and J. V. Lebesque, “Dose-wall histograms and normalized dose-surface histograms for the rectum: A new method to analyze the dose distribution over the rectum in conformal radiotherapy,” *International Journal of Radiation Oncology Biology Physics*, 1999.
- [93] E. Onjukka, C. Fiorino, A. Cicchetti, F. Palorini, I. Improta, G. Gagliardi, C. Cozzarini, C. Degli Esposti, P. Gabriele, R. Valdagni, and T. Rancati, “Patterns in ano-rectal dose maps and the risk of late toxicity after prostate IMRT,” *Acta Oncologica*, 2019.
- [94] S. M. Zhou, L. B. Marks, G. S. Tracton, G. S. Sibley, K. L. Light, P. D. Maguire, and M. S. Anscher, “A new three-dimensional dose distribution reduction scheme for tubular organs,” *Medical Physics*, 2000.
- [95] G. Wang, E. G. McFarland, B. P. Brown, and M. W. Vannier, “GI tract unraveling with curved cross sections,” *IEEE Transactions on Medical Imaging*, 1998.
- [96] S. L. Tucker, M. Zhang, L. Dong, R. Mohan, D. Kuban, and H. D. Thames, “Cluster model analysis of late rectal bleeding after IMRT of prostate cancer: A case-control study,” *International Journal of Radiation Oncology Biology Physics*, 2006.
- [97] E. N. van Lin, J. Kristinsson, M. E. Philippens, D. J. de Jong, L. P. van der Vight, J. H. Kaanders, J. W. Leer, and A. G. Visser, “Reduced late rectal mucosal changes after prostate three-dimensional conformal radiotherapy with endorectal balloon as observed in repeated endoscopy,” *International Journal of Radiation Oncology Biology Physics*, 2007.
- [98] J. Murray, D. McQuaid, A. Dunlop, F. Buettner, S. Nill, E. Hall, D. Dearnaley, and S. Gulliford, “SU-E-J-14: A Novel Approach to Evaluate the Dosimetric Effect of Rectal Variation During Image Guided Prostate Radiotherapy,” *Medical Physics*, 2014.

- [99] B. G. Vanneste, F. Buettner, M. Pinkawa, P. Lambin, and A. L. Hoffmann, “Ano-rectal wall dose-surface maps localize the dosimetric benefit of hydrogel rectum spacers in prostate cancer radiotherapy,” *Clinical and Translational Radiation Oncology*, 2019.
- [100] C. R. Moulton, M. J. House, V. Lye, C. I. Tang, M. Krawiec, D. J. Joseph, J. W. Denham, and M. A. Ebert, “Spatial features of dose-surface maps from deformably-registered plans correlate with late gastrointestinal complications,” *Physics in Medicine and Biology*, 2017.
- [101] F. Buettner, S. L. Gulliford, S. Webb, and M. Partridge, “Modeling late rectal toxicities based on a parameterized representation of the 3D dose distribution,” *Physics in Medicine and Biology*, 2011.
- [102] C. Chen, M. Witte, W. Heemsbergen, and M. V. Herk, “Multiple comparisons permutation test for image based data mining in radiotherapy,” *Radiation Oncology*, 2013.
- [103] W. D. Heemsbergen, L. Incrocci, F. J. Pos, B. J. Heijmen, and M. G. Witte, “Local Dose Effects for Late Gastrointestinal Toxicity After Hypofractionated and Conventionally Fractionated Modern Radiotherapy for Prostate Cancer in the HYPRO Trial,” *Frontiers in Oncology*, 2020.
- [104] O. D’Aquino, *Assessment of delivered rectal dose in prostate cancer radiotherapy*. PhD thesis, Sutton UK, 2021.
- [105] L. E. A. Shelley, *Accumulating delivered dose to the rectum to improve toxicity prediction in prostate radiotherapy*. PhD thesis, Cambridge UK, 2020.
- [106] X. Liu, J. Zhang, D. Ruan, A. S. Yu, V. Sehgal, X. S. Qi, M. C. Barker, Z. L. Shen, and S. Goetsch, “Radiation therapy practice changes in the covid-19 pandemic era: A pilot study in california,” *Journal of Applied Clinical Medical Physics*, vol. 23, no. 11, p. e13770, 2022.

Chapter 3

Technical Note: *rtdsm* — An open-source software for radiotherapy dose-surface map generation and analysis

Haley M Patrick and John Kildea

Article published in: *Medical Physics*. 2022; 49: 7327– 7335

doi: <https://doi.org/10.1002/mp.15900>

3.1 Preface

In this chapter, we describe how we developed an open-source code package for the calculation and analysis of dose-surface maps (DSMs). DSM calculation is a complex and variable process, with many different existing techniques to perform each stage of the process. To date, individual researchers have typically developed their own proprietary codes that may entail different slice orientations, sampling resolutions, or unwrapping techniques, and may also utilize different DSM analysis techniques. However, the

variability introduced by such compartmentalized code bases hinders inter-study comparisons and meta-analyses. Furthermore, it also increases the barrier to entry into the DSM research field, especially for individuals who are not proficient in computer programming. We therefore determined that the creation and release of an open-source code package for DSM calculation would be a major step in addressing both issues.

In this chapter, we outline the design and testing of a customizable DSM calculation code package, *rtdsm*, to replicate the existing dominant DSM calculation techniques in the literature. We used the popular and beginner-friendly open-source computation language, Python, to construct a package of modular functions for DSM generation and analysis. For the first time in the literature, we reported implementation methodologies that calculate both planar and non-coplanar DSMs in a single code base. Our package also includes support for post-processing and popular DSM analysis techniques. Finally, we demonstrated the package's usefulness with an end-to-end comparison study of DSMs from two prostate radiotherapy treatment courses. This work serves not only as a basis for the DSM studies that are described later in this thesis, but also provides an open-source, ready-to-use platform for DSM research that is available to the broader scientific community.

3.2 Abstract

Background: Dose-outcome studies in radiation oncology have historically excluded spatial information due to dose-volume histograms being the most dominant source of dosimetric information. In recent years, dose-surface maps (DSMs) have become increasingly popular for characterization of spatial dose distributions and identification of radiosensitive subregions for hollow organs. However, methodological variations and lack of open-source, publicly offered code-sharing between research groups has limited reproducibility and wider adoption.

Purpose: This article presents *rtdsm*, an open-source software for DSM calculation with the intent to improve the reproducibility of and the access to DSM-based research in medical physics and radiation oncology.

Methods: A literature review was conducted to identify essential functionalities and

prevailing calculation approaches to guide development. The described software has been designed to calculate DSMs from DICOM data with a high degree of user customizability and to facilitate DSM feature analysis. Core functionalities include DSM calculation, equivalent dose conversions, common DSM feature extraction, and simple DSM accumulation.

Results: A number of use cases were used to qualitatively and quantitatively demonstrate the use and usefulness of *rtdsm*. Specifically, two DSM slicing methods, planar and non-coplanar, were implemented and tested, and the effects of method choice on output DSMs were demonstrated. An example comparison of DSMs from two different treatments was used to highlight the use cases of various built-in analysis functions for equivalent dose conversion and DSM feature extraction.

Conclusions: We developed and implemented *rtdsm* as a standalone software that provides all essential functionalities required to perform a DSM-based study. It has been made freely accessible under an open-source license on Github to encourage collaboration and community use.

3.3 Introduction

Achieving balance between sufficient tumour irradiation and normal tissue sparing is a longstanding challenge in the field of radiotherapy. For this reason, dosimetric constraints for organs-at-risk (OARs) derived from dose-volume histograms (DVHs) have been an essential component of treatment planning since their introduction in the 1990s [1]. However, DVHs are limited by their lack of spatial information and by the inherent assumption that OARs respond homogeneously to radiation [2]. This assumption is counter to evidence supporting regional variations in OAR sensitivity, which have been reported for multiple organs such as salivary glands, bladder, rectum, and lung [3–7] using a variety of techniques to study spatial dose information [8].

A popular instrument used to help identify radiosensitive subregions in hollow organs is the dose-surface map (DSM), which projects a region of interest's (ROI) surface dose onto a 2D grid. This is typically achieved by sampling the surface dose across multiple slices of the ROI and cutting and unwrapping its hollow structure to create the map, though some

alternative surface dose sampling methods exist [9, 10]. To date, DSMs have been used to identify spatial dose features predictive of toxicities for the rectum, bladder, vagina, and heart [11–15].

Despite their growing popularity, there is a lack of consensus about how to generate DSMs. Most notable is the ways in which the ROI slices are defined: some groups opt to use axial slices parallel to the slices of the treatment planning image [6, 16, 17] (planar slicing), while others define them orthogonal to the ROI’s central-axis path [12, 18, 19] (non-coplanar slicing), arguing that this better represents ROIs with irregular curvatures. Sampling resolutions and unwrapping approaches can also be disparate. For example, reported rectum DSM resolutions vary from 21×21 [6] to 120×200 [20] pixels with minimal justification as to why. This issue is further compounded by inconsistent methodological reporting, as not all groups provide sufficient information on their sampling or unwrapping approaches for their work to be reproducible. The lack of consensus, paired with a lack of code and data-sharing within the broader community makes it challenging to replicate and validate earlier findings.

In this work, we present a technical description of a new open-source software package that we have developed for DSM calculations called *rtdsm*. To the best of our knowledge, our package is the first open-source DSM software that is highly customizable and capable of implementing both slicing methods, as well providing functionalities for dose accumulation, equivalent dose conversions, and DSM feature extraction. As such, we believe it offers the potential to reduce software development barriers for DSM implementation and provides a referenceable calculation framework to facilitate reproducibility and consistent DSM reporting.

3.4 Methods

3.4.1 Planning Phase

In order to inform the development of *rtdsm*, we conducted a literature review [21] of the DSM literature and used it to guide our design considerations and identify the needs of the community. The Pubmed and Google Scholar databases were queried using the phrase

”dose surface map*” to identify English language articles published between 2000 and 2021. The references sections of relevant identified articles were also used to identify further relevant papers. Only papers that included a description of the DSM calculation process were included in the final list, which consisted of 30 publications [5, 6, 9–20, 22–37]. The contents of each paper were read in detail to identify the methodology and parameters used for DSM calculation, and if the authors had made their code publically available. This was followed by a search for dose-surface mapping software packages using the Google search engine and on the GitHub and pypi online code repositories. An overview of the review’s findings is presented in [Table S-1](#).

Our review found just a single open-source package for non-coplanar DSM calculation (which was not mentioned in its associated journal article) [19] and nothing for planar DSMs. Furthermore, the code bases described in the literature were each restricted to single slicing approaches [6, 18, 19], limiting calculation flexibility. Based on these limitations and the commonalities in the calculation and analysis methodologies identified during the review, we determined that *rtdsm* must:

1. **Be open-source** to help reduce software development barriers and facilitate standards.
2. **Support planar and non-coplanar slicing methods** to facilitate the most appropriate DSM calculation strategy for each ROI.
3. **Permit customizable slicing and sampling resolutions** to allow for reproduction of previous and future studies.
4. **Calculate and report common DSM features** for use in clinical decision-making and dose-outcomes research.
5. **Be modularly designed** to enable easy implementation of alternative or improved calculation methods in the future.

We opted not to use the existing open-source package of Witztum et al. (2016, <https://github.com/bgeorge0/dsm>) [19], which provides functionality solely for non-coplanar DSM calculations as it is written for the Matlab platform (The MathWorks, Inc., Natick, MA), and we desired a solution for both slicing methods and DSM analysis

that would be fully accessible to the community with minimal cost overhead. After considering the various programming language options available, we selected Python for *rtdsm* due to its popularity, variety of libraries, and open-source nature. The Python libraries utilized by *rtdsm* are listed in Table 3.1 The software is designed to be imported and used like other python packages, and is available on GitHub (<https://github.com/McGillMedPhys/rtdsm>). Detailed documentation and tutorials on software implementation are included with the repository in accordance with Python code style standards. The remainder of the Methods therefore focuses on describing the process used by *rtdsm* to calculate a DSM following the four key stages identified in the literature (Fig 3.1).

Table 3.1: Python library dependencies of *rtdsm*.

Library	Usage
pyvista [38]	Create and operate on 3D mesh objects
pydicom [39]	Read DICOM formatted data files
scikit-image [40]	Mesh generation, cluster analysis
scipy [41]	Interpolation operations

3.4.2 Stage 1: Mesh Creation

ROI contour information is provided by way of a point-cloud stored in RTStructureSet DICOM-RT format that is generated by the treatment planning system. Once read into *rtdsm*, a surface mesh of the ROI is generated from its point-cloud using a smoothed marching cubes algorithm [42]. While we also tested Delaunay triangulation [43], we found that it was too sensitive to irregular point-cloud resolution and open surfaces and subsequently produced inadequate meshes for commonly-studied ROIs.

In addition to generating the surface mesh, the point-cloud read-in process also calculates the geometric central axis path (CAP) of the ROI using the centroids of the CT slices delineated in the RTStructure file. Alternative user-created CAPs can also be swapped in to replace the default geometric CAP.

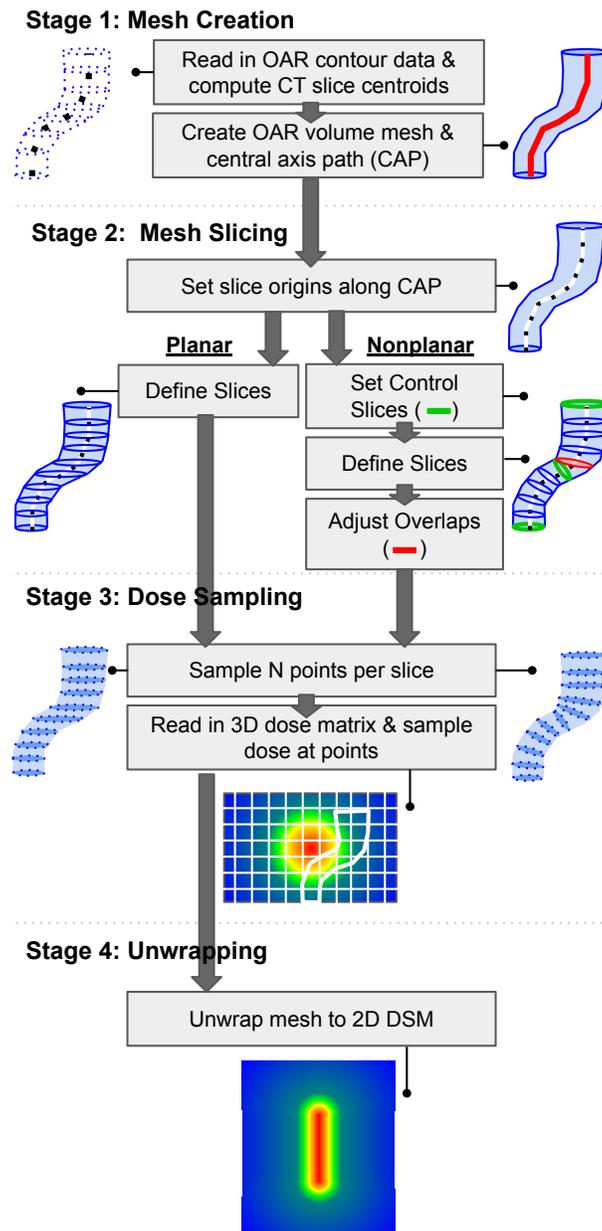


Figure 3.1: The four key stages of a typical dose-surface map generation workflow for an organ at risk, as identified in our literature review and implemented in *rtdsm*.

3.4.3 Stage 2: Mesh Slicing

A series of slice origins are specified along the length of the CAP based on the preferred slicing method for the DSM (planar or non-coplanar) and the preferred mapping along the

vertical axis. Both 1:1 (where slices are spaced by a constant absolute distance) and scaled mapping (where the number of slices remains constant and spacing adjusts accordingly) can be used to define slice origins in *rtdsm*. As described below, slice planes are then defined at each slice origin according to the chosen slicing method and the centroid of each mesh slice is approximated using the method shown in Figure 3.2a. This approximated slice centroid is used as the starting point for equiangular ray-casting of a user-defined number of rays. The intersection points of the rays with the surface mesh are then stored in a Python dictionary and used in stage 3 to sample the dose (dose sampling points).

Planar Slicing

Planar slicing is conducted by defining equally-spaced parallel slice planes along a single linear axis of the ROI [6, 16, 24], the orientations of which are set using a user-specified slice normal vector. If a user-specified slice normal is not provided, *rtdsm* creates axial planes by default.

Non-coplanar Slicing

Unlike planar slicing, non-coplanar slicing uses slice planes that are each individually orthogonal to the ROI's CAP. Each of these slice planes are defined by a tangent vector to the CAP at the slice origin point, which is determined using the preceding and succeeding points and the CAP's gradient (Fig 3.2b). Because of these slice planes, the non-coplanar slices may overlap with one another, requiring additional steps for overlap detection and correction. To facilitate these additional steps, our implementation of non-coplanar slicing begins by defining "control slices" along the CAP. A control slice is a slice at the start, end, or a point of direction change along the CAP where overlapping slices are more common. They serve to quickly identify slice planes that are angled such that they exit the mesh or overlap many other slice planes, thereby simplifying the level of slice corrections needed. Once the control slices are created, the code checks the proposed slices in ascending order for overlaps with the closest control slices and the preceding neighbouring slice. If no overlap is found, the proposed slice is retained, otherwise it is flagged for adjustment. This approach, paired with the selected slice adjustment methodology, removes the need to

iteratively check all slices for collisions.

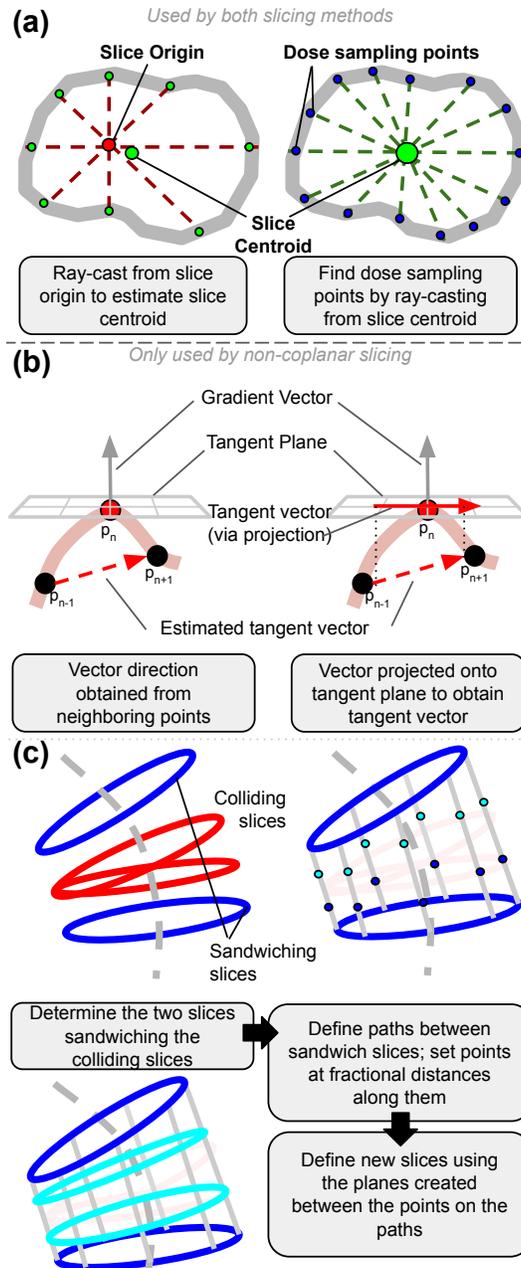


Figure 3.2: Visual explanations of specific operations performed during planar and non-coplanar mesh slicing. (a) Two-step ray-casting approach to acquire dose sampling points used by both slicing methods. (b) Calculation of CAP tangent vectors to define planes for non-coplanar slicing. (c) Correction method to resolve overlapping non-coplanar slices.

Slice adjustments are performed following the methodology of Witztum *et al* [19]. To briefly summarise, adjacent flagged slices are grouped together with the two non-flagged slices that sandwich them (Fig 3.2c). Matching angular vertices of the sandwiching slices are connected with one another and the connecting lines are segmented at N equidistant points, where N is equal to the number of flagged slices between them. Corresponding segmentation points are then used to define new non-overlapping sampling planes for the flagged slices, and ray-casting of the dose sampling points proceeds as normal.

3.4.4 Stages 3 and 4: Dose Sampling and Unwrapping

As is the case for contour data input, *rtdsm* accepts DICOM-RT files with dose information. Dose matrices are read in from RTDose files and are used to sample dose by means of linear interpolation at the dose sampling points identified during mesh slicing. The cutting open and unwrapping of the DSM is straightforward due to the way in which *rtdsm* defines dose sampling points. This is because the ray-casting process defines rays in clockwise order using a slice-specific orthonormal basis, wherein the j axis (along which the first ray is cast) is always defined to point to the ROI's posterior wall, regardless of slice orientation. If needed, post-processing can be performed on the output DSM to change the cutpoint.

3.4.5 Post-processing Functionalities

Cluster and ellipse-based spatial dose features are calculated by finding the largest contiguous cluster of DSM pixels above a given dose level and fitting an ellipse to them [6, 16, 20, 32]. Table 3.2 outlines the features developed by Buettner *et al.* and Moulton *et al.* that *rtdsm* supports. Additionally, we included support for equivalent dose (EQD) conversion [44] and DSM aggregation (addition, subtraction, and averaging) to facilitate common analysis and visualization strategies.

Table 3.2: Spatial dose features supported by *rtdsm*.

Feature	Definition
Cluster Area	The percent area of the DSM covered by the cluster.
Cluster Centroid	The center of mass of the cluster. Provided in units of array indices and percent of the lateral and longitudinal spans.
Cluster Lateral Extent	The percent of the lateral span of the DSM covered by the cluster.
Cluster Longitudinal Extent	The percent of the longitudinal span of the DSM covered by the cluster.
Ellipse Area	The percent area of the DSM covered by the ellipse.
Ellipse Angle	The rotation of the ellipse, in radians.
Ellipse Eccentricity	The eccentricity of the ellipse.
Ellipse Lateral projection	The percent of the lateral span of the DSM covered by a projection of the ellipse's lateral axis onto the DSM's lateral axis.
Ellipse Longitudinal projection	The percent of the longitudinal span of the DSM covered by a projection of the ellipse's longitudinal axis onto the DSM's longitudinal axis.

3.5 Results

3.5.1 Dose-Surface Maps

The performance and capabilities of *rtdsm* were tested using using retrospective data from 36.25 Gy in 5 fraction VMAT plans created in Eclipse (Varian Medical Systems, Palo Alto, CA). 40 planar and non-coplanar rectum DSMs were calculated using 3 mm stepsize 1:1 vertical mapping, with 45 points per slice. On average, calculation of the planar DSMs took 2.8 minutes (≈ 4.2 sec/DSM) and 5.7 minutes (≈ 8.5 sec/DSM) for non-coplanar on an Intel® Xeon® CPU X3440 (2.53 GHz) with 4 GB of RAM. Figure 3.3 shows the rectum DSMs obtained for several patients and illustrates how rectum shape and slicing choice influence the final DSM. For ROIs with CAPs that closely follow the longitudinal axis of the contoured image, the planar and non-coplanar slices are similar, resulting in similar DSMs. However, if the CAP significantly traverses anteriorly-posteriorly or left-right, the slices and subsequent DSMs are quite different. As seen in Figure 3.3, the orientation of planar slices

in these example cases leads to the inclusion of more anterior points that lie in the high-dose region, increasing the size of the hotspot relative to the non-coplanar DSM.

3.5.2 DSM Conversions and Combinations

In many studies it can be beneficial to combine multiple DSMs in order to visually compare patient cohorts. *rtdsm*'s DSM combination function provides a built-in method to do this, under the assumption all DSMs are aligned at the first (inferior-most) slice and use the same vertical sampling approach. As an example of its possible use cases, non-coplanar rectum DSMs were calculated using scaled vertical mapping to produce normalized 30×30 pixel DSMs for two cohorts of ten prostate cancer patients each, who were either prescribed 60 Gy in 20 fractions (hypofractionated IMRT) or 36.25 Gy in 5 fractions (SBRT). Figure 3.4 shows the average DSM for each cohort, as well as their difference calculated by *rtdsm*.

In order to properly compare them, the DSMs were converted to EQD2 Gy using *rtdsm*'s built-in conversion function with an α/β ratio of 2.3 [9]. Through the comparison it is made apparent that the SBRT treatment regularly delivers doses exceeding 80 Gy to a small anterior region of the rectum (mean area: 4.8%), whereas this is less common for the IMRT treatment (mean area: 0.7%). However, the IMRT treatment delivers doses of ≥ 40 Gy to noticeably larger portions of the rectum than SBRT treatment (32.5% vs. 15.7%).

3.5.3 Spatial Features

A quantitative comparison of the example cohorts from the previous section was conducted by calculating spatial features for 15, 30, 45, 60, and 75 Gy clusters using the EQD2 Gy converted DSMs (Fig 3.5). As noted visually, the IMRT treatment (red points) delivered dose to systematically larger areas of the rectum than the SBRT treatment (blue points), which is also apparent from the DVH and the area-based features. However, spatial features also revealed that the investigated dose levels spanned systematically larger longitudinal proportions of the rectum in the IMRT group than the SBRT group for all dose levels, but only for those levels below 60 Gy when examining lateral span. These patterns may be relevant to differences in outcomes between the two groups and are not immediately apparent

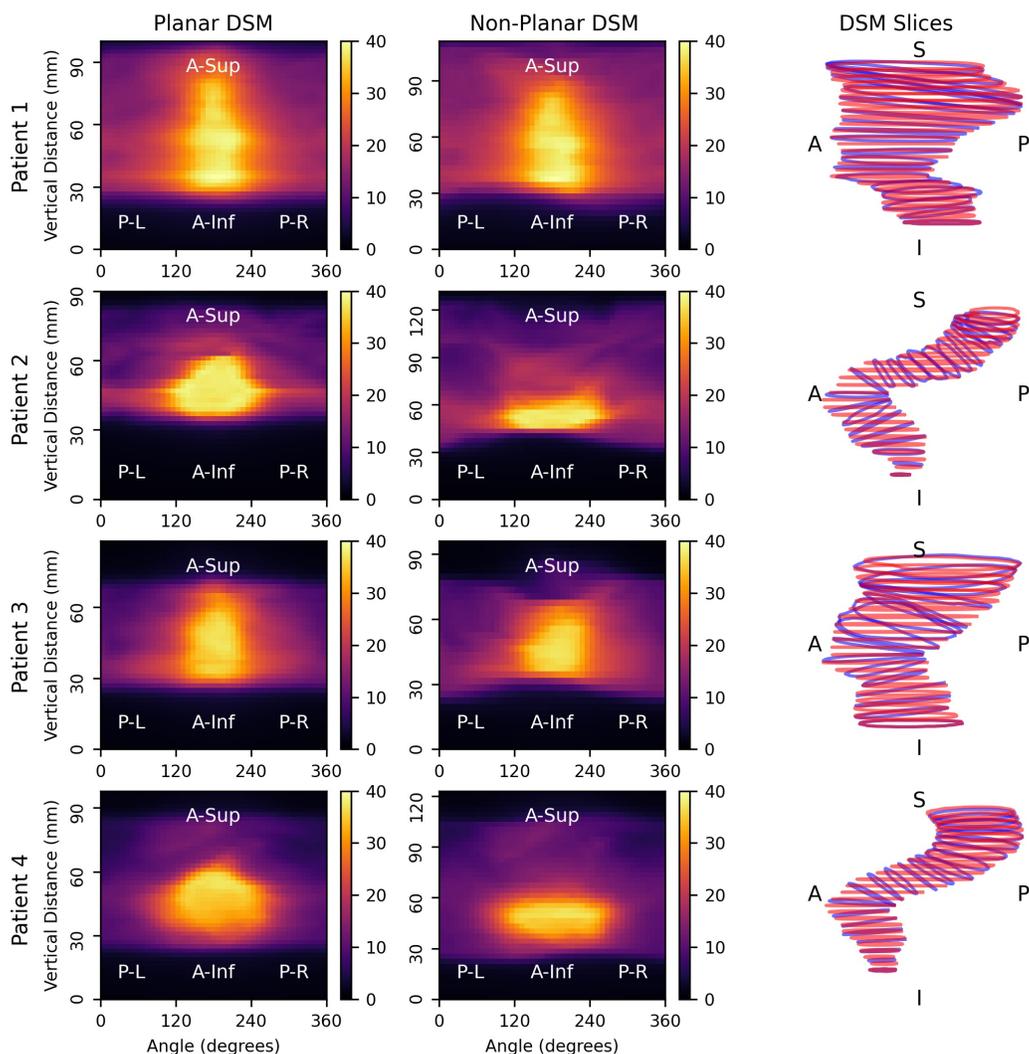


Figure 3.3: Planar DSMs (left), non-coplanar DSMs (centre), and the slices of the ROI used to construct the DSMs (right) of four example patients (units: Gy). Individual vertical axes are used for each DSM in order to illustrate how choice of slicing method influences sampling path length and DSM shape. In the right column, planar slices are shown in pale red and non-coplanar slices in dark blue. For clarity, only every second slice is shown.

based on volume or area information alone, thus highlighting the value of the spatial-dose features. Features calculated from planar DSMs are shown in [Figure S-2](#) for interested readers.

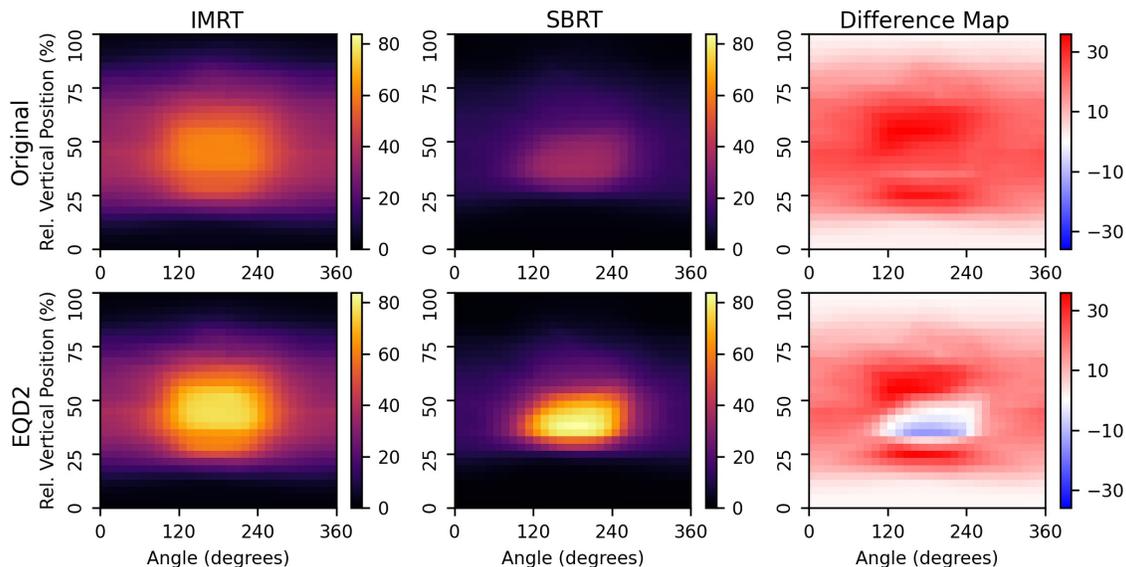


Figure 3.4: Average dose-surface maps for the IMRT and SBRT cohorts (10 patients each), before and after EQD2 Gy conversion, in units of Gy. Difference maps are also shown with contours indicating specific dose thresholds.

3.6 Discussion

We have presented a technical overview of a new open-source package for DSM calculation that we have developed called *rtdsm*. Using retrospective data, we have demonstrated *rtdsm*'s ability to calculate planar and non-coplanar DSMs, combine and convert cohort data, and extract common spatial-dose features to enable users to perform standard DSM-based studies.

