7-28-2022 2:00 PM

Software and Hardware-based Tools for Improving Ultrasound Guided Prostate Brachytherapy

Nathan Orlando, *The University of Western Ontario*

Supervisor: Fenster, Aaron, *The University of Western Ontario*
Co-Supervisor: Hoover, Douglas A., *London Health Sciences Centre and Western University*
A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Medical Biophysics

Abstract

Minimally invasive procedures for prostate cancer diagnosis and treatment, including biopsy and brachytherapy, rely on medical imaging such as two-dimensional (2D) and three-dimensional (3D) transrectal ultrasound (TRUS) and magnetic resonance imaging (MRI) for critical tasks such as target definition and diagnosis, treatment guidance, and treatment planning. Use of these imaging modalities introduces challenges including time-consuming manual prostate segmentation, poor needle tip visualization, and variable MR-US cognitive fusion. The objective of this thesis was to develop, validate, and implement software- and hardware-based tools specifically designed for minimally invasive prostate cancer procedures to overcome these challenges.

First, a deep learning-based automatic 3D TRUS prostate segmentation algorithm was developed and evaluated using a diverse dataset of clinical images acquired during prostate biopsy and brachytherapy procedures. The algorithm significantly outperformed state-of-the-art fully 3D CNNs trained using the same dataset while a segmentation time of 0.62 s demonstrated a significant reduction compared to manual segmentation. Next, the impact of dataset size, image quality, and image type on segmentation performance using this algorithm was examined. Using smaller training datasets, segmentation accuracy was shown to plateau with as little as 1000 training images, supporting the use of deep learning approaches even when data is scarce. The development of an image quality grading scale specific to 3D TRUS images will allow for easier comparison between algorithms trained using different datasets. Third, a power Doppler (PD) US-based needle tip localization method was developed and validated in both phantom and clinical cases, demonstrating reduced tip error and variation for obstructed needles compared to conventional US. Finally, a surface-based MRI-3D TRUS deformable image registration algorithm was developed and implemented clinically, demonstrating improved registration accuracy compared to manual rigid registration and reduced variation compared to the current clinical standard of physician cognitive fusion. These generalizable and easy-to-implement tools have the potential to improve workflow efficiency and accuracy for minimally invasive prostate procedures.

Keywords: prostate cancer, 3D ultrasound, magnetic resonance imaging, prostate brachytherapy, prostate biopsy, deep learning, image segmentation, Doppler ultrasound, image registration

Summary for Lay Audience

Prostate cancer is the most common non-skin cancer in Canadian men. The management of prostate cancer often includes the use of needles, such as in diagnosis to extract tissue samples and test for the presence of cancer, known as biopsy, and in treatment to deliver radiation from inside the body, known as brachytherapy. These procedures rely on medical imaging for critical tasks including target definition, creating patient-specific treatment plans, and accurately guiding needles into the body. This also introduces challenges as time-consuming and difficult manual tasks must be completed in the operating room such as accurately identifying the prostate and needle tip locations and mentally combining information from multiple imaging types. This thesis is focused on developing innovative software and hardware solutions to overcome these challenges and improve treatment efficiency and accuracy.

First, artificial intelligence was used to train an algorithm to locate the prostate boundary (or 'segment') in three-dimensional ultrasound images in under 1 second, demonstrating equal segmentation accuracy and greatly reducing time compared to manual segmentation, which can take up to 10 minutes. Next, this same algorithm was tested using smaller datasets, demonstrating equivalent performance with as little as 7% of the full dataset, potentially increasing access to artificial intelligence methods even if data is scarce. Third, an alternative ultrasound technique known as power Doppler ultrasound was used to improve needle tip visibility during the live brachytherapy procedure, demonstrating reduced variability compared to standard ultrasound. Finally, an automated image registration algorithm was developed to overlay magnetic resonance images on ultrasound images, facilitating the targeting of the previously invisible tumour in the operating room.

Co-authorship Statement

This thesis is an integration of four original research articles which are either published in peer-reviewed scientific journals or are submitted and under peer-review. As the first author of each of these manuscripts, I had a significant role in all aspects of the work including the proposal of research questions, experimental design, data acquisition, data analysis, statistics, manuscript drafting, and manuscript revision. All authors contributed editorial feedback prior to manuscript submission. The specific contributions of all co-authors for each manuscript are described below.

Chapter 2 is an original research article entitled "Automatic prostate segmentation using deep learning on clinically diverse 3D transrectal ultrasound images" which was published in *Medical Physics* in 2020. This manuscript was co-first authored by Nathan Orlando and Derek J. Gillies with co-authorship from Igor Gyacskov, Cesare Romagnoli, David D'Souza, and Aaron Fenster. Derek Gillies and I equally split contributions for experimental design, data collection and curation, development of image analysis software, 3D CNN implementation (model training, optimization, validation, and testing), data analysis, interpretation, and statistics, and manuscript drafting, editing, and revising. Igor Gyacskov completed all manual prostate segmentations and was responsible for the 2D U-Net implementation (model training, optimization, validation, and testing) and 3D reconstruction algorithm development. Cesare Romagnoli (now deceased) and David D'Souza were the Radiologist and Radiation Oncologist, respectively, who assisted with clinical 3D TRUS image acquisition. Aaron Fenster contributed to the conception and design of the study and supervised the data interpretation and manuscript drafting and revision.

Chapter 3 is an original research article entitled "Effect of dataset size, image quality, and image type on deep learning-based automatic prostate segmentation in 3D ultrasound" which was published in *Physics in Medicine & Biology* in 2022. This manuscript was co-authored by Nathan Orlando, Igor Gyacskov, Derek J. Gillies, Fumin Guo, Cesare Romagnoli, David D'Souza, Derek W. Cool, Douglas A. Hoover, and Aaron Fenster. My contributions included experimental design, dataset collection and curation, image and data

analysis, data interpretation and statistics, and manuscript drafting and revision, in addition to assisting with the U-Net++ implementation. Igor Gyacskov was responsible for 2D neural network implementation and 3D reconstruction algorithm development. Derek J. Gillies assisted in the development of segmentation evaluation software. Fumin Guo provided the U-Net++ python code and assisted with implementation. Cesare Romagnoli (now deceased) and David D'Souza were the Radiologist and Radiation Oncologist, respectively, who assisted with clinical 3D TRUS image acquisition. Derek W. Cool was the interventional radiologist who developed the 3D TRUS image quality grading scale and subsequently graded the testing dataset images. Douglas A. Hoover and Aaron Fenster contributed to the conception and design of the study and supervised the data interpretation and manuscript drafting and revision.

Chapter 4 is an original research article entitled "A power Doppler ultrasound method for improving intraoperative needle tip localization in interstitial prostate brachytherapy", which was submitted for publication in *Medical Physics* on March 28, 2022 and is currently under peer-review. This manuscript was co-authored by Nathan Orlando, Jonatan Snir, Kevin Barker, David D'Souza, Vikram Velker, Lucas C. Mendez, Aaron Fenster, and Douglas A. Hoover. My specific responsibilities included contributing to the initial conception of the power Doppler (PD) US method and oscillator design in addition to experimental design, phantom construction and image acquisition, computed tomography-US rigid registration and evaluation, oscillator operation and needle tip localization in phantom and clinical brachytherapy procedures, data analysis, interpretation, and statistics, and manuscript drafting. Jonatan Snir contributed to the initial conception of the PD US method and oscillator design. Kevin Barker was responsible for the wireless oscillator design and manufacture. David D'Souza, Vikram Velker, and Lucas C. Mendez were physicians who operated the mechanical oscillator as part of the prospective feasibility clinical trial in addition to grading needle tip visibility. Aaron Fenster and Douglas A. Hoover contributed to the conception and design of the study and supervised the data interpretation and manuscript drafting. In addition, Douglas A. Hoover assisted with phantom experimentation and clinical validation including needle tip localization, clinical software operation, and needle tip visibility grading.

Chapter 5 is an original research article entitled "Validation of a surface-based deformable MRI-3D ultrasound image registration algorithm towards clinical implementation for interstitial prostate brachytherapy", which was submitted for publication in *Brachytherapy* on June 15, 2022. This manuscript is co-authored by Nathan Orlando, Chandima Edirisinghe, Igor Gyacskov, Jason Vickress, Robin Sachdeva, Jose A. Gomez, David D'Souza, Vikram Velker, Lucas C. Mendez, Glenn Bauman, Aaron Fenster, and Douglas A. Hoover. My contributions included phantom image acquisition and retrospective clinical data collection and curation. I was responsible for the validation and operation of the deformable image registration algorithm, including intraoperative implementation, as well as for image analysis to assess registration accuracy, organ contouring in phantom, manual landmark localization, data analysis, interpretation, and statistics, and manuscript drafting. Chandima Edirisinghe was responsible for the registration algorithm development and implementation. Igor Gyacskov assisted with the initial design of the registration algorithm. Jason Vickress assisted with study design registration accuracy evaluation strategy. Robin Sachdeva was a part of the clinical team which facilitated implementation of our algorithm in a prospective clinical trial. Jose A. Gomez was the pathologist who completed histological evaluation of the targeted biopsy tissue sample. David D'Souza, Vikram Velker, and Lucas C. Mendez were the physicians who completed the manual MR-US cognitive fusions as part of the retrospective clinical validation. In addition, Lucas C. Mendez completed all contouring for the retrospective and prospective clinical cases, and was the co-principal investigator along with Glenn Bauman for the prospective clinical trial that used our algorithm was implemented in. Aaron Fenster and Douglas A. Hoover contributed to the conception and design of the study and supervised the data interpretation and manuscript drafting. In addition, Douglas A. Hoover completed manual rigid registrations and assisted with clinical implementation.

Dedication

To Mom and Dad,

for your never-ending support

Acknowledgments

First and foremost, I would like to thank my supervisor, Dr. Aaron Fenster. From the very start you have given me every chance to succeed. It's been a true pleasure working in such a high-energy lab environment filled with brilliant people. Behind all of this, however, is your unrelenting work ethic and passion for science, which have been a huge inspiration for me and have driven me to accomplish more than I ever imagined. It is very motivating to have a supervisor who "walks the talk", as I knew that on any early morning, late night, or weekend when I was working, there was a good chance you were also working (likely writing a grant!). By showing confidence in me and giving me the space to work through problems and plan out my own projects, I feel I have also grown immensely as a person since joining your lab, and I can't thank you enough for that opportunity. You were the reason I came to Western, and I haven't regretted that decision for one second.

To my co-supervisor Dr. Douglas Hoover, I am truly grateful for the dedication of your time and energy to being a fantastic supervisor. While you may have officially signed on midway through my studies, you really were acting as my co-supervisor the whole time. Your integrity, clinical prowess, drive, and keen attention to detail are all things I will strive to emulate as I continue my journey in medical physics. Beyond your supervisory role, I also want to thank you for your willingness to get your hands dirty and assist with my experiments. Some of my fondest memories are the phantom experiments in the OR that carried on late into the evening. I cherished the time I spent at the cancer centre, and I feel it played a key role in increasing both my knowledge and confidence in the clinical setting. This is largely due to your willingness to explain the clinical workflow, even when it went beyond what we were working on. I wouldn't be the same researcher or person if it wasn't for your guidance.

Thank you to my advisory committee, Dr. Aaron Ward and Dr. Eugene Wong. I was lucky to have such an insightful, kind, and dedicated committee. I came away from each meeting with (too many!) new ideas and a clear sense of what areas I needed to review. My thesis would not be the same if it wasn't for your mentorship.

I would also like to thank the clinical team at the London Regional Cancer Program, and in particular the brachytherapy radiation oncology team, Dr. David D'Souza, Dr.

Vikram Velker, and Dr. Lucas Mendez. Your contributions to each chapter of this thesis have been critical to the completion of impactful translational research. The chance to implement some of my developments clinically has been one of the biggest highlights of my time here. Thank you to the patients who consented to our studies, this work would not be possible without your selflessness.

Thank you to my examiners, Dr. Sarah Mattonen, Dr. Donna Murrell, Dr. Jinka Sathya, and Dr. Septimiu Salcudean, for participating in my defense and critically reviewing my thesis.

The work described in this thesis would not be possible without support from the following funding sources: Cancer Research and Technology Transfer strategic training program, the Government of Ontario, the National Sciences and Engineering Research Council, the Canadian Institutes of Health Research, the Ontario Institute for Cancer Research, and the London Health Sciences Foundation.

Thank you to the past and present members of the Fenster lab whose contributions have made this work possible. I have been incredibly lucky to work with such a talented group of people. Thank you to Kevin Barker, Lori Gardi, Jeff Bax, Jacques Montreuil, and many more. To David, I don't think anything would get done in our lab without you, so thank you for your tireless work and kindness. To Chandima and Igor, it's been a pleasure working with you, your expertise and willingness to help is the reason I got to this point.
To my labmates, thank you making my time at Western so special. Derek, you have been an amazing mentor and a true friend. Thank you for the countless laughs, heart-to-heart chats after hockey, and impromptu Toronto concerts. I'm eternally grateful I got to work with you. To Claire, I feel like we have been on this graduate school journey together. Thank you for motivating me, inspiring me, and always pushing me to be my best, even when that means going out of my comfort zone. Jessica, thank you for being such an inspiration and role model. To Sam, thanks for always being a bright spot in my day, I know I can always count on you for a laugh. To Carla and Tiana, although it still feels like you just started, thanks for bringing so much joy into the lab.

I have had the pleasure of making some amazing friends during my time at Western, thank you to Mark, Andrew, Hossein, Leah, Harley, Natasha, and countless others for making this experience so memorable. To Chris, thanks for being a truly amazing friend.

You have always had my back and never hesitate to make the time to chat, whether it's about prostate cancer or anything else. I wouldn't want to go through this PhD with anyone else. To Gabe, I'll forever be grateful we met on our very first day at Robarts, thank you for the countless lunches, tennis matches, and laughs. TD, you are one of the most genuine people I have had the pleasure to meet. Thank you for always inspiring me with your brilliance. To Jordan, thanks for always saying yes to the countless concerts, hockey games, and other random events I invite you to, I'll always cherish those memories and your friendship. To the boys back home, Jordan, Branden, Brandon, Chad, and Austin, thanks for always making time for me. Here's to another decade of late nights, lake weekends, and Mt. Fuji lunches.

Jacob, thanks for being the best brother a guy could ask for, I can always rely on you for anything (unless its earlier than 11am on a Saturday morning). To Grandma Donna and Grandpa Gary, thank you for your endless encouragement and support throughout all of my studies. I'm forever grateful to have you in my corner. To Dave, Tessa, Jenna, Gracie, and Bradley, thanks for accepting me into your family and always supporting me in everything I do.

To my parents, Karen and Martino, I'm not sure what else I can say other than thank you. I absolutely would not be in this position without you. Everything I have accomplished is thanks to the 26 years of unwavering love and support you have given me.

To Nicole, my best friend, fiancée, and everything in between. We started our journey together right as I started graduate school, and I'm not sure how I would have made it without you. You know exactly how to cheer me up when I've had a tough day and how to motivate me when I'm feeling stuck. You're always there when I need support or just a laugh. I am so lucky to have found the most loving, generous, thoughtful, and intelligent partner. I can't wait to begin our next adventure together.

Table of Contents

Abstract i

Summary for Lay Audience iii

Co-authorship Statement iv

Dedication vii

Acknowledgments viii

Table of Contents xi

List of Tables xvii

List of Figures xx

List of Abbreviations xxv

Chapter 1 1

1 Introduction 1

1.1 Prostate Cancer 1

1.2 Prostate Cancer Diagnosis 3

1.2.1 Initial Screening Methods 3

1.2.2 Prostate Biopsy 4

1.2.3 Prostate Cancer Staging 5

1.3 Prostate Cancer Treatment 6

1.3.1 Surgery 7

1.3.2 Systemic Treatments 8

1.3.3 Radiation Therapy 9

1.3.3.1 External Beam Radiation Therapy 10

1.3.3.2 Brachytherapy 12

1.3.4 Focal Therapy 17

1.4 Imaging in Prostate Cancer Diagnosis and Treatment 18
1.4.1 X-ray Computed Tomography (CT) 18
1.4.2 Two-Dimensional Ultrasound (US) 19
1.4.2.1 Doppler Ultrasound 20
1.4.2.2 Three-Dimensional Ultrasound 21
1.4.3 Magnetic Resonance (MR) Imaging 23
1.4.4 Nuclear Imaging 25
1.5 Challenges in Minimally Invasive Prostate Procedures 26
1.5.1 Prostate Segmentation in Three-Dimensional Ultrasound Images 26
1.5.2 Variability in Dataset Size, Quality, and Diversity for Deep Learning Segmentation Applications 28
1.5.3 Intraoperative Needle Tip Localization in Transrectal Ultrasound 30
1.5.4 Deformable Magnetic Resonance Imaging to Three-Dimensional Ultrasound Image Registration 31
1.6 Thesis Objectives and Outline 32
1.6.1 Thesis Outline 34
1.7 References 38
Chapter 2 53
2 Automatic prostate segmentation using deep learning on clinically diverse 3D transrectal ultrasound images 53
2.1 Introduction 53
2.2 Materials and Methods 56
2.2.1 Clinical dataset 56
2.2.2 3D segmentation algorithm 58
2.2.2.1 Training dataset for modified U-Net 58
2.2.2.2 Modified U-Net 58

2.2.2.3 3D reconstruction 59
2.2.3 Evaluation and comparison 60
2.2.4 Statistical analysis 62
2.3 Results 63
2.3.1 Reconstructed modified U-Net 63
2.3.2 3D CNNs and V-Net optimization 65
2.3.3 Comparison of reconstructed modified U-Net and 3D CNNs 66
2.4 Discussion 71
2.4.1 Reconstructed modified U-Net 71
2.4.2 3D CNNs and V-Net Optimization 72
2.4.3 Comparison with 3D V-Net and previously published methods 73
2.4.4 Limitations and future work 76
2.5 Conclusions 77
2.6 References 77
2.7 Supplement A 80
Chapter 3 81
3 Effect of dataset size, image quality, and image type on deep learning-based automatic prostate segmentation in 3D ultrasound 81
3.1 Introduction 81
3.2 Materials and methods 84
3.2.1 Complete ultrasound dataset 84
3.2.2 Reduced-size datasets 85
3.2.2.1 Split end-fire and side-fire datasets 85
3.2.2.2 Smaller end-fire, side-fire, and mixed datasets 86
3.2.3 Image quality assessment 87
3.2.4 3D segmentation algorithm 88

3.2.5 2D neural networks 88
3.2.5.1 Modified U-Net 89
3.2.5.2 U-Net++ 90
3.2.6 Evaluation and comparison 90
3.3 Results 92
3.3.1 Complete dataset 92
3.3.2 Split end-fire and side-fire datasets 94
3.3.3 Smaller end-fire, side-fire, and mixed datasets 96
3.3.4 Image quality 99
3.4 Discussion 101
3.4.1 Complete dataset 101
3.4.2 Split end-fire and side-fire datasets 103
3.4.3 Smaller end-fire, side-fire, and mixed datasets 104
3.4.4 Image quality 105
3.4.5 Limitations and future work 107
3.5 Conclusions 108
3.6 References 108
3.7 Supplement A 112
3.8 Supplement B 113
Chapter 4 116
4 A power Doppler ultrasound method for improving intraoperative needle tip localization in interstitial prostate brachytherapy 116
4.1 Introduction 116
4.2 Materials and Methods 118
4.2.1 Wireless mechanical oscillator design 118
4.2.2 Phantom validation 120

4.2.2.1 Mock HDR-BT procedure 122
4.2.2.2 Explicit needle shadowing 123
4.2.2.3 Evaluation of tip localization accuracy 124
4.2.3 Clinical validation 125
4.2.4 Statistical analysis 126
4.3 Results 127
4.3.1 Phantom validation 127
4.3.2 Clinical validation 131
4.4 Discussion 135
4.4.1 Phantom validation 135
4.4.2 Clinical validation 138
4.5 Conclusions 142
4.6 References 142
Chapter 5 145
5 Validation of a surface-based deformable MRI-3D ultrasound image registration algorithm towards clinical implementation for interstitial prostate brachytherapy ... 145
5.1 Introduction 145
5.2 Methods and materials 147
5.2.1 Deformable surface-based image registration 147
5.2.2 Phantom validation 149
5.2.3 Retrospective clinical validation 150
5.2.4 Evaluation of registration accuracy 151
5.2.5 Clinical implementation 152
5.2.6 Statistical analysis 153
5.3 Results 154
5.3.1 Phantom validation 154

5.3.2 Retrospective clinical validation 157

5.3.3 Clinical implementation 159

5.4 Discussion 161

5.5 Conclusions 163

5.6 References 163

Chapter 6 167

6 Conclusions, Limitations, and Future Work 167

6.1 Summary and Conclusions 167

6.2 Limitations 170

6.2.1 General Limitations 170

6.2.2 Study-Specific Limitations 171

6.3 Future Work 175

6.3.1 Multi-Institution and Multi-Physician Validation of Deep Learning Prostate Segmentation Towards Clinical Implementation 175

6.3.2 Automatic Prostate Segmentation in Post-Implant 3D TRUS Images 176

6.3.3 Power Doppler Needle Localization Extended to other Minimally Invasive Procedures 177

6.3.4 Deep Learning-Based Automatic MRI-3D US Image Registration Algorithm 178

6.4 References 179

APPENDIX A – Copyright Releases 182

APPENDIX B – Research Ethics and Health Canada Approvals 189

Curriculum Vitae 195

List of Tables

Table 2.1. Clinical 3D TRUS dataset split based on end-fire and side-fire scan geometries and resulting training, validation, and testing datasets used for deep learning..... 57

Table 2.2. Hyperparameter selection when employing the V-Net in NiftyNet. 62

Table 2.3. Absolute median [Q1, Q3] results comparing 2D radial slice segmentation to 3D reconstructed segmentation on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate. 64

Table 2.4. Signed median [Q1, Q3] results comparing 2D radial slice segmentation to 3D reconstructed segmentation on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate. 64

Table 2.5. Absolute median [Q1, Q3] results comparing a standard 3D V-Net with a Dice similarity coefficient loss function to a Dice similarity plus cross-entropy (DiceXEnt) loss function on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate. 67

Table 2.6. Signed median [Q1, Q3] results comparing a standard 3D V-Net with a Dice loss function to a Dice plus cross-entropy (DiceXEnt) loss function on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate. 67

Table 2.7. Absolute median [Q1, Q3] results comparing a standard 3D V-Net to our proposed reconstructed modified (rm) U-Net on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate. 68

Table 2.8. Signed median [Q1, Q3] results comparing a standard 3D V-Net to our proposed reconstructed modified (rm) U-Net on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate. 68

Table 2.A1. Absolute median [Q1, Q3] results comparing a Dense V-Net (DenseNet) and High-resolution 3D Network (HighRes3dNet) on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate. 80

Table 2.A2. Signed median [Q1, Q3] results comparing a Dense V-Net (DenseNet) and High-resolution 3D Network (HighRes3dNet) on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate. 80

Table 3.1. Breakdown of the clinical 3D TRUS training dataset of 206 volumes resliced into 2D images for training 2D neural networks. Rows from top to bottom show the complete and reduced-size side-fire, end-fire datasets, and mixed datasets. 86

Table 3.2. Image quality grading scale for 3D TRUS images of the prostate. 88

Table 3.3. List of comparisons and corresponding statistical tests. Parametric tests are shown with corresponding nonparametric alternative tests presented in parentheses. 91

Table 3.4. Median [Q1, Q3] 3D results showing the modified U-Net and U-Net++ trained using the complete dataset and tested on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS volumes. 93

Table 3.5. Median [Q1, Q3] 3D results showing the modified U-Net and U-Net++ trained using only end-fire images and tested on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS volumes. 96

Table 3.6. Median [Q1, Q3] 3D results showing the modified U-Net and U-Net++ trained using only side-fire images and tested on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS volumes. 97

Table 3.7. Mean plus/minus standard deviation image quality grading results for side-fire and end-fire 3D TRUS images of the prostate. Mean grades are provided out of five for each individual image quality factor and for the total image quality. 99

Table 3.B1. Median [Q1, Q3] 2D results for the U-Net++ trained using end-fire datasets of varying size, from 2738 (full end-fire set) to 100 images. The networks were evaluated on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate.113

Table 3.B2. Median [Q1, Q3] 2D results for the U-Net++ trained using side-fire datasets of varying size, from 4023 (full end-fire set) to 100 images. The networks were evaluated on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate. 114

Table 3.B3. Median [Q1, Q3] 2D results for the U-Net++ trained using mixed (end-fire and side-fire) datasets of varying size, from 6761 (full dataset) to 100 images. The networks were evaluated on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images. 115

Table 4.1. Mean ± standard deviation absolute and signed tip localization errors based on ideal reference needles for visualization with standard B-mode US, PD US alone, and PD plus B-mode US in three phantom experiments with two needle types. Negative values represent an underprediction relative to the expected tip location. 129

Table 4.2. Mean ± standard deviation absolute and signed tip localization errors compared to the CT gold standard for needle visualization with standard B-mode US, PD US alone, and PD plus B-mode US in three phantom experiments with two needle types. 130

Table 4.3. Mean ± standard deviation absolute tip localization errors comparing needle visualization with standard B-mode US to PD US with perturbation for clinical cases. Tip errors for each individual patient and averaged for each physician are shown in the left and right halves of the table, respectively. Overall tip error is shown in the bottom row..... 133

Table 5.1. Mean ± standard deviation registration error metrics for our deformable image registration (DIR) algorithm and manual rigid registration evaluated using 1.5T and 3T MRI phantom images with three deformation levels. Dice similarity coefficient (DSC) scores are provided for the prostate, urethra, and three spherical landmarks. Target registration error (TRE), in mm, is provided for the three spherical landmark centroids. 156

Table 5.2. Mean ± standard deviation Dice similarity coefficient (DSC) values averaged from three clinical cases registered using our deformable image registration (DIR) algorithm and manual rigid registration. Columns show DSC values for the clinical target volume (CTV) and the left and right neurovascular bundles (NVB) compared to cognitive fusion results for each physician individually, averaged DSC results compared to all physicians, and inter-physician variability, from left to right respectively. 158

List of Figures

Figure 1.1: Example sagittal brightness-mode transrectal ultrasound image showing three inserted needles in a clinical high-dose-rate prostate brachytherapy procedure. On the needle closest to the ultrasound transducer, the dwell positions are shown, which are the locations where the radioactive source will sit for specific amounts of time. 14

Figure 1.2: Example axial brightness-mode transrectal ultrasound image and corresponding radiation treatment plan isodose line overlay for a clinical high-dose-rate prostate brachytherapy procedure. The prostate was contoured and shaded in purple and the urethra is shown as the yellow circle. Isodose line dose levels are provided in the legend. Each implanted needle is shown as a small horizontal line. 15

Figure 1.3: Example (a) axially reconstructed step-back, (b) sagittally reconstructed side-fire, and (c) end-fire three-dimensional transrectal ultrasound images. The dotted lines in each image represent the original ultrasound acquisition planes that compose the 3D image. Arrows represent the direction of transducer motion during 3D acquisition..... 21

Figure 1.4: Example (a) three-dimensional transrectal ultrasound image and corresponding prostate contour in the axial plane and (b) axial T2-weighted 1.5T magnetic resonance image with physician-drawn prostate, urethra, clinical target volume (CTV), and neurovascular bundle (NVB) contours. 25

Figure 1.5: Example transrectal ultrasound images highlighting the variability in image quality and prostate boundary visibility which can make manual prostate segmentation time consuming and difficult. (A) Example high quality image with a clearly defined prostate boundary. (B) Example poor quality image with severe shadowing artifacts that limit the prostate boundary visibility..... 27

Figure 2.1. Mechanical scanning approaches for acquiring 3D TRUS prostate images using end-fire (left) and side-fire (right) TRUS transducers. 2D images are acquired by rotating around the long axis of the transducer at known sample spacings to create 3D TRUS images. Example 3D TRUS images are shown in the bottom row, with the front face demonstrating the reconstructed image plane and the white lines showing representative acquisition planes. 57

Figure 2.2. Proposed 3D prostate segmentation workflow. A 3D TRUS prostate image is used as input, followed by radial sampling to generate 12 2D image planes. Each image plane was used to predict a prostate boundary with a trained modified U-Net prior to reconstruction into the 3D prostate surface. 58

Figure 2.3. (a) Method for acquiring radially sliced 2D TRUS image planes (dotted lines) from previously acquired 3D TRUS images. The axis of rotation (red) is approximately at

the central axis of the prostate and in the anterior-posterior direction. (b) 12 image slices were obtained to perform prostate segmentation. For an end-fire 3D TRUS image, 2D images are about the axis of image acquisition. (c) The majority of image slices obtained from side-fire 3D TRUS images are across many acquisition slices and are subject to increased interpolation, except for one slice, which matched the original acquired 2D TRUS image..... 60

Figure 2.4. End-fire prostate segmentation results comparing manual (red), our proposed reconstructed modified (rm) U-Net (blue), and V-Net (yellow) 3D surfaces. The columns from left to right show the 25th, 50th, and 75th percentile results, respectively, based on DSC metrics. Segmentations in the axial plane, sagittal plane, 45° oblique radial plane, and reconstructed 3D surface error are shown in the respective rows from top to bottom. 69

Figure 2.5. Side-fire prostate segmentation results comparing manual (red), our proposed reconstructed modified (rm) U-Net (blue), and V-Net (yellow) 3D surfaces. The columns from left to right show the 25th, 50th, and 75th percentile results, respectively, based on DSC metrics. Segmentations in the axial plane, sagittal plane, 45° oblique radial plane, and reconstructed 3D surface error are shown in the respective rows from top to bottom. 70

Figure 2.6. A sample end-fire prostate segmentation result comparing manual (red), our proposed algorithm (blue), and V-Net (yellow) 3D surfaces in the presence of a hyperechoic calcification image artifact. 74

Figure 3.1. 3D prostate segmentation workflow using an example end-fire 3D TRUS volume. The input 3D TRUS volume was resliced radially at 15° spacing to generate 12 2D TRUS images with similar size and shape. A trained 2D neural network was used to predict the prostate boundary locations in 2D binary masks, which were used to reconstruct the 3D prostate surface..... 89

Figure 3.2. Example end-fire (top row) and side-fire (bottom row) prostate segmentation results comparing manual (red), modified U-Net (blue), and U-Net++ (yellow) 3D surfaces for median cases based on DSC. The columns from left to right show the prostate surface in the axial plane, sagittal plane, and an oblique radial plane, respectively. 93

Figure 3.3. Example end-fire (top row) and side-fire (bottom row) median DSC prostate segmentation results comparing manual (red), modified U-Net (blue), and U-Net++ (yellow) 3D surfaces for networks trained only using end-fire images. The columns from left to right show the prostate surface in the axial plane, sagittal plane, and an oblique radial plane, respectively..... 94

Figure 3.4. Example side-fire (top row) and end-fire (bottom row) median DSC prostate segmentation results comparing manual (red), modified U-Net (blue), and U-Net++ (yellow) 3D surfaces for networks trained only using side-fire images. The columns from

left to right show the prostate surface in the axial plane, sagittal plane, and an oblique radial plane, respectively..... 95

Figure 3.5. Plot of median DSC for the modified U-Net and U-Net++ trained using only end-fire (left) and only side-fire (right) images and tested on both end-fire and side-fire images shown on the left and right half of each graph, respectively. 95

Figure 3.6. Example results for the U-Net++ trained with varying-size end-fire (top row), side-fire (middle row), and mixed (bottom row) datasets. The full dataset corresponds to 2738 images for the end-fire network, 4023 images for the side-fire network, and 6761 for the mixed network..... 97

Figure 3.7. Median DSC for the U-Net++ trained using varying-sized end-fire (left), side-fire (middle), and mixed (right) datasets, and tested on both end-fire (solid red) and side-fire (dashed line) images. The number of training images are provided along the x-axis. The dotted line denotes the maximum performance achieved by the U-Net++ trained with the full 6761 image dataset. 98

Figure 3.8. Mean DSC as a factor of acquisition quality (left), anatomy artifacts severity (middle), and boundary visibility (right) for the modified U-Net (top row) and U-Net++ (bottom row) trained on the full dataset and tested on end-fire and side-fire images shown on the left and right side of each graph, respectively. 100

Figure 3.9. DSC as a function of total averaged image quality for the modified U-Net (left) and U-Net++ (right) trained on the full dataset and tested on end-fire (red) and side-fire (black) images. A linear regression is provided for each dataset. 101

Figure 3.A1. Network diagram for the modified U-Net. 112

Figure 3.A2. Network diagram for the U-Net++ based on Zhou et al. 2018.....112

Figure 4.1. (A) Photograph of our wireless mechanical oscillator. (B) Front view showing the magnetically attached cylindrical end-pieces and example plastic needle with the corresponding mandrin. Two end-piece sizes were used corresponding to plastic and metal Varian needles. (C) The oscillation frequency is controlled using a 16-position rotary dial at the bottom of the device. Frequency values are printed on the device for ease of use. 120

Figure 4.2. Photograph of our experimental set-up showing needle implant patterns for a simulated HDR-BT procedure with metal needles (A), and a simulated case with extensive needle shadowing with plastic needles (B). The US transducer, needle template, and varying size needle mandrins can be seen in both experiments..... 122

Figure 4.3. Example needle visualization results in phantom, showing (A) a standard B-mode US image, (B) PD US image with perturbation from our mechanical oscillator, and (C) PD US only with B-mode context blinded to the observer. 128

Figure 4.4. Box and whisker plots of absolute needle tip localization errors, in mm, comparing needle visualization with standard B-mode US, PD US alone, and PD plus B-mode US. Plots from left to right show results for the mock HDR-BT procedure with (A) metal needles (N=12) and the explicit needle shadowing implant patterns with (B) metal (N=15) and (C) plastic needles (N=13). Boxes denote interquartile range from the 25th to 75th percentile while whiskers show the minimum and maximum values. Each individual tip error is shown with a bar representing the median. 129

Figure 4.5. Example images from the clinical trial showing needle visualization in patients using standard B-mode US (left) and PD US with perturbation (right). The rows show an example (A) unobstructed needle, (B) partially obstructed needle, and (C) obstructed needle. Arrows point to the needle in question. 132

Figure 4.6. Box and whisker plot of absolute needle tip localization errors comparing needle visualization with standard B-mode US to PD US with perturbation in clinical cases. Tip error for each individual patient and the total tip error are shown from left to right. Boxes denote interquartile range from the 25th to 75th percentile while whiskers show the minimum and maximum values. Each individual tip error is shown with a bar representing the median. 133

Figure 4.7. Scatter plot of PD US absolute tip error as a function of B-mode US absolute tip error for all 63 needles in the feasibility clinical trial. 134

Figure 4.8. Plot of the mean ± standard deviation absolute tip localization errors comparing B-mode US to PD US with perturbation in clinical cases with varying needle visibility. Mean tip error for unobstructed, partially obstructed, and obstructed needles are shown from left to right. 135

Figure 4.9. (A) Example axial B-mode US image of a tissue-mimicking agar phantom with an embedded agar prostate showing 12 implanted needles. (B) Example PD US image after vibration of the needle template, showing visualization of all needles. 141

Figure 5.1. Summary of proposed deformable image registration algorithm workflow. 148

Figure 5.2. Example phantom registration results for three deformation levels ranging from no deformation to large deformation shown in columns from left to right, respectively. The top row shows an example 3T MR image, 3D surface view of the phantom highlighting landmark distribution, and contour legend. The middle and bottom rows show MR-to-US registration results generated using our DIR algorithm and manual rigid registration, respectively. Each deformation level examines a different landmark location. 154

Figure 5.3. Box and whisker plots showing (A) landmark Dice similarity coefficient (DSC) and (B) target registration error (TRE) for 1.5T and 3T MRI images registered to 3D US using our DIR algorithm and manual rigid registration in phantom, averaged across three

deformation levels. Boxes denote the interquartile range while whiskers show the minimum and maximum values. A bar and plus sign represent the median and mean, respectively. Statistically significant differences (p<0.05) are denoted by an asterisk. . 156

Figure 5.4. Example clinical registration results for one retrospective HDR-BT case. The top row shows the physician-contoured MRI and US as well as the contour legend. The bottom row shows the MR-to-US registration using our DIR algorithm, manual rigid registration, and cognitive fusion completed by three physicians, respectively. 157

Figure 5.5. Box and whisker plots showing the Dice similarity coefficient (DSC) score for clinical target volume (CTV) and neurovascular bundle (NVB) contours registered using our DIR algorithm, manual rigid registration, and manual cognitive fusion by the physician, presented as inter-physician variability. Boxes denote the interquartile range while whiskers show the minimum and maximum values. A bar and plus sign represent the median and mean, respectively. 159

Figure 5.6. Example intraoperative clinical registration results in a prospective clinical case. The physician-contoured MR image is shown in the top left pane. The pre- and post-implant US image and registration results generated using our DIR algorithm are shown in the top and bottom row, respectively. The contour legend is provided in the bottom left pane. 160

Figure 5.7. Histology slide demonstrating the presence of cancer for the GTV-targeted biopsy result obtained using our DIR algorithm. 160

List of Abbreviations

^{125}I	Iodine-125
^{192}Ir	Iridium-192
2D	Two-dimensional
3D	Three-dimensional
A/VPD	Area/Volume Percent Difference
ABS	American Brachytherapy Society
ADT	Androgen Deprivation Therapy
ANOVA	Analysis of Variance
ASCO	The American Society of Clinical Oncology
BT	Brachytherapy
CBCT	Cone-beam Computed Tomography
CD	Colour Doppler
CNN	Convolutional Neural Network
CPU	Central Processing Unit
CT	Computed Tomography
CTV	Clinical Target Volume
DCE	Dynamic Contrast-Enhanced
DiceXEnt	Dice Plus Cross-Entropy
DIL	Dominant Intraprostatic Legion
DIR	Deformable Image Registration
DNA	Deoxyribonucleic Acid
DRE	Digital Rectal Examination
DSC	Dice Similarity Coefficient
DWI	Diffusion-weighted Imaging
EBRT	External Beam Radiation Therapy
FLE	Fiducial Localization Error
FSE	Fast Spin Echo
GPU	Graphics Processing Unit
GTV	Gross Tumour Volume
HD	Hausdorff Distance
HDR	High-dose-rate
ICP	Iterative Closest Point
IMRT	Intensity Modulated Radiation Therapy
LDR	Low-dose-rate
MLC	Multileaf Collimator
mpMRI	Multiparametric Magnetic Resonance Imaging
MR	Magnetic Resonance

MRI	Magnetic Resonance Imaging
MRS	MR Spectroscopy
MSD	Mean Surface Distance
NCCN	National Comprehensive Cancer Network
NVB	Neurovascular Bundle
OAR	Organ-at-risk
PCa	Prostate Cancer
PD	Power Doppler
PET	Positron Emission Tomography
PI-RADS	Prostate Imaging - Reporting and Data System
PRF	Pulse Repetition Frequency
PSA	Prostate Specific Antigen
PSMA	Prostate Specific Membrane Antigen
Q1	First Quartile (25%)
Q3	Third Quartile (75%)
ReLU	Rectified Linear Unit
rmU-Net	Reconstructed Modified U-Net
RP	Radical Prostatectomy
sA/VPD	Signed Area/Volume Percent Difference
SBRT	Stereotactic Body Radiation Therapy
SD	Standard Deviation
sHD	Signed Hausdorff Distance
sMSD	Signed Mean Surface Distance
SR	Sagittally Reconstructed
TNM	Tumour-Node-Metastasis
TPS	Thin Plate Spline
TRE	Target Registration Error
TRUS	Transrectal Ultrasound
US	Ultrasound
VMAT	Volumetric Modulated Arc Therapy
VTK	Visualization Toolkit

Chapter 1

1 Introduction

Prostate cancer is the third highest cause of cancer death among Canadian men.[1] Prostate cancer management, including diagnosis and treatment, commonly utilizes minimally invasive procedures such as biopsy and brachytherapy. These procedures rely on medical imaging, particularly ultrasound and magnetic resonance imaging, for target definition and diagnosis, treatment guidance, and treatment planning. Use of these imaging methods for minimally invasive procedures also introduces challenges including time-consuming manual prostate segmentation, difficulty localizing implanted needle tips, and variable cognitive fusion of magnetic resonance and ultrasound images. This thesis is focused on overcoming these challenges through the development, validation, and implementation of software- and hardware-based tools designed specifically for minimally invasive prostate cancer procedures.

Chapter 1 provides foundational background knowledge including a description of prostate cancer epidemiology (1.1), an overview of current prostate cancer diagnosis (1.2) and treatment (1.3) techniques, and a description of how medical imaging is used in prostate cancer disease management (1.4). The remaining sections motivate the work in Chapters 2-5 through a description of current challenges in minimally invasive prostate procedures (1.5) and an outline of the thesis including specific objectives for each study (1.6).

1.1 Prostate Cancer

Prostate cancer is the most commonly diagnosed non-cutaneous cancer in Canadian men, with 24,000 cases or 20.3% of all new cancer cases in men projected for 2021.[1] This corresponds to 1 in 8 Canadian males expected to be diagnosed with prostate cancer in their lifetime.[1] Worldwide, the incidence of prostate cancer is second only to lung cancer

accounting for 13.5% of new cancer cases in men and an estimated 1.3 million cases in 2018.[2] Prostate cancer incidence increases as age increases with an average age of 66 years at initial diagnosis.[3]

In contrast to the high incidence rate of prostate cancer, the mortality rate is considerably lower with only 1 in 29 Canadian men expected to die from the disease. Prostate cancer is the third highest cause of cancer death in Canadian men with an estimated 10.1% of total cancer deaths behind only lung and colorectal cancer.[1] Prostate cancer represents an even lower share of mortality rate due to cancer in men worldwide at only 6.7% in 2018.[2] Reduced mortality rates may be due to earlier detection with the prevalence of prostate specific antigen (PSA) testing and improved management of the disease after detection.[4] Ultimately, men diagnosed with prostate cancer will most likely die of other causes.[5]

The prostate is an accessory gland in the male reproductive system that lies inferior to the bladder and anterior to the rectum, with the urethra passing through it.[6] The prostate has been defined using a zonal structure,[7] with an anterior nonglandular zone, transitional zone surrounding the urethra, central zone surrounding the ejaculatory ducts, and a large peripheral zone, collectively composed of 30 to 40 individual glands lined with epithelial cells.[6] Secretions from the prostate and neighbouring seminal vesicles contribute to the formation of semen, with the prostate contributing nutrients, enzymes, and buffers.[6,8] Due to its glandular structure, most prostate tumours are adenocarcinomas, occurring primarily in the peripheral zone, while benign prostatic hyperplasia (BPH) or nonmalignant overgrowth occurs primarily in the transitional zone.[9] The presence of prostate cancer may cause adverse symptoms including blood in the urine, frequent urination, and pain.[10]

As the prostate is a non-vital accessory gland, the 5-year survival rate for men diagnosed with prostate cancer is high at approximately 98% in the USA,[11] with the outcome of prostate cancer driven largely by metastases, or the spread of cancer to nearby organs.[12] The most common sites for prostate cancer metastasis include bone, lung, liver, pleura, and adrenals, with more than 30% of prostate cancer patients expected to harbour metastatic tumours at the time of death.[12] This stresses the importance of early detection for prostate cancer prognosis and the facilitation of safe increases in treatment aggressiveness for high-risk disease.

1.2 Prostate Cancer Diagnosis

Prostate cancer diagnosis and staging are required prior to deciding on a course of treatment. The diagnosis pathway includes initial screening (1.2.1), definitive diagnosis with biopsy (1.2.2), and disease staging (1.2.3).

1.2.1 Initial Screening Methods

Initial prostate cancer screening is typically completed using two techniques in combination: digital rectal examination (DRE) and prostate specific antigen (PSA) blood test.[13]

The DRE procedure involves the physician palpating the prostate through the rectal wall using a finger inserted into the patient's rectum. A tumorous prostate may feel hard or asymmetric relative to healthy prostate and lumps may be present on the prostate surface, all of which may be felt by the physician during a DRE.[6] The efficacy of DRE has not been definitively demonstrated, with Schröder *et al.* showing that DRE alone could detect 55.8% of confirmed cancers and was strongly dependent on PSA level with low performance in low PSA ranges (< 4.0 ng/mL).[14] As such, an abnormal DRE examination is typically paired with in subsequent PSA blood testing as well as a potential biopsy.

A blood test to determine PSA levels is a common screening technique for prostate cancer. PSA is a protein produced exclusively by epithelial cells in the prostate and can be detected in a blood sample.[15] PSA is produced in both healthy and cancerous prostates, with elevated concentrations of PSA observed in 25 to 92% of prostate cancer patients depending on prostate size.[16] PSA is typically measured in nanograms (ng) per millilitre (mL), with PSA concentration above 4.0 ng/mL traditionally indicating suspicions of prostate cancer.[17] One of the difficulties with PSA screening is the elevated levels of PSA observed in non-cancerous prostate conditions, such as in 30 to 50% of benign prostatic hyperplasia patients, limiting the specificity of the test.[16] While PSA testing is currently the most sensitive biomarker for monitoring prostate cancer, its utility in screening and diagnosis is not clear. In particular, the intensification of PSA testing has led to a dramatic increase in the diagnosis of prostate cancer since its introduction, peaking in 1993 in Canada,[1] resulting in the identification of low-risk cancer that may not pose a risk to the

patient. This has led to the overdiagnosis and overtreatment of prostate cancer, exposing patients to the adverse side effects of biopsy or radical treatment without clearly defined benefit. To combat the trend of overdiagnosis in prostate cancer, the US Preventive Services Task Force advised against PSA screening for men over 65 and since 2014, PSA screening is no longer suggested for asymptomatic men of any age, with Canada adopting similar guidelines in 2014.[18,19]

Currently, the decision to use asymptomatic PSA screening is an individual choice based on discussion with a physician. Screening is typically used in men suspected of having prostate cancer, with optimal results obtained using a combination of DRE and PSA testing, with transrectal ultrasound imaging to detect hyperechoic lesions included in certain cases as well.[14,16] Abnormal DRE, abnormal ultrasound imaging results, PSA levels greater than 4.0 ng/mL, and clinical suspicion of prostate cancer are indications for prostate biopsy to definitively confirm the presence of cancer.

PSA testing is also routinely employed after therapy to assess treatment outcome, specifically looking for rising PSA levels, known as biochemical recurrence, which may be a predictor of subsequent cancer recurrence and/or metastasis.[20] Biochemical disease-free survival is a common metric reported for prostate cancer treatment assessment in clinical trials, defined as patient survival time with little to no rise in PSA levels and no evidence of local recurrence or global metastasis.

1.2.2 Prostate Biopsy

Prostate biopsy is the gold standard for definitive diagnosis of prostate cancer after suspicions arise during screening.[21] Core needle biopsy procedures involve the removal of small samples, or so-called cores, of prostate tissue throughout the gland using a hollow needle inserted into through the rectum or perineum. Core samples are stained and analyzed under microscope by a pathologist to determine if cancer cells are present. Biopsy needle insertion is typically guided by transrectal ultrasound (TRUS) imaging, offering visualization of the needle as its inserted into the prostate. The selection of core location and number of cores is variable between biopsy techniques.

The original gold standard prostate biopsy method used the sextant technique, a systematic approach where six cores were extracted from the base, middle, and apex of the

prostate.[22] Levine *et al*. demonstrated that the systematic sextant technique had false negative rates of approximately 30%, meaning patients with prostate cancer were improperly diagnosed as cancer-free in 30% of biopsies.[23] To address this limitation and decrease the false negative rate, increasing the number of biopsy cores to 8 or 11 was proposed, demonstrating increased detection rates.[24,25] The completion of two consecutive sextant biopsies in the same procedure was also shown to increase the number of cancer detected by 30%.[23] Due to the increase number of core samples removed from the prostate, these approaches also increased the patient burden. Alternative approaches have proposed targeted biopsies using imaging. TRUS approaches targeting hyperechoic or hypoechoic regions in addition to the standard sextant biopsy locations demonstrated increased sensitivity.[26] More recently, magnetic resonance imaging (MRI) targeted biopsy methods have been proposed,[27] and will be described in further detail in section 1.4.3.

Regardless of the biopsy technique used, the pathology report generated based on the extracted samples is required for prostate cancer staging.

1.2.3 Prostate Cancer Staging

As prostate cancer is a heterogenous disease,[9] prostate cancer staging to determine the cancer extent and aggressiveness is required to determine the appropriate treatment response. Prostate cancer staging and risk assessment most commonly includes consideration of Gleason score,[28] PSA level,[16] and TNM-stage.[29]

Gleason score is determined based on the histologic arrangement of cancer cells in H&E stained prostate tissue slides assessed under light microscopy.[28,30] The Gleason grading system defines five grades, labelled 1 through 5, based on the differentiation in the pattern of growth of the tumour compared to healthy prostate cells. Prostate cancer cells are less differentiated than healthy prostate cells, and this differentiation scales with cancer aggressiveness. The Gleason score is then reported as the most common or primary pattern grade plus the second most common pattern grade, thus ranging from 2-10. If only one pattern is present, the grade is multiplied by two to give the Gleason score. For example, if the predominant grade was 3 with a second most common grade of 4, the Gleason score would be 7, alternatively reported as 3+4.

PSA concentration in the patient's blood, as described in section 1.2.1, is also included in assessment of risk level. PSA levels less than 10 ng/mL are considered low risk, PSA levels between 10 and 20 ng/mL are considered intermediate risk, and PSA levels above 20 ng/mL are considered high risk.[31]

The TNM system for cancer staging, published by the Union for International Cancer Control, is one of the most common systems for classification of malignant tumours.[29] This system characterizes the extent of the cancer in three categories. First, the extent of the primary tumour (T), second, whether the cancer has spread to nearby lymph nodes (N), and third, whether the cancer has metastasized to other parts of the body (M). The T-category ranges from 1 to 4 as primary tumour extent increases, while N- and T-category are yes or no categories, presented as 1 or 0, respectively. TNM characterization along with Gleason grade and PSA level are also used to define the overall stage of prostate cancer, ranging from I through IV, with the higher stage denoting higher spread of cancer.[29]

Overall risk assessment of prostate cancer considers all available diagnostic information including Gleason score, PSA level, and TNM-stage, as defined by the National Comprehensive Cancer Network (NCCN).[32] The defined risk levels stratify patients into either very low risk, low risk, intermediate risk, or high risk based on cancer aggressiveness and risk of early mortality. Optimal treatment courses differ depending on risk level assessment for each patient, making cancer staging a critical component of the treatment process. Treatment options will be described in detail in the following section.

1.3 Prostate Cancer Treatment

As described in section 1.2.3, prostate cancer risk level plays a critical role in deciding the optimal course of treatment. This may include a combination of surgery (1.3.1), systemic treatments (1.3.2), radiation therapy (1.3.3), or focal therapy (1.3.4), all of which are described in detail in this section. The widespread use of PSA testing has led to increased prostate cancer diagnosis rates, and thus overtreatment of patients where the cancer poses little risk to early mortality.[33]

Corresponding to the need to reduce overtreatment of prostate cancer, the witholding of radical treatment is a common method of disease management for many low-

risk cases, so-called active surveillance.[34,35] This allows patients to avoid the adverse side effects of radical therapy until the potential benefits of treatment outweigh the side effects. Active surveillance involves the careful monitoring of prostate cancer patients using diagnostic tests such as imaging, PSA testing, and biopsies, deferring treatment until the first signs of higher-risk disease.[35] This relies on the asumption that the time between initial diagnosis and disease progression is long for low-risk disease.[36] Criteria for active surveillance is variable, but commonly includes patients with Gleason score less than 6, clinical stage less than T2a, and PSA levels less than 10 ng/mL.[34] Based on this criteria, up to 36% of men diagnosed with prostate cancer may be suitable for active surveillance.[34] At the first signs of higher-risk disease, radical treatment can be initiated using techniques described in sections 1.3.1-1.3.4. In contrast, for elder men diagnosed with prostate cancer, so-called watchful waiting is common, where treatment is deferred without diagnostic monitoring until symptoms develop at which point palliative treatment is initiated.[35]

1.3.1 Surgery

Surgery in the form of radical prostatectomy (RP) is a common treatment approach for clinically organ-confined prostate cancer.[37] RP involves the complete removal of the prostate gland, and may include the removal of lymph nodes around the prostate gland if there is suspicion of cancer spreading to the nodes. Retropubic RP, an open surgical approach where an incision is made in the lower abdomen is the most common technique.[38] Advances in RP surgical technique have since focused on increased nerve-sparing to preserve patient urinary and erectile function.[37]

Laparoscopic RP has been proposed, demonstrating reduced operative and postoperative morbidity as well as equivalent short-term outcomes compared to standard retropubic RP.[39] The difficulty of laparoscopic surgeries for inexperienced physicians has prompted the development of robot-assisted laparoscopic surgeries utilizing high degree-of-freedom robots such as the da Vinci system that can mimic hand movements with high accuracy.[40] Although laparoscopic and robot-assisted RP may reduce blood loss resulting from surgery, there is no evidence of the superiority of any surgical approach in terms of oncologic outcome.[37]

Survival outcomes following RP are good, with 5-year biochemical disease-free survival of up to 92% for low-risk prostate cancer and 35% for high-risk prostate cancer.[41] Severe toxicities associated with RP including erectile disfunction and incontinence can result in decreased quality of life.[42] Recent advances in radiotherapy have improved oncologic outcome to levels comparable to radical prostatectomy for organ-confined cancer, often with decreased severity of toxicities.[43] As such, the treatment of localized prostate cancer has shifted focus to radiation therapy and focal therapy approaches, described in sections 1.3.3-1.3.4.

1.3.2 Systemic Treatments

Systemic therapy targeting the whole body is another prostate cancer treatment technique, most commonly in the form of androgen deprivation therapy (ADT)[44] or chemotherapy,[45] often employed for advanced prostate cancer that has metastasized to other organs.

Androgens such as testosterone and dihydrotestosterone are made primarily by the testicles and excess production can stimulate prostate cancer cells to grow. ADT aims to lower androgen levels, depriving the prostate cancer cells which can cause the prostate cancers to shrink or grow more slowly.[46] Orchiectomy, the surgical removal of the testicles, was the most common form of ADT, but was shown to have a large psychological burden and thus has been rejected in favor of medical castration in recent years.[47] Medical castration involves the delivery of hormonal agents such as estrogens, antiandrogens, and luteinizing hormone-releasing hormone (LHRH) agonists.[47] While ADT is an effective palliative treatment, it has several adverse side effects including decreased libido, impotence, hot flashes, metabolic alterations, and changes in cognition and mood.[46] Not all cancers respond the same to ADT, so treatment of hormone-resistant prostate cancer is an active area of study.[48] In addition to palliative treatment for patients with advanced prostate cancer, ADT has also been proposed as a neo-adjuvant therapy prior to radiation therapy to shrink the cancer or after therapy if PSA levels rise.[44]

Chemotherapy involves the delivery of cytotoxic agents to kill cancer cells, stopping further growth and preventing cell division. For the management of prostate cancer, chemotherapy is commonly employed for palliative care of hormone-resistant disease.[45] The use of chemotherapy as an adjuvant and neoadjuvant treatment modality has

been proposed.[45] As with ADT, chemotherapy has several adverse side effects including hair loss, loss of appetite, nausea, vomiting, diarrhea, and fatigue.

1.3.3 Radiation Therapy

Radiation therapy, also known as radiotherapy, relies on the delivery of high energy photons or particles to destroy cancer cells. For photon-based radiation therapy, so-called indirectly ionizing radiation, the incoming photons first interact with atoms in the patient's body to produce charged particles, which can then damage the cancer cell's deoxyribonucleic acid (DNA).[49] In contrast, for particle-based or directly-ionizing radiation therapy, charged particles such as electrons or protons are delivered, damaging the DNA without requiring prior interactions. Charged particles damage DNA through either direct action, where the charged particle interacts directly with the path of DNA to damage the strands or base pairs, or indirect action, where the charged particles break down water molecules near the DNA, producing free radicals which diffuse toward the DNA and damage the strands or base pairs.[49] DNA damage can be in the form of base damage or strand breaks. Single-strand breaks in the sugar/phosphate DNA backbone due to damage from charged particles can combine with single-strand breaks on the opposite side of the DNA molecule, leading to double-strand breaks. The majority of DNA damage is repaired by the cell but misrepair or incomplete repair can lead to chromosome aberrations, which have been shown to correlate directly with cell survival.[50]

Radiotherapy treatment is typically fractionated, thus delivered over numerous doses or fractions. The outcome of fractionated radiotherapy treatment is influenced by five factors: cell repair of sublethal damage, redistribution through the cell cycle, repopulation between fractions, reoxygenation of the tumour, and the inherent radiosensitivity of the tissue in question, the so-called 5 R's of radiotherapy.[51] Cell survival after fractionated radiotherapy can be described using a linear quadratic model.[52] The linear alpha term represents the susceptibility to single-track damage, or the sensitivity at low dose. The quadratic beta term represents the sensitivity to dual-track damage, which is affected by dose rate and interim chromosome repair. The effectiveness of the fractionation scheme can thus be evaluated using the tumour's alpha over beta (α/β) ratio, which

describes the tissue susceptibility to different types of cell damage.[53] The α/β ratio, with units of Gy, is the dose at which the two cell killing effects, single-track and dual-track damage, are equal, thus making it a measure of insensitivity to dose fractionation scheme. Cells with a large α/β ratio are less sensitive to the effects of fractionation. Most healthy tissue is classified as late reacting with α/β ratio on the order of 3 Gy, while most human tumours are classified as early reacting tissue with α/β ratio on the order of 10 Gy.[53] In this case where the α/β ratio of the tumour is much larger than the healthy tissue, dose hyperfractionation offers greater healthy tissue sparing and a high therapeutic index as the tumour response is higher than the normal tissue response for a given dose. Prostate cancer is unique as it has an estimated α/β ratio of 1.2-1.5 Gy, making it very sensitive to fractionation.[53] Thus, for prostate cancer treatment, since the α/β ratio of the tumour is lower than the healthy tissue, hyperfractionation is not ideal as it spares the tumour more than healthy tissue. This suggests that decreased fractionation, so-called hypofractionated radiotherapy, may offer an improved therapeutic index for prostate cancer.

Radiotherapy is commonly classified into two approaches: external beam radiation therapy and brachytherapy, described in detail in the following sections. Radiotherapy can be delivered as a monotherapy but is commonly included as a follow-up to or concurrent with other treatment modalities described in previous sections such as ADT.