DSMs have been used in dose-outcome research since the early 2000s, but calculation tools have largely remained as custom in-house developments with little exchange between groups. For example, from our reading of the author lists, it appears that most rectal DSM papers using the more complex non-coplanar approach include an author or associate of the original 2004 method paper by Hoogeman *et al* [11, 12, 18, 33, 34]. Non-associated groups appear to have largely opted to use the simpler-to-develop planar approach, despite non-coplanar DSMs being arguably a more appropriate representation of the rectum structure. *rtdsm* is the first DSM codebase, to our knowledge, to support both calculation methods

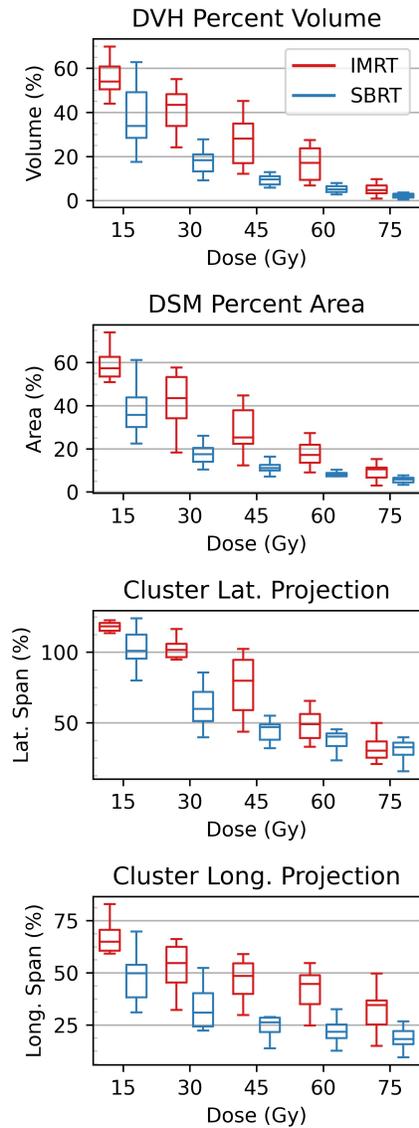


Figure 3.5: Boxplots of the dose features of the IMRT and SBRT cohorts, converted to EQD2 Gy: (a) percent volume (from DVHs), (b) percent area (from DSMs), (c) lateral projection (from DSMs), and (d) longitudinal projection (from DSMs).

that is also an open-source release, inline with our goal to remove the large development barrier to the non-coplanar approach and increasing its accessibility to more researchers.

In addition to removing programming barriers, the accessibility of *rtdsm* has the potential to improve DSM reproducibility between groups, which is becoming an area of concern as the

methodology gains in popularity. Mylona *et al.* [35] recently reported poor reproducibility of toxicity-predictive subregions from bladder DSMs of prostate patients when comparing their results to three other studies. Despite similar DSM construction and analysis approaches, it appears variations in vertical mapping methods between groups (scaled mapping by Mylona, 1:1 mapping truncated at 25 or 45 mm by others [5, 13, 36]) influenced the reproducibility of results. Similar reproducibility issues may also exist for rectal DSMs, especially considering the greater diversity of slicing methods, vertical mapping schemes, and analysis approaches used by different researchers. The result of Mylona’s study is an important example of how the results of DSM analyses depend on the calculation methods used and how the lack of methodological standardisation over the past two decades may be impeding confirmation of important results. We present *rtdsm* as an accessible open-source package and as an opportunity to begin the discussion around DSM standardization. It is designed to be a software with which research groups can easily test the methods and findings of each other by applying custom settings.

rtdsm is published at github.com and will continue to be updated and built upon as the needs of the DSM field evolve, either by its original developers or by new contributors who are welcome to support the project. For example, we have already identified different CAP-generating approaches, such as the racecar [45] or electric field path [22] methods, as areas for potential future improvement, along with implementation of more advanced unwrapping approaches [19]. Support for additional analysis approaches akin to the significance testing popularised by Chen *et al.* [46] are other areas of active development. While a potential limitation of the current version *rtdsm* is that testing has largely focused on rectum structures to date, further generalizability tests using other organs are planned. Further investigations into additional uses of *rtdsm* are also of interest, including the creation of surface maps from other 3D medical data like PET or ECG images, as well as the extraction of “dosiomic” features from DSMs [47].

3.7 Conclusions

We have presented a technical overview of *rtdsm*, a new software for the calculation and evaluation of dose-surface maps. *rtdsm* is a python package that works by computing the dose to the surface of a 3D contoured object and unwrapping it to a 2D map according to user specifications. It is highly flexible and extensible with sufficiently small calculation times to facilitate analysis of large datasets. The results presented in this work demonstrate how *rtdsm* can be used to (1) create multiple types of DSMs, (2) calculate DSM features, and (3) evaluate spatial-dose variations between cohorts. *rtdsm* has been made publicly accessible through GitHub (<https://github.com/McGillMedPhys/rtdsm>) with detailed examples and documentation and can be freely used or contributed to by any user.

3.8 Acknowledgements

We gratefully appreciate the insights and support of Luc Galarneau for valuable discussions regarding the construction and packaging of *rtdsm*. H.M.P. acknowledges funding from the Fonds de recherche du Québec - Santé in the form of a Doctoral Training Award, as well as partial support by the CREATE Responsible Health and Healthcare Data Science (SDRDS) grant of the Natural Sciences and Engineering Research Council. The authors have no conflicts to disclose.

Bibliography

- [1] B. Emami, J. Lyman, A. Brown, L. Cola, M. Goitein, J. E. Munzenrider, B. Shank, L. J. Solin, and M. Wesson, “Tolerance of normal tissue to therapeutic irradiation,” *International Journal of Radiation Oncology, Biology, Physics*, vol. 21, pp. 109–122, may 1991.
- [2] D. A. Jaffray, P. E. Lindsay, K. K. Brock, J. O. Deasy, and W. A. Tomé, “Accurate accumulation of dose for improved understanding of radiation effects in normal tissue.”

- International journal of radiation oncology, biology, physics*, vol. 76, pp. S135–9, mar 2010.
- [3] W. Jiang, P. Lakshminarayanan, X. Hui, P. Han, Z. Cheng, M. Bowers, I. Shpitser, S. Siddiqui, R. H. Taylor, H. Quon, and T. McNutt, “Machine Learning Methods Uncover Radiomorphologic Dose Patterns in Salivary Glands that Predict Xerostomia in Patients with Head and Neck Cancer,” *Advances in Radiation Oncology*, 2019.
- [4] P. Van Luijk, S. Pringle, J. O. Deasy, V. V. Moiseenko, H. Faber, A. Hovan, M. Baanstra, H. P. Van Der Laan, R. G. Kierkels, A. Van Der Schaaf, M. J. Witjes, J. M. Schippers, S. Brandenburg, J. A. Langendijk, J. Wu, and R. P. Coppes, “Sparing the region of the salivary gland containing stem cells preserves saliva production after radiotherapy for head and neck cancer,” *Science Translational Medicine*, 2015.
- [5] F. Palorini, C. Cozzarini, S. Gianolini, A. Botti, V. Carillo, C. Iotti, T. Rancati, R. Valdagni, and C. Fiorino, “First application of a pixel-wise analysis on bladder dose-surface maps in prostate cancer radiotherapy,” *Radiotherapy and Oncology*, 2016.
- [6] F. Buettner, S. L. Gulliford, S. Webb, M. R. Sydes, D. P. Dearnaley, and M. Partridge, “Assessing correlations between the spatial distribution of the dose to the rectal wall and late rectal toxicity after prostate radiotherapy: An analysis of data from the MRC RT01 trial (ISRCTN 47772397),” *Physics in Medicine and Biology*, 2009.
- [7] Y. Seppenwoolde, K. De Jaeger, L. J. Boersma, J. S. Belderbos, and J. V. Lebesque, “Regional differences in lung radiosensitivity after radiotherapy for non-small-cell lung cancer,” *International Journal of Radiation Oncology Biology Physics*, 2004.
- [8] M. A. Ebert, S. Gulliford, O. Acosta, R. De Crevoisier, T. McNutt, W. D. Heemsbergen, M. Witte, G. Palma, T. Rancati, and C. Fiorino, “Spatial descriptions of radiotherapy dose: Normal tissue complication models and statistical associations,” 2021.
- [9] L. E. Shelley, M. P. Sutcliffe, S. J. Thomas, D. J. Noble, M. Romanchikova, K. Harrison, A. M. Bates, N. G. Burnet, and R. Jena, “Associations between voxel-level accumulated

- dose and rectal toxicity in prostate radiotherapy,” *Physics and Imaging in Radiation Oncology*, 2020.
- [10] R. Munbodh, A. Jackson, J. Bauer, C. Ross Schmidlein, and M. J. Zelefsky, “Dosimetric and anatomic indicators of late rectal toxicity after high-dose intensity modulated radiation therapy for prostate cancer,” *Medical Physics*, 2008.
- [11] W. D. Heemsbergen, M. S. Hoogeman, G. A. Hart, J. V. Lebesque, and P. C. Koper, “Gastrointestinal toxicity and its relation to dose distributions in the anorectal region of prostate cancer patients treated with radiotherapy,” *International Journal of Radiation Oncology Biology Physics*, 2005.
- [12] R. C. Wortel, M. G. Witte, U. A. van der Heide, F. J. Pos, J. V. Lebesque, M. van Herk, L. Incrocci, and W. D. Heemsbergen, “Dose–surface maps identifying local dose–effects for acute gastrointestinal toxicity after radiotherapy for prostate cancer,” *Radiotherapy and Oncology*, 2015.
- [13] N. Yahya, M. A. Ebert, M. J. House, A. Kennedy, J. Matthews, D. J. Joseph, and J. W. Denham, “Modeling Urinary Dysfunction After External Beam Radiation Therapy of the Prostate Using Bladder Dose-Surface Maps: Evidence of Spatially Variable Response of the Bladder Surface,” *International Journal of Radiation Oncology Biology Physics*, 2017.
- [14] M. Serban, A. A. de Leeuw, K. Tanderup, and I. M. Jürgenliemk-Schulz, “Vaginal dose-surface maps in cervical cancer brachytherapy: Methodology and preliminary results on correlation with morbidity,” *Brachytherapy*, 2021.
- [15] A. McWilliam, C. Dootson, L. Graham, K. Banfill, A. Abravan, and M. van Herk, “Dose surface maps of the heart can identify regions associated with worse survival for lung cancer patients treated with radiotherapy,” *Physics and Imaging in Radiation Oncology*, 2020.
- [16] C. R. Moulton, M. J. House, V. Lye, C. I. Tang, M. Krawiec, D. J. Joseph, J. W. Denham, and M. A. Ebert, “Spatial features of dose-surface maps from deformably-

- registered plans correlate with late gastrointestinal complications,” *Physics in Medicine and Biology*, 2017.
- [17] J. Chen, H. Chen, Z. Zhong, Z. Wang, B. Hrycushko, L. Zhou, S. Jiang, K. Albuquerque, X. Gu, and X. Zhen, “Investigating rectal toxicity associated dosimetric features with deformable accumulated rectal surface dose maps for cervical cancer radiotherapy,” *Radiation Oncology*, 2018.
- [18] M. S. Hoogeman, M. Van Herk, J. De Bois, P. Muller-Timmermans, P. C. Koper, and J. V. Lebesque, “Quantification of local rectal wall displacements by virtual rectum unfolding,” *Radiotherapy and Oncology*, vol. 70, pp. 21–30, jan 2004.
- [19] A. Witztum, B. George, S. Warren, M. Partridge, and M. A. Hawkins, “Unwrapping 3D complex hollow organs for spatial dose surface analysis,” *Medical Physics*, 2016.
- [20] E. Onjukka, C. Fiorino, A. Cicchetti, F. Palorini, I. Improta, G. Gagliardi, C. Cozzarini, C. Degli Esposti, P. Gabriele, R. Valdagni, and T. Rancati, “Patterns in ano-rectal dose maps and the risk of late toxicity after prostate IMRT,” *Acta Oncologica*, 2019.
- [21] M. J. Grant and A. Booth, “A typology of reviews: an analysis of 14 review types and associated methodologies: A typology of reviews, Maria J. Grant & Andrew Booth,” *Health Info. Libr. J.*, 2009.
- [22] G. Wang, E. G. McFarland, B. P. Brown, and M. W. Vannier, “GI tract unraveling with curved cross sections,” *IEEE Transactions on Medical Imaging*, 1998.
- [23] G. J. Meijer, M. Van Den Brink, M. S. Hoogeman, J. Meinders, and J. V. Lebesque, “Dose-wall histograms and normalized dose-surface histograms for the rectum: A new method to analyze the dose distribution over the rectum in conformal radiotherapy,” *International Journal of Radiation Oncology Biology Physics*, 1999.
- [24] S. L. Tucker, M. Zhang, L. Dong, R. Mohan, D. Kuban, and H. D. Thames, “Cluster model analysis of late rectal bleeding after IMRT of prostate cancer: A case-control study,” *International Journal of Radiation Oncology Biology Physics*, 2006.

- [25] E. N. van Lin, J. Kristinsson, M. E. Philippens, D. J. de Jong, L. P. van der Vight, J. H. Kaanders, J. W. Leer, and A. G. Visser, “Reduced late rectal mucosal changes after prostate three-dimensional conformal radiotherapy with endorectal balloon as observed in repeated endoscopy,” *International Journal of Radiation Oncology Biology Physics*, 2007.
- [26] F. Buettner, S. L. Gulliford, S. Webb, and M. Partridge, “Using dose-surface maps to predict radiation-induced rectal bleeding: A neural network approach,” *Physics in Medicine and Biology*, 2009.
- [27] F. Buettner, S. L. Gulliford, S. Webb, and M. Partridge, “Modeling late rectal toxicities based on a parameterized representation of the 3D dose distribution,” *Physics in Medicine and Biology*, 2011.
- [28] J. Murray, D. McQuaid, A. Dunlop, F. Buettner, S. Nill, E. Hall, D. Dearnaley, and S. Gulliford, “SU-E-J-14: A Novel Approach to Evaluate the Dosimetric Effect of Rectal Variation During Image Guided Prostate Radiotherapy,” *Medical Physics*, 2014.
- [29] J. E. Scaife, S. J. Thomas, K. Harrison, M. Romanchikova, M. P. F. Sutcliffe, J. R. Forman, A. M. Bates, R. Jena, M. A. Parker, and N. G. Burnet, “Accumulated dose to the rectum, measured using dose–volume histograms and dose-surface maps, is different from planned dose in all patients treated with radiotherapy for prostate cancer,” *The British Journal of Radiology*, vol. 88, p. 20150243, oct 2015.
- [30] O. Casares-Magaz, L. P. Muren, V. Moiseenko, S. E. Petersen, N. J. Pettersson, M. Høyer, J. O. Deasy, and M. Thor, “Spatial rectal dose/volume metrics predict patient-reported gastro-intestinal symptoms after radiotherapy for prostate cancer,” *Acta Oncologica*, 2017.
- [31] L. E. Shelley, J. E. Scaife, M. Romanchikova, K. Harrison, J. R. Forman, A. M. Bates, D. J. Noble, R. Jena, M. A. Parker, M. P. Sutcliffe, S. J. Thomas, and N. G. Burnet, “Delivered dose can be a better predictor of rectal toxicity than planned dose in prostate radiotherapy,” *Radiotherapy and Oncology*, 2017.

- [32] B. G. Vanneste, F. Buettner, M. Pinkawa, P. Lambin, and A. L. Hoffmann, “Ano-rectal wall dose-surface maps localize the dosimetric benefit of hydrogel rectum spacers in prostate cancer radiotherapy,” *Clinical and Translational Radiation Oncology*, 2019.
- [33] D. M. Reijtenbagh, J. Godart, J. W. M. Mens, S. T. Heijkoop, W. D. Heemsbergen, and M. S. Hoogeman, “Patient-reported acute GI symptoms in locally advanced cervical cancer patients correlate with rectal dose,” *Radiotherapy and Oncology*, 2020.
- [34] W. D. Heemsbergen, L. Incrocci, F. J. Pos, B. J. Heijmen, and M. G. Witte, “Local Dose Effects for Late Gastrointestinal Toxicity After Hypofractionated and Conventionally Fractionated Modern Radiotherapy for Prostate Cancer in the HYPRO Trial,” *Frontiers in Oncology*, 2020.
- [35] E. Mylona, A. Cicchetti, T. Rancati, F. Palorini, C. Fiorino, S. Supiot, N. Magne, G. Crehange, R. Valdagni, O. Acosta, and R. de Crevoisier, “Local dose analysis to predict acute and late urinary toxicities after prostate cancer radiotherapy: Assessment of cohort and method effects,” *Radiotherapy and Oncology*, 2020.
- [36] I. Improta, F. Palorini, C. Cozzarini, T. Rancati, B. Avuzzi, P. Franco, C. Degli Esposti, E. Del Mastro, G. Girelli, C. Iotti, V. Vavassori, R. Valdagni, and C. Fiorino, “Bladder spatial-dose descriptors correlate with acute urinary toxicity after radiation therapy for prostate cancer,” *Physica Medica*, 2016.
- [37] F. Dankers, R. Wijsman, E. G. Troost, R. Monshouwer, J. Bussink, and A. L. Hoffmann, “Esophageal wall dose-surface maps do not improve the predictive performance of a multivariable NTCP model for acute esophageal toxicity in advanced stage NSCLC patients treated with intensity-modulated (chemo-)radiotherapy,” *Physics in Medicine and Biology*, 2017.
- [38] C. Sullivan and A. Kaszynski, “PyVista: 3D plotting and mesh analysis through a streamlined interface for the Visualization Toolkit (VTK),” *Journal of Open Source Software*, 2019.

- [39] D. Mason, “SU-E-T-33: Pydicom: An Open Source DICOM Library,” in *Medical Physics*, 2011.
- [40] S. Van Der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu, “Scikit-image: Image processing in python,” *PeerJ*, 2014.
- [41] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, A. Vijaykumar, A. P. Bardelli, A. Rothberg, A. Hilboll, A. Kloeckner, A. Scopatz, A. Lee, A. Rokem, C. N. Woods, C. Fulton, C. Masson, C. Häggström, C. Fitzgerald, D. A. Nicholson, D. R. Hagen, D. V. Pasechnik, E. Olivetti, E. Martin, E. Wieser, F. Silva, F. Lenders, F. Wilhelm, G. Young, G. A. Price, G. L. Ingold, G. E. Allen, G. R. Lee, H. Audren, I. Probst, J. P. Dietrich, J. Silterra, J. T. Webber, J. Slavič, J. Nothman, J. Buchner, J. Kulick, J. L. Schönberger, J. V. de Miranda Cardoso, J. Reimer, J. Harrington, J. L. C. Rodríguez, J. Nunez-Iglesias, J. Kuczynski, K. Tritz, M. Thoma, M. Newville, M. Kümmerer, M. Bolingbroke, M. Tartre, M. Pak, N. J. Smith, N. Nowaczyk, N. Shebanov, O. Pavlyk, P. A. Brodtkorb, P. Lee, R. T. McGibbon, R. Feldbauer, S. Lewis, S. Tygier, S. Sievert, S. Vigna, S. Peterson, S. More, T. Pudlik, T. Oshima, T. J. Pingel, T. P. Robitaille, T. Spura, T. R. Jones, T. Cera, T. Leslie, T. Zito, T. Krauss, U. Upadhyay, Y. O. Halchenko, and Y. Vázquez-Baeza, “SciPy 1.0: fundamental algorithms for scientific computing in Python,” *Nature Methods*, 2020.
- [42] T. Lewiner, H. Lopes, A. W. Vieira, and G. Tavares, “Efficient Implementation of Marching Cubes’ Cases with Topological Guarantees,” *Journal of Graphics Tools*, 2003.
- [43] T. J. Baker, “Automatic mesh generation for complex three-dimensional regions using a constrained Delaunay triangulation,” *Engineering with Computers: An International*

Journal of Simulation-Based Engineering, 1989.

- [44] S. M. Bentzen, W. Dörr, R. Gahbauer, R. W. Howell, M. C. Joiner, B. Jones, D. T. Jones, A. J. Van Der Kogel, A. Wambersie, and G. Whitmore, “Bioeffect modeling and equieffective dose concepts in radiation oncology-Terminology, quantities and units,” *Radiotherapy and Oncology*, 2012.
- [45] J. Ulén, P. Strandmark, and F. Kahl, “Shortest Paths with Higher-Order Regularization,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015.
- [46] C. Chen, M. Witte, W. Heemsbergen, and M. V. Herk, “Multiple comparisons permutation test for image based data mining in radiotherapy,” *Radiation Oncology*, 2013.
- [47] H. S. Gabryś, F. Buettner, F. Sterzing, H. Hauswald, and M. Bangert, “Design and selection of machine learning methods using radiomics and dosiomics for normal tissue complication probability modeling of xerostomia,” *Frontiers in Oncology*, 2018.

3.9 Supplemental Materials

Table S-1: Overview of DSM methods published in the literature

Citation		Organ Studied	Software Used	DSM style	Centerline type	Slice Determination	Points on slices	Cut location	Slice Spacing	Samples per slice
1	Wang <i>et al.</i>	Colon	Custom IDL code	Non-planar	spline fit to n points	Electric field model	Equi-angular	Unspecified	unclear	unclear
2	Hoogeman <i>et al.</i>	Rectum	unclear	Non-planar	Central axis	Orthogonal to centerline	Equi-distant	Posterior to slicewise COM	3.5 mm	50
3	Meijer <i>et al.</i>	Rectum	unclear	Biomechanical model-based	From FE modeling	From FE modeling	From FE modeling	unclear	unclear	unclear
4	Heemsbergen <i>et al.</i>	Anorectum	unclear	Non-planar	Central axis	Orthogonal to centerline	Equi-distant	Posterior to slicewise COM	5 mm	unclear
5	Tucker <i>et al.</i>	Rectum	unclear	Planar	slicewise centroid	Parallel, double the slices from CT image	Equi-angular	Posterior-most per slice	1.5 mm (half the slice thickness)	72
6	van Lin <i>et al.</i>	Rectum	Unclear. May have replicated methods of Tucker <i>et al.</i> based on use of in-text citations							
7	Munbodh <i>et al.</i>	Rectum	unclear	Non-planar	From FE modeling	From FE modeling	From FE modeling	Custom method	unclear	unclear
8	Buettner <i>et al.</i>	Rectum	unclear	Planar	Slicewise centroid	On CT slices	Equi-distant	Posterior-most per slice	5 mm	21
9	Buettner <i>et al.</i>									
10	Buettner <i>et al.</i>									
11	Murray* <i>et al.</i>	Rectum	unclear	Planar	Slicewise centroid	On CT slices	Equi-distant	Posterior-most per slice	Slice thickness (unspecified)	21
12	Scaife <i>et al.</i>	Rectum	MATLAB	Planar	Slicewise centroid	On CT slices	Equi-distant	Posterior to slicewise COM	Slice thickness	21

									(unspecified)	
13	Wortel et al.	Anorectum	unclear	Non-planar	Central axis	Orthogonal to centerline (#:50)	unclear	Unspecified "posterior" point	Normalized to 50 slices	45
14	Witztum et al.	Duodenum, Esophagus	MATLAB (public on Github)	Non-planar	"racecar path" central axis	Orthogonal to centerline (#: 200)	Equi-angular	Custom method for duodenum	Normalized to 200 slices	30
15	Palorini et al.	Bladder	VODCA	Planar	Slicewise centroid	Parallel, not on CT slices (max 25)	unclear	Anterior to global COM	1 mm	unclear
16	Moulton et al.	Rectum	MATLAB + CERR	Planar	Slicewise centroid	On CT slices (but then normalized)	Equi-angular	Posterior to slicewise COM	Slice thickness (unspecified)	45
17	Casares-Magaz et al.	Rectum	unclear	Planar	Slicewise centroid	On CT slices	Equi-distant for 8 subregions	Posterior to slicewise COM	Slice thickness (unspecified)	200
18	Shelley et al.	Rectum	MATLAB? (unclear)	Planar	Slicewise centroid	On CT slices	equi-distant	Posterior to slicewise COM	3 mm	unclear
19	Onjukka et al.	Anorectum	VODCA	unclear	unclear	unclear	unclear	Posterior to global COM	unclear	200
20	Vanneste et al.	Anorectum	unclear	Planar	slicewise centroid	On CT slices	equi-angular	Posterior-most per slice	Slice thickness (unspecified)	100
21	Reijtenbagh et al.	Rectum	unclear	Non-planar	Central axis	Orthogonal to centerline	equi-distant	Unspecified "posterior" point	unclear	unclear
22	Serban et al.	Vagina	MATLAB	Non-planar	"fanned path" central axis	Orthogonal to centerline	equi-distant for 3 subregions	Anterior to slicewise COM	unclear	360
23	McWilliam et al.	Heart	unclear	Planar	Slicewise centroid	Parallel, not on CT slices (#: 50)	equiangular	Anterior to global COM	Normalized to 50 slices	360

24	Heemsbergen <i>et al.</i>	Rectum	unclear	Non-planar	Central axis	Orthogonal to centerline	equiangular	unclear	Normalized to 100 slices	90
25	Chen** <i>et al.</i>	Rectum	unclear	unclear	unclear	unclear	Equi-distant (1 mm)	Posterior-most per slice	1 mm	variable
26	Shelley <i>et al.</i>	Rectum	Abaqus + MATLAB	Non-planar	From FE modeling	From FE modeling	From FE modeling	Posterior to slice-wise COM	Normalized to 80 slices	30
27	Mylona <i>et al.</i>	Bladder	VODCA	Planar	unclear	Parallel, not on CT slices (unspec. #)	unclear	Anterior to global COM	Normalized to set # slices	unclear
28	Yahya <i>et al.</i>	Bladder	unclear	Planar	slicewise centroid	Parallel, not on CT slices (max 45)	unclear	Anterior to global COM	1 mm	201
29	Improta <i>et al.</i>	Bladder	VODCA	Planar	Slicewise centroid	Parallel, not on CT slices (max 25)	unclear	Anterior to global COM	1 mm	unclear
30	Dankers <i>et al.</i>	Esophagus	unclear	Planar	Slicewise centroid	On CT slices	equiangular	Unspecified "posterior" point	3 mm	360

* Conference abstract

** Description of DSM calculation methods provided as supplementary material instead of in paper body

References

- 1) Wang *et al.* 1999. <https://doi.org/10.1109/42.700745>
- 2) Hoogeman *et al.* 2004. <https://doi.org/10.1016/j.radonc.2003.11.015>
- 3) Meijer *et al.* 1999. [https://doi.org/10.1016/S0360-3016\(99\)00270-9](https://doi.org/10.1016/S0360-3016(99)00270-9)
- 4) Heemsbergen *et al.* 2005. <https://doi.org/10.1016/j.ijrobp.2004.07.724>
- 5) Tucker *et al.* 2006. <https://doi.org/10.1016/j.ijrobp.2005.10.029>
- 6) van Lin *et al.* 2007. <https://doi.org/10.1016/j.ijrobp.2006.09.034>
- 7) Mundodh *et al.* 2008. <https://doi.org/10.1118/1.2907707>
- 8) Buettner *et al.* 2009. <https://doi.org/10.1088/0031-9155/54/17/005>
- 9) Buettner *et al.* 2009. <https://doi.org/10.1088/0031-9155/54/21/006>
- 10) Buettner *et al.* 2011. <https://doi.org/doi:10.1088/0031-9155/56/7/013>
- 11) Murray *et al.* 2014. <https://doi.org/10.1118/1.4888065>
- 12) Scaife *et al.* 2015. <https://doi.org/10.1259/bjr.20150243>
- 13) Wortel *et al.* 2015. <https://doi.org/10.1016/j.radonc.2015.10.020>
- 14) Witzum *et al.* 2016. <https://doi.org/10.1118/1.4964790>
- 15) Palorini *et al.* 2016. <https://doi.org/10.1016/j.radonc.2016.02.025>
- 16) Moulton *et al.* 2017. <https://doi.org/10.1088/1361-6560/aa663d>
- 17) Casares-Magaz *et al.* 2017. <https://doi.org/10.1080/0284186X.2017.1370130>
- 18) Shelley *et al.* 2017. <https://doi.org/10.1016/j.radonc.2017.04.008>
- 19) Onjukka *et al.* 2019. <https://doi.org/10.1080/0284186X.2019.1635267>
- 20) Vanneste *et al.* 2019. <https://doi.org/10.1016/j.ctro.2018.10.006>
- 21) Reijtenbagh *et al.* 2020. <https://doi.org/10.1016/j.radonc.2020.03.035>
- 22) Serban *et al.* 2020. <https://doi.org/10.1016/j.brachy.2021.02.004>
- 23) McWilliam *et al.* 2020. <https://doi.org/10.1016/j.phro.2020.07.002>
- 24) Heemsbergen *et al.* 2020. <https://doi.org/10.3389/fonc.2020.00469>
- 25) Chen *et al.* 2018. <https://doi.org/10.1186/s13014-018-1068-0>
- 26) Shelley *et al.* 2020. <https://doi.org/10.1016/j.phro.2020.05.006>
- 27) Mylona *et al.* 2020. <https://doi.org/10.1016/j.radonc.2020.02.028>
- 28) Yahya *et al.* 2016. <https://doi.org/10.1016/j.ijrobp.2016.10.024>
- 29) Improta *et al.* 2016. <https://doi.org/10.1016/j.ejimp.2016.08.013>
- 30) Dankers *et al.* 2017. <https://doi.org/10.1088/1361-6560/aa5e9e>

Figure S-2: DSM features derived from Planar and Non-coplanar DSMs

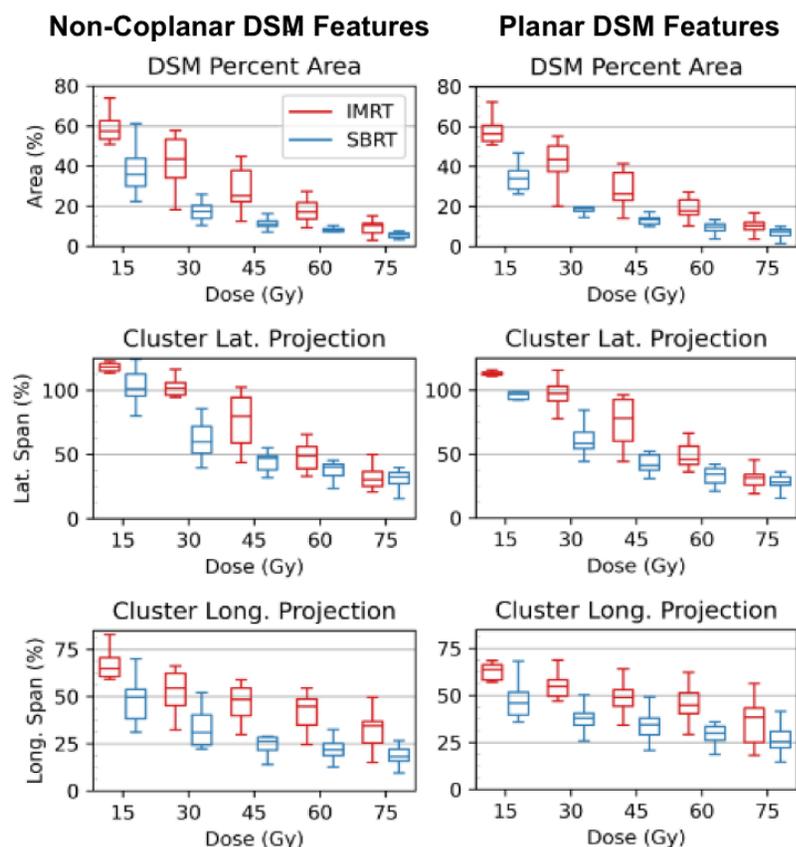


Figure S-2: Boxplots of the dose features of the IMRT (red) and SBRT (blue) cohorts, converted to EQD2 Gy. Results for non-coplanar DSMs (as seen in the paper body) are shown on the left and results for planar DSMs (new) shown on the right.

Methods

Patient Datasets

- 20 patients prescribed 60 Gy in 20 fractions (IMRT)
- 20 patients prescribed 36.25 Gy in 5 fractions (SBRT)

Rectal DSM and DSM Feature Calculation

- DSMs were calculated twice: once with the planar and once with the non-coplanar method
- All DSMs were calculated using scaled vertical mapping (number of slices held constant between pts)
- Each DSM was composed of 30 slices sampled at 30 equiangular points (30x30 maps)
- DSMs were converted to EQD2_{Gy} using $\alpha/\beta = 2.3$
- Features were calculated using *rtdsm*'s *cluster_features* function

Chapter 4

More than one way to skin a dose volume: the impact of dose-surface map calculation approach on study reproducibility

Haley M Patrick and John Kildea

Manuscript submitted to *Physics in Medicine & Biology*.

4.1 Preface

In chapter 3, the construction of a flexible DSM-calculation software package was described. Here we use this package to examine consistency between DSMs calculated in different manners. Many previous studies have used a variety of different DSM calculation techniques to identify statistically significant spatial dose patterns predictive of various radiation toxicities, especially for the rectum and bladder. However, their findings are highly variable and few predictive spatial dose patterns have been reproduced between different research groups. Understanding the reason for this lack of reproducibility and its relation to DSM calculation technique is important to determine if standardization may be

required for this research field.