1.3.3.1 External Beam Radiation Therapy

External beam radiation therapy (EBRT) involves the delivery of ionizing radiation from a source outside of the body. Most commonly, the ionizing radiation source is high energy (6 MV or higher) x-rays generated using a linear accelerator (linac). Linac-based EBRT delivery utilizes several techniques to deliver dose to the prostate while limiting exposure for nearby organs at risk (OAR) such as the rectum and bladder. This includes a rotating gantry allowing radiation beams to enter the patient from different angles to reduce entry dose as well as collimation in the form of moveable perpendicular jaws and highly controllable multileaf collimators (MLCs) to ensure dose is conformal to the prostate. In the pursuit of reducing dose delivered outside of the prostate, EBRT techniques for prostate cancer treatment have progressed from three-dimensional conformal radiation therapy

(3DCRT),[54] where multiple stationary, shaped beams deliver radiation, to intensity modulated radiation therapy (IMRT),[55] with dynamic radiation beams that vary in intensity, to volumetric modulated arc therapy (VMAT),[56] where the MLC, gantry speed, and dose rate are simultaneously varied while the linac gantry moves in a continuous arc. These delivery techniques in addition to advances such as image-guided radiotherapy and implanted target fiducials allow for highly conformal radiation treatment plans and have helped limit dose to healthy tissue, reducing treatment toxicities.[57] Traditionally, EBRT dose fractionation for prostate cancer treatment involved the delivery of conventional 2 Gy fractions for a total of 70-80 Gy dose, delivered five days per week over several weeks, with higher total dose corresponding to superior disease control but also increased toxicity.[57] The standard of care has recently shifted towards the use of moderate hypofractionation with 2.5-3 Gy fractions after several large randomized phase 3 trials demonstrated similar efficacy and toxicity compared to conventional fractionation.[58–60] Acute adverse side effects following radiotherapy include urinary symptoms such as increased urination frequency and nocturia in 28-50% of patients, and bowel symptoms such as increased bowel movement frequency and loose stools in 9-26% of patients, relative to baseline function.[61] Survival outcomes following EBRT approximately match or exceed survival following radical prostatectomy, with 5-year biochemical disease-free survival of up to 90% for low-risk prostate cancer and 55% for high-risk prostate cancer.[41]

Radiotherapy, including EBRT is a recommended treatment option for localized intermediate- and high-risk disease based on The American Society of Clinical Oncology (ASCO) clinical practice guidelines.[62] EBRT can be delivered as a monotherapy, but the inclusion of concurrent ADT is suggested by the ASCO, in particular for high-risk disease.[62] Radiotherapy delivery to the prostatic bed after radical prostatectomy has been shown to reduce the risk of cancer recurrence in patients with aggressive prostate cancer, while also resulting in increased adverse side effects.[63]

As described in the previous section, the standard 2 Gy fractionation scheme may not be optimal for prostate cancer treatment due to the low α/β ratio. Hypofractionated EBRT has been suggested and may offer an improved therapeutic index. Hypofractionated EBRT includes moderate fractionation with fraction size between 2.4-3.4 Gy and ultra-hypofractionation, also known as stereotactic body radiation therapy (SBRT), with fraction

size greater than 5 Gy.[64] Several phase 3 trials have shown that moderate hypofractionation offers similar efficacy and toxicity compared to standard fractionation.[58,59] Results of the phase 3 HYPO-RT-PC trial comparing SBRT (42.7 Gy total with 6.1 Gy/fraction) to standard fractionation (78 Gy total with 2 Gy/fraction) demonstrated non-inferiority for the ultra-hypofractionated approach.[64] These results along with the reduced patient burden and treatment time make hypofractionated EBRT a promising treatment method for localized prostate cancer.

Proton-based EBRT has recently been proposed for prostate cancer treatment, taking advantage of the unique properties of protons including increased biological effectiveness in cell killing compared to photons and sharp peak in the depth dose profile allowing for the creation of highly conformal treatment plans.[57] Proton-based EBRT has been to shown to allow increased whole-gland dose to the prostate compared to photon-based EBRT.[65] However, the lack of evidence for improved outcomes compared to photon-based approaches and the high cost of a initiating and maintaining a proton therapy facility has limited widespread adoption for prostate cancer.

1.3.3.2 Brachytherapy

Brachytherapy (BT) involves the delivery of ionizing radiation from a radioactive source or sources placed temporarily or permanently inside or adjacent to the treatment target. This close proximity to the treatment target allows for highly conformal treatment plans that can deliver a high dose to the target while limiting dose to the surrounding healthy tissue.[65] BT can be delivered as a monotherapy or in conjunction with EBRT.[66,67] BT treatments can be separated into low-dose-rate permanent implant BT and high-dose-rate temporary implant BT and are described in detail in the following sections.

High-Dose-Rate Brachytherapy

High-dose-rate (HDR)-BT for prostate cancer treatment is the primary focus of this thesis. In this treatment, approximately 12-20 hollow needle applicators are inserted into the prostate through the perineum, guided by a needle template. The needle template is a rigid grid of evenly spaced holes, typically 5 mm apart, helping to ensure spacing between the

needles and aid in insertion into the prostate. In addition, the needle template can lock the needle position in place either individually or collectively to ensure there is no movement prior to treatment delivery. Needle insertion is typically guided using transrectal ultrasound, described in detail in section 1.4.2. Using the needle template and intraoperative ultrasound imaging, the needles are inserted throughout the prostate to maximize radiation coverage while avoiding dose hotspots at the urethra. Following needle insertion, a single high-activity radioactive source is moved through the inserted needles by a remote afterloader to irradiate the prostate. HDR-BT procedures typically utilize iridium-192 (^{192}Ir), which has a half-life of 73.8 days and decays by β^- decay 95% of the time, leading to numerous gamma emissions from 0.11 to 1.378 MeV as well as electrons with energy up to 1.377 MeV.[68] The mean photon energy from ^{192}Ir decay is approximately 0.3 MeV, capable of delivering more than 100 Gy per hour. Treatment follows a personalized radiation plan, which specifies how long the radioactive source will remain in specific dwell positions, known as the dwell time. These specific source dwell positions are based on needle tip locations identified using imaging during the procedure. An example transrectal ultrasound image with highlighted dwell positions overlaid on an inserted needle is shown in Fig. 1.1. The treatment plan is created based on the prostate and OAR anatomy (rectum, urethra, bladder) as defined using intraoperative medical imaging to ensure the prescription dose is delivered to the entire prostate while limiting dose to the defined OARs.

An example axial transrectal ultrasound image and the corresponding radiation treatment plan for a clinical high-dose-rate prostate brachytherapy procedure is shown in Fig. 1.2. Isodose levels are provided in the legend, showing the hotspots surrounding each needle as well as the dose avoidance for the urethra. As described in the previous paragraph, the treatment plan is defined such that the 100% isodose line in red corresponding to 15 Gy covers the whole prostate.

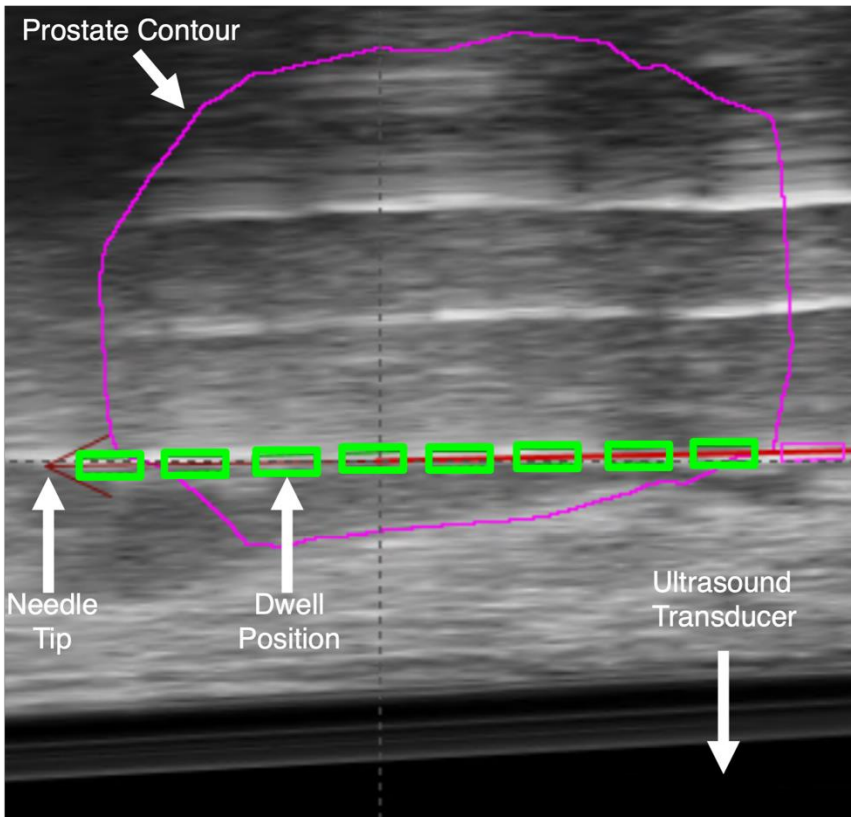

Figure 1.1: Example sagittal brightness-mode transrectal ultrasound image showing three inserted needles in a clinical high-dose-rate prostate brachytherapy procedure. On the needle closest to the ultrasound transducer, the dwell positions are shown, which are the locations where the radioactive source will sit for specific amounts of time.

HDR-BT is commonly delivered as a 15 Gy boost dose in conjunction with EBRT for treating intermediate to high-risk prostate cancer patients.[69] In a large study of more than 500 patients, 5-year freedom from biochemical failure was 91% for intermediate-risk prostate cancer patients.[69] HDR-BT as a monotherapy for treating intermediate to high-risk prostate cancer patients has been proposed.[67,70,71] HDR-BT monotherapy dose prescriptions range from seven 6.5 Gy fractions[67] to two 13.5 Gy fractions,[71] with biochemical disease-free survival at 5 years ranging from 81% for high-risk patients[67] to 93-98.6% for intermediate-risk patients.[67,70,71] The demonstrated disease-free survival for HDR-BT monotherapy as well as the reduction in patient burden due to greatly reduced treatment time make this a promising option for treating localized intermediate to high-risk prostate cancer. In addition, the hypofractionated nature of HDR-BT takes advantage of

the low α/β ratio for prostate cancer. Compared to other radiotherapy treatment options, HDR-BT is able to deliver a prescribed whole-gland dose with decreased dose to surrounding healthy tissue compared to intensity-modulated radiation therapy, proton therapy, and LDR-BT.[65]

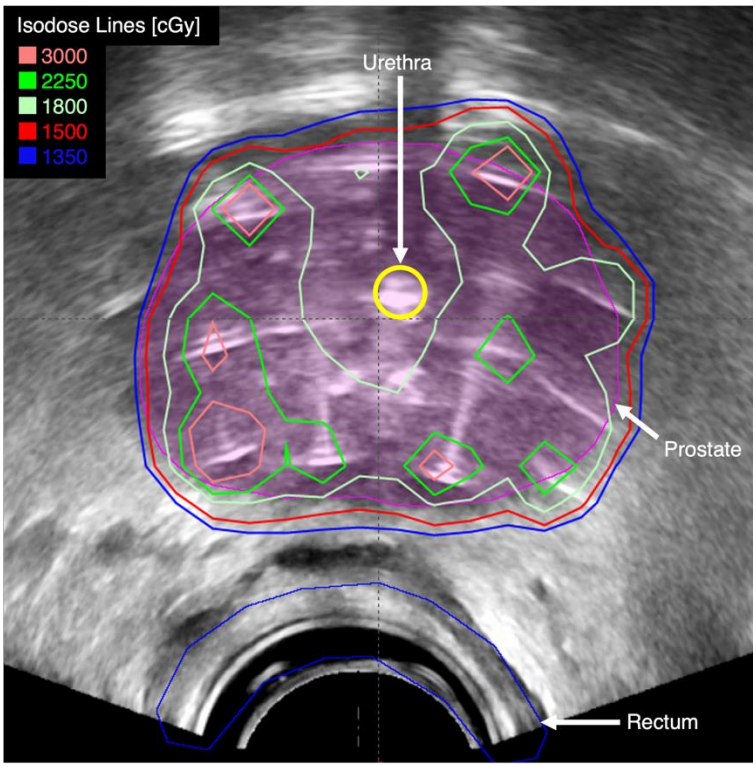

Figure 1.2: Example axial brightness-mode transrectal ultrasound image and corresponding radiation treatment plan isodose line overlay for a clinical high-dose-rate prostate brachytherapy procedure. The prostate was contoured and shaded in purple and the urethra is shown as the yellow circle. Isodose line dose levels are provided in the legend. Each implanted needle is shown as a small horizontal line.

Low-Dose-Rate Brachytherapy

Low-dose-rate (LDR)-BT involves the permanent implantation of numerous low activity radioactive sources within the prostate through inserted hollow needle catheters. Similar to the HDR-BT procedure, needles are typically inserted through a needle template while guidance is completed with transrectal ultrasound. Commonly used LDR radioactive sources include iodine-125 (^{125}I) and palladium-103 (^{103}Pd), which have half-lives of 59.4

days and 17 days and decay by electron capture with energy of approximately 28 keV and 21 keV, respectively.[72] Whole gland prescription dose is 144 Gy and 115-120 Gy for LDR-BT monotherapy and 100-110 Gy and 80-90 Gy for LDR-BT boost following EBRT according to the American Brachytherapy Society (ABS) guidelines for ^{125}I and ^{103}Pd, respectively, typically requiring 50-125 radioactive seeds.[73] LDR-BT is commonly used as a monotherapy for low-risk prostate cancer patients, with 5-year biochemical disease-free survival of up to 94%.[41] LDR-BT has also been demonstrated to improve disease control in intermediate- and high-risk prostate cancer patients when delivered as a boost dose following EBRT.[74] Treatment planning simulations have shown that HDR-BT can deliver more conformal plans compared to LDR-BT, making it favored for intermediate to high-risk localized prostate cancer treatment.[65] Furthermore, the constant irradiation delivered in LDR-BT treatments does not take advantage of low α/β ratio for prostate cancer.

Tumour-Targeted Brachytherapy

As described in the previous section, conventional BT aims to deliver a uniform dose to the whole gland. The maximum deliverable dose is limited by the surrounding OARs, including the rectum, bladder, and urethra, with the frequency of grade 3 toxicities increasing for whole-gland HDR-BT dose escalation beyond the standard 15 Gy.[75] Previous studies have shown that the site of local recurrence following radiation therapy is often the dominant intraprostatic lesion (DIL), which is the site of highest cancer burden prior to treatment.[76–79] To take advantage of this, tumour-targeted HDR-BT has been proposed to escalate dose to the DIL while maintaining the conventional whole-gland dose, which may lead to improved tumour control while maintaining acceptable rates of toxicity. With standard TRUS-guided BT, however, the DIL is not visible, necessitating the use of a supplementary imaging modality to localize the DIL, such as MRI or positron emission tomography (PET). These alternative imaging modalities are described in sections 1.4.3 and 1.4.4, respectively, including discussion of the registration methods required to integrate this supplementary information into the intraoperative US imaging. By maintaining the conventional whole-gland dose, the risk of delivering reduced dose to the cancer not identified using imaging is minimized. Boosted dose to the DIL can be achieved

with EBRT approaches as well as both HDR- and LDR-BT approaches, allowing strategic placement of hotspots to overlap with the defined DIL.[80,81] In HDR-BT focal boost procedures, elevated dose to the DIL can be achieved by increasing the dwell times for specific dwell positions near the DIL, or additional needles can be inserted, directly targeting the identified DIL. Combined HDR- and LDR-BT dose escalation, utilizing whole-gland HDR-BT followed by targeted DIL boost with LDR-BT seed implantation has been proposed in the DELIGHT trial which began recruiting at Sunnybrook Health Sciences Centre and is now recruiting at the London Regional Cancer Program (NCT03323879). The level of dose boost to the DIL is highly variable depending on the treatment centre and clinical trial with no current accepted standard.[80,81] Limitations include lack of standardized procedures for DIL localization and selection of dose boost level as well as a lack of randomized trials demonstrating efficacy, suggesting that focal boosting cannot be considered the standard of care at this time.[80]

1.3.4 Focal Therapy

In contrast with tumour-targeted BT, focal therapy is used to exclusively treat the tumour(s), thus sparing all surrounding healthy prostate tissue and potentially reducing morbidity while maintaining cancer control.[82] For focal therapies, a supplementary imaging modality is once again required to identify the tumour/DIL, such as MRI or PET. The most common focal therapy techniques include cryotherapy, high-intensity focused ultrasound, and thermal ablation.[82] Focal therapy approaches rely on accurate identification and localization of the tumour(s) using imaging, as any missed cancer will not be treated in a focal approach. As such, current ASCO guidelines do not recommend focal therapy for high-risk cancer.[62] While focal therapy is a viable treatment option for intermediate-risk disease, it is not standard care as the comparative outcome evidence is lacking relative to surgery or whole-gland radiotherapy, leaving focal therapy to be most common in the salvage setting outside of a clinical trial.[62]

1.4 Imaging in Prostate Cancer Diagnosis and Treatment

Medical imaging plays a critical role in nearly all steps of prostate cancer management. This includes the use of imaging for initial disease detection and diagnosis, cancer grading, treatment planning and target definition, intraoperative treatment guidance, and assessing treatment response. The most common medical imaging modalities used in prostate cancer diagnosis and treatment planning are x-ray computed tomography (CT) (1.4.1), US (1.4.2), MRI (1.4.3), and nuclear imaging (1.4.4), all of which are described in further detail in this section.

1.4.1 X-ray Computed Tomography (CT)

In CT imaging, a low energy (70-150 keV) x-ray source is directed toward the patient, typically in a fan beam geometry. In line with the x-ray source on the opposing side of the patient, an array of detectors is arranged to detect the transmitted photons. By rotating the x-ray source and detector array around the patient, x-ray images at different angles can be acquired and reconstructed to form a 3D CT image. The difference in the photon attenuation of materials within the body determines the number of photons reaching the detector, thus providing image contrast to distinguish different anatomy in the image. Although the use of CT has limitations for cancer detection,[83] it is commonly used in prostate cancer treatment, specifically for treatment planning in EBRT[84] and in some cases for treatment guidance and planning in BT.[85]

The most common use of CT imaging is for treatment planning in EBRT. At megavoltage energies typically used for EBRT treatment, Compton scatter is the dominant photon interaction with photon attenuation directly influenced by electron density. Hounsfield units obtained from a CT image are related to electron density allowing for accurate modelling of megavoltage beam dose, necessary for treatment planning.[84] In addition, the CT image is required for the contouring of both target organs and organs-at-risk. Prior to the delivery of EBRT, on-board cone-beam CT (CBCT) imaging is typically used to ensure accurate and consistent patient positioning relative to the planning CT.[86]

CT imaging can also be utilized for both LDR and HDR-BT treatment guidance, needle position verification, and dose planning.[85,87,88] For LDR-BT, CT is the standard-of-care technique to visualize implanted radioactive seeds.[87] HDR-BT treatments employing CT-based workflows typically require post-implant imaging outside of the operating room to verify needle positions and identify organs-at-risk, necessitating patient movement that has been shown to cause needle shifts.[89] Intraoperative imaging required for needle guidance is difficult and costly with CT as it requires a specialized operating suite with an integrated in-room CT. In addition, CT suffers from a lack of real-time imaging. Furthermore, soft tissue contrast is poor in CT imaging due to the similarity in electron density between different tissue types, making prostate and tumour delineation difficult.[83] CT imaging also utilizes ionizing radiation, adding to the patient dose.

1.4.2 Two-Dimensional Ultrasound (US)

Conventional brightness (B)-mode US imaging involves the propagation and detection of ultrasonic (> 20 kHz) sound waves that pass through the body and reflect at an interface between materials to produce 2D images of internal structures. Sound wave propagation speed through a medium depends on the medium tissue properties such as density and stiffness. Computing the time between sound wave emission and reflection back to the transducer, which differs based on the medium, provides information on the location of the reflecting interface. The mechanical, longitudinal sound waves typically used in clinical US imaging range from 2 MHz to 15 MHz, with higher frequency resulting in increased axial resolution and decreased penetration. For prostate cancer imaging, transrectal ultrasound (TRUS) imaging is typically used, where the US transducer is inserted in the patient's rectum. TRUS transducers include end-fire and side-fire geometries, with modern iterations of the latter employing a bi-plane transducer with both axial and sagittal planes. TRUS imaging is primarily used intraoperatively in interventional procedures for prostate cancer diagnosis (biopsy) and treatment, where advantages over CT imaging include real-time imaging for needle insertion guidance, high portability, accessibility, cost effectiveness, and the lack of ionizing radiation. Specifically, TRUS is utilized in prostate

biopsy targeting and needle guidance[26] and in prostate brachytherapy needle guidance and treatment planning.[90,91]

TRUS imaging has been proposed as a tool for cancer screening via the detection of hypoechoic or hyperechoic lesions and abnormalities, but lack of sensitivity compared to imaging modalities such as MRI limits the clinical utility.[16] For prostate biopsy, TRUS is commonly used for guiding needle placement intraoperatively, including the use of TRUS to target hyperechoic or hypoechoic regions in addition to the standard sextant biopsy locations.[26] In prostate BT, TRUS imaging is the standard of care for guiding needle insertion intraoperatively.[90,91] Conventional TRUS imaging limitations include small field-of-view, operator dependence, and confinement to 2D image acquisition. This requires the physician to mentally relate each 2D image to the inherently 3D anatomy, further increasing operator dependence. Furthermore, conventional TRUS has poor ability to detect and localize cancer within the prostate.[92]

1.4.2.1 Doppler Ultrasound

Doppler US is an alternative US imaging technique that displays a visual representation of motion in the plane of the US image, presented as a coloured overlay on the conventional B-mode image. This relies on the Doppler shift or variation in the frequency of sound as a result of motion between the US source and receiver, with greater velocities resulting in greater frequency shifts. Doppler US techniques vary based on the method used to calculate and present the Doppler shift, with two common techniques being colour Doppler (CD) and power Doppler (PD).[93] CD can display the direction of motion in the US plane by coding Doppler shifts into different colours that are then superimposed on the B-mode image. PD provides no directionality, displaying the total integrated Doppler signal power, and thus demonstrating increased sensitivity compared to CD.[93] Doppler US can be improved with the addition of contrast agents that improve blood flow visibility.[94]

In the management of prostate cancer, Doppler US imaging has been proposed for use in cancer detection[92,94,95] as well as biopsy targeting.[96] Contrast-enhanced Doppler US imaging has shown promise in localizing prostate cancer by visualizing microvessels, which have increased density in malignant tissue.[94] Although there is promise, Doppler US cannot replace biopsies for the early detection of prostate cancer, and is currently not used

as a standard-of-care imaging technique in prostate cancer diagnosis. Contrast-enhanced CD US has also been proposed for use in targeted biopsy.[96] By targeting hypervascular regions intraoperatively, the CD-targeted biopsy approach demonstrated significantly increased detection rate compared to systematic biopsy.[96] Conventional Doppler US is still limited to 2D acquisitions, however.

1.4.2.2 Three-Dimensional Ultrasound

To overcome the limitations of conventional 2D TRUS imaging, 3D TRUS techniques have been proposed.[97–100] These techniques provide a 3D view of the internal anatomy, removing the need for physicians to mentally relate each 2D slice to the underlying 3D anatomy, thus reducing operator dependence. In addition, these techniques remove the need for a sonographer to manually manipulate the US transducer, further reducing operator dependence. While matrix array 3D US transducers exist, mechanical 3D scanning approaches, which involve the tracking and manipulation of a conventional 2D transducer through a series of angles or steps to reconstruct a 3D image are typically used due to increased field-of-view and resolution. There are numerous mechanical 3D US scanning approaches, but axially reconstructed step-back, sagittally reconstructed side-fire, and end-fire 3D TRUS imaging techniques are most commonly used in prostate cancer imaging. An example of each 3D TRUS image acquisition is shown in Fig. 1.3.

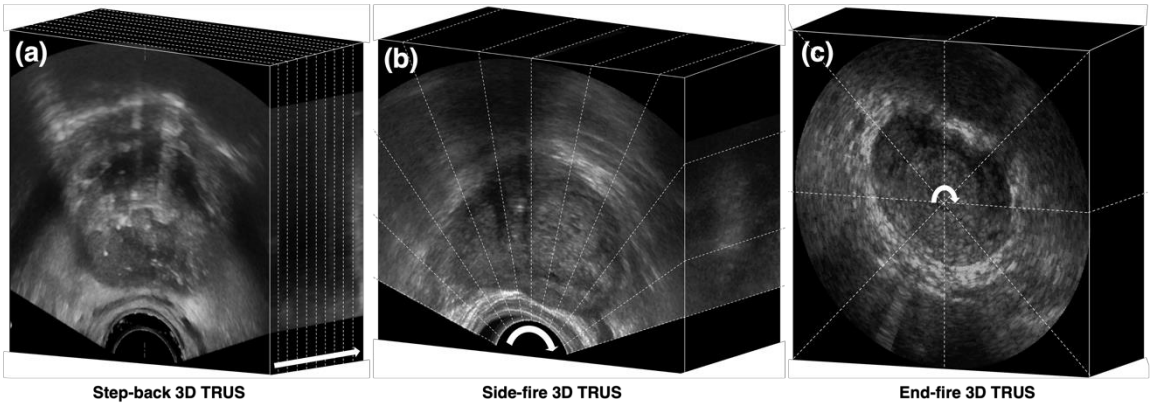

Figure 1.3: Example (a) axially reconstructed step-back, (b) sagittally reconstructed side-fire, and (c) end-fire three-dimensional transrectal ultrasound images. The dotted lines in each image represent the original ultrasound acquisition planes that compose the 3D image. Arrows represent the direction of transducer motion during 3D acquisition.

In end-fire and side-fire 3D TRUS imaging, a TRUS transducer is mechanically rotated about its long axis. 2D TRUS images are acquired at set angular intervals, which are then reconstructed to generate 3D TRUS images. End-fire 3D TRUS, commonly used for prostate biopsy, utilizes an end-fire transducer that is rotated 180° while 2D TRUS images are acquired at 1.0° intervals and reconstructed into a 3D volume.[99] In contrast, sagittally reconstructed side-fire 3D TRUS imaging, commonly used for prostate brachytherapy, utilizes the sagittal plane on the bi-plane transducer.[98,100,101] By rotating the transducer 140° while 2D TRUS images are acquired at 0.5° intervals, a fan-shaped 3D TRUS image can be reconstructed. Axially reconstructed step-back 3D TRUS images, also used for prostate brachytherapy, utilize the axial plane on the bi-plane transducer. A series of axial 2D US images is acquired by translating the transducer in steps inferiorly, allowing for 3D image reconstruction.[90,102,103]

3D TRUS imaging is used in both biopsy and brachytherapy procedures. For prostate biopsy, 3D TRUS imaging is used in MRI-3D TRUS fusion-guided biopsy, demonstrating improved tumour sampling compared to 2D TRUS guidance alone by combining the high soft-tissue contrast of MRI with real-time needle guidance capabilities of TRUS.[104] For prostate brachytherapy, axially reconstructed step-back 3D TRUS imaging is most commonly employed.[90,102,103] Intraoperative 3D TRUS imaging allows for the visualization and localization of the prostate as well as nearby organs-at-risk including the rectum, urethra, and bladder, providing improved spatial context compared to 2D TRUS. In addition, 3D TRUS allows for the verification of implanted needle positions. Through visualization of both anatomy and implanted needles, 3D TRUS imaging facilitates accurate treatment planning without the need for additional modalities such as CT. Sagittally reconstructed side-fire 3D TRUS imaging has also been proposed for prostate brachytherapy, improving needle tip identification accuracy compared to step-back 3D TRUS imaging due to increased resolution in the insertion direction.[101] 3D TRUS-guided brachytherapy procedures allow for all aspects of the treatment workflow to occur in the operating room, including intraoperative needle guidance and treatment planning, allowing the patient to remain stationary.[90,91] TRUS is a widely available and low-cost modality, offering real-time imaging and comparable needle tip identification accuracy to CT,[90] making it the preferred method for prostate HDR-BT guidance at many centers. 3D

TRUS imaging is still limited in its ability to detect and localize cancer within the prostate, so supplemental imaging modalities such as MRI and nuclear imaging are required to facilitate lesion-targeted biopsy and treatment.[92]

1.4.3 Magnetic Resonance (MR) Imaging

MR imaging uses magnetic fields and radiofrequency waves to manipulate and measure changes in the alignment of hydrogen nuclei (protons) within the patient's body to generate signals and produce 3D images of the internal anatomy. MR images have very high soft tissue contrast compared to other modalities such as CT and TRUS imaging, making them invaluable for both organ contouring and tumour identification and localization.[105–107] MRI in prostate cancer largely uses magnetic field strengths of 1.5T and 3T, with either a phased array surface coil or an endorectal coil.[83,107] Endorectal coils offer a significant increase in signal-to-noise ratio when used with 1.5T MR scanners, however they are expensive, uncomfortable for the patient, and can deform the prostate.[83] The benefit of endorectal coils compared to phased array body coils at 3T magnetic field strength is still debated, but image resolution is higher with either coil compared to 1.5T MR.[83,107] Conventional anatomical MR imaging for prostate cancer commonly uses T2-weighted and T1-weighted images.[83,107] While T1-weighted images can be used to detect post-biopsy hemorrhage, T2-weighted images are the most common pulse sequence for lesion detection as it offers improved soft tissue resolution including depiction of zonal anatomy.[83] In T2-weighted images, tumours commonly appear hypointense compared to healthy prostate tissue due to tissue density differences.[108] To improve lesion detection and localization, functional MR imaging techniques such as MR spectroscopy (MRS), dynamic contrast-enhanced MRI (DCE-MRI), and diffusion-weighted imaging (DWI) can be used.[105,106,109] The combination of anatomical T2-weighted MR imaging with the functional imaging techniques listed above is known as multiparametric MRI (mpMRI). mpMRI offers higher cancer detection rates compared to any one MR technique alone according to histological validation, demonstrating detection rates greater than 90% for moderate to high risk tumours.[105,106]

The Prostate Imaging – Reporting and Data System (PI-RADS) was developed to globally standardize the acquisition, interpretation, and reporting of prostate mpMRI examination for prostate cancer diagnosis.[108] The PI-RADS system establishes mpMRI acquisition guidelines and provides assessment categories to report the risk of clinically significant prostate cancer for each suspicious location.[108] PI-RADS assessment categories range from one to five, with a higher number denoting increased likelihood of clinically significant cancer.[108]

As the high soft tissue contrast of MRI allows for accurate tumour localization, MR imaging is typically used to target the tumour for prostate cancer diagnosis and treatment. In prostate biopsy, MR imaging can be added in an MRI-3D TRUS fusion guided biopsy approach.[104] By combining the superior soft tissue contrast of MR with the real-time needle guidance of TRUS imaging, improved tumour sampling can be achieved compared to conventional TRUS-guided biopsy.[104] In prostate cancer treatment, MR imaging can facilitate tumour-targeted treatment in the form of dose escalation or focal therapy, reducing dose to nearby organs-at-risk while increasing dose to the most likely site of cancer recurrence.[82,106] MRI can also visualize sensitive organs-at-risk such as the neurovascular bundles (NVB), which are not typically visible on CT or TRUS images, allowing for dose sparing and thus reduced probability of adverse side effects.[83] An example US and 1.5T MRI with corresponding physician contours is shown in Fig. 1.4. In all the described use cases, the MRI information is obtained from a pre-procedural MRI scan, necessitating MRI-to-US image registration techniques to integrate this information into the operating room environment where TRUS imaging is typically used. MR-only brachytherapy workflows have been proposed, utilizing MR imaging for both needle guidance and treatment planning.[110–112] MR-only workflows are limited by the lack of intraoperative real-time imaging if the MR scanner is not located in the operating room, necessitating patient movement post-implant, which may cause needle shifts.[89] Dedicated MR operating room suites are available, but require specialized BT tools and procedures, resulting in an expensive and challenging procedure. The prohibitive cost of MRI systems is still a barrier to widespread use, even in developed countries such as Canada, resulting in TRUS and CT-guided BT to still be the standard of care.

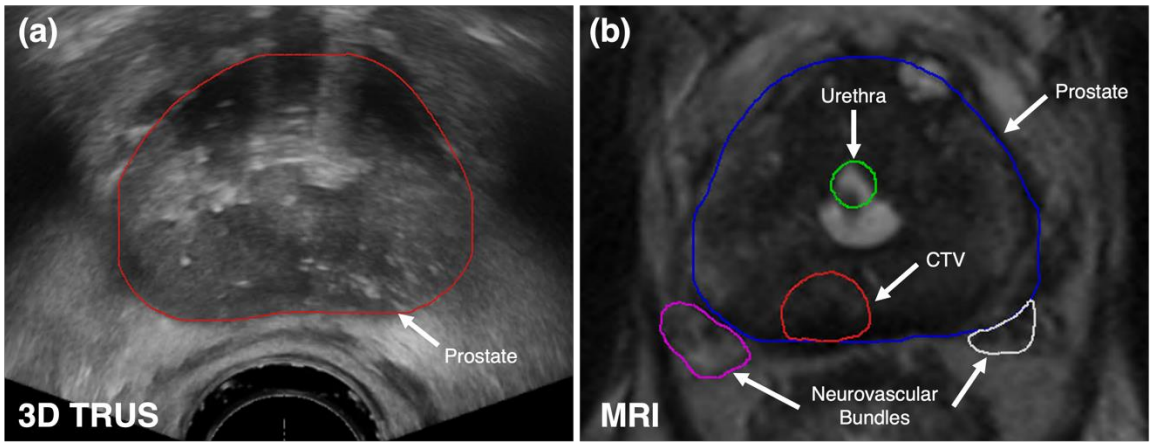

Figure 1.4: Example (a) three-dimensional transrectal ultrasound image and corresponding prostate contour in the axial plane and (b) axial T2-weighted 1.5T magnetic resonance image with physician-drawn prostate, urethra, clinical target volume (CTV), and neurovascular bundle (NVB) contours.

1.4.4 Nuclear Imaging

Nuclear imaging in the form of PET is an alternative functional imaging technique which allows for visualization of the tumour location within the prostate.[83] PET imaging involves the injection of a positron emitting radiotracer, which collects in organs with increased metabolic activity, such as cancerous tissue, emitting radiation which can be subsequently detected and used to produce an image characterizing the tracer location. Historically, fluorodeoxyglucose (^{18}F-FDG) is the most common PET radiotracer, as it serves as a marker for glucose uptake in tissue, which is correlated with cell metabolism that is typically higher in cancer cells.[113,114] Recently, prostate specific membrane antigen (PSMA)-PET has been proposed, demonstrating increased specificity and sensitivity in prostate cancer detection compared to alternative imaging modalities.[115] Gallium-based PSMA radiotracers attach to PSMA proteins that are often overexpressed in prostate cancer tissue allowing for highly specific tumour localization. Although PSMA is highly specific to prostate cancer, up to 10% of prostate cancers do not overexpress PSMA.[116] PSMA-PET can be utilized alone or in combination with MRI, which has shown to increase cancer detection rate.[115] Aside from its utility in the initial diagnosis and monitoring of prostate cancer, PSMA-PET imaging has also been proposed as an alternative to MRI for use in

targeted biopsy as well as lesion-targeted dose escalation in HDR-BT.[117] Limitations of PET include low spatial resolution, lack of structural information necessitating registration with another modality such as CT or MRI, and the increased cost and challenges involved in producing, handling, and administering radiotracers. Recently developed hybrid PET-MR scanners aim to overcome the first two limitations, offering simultaneous co-registered PET and MR image acquisition.[116]

1.5 Challenges in Minimally Invasive Prostate Procedures

Ultrasound-guided minimally invasive prostate procedures such as biopsy and HDR-BT offer several advantages in terms of detection and treatment accuracy, respectively; however, there are still numerous challenges involved in these procedures. In both procedure types, imaging is required to guide needles to precise locations within the prostate, so accurate identification of the needle and the target in the image is critical. A common challenge across procedures is that these steps are commonly completed manually during the procedure in the operating room, lengthening the procedure time as well as adding observer variability. Specific challenges are described in sections 1.5.1-1.5.4 with description of the motivation, previous work, and unmet need for each case.

1.5.1 Prostate Segmentation in Three-Dimensional Ultrasound Images

Motivation

3D US imaging is a valuable tool in both prostate biopsy and prostate BT, allowing for improved needle insertion accuracy and target definition as well as facilitating the integration of other 3D imaging modalities such as MRI intraoperatively. To effectively utilize 3D TRUS images, segmentation of the prostate gland in 3D is required. In HDR-BT, the prostate is the treatment target, so accurate segmentation is critical for target visualization during needle insertion, for glandular volume measurements,[118,119] and for dose-volume calculations during treatment planning.[120] In 3D TRUS-MRI fusion-guided biopsy and tumour-targeted HDR-BT, accurate prostate segmentation is required for use in

surface-based registration approaches.[121,122] These segmentations are typically performed manually by the physician during the clinical procedure while the patient is under sedation or anesthetic. Manual segmentation in 3D TRUS images is time-consuming as the physician must segment the 3D volume slice-by-slice which can take up to 15 minutes, extending procedure times and increasing patient risk due to anesthesia exposure.[123] This is compounded by the fact that variation in image quality can make prostate segmentation a difficult task. Fig. 1.5 shows an example high quality 3D TRUS with a clearly visible prostate boundary and an example poor quality image where image artifacts limit the visibility of the prostate boundary making segmentation difficult. Manual segmentation also leads to intra- and inter-observer variability which has been shown to have a volume percent difference of 5.4% and 11.4%, respectively, for side-fire 3D TRUS imaging.[124] Investigation into automated segmentation approaches that may decrease procedure time and reduce operator dependence is thus critical.

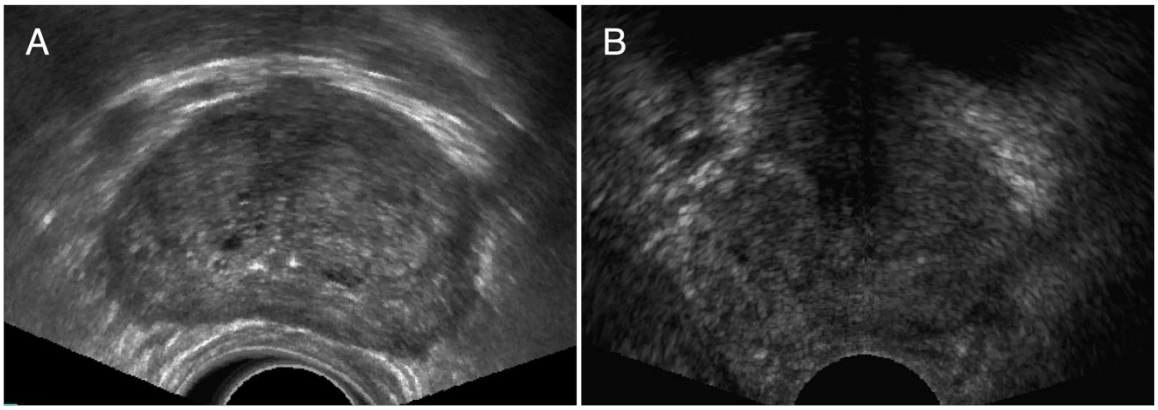

Figure 1.5: Example transrectal ultrasound images highlighting the variability in image quality and prostate boundary visibility which can make manual prostate segmentation time consuming and difficult. (A) Example high quality image with a clearly defined prostate boundary. (B) Example poor quality image with severe shadowing artifacts that limit the prostate boundary visibility.

Previous Work

Automated or semi-automated 3D TRUS prostate segmentation approaches using conventional image processing techniques have been previously investigated;[125–128]

however, these techniques have lacked clinical translation due to difficult generalizing to diverse datasets and limitations related to computational complexity and computation time. Deep learning, specifically convolutional neural networks (CNNs), have been widely investigated for use in image processing tasks, including image segmentation, promising fast and accurate automated segmentation. Prostate segmentation in both 2D TRUS[129] and 3D TRUS[130,131] using CNN-based approaches have been proposed. While the reported segmentation accuracy was high, key limitations include a lack of generalizability due to use of image datasets obtained using a single US system with matched voxel dimensions and sizes. Furthermore, the proposed approaches investigated segmentation of only one 3D TRUS image type, further limiting generalizability and necessitating the use of multiple networks for different procedure types. The proposed approaches have also relied on cross-validation where information bleeding could influence the reported segmentation performance.[132,133]

Unmet Need

While deep learning-based prostate segmentation in 3D TRUS images have been investigated, there is an unmet need for clinically translatable segmentation techniques that are generalizable and robust to diverse clinical data obtained using different US systems. In addition, the use of different 3D TRUS acquisition methods across different minimally invasive ultrasound-guided procedures leaves the unmet need to develop a consolidated segmentation approach that is generalizable to different 3D TRUS acquisition methods.

1.5.2 Variability in Dataset Size, Quality, and Diversity for Deep Learning Segmentation Applications

Motivation

Deep learning approaches for automated image segmentation tasks have been widely proposed; however, the segmentation performance and generalizability of these approaches rely heavily on the quality of the training dataset. For 3D TRUS prostate segmentation in particular, the dataset size, image quality, and diversity in terms of 3D TRUS acquisition type and US system may influence the segmentation performance and clinical translation

potential. The generation of large clinical datasets is both time-consuming and costly, making access to such datasets rare, especially for ultrasound. This is a major limitation for the widespread clinical translation of deep learning approaches, especially for smaller hospitals where data may be scarce. As described in section 1.5.1, different minimally invasive ultrasound-guided procedures utilize different TRUS acquisition methods. Datasets and segmentation methods that are effective for multiple TRUS acquisition methods would increase clinical translation efficiency and reduce the duplication of work. Finally, image quality, although critical for prostate segmentation, is highly variable based on acquisition type, presence of artifacts, and prostate visibility. These differences also make comparison of segmentation performance for methods trained and tested using different datasets challenging. Investigation into the effect of these dataset considerations on segmentation performance and generalizability may increase the clinical translation potential of deep learning segmentation approaches in addition to improving transparency in reporting and comparing segmentation results across methods and datasets.

Previous Work

Proposed deep learning-based segmentation approaches have utilized small datasets ranging in size from 40 to 109 3D TRUS images;[129–131,134,135] however, these methods are still limited to one TRUS acquisition type, typically acquired using one US system, which reduces the potential clinical translation. There is also no investigation of how performance varies as the dataset size changes. Although image quality is critical to segmentation performance, there is currently no grading scale to quantify 3D TRUS image quality.

Unmet Need

While deep learning segmentation approaches show great promise, there is an unmet need to rigorously evaluate the impact of the training dataset in terms of size, image type, and image quality. Understanding of how these factors influence segmentation performance could improve the clinical translation potential of deep learning segmentation methods. In addition, a segmentation method that is generalizable to different 3D TRUS acquisition types and able to maintain performance while utilizing small training datasets may be suitable for widespread use even if data is scarce. There is also an unmet need for a method

to quantitatively grade the image quality of 3D TRUS images to assess the impact on segmentation performance as well as improve the ability to compare segmentation performance for methods trained using different datasets.

1.5.3 Intraoperative Needle Tip Localization in Transrectal Ultrasound

Motivation

HDR-BT involves the irradiation of the prostate by passing a high-activity radioactive source through multiple hollow needles inserted into the prostate under TRUS guidance.[90] Intraoperative needle tip localization is critical for safe and effective HDR-BT treatment, as it directly influences the radiation treatment plan.[102,136,137] As shown in Fig. 1.1, the location of the dwell positions are set based on the identified tip location, so any error in needle tip localization could lead to radiation delivery that deviates from the planned dose, potentially exposing the patient to increased radiation or underdosing the tumour. Needle tip localization error is a primary component of uncertainty in HDR-BT setup, with errors greater than 3 mm potentially leading to adverse outcomes.[136] In TRUS-guided HDR-BT procedures, needle tip localization is completed using standard B-mode US; however, image artifacts including shadowing and reverberation artifacts have been shown to limit needle visibility intraoperatively.[138] Investigation into approaches to improve intraoperative needle tip visualization are critical for improving HDR-BT treatment accuracy and safety.

Previous Work

Software-based automated needle tip localization methods using both conventional image processing techniques[139–142] and deep learning[143,144] have been proposed; however, these approaches are still limited by their use of B-mode US images, and are thus sensitive to image artifacts that may limit needle visibility.

Doppler US has been proposed as an alternative imaging approach to improve surgical instrument visibility intraoperatively.[99-110] As described in section 1.4.2.1, Doppler US displays the motion in the US plane as a coloured overlay on the B-mode image, so by perturbating a surgical instrument, motion is generated that can then be imaged using

Doppler US. Previously published Doppler US-based visualization approaches include applications in biopsy, pericardiocentesis, and radiofrequency ablation, with none considering BT. Furthermore, these approaches are often cumbersome, requiring modifications or additions to the standard clinical equipment including the use of vibrating devices clipped onto the needle,[145–147,153] the attachment of piezoelectric crystals,[148,150,156] or electromagnetic actuation of small permanent magnets inside the needle.[155] In HDR-BT where more than 10 needles are typically required, these approaches are not suitable. There is also no evidence of Doppler US-based needle visualization in clinical HDR-BT procedures.

Unmet Need

While Doppler US-based needle visualization has been proposed in the past, there is an unmet need for a perturbation technique designed specifically for BT applications and requiring minimal modifications or additions to the operating room and the standard clinical equipment. In addition, there is an unmet need to validate Doppler US as a needle visualization tool intraoperatively in clinical HDR-BT procedures. Validation of a Doppler US technique may improve needle visualization intraoperatively, potentially improving both patient safety and treatment accuracy in HDR-BT and could be applied to other minimally invasive procedures where similar challenges with intraoperative needle visibility are prevalent, for example gynecologic interstitial BT or radiofrequency ablation.

1.5.4 Deformable Magnetic Resonance Imaging to Three-Dimensional Ultrasound Image Registration

Motivation

As the DIL is the most likely site of prostate cancer recurrence following radiotherapy, tumour-targeted BT featuring escalation of dose to the DIL may lead to improved tumour control while maintaining low dose to healthy tissues.[76] Since the DIL is not easily visible in TRUS images, the addition of supplementary imaging capable of visualizing the DIL, such as MRI, is required, necessitating 3D TRUS-MRI image registration. Currently, translation of targets identified in the MR image to the intraoperative TRUS image is

typically done manually by the physician using cognitive fusion or rigid registration, introducing inter- and intra-observer variability. Accurate registration is critical for effective targeting of the DIL, so automated registration methods may reduce observer dependence while also reducing registration time. Furthermore, registration of radiosensitive organs-at-risk visible in MR imaging into the 3D TRUS image, such as the neurovascular bundles, may allow dose avoidance during treatment planning which could reduce acute side effects following radiotherapy.

Previous Work

Deformable image registration methods have been proposed, commonly focusing on MR-TRUS fusion-guided biopsy.[157–162] As described in section 1.4.2.2, TRUS-guided prostate biopsy typically utilizes end-fire 3D TRUS imaging which uses a different acquisition technique compared to axially reconstructed step-back 3D TRUS commonly used in prostate BT. As such, validation of registration approaches seeking to translate MR-defined treatment targets such as the DIL or anatomical structures to avoid during therapy into the intraoperative step-back 3D TRUS images is required. Furthermore, prospective clinical integration in tumour-targeted BT procedures is limited.

Unmet Need

While deformable 3D TRUS-MR image registration approaches have been proposed, there is an unmet need for the development and validation of a registration algorithm designed for tumour-targeted BT, including avoidance of radiosensitive anatomy such as the neurovascular bundle. There is also an unmet need for prospective clinical validation and workflow efficiency assessment in prostate BT procedures.

1.6 Thesis Objectives and Outline

The **central objective** of the research described in this thesis was to improve the efficiency and accuracy of minimally invasive procedures for prostate cancer diagnosis and treatment, including image-guided biopsy and brachytherapy, through the development and validation of software- and hardware-based tools. The focus is on reducing lengthy procedure times,

improving treatment accuracy compared to standard-of-care techniques, and reducing manual operator variability. This objective is motivated by the time-consuming and difficult manual intraoperative work required in minimally invasive prostate procedures, including prostate segmentation, needle tip identification, and image registration. From the perspective of software development, we hypothesize that the integration of a generalizable deep learning-based prostate segmentation algorithm in prostate biopsy or brachytherapy procedures would allow for reduced procedure time, while the integration of a deformable image registration algorithm would allow for safe and effective targeting of MR-defined tumours as well as avoidance of MR-defined organs-at-risk, decreasing variation due to manual registration. From the perspective of hardware-based solutions, we hypothesize that a PD-based needle localization method utilizing needle perturbation from a mechanical oscillator may allow for improved needle tip visualization compared to the standard of care B-mode US, potentially improving treatment accuracy. We propose the tools in this thesis with the aim of providing easy-to-implement, generalizable solutions to common problems plaguing minimally invasive prostate cancer procedures with the end goal of improving care for men with prostate cancer through reduced procedure time and improved treatment accuracy.

The specific objectives of this thesis are to:

1. Develop an accurate and generalizable deep learning-based automatic prostate segmentation algorithm for 3D TRUS images, trained using a clinically diverse dataset of images from prostate biopsy and brachytherapy procedures.
2. Examine the effect of dataset size, image quality, and image type on deep learning-based automatic prostate segmentation in 3D TRUS images through the use of smaller training datasets and the development of an image quality grading scale.
3. Develop a novel wireless mechanical oscillator for use in a power Doppler-based needles tip localization method to improve intraoperative needle visualization in interstitial prostate brachytherapy.
4. Develop and clinically validate a surface-based deformable MR-3D TRUS image registration algorithm for interstitial prostate brachytherapy.

1.6.1 Thesis Outline

The overarching goal of this thesis is to address these specific objectives in four manuscripts (Chapters 2 to 5), briefly described in the following thesis outline.

Chapter 2: Automatic prostate segmentation using deep learning on clinically diverse 3D transrectal ultrasound images

Minimally invasive procedures for diagnosing and treating prostate cancer, such as biopsy and brachytherapy, have incorporated 3D TRUS imaging to improve needle guidance and target definition as well as for registration to other modalities such as MRI. To effectively utilize these images, manual segmentations of the prostate are required, which can be time-consuming and difficult, often occurring while the patient is under sedation or anesthetic. This chapter describes our work to develop a novel deep learning-based automatic prostate segmentation algorithm. Our algorithm utilized 2D prediction with a modified version of the U-Net architecture[163] on 12 2D US slices sampled radially about the approximate central axis of the prostate before reconstruction into a 3D surface, exploiting the symmetrical nature of the prostate. The algorithm was trained using a diverse dataset of clinical images from clinical biopsy and brachytherapy procedures acquired using end-fire and side-fire 3D TRUS acquisition geometries, respectively.

Our radial 2D plus 3D reconstruction method significantly outperformed a fully 3D network trained using the same dataset as well as state-of-the-art algorithms published in the literature. This work was the first deep learning algorithm to successfully segment the prostate in multiple TRUS acquisition geometries. Prostate segmentation was fast and accurate, with demonstrated generalizability and robustness to different ultrasound systems and acquisition types. This method has the potential to decrease procedure time and reduce operator dependence in minimally invasive procedures for prostate cancer diagnosis and treatment.

Chapter 3: Effect of dataset size, image quality, and image type on deep learning-based automatic prostate segmentation in 3D ultrasound

Deep learning-based segmentation approaches offer the potential for fast and accurate segmentation results. However, these approaches are sensitive to the training dataset used,

and as large clinical datasets are rare, widespread adoption of deep learning-based automatic segmentation has been limited. Using the 2D radial plus 3D reconstruction segmentation method described in Chapter 2, the effect of dataset size, image quality, and image type on segmentation accuracy was examined. Smaller training datasets were achieved by splitting the complete dataset into different image types and subsequently reducing the number of images in steps. For this work, a modified version of the U-Net architecture[163] as well as a U-Net++ architecture were used, as the latter has been shown to perform well with small datasets.[164,165] To assess image quality, a 3D TRUS image quality grading scale with three factors (acquisition quality, artifact severity, and prostate boundary visibility) was developed, the first of its kind to our knowledge.

For training datasets split based on acquisition geometry, the U-Net++ significantly improved performance compared to the modified U-Net, while for smaller datasets, the U-Net++ offered equivalent performance down to as small as 500 training images. For our dataset of TRUS images, image quality did not have a significant impact on segmentation performance for end-fire images, while boundary visibility grade has a significant effect for side-fire images. The image quality grading scale provides a quantitative tool for assessing segmentation performance while allowing for easier comparison between networks trained using different datasets. High performance of our approach with small training datasets supports the potential for widespread use of deep learning for segmentation tasks, even if data is scarce.

Chapter 4: A power Doppler ultrasound method for improving intraoperative needle tip localization in interstitial prostate brachytherapy

Standard B-mode US is typically used for needle guidance in HDR-BT, including localization of needle tip positions which is critical for treatment planning. Image artifacts can limit needle tip visualization in B-mode US which can cause the delivered dose to deviate from the treatment plan. This chapter describes the development and validation of a wireless mechanical oscillator for use in a novel PD-based needle localization method designed specifically for prostate brachytherapy. Our PD method was evaluated in a series of tissue-mimicking phantom experiments using a simulated HDR-BT needle implant as well as a needle implant designed to maximize shadowing artifacts. Clinical validation was

completed in five patients who underwent standard HDR-BT as part of a feasibility clinical trial.

In tissue-mimicking phantom experiments, our PD method offered identical tip localization accuracy when needles were clearly visible and demonstrated increased accuracy for shadowed needles, including providing the ability to visualize needles previously not visible using B-mode US alone. Across all five patients, our PD method demonstrated tip localization error of 0.8 ± 0.5 mm compared to 0.9 ± 0.7 mm for B-mode US, showing reduced variation and a reduction in tip error for difficult to see outlier needles. This easy-to-implement method requiring no modifications to the clinical equipment or workflow has the potential to improve needle visualization and thus treatment accuracy in HDR-BT and may be broadly applied in other minimally invasive needle-based procedures.

Chapter 5: Validation of a surface-based deformable MRI-3D ultrasound image registration algorithm towards clinical implementation for interstitial prostate brachytherapy

MR imaging offers superior soft tissue contrast compared to ultrasound, providing the ability to visualize the tumour and other points of interest such as the neurovascular bundles. To utilize this information intraoperatively to target treatment or avoid organs-at-risk, registration between MRI and 3D TRUS is required. This chapter describes the development and validation of a surface-based MRI-3D TRUS deformable image registration (DIR) algorithm, which we hypothesize will reduce variability compared to the manual cognitive fusion technique that is currently employed. Our registration algorithm was evaluated using a deformable tissue-mimicking prostate phantom comparing registration accuracy at different levels of deformation versus manual rigid registration. Clinical validation was completed in three HDR-BT clinical cases where algorithmic registrations were compared to manual cognitive fusion registrations completed by three different physicians. Our DIR algorithm was implemented intraoperatively in a prospective HDR-BT clinical case.

In phantom experiments our DIR algorithm demonstrated a mean Dice similarity coefficient (DSC) and target registration error (TRE) across all deformation levels of

0.74±0.08 and 0.94±0.49 mm, respectively, significantly improving the performance compared to manual rigid registration with 0.64±0.16 and 1.88±1.24 mm, respectively. Across the three clinical cases, our algorithm demonstrated reduced variability in DSC score compared to the inter-physician variability of the cognitive fusion approach, offering the potential to reduce operator dependence, decrease procedure time, and standardize the registration process between physicians. Clinical implementation of our algorithm allowed for a successful PSMA-PET-MR targeted biopsy, confirmed by histology, and facilitated dose scalation to the registered clinical target volume.

Chapter 6: Conclusions, Limitations, and Future Work

This chapter focuses on the overall conclusions of the research described in this thesis and will discuss the future work that could expand the presented research or address the limitations and remaining unmet clinical needs.

1.7 References

1. Canadian Cancer Statistics Advisory Committee. *Canadian Cancer Statistics* 2021. Toronto, ON: Canadian Cancer Society; 2021. cancer.ca/Canadian-Cancer-Statistics-2021-EN. Accessed March 29, 2021.

2. Wild CP, Weiderpass E, Stewart BW. *World Cancer Report: Cancer Research for Cancer Prevention*. Lyon; 2020.

3. Rawla P. Epidemiology of Prostate Cancer. *World J Oncol*. 2019;10(2):63. doi:10.14740/WJON1191

4. Wong MCS, Goggins WB, Wang HHX, et al. Global Incidence and Mortality for Prostate Cancer: Analysis of Temporal Patterns and Trends in 36 Countries. *Eur Urol*. 2016;70(5):862-874. doi:10.1016/J.EURURO.2016.05.043

5. Cooperberg MR, Broering JM, Carroll PR. Risk Assessment for Prostate Cancer Metastasis and Mortality at the Time of Diagnosis. *JNCI J Natl Cancer Inst*. 2009;101(12):878-887. doi:10.1093/JNCI/DJP122

6. Drake RL, Vogl AW, Mitchell AW. *Gray's Anatomy for Students*. 3rd ed. Churchhill Livingstone; 2015.

7. McNeal JE. Origin and development of carcinoma in the prostate. *Cancer*. 1969;23(1):24-34. doi: 10.1002/1097-0142(196901)23:1<24::aid-cncr2820230103>3.0.co;2-1.

8. Silverthorn DU. *Human Physiology: An Integrated Approach*. 7th ed. Pearson Education; 2016.

9. Shen MM, Abate-Shen C. Molecular genetics of prostate cancer: new prospects for old challenges. *Genes Dev*. 2010;24(18):1967-2000. doi:10.1101/gad.819500

10. Jønler M, Nielsen OS, Wolf H. Urinary symptoms, potency, and quality of life in patients with localized prostate cancer followed up with deferred treatment. *Urology*. 1998;52(6):1055-1062. doi:10.1016/S0090-4295(98)00449-X

11. Howlader N, Noone AM, Krapcho M, et al. SEER Cancer Statistics Review, 1975-2017, Bethesda, MD: National Cancer Institute; 2020. https://seer.cancer.gov/csr/1975_2017/. Accessed March 30, 2021.

12. Bubendorf L, Schöpfer A, Wagner U, et al. Metastatic patterns of prostate cancer: An autopsy study of 1,589 patients. *Hum Pathol*. 2000;31(5):578-583. doi:10.1053/HP.2000.6698

13. Rendon RA, Mason RJ, Marzouk K, et al. Canadian Urological Association recommendations on prostate cancer screening and early diagnosis. *Can Urol Assoc*

J. 2017;11(10):298. doi:10.5489/CUAJ.4888

14. Schröder FH, Van Der Maas P, Beemsterboer P, et al. Evaluation of the Digital Rectal Examination as a Screening Test for Prostate Cancer. *JNCI J Natl Cancer Inst*. 1998;90(23):1817-1823. doi:10.1093/JNCI/90.23.1817

15. Nadji M, Tabei SZ, Castro A, et al. Prostatic-specific antigen: An immunohistologic marker for prostatic neoplasms. *Cancer.* 1981;48(5):1229-1232. doi: 10.1002/1097-0142(19810901)48:5<1229::AID-CNCR2820480529>3.0.CO;2-L.