Here we quantitatively compare rectum and bladder DSMs produced using different calculation techniques to determine the role of calculation technique in study reproducibility. First, we directly compared DSMs produced by different techniques using a self-developed paired implementation of the popular multiple-comparisons permutation test used in DSM research [1]. Second, we evaluated the equivalence of DSM features calculated for the different DSMs. Finally, we assessed the stability of cohort comparison results to DSM calculation technique to determine whether choice of technique can influence the results of a DSM study. The results of our analysis show that DSM calculation technique can significantly impact the information presented in a DSM. This finding indicates that inter-researcher calculation variations are likely contributing to the limited reproducibility in the field and confirm that discussions on standardization efforts are warranted.

4.2 Abstract

Objective: Dose-surface maps (DSMs) provide spatial representations of the radiation dose to organ surfaces during radiotherapy and are a valuable tool for identifying dose deposition patterns that are predictive of radiation toxicities. Over the years, many different DSM calculation approaches have been introduced and used in dose-outcome studies. However, little consideration has been given to how these calculation approaches may be impacting the reproducibility of studies in the field. Therefore, we conducted an investigation to determine the level of equivalence of DSMs calculated with different approaches and their subsequent impact on study results.

Approach: Bladder and rectum DSMs were calculated for 20 prostate radiotherapy patients using combinations of the most common slice orientation and spacing styles in the literature. Equivalence of differently calculated DSMs was evaluated using pixel-wise comparisons and DSM features (rectum only). Finally, mock cohort comparison studies were conducted with DSMs calculated using each approach to determine the level of dosimetric study reproducibility between calculation approaches.

Main Results: We found that rectum DSMs calculated using the planar and

non-coplanar orientation styles were non-equivalent in the posterior rectal region and that equivalence of DSMs calculated with different slice spacing styles was conditional on the choice of inter-slice distance used. DSM features were highly sensitive to choice of slice orientation style and DSM sampling resolution. Finally, while general result trends were consistent between the comparison studies performed using different DSMs, statistically significant subregions and features could vary greatly in position and magnitude.

Significance: We have determined that DSMs calculated with different calculation approaches are frequently non-equivalent and can lead to differing conclusions between studies performed using the same dataset. We recommend that the DSM research community work to establish consensus calculation approaches to ensure reproducibility within the field.

4.3 Introduction

Proper understanding of the dose-outcome responses of normal tissues is essential in order to be able to design radiotherapy treatment plans that minimize the likelihood of radiation toxicity. Traditionally, dose-volume histograms (DVHs) have been the primary tool used to derive dose-outcome relationships and dosimetric constraints for organs at risk (OARs) in radiotherapy research studies. These constraints may end up used in clinical practice to guide and evaluate the quality of individual treatment plans [2, 3]. However, DVH-based dose-outcome models lack spatial information and assume OARs have homogenous radiation sensitivities, potentially masking the existence of important radiosensitive subregions [4, 5]. Therefore, for certain OARs, alternative dose-outcome analysis tools are of interest to the radiation oncology community.

One alternative to the DVH that preserves spatial information is the dose-surface map (DSM): a 2D projection of the dose to an organ's 3D surface. DSMs have mainly been used to study dose to the rectum and bladder [6–8], though several studies have also been published for other hollow organs such as the vagina, esophagus, duodenum, and heart [9–11]. To date, DSMs have been used to identify spatial dose features and organ subregions predictive of early and late toxicities. In some cases, DSMs have been shown to be more predictive of

radiation toxicities than DVHs [4, 8, 12, 13].

Although promising as a dosimetric tool, it is important to note that the published methods of calculation and analysis of DSMs are much more diverse than is the case for DVHs. While nearly all DSMs are created by (1) defining slices of the organ of interest, (2) defining points around the surface of each slice to sample dose at, and (3) cutting open and unfurling the surface to create a 2D dose map, individual DSM implementations may use different approaches for each step. For instance, the DSM slices may all be oriented parallel to those of the treatment planning image (planar slicing) [6, 7, 14] or individually angled such that each slice is orthogonal to the organ's central axis path (non-coplanar slicing) [15, 16]. Slices may also be separated using a set spacing for all patients (fixed spacing) [8], or with different spacing for each patient to ensure all DSMs contain the same number of slices (scaled spacing) [6, 13]. Analysis techniques are similarly diverse, with different groups comparing DSMs either in a pixel-wise manner or based on features. This diversity of calculation and analysis approaches can make it difficult to compare results between research studies and may be impacting the reproducibility of results in the field.

To date, the only DSM-based toxicity metrics that have been reproduced in the literature have been for late rectal bleeding [6, 17, 18] and late bladder dysuria [13, 19], despite many unconfirmed reports of other predictive metrics. Although cohort effects may play a role in the lack of reproducibility across studies, it is possible that variations in DSM calculation approaches may also be responsible. Determining the influence of calculation approaches on DSM-based findings is important, not only to help facilitate the consolidation of findings across DSM studies to firmly establish spatially-informed dosimetric constraints, but also to determine how dependent the clinical validity of these constraints is on the level of concordance between the DSM calculation approaches used in the clinic and in the research that was used to derive the constraints in the first place. With this in mind, the purpose of the present study was to determine the impact of DSM calculation approach on DSM topography and analysis for bladder and rectum structures. Specifically, we aimed to determine if:

1. Choice of slice orientation style has a statistically significant effect on DSM topography and features (rectum only);

2. Choice of slice spacing style has a statistically significant effect on DSM topography and features (rectum and bladder);
3. Choice of DSM calculation approach impacts the conclusions one draws when comparing the average DSMs of two different cohorts of patients to one another (rectum and bladder).

4.4 Methods

4.4.1 Patient Cohort

To evaluate the effect of DSM calculation methodology on DSM topology, our analyses for aims #1 and #2 were conducted at a population level, using a benchmark cohort of patients (to represent a retrospective research study), and at an individual representative patient level (to represent a single clinical case). The treatment plans of 20 moderate-risk prostate cancer patients treated at our centre between 2016 and 2017 were used as our retrospective patient cohort. One patient from the cohort with a rectum of median length was chosen as the representative patient. Simulation CT images, acquired on a Philips Big Bore CT scanner using a 3.0 mm slice thickness, were contoured according to RTOG guidelines for the male pelvis [20]. All patients were prescribed 60 Gy in 20 fractions to the prostate alone, plus 7.0 mm isotropic PTV margins, using a two-arc VMAT approach. Plans were generated in the Eclipse treatment planning system (Varian Medical Systems, Palo Alto, CA) using previously-published treatment-planning constraints [21].

In order to facilitate the investigation of aim #3, a second comparison cohort was artificially created by calculating new dose distributions for each patient using 5.0 mm isotropic PTV margins. This yielded a paired cohort with predictable dose distribution differences from the benchmark cohort, making for easy assessment of how the comparison of two cohorts is affected by DSM calculation approach.

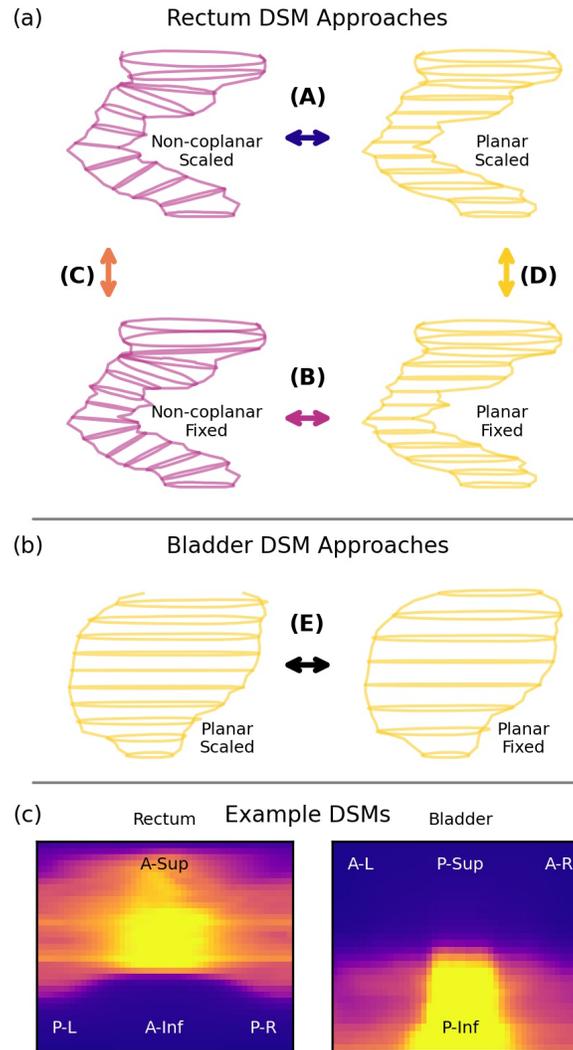


Figure 4.1: Visualization of all DSM calculation approaches and comparisons used in this work for (a) rectum and (b) bladder structures. Individual comparisons are indicated with capital letters and colors that are consistent across all figures in this article. Examples of typical bladder and rectum DSMs and how to interpret them are shown in (c).

4.4.2 DSM Calculation Workflow

As stated in the introduction, the two aspects of DSM calculation approach examined in this study were choice of slice orientation (planar or non-coplanar) and choice of slice spacing (scaled or fixed). Rectum DSMs were calculated using all four possible combinations of these aspects (Fig 4.1a), whereas bladder DSMs were calculated with planar slice orientation only and the two different slice spacing approaches (Fig 4.1b). This allowed us to reproduce the breadth of calculation approaches present in current DSM literature. All DSM calculations were performed using *rtdsm*, an open-source Python package previously developed and published by our research group for straightforward and accessible DSM calculation and analysis [22]. *rtdsm* can calculate planar or non-coplanar DSMs using the standard RT-Structure and RT-Dose files from a DICOM-RT-compliant radiotherapy treatment planning system as input. For the calculation of fixed-spacing DSMs, a slice separation of 3.0 mm (CT slice thickness) was used, whereas scaled-spacing DSMs fixed the total number of slices to the median number of CT slices (n_{slices}) the organ spanned for all patients in the cohort ($n_{slices} = 35$ for rectum, $n_{slices} = 25$ for bladder). When unfurling the surface doses to form the DSMs, rectum DSMs were cut open on the posterior side, and bladder DSMs on the anterior. These cut locations are typical in DSM research as they allow for the anticipated dose hotspots of these organs to be centered in their DSMs (Fig 4.1c). All DSMs used a sampling resolution of 45 equiangular points per slice.

4.4.3 DSM Analysis Technique

The average DSM of the benchmark cohort and the representative patient’s DSM for each calculation approach were calculated and compared between approaches. To enable direct visual comparison of the effects of DSM calculation approach, dose difference maps (DDMs) were calculated for each comparison shown in Figure 4.1 by subtracting the comparator DSMs. Because the DSMs in the benchmark cohort did not all contain the same number of slices when using fixed slicing, the average DSMs and DDMs in the population-level comparisons were truncated to the height of the shortest DSM in the

cohort. Differences between DSMs owing to the different calculation approaches were quantified in two ways: (1) using pixel-wise comparisons through multiple comparisons permutation (MCP) testing, which is a commonly-used method to compare dose maps, and (2) using feature-based comparisons, as is popular for rectum DSMs. This dual analysis was performed in order to make findings easily translatable to the existing body of literature.

Pixel-wise comparison has been used in both bladder and rectum DSM research to identify subregions of either organ where statistically-meaningful differences in dose exist between two cohorts. While pixel-wise DSM comparisons are possible with pixel-wise t-tests and can be used to identify general areas where dose differences exist, it is good practice to apply a correction for multiple comparisons, such as with MCP testing, to reduce sensitivity to false positives [1]. The standard MCP test determines the similarity of two unpaired image-type datasets of the same anatomy and identifies pixels that vary significantly between the datasets while accounting for pixel-wise variance. As the datasets in our study were paired, we developed a paired implementation of the MCP test by modifying the permutation process to keep the labels of data pairs linked. Full details are provided in [Supplement A](#).

Feature-based comparison is an analysis technique used to identify statistically-meaningful differences in isodose cluster characteristics that was developed for rectum DSMs [6, 14]. Features are derived by first creating a mask of a cluster of pixels for a given dose level and then extracting size, position, and shape metrics from either the mask itself or from an ellipse fitted to it, as initially performed by Buettner *et al* [6]. For this study, we opted to calculate the five most common features that have been reported in the literature: cluster area, cluster lateral and longitudinal extent, and ellipse lateral and longitudinal extent (Fig 4.2) for four dose levels: 15, 35, 45, and 55 Gy. These dose levels were selected as they covered the full dose range of our data, and they matched the dose levels of toxicity-predictive features reported in other 2 Gy-per-fraction (or equivalent) studies [6, 14, 23]. Once calculated, we compared features between the cohorts using Wilcoxon signed-rank testing.

Statistical significance for both types of analysis was defined as $p \leq 0.05$. A Bonferroni correction of 4.0 was applied when comparing pairs of rectum DSM features to reduce false positives. In addition to comparing DSMs between calculation approaches for the same

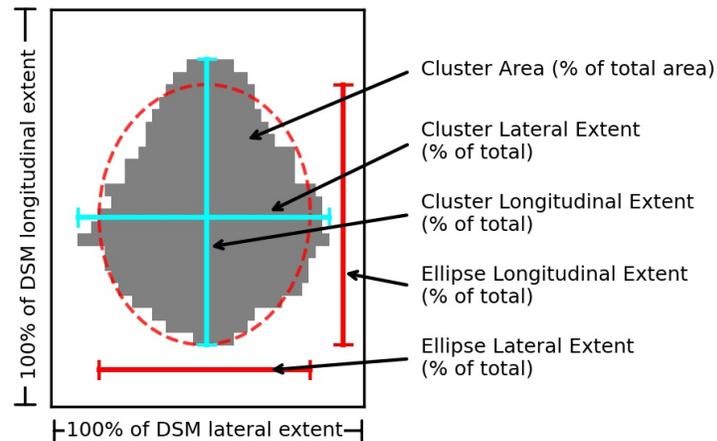


Figure 4.2: Visualization of the DSM features used in this work, calculated as originally defined in the literature [6, 14]. As is convention, each feature is reported as a percentage of either the total area or total dimensional extent (lateral or longitudinal) of the DSM used to calculate it.

cohort, the analysis was also employed to compare average DSMs between the benchmark cohort and the artificially-generated comparison cohort (7.0 mm versus 5.0 mm PTV margins) for each of the DSM calculation approaches investigated. For each calculation approach, the average DSMs of the two cohorts were compared to assess how DSM calculation approach affects the ability to examine dosimetric differences between distinct cohorts.

4.5 Results

4.5.1 Influence of Calculation Approach on Pixel-wise Comparisons

Aim 1: Influence of Slice Orientation Style

Side-by-side comparisons of the average rectum DSMs of the benchmark cohort produced using non-coplanar and planar slicing are shown in the first two rows of Figure 4.3. Differences in the shapes of the 35, 45 and 55 Gy isodose clusters are qualitatively

observable between the two slice orientation styles (first column compared to second column), illustrating the influence of slice angling on DSMs. When scaled spacing is used (i.e. constant number of slices, row A, purple), the two DSM styles disagree significantly on dose to the posterior-inferior wall as shown in the corresponding DDMs and p-value maps. Similarly, significant disagreement is also observed when using fixed spacing (row B, magenta), only this time it occurs over a much larger area of the superior-posterior wall. In both cases, as indicated by the DDMs, the disagreement appears to be caused by the non-coplanar DSMs measuring higher doses than their planar counterparts.

Similar patterns of disagreement are observable in the DSMs of the representative patient (Fig 4.4) and provide further insight into probable underlying causes. There is a clear fundamental difference in the shape of the moderate (35 Gy) and high dose (55 Gy) regions between the non-coplanar and planar slicing methods (first column compared to second column in Figures 4.3 and 4.4), which is likely related to slice angling and may explain the inferior differences. However, in the case of fixed spacing, an additional factor is introduced: difference in DSM length (i.e. number of slices) between the non-coplanar and planar calculation methods. The longer sampling path of the non-coplanar method requires more slices than the planar, stretching out the DSM and effectively desynchronizing information between the two styles the more superior in the DSM we go from the common inferior-most starting point. This may explain the difference between the non-coplanar and planar DSMs at the superior end when fixed spacing is used.

Aim 2: Influence of Slice Spacing Style

Comparisons of rectum DSMs created with fixed and scaled spacing are presented in columns C (orange) and D (yellow) of Figure 4.3, as well as for the representative patient in Figure 4.4. Corresponding DDMs (scaled minus fixed) and p-value maps are found in the rows beneath. Overall, slice spacing style has much less of an effect on rectum DSM topography than slice orientation style as evidenced by the lack of significant regions in the p-value maps. For the non-coplanar DSMs (column C, orange), the DDM suggests a shift or rescaling of the DSM topography in the superior-inferior direction between the two spacing styles. This rescaling effect is also visible in non-coplanar scaled-minus-fixed DDM of the representative

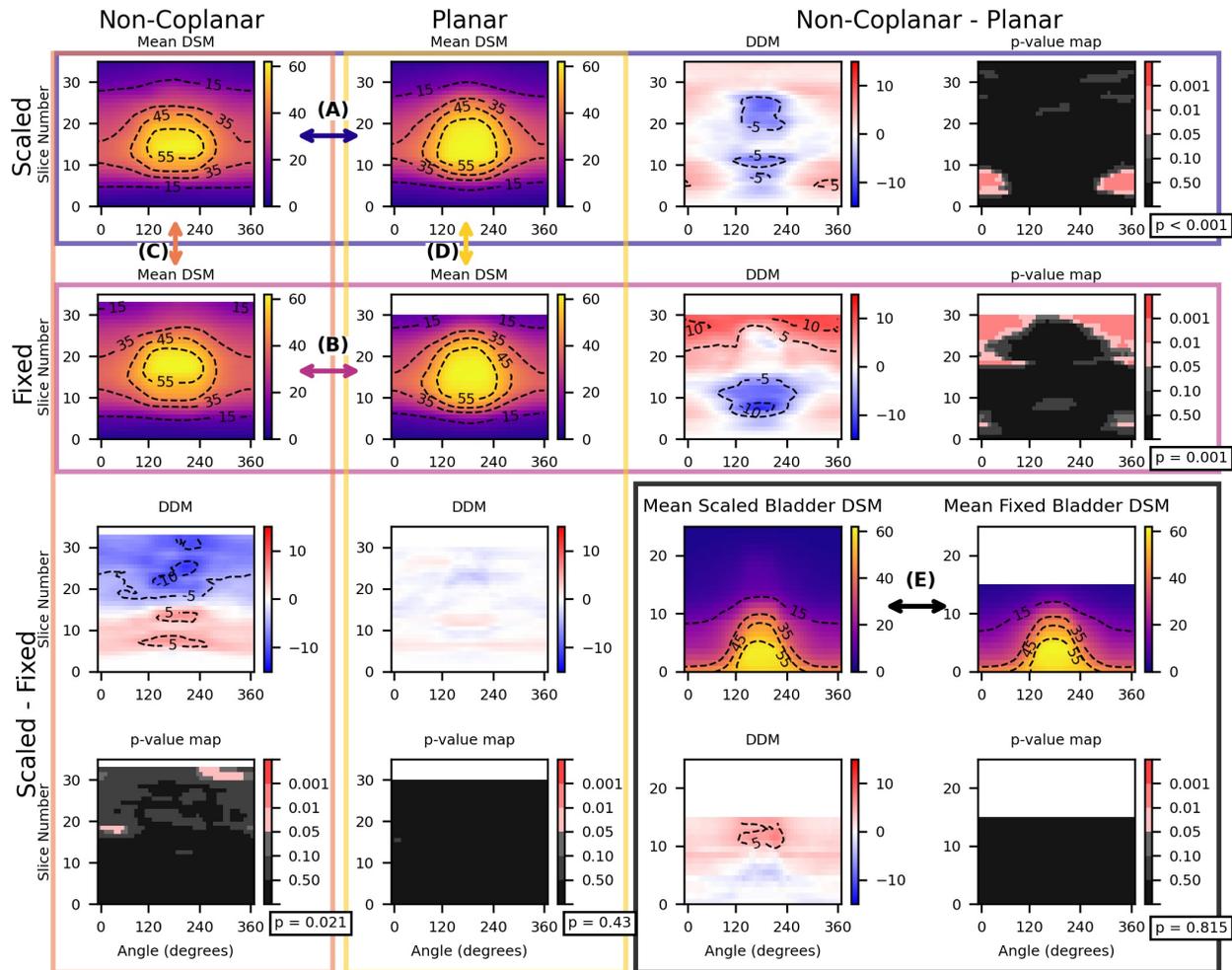


Figure 4.3: Quantitative comparisons of DSM calculation approaches. Each comparison is indicated as a row or column with its letter label and arrow from Figure 1, and a correspondingly colored box. Each comparison box contains: the cohort-mean DSMs of the two approaches being compared, their DDM, and the map of significantly different pixels between the two as determined by MCP testing. All maps are shown in units of Gy except for the MCP test result maps, which show p-values.

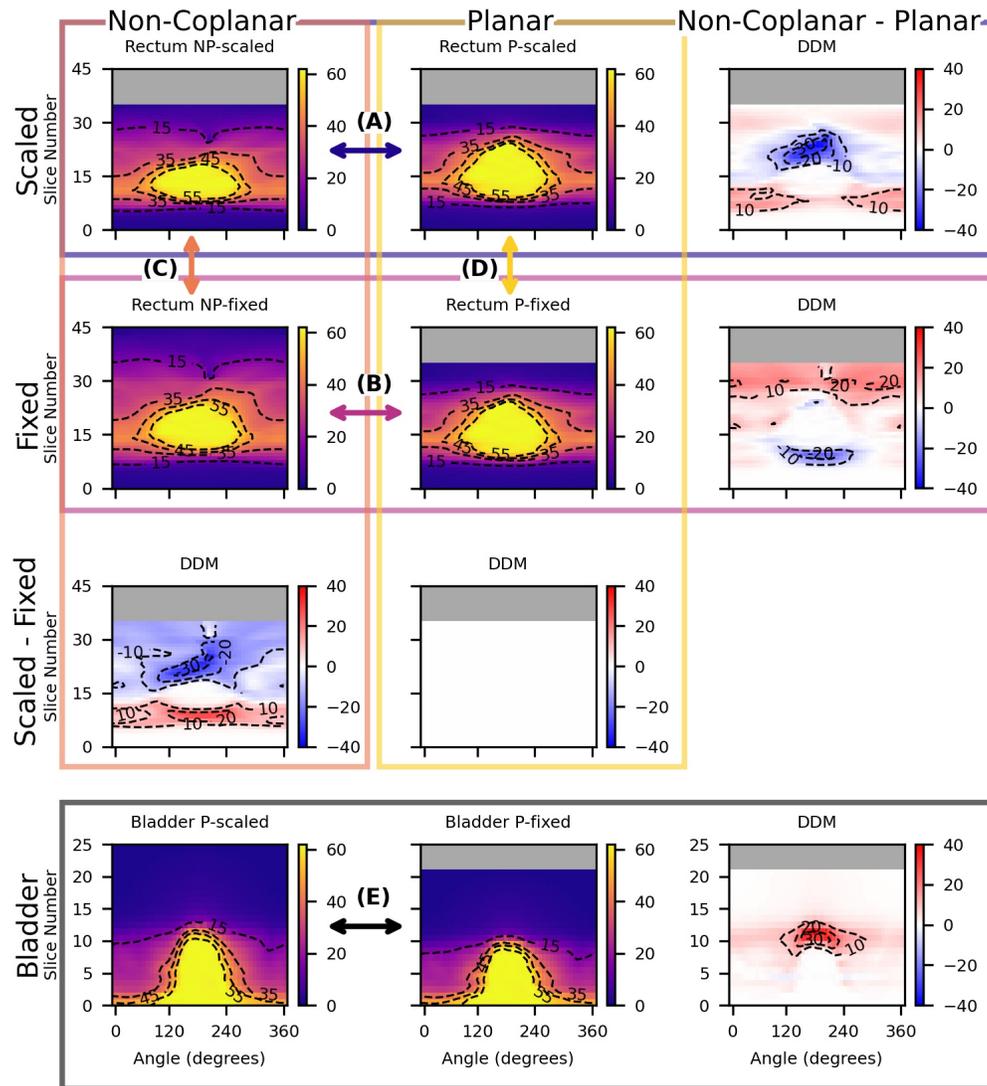


Figure 4.4: Comparisons of DSM calculation approaches for our representative patient. Each comparison is indicated with its respective label, arrow, and colored box and includes the individual DSMs calculated by each approach along with their DDM. All maps are in units of Gy.

patient (Fig 4.4C), which also clearly demonstrates how fixed slice spacing requires more slices to represent the entire rectum. Despite this clearly introducing a desynchronization effect, similar to that between non-coplanar and planar fixed spacing DSMs, significant disagreement is only found for two small patches of the superior rectum. Planar DSMs are even less affected by choice of slice spacing style, as no significant sites of disagreement were observed between the planar scaled and planar fixed DSMs. We note that, based on the DDMs of the representative patient, this result may have occurred due to our choice to use an n_{slices} for scaled DSMs that equaled the number of slices in our median rectum, causing the median effective slice distance to be approximately equal to the fixed DSM spacing (3 mm).

In the case of bladder DSMs, those created with scaled spacing contain noticeably more slices than those created with fixed spacing (Fig 4.3 and 4.4, box E), which can be attributed to the need to truncate the fixed spacing DSMs to the height of the shortest bladder in the cohort. Nevertheless, while the DDM suggests slice desynchronization akin to the rectum DSMs, no significant disagreement was observed between the two DSM slice spacing approaches across the slices they had in common. Once again, this is most likely attributed to our choice of n_{slices} for scaled spacing being similar to the number of slices required to construct a fixed spacing DSM for the average patient.

4.5.2 Influence of Calculation Approach on DSM Features (Aims 1 and 2)

DSM features were calculated and compared for rectum DSMs for all four calculation approaches, and the results for the 15 and 55 Gy clusters are shown in Figure 4.5 (figures of the 35 and 45 Gy clusters, and the representative patient are available in Supplement B). Features differ between calculation approaches, particularly between the planar (magenta) and non-coplanar (yellow) slicing styles. Longitudinal features are generally larger for planar DSMs at higher dose levels (Fig 4.5c,e), consistent with the pixel-wise findings when comparing planar and non-coplanar mean DSMs (Fig 4.3), whereas they are larger for non-coplanar DSMs at the 15 Gy isodose level. Interestingly, this trend of

features being larger for non-coplanar DSMs at low doses also extends to ellipse lateral extent (Fig 4.5d). While pixel-wise comparisons of mean DSM isodose regions do not suggest particularly notable differences in lateral spans between slice orientation approaches (Fig 4.3), it is possible that choice of slice orientation approach subtly influences the shape of the clusters and fitted ellipses used in feature-based analysis.

While scaled and fixed DSM features are generally in more agreement with one another for a given slice orientation style, disagreements do occur for certain area and ellipse-based features (Fig 4.5a,d,e). This is somewhat unexpected as DSM features are conventionally reported as percentages and thereby should not disagree significantly when DSMs are only effectively rescaled. As the primary difference between our scaled and fixed DSMs was longitudinal sampling resolution (fixed: 3 mm, scaled: variable per patient), these results may indicate that DSM features may not be stable between different sampling resolutions.

4.5.3 Influence of Calculation Approach on the Conclusions of a Cohort Comparison (Aim 3)

Rectum and bladder doses were compared between the benchmark cohort with 7.0 mm margins and the comparison cohort with 5.0 mm margins using DSMs calculated with all calculation approach variants, and the consistency of the results was assessed across approaches. As expected from the study design, the cohorts were found to be statistically dissimilar using all styles of DSMs, but quantitative differences did exist between the comparisons provided by the different approaches.

For rectum DSMs, the higher dose ring present in all DDMs (7.0 mm minus 5.0 mm margins) changed in both shape and magnitude depending on calculation approach (Fig 4.6a-h). MCP testing p-value maps also indicated that the locations of statistically significant subregions (SSRs) depended on DSM type (Fig 4.6a). Rectum DSM features were relatively consistent between calculation approaches (Fig 57), with all four styles of DSM generally agreeing on whether or not a feature differed significantly between the two cohorts. However, feature magnitudes did differ between approaches, especially for the longitudinal features. For example, the mean difference in the 15 Gy cluster longitudinal

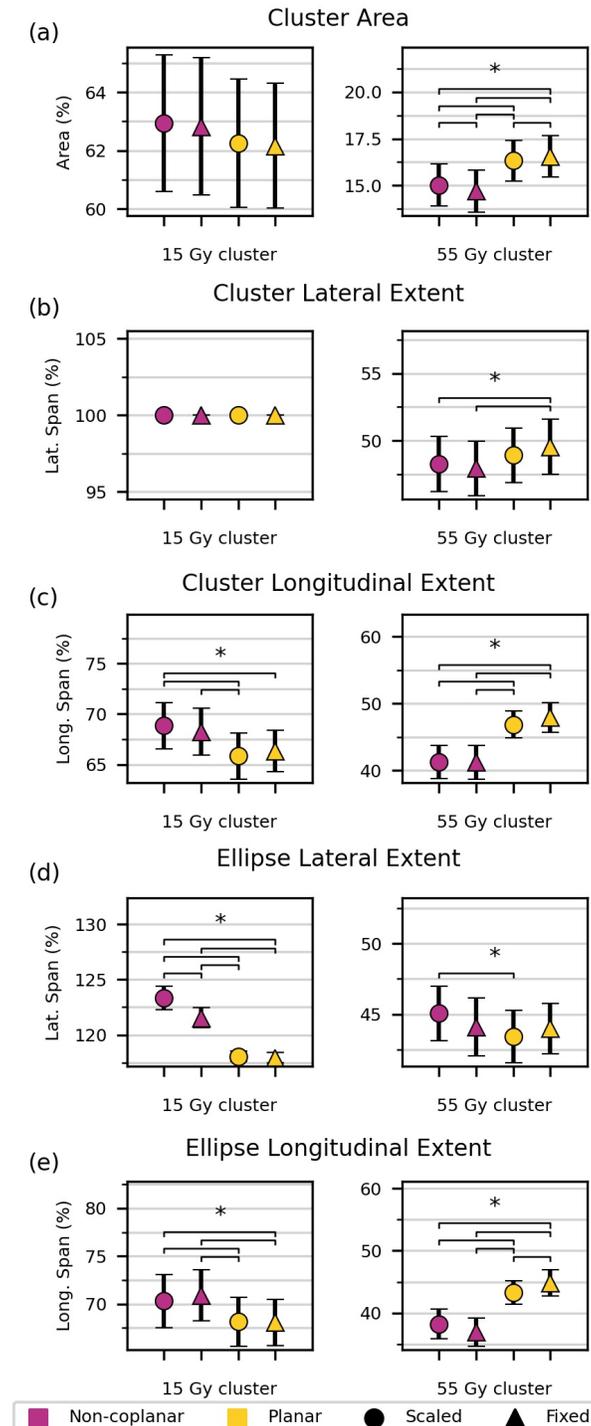


Figure 4.5: Mean DSM feature values (and standard uncertainty of the mean) from the four rectum DSM calculation approaches. In order, they are (a) cluster area, (b) cluster lateral extent, (c) cluster longitudinal extent, (d) ellipse lateral extent, and (e) ellipse longitudinal extent. Significantly different pairs of features are indicated with a bracket and an asterisk. Figures of the other dose levels are provided in the supplemental material.

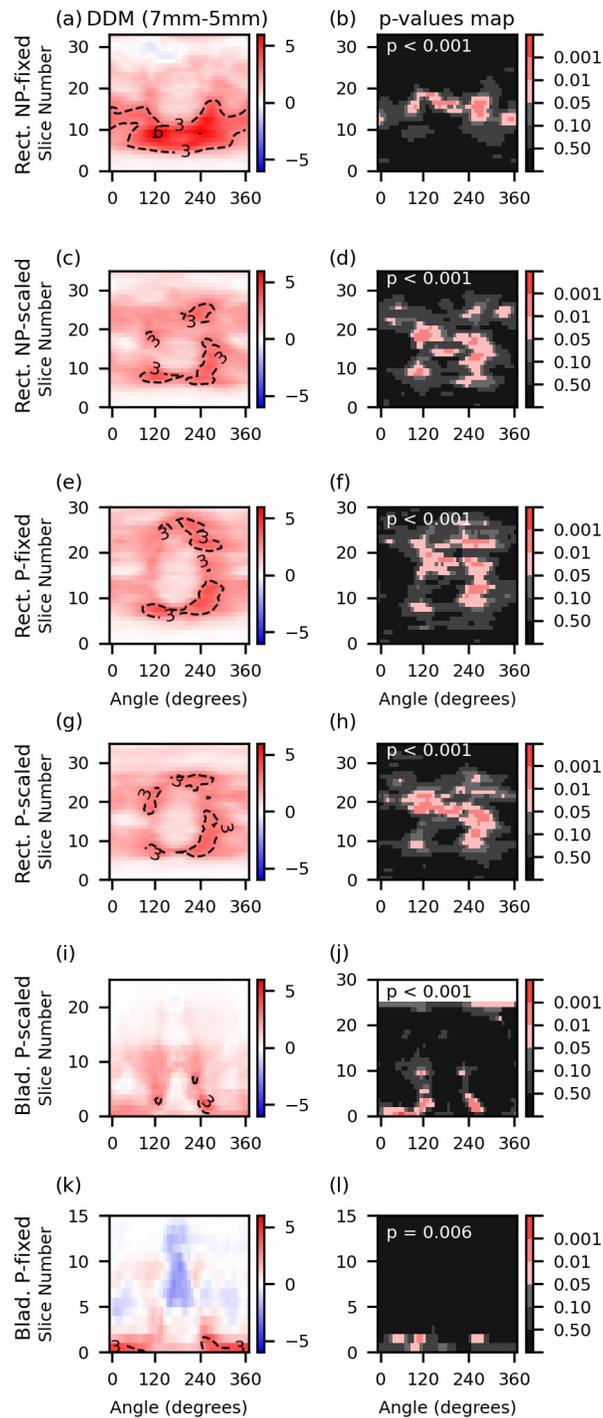


Figure 4.6: DDMs (left) and maps of significantly different pixels (right) between the benchmark and comparison cohorts for each of the DSM calculation approaches tested. “P” refers to planar slice orientations and “NP” to non-coplanar. DDMs are in units of Gy.