16. Catalona WJ, Smith DS, Ratliff TL, et al. Measurement of Prostate-Specific Antigen in Serum as a Screening Test for Prostate Cancer. 2010;324(17):1156-1161. doi:10.1056/NEJM199104253241702

17. Wolf AMD, Wender RC, Etzioni RB, et al. American Cancer Society Guideline for the Early Detection of Prostate Cancer: Update 2010. *CA Cancer J Clin*. 2010;60(2):70-98. doi:10.3322/CAAC.20066

18. Bell N, Connor Gorber S, Shane A, et al. Recommendations on screening for prostate cancer with the prostate-specific antigen test. *CMAJ*. 2014;186(16):1225-1234. doi:10.1503/CMAJ.140703/-/DC1

19. Fenton JJ, Weyrich MS, Durbin S, Liu Y, Bang H, Melnikow J. Prostate-Specific Antigen–Based Screening for Prostate Cancer: Evidence Report and Systematic Review for the US Preventive Services Task Force. *JAMA*. 2018;319(18):1914-1931. doi:10.1001/JAMA.2018.3712

20. Stephenson AJ, Kattan MW, Eastham JA, et al. Defining biochemical recurrence of prostate cancer after radical prostatectomy: a proposal for a standardized definition. *J Clin Oncol*. 2006;24(24):3973-3978. doi:10.1200/JCO.2005.04.0756

21. Matlaga BR, Eskew LA, McCullough DL. Prostate biopsy: Indications and technique. *J Urol*. 2003;169(1):12-19. doi:10.1016/S0022-5347(05)64024-4

22. Hodge KK, McNeal JE, Terris MK, Stamey TA. Random Systematic Versus Directed Ultrasound Guided Transrectal Core Biopsies of the Prostate. *J Urol*. 1989;142(1):71-74. doi:10.1016/S0022-5347(17)38664-0

23. Levine MA, Ittman M, Melamed J, Lepor H. Two Consecutive Sets of Transrectal Ultrasound Sextant Biopsies of the Prostate for the Detection of Prostate Cancer. *J Urol*. 1998;159(2):471-476. doi:10.1016/S0022-5347(01)63951-X

24. Presti JC, Chang JJ, Bhargava V, Shinohara K. The Optimal Systematic Prostate Biopsy Scheme Should Include 8 Rather than 6 Biopsies: Results of a Prospective Clinical Trial. *J Urol*. 2000;163(1):163-167. doi:10.1016/S0022-5347(05)67995-5

25. Babaian RJ, Toi A, Kamoi K, et al. A Comparative Analysis of Sextant and an Extended 11-Core Multisite Directed Biopsy Strategy. *J Urol*. 2000;163(1):152-

157. doi:10.1016/S0022-5347(05)67993-1

26. Norberg M, Egevad L, Holmberg L, Sparén P, Norlén BJ, Busch C. The sextant protocol for ultrasound-guided core biopsies of the prostate underestimates the presence of cancer. *Urology*. 1997;50(4):562-566. doi:10.1016/S0090-4295(97)00306-3

27. Streicher J, Meyerson BL, Karivedu V, Sidana A. A review of optimal prostate biopsy: indications and techniques. *Ther Adv Urol*. 2019;11:1756287219870074. doi:10.1177/1756287219870074

28. Gleason DF. Histologic grading of prostate cancer: A perspective. *Hum Pathol*. 1992;23(3):273-279. doi:10.1016/0046-8177(92)90108-F

29. Brierley J, Gospodarowicz MD, Wittekind CT. *TNM Classification of Malignant Tumors*. 8th ed. Wiley; 2017.

30. Humphrey PA. Gleason grading and prognostic factors in carcinoma of the prostate. *Mod Pathol.* 2004;17(3):292-306. doi:10.1038/modpathol.3800054

31. D'Amico A V., Whittington R, Bruce Malkowicz S, et al. Biochemical Outcome After Radical Prostatectomy, External Beam Radiation Therapy, or Interstitial Radiation Therapy for Clinically Localized Prostate Cancer. *JAMA*. 1998;280(11):969-974. doi:10.1001/JAMA.280.11.969

32. Carroll PH, Mohler JL. NCCN Guidelines Updates: Prostate Cancer and Prostate Cancer Early Detection. *J Natl Compr Cancer Netw*. 2018;16(5S):620-623. doi:10.6004/JNCCN.2018.0036

33. Loeb S, Bjurlin MA, Nicholson J, et al. Overdiagnosis and Overtreatment of Prostate Cancer. *Eur Urol*. 2014;65(6):1046. doi:10.1016/J.EURURO.2013.12.062

34. Dall'Era MA, Albertsen PC, Bangma C, et al. Active Surveillance for Prostate Cancer: A Systematic Review of the Literature. *Eur Urol*. 2012;62(6):976-983. doi:10.1016/J.EURURO.2012.05.072

35. Cooperberg MR, Carroll PR, Klotz L. Active surveillance for prostate cancer: Progress and promise. *J Clin Oncol*. 2011;29(27):3669-3676. doi:10.1200/JCO.2011.34.9738

36. Draisma G, Etzioni R, Tsodikov A, et al. Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context. *J Natl Cancer Inst*. 2009;101(6):374-383. doi:10.1093/JNCI/DJP001

37. Ficarra V, Novara G, Artibani W, et al. Retropubic, Laparoscopic, and Robot-Assisted Radical Prostatectomy: A Systematic Review and Cumulative Analysis of Comparative Studies. *Eur Urol*. 2009;55(5):1037-1063. doi:10.1016/J.EURURO.2009.01.036

38. Walsh PC, Donker PJ. Impotence Following Radical Prostatectomy: Insight Into Etiology and Prevention. *J Urol*. 1982;128(3):492-497. doi:10.1016/S0022-5347(17)53012-8

39. Guillonneau B, Cathelineau X, Barret E, Rozet F, Vallancien G. Laparoscopic Radical Prostatectomy: Technical and Early Oncological Assessment of 40 Operations. *Eur Urol*. 1999;36(1):14-20. doi:10.1159/000019921

40. Binder J, Kramer W. Robotically-assisted laparoscopic radical prostatectomy. *BJU Int*. 2001;87(4):408-410. doi:10.1046/J.1464-410X.2001.00115.X

41. Peschel RE, Colberg JW. Surgery, brachytherapy, and external-beam radiotherapy for early prostate cancer. *Lancet Oncol*. 2003;4(4):233-241. doi: 10.1016/s1470-2045(03)01035-0.

42. Adam M, Tennstedt P, Lanwehr D, et al. Functional Outcomes and Quality of Life After Radical Prostatectomy Only Versus a Combination of Prostatectomy with Radiation and Hormonal Therapy. *Eur Urol*. 2017;71(3):330-336. doi:10.1016/j.eururo.2016.11.015

43. Moghanaki D, Freedland SJ, Anscher M. Re: Comparison of mortality outcomes after radical prostatectomy versus radiotherapy in patients with localized prostate cancer: A population-based analysis. *Int J Urol*. 2013;20(5):547-548. doi:10.1111/j.1442-2042.2012.03212.x

44. Sharifi N, Gulley JL, Dahut WL. An update on androgen deprivation therapy for prostate cancer. *Endocr Relat Cancer*. 2010;17(4):R305-R315. doi:10.1677/ERC-10-0187

45. Gilligan T, Kantoff PW. Chemotherapy for prostate cancer. *Urology*. 2002;60(3):94-100. doi:10.1016/S0090-4295(02)01583-2

46. Sharifi N, Gulley JL, Dahut WL. Androgen Deprivation Therapy for Prostate Cancer. *JAMA*. 2005;294(2):238-244. doi:10.1001/JAMA.294.2.238

47. McLeod DG. Hormonal therapy: historical perspective to future directions. *Urology*. 2003;61(2):3-7. doi:10.1016/S0090-4295(02)02393-2

48. Di Lorenzo G, Buonerba C, Autorino R, De Placido S, Sternberg CN. Castration-Resistant Prostate Cancer: Current and Emerging Treatment Strategies. *Drugs* 2010;70(8):983-1000. doi:10.2165/10898600-000000000-00000

49. McBride W, Withers H, Schaue D. Biologic Basis of Radiation Therapy. In: *Perez & Brady's Principles and Practice of Radiation Oncology*. Halperin EC, Wazer DE, Perez CA, Brady LW, eds. 7th ed. Wolters Kluwer Health; 2018.

50. Cornforth MN, Bedford JS. A Quantitative Comparison of Potentially Lethal Damage Repair and the Rejoining of Interphase Chromosome Breaks in Low

Passage Normal Human Fibroblasts. *Radiat Res*. 1987;111(3):385-405. doi:10.2307/3576926

51. Withers HR. The Four R's of Radiotherapy. *Adv Radiat Biol.* 1975;5:241-271. doi:10.1016/B978-0-12-035405-4.50012-8

52. Fowler JF. The linear-quadratic formula and progress in fractionated radiotherapy. *Br J Radiol.* 1989;62(740):679-694. doi:10.1259/0007-1285-62-740-679

53. Fowler JF. The radiobiology of prostate cancer including new aspects of fractionated radiotherapy. *Acta Oncol*. 2005;44(3):265-276. doi:10.1080/02841860410002824

54. Soffen EM, Hanks GE, Hunt MA, Epstein BE. Conformal static field radiation therapy treatment of early prostate cancer versus non-conformal techniques: a reduction in acute morbidity. *Int J Radiat Oncol Biol Phys*. 1992;24(3):485-488. doi:10.1016/0360-3016(92)91063-S

55. Ling CC, Burman C, Chui CS, et al. Conformal radiation treatment of prostate cancer using inversely-planned intensity-modulated photon beams produced with dynamic multileaf collimation. *Int J Radiat Oncol Biol Phys*. 1996;35(4):721-730. doi:10.1016/0360-3016(96)00174-5

56. Olivera J, Penedo J, Marin J, et al. Localized Prostate Cancer: Volumetric Modulated Arc Therapy (VMAT) Versus Intensity Modulated Radiation Therapy (IMRT)–Which One is Better? *Int J Radiat Oncol Biol Phys*. 2014;90(1):S863. doi:10.1016/J.IJROBP.2014.05.2467

57. Martin NE, D'Amico A V. Progress and controversies: Radiation therapy for prostate cancer. *CA Cancer J Clin*. 2014;64(6):389-407. doi:10.3322/CAAC.21250

58. Dearnaley D, Syndikus I, Mossop H, et al. Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: 5-year outcomes of the randomised, non-inferiority, phase 3 CHHiP trial. *Lancet Oncol*. 2016;17(8):1047-1060. doi:10.1016/S1470-2045(16)30102-4

59. Incrocci L, Wortel RC, Alemayehu WG, et al. Hypofractionated versus conventionally fractionated radiotherapy for patients with localised prostate cancer (HYPRO): final efficacy results from a randomised, multicentre, open-label, phase 3 trial. *Lancet Oncol*. 2016;17(8):1061-1069. doi:10.1016/S1470-2045(16)30070-5

60. Catton CN, Lukka H, Gu CS, et al. Randomized Trial of a Hypofractionated Radiation Regimen for the Treatment of Localized Prostate Cancer. *J Clin Oncol*. 2017;35(17):1884-1890. doi:10.1200/JCO.2016.71.7397

61. Chen RC, Zhang Y, Chen MH, et al. Patient-reported quality of life during radiation treatment for localized prostate cancer: results from a prospective phase II trial. *BJU Int*. 2012;110(11):1690-1695. doi:10.1111/J.1464-410X.2012.11117.X

62. Bekelman JE, Rumble RB, Chen RC, et al. Clinically Localized Prostate Cancer: ASCO Clinical Practice Guideline Endorsement of an American Urological Association/American Society for Radiation Oncology/Society of Urologic Oncology Guideline. *J Clin Oncol*. 2018;36(32):3251-3258. doi:10.1200/JCO.18.00606

63. Gandaglia G, Briganti A, Clarke N, et al. Adjuvant and Salvage Radiotherapy after Radical Prostatectomy in Prostate Cancer Patients. *Eur Urol*. 2017;72(5):689-709. doi:10.1016/J.EURURO.2017.01.039

64. Widmark A, Gunnlaugsson A, Beckman L, et al. Ultra-hypofractionated versus conventionally fractionated radiotherapy for prostate cancer: 5-year outcomes of the HYPO-RT-PC randomised, non-inferiority, phase 3 trial. *Lancet*. 2019;394(10196):385-395. doi:10.1016/S0140-6736(19)31131-6

65. Georg D, Hopfgartner J, Gòra J, et al. Dosimetric considerations to determine the optimal technique for localized prostate cancer among external photon, proton, or carbon-ion therapy and high-dose-rate or low-dose-rate brachytherapy. *Int J Radiat Oncol Biol Phys*. 2014;88(3):715-722. doi:10.1016/j.ijrobp.2013.11.241

66. Yamada Y, Rogers L, Demanes DJ, et al. American Brachytherapy Society consensus guidelines for high-dose-rate prostate brachytherapy. *Brachytherapy*. 2012;11(1):20-32. doi:10.1016/J.BRACHY.2011.09.008

67. Yoshioka Y, Suzuki O, Isohashi F, et al. High-Dose-Rate Brachytherapy as Monotherapy for Intermediate- and High-Risk Prostate Cancer: Clinical Results for a Median 8-Year Follow-Up. *Int J Radiat Oncol*. 2016;94(4):675-682. doi:10.1016/J.IJROBP.2015.05.044

68. Pouliot J, Beaulieu L. Modern Principles of Brachytherapy Physics: From 2-D to 3-D to Dynamic Planning and Delivery. In: *Leibel and Philips Textbook of Radiation Oncology*. Philips TL, Hoppe R, Roach M, Leibel SA, eds. 3rd ed. Elsevier Inc.; 2010.

69. Martell K, Mendez LC, Chung HT, et al. Results of 15 Gy HDR-BT boost plus EBRT in intermediate-risk prostate cancer: Analysis of over 500 patients. *Radiother Oncol*. 2019;141:149-155. doi:10.1016/j.radonc.2019.08.017

70. Hauswald H, Kamrava MR, Fallon JM, et al. Clinical Investigation High-Dose-Rate Monotherapy for Localized Prostate Cancer: 10-Year Results Radiation Oncology. *Int J Radiat Oncol Biol Phys*. 2016;94(4):667-674. doi:10.1016/j.ijrobp.2015.07.2290

71. Morton G, McGuffin M, Chung HT, et al. Prostate high dose-rate brachytherapy as monotherapy for low and intermediate risk prostate cancer: Efficacy results from a randomized phase II clinical trial of one fraction of 19 Gy or two fractions of 13.5 Gy. *Radiother Oncol*. 2020;146:90-96. doi:10.1016/j.radonc.2020.02.009

72. Gunderson LL, Tepper JE. *Clinical Radiation Oncology*. 3rd ed. Elsevier Inc.; 2011.

73. Nag S, Beyer D, Friedland J, Grimm P, Nath R. American brachytherapy society (ABS) recommendations for transperineal permanent brachytherapy of prostate cancer. *Int J Radiat Oncol*. 1999;44(4):789-799. doi:10.1016/S0360-3016(99)00069-3

74. Morris WJ, Tyldesley S, Rodda S, et al. Androgen Suppression Combined with Elective Nodal and Dose Escalated Radiation Therapy (the ASCENDE-RT Trial): An Analysis of Survival Endpoints for a Randomized Trial Comparing a Low-Dose-Rate Brachytherapy Boost to a Dose-Escalated External Beam Boost for High- and Intermediate-risk Prostate Cancer. *Int J Radiat Oncol Biol Phys*. 2017;98(2):275-285. doi:10.1016/J.IJROBP.2016.11.026

75. Hoskin P, Rojas A, Ostler P, et al. High-dose-rate brachytherapy alone given as two or one fraction to patients for locally advanced prostate cancer: Acute toxicity. *Radiother Oncol*. 2014;110(2):268-271. doi:10.1016/j.radonc.2013.09.025

76. Pucar D, Hricak H, Shukla-Dave A, et al. Clinically Significant Prostate Cancer Local Recurrence After Radiation Therapy Occurs at the Site of Primary Tumor: Magnetic Resonance Imaging and Step-Section Pathology Evidence. *Int J Radiat Oncol Biol Phys*. 2007;69(1):62-69. doi:10.1016/j.ijrobp.2007.03.065

77. Arrayeh E, Westphalen AC, Kurhanewicz J, et al. Does local recurrence of prostate cancer after radiation therapy occur at the site of primary tumor? Results of a longitudinal MRI and MRSI study. *Int J Radiat Oncol Biol Phys*. 2012;82(5):e787-93. doi:10.1016/j.ijrobp.2011.11.030

78. Cellini N, Morganti AG, Mattiucci GC, et al. Analysis of intraprostatic failures in patients treated with hormonal therapy and radiotherapy: implications for conformal therapy planning. *Int J Radiat Oncol Biol Phys*. 2002;53(3):595-599. doi:10.1016/S0360-3016(02)02795-5

79. Chopra S, Toi A, Taback N, et al. Pathological Predictors for Site of Local Recurrence After Radiotherapy for Prostate Cancer. *Int J Radiat Oncol*. 2012;82(3):e441-e448. doi:10.1016/J.IJROBP.2011.05.035

80. Bauman G, Haider M, Van der Heide UA, Ménard C. Boosting imaging defined dominant prostatic tumors: A systematic review. *Radiother Oncol*. 2013;107(3):274-281. doi:10.1016/J.RADONC.2013.04.027

81. Von Eyben FE, Kiljunen T, Kangasmaki A, Kairemo K, Von Eyben R, Joensuu T. Radiotherapy Boost for the Dominant Intraprostatic Cancer Lesion - A Systematic Review and Meta-Analysis. *Clin Genitourin Cancer*. 2016;14(3):189-197. doi:10.1016/j.clgc.2015.12.005

82. Kasivisvanathan V, Emberton M, Ahmed HU. Focal therapy for prostate cancer: Rationale and treatment opportunities. *Clin Oncol*. 2013;25(8):461-473.

doi:10.1016/j.clon.2013.05.002

83. Turkbey B, Albert PS, Kurdziel K, Choyke PL. Imaging Localized Prostate Cancer: Current Approaches and New Developments. *AJR Am J Roentgenol*. 2009;192(6):1471. doi:10.2214/AJR.09.2527

84. Seco J, Evans PM. Assessing the effect of electron density in photon dose calculations. *Med Phys*. 2006;33(2):540-552. doi:10.1118/1.2161407

85. Martin T, Kolotas C, Dannenberg T, et al. New interstitial HDR brachytherapy technique for prostate cancer: CT based 3D planning after transrectal implantation. *Radiother Oncol*. 1999;52(3):257-260. doi:10.1016/S0167-8140(99)00113-9

86. Jaffray DA, Siewerdsen JH, Wong JW, Martinez AA. Flat-panel cone-beam computed tomography for image-guided radiation therapy. *Int J Radiat Oncol Biol Phys*. 2002;53(5):1337-1349. doi:10.1016/S0360-3016(02)02884-5

87. Nag S, Bice W, DeWyngaert K, Prestidge B, Stock R, Yu Y. The american brachytherapy society recommendations for permanent prostate brachytherapy postimplant dosimetric analysis. *Int J Radiat Oncol*. 2000;46(1):221-230. doi:10.1016/S0360-3016(99)00351-X

88. Koutrouvelis PG. Three-Dimensional Stereotactic Posterior Ischiorectal Space Computerized Tomography Guided Brachytherapy of Prostate Cancer; a Preliminary Report. *J Urol*. 1998;159(1):142-145. doi:10.1016/S0022-5347(01)64037-0

89. Holly R, Morton GC, Sankreacha R, et al. Use of cone-beam imaging to correct for catheter displacement in high dose-rate prostate brachytherapy. *Brachytherapy*. 2011;10(4):299-305. doi:10.1016/J.BRACHY.2010.11.007

90. Batchelar D, Gaztañaga M, Schmid M, Araujo C, Bachand F, Crook J. Validation study of ultrasound-based high-dose-rate prostate brachytherapy planning compared with CT-based planning. *Brachytherapy*. 2014;13(1):75-79. doi:10.1016/J.BRACHY.2013.08.004

91. Morton GC. Prostate high-dose-rate brachytherapy: Transrectal ultrasound based planning, a technical note. *Pract Radiat Oncol*. 2015;5(4):238-240. doi:10.1016/J.PRRO.2014.12.009

92. Sedelaar JPM, Vijverberg PLM, De Reijke TM, et al. Transrectal Ultrasound in the Diagnosis of Prostate Cancer: State of the Art and Perspectives. *Eur Urol*. 2001;40(3):275-284. doi:10.1159/000049787

93. R Rubin JM, Bude RO, Carson PL, Bree RL, Adler RS. Power Doppler US: a potentially useful alternative to mean frequency-based color Doppler US. *Radiology*. 1994;190(3):853-856. doi:10.1148/RADIOLOGY.190.3.8115639

94. Pallwein L, Mitterberger M, Pelzer A, et al. Ultrasound of prostate cancer: Recent advances. *Eur Radiol*. 2008;18(4):707-715. doi:10.1007/S00330-007-0779-7/FIGURES/4

95. Goossen TEB, De la Rosette JJMCH, Hulsbergen-van de Kaa CA, Van Leenders GJLH, Wijkstra H. The Value of Dynamic Contrast Enhanced Power Doppler Ultrasound Imaging in the Localization of Prostate Cancer. *Eur Urol*. 2003;43(2):124-131. doi:10.1016/S0302-2838(02)00582-1

96. Mitterberger MJ, Aigner F, Horninger W, et al. Comparative efficiency of contrast-enhanced colour Doppler ultrasound targeted versus systematic biopsy for prostate cancer detection. *Eur Radiol 2010 2012*. 2010;20(12):2791-2796. doi:10.1007/S00330-010-1860-1

97. Fenster A, Parraga G, Bax J. Three-dimensional ultrasound scanning. *Interface Focus*. 2011;1(4):503-519. doi:10.1098/rsfs.2011.0019

98. Tong S, Downey DB, Cardinal HN, Fenster A. A three-dimensional ultrasound prostate imaging system. *Ultrasound Med Biol*. 1996;22(6):735-746. doi:10.1016/0301-5629(96)00079-8

99. Bax J, Cool D, Gardi L, et al. Mechanically assisted 3D ultrasound guided prostate biopsy system. *Med Phys*. 2008;35(12):5397. doi:10.1118/1.3002415

100. Bax J, Smith D, Bartha L, et al. A compact mechatronic system for 3D ultrasound guided prostate interventions. *Med Phys*. 2011;38(2):1055-1069. doi:10.1118/1.3531540

101. Hrinivich WT, Hoover DA, Surry K, et al. Three-dimensional transrectal ultrasound guided high-dose-rate prostate brachytherapy: A comparison of needle segmentation accuracy with two-dimensional image guidance. *Brachytherapy*. 2016;15(2):231-239. doi:10.1016/j.brachy.2015.12.005

102. Siebert F-A, Hirt M, Niehoff P, Kovács G. Imaging of implant needles for real-time HDR-brachytherapy prostate treatment using biplane ultrasound transducers. *Med Phys*. 2009;36(8):3406-3412. doi:10.1118/1.3157107

103. Schmid M, Crook JM, Batchelar D, et al. A phantom study to assess accuracy of needle identification in real-time planning of ultrasound-guided high-dose-rate prostate implants. *Brachytherapy*. 2013;12:56-64. doi:10.1016/j.brachy.2012.03.002

104. Cool DW, Romagnoli C, Izawa JI, et al. Comparison of prostate MRI-3D transrectal ultrasound fusion biopsy for first-time and repeat biopsy patients with previous atypical small acinar proliferation. *Can Urol Assoc J*. 2016;10(9-10):342-348. doi:10.5489/cuaj.3831

105. Thompson J, Lawrentschuk N, Frydenberg M, Thompson L, Stricker P. The role of

magnetic resonance imaging in the diagnosis and management of prostate cancer. *BJU Int*. 2013;112(SUPPL. 2):6-20. doi:10.1111/BJU.12381

106. Bauman G, Haider M, Van Der Heide UA, Ménard C. Boosting imaging defined dominant prostatic tumors: A systematic review. *Radiother Oncol*. 2013;107(3):274-281. doi:10.1016/j.radonc.2013.04.027

107. Bouchelouche K, Turkbey B, Choyke P, Capala J. Imaging prostate cancer: An update on positron emission tomography and magnetic resonance imaging. *Curr Urol Rep*. 2010;11(3):180-190. doi:10.1007/S11934-010-0105-9/FIGURES/5

108. Weinreb JC, Barentsz JO, Choyke PL, et al. PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2. *Eur Urol*. 2016;69(1):16-40. doi:10.1016/J.EURURO.2015.08.052

109. Sciarra A, Barentsz J, Bjartell A, et al. Advances in Magnetic Resonance Imaging: How They Are Changing the Management of Prostate Cancer. *Eur Urol*. 2011;59(6):962-977. doi:10.1016/J.EURURO.2011.02.034

110. Ménard C, Susil RC, Choyke P, et al. MRI-guided HDR prostate brachytherapy in standard 1.5T scanner. *Int J Radiat Oncol*. 2004;59(5):1414-1423. doi:10.1016/J.IJROBP.2004.01.016

111. Murgic J, Chung P, Berlin A, et al. Lessons learned using an MRI-only workflow during high-dose-rate brachytherapy for prostate cancer. *Brachytherapy*. 2016;15(2):147-155. doi:10.1016/J.BRACHY.2015.12.004

112. Buus S, Rylander S, Hokland S, et al. Learning curve of MRI-based planning for high-dose-rate brachytherapy for prostate cancer. *Brachytherapy*. 2016;15(4):426-434. doi:10.1016/J.BRACHY.2016.03.011

113. Effert PJ, Bares R, Handt S, Wolff JM, Büll U, Jakse G. Metabolic Imaging of Untreated Prostate Cancer by Positron Emission Tomography with sup 18 Fluorine-Labeled Deoxyglucose. *J Urol*. 1996;155(3):994-998. doi:10.1016/S0022-5347(01)66366-3

114. Shreve PD, Grossman HB, Gross MD, Wahl RL. Metastatic prostate cancer: initial findings of PET with 2-deoxy-2-[F-18]fluoro-D-glucose. *Radiology*. 1996;199(3):751-756. doi:10.1148/RADIOLOGY.199.3.8638000

115. Maurer T, Eiber M, Schwaiger M, Gschwend JE. Current use of PSMA–PET in prostate cancer management. *Nat Rev Urol 2016 134*. 2016;13(4):226-235. doi:10.1038/nrurol.2016.26

116. Eiber M, Weirich G, Holzapfel K, et al. Simultaneous 68 Ga-PSMA HBED-CC PET/MRI Improves the Localization of Primary Prostate Cancer. *Eur Urol*. 2016;70(5):829-836. doi:10.1016/J.EURURO.2015.12.053

117. Smith CW, Alfano R, Hoover D, et al. Prostate specific membrane antigen positron emission tomography for lesion-directed high-dose-rate brachytherapy dose escalation. *Phys Imaging Radiat Oncol*. 2021;19:102-107. doi:10.1016/J.PHRO.2021.07.001

118. Choi YJ, Kim JK, Kim HJ, Cho K-S. Interobserver Variability of Transrectal Ultrasound for Prostate Volume Measurement According to Volume and Observer Experience. *Am J Roentgenol*. 2009;192(2):444-449. doi:10.2214/AJR.07.3617

119. Murciano-Goroff YR, Wolfsberger LD, Parekh A, et al. Variability in MRI vs. ultrasound measures of prostate volume and its impact on treatment recommendations for favorable-risk prostate cancer patients: A case series. *Radiat Oncol*. 2014;9(1):200. doi:10.1186/1748-717X-9-200

120. Mahdavi SS, Spadinger I, Chng N, Salcudean SE, Morris WJ. Semiautomatic segmentation for prostate brachytherapy: Dosimetric evaluation. *Brachytherapy*. 2013;12(1):65-76. doi:10.1016/j.brachy.2011.07.007

121. Narayanan R, Kurhanewicz J, Shinohara K, Crawford ED, Simoneau A, Suri JS. MRI-ultrasound registration for targeted prostate biopsy. In: *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. 2009:991-994. doi:10.1109/ISBI.2009.5193221

122. Karnik V V, Fenster A, Bax J, et al. Assessment of image registration accuracy in three-dimensional transrectal ultrasound guided prostate biopsy. *Med Phys*. 2010;37(2):802-813. doi:10.1118/1.3298010

123. Reich DL, Hossain S, Krol M, et al. Predictors of Hypotension After Induction of General Anesthesia. *Anesth Analg*. 2005;101(3):622-628. doi:10.1213/01.ANE.0000175214.38450.91

124. Tong S, Cardinal HN, McLoughlin RF, Downey DB, Fenster A. Intra- and inter-observer variability and reliability of prostate volume measurement via two-dimensional and three-dimensional ultrasound imaging. *Ultrasound Med Biol*. 1998;24(5):673-681. doi:10.1016/S0301-5629(98)00039-8

125. Qiu W, Yuan J, Ukwatta E, Sun Y, Rajchl M, Fenster A. Prostate Segmentation: An Efficient Convex Optimization Approach With Axial Symmetry Using 3-D TRUS and MR Images. *IEEE Trans Med Imaging*. 2014;33(4):947-960. doi:10.1109/TMI.2014.2300694

126. Qiu W, Rajchl M, Guo F, et al. 3D prostate TRUS segmentation using globally optimized volume-preserving prior. In: *Lect Notes in Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2014;8673:796-803. doi:10.1007/978-3-319-10404-1_99

127. Yuan J, Qiu W, Rajchl M, Ukwatta E, Tai X-C, Fenster A. Efficient 3D Endfiring TRUS Prostate Segmentation with Globally Optimized Rotational Symmetry. In:

2013 IEEE Conference on Computer Vision and Pattern Recognition. 2013:2211-2218. doi:10.1109/CVPR.2013.287

128. Qiu W, Yuan J, Ukwatta E, Fenster A. Rotationally resliced 3D prostate TRUS segmentation using convex optimization with shape priors. *Med Phys*. 2015;42(2):877-891. doi:10.1118/1.4906129