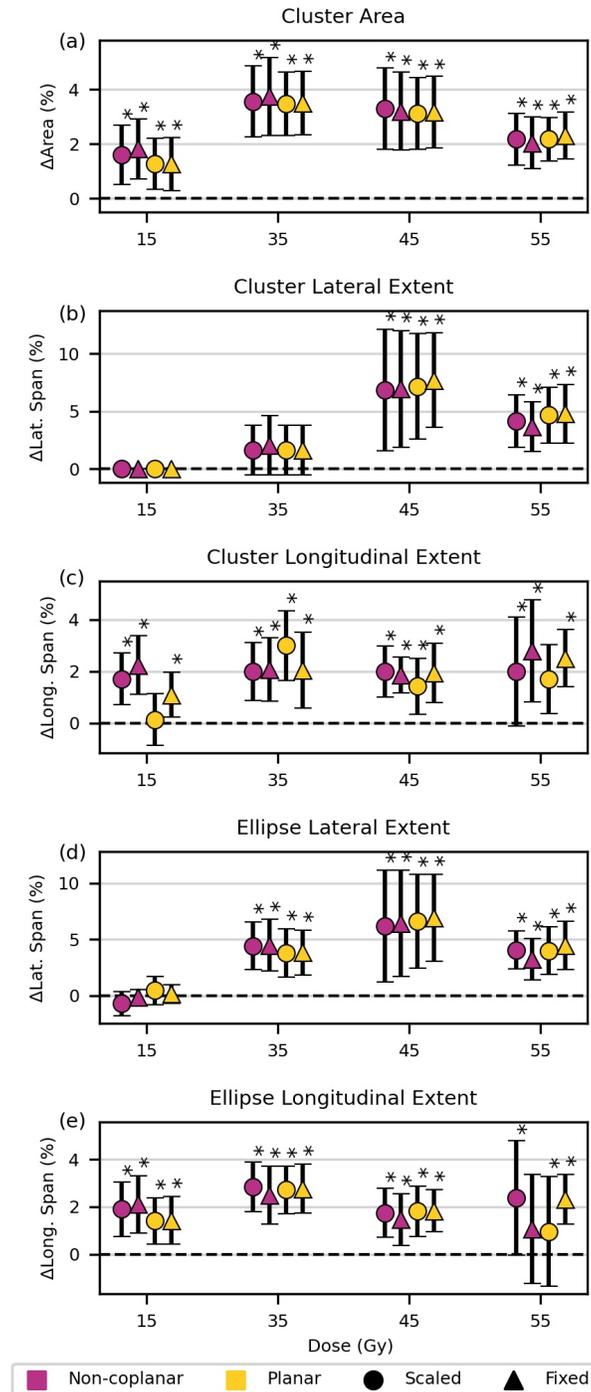


Figure 4.7: Mean differences in DSM feature values between the benchmark and comparison cohorts for each of the four rectum DSM calculation approaches. Error bars indicate 95% confidence intervals and statistically significant differences between the cohorts are indicated with asterisks.

extent for planar scaled DSMs was notably different in magnitude from the other DSM approaches (Fig 4.7c). Similarly, there was reduced consistency in ellipse longitudinal extent mean differences between DSM calculation approaches at the 55 Gy dose level (Fig 4.7e).

Bladder DSM findings were much less consistent between DSM styles when the two artificially different cohorts were compared. DDMs show cold spots in the fixed spacing comparison that are not present in the scaled spacing comparison (Fig 4.6i-l). While MCP testing does identify significantly differing pixels in the right and left inferior bladder for both DSM styles, the size and shape of these regions are different for the two calculation approaches. Tests conducted with the fixed-spacing dataset also failed to identify the SSR found by the scaled-spacing dataset in the superior bladder, likely a product of DSM truncation (Fig 4.6j,l).

4.6 Discussion

DSM calculation approaches are diverse and can differ considerably between research groups. Although all DSM research studies share the same general goal of identifying dosimetric spatial factors that are predictive of radiation toxicities, little to no work has been done to assess the reproducibility of these spatial factors between different DSM calculation approaches. To the best of our knowledge, this paper is the first quantitative investigation of analysis sensitivity to DSM calculation approach. We have identified that significant disagreement between DSMs can occur when different calculation approaches are used. In the discussion of our findings below, we refer to Figures 4.8 and 4.9, in which we have attempted to graphically illustrate the influences of the DSM calculation approaches we examined in this work.

4.6.1 Aim 1: Equivalence of planar and non-coplanar DSMs

Although they have existed side-by-side in the literature for nearly two decades, we found that rectum DSMs calculated using planar and non-coplanar slicing approaches are non-equivalent in two specific regions. Namely, the inferior-posterior wall when using scaled

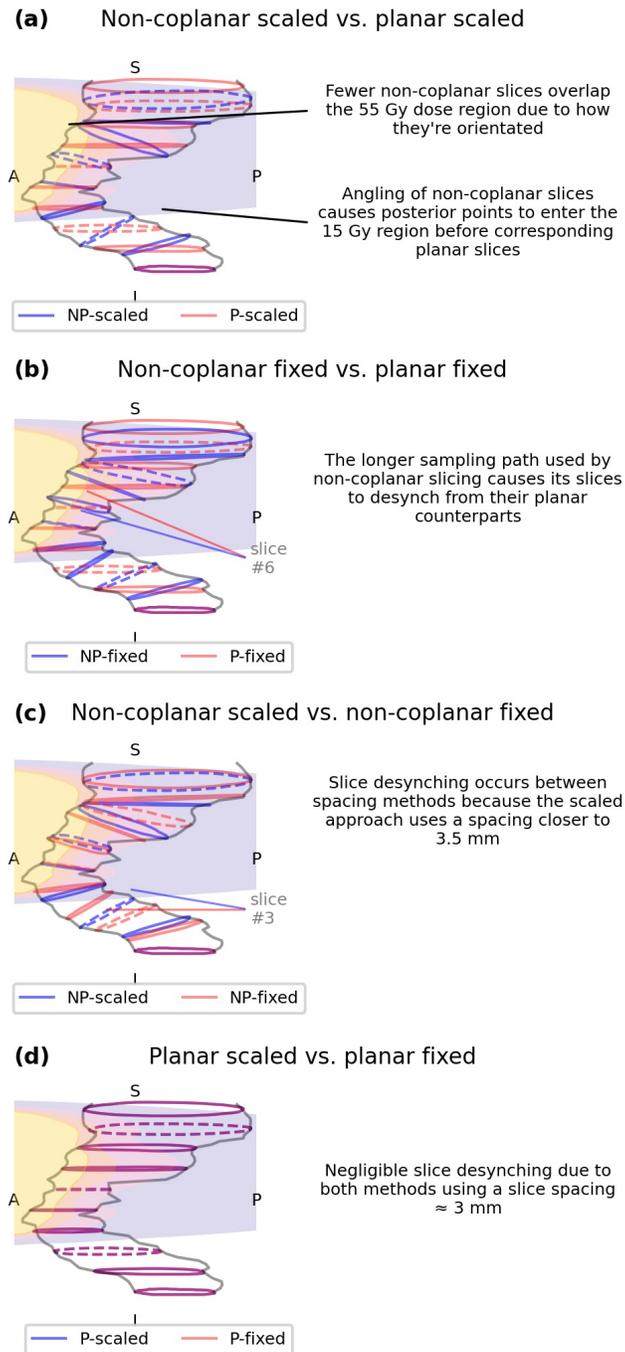


Figure 4.8: Examples of how the choice of DSM calculation approach influences the slices of the rectum (and subsequent appearance of the DSM) of the representative patient. Subfigures cover each comparison made in this work. The 15, 35, and 55 Gy isodose regions are included as purple, pink, and yellow overlays, respectively, for dosimetric context. Additionally, for visual clarity, only every fourth slice is shown. Dashed lines are used to represent every 12th slice to make visual comparisons of relative slice positions between slicing styles easier.

slice spacing and the superior-posterior wall when using fixed spacing (Fig 4.3). These non-equivalent regions stem from differences in sampling point locations introduced by the slicing methods themselves, causing the same slice to sample dose in different locations between the two approaches (as illustrated in Fig 4.8a-b). In general, one can expect that the less linear a rectum's path is the more its planar and non-coplanar DSMs will disagree. This would suggest that although the rectum has sometimes been considered a relatively simple organ that can be represented with planar slices [10], this cannot be assumed to be the case without careful review of all rectum structures in a given cohort. For this reason, we recommend the use of non-coplanar DSM slicing approaches where possible and to be aware and investigate the influence of the planar approach on DSM topology before proceeding otherwise.

4.6.2 Aim 2: Equivalence of scaled and fixed slice spacing DSMs

In our testing, we found that scaled and fixed slice spacing approaches were roughly equivalent to one another for our chosen comparison scenarios, wherein the fixed and median effective scaled slice spacings were both approximately equal to 3 mm. As some evidence of slice desynching was still observed between spacing approaches (Fig 4.3C,E), we suspected that these results may be conditional on our choice of matching slice spacing distances and could change if we used a different combination (Fig 4.8c-d). To demonstrate the possible influence of this, Figure 4.9 shows what the results of a comparison of fixed bladder DSMs (with a spacing of 3 mm) versus scaled bladder DSMs with 5 additional slices (meaning a median effective slice spacing 2.5 mm instead of 3 mm) would be. As shown, when effective slice spacing resolutions are not equal slice desynching is much greater, causing the DSMs to be more dissimilar. Based on this apparent sensitivity, we would heavily advise against direct comparisons of DSMs that use different slice spacing resolutions and encourage relative, scaled comparisons instead. We note, however, that the superior-inferior limits of DSMs should match when doing this in order to ensure anatomical alignment is maintained.

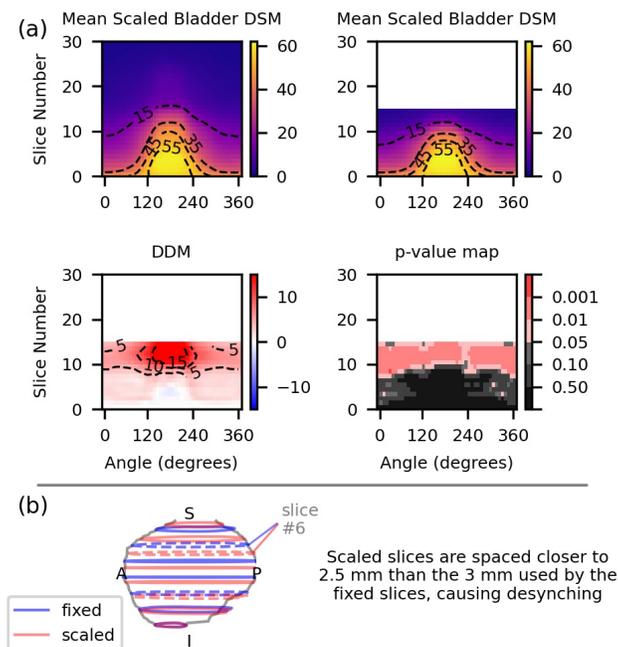


Figure 4.9: (a) Comparison of scaled bladder DSMs calculated with 30 scaled spacing slices (median 2.5 mm) and fixed slice spacing DSMs. Mean DSMs are shown for both approaches along with their DDM and p-value map. (b) Visual demonstration of the slice desynchroning that occurs between scaled and planar bladder DSMs. Only every 4th slice is shown for visual clarity, with every 12th slice represented with dashed lines.

4.6.3 Aim 3: DSM Feature robustness again DSM calculation method

Rectal DSM features were originally designed for planar DSMs and their specific patterns of dose topography [6]. As such, it is not surprising that features calculated from non-coplanar DSMs are non-equivalent to their planar counterparts (Fig 4.5). Due to their sampling approach, non-coplanar DSMs have smaller high dose isodose clusters and less elliptical low dose isodose clusters (Fig 4.4), impacting the calculation of ellipse-based features. For these reasons, we recommend that DSM features continue to be calculated from planar DSMs only to facilitate reproducibility. However, slice spacing approach must also be considered. We observed significant differences between features calculated with fixed and scaled slice spacing approaches, suggesting a possible resolution effect as well. This is concerning, as DSM resolution is one of the most variable factors between studies. From what we have observed in the literature, reported resolutions vary from 21×21 to 200×200 pixels [6, 23] and can be achieved through either direct sampling [7, 8] or interpolation methods [14, 18, 24], which may introduce their own effects. Because of this, we strongly recommend that new feature-based studies choose DSM sampling resolutions that are consistent with the previous studies they plan to compare to.

4.6.4 Relation to current state of reproducibility in the literature

To date, few studies in the DSM literature have agreed on what DSM information is predictive of radiation toxicities. Most discussions focus more on general trends that persist across studies, such as increased dose to the posterior rectum [14, 16] and the bladder trigone [8] causing increasing toxicity risk, and usually point to cohort effects to explain differences [13]. While the impacts of cohort, fractionation scheme, and analysis techniques on reproducibility cannot be discounted, our findings highlight that is also important to consider DSM calculation approaches when trying to understand the similarities (or dissimilarities) between published results.

Reproduced rectal DSM-toxicity results exist only for rectal bleeding and are limited to the 51 Gy cluster area [6, 14], the 40-60 Gy lateral ellipse extents [6, 7], and dose to the

inferior quarter of the rectum [14, 17, 18]. Corroborated feature results were all obtained using planar DSMs with resolutions between 21×21 and 51×45 pixels, whereas papers using higher resolution DSMs ($\geq 200 \times 200$) reported no reproduced features [23, 24]. It is also worthwhile to note that the reproduced rectal bleeding SSR was located inferiorly, where slice desynching effects are expected to be minimal between the two different slice orientation styles used by these authors (planar and non-coplanar). In contrast, other non-reproduced SSRs, like those for proctitis [14, 16, 17] and incontinence [17, 18, 23], were distributed in different locations in the superior half of the rectum, which we have observed to be more prone to slice desynching between calculation methods (Fig 4.8).

Bladder DSM SSR reproducibility has been limited to dose to the inferior-anterior bladder being predictive of late dysuria and is the subject of a study by Mylona *et al* [13]. Once again, we note that this SSR is located in the inferior organ, where we would expect the least discordance between Mylona’s planar scaled slice spacing DSMs and the planar fixed slice spacing DSMs of other published studies [8, 19]. However, we would also like to highlight the two SSRs Mylona found for acute and late retention. These were located in the superior half of the bladder, in the region of more significant slice desynching and above the level at which the fixed slice spacing DSM cohorts of the other studies truncated the maximum extent of their bladders. This truncation handicapped the comparability of these studies and is a noteworthy example of why consensus DSM calculation approaches are needed.

Although variations in DSM calculation methodology can help explain the state of reproducibility in our field, we recognize other factors do need to be considered as well. In addition to commonly discussed cohort or analysis differences, it is worth noting the role that different outcome reporting metrics may play. Choice of toxicity scoring instrument also varies greatly between studies (e.g. CTCAE, UCLA-QoL, IPSS, custom patient-reported outcome measures, [14, 18, 25]), as do the timepoints at which outcomes are collected, especially for late effects (first timepoint range: 3-27 months, [14, 24]). Considering that toxicity scoring concordance has been shown to be limited between observers and scoring instruments [26, 27], further investigations into their effect are warranted.

4.7 Conclusion

We have determined that different DSM calculation approaches produce non-equivalent DSMs that can impact the conclusions of a given study. This has the potential to limit clinical translation of DSM-based research unless measures are taken. Ideally, the community should establish standardized methodologies to calculate DSMs for each organ of interest, and at a minimum better awareness of DSM non-equivalencies is needed. While further discussions within the community are required to establish any sort of consensus, we wish to present the following recommendations for consideration:

1. **A planar scaled slice spacing approach should be used to calculate DSMs of sphere-like organs.** This is especially true for organs, such as the bladder, that exhibit significant isotropic volume changes between subjects, as the scaling ensures the same anatomical regions are represented by the same DSM slices.
2. **Non-coplanar slicing should be used for tubular organs with significant curvatures.** This includes organs like the rectum and duodenum where planar slicing cannot accurately account for the organ's trajectory.
3. **The planar slicing approach is acceptable for straight tubular organs or specific situations for organs with curvature.** These include the esophagus or spinal cord, or rectum DSMs in the context of feature-based analysis.
4. **Consensus calculation approaches should be developed for each organ by the DSM research community.** While some approaches may be more straightforward to decide, such as for the bladder, discussions will be necessary for organs with more complex geometries like the rectum.
5. **Data sharing should be encouraged within the community to better evaluate study reproducibility between different DSM code implementations.** This could be facilitated through either the sharing of anonymized DICOM files, or arrays of surface vertices and dose matrices.

6. **Open-source DSM calculation codes should be encouraged. This can be facilitated by code-sharing sites such as Github.** Our own DSM calculation code, *rtdsm*, is an example.

4.8 Acknowledgements

H.M.P. acknowledges funding from the Fonds de recherche du Québec - Santé in the form of a Doctoral Training Award, as well as partial support by the CREATE Responsible Health and Healthcare Data Science (SDRDS) grant of the Natural Sciences and Engineering Research Council. J.K. acknowledges support from the Canada Foundation for Innovation John R. Evans Leaders Fund. Logistical support provided by the Medical Physics Department at the McGill University Health centre was invaluable for this work.

Bibliography

- [1] C. Chen, M. Witte, W. Heemsbergen, and M. V. Herk, “Multiple comparisons permutation test for image based data mining in radiotherapy,” *Radiation Oncology*, 2013.
- [2] S. M. Bentzen, L. S. Constine, J. O. Deasy, A. Eisbruch, A. Jackson, L. B. Marks, R. K. Ten Haken, and E. D. Yorke, “Quantitative Analyses of Normal Tissue Effects in the Clinic (QUANTEC): an introduction to the scientific issues.,” *International journal of radiation oncology, biology, physics*, vol. 76, pp. S3–9, mar 2010.
- [3] B. Emami, J. Lyman, A. Brown, L. Cola, M. Goitein, J. E. Munzenrider, B. Shank, L. J. Solin, and M. Wesson, “Tolerance of normal tissue to therapeutic irradiation,” *International Journal of Radiation Oncology, Biology, Physics*, vol. 21, pp. 109–122, may 1991.
- [4] O. Acosta, G. Drean, J. D. Ospina, A. Simon, P. Haigron, C. Lafond, and R. De Crevoisier, “Voxel-based population analysis for correlating local dose and rectal toxicity in prostate cancer radiotherapy,” *Physics in Medicine and Biology*, 2013.

- [5] D. A. Jaffray, P. E. Lindsay, K. K. Brock, J. O. Deasy, and W. A. Tomé, “Accurate accumulation of dose for improved understanding of radiation effects in normal tissue.,” *International journal of radiation oncology, biology, physics*, vol. 76, pp. S135–9, mar 2010.
- [6] F. Buettner, S. L. Gulliford, S. Webb, M. R. Sydes, D. P. Dearnaley, and M. Partridge, “Assessing correlations between the spatial distribution of the dose to the rectal wall and late rectal toxicity after prostate radiotherapy: An analysis of data from the MRC RT01 trial (ISRCTN 47772397),” *Physics in Medicine and Biology*, 2009.
- [7] L. E. Shelley, J. E. Scaife, M. Romanchikova, K. Harrison, J. R. Forman, A. M. Bates, D. J. Noble, R. Jena, M. A. Parker, M. P. Sutcliffe, S. J. Thomas, and N. G. Burnet, “Delivered dose can be a better predictor of rectal toxicity than planned dose in prostate radiotherapy,” *Radiotherapy and Oncology*, 2017.
- [8] F. Palorini, C. Cozzarini, S. Gianolini, A. Botti, V. Carillo, C. Iotti, T. Rancati, R. Valdagni, and C. Fiorino, “First application of a pixel-wise analysis on bladder dose-surface maps in prostate cancer radiotherapy,” *Radiotherapy and Oncology*, 2016.
- [9] M. Serban, A. A. de Leeuw, K. Tanderup, and I. M. Jürgenliemk-Schulz, “Vaginal dose-surface maps in cervical cancer brachytherapy: Methodology and preliminary results on correlation with morbidity,” *Brachytherapy*, 2021.
- [10] A. Witztum, B. George, S. Warren, M. Partridge, and M. A. Hawkins, “Unwrapping 3D complex hollow organs for spatial dose surface analysis,” *Medical Physics*, 2016.
- [11] A. McWilliam, C. Dootson, L. Graham, K. Banfill, A. Abravan, and M. van Herk, “Dose surface maps of the heart can identify regions associated with worse survival for lung cancer patients treated with radiotherapy,” *Physics and Imaging in Radiation Oncology*, 2020.
- [12] F. Buettner, S. L. Gulliford, S. Webb, and M. Partridge, “Modeling late rectal toxicities based on a parameterized representation of the 3D dose distribution,” *Physics in Medicine and Biology*, 2011.

- [13] E. Mylona, A. Cicchetti, T. Rancati, F. Palorini, C. Fiorino, S. Supiot, N. Magne, G. Crehange, R. Valdagni, O. Acosta, and R. de Crevoisier, “Local dose analysis to predict acute and late urinary toxicities after prostate cancer radiotherapy: Assessment of cohort and method effects,” *Radiotherapy and Oncology*, 2020.
- [14] C. R. Moulton, M. J. House, V. Lye, C. I. Tang, M. Krawiec, D. J. Joseph, J. W. Denham, and M. A. Ebert, “Spatial features of dose-surface maps from deformably-registered plans correlate with late gastrointestinal complications,” *Physics in Medicine and Biology*, 2017.
- [15] W. D. Heemsbergen, M. S. Hoogeman, G. A. Hart, J. V. Lebesque, and P. C. Koper, “Gastrointestinal toxicity and its relation to dose distributions in the anorectal region of prostate cancer patients treated with radiotherapy,” *International Journal of Radiation Oncology Biology Physics*, 2005.
- [16] R. C. Wortel, M. G. Witte, U. A. van der Heide, F. J. Pos, J. V. Lebesque, M. van Herk, L. Incrocci, and W. D. Heemsbergen, “Dose–surface maps identifying local dose–effects for acute gastrointestinal toxicity after radiotherapy for prostate cancer,” *Radiotherapy and Oncology*, 2015.
- [17] L. E. Shelley, M. P. Sutcliffe, S. J. Thomas, D. J. Noble, M. Romanchikova, K. Harrison, A. M. Bates, N. G. Burnet, and R. Jena, “Associations between voxel-level accumulated dose and rectal toxicity in prostate radiotherapy,” *Physics and Imaging in Radiation Oncology*, 2020.
- [18] W. D. Heemsbergen, L. Incrocci, F. J. Pos, B. J. Heijmen, and M. G. Witte, “Local Dose Effects for Late Gastrointestinal Toxicity After Hypofractionated and Conventionally Fractionated Modern Radiotherapy for Prostate Cancer in the HYPRO Trial,” *Frontiers in Oncology*, 2020.
- [19] N. Yahya, M. A. Ebert, M. J. House, A. Kennedy, J. Matthews, D. J. Joseph, and J. W. Denham, “Modeling Urinary Dysfunction After External Beam Radiation Therapy of the Prostate Using Bladder Dose-Surface Maps: Evidence of Spatially Variable Response

- of the Bladder Surface,” *International Journal of Radiation Oncology Biology Physics*, 2017.
- [20] H. A. Gay, H. J. Barthold, E. O’Meara, W. R. Bosch, I. El Naqa, R. Al-Lozi, S. A. Rosenthal, C. Lawton, W. R. Lee, H. Sandler, A. Zietman, R. Myerson, L. A. Dawson, C. Willett, L. A. Kachnic, A. Jhingran, L. Portelance, J. Ryu, W. Small, D. Gaffney, A. N. Viswanathan, and J. M. Michalski, “Pelvic normal tissue contouring guidelines for radiation therapy: A radiation therapy oncology group consensus panel atlas,” *International Journal of Radiation Oncology Biology Physics*, vol. 83, pp. 1–86, jul 2012.
- [21] O. Barbosa Neto, L. Souhami, and S. Faria, “Hypofractionated radiation therapy for prostate cancer: The McGill University Health Center experience,” *Cancer/Radiotherapie*, vol. 19, pp. 431–436, oct 2015.
- [22] H. M. Patrick and J. Kildea, “Technical note: rtdsm—an open-source software for radiotherapy dose-surface map generation and analysis,” *Medical Physics*, vol. 49, no. 11, pp. 7327–7335, 2022.
- [23] E. Onjukka, C. Fiorino, A. Cicchetti, F. Palorini, I. Improta, G. Gagliardi, C. Cozzarini, C. Degli Esposti, P. Gabriele, R. Valdagni, and T. Rancati, “Patterns in ano-rectal dose maps and the risk of late toxicity after prostate IMRT,” *Acta Oncologica*, 2019.
- [24] O. Casares-Magaz, L. P. Muren, V. Moiseenko, S. E. Petersen, N. J. Pettersson, M. Høyer, J. O. Deasy, and M. Thor, “Spatial rectal dose/volume metrics predict patient-reported gastro-intestinal symptoms after radiotherapy for prostate cancer,” *Acta Oncologica*, 2017.
- [25] I. Improta, F. Palorini, C. Cozzarini, T. Rancati, B. Avuzzi, P. Franco, C. Degli Esposti, E. Del Mastro, G. Girelli, C. Iotti, V. Vavassori, R. Valdagni, and C. Fiorino, “Bladder spatial-dose descriptors correlate with acute urinary toxicity after radiation therapy for prostate cancer,” *Physica Medica*, 2016.

- [26] F. Denis, P. Garaud, E. Bardet, M. Alfonsi, C. Sire, T. Germain, P. Bergerot, B. Rhein, J. Tortochaux, P. Oudinot, and G. Calais, “Late toxicity results of the GORTEC 94-01 randomized trial comparing radiotherapy with concomitant radiochemotherapy for advanced-stage oropharynx carcinoma: Comparison of LENT/SOMA, RTOG/EORTC, and NCI-CTC scoring systems,” *International Journal of Radiation Oncology Biology Physics*, 2003.
- [27] T. M. Atkinson, S. J. Ryan, A. V. Bennett, A. M. Stover, R. M. Saracino, L. J. Rogak, S. T. Jewell, K. Matsoukas, Y. Li, and E. Basch, “The association between clinician-based common terminology criteria for adverse events (CTCAE) and patient-reported outcomes (PRO): a systematic review,” 2016.

4.9 Supplemental Materials

4.9.1 Supplement A: Multiple Comparison Permutation Testing

The multiple comparisons permutation (MCP) test was introduced by Chen *et al* in 2013 as a means to compare images of dose distributions (like DSMs) between two groups of radiotherapy patients with strong control for type I (false positive) errors. It allows for the calculation of a single p-value to describe the discrepancy of the dose distributions between the two groups, as well the identification of subregions of the dose distributions that differ significantly between the groups.

Like any permutation test, the MCP test centers around computing the sampling distribution of a test statistic under the null hypothesis and determining where the test statistic of the observed sample sits in that distribution. In order to estimate this sampling distribution, many samples of the test statistic under the null hypothesis need to be calculated. This can be achieved by permuting (i.e., rearranging) the data of the observed sample, calculating the test statistic for that permutation, and repeating the process until a sampling distribution is achieved (Fig 4.10). This is possible because under the null hypothesis all possible permutations are equally likely, meaning we can relate our permutations to the sampling distribution.

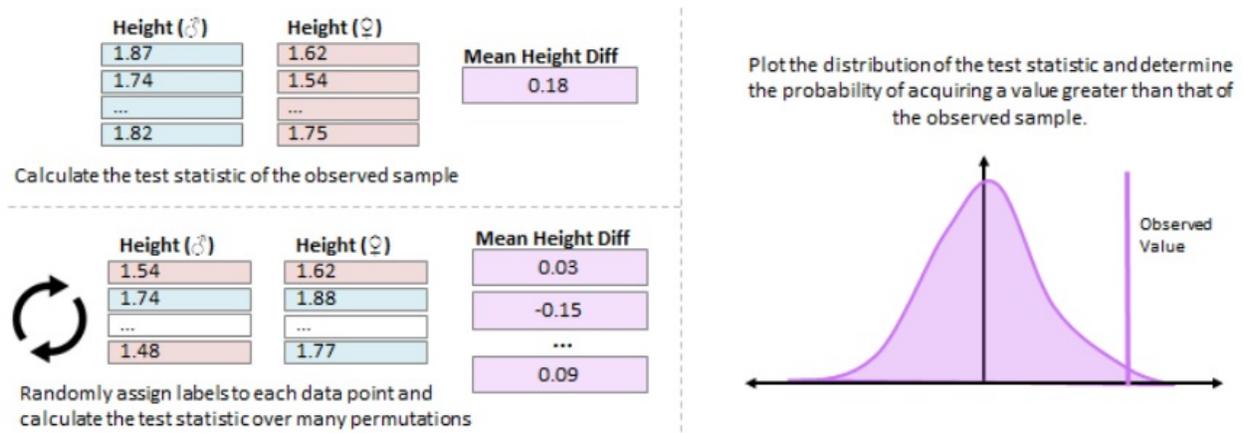


Figure 4.10: An example of a basic permutation test. We hypothesize the average male is taller than the average female, therefore under the null hypothesis we would expect the mean height difference (our test statistic) to be zero. To test this, we calculate the mean height difference many times, randomly grouping our data into "male" and "female" categories each time. Once we have a distribution we can determine where the test statistic of our observed sample lies within it and determine if the null hypothesis can be rejected.

The main modification introduced in Chen's implementation of permutation testing is the way in which they address the multiple comparisons problem. Consider the typical α value of 0.05 used by most scientists when conducting statistical tests: we deem $\alpha \leq 5\%$ chance that the observed data could have occurred under the null hypothesis to be small enough that we can reject it. When conducting a single test, there is a 1/20 chance that this occurs, but when conducting many tests in tandem (such as when comparing doses in many voxels) it becomes much more likely that we falsely reject the null hypothesis for some of the tests. This is the multiple comparisons problem, which we can account for by performing a single test for all voxels. An easy way to do this for permutation testing is to create a single test statistic distribution using the maximum test statistic value of all voxels from each permutation. However, Chen recognized that the typical test statistic, the mean or median difference in dose value of a given permutation, would have different variance in different voxels. If unaccounted for, it would be more likely that the maximum test statistic of a given permutation would be selected from a voxel with a high level of variance that occurred by chance, rather than a lower statistic from a voxel with smaller variance that is much less

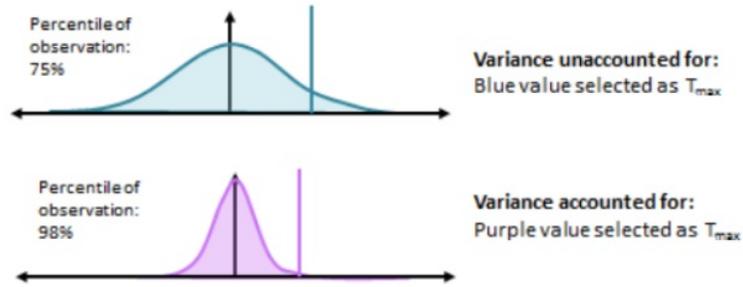


Figure 4.11: Visual explanation of how accounting for variance in test statistic value impacts the choice of T_{max} value. If variance is not taken into consideration, a value from the 75th percentile of one voxel’s distribution of values can be selected over a value in another voxel’s 98th percentile, which is arguably a greater outlier.

likely (Fig 4.11). For this reason, Chen’s MCP test statistic for a voxel is normalized using an estimate of its standard deviation derived from many permutations.