129. Anas EMA, Mousavi P, Abolmaesumi P. A deep learning approach for real time prostate segmentation in freehand ultrasound guided biopsy. *Med Image Anal*. 2018;48:107-116. doi:10.1016/j.media.2018.05.010

130. Ghavami N, Hu Y, Bonmati E, et al. Integration of spatial information in convolutional neural networks for automatic segmentation of intraoperative transrectal ultrasound images. *J Med Imaging*. 2018;6(1):011003. doi:10.1117/1.jmi.6.1.011003

131. Lei Y, Tian S, He X, et al. Ultrasound prostate segmentation based on multidirectional deeply supervised V-Net. *Med Phys*. 2019;46(7):3194-3206. doi:10.1002/mp.13577

132. Dwork C, Feldman V, Hardt M, Pitassi T, Reingold O, Roth A. The reusable holdout: Preserving validity in adaptive data analysis. *Science*. 2015;349(6248):636-638. doi:10.1126/science.aaa9375

133. Valdes G, Interian Y. Comment on 'Deep convolutional neural network with transfer learning for rectum toxicity prediction in cervical cancer radiotherapy: a feasibility study.' *Phys Med Biol*. 2018;63(6):068001. doi:10.1088/1361-6560/aaae23

134. Wang Y, Ni D, Dou H, et al. Deep Attentive Features for Prostate Segmentation in 3D Transrectal Ultrasound. *IEEE Trans Med Imaging*. 2019;38(12):2768-2778. doi:10.1109/TMI.2019.2913184

135. Lei Y, Wang T, Roper J, et al. Male pelvic multi-organ segmentation on transrectal ultrasound using anchor-free mask CNN. *Med Phys*. 2021;48(6):3055-3064. doi:10.1002/mp.14895

136. Tiong A, Bydder S, Ebert M, et al. A Small Tolerance for Catheter Displacement in High–Dose Rate Prostate Brachytherapy is Necessary and Feasible. *Int J Radiat Oncol*. 2010;76(4):1066-1072. doi:10.1016/J.IJROBP.2009.03.052

137. Mason J, Al-Qaisieh B, Bownes P, Thwaites D, Henry A. Dosimetry modeling for focal high-dose-rate prostate brachytherapy. *Brachytherapy*. 2014;13(6):611-617. doi:10.1016/j.brachy.2014.06.007

138. Hamper UM, Savader BL, Sheth S. Improved needle-tip visualization by color Doppler sonography. *Am J Roentgenol*. 1991;156(2):401-402. doi:10.2214/ajr.156.2.1898823

139. Ding M, Fenster A. A real-time biopsy needle segmentation technique using Hough Transform. *Med Phys*. 2003;30(8):2222-2233. doi:10.1118/1.1591192

140. Qiu W, Zhou H, Ding M, Zhang S. New real time needle segmentation technique using grayscale Hough transformation. *Proc. SPIE 6789, Medical Imaging 2007: Parallel Processing of Images and Optimization Techniques.* 2007;67890Q. doi:10.1117/12.749277

141. Hrinivich WT, Hoover DA, Surry K, et al. Simultaneous automatic segmentation of multiple needles using 3D ultrasound for high-dose-rate prostate brachytherapy. *Med Phys.* 2017;44(4):1234-1245. doi: 10.1002/mp.12148

142. Rodgers JR, Hrinivich WT, Surry K, Velker V, D'Souza D, Fenster A. A semiautomatic segmentation method for interstitial needles in intraoperative 3D transvaginal ultrasound images for high-dose-rate gynecologic brachytherapy of vaginal tumors. *Brachytherapy*. 2020;19(5):659-668. doi:10.1016/j.brachy.2020.05.006

143. Gillies DJ, Rodgers JR, Gyacskov I, et al. Deep Learning Segmentation of General Interventional Tools in Two-dimensional Ultrasound Images. *Med Phys*. 2020;47(10):4956-4970. doi:10.1002/mp.14427

144. Pourtaherian A, Farhad ·, Zanjani G, et al. Robust and semantic needle detection in 3D ultrasound using orthogonal-plane convolutional neural networks. *Int J Comput Assist Radiol Surg*. 2018;13:1321-1333. doi:10.1007/s11548-018-1798-3

145. Feld R, Needleman L, Goldberg BB. Use of a needle-vibrating device and color Doppler imaging for sonographically guided invasive procedures. *Am J Roentgenol*. 1997;168(1):255-256. doi:10.2214/ajr.168.1.8976955

146. Jones CD, McGahan JP, Clark KJ. Color Doppler ultrasonographic detection of a vibrating needle system. *J Ultrasound Med*. 1997;16(4):269-274. doi:10.7863/jum.1997.16.4.269

147. Armstrong G, Cardon L, Vilkomerson D, et al. Localization of needle tip with color Doppler during pericardiocentesis: In vitro validation and initial clinical application. *J Am Soc Echocardiogr*. 2001;14(1):29-37. doi:10.1067/mje.2001.106680

148. Fronheiser MP, Wolf PD, Idriss SF, Nelson RC, Lee W, Smith SW. Real-time 3D color flow doppler for guidance of vibrating interventional devices. *Ultrason Imaging*. 2004;26(3):173-184. doi:10.1177/016173460402600304

149. Harmat A, Rohling RN, Salcudean SE. Needle tip localization using stylet vibration. *Ultrasound Med Biol*. 2006;32(9):1339-1348. doi:10.1016/j.ultrasmedbio.2006.05.019

150. Klein SM, Fronheiser MP, Reach J, Nielsen KC, Smith SW. Piezoelectric vibrating needle and catheter for enhancing ultrasound-guided peripheral nerve blocks. *Anesth*

Analg. 2007;105(6):1858-1860. doi:10.1213/01.ane.0000286814.79988.0a

151. Fronheiser MP, Idriss SF, Wolf PD, Smith SW. Vibrating interventional device detection using real-time 3-D color Doppler. *IEEE Trans Ultrason Ferroelectr Freq Control*. 2008;55(6):1355-1362. doi:10.1109/TUFFC.2008.798

152. Reddy KE, Light ED, Rivera DJ, Kisslo JA, Smith SW. Color Doppler imaging of cardiac catheters using vibrating motors. *Ultrason Imaging*. 2008;30(4):247-250. doi:10.1177/016173460803000408

153. Adebar TK, Fletcher AE, Okamura AM. 3-D ultrasound-guided robotic needle steering in biological tissue. *IEEE Trans Biomed Eng*. 2014;61(12):2899-2910. doi:10.1109/TBME.2014.2334309

154. Greer JD, Adebar TK, Hwang GL, Okamura AM. Real-Time 3D Curved Needle Segmentation Using Combined B-Mode and Power Doppler Ultrasound. *Med Image Comput Comput Assist Interv*. 2014;17(2):381-388. doi:10.1007/978-3-319-10470-6_48

155. Cabreros SS, Jimenez NM, Greer JD, Adebar TK, Okamura AM. Remote electromagnetic vibration of steerable needles for imaging in power Doppler ultrasound. *Robot Autom (ICRA), 2015 IEEE Int Conf*. 2015:2244-2249. doi:10.1109/ICRA.2015.7139496

156. Kuang Y, Hilgers A, Sadiq M, Cochran S, Corner G, Huang Z. Modelling and characterisation of a ultrasound-actuated needle for improved visibility in ultrasound-guided regional anaesthesia and tissue biopsy. *Ultrasonics*. 2016;69:38-46. doi:10.1016/J.ULTRAS.2016.02.018

157. Sun Y, Qiu W, Romagnoli C, Fenster A. 3D non-rigid surface-based MR-TRUS registration for image-guided prostate biopsy. *Proc. SPIE 9036, Medical Imaging 2014: Image-Guided Procedures, Robotic Interventions, and Modeling*. 2014;90362J. doi:10.1117/12.2043662

158. Narayanan R, Kurhanewicz J, Shinohara K, Crawford ED, Simoneau A, Suri JS. Mri-ultrasound registration for targeted prostate biopsy. *Proc - 2009 IEEE Int Symp Biomed Imaging From Nano to Macro, ISBI 2009*. 2009:991-994.

159. Hu Y, Ahmed HU, Taylor Z, et al. MR to ultrasound registration for image-guided prostate interventions. *Med Image Anal*. 2012;16(3):687-703. doi:10.1016/J.MEDIA.2010.11.003

160. Baco E, Ukimura O, Rud E, et al. Magnetic Resonance Imaging–Transectal Ultrasound Image-fusion Biopsies Accurately Characterize the Index Tumor: Correlation with Step-sectioned Radical Prostatectomy Specimens in 135 Patients. *Eur Urol*. 2015;67(4):787-794. doi:10.1016/J.EURURO.2014.08.077

161. Meng X, Rosenkrantz AB, Mendhiratta N, et al. Relationship Between Prebiopsy

Multiparametric Magnetic Resonance Imaging (MRI), Biopsy Indication, and MRI-ultrasound Fusion–targeted Prostate Biopsy Outcomes. *Eur Urol*. 2016;69(3):512-517. doi:10.1016/J.EURURO.2015.06.005

162. Ukimura O, Marien A, Palmer S, et al. Trans-rectal ultrasound visibility of prostate lesions identified by magnetic resonance imaging increases accuracy of image-fusion targeted biopsies. *World J Urol*. 2015;33(11):1669-1676. doi:10.1007/s00345-015-1501-z

163. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2015;9351:234-241. doi:10.1007/978-3-319-24574-4_28

164. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2018;11045:3-11. http://arxiv.org/abs/1807.10165

165. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation. *IEEE Trans Med Imaging*. 2020;39(6):1856-1867. doi:10.1109/TMI.2019.2959609

Chapter 2

2 Automatic prostate segmentation using deep learning on clinically diverse 3D transrectal ultrasound images

Manual prostate segmentation in 3D TRUS images is a time-consuming and difficult task that must be completed in the operating room. An accurate automatic segmentation algorithm may reduce procedure and increase workflow efficiency in minimally invasive prostate cancer procedures. The purpose of Chapter 2 is to describe the development and validation of a deep learning-based automatic prostate segmentation algorithm for 3D TRUS images.

The contents of this chapter have been previously published in *Medical Physics*: Orlando N and Gillies DJ, Gyacskov I, Romagnoli C, D'Souza D, and Fenster A. *Medical Physics*. 2020;47(6):2413-2426. Permission to reproduce this article was granted by John Wiley and Sons and is provided in Appendix A – Copyright Releases.

2.1 Introduction

Diagnosing and treating prostate cancer continues to burden global populations as it is the second most common noncutaneous cancer among men worldwide.[1] Investigation into methods to diagnose and treat prostate cancer has shifted towards improved needle-based approaches that utilize three-dimensional (3D) information intraoperatively. Magnetic resonance imaging (MRI)-3D transrectal ultrasound (TRUS) guided biopsy is one diagnostic method that fuses the superior soft-tissue contrast of MRI to accurately localize, target, and sample suspicious tissue regions for prostate cancer with the real-time, low-cost, and portable capabilities of ultrasound. In the treatment of prostate cancer, high- and low-dose-rate (H/LDR) brachytherapy (BT) procedures offer therapeutic benefits for patients by exploiting radiobiological effects and offering shorter treatment times, while minimizing adverse side effects. By adding 3D ultrasound intraoperatively, improved spatial context and targeting

can be achieved to further improve the conventional clinical workflow of biopsy and brachytherapy procedures; however, both procedures rely on accurate segmentation of the prostate in 3D TRUS images to perform necessary clinical tasks. This includes surface-based registration approaches with MRI[2] and subsequent 3D TRUS images,[3] glandular volume measurements,[4,5] and dose-volume calculations,[6] especially when using a commercial guidance system. These segmentations are predominantly performed manually during the procedure, which is time-consuming, variable, and often difficult, which can lead to increased patient risk due to increased anesthesia exposure.[7]

Minimizing procedure time through fully or semi-automatic 3D TRUS prostate segmentation has been previously investigated.[8–11] Many methods have been shown to be promising, but have lacked clinical translation due to computational complexity, computation time, and robustness to diverse clinical datasets. Convolutional neural networks (CNNs) have received widespread attention in many image processing applications with much work investigating their accuracy and speed in medical imaging tasks. Prostate segmentation in 3D TRUS is an image processing task that could be an ideal candidate for the data-driven predictions provided by CNNs, although most existing work has investigated their application in MRI,[12,13] as ultrasound is considered more challenging due to noise and image artifacts.

Recent work has investigated and provided promising results for the use of CNNs in 2D TRUS[14] and 3D TRUS[15,16] prostate segmentation. Ghavami *et al*.,[15] evaluated the performance of an adapted U-Net on 109 side-fire sagittally-reconstructed (SR) 3D TRUS images. Predictions were performed on acquired 2D images and varying adjacent neighboring slices were also investigated to evaluate accuracy due to increasing spatial 3D context. The best results reported for the 2D Dice similarity coefficient (DSC), 3D DSC, and boundary distance were 89 ± 12 %, 89 ± 5 %, and 1.68 ± 1.57 mm, respectively. Lei *et al*.,[16] investigated a multidirectional deeply supervised 3D V-Net with contour refinement on 44 patient 3D TRUS images. Their method was shown to improve performance when segmenting the apex and base of the prostate, which is often difficult due to low image contrast, and reported overall segmentation results for a 3D DSC, Hausdorff distance (HD), mean surface distance (MSD), and residual mean surface distance (RMSD) of 92 ± 3 %, 3.94 ± 1.55 mm, 0.60 ± 0.23 mm, and 0.90 ± 0.38 mm,

respectively. However, these methods were tested using cross-validation approaches on 3D TRUS images from a single ultrasound machine with matched voxel dimensions and sizes, so further testing is still required on an unseen and variable dataset to provide a complete understanding of performance while avoiding potential limitations due to information bleeding.[17,18] Furthermore, investigations into generalizability across procedures and acquisition geometries have been limited, to our knowledge, which could restrict usability when applying these techniques across applications.

Our work aims to demonstrate that a diverse image dataset can train a supervised CNN to provide an accurate, fast, automated, and generalizable 3D prostate segmentation prediction. We used 206 3D TRUS patient images from two different procedures and acquisition geometries, two facilities, and four transducers used with three different ultrasound machine models to modify and train a deep learning-based 2D segmentation method followed by reconstruction into a 3D surface. Since deep learning approaches often improve in performance when using large datasets (*i.e.*, >1000 images), we chose to reslice each 3D image to increase the amount of usable data for prostate segmentation training and prediction. Testing was performed on 40 unseen 3D TRUS patient images and segmentation performance was compared to state-of-the-art fully 3D approach's for assessing the impact of reducing spatial context. Various metrics are reported in the literature, and typically vary in choice and quantity between studies, but many metrics are required to obtain a complete understanding of segmentation performance and to allow for comparison with previous studies. By using a clinically diverse dataset with variable image representation and image quality of the prostate, we intend to provide a thorough analysis of performance for a broader scope of comparison. Once this method is evaluated, completion of required intraoperative image-guidance tasks can be facilitated for different needle-based prostate cancer procedures and potentially decrease overall clinical procedure times and anesthesia risks to patients.

2.2 Materials and Methods

2.2.1 Clinical dataset

3D images of the prostate were acquired using end-fire (as used in prostate biopsy) and side-fire SR (as used in some HDR-BT) mechanical scanning approaches (Fig. 2.1).[19] Both methods rotate a TRUS transducer around the long-axis to create geometrically different reconstructed 3D images that are influenced by the transducer array configuration. The images used in this study were acquired with the C9-5 transducer with the iU22 (Philips, Amsterdam, the Netherlands), the C9-5 and BPTRT9-5 transducers with the ATL HDI-5000 (Philips, Amsterdam, the Netherlands), and the 8848 transducer with the Profocus 2202 (BK Medical, Peabody, MA, United States) ultrasound machine models. The total dataset of 246 3D TRUS images consisted of 104 end-fire and 142 side-fire 3D TRUS images and was split into training, validation, and testing datasets as shown in Table 2.1. Manual 3D prostate segmentations (excluding the seminal vesicles) were performed by an observer (IG) with approximately 15 years of TRUS prostate image analysis experience. 3D image sizes ranged from [300×400×784] to [408×441×870] voxels with dimensions of [0.094×0.154×0.154] to [0.183×0.186×0.186] mm^3/voxel for side-fire images and from [448×350×448] to [692×520×692] voxels with dimensions of [0.115×0.115×0.115] to [0.190×0.195×0.190] mm^3/voxel for end-fire images.

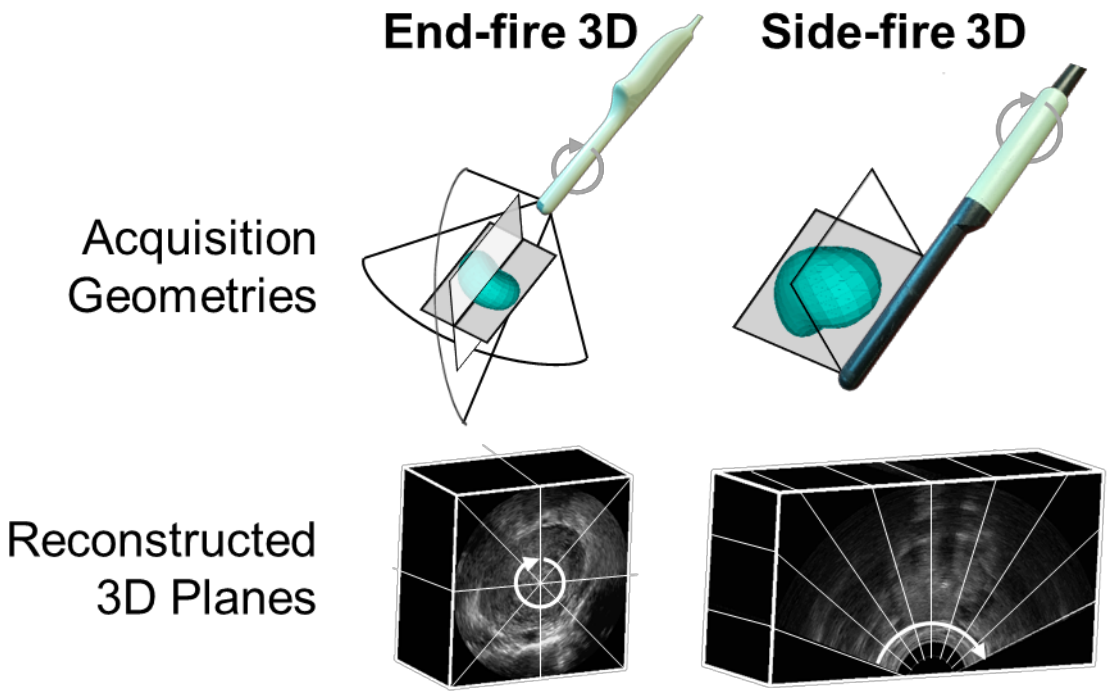

Figure 2.1. Mechanical scanning approaches for acquiring 3D TRUS prostate images using end-fire (left) and side-fire (right) TRUS transducers. 2D images are acquired by rotating around the long axis of the transducer at known sample spacings to create 3D TRUS images. Example 3D TRUS images are shown in the bottom row, with the front face demonstrating the reconstructed image plane and the white lines showing representative acquisition planes.

Table 2.1. Clinical 3D TRUS dataset split based on end-fire and side-fire scan geometries and resulting training, validation, and testing datasets used for deep learning.

Image	Training	Validation	Testing	Total
End-fire	67	17	20	104
Side-fire	98	24	20	142
Total	165	41	40	246

2.2.2 3D segmentation algorithm

A workflow diagram of our proposed method is shown in Fig. 2.2. This includes 3D TRUS prostate image input, radial sampling, prediction using a trained modified U-Net, 3D reconstruction, and 3D prostate surface output.

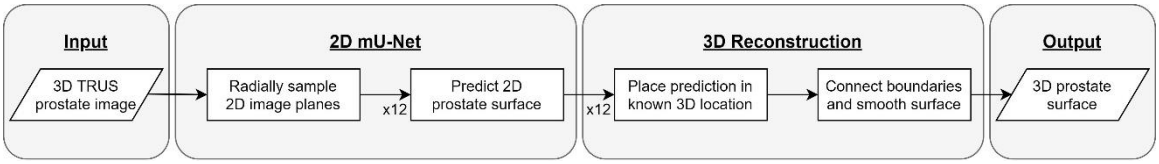

Figure 2.2. Proposed 3D prostate segmentation workflow. A 3D TRUS prostate image is used as input, followed by radial sampling to generate 12 2D image planes. Each image plane was used to predict a prostate boundary with a trained modified U-Net prior to reconstruction into the 3D prostate surface.

2.2.2.1 Training dataset for modified U-Net

Images from the training and validation split were used to obtain resliced 2D images of the prostate. These 2D images were obtained at randomized axial, sagittal, coronal, radial, and oblique image planes with varying rotations and zooms. This resulted in a dataset of 6,773 2D TRUS images with matched manual contours. All 2D images were resized to 256×256 pixels with no preprocessing (*i.e.*, despeckling or bias correction) and were separated into an 80/20 training/validation split for deep learning, resulting in 5418 training and 1355 validation 2D TRUS images.

2.2.2.2 Modified U-Net

The previously published U-Net[20] was implemented using Keras[21] with TensorFlow[22] and modified by adding 50% dropouts at every block on the expansion section of the network to increase regularization and prevent overfitting. In addition, transpose convolutions were used at each step in the expansion section instead of the standard upsampling followed by convolution, as this allowed for improved performance

in preliminary experiments. Data augmentation from random combinations of horizontal flips, 2D shifts (up to 20%), rotations (up to 20°), and zooms (up to 20%) were employed to double the training dataset to 10,836 2D TRUS images. Preliminary experiments led to the selection of an Adam optimizer, 0.0001 learning rate, Dice-coefficient loss function, 200 epochs, and 200 steps per epoch. This network was trained and used for predicting unseen data on a personal computer with two Xeon E5645 central processing units at 2.40 GHz (Intel Corporation, Santa Clara, CA, USA), 24.0 GB of memory, and a 6 GB Ge-Force GTX TITAN (NVIDIA Corporation, Santa Clara, CA, USA) graphics processing unit (GPU).

2.2.2.3 3D reconstruction

Predicted 3D prostate segmentations were obtained by segmenting multiple 2D radial frames generated by rotation around a central axis, followed by reconstruction to a 3D surface following a reconstruction method similar to Qiu *et al*.[11] Previous observations have noted that segmenting the prostate on slices near the apex and base of the prostate can be challenging due to boundary incompleteness,[15] so we chose to radially slice the 3D prostate image as opposed to transverse slicing in an attempt to improve segmentations at all boundaries. This choice was motivated by the experience of segmenting the prostate when the center of the gland is in-plane, which typically presents as an easier image to accurately define and segment the boundaries on the left and right sides of the 2D image. In contrast to this, a transverse slicing approach would result in 2D images with the prostate appearing as a different size and shape, with this difference more pronounced at the prostate apex and base, and when comparing end-fire and side-fire image geometries. Difficulty arises, predominantly in side-fire geometries, when segmenting the ends of the prostate along the axis of acquisition (when using a 2D approach) due to the changes in prostate appearance and size. Thus, radially slicing and segmenting the prostate in these views allows for a method that can train and predict on images containing similar structural shapes, across different acquisition geometries.

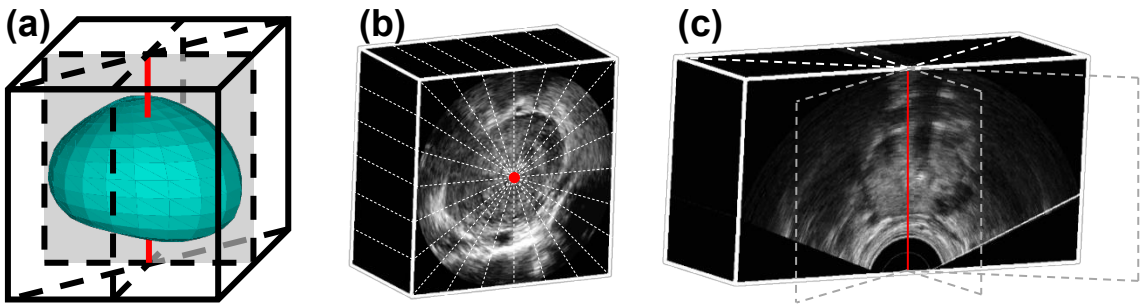

Figure 2.3. (a) Method for acquiring radially sliced 2D TRUS image planes (dotted lines) from previously acquired 3D TRUS images. The axis of rotation (red) is approximately at the central axis of the prostate and in the anterior-posterior direction. (b) 12 image slices were obtained to perform prostate segmentation. For an end-fire 3D TRUS image, 2D images are about the axis of image acquisition. (c) The majority of image slices obtained from side-fire 3D TRUS images are across many acquisition slices and are subject to increased interpolation, except for one slice, which matched the original acquired 2D TRUS image.

Reconstructing a 3D contour was accomplished by radially slicing a 3D image in equal 15° spacings around the approximate central axis of the prostate (Fig. 2.3) to produce 12 2D images for prediction. These 12 images were predicted using the trained modified U-Net to produce 2D prostate segmentations, and 204 equally spaced points were sampled around the boundary of each 2D image. Since the original spatial location of the input 2D image in the 3D volume was known, each predicted 2D segmentation was placed appropriately back into the 3D volume and the boundary points on each segmentation were connected to the adjacent slices. The intermediate surface was smoothed by a windowed *sinc* filter, resulting in a final reconstructed 3D contour.

2.2.3 Evaluation and comparison

Our algorithm was evaluated on a test data set of 20 unseen end-fire and 20 unseen side-fire 3D TRUS images of the prostate. Standard pixel map comparisons (DSC, recall, precision) were computed for both 2D radial slice segmentations and the reconstructed 3D segmentation for each prostate to obtain an understanding of the prediction quality and reconstruction accuracy. We also computed absolute area/volume percent differences

(A/VPD), absolute mean surface distances (MSD), and absolute Hausdorff distances (HD), as well as signed A/VPD (sA/VPD), signed MSD (sMSD), and signed HD (sHD) for all 2D predictions and reconstructed 3D segmentations. The signed metrics, while not commonly reported, are important as they provide information on the segmentation bias and whether the prostate boundary is over or underpredicted. The inclusion of these metrics will provide a more complete understanding of the performance of our proposed method. All pixel map comparisons and 2D MSD/sMSD and HD/sHD distance metrics were computed using MATLAB R2019a (MathWorks, Natick, MA, United States). The 3D MSD/sMSD and HD/sHD metrics were computed by measuring the distances between all points of the automatically predicted segmentation to the closest point on the manual gold-standard segmentation (CloudCompare v2.10.2).[23] For comparison purposes, the MSD reported here is similar to the boundary distance[15] and mean surface distance[16] presented previously. Computation times were recorded for 2D slice segmentation, 3D reconstruction, and overall 3D segmentation time.

The performance of our algorithm was compared against three state-of-the-art fully 3D predicting CNNs (V-Net,[24] Dense V-Net,[25] and High-resolution 3D-Net[26]) using an open-source implementation on the NiftyNet platform.[27] It is often assumed that performing a prediction based on 3D information allows for an improved result due to increased spatial context, so we completed a direct comparison on the same test dataset to investigate this hypothesis. Similar to our proposed method, the same 165/41 3D TRUS images (Table 2.1) were used for training/validation, respectively. The 3D V-Net was chosen to optimize hyperparameters, including loss function, due to its widespread use and performance in preliminary experiments. For simplicity, these hyperparameters were also used for the Dense V-Net and High-resolution 3D-Net. Parameters were chosen to maximize the spatial window size and usable memory on the GPU with optimized hyperparameters shown in Table 2.2. Previous work has shown improved performance with a hybrid loss function,[16] so we compared performance between a Dice loss function and a Dice plus cross-entropy (DiceXEnt) loss function, as provided in NiftyNet, using the 3D V-Net. Although NiftyNet offers a patch-based analysis, preliminary experiments resulted in 3D segmentations with many flat surfaces throughout the prediction corresponding to patch edges. Since we had one structure of interest (*i.e.*, the prostate), we

did not perform a patch-based analysis and predictions were performed on a resized image to match the spatial window. Data augmentation was employed to double the training dataset to 330 3D TRUS images. The chosen hyperparameters for the Dense V-Net and High-resolution 3D-Net were the same as shown in Table 2.2 (with the DiceXEnt loss function). These networks were trained and used for predicting unseen data on a personal computer with an Intel Core i7-4770 central processing unit at 3.40 GHz (Intel Corporation, Santa Clara, CA, USA), 32.0 GB of memory, and a 6 GB Ge-Force GTX TITAN (NVIDIA Corporation, Santa Clara, CA, USA) graphics processing unit (GPU). Training and 3D segmentation computation times were recorded.

Table 2.2. Hyperparameter selection when employing the V-Net in NiftyNet.

Hyperparameter	Value
Optimizer	Adam
Loss function	Dice and Dice + cross-entropy
Activation function	PReLU
Learning rate	0.0001
Spatial window size	[64, 64, 64]
Mini-batch size	2
Weighted L2-decay	0.0001

2.2.4 Statistical analysis

Statistical calculations were performed in GraphPad Prism 8.3 (Graphpad Software, Inc., San Diego, CA, USA). The normality of distributions was evaluated using the Shapiro-Wilk test and led to the use of nonparametric statistical tests when the assumption was violated. The corresponding nonparametric alternative tests are presented in parentheses for the remainder of the section. The significance level for statistical analysis was chosen such that the probability of making a type I error was less than 5% ($p < 0.05$), with statistically significant differences denoted simply as significant for the remainder of this manuscript.

2D radial slice segmentation and 3D reconstructed segmentation accuracy as well as Dice and DiceXEnt 3D V-Net loss functions were compared using two-tailed paired t-tests (Wilcoxon matched-pairs signed-rank tests). Comparisons between our proposed

algorithm and three fully 3D CNNs were performed using two-tailed paired t-tests (Wilcoxon matched-pairs signed-rank tests) with a Bonferroni multiple-comparison correction, which adjusted the significance level to $p < 0.0167$. Comparisons between segmentation accuracy for end-fire and side-fire 3D TRUS images on each network were completed using two-tailed unpaired t-tests (Mann-Whitney tests).

2.3 Results

2.3.1 Reconstructed modified U-Net

The results of our modified U-Net for 2D prostate segmentation and the effects of reconstruction on 3D surface generation are shown in Tables 2.3 and 2.4 for the absolute and signed evaluation metrics, respectively. Overall, our proposed method generated 3D surfaces with a median [first quartile (Q1), third quartile (Q3)] 3D DSC, recall, and precision of 94.1 [92.6, 94.9] %, 96.0 [93.1, 98.5] %, and 93.2 [88.8, 95.4] %, respectively, for the pixel map comparison metrics. Absolute VPD, MSD, and HD metrics resulted in 5.78 [2.49, 11.5] %, 0.89 [0.73, 1.09] mm, and 2.89 [2.37, 4.35] mm with signed metrics of 2.38 [-2.98, 11.0] %, 0.11 [-0.24, 0.58] mm, and 2.02 [-3.34, 2.88] mm, respectively. All metrics, aside from the absolute and signed HD metrics, showed significant differences between the 2D predictions and 3D reconstructed segmentations. Interestingly, recall and MSD metrics were observed to significantly improve in performance after 3D reconstruction, with the HD metric improving as well when evaluating all unseen images. These findings agreed when splitting the results into end-fire and side-fire 3D TRUS images, other than end-fire A/VPD and the signed metrics. For end-fire images, absolute VPD increased after 3D reconstruction, although this was not significant, while the signed metrics significantly improved after 3D reconstruction. For side-fire images, the opposite was true, with signed metrics significantly improved for 2D slice segmentations. When comparing the performance of our proposed method between end-fire and side-fire 3D TRUS images, we found there was no significant difference in any metric for both 2D radial segmentations and 3D reconstructed segmentations. Mean computation times were observed to be 0.029 s for each 2D segmentation (*i.e.*, 12 images) and 0.27 s for

reconstruction into a 3D surface, resulting in a total throughput time of 0.62 s from 3D image input to generated 3D surface.

Table 2.3. Absolute median [Q1, Q3] results comparing 2D radial slice segmentation to 3D reconstructed segmentation on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate.

Acquisition	Segmentation	DSC (%)	Recall (%)	Precision (%)	A/VPD (%)	MSD (mm)	HD (mm)
End-fire	2D Radial	**95.0 [93.6, 95.6]**	94.5 [92.7, 97.2]	**95.9 [92.2, 97.4]**	**4.71 [1.71, 7.32]**	1.16 [0.95, 1.37]	3.64 [3.11, 4.47]
	3D Reconstruction	94.3 [93.1, 95.2]	**96.0 [93.2, 98.7]**	94.6 [88.8, 95.8]	5.18 [1.62, 11.2]	**0.99 [0.78, 1.18]**	**3.41 [2.49, 4.41]**
	p-value	0.0052*	0.0102*	<0.0001	0.0532	<0.0001*	0.5217
Side-fire	2D Radial	**94.6 [92.7, 95.4]**	95.3 [90.6, 96.9]	**94.9 [92.6, 96.4]**	**4.05 [1.07, 6.23]**	0.95 [0.82, 1.26]	3.15 [2.51, 4.27]
	3D Reconstruction	93.5 [91.1, 94.6]	**96.2 [92.5, 98.4]**	91.6 [87.8, 94.8]	5.89 [3.17, 11.9]	**0.78 [0.67, 0.98]**	**2.61 [2.32, 4.01]**
	p-value	0.0037	0.0215	<0.0001*	0.0441	<0.0001	0.3683
Overall	2D Radial	**94.9 [93.2, 95.5]**	94.9 [91.9, 97.0]	**95.6 [92.6, 96.7]**	**4.34 [1.60, 6.77]**	1.06 [0.85, 1.32]	3.34 [2.61, 4.41]
	3D Reconstruction	94.1 [92.6, 94.9]	**96.0 [93.1, 98.5]**	93.2 [88.8, 95.4]	5.78 [2.49, 11.5]	**0.89 [0.73, 1.09]**	**2.89 [2.37, 4.35]**
	p-value	<0.0001	0.0005	<0.0001	0.0061	<0.0001*	0.2766

DSC, Dice similarity coefficient; A/VPD, area/volume percent diff.; MSD, mean surface dist.; HD, Hausdorff dist.
*Normal distribution = paired t-test; Bolded metrics highlight the method with reduced relative error.

Table 2.4. Signed median [Q1, Q3] results comparing 2D radial slice segmentation to 3D reconstructed segmentation on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate.

Acquisition	Segmentation	sA/VPD (%)	sMSD (mm)	sHD (mm)
End-fire	2D Radial	-1.39 [-3.82, 5.82]	-0.13 [-0.50, 0.71]	-0.91 [-2.74, 2.81]
	3D Reconstruction	**-0.05 [-2.98, 11.2]**	**0.06 [-0.38, 0.85]**	**-0.34 [-3.82, 3.28]**
	p-value	0.0011*	<0.0001*	0.9563
Side-fire	2D Radial	**-0.57 [-5.24, 3.27]**	**0.09 [-0.46, 0.31]**	**-0.31 [-1.83, 1.72]**
	3D Reconstruction	3.20 [-2.96, 10.1]	0.20 [-0.22, 0.46]	2.25 [-2.53, 2.81]
	p-value	0.0001*	<0.0001*	0.2305
Overall	2D Radial	**-0.91 [-4.93, 4.20]**	**-0.09 [-0.46, 0.37]**	**-0.91 [-2.15, 2.02]**
	3D Reconstruction	2.38 [-2.98, 11.0]	0.11 [-0.24, 0.58]	2.02 [-3.34, 2.88]
	p-value	<0.0001*	<0.0001*	0.3611

sA/VPD, signed area/volume percent diff.; sMSD, signed mean surface dist.; sHD, signed Hausdorff dist.
*Normal distribution = paired t-test; Bolded metrics highlight the method with reduced relative error.

2.3.2 3D CNNs and V-Net optimization

Results of the NiftyNet 3D V-Net with a Dice and DiceXEnt loss function on 20 unseen end-fire and 20 unseen side-fire 3D TRUS images are shown in Tables 2.5 and 2.6 for the absolute and signed evaluation metrics, respectively. When comparing 3D V-Net performance with Dice and DiceXEnt loss functions on the full testing dataset, all metrics, aside from DSC, showed significant differences. Precision, VPD/sVPD, MSD/sMSD, and HD/sHD were significantly improved with the DiceXEnt loss function, while recall was significantly improved with the Dice loss function. Although there was no significant difference in the DSC metric, the DiceXEnt loss function showed an improved median DSC. When considering end-fire and side-fire images individually, identical trends were observed for precision, recall, sMSD, and sHD. For the DSC, VPD, MSD, and HD metrics, we observed a significant and nonsignificant increase in performance with the DiceXEnt loss function for end-fire images and side-fire images, respectively. For end-fire images, the sVPD metric improved significantly with the DiceXEnt loss function, while the sVPD metric improved significantly with the Dice loss function for side-fire images. When comparing the 3D V-Net performance with DiceXEnt between end-fire and side-fire 3D TRUS images, we found no significant differences in any metric except HD, where side-fire images had a significantly reduced median value compared to end-fire images. Overall, the 3D V-Net showed improved performance with the DiceXEnt loss function and produced 3D segmentations with median [Q1, Q3] 3D DSC, recall, and precision results of 91.3 [88.6, 93.1] %, 90.0 [85.6, 93.3] %, and 94.5 [90.0, 96.5] %, respectively, for the pixel map comparison metrics. Absolute VPD, MSD, and HD metrics resulted in 7.94 [3.55, 13.4] %, 1.27 [0.92, 1.61] mm, and 6.18 [4.51, 7.82] mm with signed variants of -3.66 [-9.25, 3.34] %, -0.13 [-0.73, 0.26] mm, and -4.16 [-7.04, 4.76] mm, respectively. Mean computation times were observed to be 3.43 s for a full 3D segmentation.

Results of the Dense V-Net and High-resolution 3D-Net with a DiceXEnt loss function on 20 unseen end-fire and 20 unseen side-fire 3D TRUS images are shown in Tables 2.A1 and 2.A2 in Supplement A for the absolute and signed evaluation metrics, respectively. Compared to the 3D V-Net, the High-resolution 3D-Net showed a reduction in median performance for all metrics, while the Dense V-Net showed a reduction in

performance for all metrics except recall. In contrast with our proposed method and the 3D V-Net, we observed significant differences in segmentation performance for several metrics when comparing end-fire and side-fire 3D TRUS images using the Dense V-Net and High-resolution 3D-Net. For the Dense V-Net, improved performance was observed on side-fire images for every metric except recall, with significant differences observed for the precision, recall, VPD/sVPD, and sMSD metrics. For the High-resolution 3D-Net, improved performance was observed on end-fire images for every metric except precision and HD, with significant differences observed for the DSC, precision, recall, VPD/sVPD, and sMSD metrics. Mean 3D segmentation times for the Dense V-Net and High-resolution 3D-Net were observed to be 2.98 s and 2.83 s, respectively.

2.3.3 Comparison of reconstructed modified U-Net and 3D CNNs

Sample segmentation results from the 20 unseen end-fire and 20 unseen side-fire 3D TRUS images from our proposed method compared against the 3D V-Net with a DiceXEnt loss function and manual segmentations are shown in Fig. 2.4 and Fig. 2.5, respectively. A comparison of segmentation performance between our proposed method and a standard 3D V-Net is shown in Tables 2.7 and 2.8 for the absolute and signed evaluation metrics, respectively. Overall, our proposed method had significantly improved DSC, Recall, sVPD, MSD/sMSD, and HD when compared to the 3D V-Net. Absolute VPD and sHD, while not significantly different, were reduced for our proposed method. The only evaluation metric where the 3D V-Net outperformed our proposed method was precision, where the 3D V-Net showed a nonsignificant increase. Considering segmentation performance for end-fire and side-fire 3D TRUS images separately, similar trends hold. For end-fire images, our proposed method had better performance in all metrics except sHD, with DSC, recall, sVPD, MSD, and HD showing significant differences, and precision, VPD, and sMSD showing nonsignificant improvements. However, for side-fire images our proposed method was superior in all metrics except precision. Significant improvements were shown for DSC, recall, sVPD, MSD/sMSD, and HD/sHD, while the 3D V-Net had significantly improved precision. As was observed in

the overall case, our proposed method showed a nonsignificant decrease in VPD compared to the 3D V-Net for both end-fire and side-fire images.

Table 2.5. Absolute median [Q1, Q3] results comparing a standard 3D V-Net with a Dice similarity coefficient loss function to a Dice similarity plus cross-entropy (DiceXEnt) loss function on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate.

Acquisition	Loss function	DSC (%)	Recall (%)	Precision (%)	VPD (%)	MSD (mm)	HD (mm)
End-fire	Dice	89.5 [84.6, 92.0]	**97.8 [94.6, 98.3]**	83.7 [75.1, 88.4]	17.8 [7.57, 30.7]	1.79 [1.43, 2.47]	8.64 [7.53, 10.8]
	DiceXEnt	**91.7 [89.0, 93.2]**	91.7 [86.8, 94.6]	**94.3 [87.1, 95.8]**	**7.94 [2.95, 12.5]**	**1.32 [0.99, 1.77]**	**6.95 [5.06, 9.10]**
	p-value	0.0037	<0.0001	<0.0001	0.0021	0.0009	0.0172
Side-fire	Dice	90.6 [89.1, 93.2]	**94.5 [91.1, 96.0]**	92.8 [89.9, 96.9]	9.08 [4.35, 14.0]	1.16 [0.89, 1.46]	5.81 [3.85, 9.61]
	DiceXEnt	**91.2 [87.4, 92.8]**	89.5 [80.7, 92.9]	**95.0 [90.8, 97.5]**	**7.71 [3.55, 15.9]**	**1.11 [0.84, 1.47]**	**4.92 [4.28, 6.55]**
	p-value	0.2943	<0.0001	0.0001	0.7012	0.7562	0.2305
Overall	Dice	90.3 [86.5, 92.1]	**95.5 [92.5, 97.8]**	87.8 [81.0, 91.7]	11.4 [4.66, 19.3]	1.46 [1.16, 2.07]	7.99 [5.34, 10.4]
	DiceXEnt	**91.3 [88.6, 93.1]**	90.0 [85.6, 93.3]	**94.5 [90.0, 96.5]**	**7.94 [3.55, 13.4]**	**1.27 [0.92, 1.61]**	**6.18 [4.51, 7.82]**
	p-value	0.1538	<0.0001	<0.0001	0.0356	0.0147	0.0067

DSC, Dice similarity coefficient; VPD, volume percent difference; MSD, mean surface dist.; HD, Hausdorff dist.
*Normal distribution = paired t-test; Bolded metrics highlight the loss function with reduced relative error.

Table 2.6. Signed median [Q1, Q3] results comparing a standard 3D V-Net with a Dice loss function to a Dice plus cross-entropy (DiceXEnt) loss function on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate.

Acquisition	Loss function	sVPD (%)	sMSD (mm)	sHD (mm)
End-fire	Dice	17.8 [7.57, 30.7]	1.51 [0.85, 2.23]	7.99 [5.37, 10.84]
	DiceXEnt	**-2.16 [-8.47, 6.10]**	**-0.07 [-0.71, 0.68]**	**-0.05 [-7.33, 6.50]**
	p-value	<0.0001*	<0.0001*	0.0009
Side-fire	Dice	**5.02 [0.93, 11.5]**	0.43 [0.16, 0.80]	5.34 [3.18, 9.61]
	DiceXEnt	-5.50 [-15.6, -0.62]	**-0.30 [-0.88, 0.02]**	**-4.52 [-6.27, -3.04]**
	p-value	<0.0001*	<0.0001*	<0.0001
Overall	Dice	9.63 [3.90, 18.9]	0.83 [0.35, 1.78]	6.91 [3.35, 10.31]
	DiceXEnt	**-3.66 [-9.25, 3.34]**	**-0.13 [-0.73, 0.26]**	**-4.16 [-7.04, 4.76]**
	p-value	<0.0001*	<0.0001*	<0.0001*

sVPD, signed volume percent difference; sMSD, signed mean surface dist.; sHD, signed Hausdorff dist.
*Normal distribution = paired t-test; Bolded metrics highlight the loss function with reduced relative error.

Overall, our proposed method significantly improved performance on all metrics when compared to the Dense V-Net and showed improved performance on all metrics when compared to the High-resolution 3D-Net, with significant differences observed for all metrics except precision, sMSD, and sHD, where our method showed a nonsignificant improvement.

Table 2.7. Absolute median [Q1, Q3] results comparing a standard 3D V-Net to our proposed reconstructed modified (rm) U-Net on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate.

Acquisition	Segmentation	DSC (%)	Recall (%)	Precision (%)	VPD (%)	MSD (mm)	HD (mm)
End-fire	V-Net	91.7 [89.0, 93.2]	91.7 [86.8, 94.6]	94.3 [87.1, 95.8]	7.94 [2.95, 12.5]	1.32 [0.99, 1.77]	6.95 [5.06, 9.10]
	rmU-Net	**94.3 [93.1, 95.2]**	**96.0 [93.2, 98.7]**	**94.6 [88.8, 95.8]**	**5.18 [1.62, 11.2]**	**0.99 [0.78, 1.18]**	**3.41 [2.49, 4.41]**
	p-value	0.0003*	<0.0001*	0.5459	0.4980	0.0003*	<0.0001
Side-fire	V-Net	91.2 [87.4, 92.8]	89.5 [80.7, 92.9]	**95.0 [90.8, 97.5]**	7.71 [3.55, 15.9]	1.11 [0.84, 1.47]	4.92 [4.28, 6.55]
	rmU-Net	**93.5 [91.1, 94.6]**	**96.2 [92.5, 98.4]**	91.6 [87.8, 94.8]	**5.89 [3.17, 11.9]**	**0.78 [0.67, 0.98]**	**2.61 [2.32, 4.01]**
	p-value	0.0073	0.0002	0.0153	0.2611	0.0027	0.0001
Overall	V-Net	91.3 [88.6, 93.1]	90.0 [85.6, 93.3]	**94.5 [90.0, 96.5]**	7.94 [3.55, 13.4]	1.27 [0.92, 1.61]	6.18 [4.51, 7.82]
	rmU-Net	**94.1 [92.6 , 94.9]**	**96.0 [93.1, 98.5]**	93.2 [88.8, 95.4]	**5.78 [2.49, 11.5]**	**0.89 [0.73, 1.09]**	**2.89 [2.37, 4.35]**
	p-value	<0.0001	<0.0001	0.1499	0.1701	<0.0001*	<0.0001

DSC, Dice similarity coefficient; VPD, volume percent difference; MSD, mean surface dist.; HD, Hausdorff dist.
*Normal distribution = paired t-test; Bolded metrics highlight the method with reduced relative error.

Table 2.8. Signed median [Q1, Q3] results comparing a standard 3D V-Net to our proposed reconstructed modified (rm) U-Net on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate.

Acquisition	Segmentation	sVPD (%)	sMSD (mm)	sHD (mm)
End-fire	V-Net	-2.16 [-8.47, 6.10]	-0.07 [-0.71, 0.68]	**-0.05 [-7.33, 6.50]**
	rmU-Net	**-0.05 [-2.98, 11.2]**	**0.06 [-0.38, 0.85]**	-0.34 [-3.82, 3.28]
	p-value	0.0030*	0.0444*	0.7942*
Side-fire	V-Net	-5.50 [-15.6, -0.62]	-0.30 [-0.88, 0.02]	-4.52 [-6.27, -3.04]
	rmU-Net	**3.20 [-2.96, 10.1]**	**0.20 [-0.22, 0.46]**	**2.25 [-2.53, 2.81]**
	p-value	0.0001*	0.0025*	0.0107
Overall	V-Net	-3.66 [-9.25, 3.34]	-0.13 [-0.73, 0.26]	-4.16 [-7.04, 4.76]
	rmU-Net	**2.38 [-2.98, 11.0]**	**0.11 [-0.24, 0.58]**	**2.02 [-3.34, 2.88]**
	p-value	<0.0001*	0.0003*	0.0408

sVPD, signed volume percent difference; sMSD, signed mean surface dist.; sHD, signed Hausdorff dist.
*Normal distribution = paired t-test; Bolded metrics highlight the method with reduced relative error.

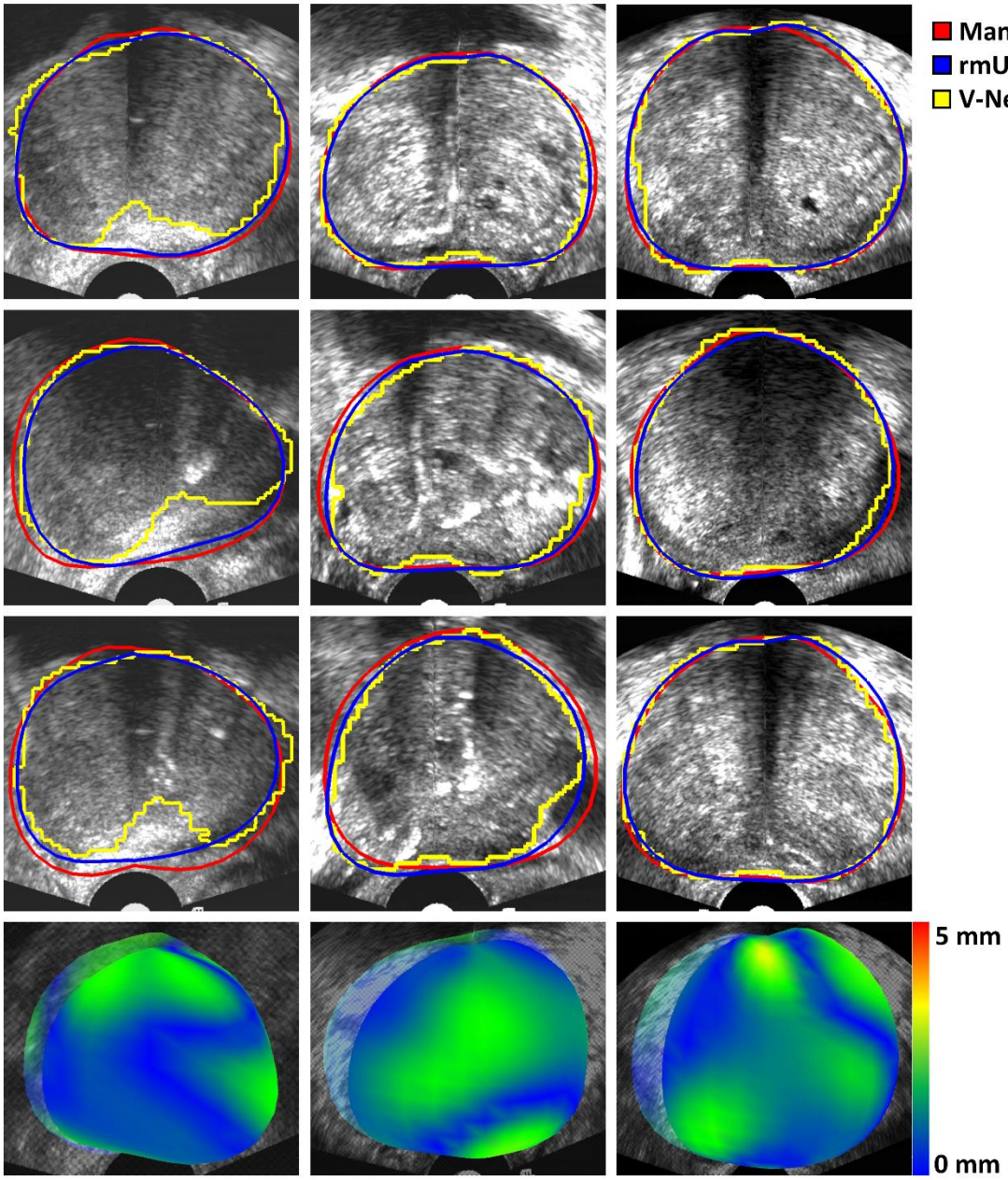

Figure 2.4. End-fire prostate segmentation results comparing manual (red), our proposed reconstructed modified (rm) U-Net (blue), and V-Net (yellow) 3D surfaces. The columns from left to right show the 25th, 50th, and 75th percentile results, respectively, based on DSC metrics. Segmentations in the axial plane, sagittal plane, 45° oblique radial plane, and reconstructed 3D surface error are shown in the respective rows from top to bottom.

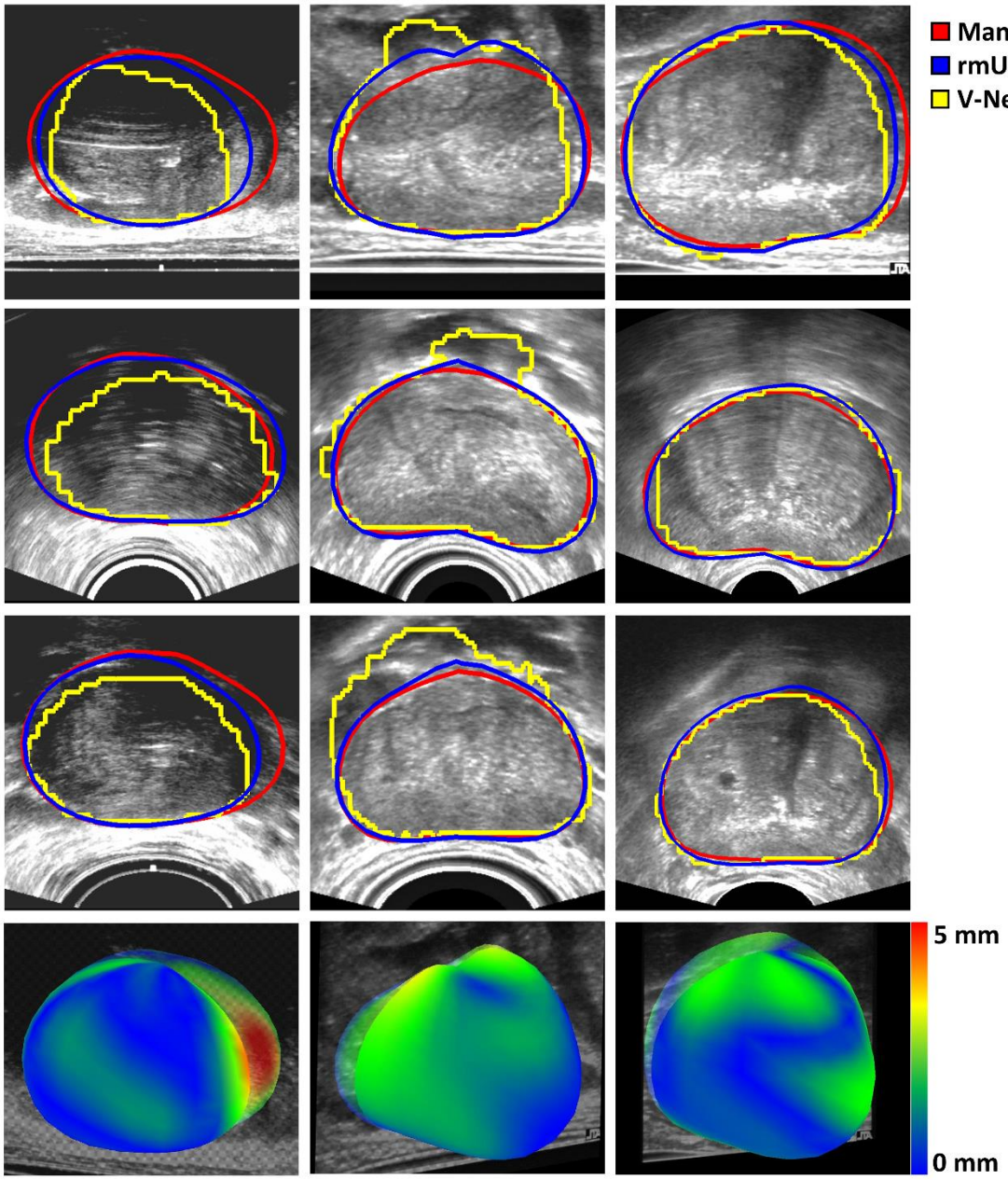

Figure 2.5. Side-fire prostate segmentation results comparing manual (red), our proposed reconstructed modified (rm) U-Net (blue), and V-Net (yellow) 3D surfaces. The columns from left to right show the 25th, 50th, and 75th percentile results, respectively, based on DSC metrics. Segmentations in the axial plane, sagittal plane, 45° oblique radial plane, and reconstructed 3D surface error are shown in the respective rows from top to bottom.

2.4 Discussion

2.4.1 Reconstructed modified U-Net

We proposed a new 3D TRUS prostate segmentation method, which utilizes a modified U-Net to segment 12 2D radial slices, which are then reconstructed into a 3D surface. We first compared segmentation accuracy in these 2D radial slices to segmentation accuracy following 3D reconstruction. In general, the performance was better on the 2D radial slice segmentations compared to the reconstructed 3D surface, but interestingly, our reconstruction method improved recall, MSD, and HD metrics when compared to 2D slice segmentation. When considered in combination, the reduced MSD and HD metrics showed better mean and irregular boundary accuracy, with the improved recall metric implying a reduction in underprediction (since decreasing underpredicted pixels, *i.e.*, false negatives, will increase recall). The reduction in underprediction is contrasted with our method tending to overpredict, supported by the slight positive bias in the signed metrics and reduced precision. In the metrics where performance was worse for our 3D reconstructed segmentations, such as DSC and VPD, the difference in median values were less than 1% and 1.5% respectively. Thus, we saw that our 3D reconstruction method did not drastically reduce performance of the evaluation metrics when compared to our 2D segmentations, with the 3D reconstruction improving performance on select metrics. Examining the signed metrics shown in Table 2.4, we observed that sVPD, sMSD, and sHD are reduced compared to their absolute metrics, with median sVPD reduced to only 2.38 %, median sMSD reduced to 0.11 mm, and median sHD reduced to 2.02 mm. This demonstrated that our algorithm was not significantly biased to over or underpredict the prostate boundary.

As our proposed network was trained and tested on both end-fire and side-fire 3D TRUS images, we directly compared the performance of our method on each image type observing no significant difference between performance for end-fire and side-fire images on any metric. Both independent image geometries also followed the same trends as the total dataset, aside from a reduced signed bias in the end-fire images following reconstruction. These results demonstrated the effectiveness of radially sampling 3D TRUS prostate images to produce similar 2D images for prediction and the ability to

accurately segment the prostate in different 3D TRUS image geometries without the need for multiple trained networks, which, to our knowledge, is the first time this has been shown.

2.4.2 3D CNNs and V-Net Optimization

The initial publication describing the V-Net architecture by Milletari *et al.*[24] proposed the use of a Dice coefficient-based loss function, but recent leave-one-out validation, described in Lei *et al.,*[16] has advocated for the use of hybrid loss functions that combine standard logistic loss, such as the cross-entropy loss metric,[20] with the Dice loss metric. We implemented a 3D V-Net with both a Dice loss function and a hybrid DiceXEnt loss function in order to compare performance between loss functions, as well as to directly compare performance to previously published V-Net implementations[16] on an unseen dataset. Our results reiterate what has been previously reported, with the hybrid DiceXEnt loss function significantly improving performance on all metrics except DSC and recall, where we observed a nonsignificant increase in performance and a significant decrease in performance, respectively. Similar trends held when examining the results for end-fire and side-fire 3D TRUS images individually. Comparing 3D V-Net performance when using a DiceXEnt loss function between end-fire and side-fire 3D TRUS images showed similar results to our proposed network, as there was no significant difference in any metric except HD.