Chen’s Original Implementation: the 2-sample unpaired test

The MCP test, as defined by Chen *et al.*, relies on the following conditions:

1. All dose images have the same size and resolution and represent the same anatomy. For DSMs, this means all maps must be arrays of size $M \times N$ and be calculated using the same contour definition of the object of interest (e.g. Rectum superior and inferior limits are consistent between patients).
2. All images are labeled as belonging to one of two groups. Chen defines their labels as “E” for event and “N” for non-event.
3. The labels are arbitrary under the null hypothesis. In other words, if there is no difference in the dose received by patients with and without adverse events, each patient is equally likely to fall into the “E” and “N” groups.

The test statistic is defined as the mean dose difference between the two groups at a given voxel, k , normalized by that voxel’s standard deviation:

$$T_{i,k} = \frac{\mu_{E,k} - \mu_{N,k}}{\sigma_k} \quad (4.1)$$

where i is a given iteration. The maximum value of the test statistic from each iteration, $T_{max,i}$, is used to create the sampling distribution required to determine likelihood of the observed sample occurring under the null hypothesis.

Overview of Test Procedure

A visual summary of the procedure is also shown in Figure 4.12a.

1. Begin by calculating the average dose difference between all voxels of the two groups of the observed sample. We will call this $d_{obs,k}$.
2. Randomly shuffle the labels assigned to each patient to produce a new permutation of the data. As the total number of possible permutations is $\frac{n_N!n_E!}{(n_N+n_E)!}$, it is acceptable to use a smaller number of permutations $N_p > 1000$ if the number of patients is sufficiently large.
3. Calculate the average dose difference between all voxels for the permutation ($d_{i,k}$), then repeat the permutation process N_p times.
4. Once you have calculated $d_{i,k}$ for all permutations, calculate the standard deviation of d_k of each voxel across all permutations, σ_k . Then calculate the test statistic $T_{i,k}$ of each permutation by dividing $d_{i,k}$ by σ_k .
5. Calculate the maximum value of $T_{i,k}$ for each permutation to get a distribution of $T_{max,i}$ values. Additionally, calculate $T_{max,obs}$ of the observed sample.
6. Finally, calculate the proportion of $T_{max,i}$ values greater than $T_{max,obs}$ to acquire the adjusted p-value of the sample. If this p-value is smaller than the significance level α the null hypothesis can be rejected.

In addition to calculating the adjusted p-value, the $T_{max,i}$ distribution can be used to calculate the threshold value of the $(1-\alpha)$ percentile, T^* . Any voxels in the observed normalized dose difference map, $T_{obs,k}$ greater than this value indicate a significant dose difference between groups E and N.

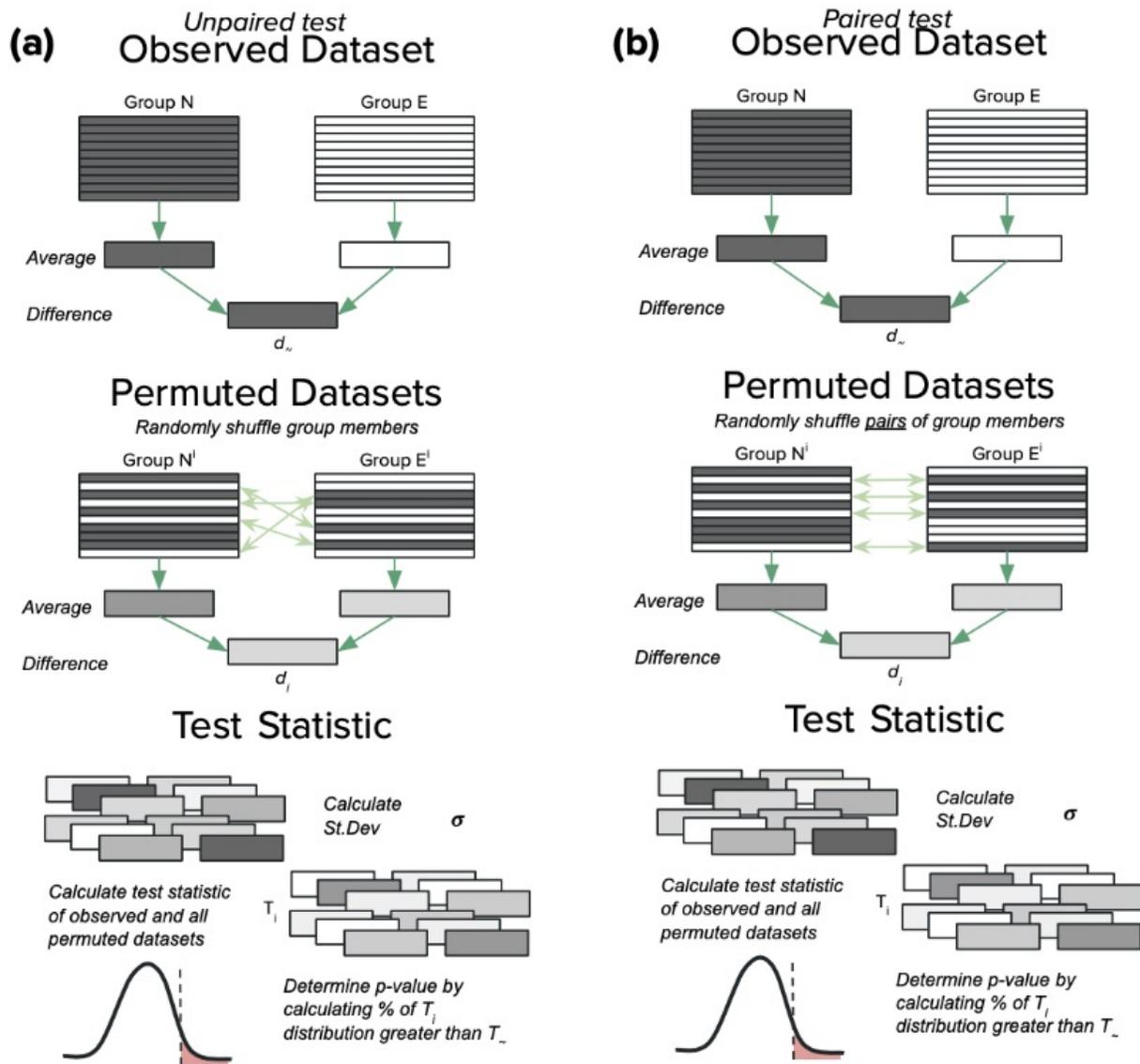


Figure 4.12: Visual depictions of the unpaired (a) and paired (b) MCP testing process.

Modifications: One sample and paired-sample tests

Chen's original implementation of MCP testing was designed to be primarily used to compare dose distributions of patients who did and did not experience adverse events in response to radiotherapy. While this approach can also be extended to other scenarios that have two sets of unpaired data, modifications are required if one wants to compare paired data (e.g. planned and delivered dose distributions or DSMs computed in different ways).

While the multiple comparisons problem remains and can be addressed as in the original

test with the normalized test statistic, we need to adjust the permutation process to ensure the data remains paired. This can be achieved by randomly determining sets of data pairs to swap labels with one another (Fig 4.12b). The process can be simplified even further if we calculate the dose differences between each pair up front, then randomly permute the signs of these differences to calculate d_k for each permutation (of 2^n possible). The value of d_k will be identical in both instances, but the second approach will be faster to implement. This implementation is also identical to that of a one-sample permutation test, meaning comparisons of a single cohort to a known mean (such as comparing DSMs from daily fractions to a planned DSM) are also possible.

As with the original test, conditions 1-3 must hold, with the only modification being the paired nature of the labels.

Overview of Paired Testing Procedure

A visual summary of the procedure is also shown in Figure 4.12b.

1. Begin by calculating the dose difference maps between each pair of data. Calculate the average of these maps to acquire $d_{obs,k}$.
2. Randomly multiply each dose difference map by +1 or -1, then calculate the average, $d_{i,k}$. Repeat this step 2^n or ≥ 1000 times, depending on the total number of pairs in the sample. If the sample size is sufficiently small, we recommend explicitly calculating the test statistic of all possible permutations.
3. Continue performing the test as normal, starting from step 4 of the original Chen implementation.

Overview of the Testing Procedure for a Single Sample

1. If you have an expected dose distribution (such as a planned DSM), subtract it from all dose distributions in your sample. If you do not, you can still proceed with the test, understanding that the null hypothesis you will be testing is that the mean dose is 0.
2. Randomly multiply each dose distribution by +1 or -1, then calculate the average, $d_{i,k}$. Repeat this step 2^n or > 1000 times, depending on the total number of pairs in the sample.

3. Continue performing the test as normal, starting from step 4 of the original Chen implementation.

Similar to the original implementation, the threshold value T^* can be calculated and used to identify pixels in the observed sample that differ significantly from their expected values under the null hypothesis.

4.9.2 Supplement B: Additional Figures

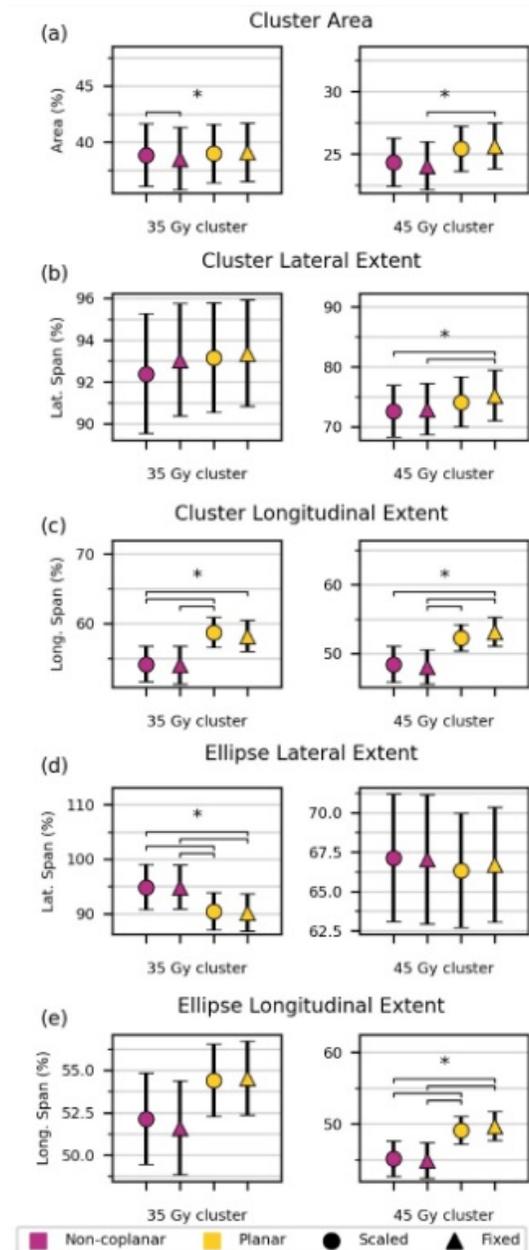


Figure 4.13: Mean DSM feature values (and standard uncertainty of the mean) from the four rectum DSM calculation approaches of the 35 and 45 Gy clusters. Significantly different pairs of features are indicated with a bracket and an asterisk.

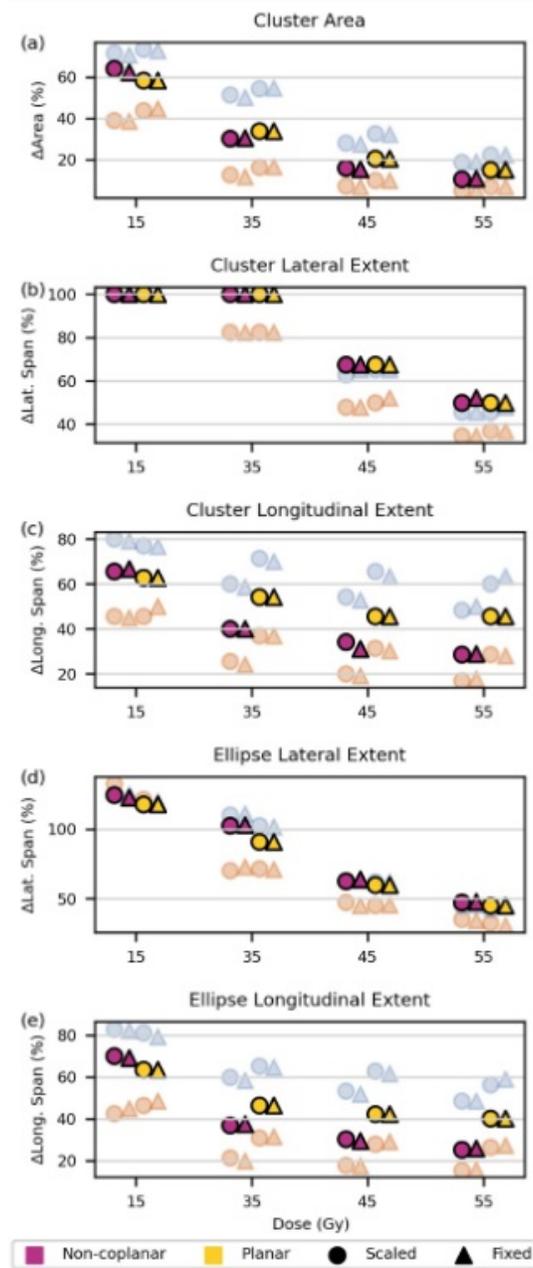


Figure 4.14: DSM feature values from the four rectum DSM calculation approaches for the representative (median) patient. Features from the patients with the shortest (blue) and tallest (red) rectums are included as well.

Chapter 5

Experimental validation of dose accumulation for the rectum using dose-surface maps

Haley M Patrick, Emily Poon, and John Kildea

Manuscript accepted to *Acta Oncologica* May 2023 and in revisions.

5.1 Preface

In this chapter, we assess whether the addition of DSMs of daily delivered dose can be used to perform dose accumulation for rectum structures. While dose accumulation is conventionally performed using deformable image registration software, results for rectum structures have historically been of poor quality. Some previous studies have therefore used daily planar DSMs to calculate accumulated rectum dose. However, no author has presented empirical evidence to indicate that DSM accumulation can yield an accurate representation of the true total delivered dose. Characterizing the accuracy of DSM accumulation is key to ensure confidence not only in the approach itself, but also in conclusions from studies employing it. Based on this need, we investigated the accuracy of DSM-based dose accumulation by comparing accumulated DSMs to physical measurements.

Following an established framework for clinical verification of treatment plan delivery, we used radiochromic films to quantify delivered dose over a multi-fraction treatment. First, we constructed a rectum phantom capable of replicating different forms of inter-fraction motion. Next, we irradiated the phantom over several different treatment scenarios, using radiochromic film to measure dose to the rectum's surface. We then calculated the accumulated rectum dose based on CBCT images of the fractions using planar and non-coplanar DSMs, as well as DIR software. Finally, we used gamma analysis to compare our calculated accumulated doses to measurements. We found excellent agreement between accumulated non-coplanar DSMs and film measurements, which outperform accumulations done using DIR or planar DSMs. These findings indicate that DSM accumulation can provide an accurate representation of rectum delivered dose, conditional on the usage of non-coplanar DSMs.

5.2 Abstract

Background: Dose-surface maps (DSMs) are an increasingly popular tool to evaluate spatial dose-outcome relationships for the rectum. Recently, DSM addition has been proposed as an alternative method of dose accumulation from deformable registration-based techniques. In this study, we experimentally investigated the accuracy at which DSM accumulation can capture the total dose delivered to a rectum's surface in the presence of interfraction motion.

Material and Methods: A custom PVC rectum phantom capable of representing typical rectum interfraction motion and filling variations was constructed for this project. The phantom allowed for the placement of EBT3 film sheets on the representative rectum surface to measure rectum surface dose. A multi-fraction prostate VMAT treatment was designed and delivered to the phantom in a water tank for a variety of interfraction motion scenarios. DSMs for each fraction were calculated in two ways using CBCT images acquired during delivery and summed to produce accumulated DSMs. Accumulated DSMs were then compared to film measurements using gamma analysis (3%/2mm criteria). Film measurements were also compared to DSMs produced by deformable image registration to

quantify accuracy of a more conventional dose accumulation method.

Results: Baseline agreement between film measurements and accumulated DSMs for a stationary rectum was 95.6%. Agreement between film and accumulated DSMs in the presence of different types of interfraction motion was $\geq 92\%$, whereas agreement with deformable registration-based DSMs was much lower (38.9-90.7%). Overall, DSM accumulation performed the best when using DSMs that accounted for changes in rectum path orientation.

Conclusion: Dose accumulation performed with DSMs was found to accurately replicate total delivered dose to a rectum phantom in the presence of interfraction motion. These results indicate that DSM accumulation can be a viable alternative to deformable registration-based dose accumulation.

5.3 Background

Dose-surface maps (DSMs) are sources of spatial dose distribution information in radiation oncology that are becoming increasingly popular for dose-outcome studies of hollow organs like the rectum [1, 2], with growing interest in using them to evaluate delivered dose to the rectum [3, 4]. However, accurate quantification of delivered dose to the rectum can be challenging. While Murray *et al* were the first to use DSMs to perform dose accumulation for the rectum several years ago [5], no attempt to experimentally test the validity of this approach has been made. Considering that interest in using DSMs for dose accumulation appears to be growing [3, 6, 7], an evaluation of its validity is warranted.

During a course of prostate radiotherapy the shapes and positions of organs within the irradiated volume can change substantially due to inter-fraction motion, with reported prostate shifts on the order of 2.7-15 mm [8–10]. Failure to account for this motion can introduce discrepancies between planned and delivered dose to the target and organs at risk (OARs) and has been shown to contribute to increased risk of relapse and radiation toxicities in conformal treatments [11–14]. While daily image-guided radiotherapy (IGRT) is able to restrict dose deviations to the prostate target to within 4%, delivered dose to the rectum and bladder can still differ substantially from what was planned [15], necessitating

a means to determine the dose that was delivered.

Dose accumulation is a process that uses a patient’s treatment plan and daily IGRT images to calculate and sum their daily delivered doses into a total, accumulated dose. As anatomical variations can exist between treatment fractions, a key element of dose accumulation is a means to combine daily doses in the same reference anatomy. Traditionally, this is achieved using deformable image registration (DIR) algorithms to deform daily images to the planning anatomy along with their daily dose distributions. However, many DIR algorithms are unable to properly handle the complex deformations and filling changes of the rectum during prostate radiotherapy [16–18]. Although some performance improvements have been reported for newer hybrid or biomechanical deformation algorithms [19–21], DIR’s historically weak performance for the rectum has led to the investigation of alternative dose accumulation strategies for this organ.

As mentioned earlier, a growing number of groups are reporting on rectum dose accumulation for dose-outcome studies using DSMs [4, 7, 22]. A DSM is a 2D representation of the dose to the surface of an organ. DSMs are typically calculated in a way that maps the dose to an organ’s surface to a standard grid, making dose accumulation as hypothetically simple as superimposing and adding daily DSMs together. However, this daily DSM addition process relies on the untested assumption that it can accurately account for inter-fraction translations and volume variations to produce accurate accumulated doses. Considering that empirical studies of DIR performance for dose deformation have revealed important real-world inconsistencies that were not observable in prior *in silico* studies [23], similar real-world validation of DSM-based dose accumulation is essential if it is to become a viable dose accumulation technique.

Motivated by this need, in this study, we attempted to validate DSM-based dose accumulation for the rectum. Inter-fraction rectal motion and distention over the course of a multi-fraction treatment were simulated using a rudimentary pelvic phantom and the accuracy of accumulated DSM dose accumulation was assessed using EBT3 film dosimetry.

5.4 Material and Methods

5.4.1 Phantom Design

A $28 \times 28 \times 29$ cm acrylic water tank was used as a simple pelvic phantom, with interchangeable 15 cm lengths of Xirtec® PVC pipe of various diameters used to represent the rectum within it (Fig 5.1a). Radiopaque BBs and etched crosshairs on the tank were used to facilitate reproducible setup of the tank’s center with in-room patient-positioning lasers. Pipe diameters of $\frac{3}{4}$ ”, 1”, and $1\frac{1}{4}$ ” were used to represent the range of rectum cross sectional areas reported in the literature [24]. In order to reproduce the rectum’s anatomical position in a supine patient, a custom stand insert was constructed to interchangeably hold each rectum pipe 9 cm above the tank’s base. This positioned the rectum pipe’s anterior wall 2.5 cm posterior to the positioning of the isocenter within the tank, replicating the average distance between the rectum and the isocenter location in prostate cancer treatment plans. Both the rectum pipes and the stand were marked with guide lines to allow for easy alignment with the lasers. Pipe positionings could also be varied by using different height stands or shimming the stand’s base plate.

5.4.2 Treatment Planning

A CT simulation image of the phantom containing the smallest diameter (reference) pipe was taken using a Big Bore CT scanner (Philips, Amsterdam, The Netherlands) with a 3.0 mm slice thickness and imported into Eclipse (v.15, Varian Medical Systems, Palo Alto, CA) for treatment planning. The tank “body” and rectum pipe were contoured and a 5.5 cm diameter spherical PTV defined at the isocenter location. A 6 MV two-arc VMAT plan was designed based on our institutional 60 Gy in 20 fractions treatment for moderate-risk prostate cancer [25]. In brief, the prescription was reduced to 3 Gy in 3 fractions and the planning constraints for the rectum and PTV were scaled down accordingly. The 3 Gy prescription dose was chosen as it provided the least noisy dose distributions in our EBT3 films of the prescriptions tested.

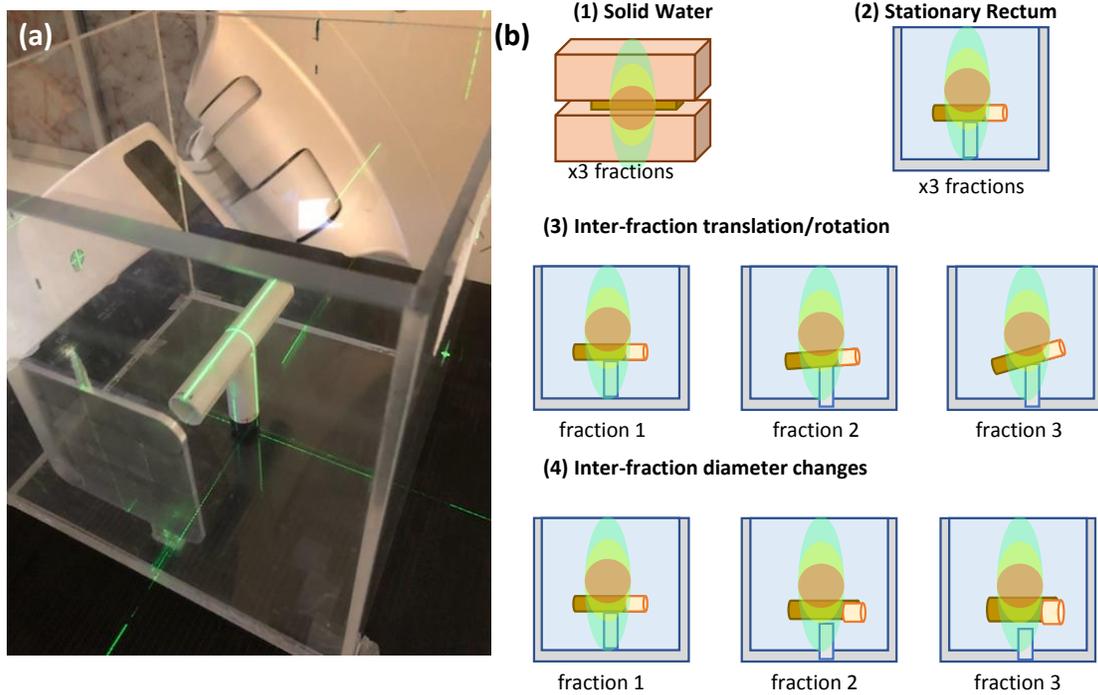


Figure 5.1: (a) Set-up of the water tank with the smallest diameter rectum pipe in place. (b) Visualizations of the four measurement set-ups tested in this study. A three-dimensional visualization is provided for the solid water set-up and a sagittal view visualization is provided for the water tank experiments.

5.4.3 Experimental Setups

All measurements were performed on a Varian TrueBeam linac. In total, four different measurements were performed to evaluate the agreement between film and calculated DSMs. Each represented a different scenario of interest as follows (Fig 5.1b)

1. **Dose plane in solid water:** used to evaluate the baseline agreement between the treatment plan and its delivery using a standard film IMRT QA protocol [26]. Two $30 \times 30 \times 6$ cm slabs of solid water were positioned on the treatment couch with a 127×203 mm sheet of film sandwiched between them, centered at the isocenter, and irradiated with the three-fraction treatment plan.
2. **Dose to a stationary rectum:** used to evaluate the baseline agreement between a DSM and film measurement. The center of the water-filled water tank was positioned at

the isocenter, a 127×85 mm sheet of film was secured tightly around the circumference of the smallest diameter rectum pipe and centered on the stand using the room lasers. The plan was delivered in its entirety to the phantom, after which the film was removed, gently towed dry, and stored to develop.

3. **Dose to a rectum with inter-fraction motion:** used to evaluate how well accumulated DSMs represent delivered dose in scenarios involving inter-fraction translation/rotation. The set-up was identical to measurement 2, except that the pipe was further secured to the stand with tape to prevent it from moving between fractions. Between the delivery of each fraction the pipe stand was shimmed, tilting the pipe 3° , then 7° from horizontal.
4. **Dose to a rectum with inter-fraction volume changes:** used to evaluate how well accumulated DSMs calculate delivered dose for scenarios that involve rectum diameter changes. As the different diameters made it impossible for a single film sheet to measure the accumulated surface dose across all three fractions, we opted to measure accumulated dose to four inferior-superior lines located at the cardinal angles along the rectum's surface. Four 127×20 mm strips of film were attached to the posterior, anterior, left, and right sides of each rectum pipe and centered in the tank using the room lasers. After each fraction the films were removed and transferred to a different diameter pipe for delivery of the next fraction.

Before each film measurement, a CBCT image of the experimental set-up was acquired with a dummy film in the place of the measurement film to facilitate the delivered dose and DSM calculation processes. In order to assess and account for any film darkening due to water, each water tank measurement had a corresponding non-irradiated reference film submerged for the same duration as the measurement film. This reference film was scanned alongside the measurement films and a piece of non-submerged film, and compared to the non-submerged film. Significant water-induced darkening could then be accounted for by subtracting the difference in optical density from measurement films [27].

5.4.4 Film Dosimetry

All measurements were carried out using Gafchromic EBT3 film (Ashland Global, Lot 03082202). Film calibration was performed by exposing film strips positioned at 6 cm depth in solid water using a 6 MV 10×10 cm, 100 cm SSD field. Six calibration doses covering a range from 0 to 500 cGy were acquired and cross-calibrated with measurements performed with an Exradin® A19 ion chamber. All films were left to develop for 42 hours to allow for full evaporation of water from submerged films before scanning in 48-bit RGB format with an Epson 11000XL scanner and glass compression plate at 0.35 mm resolution. Images were converted to dose distributions using FilmQA Pro (v.5, Ashland Global) [28]. Measurement films were scanned alongside three reference films: the aforementioned submerged film (0 cGy) and two unsubmerged films irradiated to 0 and 400 cGy in the same solid water setup as the calibration films. The two 0 cGy films were used to account for water-induced darkening, while the two non-submerged films were used to perform linear dose rescaling to correct for interscan variability.

5.4.5 Dose-Surface Map Calculation and Accumulation

For each measurement scenario, delivered dose was calculated by first registering the treatment plan to the acquired CBCT and then dose re-calculation was performed in Eclipse. For the solid water measurement, the dose plane at the position of the film was exported for comparison. For each pipe measurement, the rectum pipe was contoured on the CBCT image. RT-Structure and RT-Dose DICOM files of each setup scenario were exported to enable the DSM calculations.

DSMs were calculated using the Python package *rtdsm* [29], previously developed and published under an open-source license by our group. *rtdsm* allows for two DSM calculation styles: the “planar” approach, which samples dose using parallel axial slices, and the “non-coplanar” approach, which uses slices angled orthogonally to the rectum’s central axis path. We opted to use the non-coplanar approach as it is more appropriate for rectum structures, though we did also investigate the planar approach for comparison purposes. To begin the DSM calculation process in *rtdsm*, 91 slices were defined every 1.5 mm along the length

of each rectum pipe contour, orthogonal to the pipe's central axis, with the central slice located at the halfway point. Next, equiangular sampling points were defined every 6.3° (for constant diameter scenarios) or 6° (for the changing diameter scenario, in order to explicitly include sampling points at the four cardinal angles) around each slice's circumference and the point dose sampled from the 3D dose matrix. Finally, the tubular dose distributions were cut open posteriorly and flattened into 2D dose arrays with a point resolution on the reference (smallest) pipe of $1.61 \times 1.5 \text{ mm}$, and $1.52 \times 1.5 \text{ mm}$ for the diameter changing scenario. Accumulated doses were calculated by superimposing and adding up all single fraction DSMs for a given measurement scenario and saved as DICOM dose planes for easy import into FilmQA Pro.

In addition to the DSMs defined above, for comparison purposes, planar DSMs were also used to produce an accumulated DSM for the inter-fraction motion scenario. In this case, axial slices were simply defined every 1.5 mm along the superior-inferior axis of the CBCT image. This accumulated DSM allowed us to evaluate the accuracy of DSM accumulation using planar DSMs compared to non-coplanar ones.

In order to compare the performance of DSM-based dose accumulation to the conventional DIR-based approach, accumulated doses were also calculated using MIM Maestro (MIM Software, Cleveland, OH). Images, contours, and dose distributions of each setup scenario were imported to the software and deformably registered to the planning CT using MIM's Same Subject algorithm using a smoothing factor of 0.5 to calculate accumulated doses. Finally, DSMs of each scenario were calculated from the DIR-accumulated dose distributions to allow for comparison with films.

5.4.6 Analysis

Calculated accumulated DSMs were imported into FilmQA PRO and aligned with the film measurements using the application's alignment optimization tool. Regions of comparison excluded the outer 0.5 cm of the films where water is known to irreversibly impact optical density [27]. Gamma analysis [30] was performed on the red channel with a 3% dose difference (global normalization) and 2 mm distance-to-agreement using a dose threshold of 10% in accordance with TG-218 recommendations [31]. Performance was compared to

Table 5.1: Gamma pass rates for each combination of experimental set-up and test criterion.

Experimental Setup	2%/2mm	3%/2mm	3%/3mm
(1) Dose plane in solid water	98.2%	98.6%	99.2%
DSM-based Dose Accumulation			
(2) No inter-fraction changes	90.8%	95.6%	98.1%
(3) Inter-fraction motion	87.9%	94.9%	98.8%
(3) Inter-fraction motion - planar DSM	82.7%	92.0%	97.5%
(4) Inter-fraction volume changes – Anterior Profile	86.5%	98.1%	98.1%
(4) Inter-fraction volume changes – Posterior Profile	98.1%	100%	100%
(4) Inter-fraction volume changes – Left Profile	98.1%	100%	100%
(4) Inter-fraction volume changes – Right Profile	98.0%	100%	100%
DIR-based Dose Accumulation			
(3) Inter-fraction motion	34.8%	38.9%	43.4%
(4) Inter-fraction volume changes – Anterior Profile	65.3%	75.5%	85.7%
(4) Inter-fraction volume changes – Posterior Profile	87.0%	90.7%	90.7%
(4) Inter-fraction volume changes – Left Profile	69.8%	77.4%	88.7%
(4) Inter-fraction volume changes – Right Profile	38.0%	50.1%	82.0%

TG-218’s universal tolerance and action limits of $\geq 95\%$ and $\geq 90\%$, respectively. Additional comparisons were also performed with 2%/2 mm and 3%/3 mm criteria to evaluate performance under stricter and more lenient conditions.

5.5 Results

Results for each experimental setup are described below, graphically presented in figures 5.2 and 5.3, and tabulated in Table 5.1.

5.5.1 (1) Dose plane in solid water

Measured and calculated delivered dose distributions at the central plane of the PTV are shown in Figure 5.2a along with the corresponding gamma index map. Although gamma pass rates were within action limits (98.6%, Table 5.1), the measured and calculated delivered dose planes disagreed in a few areas. Hotspots exceeding 400 cGy existed within the PTV region of the film measurements that were not present in the calculated dose plane and were the main cause of gamma index failures.

5.5.2 (2) Dose to a stationary rectum

Measured and DSM-calculated doses to the stationary rectum had a gamma pass rate of 95.6%, within TG-218 tolerance limits (Table 5.1), and with good positional agreement of all isodose lines below the prescription dose (Fig 5.2b). Like the previous comparison, the main source of disagreement was due to the film measuring higher dose hotspots within the PTV compared to the calculated DSM.

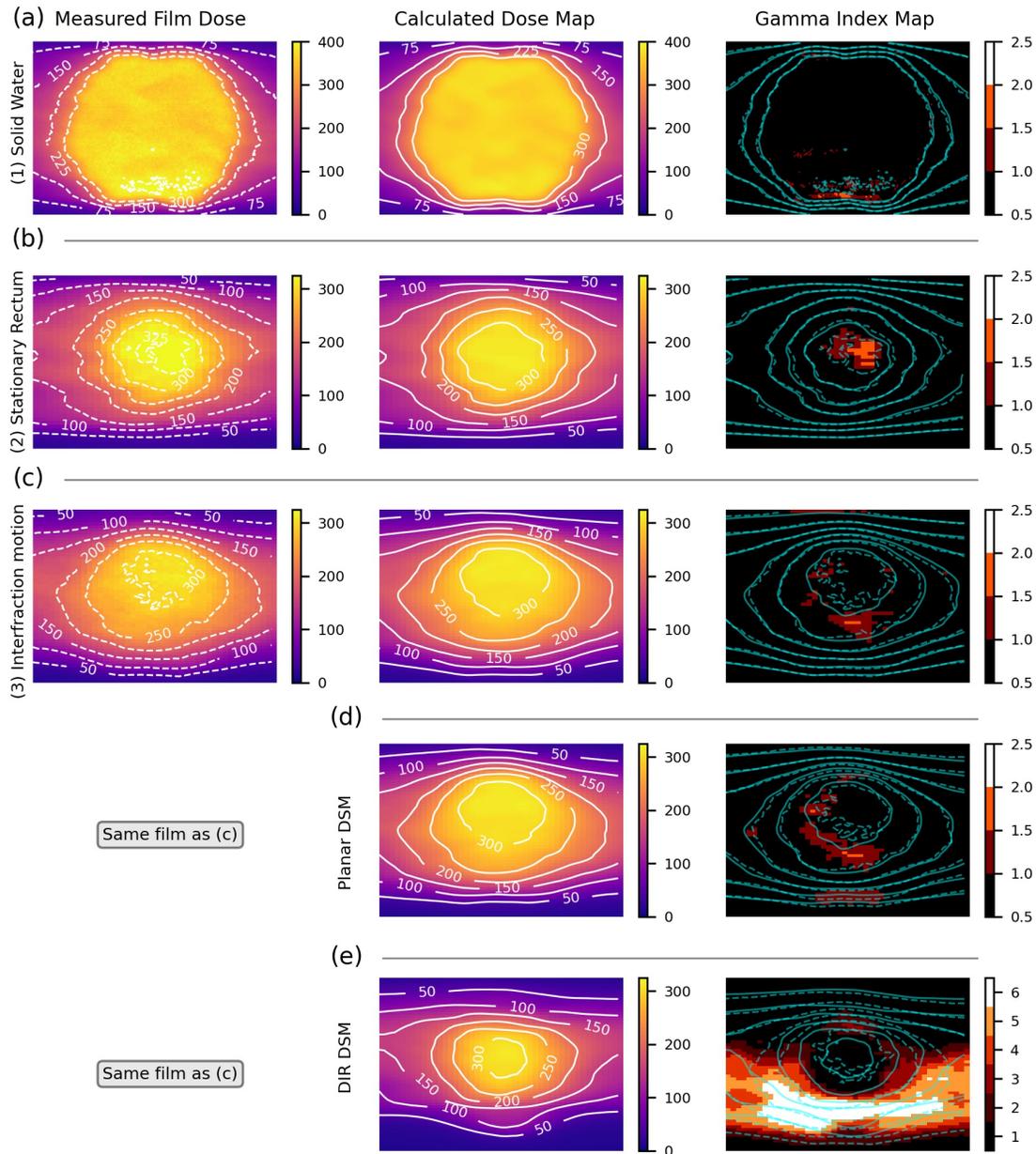


Figure 5.2: Total accumulated dose distributions for each measurement scenario, as measured by film and calculated DSMs, along with their gamma index maps for the 3%/2 mm criteria. Isodose lines are included to facilitate visual comparison of the film and DSM dose distributions.

5.5.3 (3) Dose to a rectum with inter-fraction motion

Gamma analysis of the inter-fraction motion measurements had slightly lower agreement between the film and DSM dose maps (gamma pass rate: 94.9%, Table 5.1), but still comfortably within TG-218 action limits. Film measurements reported lower doses than the accumulated DSM in this scenario, with the measured 300 cGy region being smaller and more contained than the calculation (Fig 5.2c). This influenced gamma indices in such a way that the main locations of failure occurred between the 250 and 300 cGy isodose regions, rather than within hotspots as was the case for the previous two measurements.

Film-DSM agreement decreased when using the planar style of DSM calculation (92.0% pass rate). Increased patches of disagreement were still present within the > 250 cGy area as for the non-coplanar style, but also in the anterior low-dose gradient region of the inferior rectum (Fig 5.2d).

5.5.4 (4) Dose to a rectum with inter-fraction volume changes

Dose profiles as measured with the film and extracted from the DSMs are presented in Figure 5.3 along with profiles of their gamma indices. Overall agreement was very strong for all four profiles ($\geq 98\%$, Table 5.1), with the posterior, left, and right profiles closely matching even with stricter passing criteria. Similar to the other measurements, gamma criteria failures were primarily located in the anterior rectum, where the film measured higher doses. For example, gamma failures occurred from 80-90 mm along the anterior profile for the 2%/2 mm test.

5.5.5 Accumulated dose as calculated by DIR

Dose distribution agreement was lower in both cases compared to the DSM dose accumulation (Figs 5.2d, 5.3). For scenario 3 (inter-fraction motion), the DIR-calculated dose distribution's topography differed significantly from the film. While the 300 cGy isodose region was of a similar size and position to the film's, the 100 and 200 cGy regions were much smaller and tightened in around the hotspot, especially for the superior half of the rectum. Consequently, gamma index values in this region far exceeded TG-218 action limits, leading to a pass rate

of 38.9%. Gamma pass rates were similarly reduced for scenario 4 (inter-fraction volume changes, Fig 5.3). Measurement-calculation concordance was strongest for the posterior wall profile (90.7%) and weakest for the right wall profile (50.1%) with the left and anterior wall profiles falling in the middle (77.4% and 75.5%, respectively). In the case of the right and left walls, the low pass rates were due to the DIR accumulation underestimating the dose relative to the film measurement, whereas for the anterior wall the cause was the DIR accumulation compacting the high dose region relative to the film.

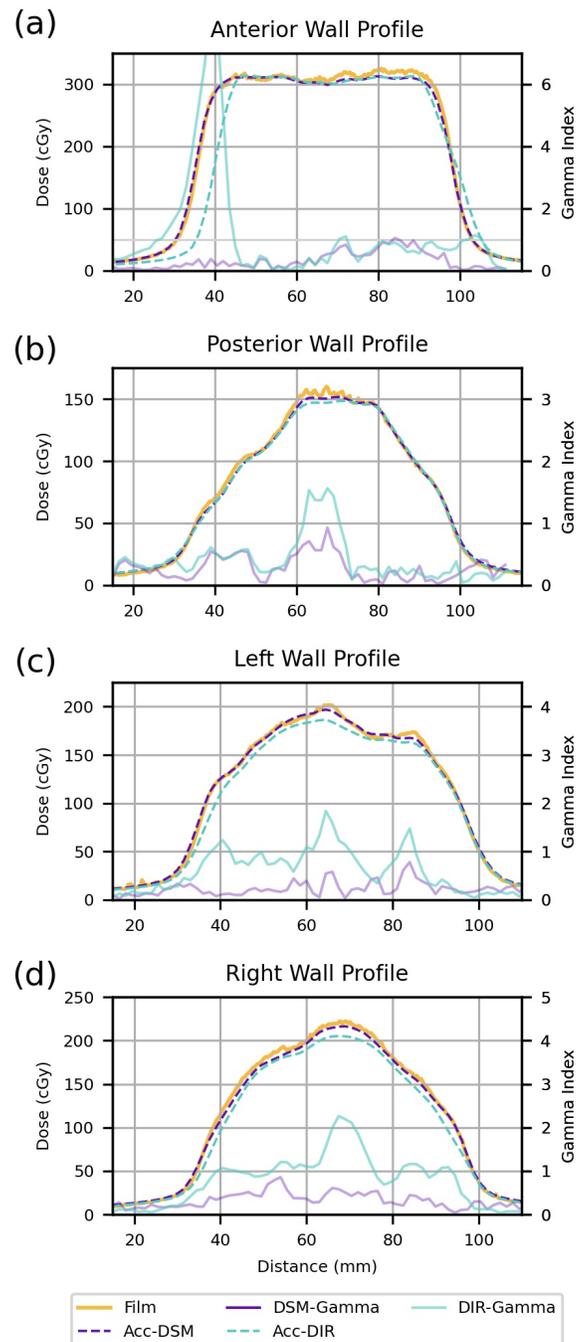


Figure 5.3: Dose profiles for the (a) anterior, (b) posterior, (c) left, and (d) right walls of the rectum phantom in the volume changing scenario. Gamma index values for the 3%/2 mm criteria are included. Both DSM- and DIR-based results are shown.

5.6 Discussion

The need for accurate dose accumulation methods has been recognized as a critical factor to improve understanding of normal tissue response to radiation [32]. In parallel, the value of including spatial-dose information in dose-outcome research has been demonstrated through the discovery of radiosensitive subregions that are otherwise masked by DVH metrics [7, 33, 34]. While a handful of studies have attempted to combine dose accumulation and spatial-dose visualization through DSM accumulation [3–7], the real-world validity of this approach has not been evaluated until now.

In this work we tested the accuracy of DSM-based dose accumulation for rectum structures against film measurements and found good agreement between measured and calculated surface doses. Gamma pass rates using the 3%/2 mm criteria ranged from 92.0-95.6% for full surfaces and exceeded 98% for 1D profiles, falling comfortably within TG-218 recommended action limits and ranges reported by recent end-to-end performance studies of MRI guided adaptive radiotherapy (MRgART) using DIR. For example, Hoffmans and Bohoudi used film to compare measured and MRgART calculated dose to the rectum’s surface for multi-fraction treatments to deformable pelvis phantoms and reported pass rates between 87.9-100% for 3%/2 mm criteria [35, 36]. A similar study by Elter *et al* employing gel dosimetry reported lower pass rates of 93.7% (3%/3 mm) for rectum dose, likely due to a full 3D volumetric comparison [37]. We found that agreement between film and accumulated DSMs was weakest for hotspots within the PTV (scenarios 2 and 4) and for locations in the high-dose gradient region with large inter-fraction motion (scenario 3). These may have been a consequence of positional uncertainties in regions with steeper dose gradients, similar to results observed by Marot *et al* using an anthropomorphic phantom [38].

We found that DSM-based dose accumulation outperformed DIR-based accumulation in both the presence of inter-fraction motion and volume changes (Table 5.1). From visual inspections we determined the cause to be the DIR algorithm’s inability to accurately handle the larger deformations of each test scenario. The DIR algorithm struggled to handle the larger positional displacements of the inferior rectum when it was tilted by 7°,

creating non-realistic deformations and causing the poor dose accumulation accuracy seen in this region (Fig 5.2e). Similarly nonsensical deformations were produced by the DIR algorithm when shrinking the largest diameter rectum to match the smallest one, pinching the anterior wall along the inferior-superior axis and impacting the accumulated dose profile. Several previous publications have observed similar nonphysical deformations when using MIM's DIR algorithm, with accuracy generally decreasing for larger deformations [21, 39–41]. This performance has been attributed to MIM using an intensity-based, restricted free-form demons' algorithm, causing it to prioritize image similarity and thereby allowing large regional deformation errors to occur. While further experimentation is required to benchmark DSM-accumulation against other algorithms, our results suggest that DSM-accumulation can outperform certain free-form intensity-based DIR algorithms.

Our results show that agreement between film and DSM reduces when using the simplified “planar” style of DSM calculation for the inter-fraction motion scenario. Due to the simplified axial orientation of the DSM slices, doses to the anterior and posterior rectal walls are sampled at different locations, producing a DSM grid that displaces measured point doses relative to their film counterparts (Fig 5.4). While this displacement was kept within 1.6 mm for the scenarios of this study, displacements exceeding 3 mm are possible for angular offsets as small as 13° . This poses a major potential issue for planar-style DSM accumulation accuracy for rectums with significant trajectory in the anterior-posterior or left-right directions and may have consequences for existing DSM accumulation studies. To date, all but one [6] of the DSM accumulation studies reported in the literature has used the planar calculation approach [3–5, 7, 22], meaning all are likely to include this effect. While we found one dissertation [22] that did provide anecdotal accounts of good correspondence between DSMs from DIR and DSM accumulation, full details of this investigation were not provided.

Although our study provides evidence for the validity of DSM-based dose accumulation for simple inter-fraction motion and deformations, further studies are warranted to investigate more advanced scenarios. The simplicity of our rudimentary pelvic phantom prevented investigations of more complex inter-fraction changes like localized motion and

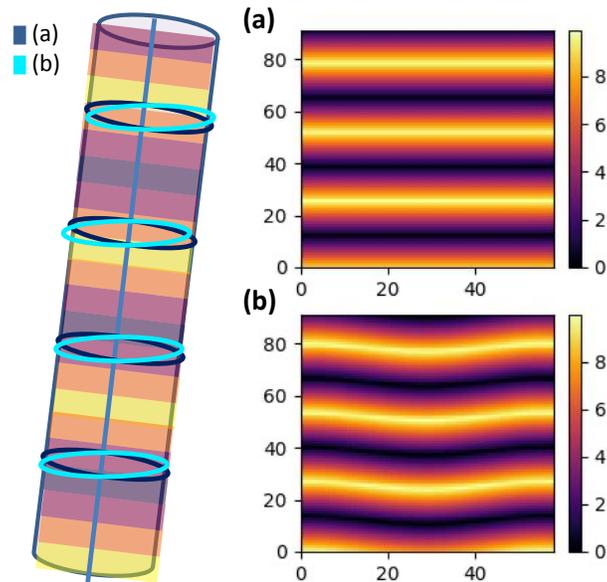


Figure 5.4: Visual comparison of DSMs created from a cylinder angled 10° in a striped dose distribution aligned with the cylinder’s central axis to demonstrate the effect that the choice of DSM calculation style has on DSM appearance. A film measurement of this scenario would feature parallel stripes of high and low dose. For calculated DSMs, the stripes will appear straight and match the film measurement for a non-coplanar DSM (a), but will be distorted and non-equivalent to the film for a planar DSM (b).

filling variations. More sophisticated deformable phantoms in the vein of the ADAM-pelvis and PETER phantoms or 3D-printed phantoms may be useful to evaluate these scenarios [37, 38]. We were also unable to assess the validity of the assumption of isotropic rectum expansion, which is a fundamental component of DSM calculation and dose accumulation. This could form the basis for future studies using deformable phantoms or biomechanical rectum models.

In summary, experimental validation of DSM-based dose accumulation for rectum structures showed good agreement with film measurements for simple inter-fraction motion scenarios. Provided the non-coplanar calculation method is used, DSM-based dose accumulation may be a viable alternative to DIR-based dose accumulation.

5.7 Acknowledgements

The authors gratefully acknowledge Joe Larkin for his assistance constructing the experimental phantom and Russel Ruo for his expertise with FilmQA Pro. HMP acknowledges funding from the Fonds de recherche du Québec - Santé in the form of a Doctoral Training Award, as well as partial support by the CREATE Responsible Health and Healthcare Data Science (SDRDS) grant of the Natural Sciences and Engineering Research Council. JK acknowledges support from the Canada Foundation for Innovation John R. Evans Leaders Fund. We also acknowledge the logistical support provided by our colleagues in the Department of Medical Physics and the Division of Radiation Oncology at the McGill University Health Centre.

The data that support the findings of this study are available from the corresponding author, HMP, upon reasonable request.

Bibliography

- [1] F. Buettner, S. L. Gulliford, S. Webb, and M. Partridge, “Using dose-surface maps to predict radiation-induced rectal bleeding: A neural network approach,” *Physics in Medicine and Biology*, 2009.
- [2] R. C. Wortel, M. G. Witte, U. A. van der Heide, F. J. Pos, J. V. Lebesque, M. van Herk, L. Incrocci, and W. D. Heemsbergen, “Dose-surface maps identifying local dose-effects for acute gastrointestinal toxicity after radiotherapy for prostate cancer,” *Radiotherapy and Oncology*, 2015.
- [3] J. E. Scaife, S. J. Thomas, K. Harrison, M. Romanchikova, M. P. F Sutcliffe, J. R. Forman, A. M. Bates, R. Jena, M. A. Parker, and N. G. Burnet, “Accumulated dose to the rectum, measured using dose-volume histograms and dose-surface maps, is different from planned dose in all patients treated with radiotherapy for prostate cancer,” *British Journal of Radiology*, vol. 88, p. 20150243, oct 2015.

-
- [4] L. E. Shelley, J. E. Scaife, M. Romanchikova, K. Harrison, J. R. Forman, A. M. Bates, D. J. Noble, R. Jena, M. A. Parker, M. P. Sutcliffe, S. J. Thomas, and N. G. Burnet, “Delivered dose can be a better predictor of rectal toxicity than planned dose in prostate radiotherapy,” *Radiotherapy and Oncology*, 2017.
- [5] J. Murray, D. McQuaid, A. Dunlop, F. Buettner, S. Nill, E. Hall, D. Dearnaley, and S. Gulliford, “SU-E-J-14: A Novel Approach to Evaluate the Dosimetric Effect of Rectal Variation During Image Guided Prostate Radiotherapy,” *Medical Physics*, 2014.
- [6] L. E. Shelley, M. P. Sutcliffe, S. J. Thomas, D. J. Noble, M. Romanchikova, K. Harrison, A. M. Bates, N. G. Burnet, and R. Jena, “Associations between voxel-level accumulated dose and rectal toxicity in prostate radiotherapy,” *Physics and Imaging in Radiation Oncology*, 2020.
- [7] O. Casares-Magaz, S. Bülow, N. J. Pettersson, V. Moiseenko, J. Pedersen, M. Thor, J. Einck, A. Hopper, R. Knopp, and L. P. Muren, “High accumulated doses to the inferior rectum are associated with late gastro-intestinal toxicity in a case-control study of prostate cancer patients treated with radiotherapy,” 2019.
- [8] M. van Herk, A. Bruce, A. P. Guus Kroes, T. Shouman, A. Touw, and J. V. Lebesque, “Quantification of organ motion during conformal radiotherapy of the prostate by three dimensional image registration,” *International Journal of Radiation Oncology, Biology, Physics*, vol. 33, pp. 1311–1320, dec 1995.
- [9] J. M. Crook, Y. Raymond, D. Salhani, H. Yang, and B. Esche, “Prostate motion during standard radiotherapy as assessed by fiducial markers,” *Radiotherapy and Oncology*, 1995.
- [10] J. Balter, H. M. Sandler, K. Lam, R. L. Bree, A. S. Lichter, and R. K. Ten Haken, “Measurement of prostate motion over the course of radiotherapy,” *International Journal of Radiation Oncology*Biological*Physics*, 1993.
- [11] M. J. Zelefsky, M. Kollmeier, B. Cox, A. Fidaleo, D. Sperling, X. Pei, B. Carver, J. Coleman, M. Lovelock, and M. Hunt, “Improved clinical outcomes with high-dose

- image guided radiotherapy compared with non-IGRT for the treatment of clinically localized prostate cancer,” *International Journal of Radiation Oncology Biology Physics*, 2012.
- [12] M. Marcello, M. Ebert, A. Haworth, A. Steigler, A. Kennedy, D. Joseph, and J. Denham, “Association between treatment planning and delivery factors and disease progression in prostate cancer radiotherapy: Results from the TROG 03.04 RADAR trial,” *Radiotherapy and Oncology*, 2018.
- [13] W. D. Heemsbergen, A. Al-Mamgani, M. G. Witte, M. Van Herk, and J. V. Lebesque, “Radiotherapy with rectangular fields is associated with fewer clinical failures than conformal fields in the high-risk prostate cancer subgroup: Results from a randomized trial,” *Radiotherapy and Oncology*, 2013.
- [14] R. C. Wortel, L. Incrocci, F. J. Pos, J. V. Lebesque, M. G. Witte, U. A. van der Heide, M. van Herk, and W. D. Heemsbergen, “Acute Toxicity After Image-Guided Intensity Modulated Radiation Therapy Compared to 3D Conformal Radiation Therapy in Prostate Cancer Patients,” *International Journal of Radiation Oncology*Biography*Physics*, vol. 91, pp. 737–744, mar 2015.
- [15] P. M. van Haaren, A. Bel, P. Hofman, M. van Vulpen, A. N. Kotte, and U. A. van der Heide, “Influence of daily setup measurements and corrections on the estimated delivered dose during IMRT treatment of prostate cancer patients,” *Radiotherapy and Oncology*, 2009.
- [16] S. Thörnqvist, J. B. Petersen, M. Høyer, L. N. Bentzen, and L. P. Muren, “Propagation of target and organ at risk contours in radiotherapy of prostate cancer using deformable image registration,” *Acta Oncologica*, 2010.
- [17] M. Thor, J. B. Petersen, L. Bentzen, M. Høyer, and L. P. Muren, “Deformable image registration for contour propagation from CT to cone-beam CT scans in radiotherapy of prostate cancer,” in *Acta Oncologica*, 2011.

- [18] S. Gao, L. Zhang, H. Wang, R. De Crevoisier, D. D. Kuban, R. Mohan, and L. Dong, “A deformable image registration method to handle distended rectums in prostate cancer radiotherapy,” *Medical Physics*, 2006.
- [19] M. Velec, J. L. Moseley, S. Svensson, B. Hårdemark, D. A. Jaffray, and K. K. Brock, “Validation of biomechanical deformable image registration in the abdomen, thorax, and pelvis in a commercial radiotherapy treatment planning system:,” *Medical Physics*, 2017.
- [20] N. Kadoya, Y. Y. Miyasaka, T. Yamamoto, Y. Kuroda, K. Ito, M. Chiba, Y. Nakajima, N. Takahashi, M. Kubozono, R. Umezawa, S. Dobashi, K. Takeda, and K. Jingu, “Evaluation of rectum and bladder dose accumulation from external beam radiotherapy and brachytherapy for cervical cancer using two different deformable image registration techniques,” *Journal of Radiation Research*, 2017.
- [21] K. Motegi, H. Tachibana, A. Motegi, K. Hotta, H. Baba, and T. Akimoto, “Usefulness of hybrid deformable image registration algorithms in prostate radiation therapy,” *Journal of Applied Clinical Medical Physics*, 2019.
- [22] A. D’Aquino, *Assessment of delivered rectal dose in prostate cancer radiotherapy*. PhD thesis, University Of London, 2021.
- [23] U. J. Yeo, M. L. Taylor, J. R. Supple, R. L. Smith, L. Dunn, T. Kron, and R. D. Franich, “Is it sensible to ”deform” dose 3D experimental validation of dose-warping,” *Medical Physics*, vol. 39, pp. 5065–5072, jul 2012.
- [24] R. De Crevoisier, S. L. Tucker, L. Dong, R. Mohan, R. Cheung, J. D. Cox, and D. A. Kuban, “Increased risk of biochemical and local failure in patients with distended rectum on the planning CT for prostate cancer radiotherapy,” *International Journal of Radiation Oncology Biology Physics*, 2005.
- [25] O. Barbosa Neto, L. Souhami, and S. Faria, “Hypofractionated radiation therapy for prostate cancer: The McGill University Health Center experience,” *Cancer/Radiotherapie*, vol. 19, pp. 431–436, oct 2015.

- [26] N. Agazaryan, “Patient specific quality assurance for the delivery of intensity modulated radiotherapy,” *Journal of Applied Clinical Medical Physics*, 2003.
- [27] S. Aldelaijan, S. Devic, H. Mohammed, N. Tomic, L. H. Liang, F. DeBlois, and J. Seuntjens, “Evaluation of EBT-2 model GAFCHROMIC™ film performance in water,” *Medical Physics*, 2010.
- [28] D. Lewis, A. Micke, X. Yu, and M. F. Chan, “An efficient protocol for radiochromic film dosimetry combining calibration and measurement in a single scan,” *Medical Physics*, vol. 39, no. 10, pp. 6339–6350, 2012.
- [29] H. M. Patrick and J. Kildea, “Technical note: rtdsm-An open-source software for radiotherapy dose-surface map generation and analysis.,” *Medical physics*, vol. 49, pp. 7327–7335, nov 2022.
- [30] D. A. Low, W. B. Harms, S. Mutic, and J. A. Purdy, “A technique for the quantitative evaluation of dose distributions,” *Medical Physics*, 1998.
- [31] M. Miften, A. Olch, D. Mihailidis, J. Moran, T. Pawlicki, A. Molineu, H. Li, K. Wijesooriya, J. Shi, P. Xia, N. Papanikolaou, and D. A. Low, “Tolerance limits and methodologies for IMRT measurement-based verification QA: Recommendations of AAPM Task Group No. 218,” *Medical Physics*, 2018.
- [32] D. A. Jaffray, P. E. Lindsay, K. K. Brock, J. O. Deasy, and W. A. Tomé, “Accurate Accumulation of Dose for Improved Understanding of Radiation Effects in Normal Tissue,” *International Journal of Radiation Oncology Biology Physics*, vol. 76, pp. S135–9, mar 2010.
- [33] O. Acosta, G. Drean, J. D. Ospina, A. Simon, P. Haigron, C. Lafond, and R. De Crevoisier, “Voxel-based population analysis for correlating local dose and rectal toxicity in prostate cancer radiotherapy,” *Physics in Medicine and Biology*, 2013.
- [34] F. Palorini, C. Cozzarini, S. Gianolini, A. Botti, V. Carillo, C. Iotti, T. Rancati, R. Valdagni, and C. Fiorino, “First application of a pixel-wise analysis on bladder dose-surface maps in prostate cancer radiotherapy,” *Radiotherapy and Oncology*, 2016.

- [35] D. Hoffmans, N. Niebuhr, O. Bohoudi, A. Pfaffenberger, and M. Palacios, “An end-to-end test for MR-guided online adaptive radiotherapy,” *Physics in Medicine and Biology*, 2020.
- [36] O. Bohoudi, F. J. Lagerwaard, A. M. Bruynzeel, N. I. Niebuhr, W. Johnen, S. Senan, B. J. Slotman, A. Pfaffenberger, and M. A. Palacios, “End-to-end empirical validation of dose accumulation in MRI-guided adaptive radiotherapy for prostate cancer using an anthropomorphic deformable pelvis phantom,” *Radiotherapy and Oncology*, 2019.
- [37] A. Elter, S. Dorsch, P. Mann, A. Runz, W. Johnen, C. K. Spindeldreier, S. Klüter, and C. P. Karger, “End-to-end test of an online adaptive treatment procedure in MR-guided radiotherapy using a phantom with anthropomorphic structures,” *Physics in Medicine and Biology*, 2019.
- [38] M. Marot, A. Elter, P. Mann, A. Schwahofer, C. Lang, W. Johnen, S. A. Körber, B. Beuthien-Baumann, and C. Gillmann, “Technical Note: On the feasibility of performing dosimetry in target and organ at risk using polymer dosimetry gel and thermoluminescence detectors in an anthropomorphic, deformable, and multimodal pelvis phantom,” *Medical Physics*, 2021.
- [39] S. Calusi, G. Labanca, M. Zani, M. Casati, L. Marrazzo, L. Noferini, C. Talamonti, F. Fusi, I. Desideri, P. Bonomo, L. Livi, and S. Pallotta, “A multiparametric method to assess the MIM deformable image registration algorithm,” *Journal of Applied Clinical Medical Physics*, 2019.
- [40] P. B. Johnson, K. R. Padgett, K. L. Chen, and N. Dogan, “Evaluation of the tool “Reg Refine” for user-guided deformable image registration,” *Journal of Applied Clinical Medical Physics*, 2016.
- [41] N. Kirby, C. Chuang, U. Ueda, and J. Pouliot, “The need for application-based adaptation of deformable image registration,” *Medical Physics*, 2013.

Chapter 6

A spatial investigation of real-world rectal dose delivery accuracy during radiotherapy for prostate cancer

Haley M Patrick and John Kildea

Manuscript submitted to *Journal of Applied Clinical Medical Physics*.

6.1 Preface

In chapter 5 we demonstrated that DSM accumulation yields accurate representations of total delivered dose to a rectum phantom when the non-coplanar calculation approach is used. Here we extend this finding to demonstrate the use of non-coplanar DSM accumulation to evaluate the delivery accuracy of dose to the rectum during hypofractionated prostate cancer radiotherapy. This study was motivated by the recent increase in the use of hypofractionated treatments since 2020 for which few studies have been reported, including for SBRT in particular. Furthermore, DSMs are currently underutilized for planning versus delivered dose studies, despite the added value they may provide over conventional DVHs. We therefore investigated the differences between planned and accumulated doses for SBRT using DSMs and their relation to changes in rectum shape over the course of treatment.

Using daily CBCT images, we calculated the daily delivered rectum doses for 40 prostate radiotherapy patients: 20 patients prescribed 36.25 Gy in 5 fraction (SBRT) and 20 patients prescribed 60 Gy in 20 fractions (IMRT). We then calculated planned, daily, and accumulated DSMs for each patient using our tested non-coplanar DSM accumulation approach. Differences between planned and delivered rectal dose were evaluated on a patient-specific and cohort-wide basis using MCP analysis. Additionally, we examined changes in rectal volume and wall positioning over the course of treatment and evaluated if they may be related to dosimetric changes. The results of our analysis showed that most of the SBRT treatments were delivered similarly to what was planned, whereas many of the IMRT treatments were delivered with significant dose deviations, especially within the posterior rectal wall region. The presence or lack of dosimetric variations in both cohorts could be explained by the stability of patients' rectal wall position over the course of treatment, which was much more informative than changes in rectal volume. These results indicate that using spatially inclusive forms of dosimetric and rectal shape information like DSMs offer additional nuance that may better explain dose difference between planning and delivery compared to just using conventional DVH metrics.

6.2 Abstract

Purpose: Prostate SBRT requires precise treatment delivery to ensure adequate target coverage and rectum sparing. However, the accuracy at which rectum dose is delivered during prostate SBRT is not well characterized, especially in a spatial context. Our aim was, therefore, to quantify the similarity of planned and delivered rectum dose in a spatially-conscious manner using dose-surface maps and to compare it to the corresponding result obtained for a well-established hypofractionated treatment scheme.

Methods: Prostate cancer patients treated with 36.25 Gy in 5 fractions (SBRT) and 60 Gy in 20 fractions (IMRT) were selected from our hospital database. Daily delivered doses for each patient were calculated using daily CBCT images. Planned and accumulated dose-surface maps of the rectal wall were then calculated and compared using multiple-comparisons permutation testing to identify regions where significant dose

differences occurred. Changes in rectal wall position between planning and delivery were also evaluated to determine possible relation to dosimetric changes.

Results: Statistically speaking, treatments were delivered consistent with planning in 14 out of 20 SBRT patients and in 8 out of 20 IMRT patients. On average, patients in the IMRT group received significantly lower doses to the posterior rectal wall during treatment than what was planned, which could be attributed to significant shifts in posterior wall position over the course of treatment. No significant dosimetric differences were observed for the average SBRT patient, who exhibited stable rectal wall positioning over the course of treatment.

Conclusions: We found that in general, delivered dose to the rectal wall during prostate SBRT is statistically the same as planned dose and any delivered dose deviations from planning during both SBRT and IMRT treatments are consistent with rectal wall motion.

6.3 Introduction

Stereotactic body radiotherapy (SBRT) is a promising treatment modality for prostate cancer delivering doses of ≥ 5 Gy per fraction to exploit prostate cancer's low α/β ratio and improve the therapeutic ratio [1]. Adoption rates for this modality have increased since 2020 in response to pandemic practice recommendations [2, 3] and are supported by reports of non-inferior 5-year failure-free survival and toxicity incidence compared to standard fractionation in clinical trials [4–6]. However, SBRT's short treatment course means each fraction heavily contributes to the total dose, making accurate fraction delivery important to ensure adequate organ at risk (OAR) sparing and to avoid radiation toxicities. Therefore, considering the potential impact on patient outcomes, it is important to quantify the level to which delivered and planned OAR doses differ in routine clinical usage of prostate SBRT.

Over the years, multiple studies have compared the delivered doses to the prostate and rectum to planning values to assess delivery accuracy for prostate radiotherapy. Using 3D anatomical image guided radiotherapy (IGRT) images to calculate daily doses, these studies have demonstrated that while IGRT is able minimize deviations in target dose to within 1 Gy over a course of conventionally-fractionated (2 Gy/fraction) treatment, large

variations in the volume of the rectum exposed to near-prescription dose can still occur [7–10]. Studies are much more limited, however, for hypofractionated treatments (> 2 Gy/fraction), and in particular for SBRT treatments. Most publications that report prostate SBRT accumulated doses center around treatments performed on MR-linacs, with a large focus on adaptive workflows [11–15]. While some investigations of SBRT treatments delivered with conventional linacs do exist [16–18], they focus more on target dose delivery and are limited in their investigations of rectal delivered dose. Also, these studies evaluated planned and delivered OAR dose differences using dose-volume histogram (DVH) metrics, thereby excluding the opportunity to compare dose differences in a spatially conscious manner.

DVHs have long been recognized to be limited by their lack of spatial information [19], making it possible for differences between dose distributions to be masked if they share similar V_xGy metrics. One alternative to the DVH that preserves spatial information is the dose-surface map (DSM). A DSM provides a 2D representation of the dose to the surface of a structure. DSMs have become a popular tool for dose-outcome studies of hollow organs, especially the rectum [20–23], and have recently been used to calculate total accumulated rectum dose for outcome studies of prostate radiotherapy [24, 25]. Although relatively unexplored for their usefulness to compare planned and delivered doses, the spatial dose information contained within DSMs means they can provide an enhanced picture of radiotherapy delivery accuracy compared to DVHs.

The purpose of this study was to evaluate dosimetric differences in real-world rectal delivered dose from planning during prostate SBRT compared to a well-established hypofractionated treatment (60 Gy/20 fractions) at our centre. Specifically, we compared planned and delivered doses using DSMs within both fractionation schemes, in order to compare dose delivery accuracy between the schemes on a spatial level, and we examined rectal shape variations to assess their contribution to any inaccuracies observed.

6.4 Methods

6.4.1 Patient Cohorts

Two retrospective, single center patient cohorts were identified for this dosimetric study. The first cohort consisted of twenty randomly selected patients with localized prostate cancer treated with 60 Gy in 20 fractions (IMRT), who were treated between September 2015-2016 following the move of our radiation oncology department to a new hospital. The second cohort consisted of twenty patients with localized prostate cancer treated consecutively with 36.25 Gy in 5 fractions (SBRT) starting from May 2020, two months after SBRT replaced the well-established 60 Gy/20 fractions hypofractionation as our standard treatment. All patients underwent a simulation CT scan in a Philips Big Bore CT scanner, with the SBRT patients also undergoing an MRI simulation scan. The target was defined as the prostate plus 7- or 5-mm isotropic margins for the IMRT and SBRT cohorts, respectively. OARs were contoured according to RTOG guidelines [26], beginning at the ischial tuberosities and finishing at the sigmoid junction. Volumetric Modulated Arc therapy (VMAT) plans consisting of two 6 MV arcs were designed for each patient in the Eclipse Treatment Planning System (Varian Medical Systems, Palo Alto, CA) in accordance with appropriate protocols: local guidelines for the IMRT cohort [27], and NRG-GU005 guidelines for the SBRT cohort [28].

Treatments were delivered daily (IMRT) or every second day (SBRT) using cone-beam CT (CBCT) IGRT guidance. For the IMRT cohort, the IGRT protocol used a single daily pre-treatment CBCT to perform soft tissue matching before treatment delivery, whereas the SBRT cohort's IGRT protocol also mandated the acquisition of a daily post-delivery CBCT to verify that no significant positional shifts had occurred during treatment. Additional pre-treatment CBCTs were allowed for both cohorts in the event that large set-up adjustments were required. To retrospectively calculate daily delivered doses for each fraction, treatment beams were registered to the last-recorded CBCT images of the fraction in Eclipse and dose calculations re-run on the daily anatomies. Rectum contours were retrospectively delineated on all daily CBCT images by a single observer to minimize interobserver contouring variations. Copies of the originally-planned rectum contours were

also reviewed and adjusted, if necessary, by the same observer to ensure retrospective contouring consistency across all patients.

6.4.2 DSM Calculation

Dose and contour data were exported as DICOM files from the treatment planning system to calculate DSMs. Generation of DSMs was achieved using *rtdsm*, an open-source Python package previously developed by our group for DSM calculation and analysis [29]. For each plan and daily image, *rtdsm* began by defining slices of the rectum contour orthogonal to its central-axis path in increments of 3 mm, correcting the orientation of intersecting slices using the approach of Witztum *et al* [23]. Next, 45 equiangular points were defined around the outer circumference of each slice and dose was sampled at these points. Finally, each contour was unwrapped along its posterior rectal wall and mapped into a 2D array. To calculate an accumulated DSM for each patient, their daily DSMs were aligned at their inferior borders and summed together, truncating longer DSMs to the length of the patient’s shortest daily rectum contour. This DSM alignment approach was also used when calculating cohort DSM averages.

6.4.3 Comparison of Planned and Delivered Doses

Planned and delivered DSMs were compared on a patient-by-patient basis and cohort-wide for both fractionation schemes. Dose differences between planned and delivered DSMs were visualized by calculating dose-difference maps (DDMs). As for the DSMs, DDMs were calculated on a patient-by-patient basis and for the cohort averages.

Statistical comparisons of planned and delivered doses were conducted using multiple comparisons permutation (MCP) testing [30]. Patient-wise comparisons of planned and accumulated DSMs were performed using a one-sample version of the test that compared the distribution of each patient’s daily-delivered DSMs (scaled up to prescription doses) to their planned DSM. The cohort-wise comparisons used a paired version of the test to compare planned and accumulated DSMs for each treatment cohort. Statistical significance for all tests was defined as $p\text{-value} \leq 0.05$.

6.4.4 Influence of Rectal Shape Metrics

As changes in rectum dose are frequently attributed to inter-fraction changes in rectum size and shape, we investigated how these factors varied over the course of treatment to determine if they played a role in the dose differences observed. To facilitate this, planned and daily rectal volumes were extracted from the treatment planning system and percent change from planning baseline was calculated for each patient to evaluate inter-fraction change in rectum volume. Localized changes in rectum shape were quantified by extracting point position information for the anterior and posterior rectum walls as identified in the DSM calculation process relative to the linac isocentre position for each day. This positional information was then used to determine daily anterior-posterior shifts in wall position relative to the planning baseline. The significance of deviations in rectal volume and wall positions from planning baseline was assessed using Wilcoxon signed-rank testing.

6.5 Results

6.5.1 Comparison of Planned and Delivered Doses

Planned and delivered doses were observed to differ for all patients in both cohorts to some extent. Examples of these differences are shown for a selection of patients in Figure 6.1 (and for all patients in the [supplementary material](#)). A majority of the IMRT patients (12/20) exhibited statistically significant delivered dose deviations from their planning baselines for at least one subregion of the rectal wall, whereas the same was true for only a minority (6/20) of the SBRT patients. In general, these subregions occurred in areas superior, inferior, or lateral to the portion of the anterior rectal wall in direct contact with the prostate target and were larger in surface area for the IMRT patients compared to the SBRT ones.

The cohort-average DDMs for the IMRT and SBRT cohorts are shown in Figure 6.2. Patients in the IMRT cohort were found to receive systematically lower doses to the posterior rectal wall than planned. No significant dose differences were observed on a cohort level for the SBRT patients.

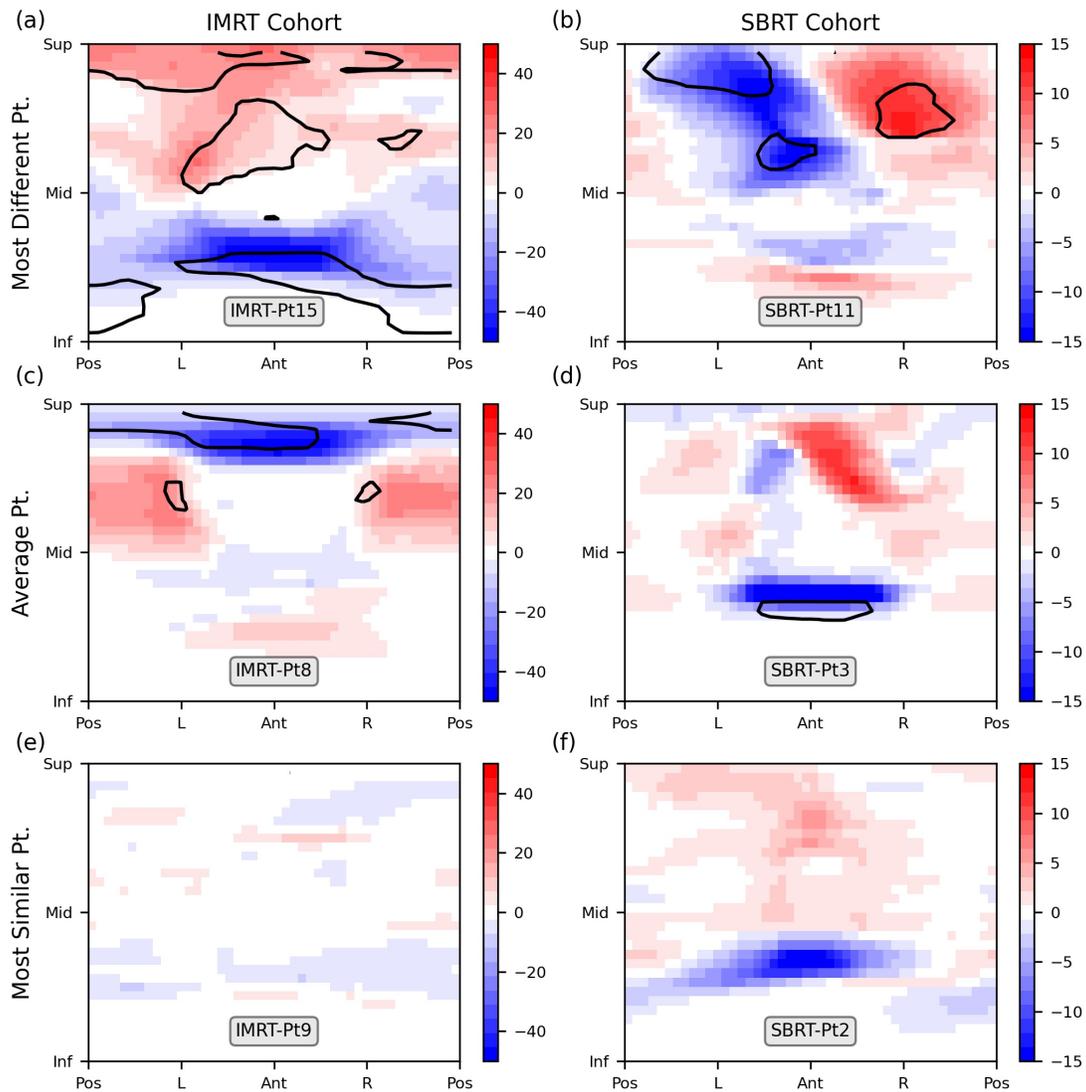


Figure 6.1: Planned minus accumulated-delivered dose difference maps (DDMs) of the patients with the most different (a-b), median different (c-d), and most similar (e-f) planned and accumulated DSMs, in units of Gy. Patients in the IMRT (60 Gy in 20 fractions) cohort are in the left column and patients in the SBRT cohort (36.25 Gy in 5 fractions) are in the right column. Subregions with statistically significant dose differences are contoured in black. Patients are identified by the grey labels.

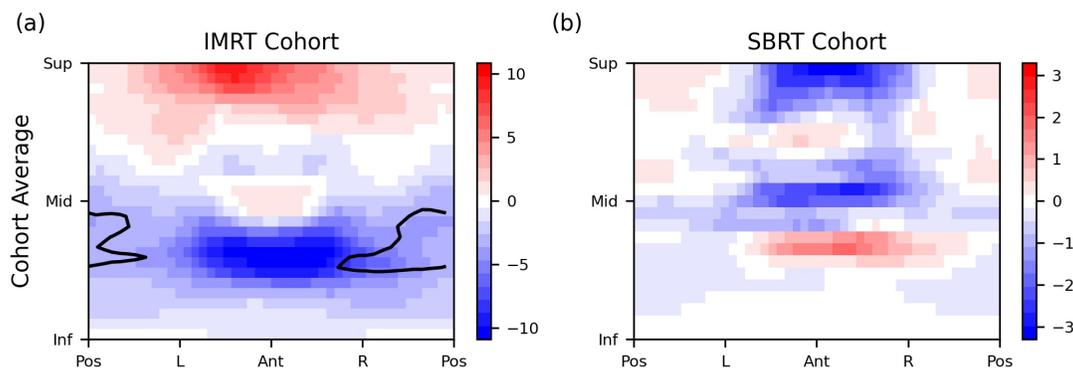


Figure 6.2: Cohort-average dose difference maps for the IMRT (a) and SBRT (b) cohorts.

6.5.2 Influence of Rectal Shape Metrics

Daily rectal volume variations were observed in all patients (Fig 6.3). In total, 18/20 IMRT and 10/20 SBRT patients had statistically significantly different rectum volumes during treatment compared to their planning volume. Interestingly, the SBRT patients tended to have smaller rectal volumes during treatment compared to at the time of planning, whereas no such cohort-wide pattern existed for the IMRT patients. It is also worth noting that not all cases of statistically significant deviations in rectal volume from planning baseline were associated with statistically significant deviations in delivered dose. For example, IMRT patient 9 had significantly smaller rectal volumes during treatment but had the most similar planned and delivered DSMs (Fig 6.1a), whereas SBRT patient 11's planning volume was not statistically different from their delivery volumes despite having the largest dose differences between planning and delivery (Fig 6.1b).

Changes in rectal wall position between planning and delivery are shown for both cohorts in Figure 6.4. Rectal wall positioning was relatively stable throughout the course of treatment for the SBRT cohort, with significant shifts only observed for small segments of the posterior wall located inferior or superior to the level of the PTV (Fig 6.1b). Rectal wall positioning was much less consistent, however, for the IMRT cohort. On average, patients' inferior posterior rectal walls were observed to have shifted significantly further from the target during treatment compared to their position at planning (Fig 6.4a). The superior-most points in this region corresponded with the region of significant dose difference (Fig 6.2a)

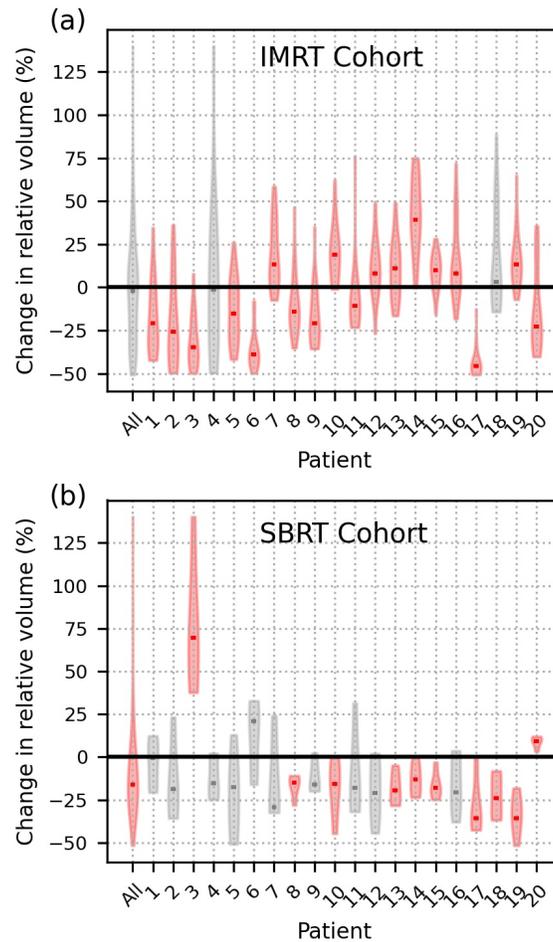


Figure 6.3: Violin plots of the relative change (planned minus delivered) in daily rectal volume from planning value for the overall IMRT and SBRT cohorts (All), as well as individually numbered patients. Statistically significant results are indicated in red. Distribution medians are shown as rectangular markers.

seen in the DSMs. This shift was present within the delivery of the first five fractions (the same duration as the complete SBRT treatment) and increased in magnitude by the last five fractions (Fig 6.4c-d).

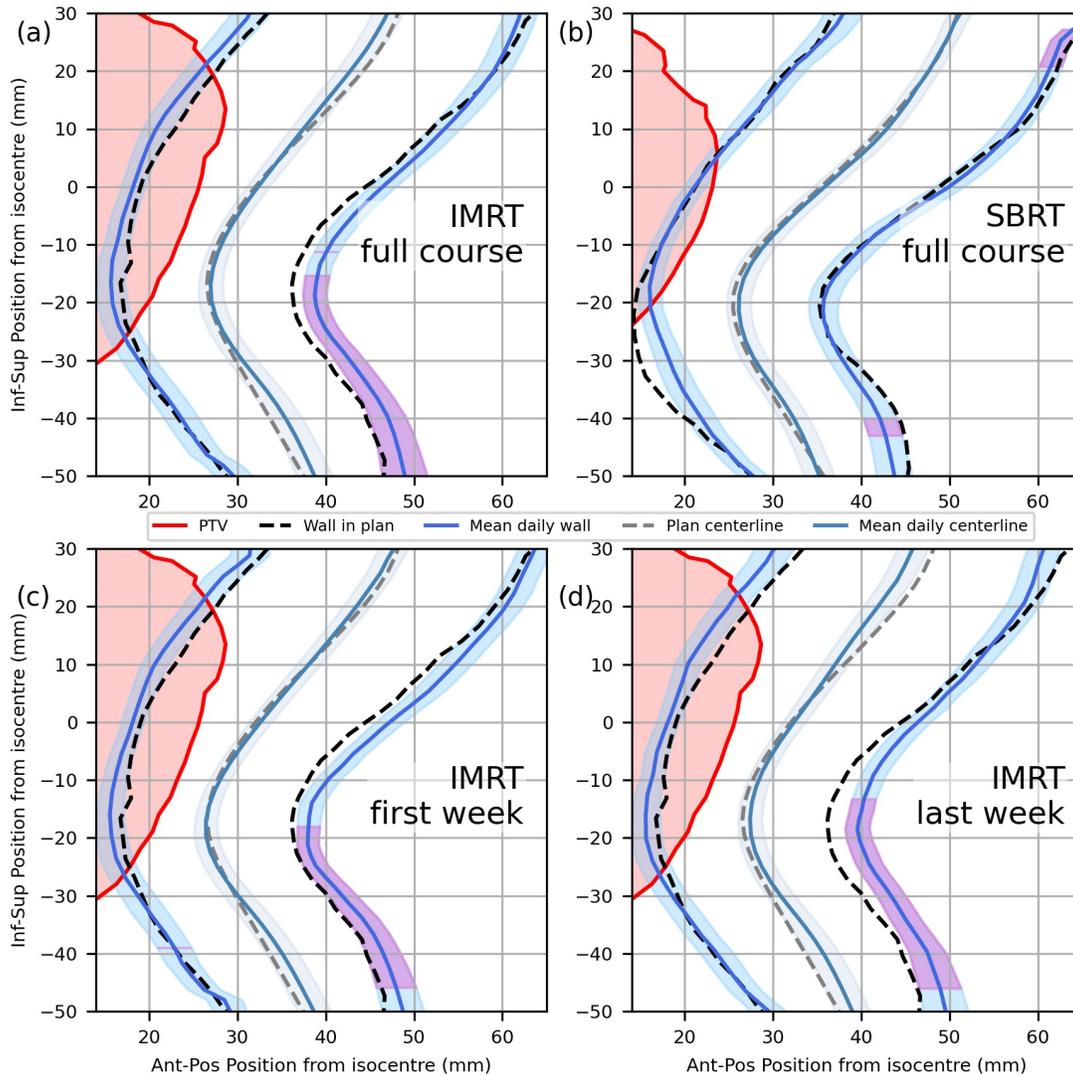


Figure 6.4: Mean rectal wall and PTV positions at planning and during treatment for the IMRT (a) and SBRT (b) cohorts, as well as their mean positions over the first (c) and last (d) five fractions for the IMRT cohort. Shaded regions indicate standard uncertainty of the mean, with purple areas indicating regions of statistically significant anterior-posterior shifts.

6.6 Discussion

The variation of daily delivered rectal dose over the course of prostate radiotherapy is a well-documented phenomenon that can lead to different total delivered dose than what was

planned. While this effect is well-characterized for conventionally-fractionated treatments, characterization for real-world SBRT treatments has been limited, especially outside of MR-linac adaptive workflows. To address this, we quantified the spatial dose differences between planned and delivered rectal wall dose for patients treated with prostate SBRT and compared them to a well-established hypofractionated IMRT regimen.

In this study, we found that most patients treated with SBRT received similar rectal doses to what was planned and that the statistically significant dose deviations that did occur in a minority of patients (6/20) were limited to small portions of the rectal wall. IMRT patients, on the other hand, were found to experience significant dose deviations much more frequently (12/20) and to larger areas. Similar patterns were also observed on a cohort level, as IMRT patients were found on average to receive statistically significant lower doses to the posterior rectal wall than what was planned whereas SBRT patients on average did not exhibit statistically significant dose deviations. This did not confirm our concern, as shared with Shelley *et al* [31], that the short-course SBRT treatments might deviate more from planned due to higher sensitivity to single-fraction dose deviations. However, our IMRT results are supported by previous studies which have reported similar patterns in dose delivery deviations for individual patients [32, 33] and on a cohort-level [34] for conventionally-fractionated treatments. Similarly, while their investigation was restricted to the V100% metric, Studenski *et al* found no statistically significant dose deviations between planned and delivered dose for the same prescription as our SBRT cohort [16]. One possible explanation may be that SBRT's higher dose-per-fraction influenced clinical staff to adhere more stringently to patient set-up, IGRT, and bladder/rectal filling protocols than they may for less hypofractionated treatments. For example, the main difference between our clinic's IMRT and SBRT IGRT protocols was the acquisition of a second post-fraction CBCT for the SBRT treatment to check for mid-treatment motion. This, paired with the generally reinforced importance of delivery accuracy for SBRT [35], could be contributing to the effect. Another potential explanation could also be that the tighter margins used for the SBRT treatment may give rise to steeper dose gradients that could lead to a different dosimetric impact for the same unit shift in rectal wall position compared to the IMRT treatment.

While changes in rectal doses have long been attributed to changes in rectal volume

[36–38], we did not observe clear evidence of this phenomenon in our data. Although patients in our more dosimetrically variable IMRT cohort were observed to exhibit more volumetric variations than the SBRT cohort, no clear pattern between volume variations and DSM differences emerged. Instead, we found that changes in rectal wall position provided a much better explanation of where and why dosimetric differences were occurring (Fig 6.4). The localized dosimetric variations present in the IMRT cohort could be easily explained by a gradual migration of the posterior rectal wall away from the PTV over the course of treatment, whereas a lack of wall migration could explain the SBRT cohort results. Our spatial investigation is, to the best of our knowledge, unique, with the closest reported finding being correlation of rectal gEUD with anterior-posterior shifts in rectum centroid positions [34], which are more proxy data. Our finding would have been masked if only DVH metrics and total rectal volumes were used.

An important limitation of our study is that our calculations of delivered dose do not account for intra-fraction motion during fraction delivery. This is a product of our SBRT IGRT protocol, which does not include real-time prostate tracking. However, conventional C-arm linac-based SBRT treatments can be delivered in a much shorter time compared to Cyberknife or Tomotherapy machines (10-15 minutes versus 20-40 minutes), reducing the window for large intra-fraction motion to occur. Intra-fraction motion has also been reported to have minimal impact on rectum dose during prostate SBRT [13, 39] over the duration of a conventional or MR- linac treatment. Additionally, we did attempt to consider the effects of intra-fraction motion by using post-treatment CBCTs for our daily dose calculation, which would reflect any positional deviations from the original soft-tissue matches.

6.7 Conclusions

In summary, we spatially evaluated rectal dose deviations over the course of SBRT and conventionally-hypofractionated radiotherapy treatments to the prostate in routine clinical practice using DSMs. SBRT treatments were shown to be delivered with fewer dosimetric deviations than the conventional hypofractionated treatments and localized dose deviations were found to be associated with localized changes in rectal wall position over the course of

treatment. These results demonstrate the accuracy to which SBRT treatments are delivered in routine practice and illustrate the value of DSMs for the identification spatial dose deviations. Future work in this area should focus on dosimetric evaluations of other fractionation schemes and OARs, as well as potential systematic factors that may have contributed to the dose deviations observed for the IMRT cohort.

6.8 Acknowledgements

HMP acknowledges funding from the Fonds de recherche du Québec - Santé in the form of a Doctoral Training Award, as well as partial support by the CREATE Responsible Health and Healthcare Data Science (SDRDS) grant of the Natural Sciences and Engineering Research Council. JK acknowledges support from the Canada Foundation for Innovation John R. Evans Leaders Fund. We also acknowledge the logistical support provided by our colleagues in the Department of Medical Physics and the Division of Radiation Oncology at the McGill University Health Centre.

Bibliography

- [1] E. F. Miles and W. Robert Lee, “Hypofractionation for Prostate Cancer: A Critical Review,” *Seminars in Radiation Oncology*, 2008.
- [2] N. G. Zaorsky, J. B. Yu, S. M. McBride, R. T. Dess, W. C. Jackson, B. A. Mahal, R. Chen, A. Choudhury, A. Henry, I. Syndikus, T. Mitin, A. Tree, A. U. Kishan, and D. E. Spratt, “Prostate Cancer Radiation Therapy Recommendations in Response to COVID-19,” *Advances in Radiation Oncology*, 2020.
- [3] K. Spencer, C. M. Jones, R. Girdler, C. Roe, M. Sharpe, S. Lawton, L. Miller, P. Lewis, M. Evans, D. Sebag-Montefiore, T. Roques, R. Smittenaar, and E. Morris, “The impact of the COVID-19 pandemic on radiotherapy services in England, UK: a population-based study,” *The Lancet Oncology*, 2021.

- [4] A. Widmark, A. Gunnlaugsson, L. Beckman, C. Thellenberg-Karlsson, M. Hoyer, M. Lagerlund, J. Kindblom, C. Ginman, B. Johansson, K. Björnlinger, M. Seke, M. Agrup, P. Fransson, B. Tavelin, D. Norman, B. Zackrisson, H. Anderson, E. Kjellén, L. Franzén, and P. Nilsson, “Ultra-hypofractionated versus conventionally fractionated radiotherapy for prostate cancer: 5-year outcomes of the HYPO-RT-PC randomised, non-inferiority, phase 3 trial,” *The Lancet*, 2019.
- [5] P. Fransson, P. Nilsson, A. Gunnlaugsson, L. Beckman, B. Tavelin, D. Norman, C. Thellenberg-Karlsson, M. Hoyer, M. Lagerlund, J. Kindblom, C. Ginman, B. Johansson, K. Björnlinger, M. Seke, M. Agrup, B. Zackrisson, E. Kjellén, L. Franzén, and A. Widmark, “Ultra-hypofractionated versus conventionally fractionated radiotherapy for prostate cancer (HYPO-RT-PC): patient-reported quality-of-life outcomes of a randomised, controlled, non-inferiority, phase 3 trial,” *The Lancet Oncology*, 2021.
- [6] D. H. Brand, A. C. Tree, P. Ostler, H. van der Voet, A. Loblaw, W. Chu, D. Ford, S. Tolan, S. Jain, A. Martin, J. Staffurth, P. Camilleri, K. Kancherla, J. Frew, A. Chan, I. S. Dayes, D. Henderson, S. Brown, C. Cruickshank, S. Burnett, A. Duffton, C. Griffin, V. Hinder, K. Morrison, O. Naismith, E. Hall, N. van As, D. Dodds, E. Lartigau, S. Patton, A. Thompson, M. Winkler, P. Wells, T. Lymberiou, D. Saunders, M. Vilarino-Varela, P. Vavassis, T. Tsakiridis, R. Carlson, G. Rodrigues, J. Tanguay, S. Iqbal, S. Morgan, A. Mihai, A. Li, O. Din, M. Panades, R. Wade, Y. Rimmer, J. Armstrong, and N. Oommen, “Intensity-modulated fractionated radiotherapy versus stereotactic body radiotherapy for prostate cancer (PACE-B): acute toxicity findings from an international, randomised, open-label, phase 3, non-inferiority trial,” *The Lancet Oncology*, 2019.
- [7] P. M. van Haaren, A. Bel, P. Hofman, M. van Vulpen, A. N. Kotte, and U. A. van der Heide, “Influence of daily setup measurements and corrections on the estimated delivered dose during IMRT treatment of prostate cancer patients,” *Radiotherapy and Oncology*, vol. 90, pp. 291–298, mar 2009.

- [8] D. J. Fraser, Y. Chen, E. Poon, F. L. Cury, T. Falco, and F. Verhaegen, “Dosimetric consequences of misalignment and realignment in prostate 3DCRT using intramodality ultrasound image guidance,” *Medical Physics*, 2010.
- [9] A. Godley, E. Ahunbay, C. Peng, and X. Allen Li, “Accumulating daily-varied dose distributions of prostate radiation therapy with soft-tissue-based KV CT guidance,” *Journal of Applied Clinical Medical Physics*, 2012.
- [10] T. C. Huang, K. T. Chou, S. N. Yang, C. K. Chang, J. A. Liang, and G. Zhang, “Fractionated changes in prostate cancer radiotherapy using cone-beam computed tomography,” *Medical Dosimetry*, 2015.
- [11] O. Bohoudi, A. M. Bruynzeel, S. Tetar, B. J. Slotman, M. A. Palacios, and F. J. Lagerwaard, “Dose accumulation for personalized stereotactic MR-guided adaptive radiation therapy in prostate cancer,” *Radiotherapy and Oncology*, 2021.
- [12] R. Ruggieri, M. Rigo, S. Naccarato, D. Gurrera, V. Figlia, R. Mazzola, F. Ricchetti, L. Nicosia, N. Giaj-Levra, F. Cuccia, C. Vitale, N. Stavreva, D. S. Pressyanov, P. Stavrev, R. Pellegrini, and F. Alongi, “Adaptive SBRT by 1.5 T MR-linac for prostate cancer: On the accuracy of dose delivery in view of the prolonged session time,” *Physica Medica*, 2020.
- [13] I. Wahlstedt, N. Andratschke, C. P. Behrens, S. Ehrbar, H. S. Gabryś, H. G. Schüler, M. Guckenberger, A. G. Smith, S. Tanadini-Lang, J. D. Tascón-Vidarte, I. R. Vogelius, and J. E. van Timmeren, “Gating has a negligible impact on dose delivered in MRI-guided online adaptive radiotherapy of prostate cancer.,” *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*, vol. 170, pp. 205–212, may 2022.
- [14] J. Winter, C. Johnstone, V. Kong, J. Dang, P. Chung, J. Padayachee, D. Letourneau, and T. Tadic, “Impact of Inter-and Intra-Fraction Motion on Delivered Dose for Prostate SBRT Patients Treated on an MRL-Linac Using a Novel Semi-Automated Dose Accumulation Framework,” *International Journal of Radiation Oncology*Biological*Physics*, 2021.

- [15] J. Yang, S. Vedam, B. Lee, P. Castillo, A. Sobremonte, N. Hughes, M. Mohammedsaid, J. Wang, and S. Choi, “Online adaptive planning for prostate stereotactic body radiotherapy using a 1.5 Tesla magnetic resonance imaging-guided linear accelerator,” *Physics and Imaging in Radiation Oncology*, 2021.
- [16] M. T. Studenski, R. Delgadillo, Y. Xu, J. Both, K. Padgett, M. Abramowitz, J. C. Ford, A. Dal Pra, A. Pollack, and N. Dogan, “Margin verification for hypofractionated prostate radiotherapy using a novel dose accumulation workflow and iterative CBCT,” *Physica Medica*, 2020.
- [17] C. G. Muurholm, T. Ravkilde, R. De Roover, S. Skouboe, R. Hansen, W. Crijns, T. Depuydt, and P. R. Poulsen, “Experimental investigation of dynamic real-time rotation-including dose reconstruction during prostate tracking radiotherapy,” *Medical Physics*, vol. 49, no. 6, pp. 3574–3584, 2022.
- [18] V. Faccenda, D. Panizza, M. C. Daniotti, R. Pellegrini, S. Trivellato, P. Caricato, R. Lucchini, E. De Ponti, and S. Arcangeli, “Dosimetric Impact of Intrafraction Prostate Motion and Interfraction Anatomical Changes in Dose-Escalated Linac-Based SBRT.,” *Cancers*, vol. 15, feb 2023.
- [19] V. Moiseenko, J. Van Dyk, J. Battista, and E. Travis, “Limitations in using dose-volume histograms for radiotherapy dose optimization,” in *The Use of Computers in Radiation Therapy*, 2000.
- [20] F. Buettner, S. L. Gulliford, S. Webb, M. R. Sydes, D. P. Dearnaley, and M. Partridge, “Assessing correlations between the spatial distribution of the dose to the rectal wall and late rectal toxicity after prostate radiotherapy: An analysis of data from the MRC RT01 trial (ISRCTN 47772397),” *Physics in Medicine and Biology*, 2009.
- [21] R. C. Wortel, M. G. Witte, U. A. van der Heide, F. J. Pos, J. V. Lebesque, M. van Herk, L. Incrocci, and W. D. Heemsbergen, “Dose–surface maps identifying local dose–effects for acute gastrointestinal toxicity after radiotherapy for prostate cancer,” *Radiotherapy and Oncology*, 2015.

- [22] L. E. Shelley, J. E. Scaife, M. Romanchikova, K. Harrison, J. R. Forman, A. M. Bates, D. J. Noble, R. Jena, M. A. Parker, M. P. Sutcliffe, S. J. Thomas, and N. G. Burnet, “Delivered dose can be a better predictor of rectal toxicity than planned dose in prostate radiotherapy,” *Radiotherapy and Oncology*, 2017.
- [23] A. Witztum, B. George, S. Warren, M. Partridge, and M. A. Hawkins, “Unwrapping 3D complex hollow organs for spatial dose surface analysis,” *Medical Physics*, 2016.
- [24] L. E. Shelley, M. P. Sutcliffe, S. J. Thomas, D. J. Noble, M. Romanchikova, K. Harrison, A. M. Bates, N. G. Burnet, and R. Jena, “Associations between voxel-level accumulated dose and rectal toxicity in prostate radiotherapy,” *Physics and Imaging in Radiation Oncology*, 2020.
- [25] O. Casares-Magaz, S. Bülow, N. J. Pettersson, V. Moiseenko, J. Pedersen, M. Thor, J. Einck, A. Hopper, R. Knopp, and L. P. Muren, “High accumulated doses to the inferior rectum are associated with late gastro-intestinal toxicity in a case-control study of prostate cancer patients treated with radiotherapy,” 2019.
- [26] H. A. Gay, H. J. Barthold, E. O’Meara, W. R. Bosch, I. El Naqa, R. Al-Lozi, S. A. Rosenthal, C. Lawton, W. R. Lee, H. Sandler, A. Zietman, R. Myerson, L. A. Dawson, C. Willett, L. A. Kachnic, A. Jhingran, L. Portelance, J. Ryu, W. Small, D. Gaffney, A. N. Viswanathan, and J. M. Michalski, “Pelvic Normal Tissue Contouring Guidelines for Radiation Therapy: A Radiation Therapy Oncology Group Consensus Panel Atlas,” *International Journal of Radiation Oncology*Biophysics*Physics*, vol. 83, pp. e353–e362, jul 2012.
- [27] O. Barbosa Neto, L. Souhami, and S. Faria, “Hypofractionated radiation therapy for prostate cancer: The McGill University Health Center experience,” *Cancer/Radiotherapie*, vol. 19, no. 6-7, pp. 431–436, 2015.
- [28] NRG Oncology, “Stereotactic Body Radiation Therapy or Intensity-Modulated Radiation Therapy in Treating Patients With Stage IIA-B Prostate Cancer (NRG-GU005),” 2017.

- [29] H. M. Patrick and J. Kildea, "Technical note: rtdsm-An open-source software for radiotherapy dose-surface map generation and analysis.," *Medical physics*, vol. 49, pp. 7327–7335, nov 2022.
- [30] C. Chen, M. Witte, W. Heemsbergen, and M. V. Herk, "Multiple comparisons permutation test for image based data mining in radiotherapy," *Radiation Oncology*, 2013.
- [31] L. E. A. Shelley, *Accumulating delivered dose to the rectum to improve toxicity prediction in prostate radiotherapy*. PhD thesis, Cambridge UK, 2020.
- [32] J. A. Hatton, P. B. Greer, C. Tang, P. Wright, A. Capp, S. Gupta, J. Parker, C. Wratten, and J. W. Denham, "Does the planning dose–volume histogram represent treatment doses in image-guided prostate radiation therapy? Assessment with cone-beam computerised tomography scans," *Radiotherapy and Oncology*, vol. 98, pp. 162–168, feb 2011.
- [33] L. Chen, K. Paskalev, X. Xu, J. Zhu, L. Wang, R. A. Price, W. Hu, S. J. Feigenberg, E. M. Horwitz, A. Pollack, and C. M. Ma, "Rectal dose variation during the course of image-guided radiation therapy of prostate cancer," *Radiotherapy and Oncology*, 2010.
- [34] Y. Akino, Y. Yoshioka, S. Fukuda, S. Maruoka, Y. Takahashi, M. Yagi, H. Mizuno, F. Isohashi, and K. Ogawa, "Estimation of rectal dose using daily megavoltage cone-beam computed tomography and deformable image registration," *International Journal of Radiation Oncology Biology Physics*, 2013.
- [35] J. B. Yu, L. D. Cramer, J. Herrin, P. R. Soulos, A. L. Potosky, and C. P. Gross, "Stereotactic body radiation therapy versus intensity-modulated radiation therapy for prostate cancer: Comparison of toxicity," *Journal of Clinical Oncology*, 2014.
- [36] P. A. Kupelian, K. M. Langen, O. A. Zeidan, S. L. Meeks, T. R. Willoughby, T. H. Wagner, S. Jeswani, K. J. Ruchala, J. Haimerl, and G. H. Olivera, "Daily variations in delivered doses in patients treated with radiotherapy for localized prostate cancer," *International Journal of Radiation Oncology Biology Physics*, 2006.

- [37] J. M. Pawlowski, E. S. Yang, A. W. Malcolm, C. W. Coffey, and G. X. Ding, “Reduction of Dose Delivered to Organs at Risk in Prostate Cancer Patients via Image-Guided Radiation Therapy,” *International Journal of Radiation Oncology Biology Physics*, 2010.
- [38] J. V. Lebesque, A. M. Bruce, A. P. Guus Kroes, A. Touw, T. Shouman, and M. van Herk, “Variation in volumes, dose-volume histograms, and estimated normal tissue complication probabilities of rectum and bladder during conformal radiotherapy of T3 prostate cancer,” *International Journal of Radiation Oncology, Biology, Physics*, vol. 33, pp. 1109–1119, dec 1995.
- [39] D. M. De Muinck Keizer, C. Kontaxis, L. G. Kerkmeijer, J. R. Van Der Voort Van Zyp, C. A. Van Den Berg, B. W. Raaymakers, J. J. Lagendijk, and J. C. De Boer, “Dosimetric impact of soft-tissue based intrafraction motion from 3D cine-MR in prostate SBRT,” *Physics in Medicine and Biology*, 2020.

6.9 Supplemental Materials

Supplementary Material

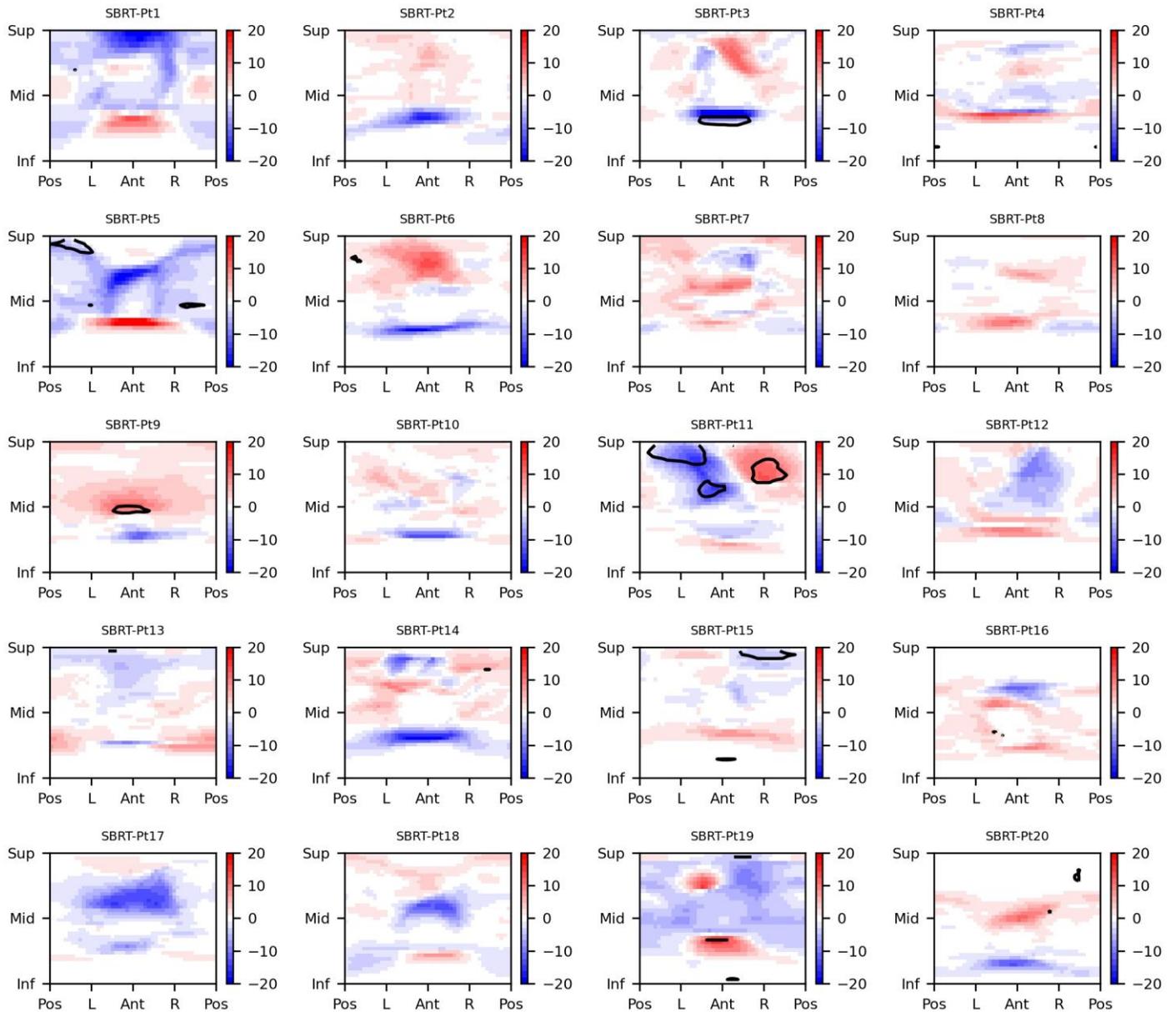


Figure S1: Dose difference maps (DDMS) of the SBRT treatment patients, in units of Gy. Subregions with statistically significant dose differences are contoured in black. Only subregions consisting of five or more continuous pixels were considered as statistically significant subregions for the purposes of this paper.

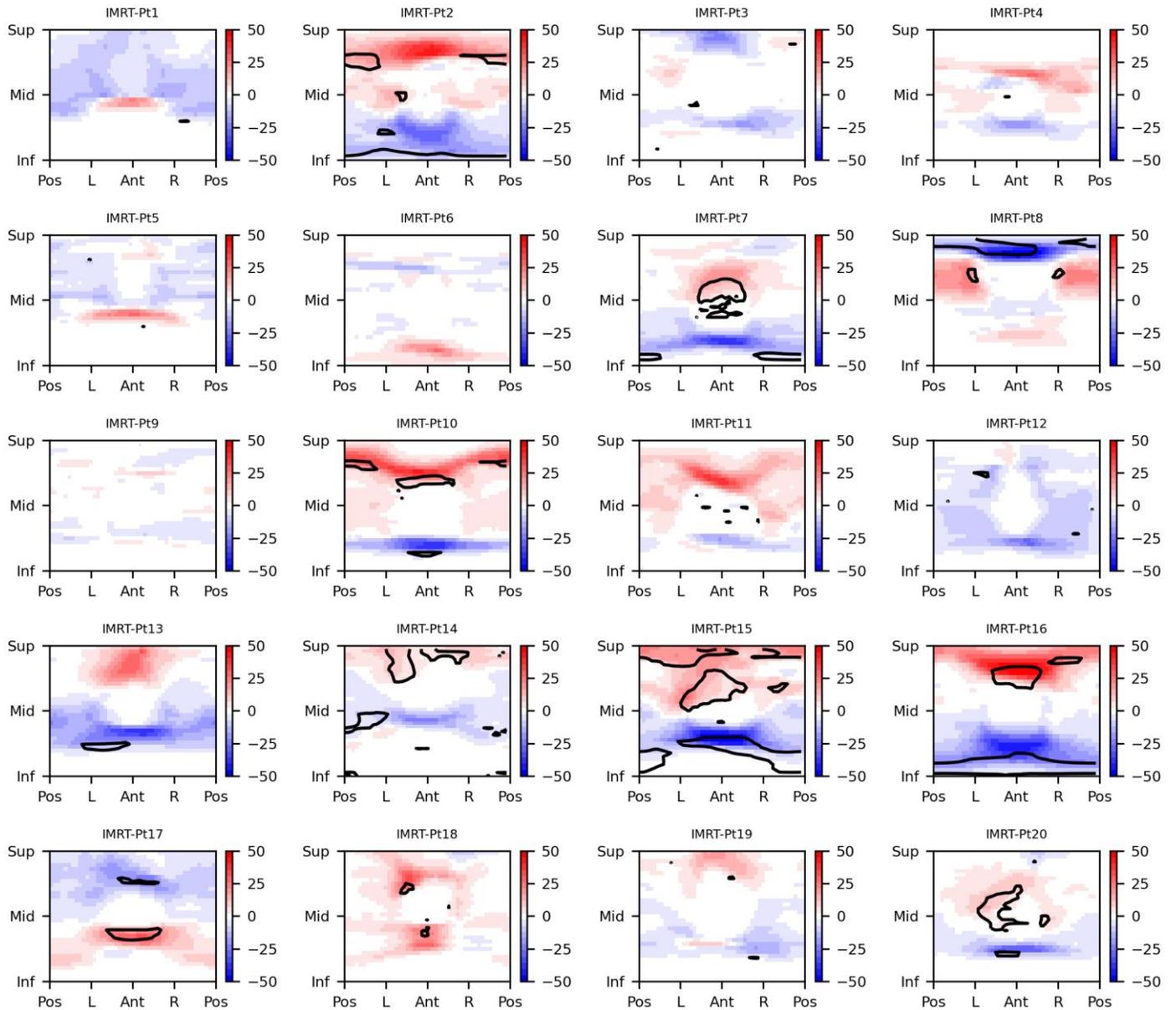


Figure S2: Dose difference maps (DDMS) of the IMRT treatment patients, in units of Gy. Subregions with statistically significant dose differences are contoured in black. Only subregions consisting of five or more continuous pixels were considered as statistically significant subregions for the purposes of this paper.

Chapter 7

Summary and Outlook

7.1 Summary

This thesis describes the development and validation of a methodology to accumulate rectal doses using DSMs. Although radiotherapy for prostate cancer has continued to become more conformal and complex, our understanding of the dose-toxicity relationship for the rectum remains limited, allowing rectal toxicity occurrences to persist and to continue to diminish patients' quality of life. Due to this ongoing issue, there is a growing need in dose-outcome research for a form of dosimetric data that better captures the radiation doses patients experience in a spatially-inclusive manner.

The creation of DSMs to visualize dose to the rectal wall was first reported near the turn of the millennium [1] based on the then newly emerging field of CT colonography [2, 3]. While usage slowly increased throughout the 2000s, the concept of DSM-based dose accumulation was not proposed until 2014 [4]. DSM accumulation was founded on the (then untested) assumption that daily rectal DSMs can be aligned and summed to yield a DSM of the total delivered rectal dose. This thesis built the necessary software (Chapter 3) to perform a comprehensive investigation of the validity of and preferable methodology for DSM accumulation (Chapters 4-5) and demonstrated a use case of the technique (Chapter 6).

7.1.1 Objective 1: Development of software for DSM calculation

While many research groups have developed code bases for the calculation of DSMs, they remain largely private and restricted to each group's preferred calculation approach. For this reason, the development of our own DSM calculation software (*rtdsm*) was required before investigations of DSM-based dose accumulation could begin. We designed *rtdsm* as a Python package with modular functions that utilizes information from RT-DICOM files to construct DSMs. Multiple interacting functions were developed to replicate popular calculation strategies from the literature for each step of DSM construction (e.g. planar and non-coplanar slicing), making *rtdsm* the first DSM calculation software explicitly designed to implement multiple calculation approaches. Several functions to perform DSM-feature analysis were also included for the benefit of end-users. *rtdsm*'s DSM calculation ability was demonstrated with a small cohort of prostate patient data. On average, rectal DSMs could be calculated in four seconds when using the planar slicing approach and in 8.5 seconds when using the non-coplanar slicing approach, making *rtdsm* a viable platform to calculate DSMs for large cohorts.

This work primarily served as the foundation to enable the development of a DSM-based dose accumulation workflow. However, the open-source release of *rtdsm* on GitHub with full documentation and tutorials also represents the first intentional public release of DSM calculation software, with the intention to encourage a more accessible and open-science atmosphere in the DSM research community.

7.1.2 Objective 2: An evaluation of DSM stability with calculation approach

Following the completion of objective 1, we next needed to identify the most promising DSM calculation method for use in rectal DSM accumulation. However, during a literature review to identify possible candidates we observed discordance in toxicity predictive metrics derived from differently-calculated DSMs, causing us to question if variable DSM calculation strategies were creating a reproducibility issue in the field. Therefore, we decided to evaluate the stability of DSM appearance with calculation approach and the subsequent impact of

DSM calculation approach on analysis for popular prostate radiotherapy DSM structures. We undertook this evaluation with the goal of characterizing the stability and identify preferable calculation methods for each approach.

In chapter 4, we calculated rectum and bladder DSMs using several different approaches. DSMs were directly compared to one another, as were their features to assess their equivalence. A cohort comparison was also performed using each style of DSM to evaluate the stability of analysis results with calculation approach. We found that rectal DSMs calculated with planar and non-coplanar slice orientation approaches were not equivalent to one another, and that they tended to identify different statistically dissimilar subregions when used for cohort comparison. Additionally, we observed DSM features to be highly sensitive to both slice orientation approach and DSM resolution, which could subsequently impact analysis results as well. This work demonstrated and confirmed our suspicion that choice of DSM calculation approach influences the results of DSM-based studies. These findings represent an important lesson for the DSM research community and highlight the need for standardization in order to improve reproducibility in the field. In service of this, the final section of chapter 4 was dedicated to recommending possible best calculation practices for the rectum and bladder and measures to improve reproducibility going forwards. Our recommended rectum calculation approach – non-coplanar slicing – was adopted as our method of choice for our subsequent DSM studies.

7.1.3 Objective 3: Experimental validation of DSM accumulation

The work presented in chapter 5 aimed to quantify the accuracy with which DSM accumulation can capture the dose delivered to a rectum structure over the course of a multi-fraction treatment. Although DSM accumulation has been previously performed and theorized to work [4–7], this study represented the first attempt to validate it experimentally.

A rectum phantom was constructed to recreate common types of rectum inter-fraction motion and delivered dose quantified using radiochromic film. Accumulated DSMs were calculated using our recommended (non-coplanar) and non-recommended (planar) DSM calculation approaches, as well as from an accumulated dose distributions calculated with

DIR. Gamma analysis found DSMs to be a good representation of rectal surface dose in a stationary scenario, as well as in the presence of both inter-fraction filling changes and translations/rotations. Film-DSM agreement was less when using planar DSMs compared to non-coplanar DSMs in situations where the angling of the rectum changed. This is noteworthy as most DSM accumulation studies to date have used planar DSMs. It was also found that DSM accumulation outperformed DIR-based dose accumulation for the specific commercial intensity-based algorithm tested in this study. This suggested that DSM accumulation may be a viable and possibly preferable means of rectal dose accumulation in cases where only intensity-based or no DIR software is available. Further work is, however, required to test the performance of DSM accumulation for more complex inter-fraction motion scenarios and for use with other organs.

7.1.4 Objective 4: Practical demonstration of DSM accumulation

As a final objective, we aimed to demonstrate the use of DSM accumulation with real-world clinical data. Inspired by our clinic's shift to prostate SBRT treatments during the COVID-19 pandemic that occurred partway through this project, we decided to compare the agreement between planned and delivered rectum doses for our previous (IMRT) and new (SBRT) standard treatments. Planned and accumulated DSMs were calculated for 20 patients from each treatment arm using the methodology developed in objective 3. In addition, we quantified global and localized changes in rectal shape over the course of treatment to determine if dosimetric differences identified with DSMs could be related to these changes.

Planned and delivered rectal doses were compared using DSMs on a patient-by-patient and cohort level. The SBRT treatment was found to have a smaller proportion of patients with statistically different planned and delivered rectum doses compared to the IMRT treatment. These differences were also limited to smaller areas of the rectal wall. On a cohort level, systematically lower posterior rectal wall doses were observed for the IMRT cohort. No statistically significant dosimetric differences were observed for the SBRT cohort, suggesting that patients on the SBRT treatment may be treated more accurately to plan than the IMRT cohort. On further analysis, the dosimetric differences observed for

the IMRT cohort were found to be well correlated with changes in rectum shape in the region where the significant dose differences occurred. Minimal changes in rectal shape over the course of treatment could also explain the SBRT cohort results. This study marks the first investigation of dose delivery accuracy of hypofractionated and SBRT prostate radiotherapy treatments with DSMs and demonstrates that spatial differences between planned and delivered doses are attributable to localized changes in rectal shape. While these results are promising, future work is required to determine whether or not a systematic factor led to the variations seen in the IMRT cohort and repeat the experiment with a larger sample size.

7.2 Future Directions

The work encompassed by this thesis provides a detailed investigation of the current state of DSM research and demonstrates several potential avenues for new applications and improvements going forwards. While the potential for DSMs is vast, we present the following areas of future research for particular consideration as they relate directly to the present work.

7.2.1 Further investigations of DSM-based dose accumulation

The work presented in chapter 5 indicates that DSM-based dose accumulation can yield accurate representations of total delivered dose to the surface of a tubular organ. As DIR-based dose accumulation for the rectum is complicated by material density changes, DSM-based accumulation may provide an alternative approach for cases where traditional DIR dose accumulation struggles. DSM accumulation may also be of interest to centres with fewer resources interested in dose accumulation that are unable to purchase clinical DIR software featuring current cutting-edge algorithms.

While all deformation scenarios shown in this work had good agreement between measured and accumulated rectum surface dose, it is important to note that they do not represent all possible motion and deformation scenarios. It is possible that certain untested scenarios may yield lower DSM accumulation accuracy. Future studies should be performed

to investigate DSM accumulation accuracy in the presence of more complex forms of inter-fraction rectal changes. One particular area to focus on in more detail is scenarios of rectal expansion and contraction that utilize more elastic phantoms. The calculation of DSMs often uses equiangular sampling, under the assumption that the rectum expands and contracts isotropically, which may or may not hold in reality. A possible avenue for future research could be testing the validity of this assumption using phantoms constructed of rubber or *ex vivo* tissue that replicate the properties of human bowel tissue. Such investigations could make use of point dosimeters such as TLDs or scintillators embedded in the phantom material to evaluate DSM accumulation, as well as other methods of dose accumulation as appropriate.

In addition to investigating DSM accumulation for more complex deformation scenarios of the rectum, it may also be of interest to investigate the use of DSM accumulation for other organs. Dose accumulation for the trachea and esophagus may be particularly promising as they are likely to present fewer complex deformations over the course of treatment and may be more straightforward to investigate. DSM accumulation for the vagina could also be explored in some capacity, as DSM addition of cervical EBRT and brachytherapy doses may be more accurate than DIR-based approaches that often struggle with the presence and removal of brachytherapy applicators [8].

7.2.2 Further clinical studies

In chapter 6 we identified systematic changes in rectal shape over the course of treatment for patients in our IMRT cohort that resulted in significantly different rectal doses from what was planned, but not in our SBRT cohort. Although we speculated that this may have been the result of more stringent application of IGRT and set-up protocols in the SBRT group, the underlying cause has yet to be determined. Identification of contributing factors will be required should clinicians wish to mitigate their effects in future practice. Future work in this area should also involve determining whether or not these findings were cohort-specific or generalizable to larger patient populations. Collaborations with clinician partners to identify and test for potentially contributing clinical factors are also a must. For example, investigations into whether or not contouring uncertainties and treatment planning

criteria influenced the results of our study are required to rule out these factors. Further follow-up studies with larger cohort sample sizes would also allow for the identification of the pervasiveness of these patterns in broader clinical practice. Only then can we determine what, if any, adjustments should be made to existing treatment delivery protocols going forwards to improve dose delivery accuracy.

In addition to further investigations into the causes of variations between planned and accumulated DSMs, we would also recommend that future DSM studies investigate the dose-outcome relationships for accumulated SBRT and hypofractionated treatments. Currently, both planned and accumulated DSM dose-outcome studies are limited to conventionally-fractionated treatments, providing an easy opportunity for discovery in SBRT studies. Accumulated DSM dose-outcome studies of SBRT treatments could allow for the identification of toxicity-predictive spatial dose patterns, better informing design of SBRT treatments. This in turn could increase clinician confidence that toxicities can be avoided during SBRT and contribute to increased prostate SBRT adoption. In an effort to facilitate such research, we have already been actively recruiting prostate SBRT patients and collecting patient-reported outcomes for a future dose-outcome study. Recruitment is expected to be completed sometime next year, at which point a future student will conduct analysis.

7.2.3 Improvements to DSM calculation software

During the course of the thesis, a code package for DSM calculation, *rtdsm*, was created. While the initial release of the package outlined in chapter 3 provided resources for key DSM calculation and analysis operations, there are still many avenues in which open-source DSM research software can be advanced. For example, in the discussion at the end of chapter 3, we briefly mentioned that the inclusion of other DSM analysis techniques into *rtdsm* beyond feature calculation should be prioritized. We are pleased to say that since that publication we have successfully developed an implementation of the popular multiple comparisons permutation test first described by Chen *et al* [9] that will be added to *rtdsm* this spring. In addition to the original implementation, we have also taken steps with the input of a statistician to create a paired version of the test (chapter 4 supplement), as well

as including the option to perform one- and two-sided versions of the test to expand the possible hypotheses that can be tested with it.

Long term, future developments should focus on improving surface mesh-generating and non-coplanar sampling methods. Currently, the method used to generate 3D surfaces in *rtdsm* consists of creating voxelated 3D objects and applying a smoothing filter to obtain an appropriate surface mesh. This solution was used as available Python implementations of the (otherwise preferred) Delaunay triangulation method to create surface meshes struggled to create continuous rectum meshes. Exploring other surface mesh generating methods or developing custom implementations of existing ones could serve to improve future structure representations. Development of alternative central-axis path computation methods [3, 10] could also stand to benefit the non-coplanar slicing method by better avoiding the creation of overlapping slices, reducing quality check and correction time and increasing overall computation speed.

Another key area for future investigation would be the development of a finite element modelling (FEM) based calculation approach. Unlike the typical methods of DSM calculation, FEM-based techniques explicitly consider the material properties of the organ of interest and deform a cylindrical mesh of sampling points to a contour. Such an approach may be better suited for dose accumulation purposes and yield better representations of total delivered dose in scenarios where isotropic rectal expansion may not exist, as touched on earlier. While an existing FEM software package has been used for DSM accumulation on one occasion [11], the development of a free, open-source version using existing FEM Python libraries could greatly increase the number of organs and ways in which DSMs could be studied.

7.2.4 Broader clinical adoption of DSMs

Currently, DSMs are primarily used in a research context to evaluate the dose-outcome relationships between organ dose and toxicity, with the aim of identifying predictive spatial dose patterns. However, the spatial dose visualization that DSMs offer has the potential to benefit both clinicians and patients in numerous ways by allowing for a straightforward review of spatial dose distributions.

One of the immediate possible clinical applications of DSMs and DSM-based dose accumulation is dose monitoring throughout the course of treatment. In a workflow similar to offline adaptive radiotherapy, clinicians could monitor total dose to specific (hollow) organs at risk over the course of treatment by contouring and calculating dose on daily images and accumulating their DSMs. These accumulated DSMs could be used to identify early indications that OAR total delivered doses may exceed tolerance limits by the end of treatment and allow time for treatment adjustments to be made. The workflow of this process could be largely automated by integrating automated contouring and planning tools into it, and the limiting of dose accumulation to a single or handful of OARs could be expected to further decrease computation time relative to full adaptive radiotherapy protocols. Based on the work of this thesis, rectum and esophageal structures would be promising first candidates to test this application, as they both exhibit frequent radiation toxicities and are structures that DSM accumulation should be valid for.

Alternatively, the same retrospective dose accumulation process could also be applied to investigate delivered organ dose in patients who experience acute or late toxicities. Oftentimes clinical evaluations of radiotherapy morbidities are limited to revisiting treatment plans and consulting IGRT images and clinical notes for possible explanations. By also including planned and accumulated DSMs in this review process, clinicians would be able to evaluate the intended and actual dose to the OAR and better identify potential reasons for the morbidity. This could be used on an individual patient basis, or as a larger retrospective practice review to guide adjustments in local practice.

A more distant clinical application of DSMs would be their integration and use in treatment planning systems for plan optimization and review. This application is relatively obvious based on the primary research application of DSMs and the main objective of DSM dose-outcome studies. However, several issues need to be addressed in the DSM field before this application can be considered feasible. As demonstrated in chapter 4, the current landscape of DSM research is highly unstandardized and subject to limited reproducibility, limiting the accrual of evidence to support specific dose-outcome responses. The establishment of DSM-based dosimetric constraints are therefore conditional on the adoption of standard DSM calculation methods for each OAR in order to identify

reproducible responses and ensure they will be applicable to the DSMs used clinically. Development of treatment optimization algorithms using DSMs instead of DVHs will also be required to demonstrate proof-of-concept for treatment planning usage and increase the likelihood of clinician and industry buy-in.

Bibliography

- [1] G. J. Meijer, M. Van Den Brink, M. S. Hoogeman, J. Meinders, and J. V. Lebesque, “Dose-wall histograms and normalized dose-surface histograms for the rectum: A new method to analyze the dose distribution over the rectum in conformal radiotherapy,” *International Journal of Radiation Oncology Biology Physics*, 1999.
- [2] A. K. Hara, C. D. Johnson, J. E. Reed, D. A. Ahlquist, H. Nelson, R. L. MacCarty, W. S. Harmsen, and D. M. Ilstrup, “Detection of colorectal polyps with CT colography: Initial assessment of sensitivity and specificity,” *Radiology*, 1997.
- [3] G. Wang, E. G. McFarland, B. P. Brown, and M. W. Vannier, “GI tract unraveling with curved cross sections,” *IEEE Transactions on Medical Imaging*, 1998.
- [4] J. Murray, D. McQuaid, A. Dunlop, F. Buettner, S. Nill, E. Hall, D. Dearnaley, and S. Gulliford, “SU-E-J-14: A Novel Approach to Evaluate the Dosimetric Effect of Rectal Variation During Image Guided Prostate Radiotherapy,” *Medical Physics*, 2014.
- [5] L. E. Shelley, J. E. Scaife, M. Romanchikova, K. Harrison, J. R. Forman, A. M. Bates, D. J. Noble, R. Jena, M. A. Parker, M. P. Sutcliffe, S. J. Thomas, and N. G. Burnet, “Delivered dose can be a better predictor of rectal toxicity than planned dose in prostate radiotherapy,” *Radiotherapy and Oncology*, 2017.
- [6] O. Casares-Magaz, S. Bülow, N. J. Pettersson, V. Moiseenko, J. Pedersen, M. Thor, J. Einck, A. Hopper, R. Knopp, and L. P. Muren, “High accumulated doses to the inferior rectum are associated with late gastro-intestinal toxicity in a case-control study of prostate cancer patients treated with radiotherapy,” 2019.

-
- [7] O. D'Aquino, *Assessment of delivered rectal dose in prostate cancer radiotherapy*. PhD thesis, Sutton UK, 2021.
- [8] J. Swamidas, C. Kirisits, M. De Brabandere, T. P. Hellebust, F. A. Siebert, and K. Tanderup, "Image registration, contour propagation and dose accumulation of external beam and brachytherapy in gynecological radiotherapy," 2020.
- [9] C. Chen, M. Witte, W. Heemsbergen, and M. V. Herk, "Multiple comparisons permutation test for image based data mining in radiotherapy," *Radiation Oncology*, 2013.
- [10] A. Witztum, B. George, S. Warren, M. Partridge, and M. A. Hawkins, "Unwrapping 3D complex hollow organs for spatial dose surface analysis," *Medical Physics*, 2016.
- [11] L. E. Shelley, M. P. Sutcliffe, S. J. Thomas, D. J. Noble, M. Romanchikova, K. Harrison, A. M. Bates, N. G. Burnet, and R. Jena, "Associations between voxel-level accumulated dose and rectal toxicity in prostate radiotherapy," *Physics and Imaging in Radiation Oncology*, 2020.