Similar segmentation performance on end-fire and side-fire 3D TRUS images for all three 3D CNNs and our proposed 3D segmentation method demonstrated that we could potentially train a single network to accurately segment the prostate in geometrically variable 3D TRUS images. This was demonstrated predominantly with our approach and the 3D V-Net as the Dense V-Net and High-resolution 3D-Net were observed to have significant differences between several metrics when comparing end-fire and side-fire segmentation performance. Interestingly, performance differed between the Dense V-Net and High-resolution 3D-Net for different image geometries, with better predictions performed on side-fire and end-fires images, respectively. Although our method performed the best, the 3D V-Net outperformed the other two 3D CNNs investigated in this study.

This improved performance on 3D US prostate segmentation could be due to the number of parameters trained by the network since the 3D V-Net has approximately two orders of magnitude more parameters relative to the Dense V-Net and High-resolution 3D-Net. Although the latter networks are more efficient and required less computation time, we found this did not benefit performance.

Although Lei *et al.*[16] reported on a deep supervision method with contour refinement, they also reported on the use of a standard 3D V-Net with a hybrid DiceXEnt loss function to segment the prostate in side-fire 3D TRUS images, showing a 3D DSC, precision, recall, HD, MSD, and RMSD of 90.5 ± 3.0 %, 88.1 ± 6.0 %, 93.5 ± 3.5 %, 4.643 ± 1.926 mm, 0.657 ± 0.270 mm, and 0.977 ± 0.410 mm, respectively. Comparatively, a standard 3D V-Net with a DiceXEnt loss function trained on our dataset and predicted on side-fire images resulted in a 3D DSC, precision, recall, HD, and MSD of 91.2 [87.4, 92.8] %, 95.0 [90.8, 97.5] %, 89.5 [80.7, 92.9] %, 4.92 [4.28, 6.55] mm, and 1.11 [0.84, 1.47] mm, showing very similar performance. Investigating the differences between mean and median values showed our V-Net implementation demonstrated improved 3D DSC and precision, while Lei *et al.*[16] demonstrated improved recall, HD, and MSD. Although VDP or any signed metrics were not reported, this demonstrated similarity in performance. Thus, we suggest future comparisons should use a standardized V-Net, like the NiftyNet open-source implementation, to provide a baseline for comparing network performance on different data sets.

2.4.3 Comparison with 3D V-Net and previously published methods

For an identical training and testing dataset, our proposed method performed significantly better than the standard 3D V-Net with a hybrid loss function, with DSC, recall, sVPD, MSD/sMSD, and HD/sHD showing significant improvement. Our proposed method also demonstrated a reduced median VPD, although this difference was not significant. Similar differences in performance were observed when considering end-fire and side-fire 3D TRUS images separately. Fig. 2.4 and 2.5 show this difference in performance qualitatively for both image geometries, with the 3D V-Net often over or underpredicting the correct prostate boundary. This difference is readily apparent in the

side-fire middle 50th percentile column of Fig. 2.5. The V-Net is shown to have incorrectly contoured part of the bladder as the prostate, drastically affecting the overall segmentation accuracy. In contrast, our proposed method was able to correctly avoid the bladder, resulting in a more accurate segmentation. A similar result is shown in the left column of Fig. 2.4 as well as in Fig. 2.6, where we show that the 3D V-Net mistakenly underpredicted the prostate boundary due to the presence of a hyperechoic calcification artifact in the TRUS image, whereas our proposed method was typically able to avoid artifacts of this nature. We demonstrated a mean 3D segmentation time of 0.62 s with our proposed method, over five times faster than the 3D V-Net, which required an average of 3.43 s per segmentation. All segmentations were completed with the same NVIDIA GeForce GTX TITAN GPU with 6 GB of memory.

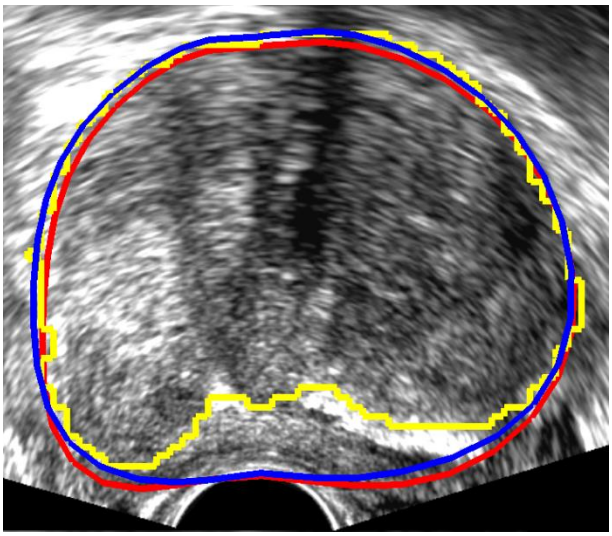

Figure 2.6. A sample end-fire prostate segmentation result comparing manual (red), our proposed algorithm (blue), and V-Net (yellow) 3D surfaces in the presence of a hyperechoic calcification image artifact.

Recent work by Ghavami *et al.*[15] and Lei *et al.*[16] report on automatic prostate segmentation in 3D TRUS images, with Ghavami *et al.*[15] reporting best results for 2D DSC, 3D DSC, and boundary distance of 89 ± 12 %, 89 ± 5 %, and 1.68 ± 1.57 mm, respectively, and Lei *et al.*[16] reporting overall segmentation results for 3D DSC, precision, recall, HD, MSD, and RMSD of 91.9 ± 2.8 %, 90.6 ± 5.5 %, 93.8 ± 4.3 %, 3.938 ± 1.550 mm, 0.599

± 0.225 mm, and 0.900 ± 0.377 mm, respectively. Our proposed method demonstrated overall segmentation results for 2D DSC, 3D DSC, precision, recall, HD, and MSD of 94.9 [93.2, 95.5] %, 94.1 [92.6, 94.9] %, 93.2 [88.08, 95.4] %, 96.0 [93.1, 98.5] %, 2.89 [2.37, 4.35] mm, and 0.89 [0.73, 1.09] mm, respectively, outperforming the network reported by Ghavami *et al*.[15] on all reported metrics, and outperforming the network reported by Lei *et al*.[16] on all metrics except MSD. In addition, our paper reports on metrics not used in the studies described here, including VPD, and signed variants of VPD, MSD, and HD, providing additional insight into the performance of our network that is otherwise not present when these metrics are excluded. Lei *et al*.[16] reported segmentation times of approximately 1-2 s for a U-Net, V-Net, and their proposed network, with segmentations completed using an NVIDIA TITAN XP GPU with 12 GB of memory. Comparatively, our V-Net implementation in NiftyNet had a mean segmentation time of 3.43 s, while our proposed method had a mean segmentation time of 0.62 s, with segmentations completed using an NVIDIA GeForce GTX TITAN GPU with 6 GB of memory. Although our V-Net had slower segmentation times, our proposed method was one and a half to three times faster, using a GPU with half the memory, demonstrating the advantage of our proposed method regarding segmentation time.

Studies reported by Ghavami *et al*.[15] and Lei *et al*.[16] are also limited by their use of cross-validation approaches, in addition to their dataset consisting of only one image geometry and a single ultrasound machine with matched voxel dimensions and sizes. In contrast, we used a clinically diverse dataset of 3D TRUS images of different image geometries, generated by several different ultrasound transducers used with different machine models, and used in distinct procedures. This dataset contains images with varying voxel dimensions, size, and image quality, and we have employed no pre-processing before training to reduce process complexity. To our knowledge, the use of a dataset with different image geometries, ultrasound transducers, ultrasound machine models, voxel dimensions, and image sizes for 3D TRUS prostate segmentation is unique and may allow for a more robust and generalizable segmentation method. In addition, we have not used a cross-validation approach, instead testing our algorithm on 3D TRUS images that were completely unseen by the network, which we believe strengthens the significance of our results and may result in improved generalizability.

2.4.4 Limitations and future work

Although we have demonstrated excellent performance with our proposed algorithm, a parameter that may be interesting to explore in future work is the step angle for radial slice generation. Specifically, it would be interesting to investigate whether a decreased step angle and thus an increased number of radial slices would significantly increase performance, or whether a decreased number of slices could be used while maintaining similar performance. An increased number of radial slices would increase computation time, which motivated our choice of a 15° step angle for the proposed method. When training the 3D CNNs for performance comparison, hyperparameters were optimized on the V-Net and were used for the other two networks. Even though we investigated other combinations and found these hyperparameters to perform the best on the Dense V-Net and High-resolution 3D-Net, a rigorous optimization was not performed and has the potential to increase performance. Another limitation of our study is the use of only one observer for providing ground truth segmentations. This meant we could not directly assess inter-observer variability for our dataset. In addition, we did not directly assess intra-observer variability over several time points. Inter- and intra-observer variability in end-fire 3D TRUS images were previously assessed by our group,[11] and are summarized here. To assess intra-observer variability, one observer segmented 15 3D images five times each, resulting in a 3D DSC of 93.0 ± 2.1 %. To assess inter-observer variability, three untrained observers segmented 15 3D images three times each, resulting in a DSC of 93.5 ± 2.1 %, 92.6 ± 3.1 %, and 92.3 ± 3.2 %, with an ANOVA demonstrating no significant difference. Inter- and intra-observer variability in side-fire 3D TRUS images were also assessed by our group,[28] reporting 5.1 % variability and 99 % reliability in intra-observer prostate volume estimates, and 11.4 % variability and 96 % reliability in inter-observer estimates, for a study of eight observers measuring 15 prostate images twice. This variability is comparable to the reported DSC of our proposed method in this work, showing that our algorithm is performing at the level of intra-observer variability in the ground truth segmentations. Due to the demonstrated variability between different observers when segmenting 3D TRUS images, segmentations from other observers should

be incorporated into our testing dataset to further improve the robustness of our proposed method.

2.5 Conclusions

This study investigated the development of a 2D deep learning with 3D reconstruction approach for automatic prostate segmentation in 3D TRUS images. Multiple facilities, ultrasound machine models, and acquisition geometries were investigated to evaluate robustness and generalizability, with comparisons performed against multiple 3D CNNs. A fast, accurate, and generalizable automatic prostate segmentation algorithm could reduce physician burden and procedure time, offering potential workflow benefits for fusion-guided prostate biopsy, tumor-targeted HDR-BT, and TRUS-guided whole-gland BT. Reducing the time a patient is under anesthesia, as in HDR-BT, also promotes a potentially safer procedure with fewer adverse side effects.

2.6 References

1. Stewart BW, Wild C. *World Cancer Report 2014*. Lyon; 2014.

2. Narayanan R, Kurhanewicz J, Shinohara K, Crawford ED, Simoneau A, Suri JS. MRI-ultrasound registration for targeted prostate biopsy. In: *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. IEEE; 2009:991-994. doi:10.1109/ISBI.2009.5193221

3. Karnik V V, Fenster A, Bax J, et al. Assessment of image registration accuracy in three-dimensional transrectal ultrasound guided prostate biopsy. *Med Phys*. 2010;37(2):802-813. doi:10.1118/1.3298010

4. Choi YJ, Kim JK, Kim HJ, Cho K-S. Interobserver Variability of Transrectal Ultrasound for Prostate Volume Measurement According to Volume and Observer Experience. *Am J Roentgenol*. 2009;192(2):444-449. doi:10.2214/AJR.07.3617

5. Murciano-Goroff YR, Wolfsberger LD, Parekh A, et al. Variability in MRI vs. ultrasound measures of prostate volume and its impact on treatment recommendations for favorable-risk prostate cancer patients: A case series. *Radiat Oncol*. 2014;9(1):200. doi:10.1186/1748-717X-9-200

6. Mahdavi SS, Spadinger I, Chng N, Salcudean SE, Morris WJ. Semiautomatic

segmentation for prostate brachytherapy: Dosimetric evaluation. *Brachytherapy*. 2013;12(1):65-76. doi:10.1016/j.brachy.2011.07.007

7. Reich DL, Hossain S, Krol M, et al. Predictors of Hypotension After Induction of General Anesthesia. *Anesth Analg*. 2005;101(3):622-628. doi:10.1213/01.ANE.0000175214.38450.91

8. Wu Qiu, Jing Yuan, Ukwatta E, Yue Sun, Rajchl M, Fenster A. Prostate Segmentation: An Efficient Convex Optimization Approach With Axial Symmetry Using 3-D TRUS and MR Images. *IEEE Trans Med Imaging*. 2014;33(4):947-960. doi:10.1109/TMI.2014.2300694

9. Qiu W, Rajchl M, Guo F, et al. 3D prostate TRUS segmentation using globally optimized volume-preserving prior. In: *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol 8673 LNCS; 2014:796-803. doi:10.1007/978-3-319-10404-1_99

10. Yuan J, Qiu W, Rajchl M, Ukwatta E, Tai X-C, Fenster A. Efficient 3D Endfiring TRUS Prostate Segmentation with Globally Optimized Rotational Symmetry. In: *2013 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE; 2013:2211-2218. doi:10.1109/CVPR.2013.287

11. Qiu W, Yuan J, Ukwatta E, Fenster A. Rotationally resliced 3D prostate TRUS segmentation using convex optimization with shape priors. *Med Phys*. 2015;42(2):877-891. doi:10.1118/1.4906129

12. Ghavami N, Hu Y, Gibson E, et al. Automatic segmentation of prostate MRI using convolutional neural networks: Investigating the impact of network architecture on the accuracy of volume measurement and MRI-ultrasound registration. *Med Image Anal*. 2019;58:101558. doi:10.1016/j.media.2019.101558

13. Litjens G, Toth R, van de Ven W, et al. Evaluation of prostate segmentation algorithms for MRI: the PROMISE12 challenge. *Med Image Anal*. 2014;18(2):359-373. doi:10.1016/j.media.2013.12.002

14. Anas EMA, Mousavi P, Abolmaesumi P. A deep learning approach for real time prostate segmentation in freehand ultrasound guided biopsy. *Med Image Anal*. 2018;48:107-116. doi:10.1016/j.media.2018.05.010

15. Ghavami N, Hu Y, Bonmati E, et al. Integration of spatial information in convolutional neural networks for automatic segmentation of intraoperative transrectal ultrasound images. *J Med Imaging*. 2018;6(1):011003. doi:10.1117/1.jmi.6.1.011003

16. Lei Y, Tian S, He X, et al. Ultrasound prostate segmentation based on multidirectional deeply supervised V-Net. *Med Phys*. 2019;46(7):3194-3206. doi:10.1002/mp.13577

17. Dwork C, Feldman V, Hardt M, Pitassi T, Reingold O, Roth A. The reusable holdout: Preserving validity in adaptive data analysis. *Science*. 2015;349(6248):636-638. doi:10.1126/science.aaa9375

18. Valdes G, Interian Y. Comment on 'Deep convolutional neural network with transfer learning for rectum toxicity prediction in cervical cancer radiotherapy: a feasibility study.' *Phys Med Biol*. 2018;63(6):068001. doi:10.1088/1361-6560/aaae23

19. Fenster A, Parraga G, Bax J. Three-dimensional ultrasound scanning. *Interface Focus*. 2011;1(4):503-519. doi:10.1098/rsfs.2011.0019

20. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2015;9351:234-241. doi:10.1007/978-3-319-24574-4_28

21. Chollet F and others. Keras. 2015. http://keras.io/.

22. Abadi M, Agarwal A, Barham P, et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. 2016; arXiv Prepr. arXiv1603.04467.

23. Girardeau-Montaut D. CloudCompare-open source project. http://www.cloudcompare.org/.

24. Milletari F, Navab N, Ahmadi S-A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. *International Conference on 3D Vision*. 2016:565-571. doi:10.1109/3DV.2016.79

25. Gibson E, Giganti F, Hu Y, et al. Automatic Multi-Organ Segmentation on Abdominal CT with Dense V-Networks. *IEEE Trans Med Imaging*. 2018;37(8):1822-1834. doi:10.1109/TMI.2018.2806309

26. Li W, Wang G, Fidon L, Ourselin S, Cardoso MJ, Vercauteren T. On the compactness, efficiency, and representation of 3D convolutional networks: Brain parcellation as a pretext task. In: *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol 10265 LNCS. Springer Verlag; 2017:348-360. doi:10.1007/978-3-319-59050-9_28

27. Gibson E, Li W, Sudre C, et al. NiftyNet: a deep-learning platform for medical imaging. *Comput Methods Programs Biomed*. 2018;158:113-122. doi:10.1016/j.cmpb.2018.01.025

28. Tong S, Cardinal HN, McLoughlin RF, Downey DB, Fenster A. Intra- and inter-observer variability and reliability of prostate volume measurement via two-dimensional and three-dimensional ultrasound imaging. *Ultrasound Med Biol*. 1998;24(5):673-681. doi:10.1016/S0301-5629(98)00039-8

2.7 Supplement A

Table 2.A1. Absolute median [Q1, Q3] results comparing a Dense V-Net (DenseNet) and High-resolution 3D Network (HighRes3dNet) on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate.

Acquisition	Segmentation	DSC (%)	Recall (%)	Precision (%)	VPD (%)	MSD (mm)	HD (mm)
End-fire	DenseNet	87.9 [84.9, 90.4]	95.4 [92.9, 97.0]	82.2 [77.4, 86.5]	15.6 [9.42, 20.8]	2.11 [1.55, 2.85]	10.0 [7.82, 12.5]
	HighRes3dNet	90.0 [87.3, 91.4]	91.7 [84.6, 94.9]	89.3 [87.0, 93.7]	6.88 [4.30, 15.7]	1.62 [1.41, 2.08]	8.95 [7.56, 11.2]
Side-fire	DenseNet	88.7 [81.6, 90.7]	89.1 [86.0, 95.1]	89.6 [82.8, 91.9]	6.38 [4.22, 12.6]	1.47 [1.26, 2.56]	8.96 [5.93, 13.8]
	HighRes3dNet	86.2 [84.3, 88.7]	80.5 [76.8, 86.0]	94.9 [90.9, 96.1]	15.3 [9.77, 20.0]	1.64 [1.36, 1.93]	7.82 [6.67, 9.31]
Overall	DenseNet	88.2 [84.3, 90.6]	93.0 [88.9, 96.1]	84.8 [79.2, 90.2]	11.0 [5.96, 19.3]	2.00 [1.36, 2.63]	9.23 [7.00, 13.2]
	HighRes3dNet	87.5 [85.5, 90.3]	85.4 [79.8, 92.1]	92.3 [88.4, 95.8]	11.8 [4.68, 18.0]	1.63 [1.40, 1.97]	8.37 [6.89, 10.9]

DSC, Dice similarity coefficient; VPD, volume percent difference; MSD, mean surface dist.; HD, Hausdorff dist.

Table 2.A2. Signed median [Q1, Q3] results comparing a Dense V-Net (DenseNet) and High-resolution 3D Network (HighRes3dNet) on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate.

Acquisition	Segmentation	sVPD (%)	sMSD (mm)	sHD (mm)
End-fire	DenseNet	15.6 [9.42, 20.8]	1.44 [0.92, 2.18]	9.13 [7.19, 12.3]
	HighRes3dNet	2.86 [-6.42, 8.86]	0.37 [-0.46, 0.93]]	6.47 [-8.97, 9.71]
Side-fire	DenseNet	0.36 [-4.91, 7.07]	0.38 [-0.15, 1.03]	8.67 [-2.86, 13.8]
	HighRes3dNet	-13.8 [-20.0, -7.55]	-0.76 [-1.23, -0.08]	-6.69 [-8.40, 4.38]
Overall	DenseNet	8.77 [-2.84, 18.3]	0.95 [0.29, 1.61]	8.87 [5.72, 13.2]
	HighRes3dNet	-6.89 [-16.4, 4.09]	-0.24 [-1.15, 0.58]	-5.84 [-8.46, 8.32]

sVPD, signed volume percent difference; sMSD, signed mean surface dist.; sHD, signed Hausdorff dist.

Chapter 3

3 Effect of dataset size, image quality, and image type on deep learning-based automatic prostate segmentation in 3D ultrasound

While deep learning-based segmentation approaches are promising, access to large clinically diverse datasets is rare, especially in US imaging. An efficient automatic segmentation algorithm that could maintain high segmentation accuracy when trained with smaller datasets may increase access to deep learning segmentation, even if data is scarce. The purpose of Chapter 3 is to examine the effect of dataset size and image quality on deep learning segmentation of the prostate in 3D TRUS images.

The contents of this chapter have been previously published in *Physics in Medicine & Biology*: Orlando N, Gyacskov I, Gillies DJ, Guo F, Romagnoli C, D'Souza D, Cool DW, Hoover D, and Fenster A. *Physics in Medicine & Biology*. 2022;67:074002. Permission to reproduce this article was granted by IOP Publishing Ltd. and is provided in Appendix A – Copyright Releases.

3.1 Introduction

Prostate biopsy is the current clinical standard for prostate cancer (PCa) diagnosis, but the conventional two-dimensional (2D) transrectal ultrasound (TRUS)-guided biopsy has been reported to have a false negative rate up to 30%.[1] Improved tumour sampling can be achieved with the addition of three-dimensional (3D) TRUS imaging using a magnetic resonance imaging (MRI)-3D TRUS fusion guided biopsy approach, which utilizes the superior soft-tissue contrast of MRI to identify suspicious tissue regions for targeting with real-time TRUS guidance.[2] For PCa treatment, high-dose-rate (HDR) brachytherapy is a common treatment modality for intermediate and high-risk localized PCa.[3,4] 3D TRUS imaging provides spatial context through visualization of the anatomy in 3D, in addition to

improving intraoperative needle tip identification and patient-specific dose optimization.[5] While 3D TRUS imaging offers many benefits, it also necessitates accurate segmentation of the prostate to perform the biopsy and brachytherapy clinical tasks. Segmentations are often completed manually by the physician during the procedure, which can be time-consuming and highly variable, extending procedure times and increasing risk due to anesthesia exposure.[6]

Multiparametric MRI is quickly becoming a standard of care imaging modality for both biopsy and brachytherapy procedures based on the results of recent randomized controlled clinical trials.[7–9] While MRI offers high soft-tissue contrast, limitations including high-cost limit widespread adoption. For MRI-guided brachytherapy, patient movement to the MRI scanner after needle implant may cause needle shifts,[10] while in-bore procedures require highly specialized magnet-safe tools. In hospitals where MRI is utilized for MRI-3D TRUS fusion-guided biopsy or brachytherapy, registration between the modalities is required, often utilizing a surface-based approach, which requires accurate segmentation of the prostate in both image modalities. Thus, even with an MRI-based approach, accurate and fast prostate segmentation in 3D TRUS images is critical.

With the increasing prevalence of deep learning in medicine,[11] and specifically of convolutional neural networks (CNNs) for medical imaging tasks, many deep learning-based automatic prostate segmentation approaches have been proposed for TRUS imaging, promising reduced procedure time and similar performance compared to manual approaches.[12–16] Recently, Lei *et al.* proposed an anchor-free mask CNN for multi-organ segmentation in 3D TRUS volumes, trained using data from 83 PCa patients with five-fold cross-validation.[17] They reported prostate segmentation accuracy with a Dice similarity coefficient (DSC) of 0.93 ± 0.03 and 95% Hausdorff distance (HD_{95}) of 2.27 ± 0.79 mm. van Sloun *et al.* propose a U-Net based approach for zonal prostate segmentation trained on a large multi-institutional dataset of 436 3D TRUS volumes from 181 men, reporting a median (95% confidence interval) accuracy of 98 (95-99)%, Jaccard index of 0.93 (0.80-0.96), and Hausdorff distance (HD) of 3.0 (1.3-8.7) mm.[18]

We previously proposed an automatic segmentation algorithm involving deep learning prediction with a modified U-Net on 2D TRUS images radially sliced from 3D TRUS volumes followed by reconstruction into a 3D surface.[19] The algorithm was trained

on a large dataset with nearly 7000 2D images resliced from 206 clinically variable 3D TRUS images from various acquisition methods, procedure types (biopsy and HDR brachytherapy), ultrasound machines, and ultrasound transducers. On a testing set of 40 unseen 3D TRUS volumes from different acquisition methods, we demonstrated high performance with a median [quartile 1, quartile 3] DSC, mean surface distance (MSD), and HD of 94.1 [92.6, 94.9] %, 0.89 [0.73, 1.09] mm, and 2.89 [2.37, 4.35] mm, respectively. This algorithm outperformed a fully 3D V-Net and state-of-the-art methods in the literature.[19]

However, large and diverse medical image datasets are rare, especially for ultrasound, with recent papers reporting utilization of datasets with 2238 2D TRUS images,[12] and 40,[16] 44,[14] 86,[17] and 109[13] 3D TRUS volumes. Generation of large clinical datasets is time-consuming and costly, so generalizable and accurate automatic segmentation approaches suitable for small datasets are critical for the widespread integration of deep learning in minimally invasive PCa procedures. The U-Net++ architecture, an evolution of the standard U-Net, has recently been proposed and may help accomplish this goal, introducing multiple CNN backbones as well as nested, dense skip connections.[20,21] These redesigned skip connections attempt to reduce the semantic gap between feature maps in the encoder and decoder sections of the network, resulting in an easier optimization problem and thus higher performance with small training datasets.

Image quality is highly variable between 3D TRUS volumes, including differences in acquisition methods, image acquisition artifacts, patient anatomy artifacts such as gas, calcifications, catheters, and prostate boundary visibility. These factors may influence prostate segmentation results, so a 3D TRUS image quality grading scale is required to aid in comparing results from different datasets and identify key image quality factors that will influence segmentation performance.

Our work aims to demonstrate that our 2D radial deep learning plus 3D reconstruction approach offers efficient utilization of training data and thus high segmentation performance when trained with smaller datasets and datasets split based on 3D TRUS acquisition type (end-fire and side-fire). To assess the impact of image quality on segmentation performance, we propose an image quality grading scale containing three distinct image quality factors. By rigorously evaluating our deep learning segmentation

approach in the context of image quality, training dataset size, and 3D TRUS acquisition type (end-fire and side-fire), we intend to provide a widely accessible, robust, generalizable, and efficient prostate segmentation algorithm. This approach may allow for reduced clinical procedure time and increased efficiency for minimally invasive PCa procedures, allowing for optimization to a clinic's local preferences, without requiring access to large ultrasound image datasets.

3.2 Materials and methods

3.2.1 Complete ultrasound dataset

The complete ultrasound dataset consisted of 246 3D TRUS volumes of the prostate.[22] This dataset contained 104 end-fire 3D TRUS volumes, obtained from clinical prostate biopsy procedures, and 142 side-fire 3D TRUS volumes, obtained from clinical prostate brachytherapy procedures. Patient clinical information such as age, stage of prostate cancer, and Gleason score were not recorded. The methods to acquire 3D TRUS volumes have been described previously, but are briefly summarized here.[23–25] To generate these images, a TRUS transducer was mechanically rotated using a motorized fixture about its long axis. 2D TRUS images were acquired at set angular intervals, which were then reconstructed to generate 3D TRUS volumes. The choice of TRUS transducer leads to geometrically variable images: the end-fire transducer used for prostate biopsy was rotated 180° while 2D TRUS images were acquired at 1.0° intervals and reconstructed into a 3D volume; the side-fire transducer used for prostate brachytherapy was rotated 140° while 2D TRUS images were acquired at 0.5° intervals and reconstructed into a fan-shaped 3D TRUS volume.

These 3D TRUS volumes were acquired with 3 transducers used with 3 different ultrasound systems of different ages and from two manufacturers. Specifically, an 8848 transducer was used with the Profocus 2202 ultrasound system (BK Medical, Peabody, MA, USA), C9-5 and BPTRT9-5 transducers were used with the ATL HDI-5000 ultrasound system (Philips, Amsterdam, the Netherlands), and a C9-5 transducer was used with the iU22 ultrasound system (Philips, Amsterdam, the Netherlands). 3D TRUS image

sizes ranged from [448×350×448] to [692×520×692] voxels with dimensions of [0.115×0.115×0.115] to [0.190×0.195×0.190] mm^3/voxel for end-fire images and from [300×400×784] to [408×441×870] voxels with dimensions of [0.094×0.154×0.154] to [0.183×0.186×0.186] mm^3/voxel for side-fire images. Manual prostate segmentations in the 3D TRUS volumes, excluding the seminal vesicles, were completed by an observer experienced with 3D TRUS imaging (IG). 20 end-fire 3D TRUS volumes and 20 side-fire 3D TRUS volumes were randomly selected from the complete dataset and reserved as a testing dataset, thus were not included during training.

As outlined in Orlando and Gillies *et al.,* the complete training dataset of 206 3D TRUS volumes was resliced at randomized axial, sagittal, coronal, radial, and oblique image planes, resulting in a final training dataset of 6761 2D TRUS images with matched manual segmentations.[19] This reslicing allowed for more efficient use of the TRUS data, demonstrating improved performance compared to a fully 3D V-Net approach.[19,26] 2D images were resampled to 256×256 pixels with no other applied preprocessing. The complete training dataset of 2D TRUS images was split for deep learning, with 80% (5409 images) used for training and 20% (1352 images) used for validation.

3.2.2 Reduced-size datasets

To evaluate our method's efficiency in utilizing the training data, we generated smaller datasets by splitting and reducing the complete dataset of 6761 2D TRUS images. In all smaller datasets, we maintained the 80/20 training/validation split for deep learning.

3.2.2.1 Split end-fire and side-fire datasets

We first split the complete dataset into an end-fire training dataset of 2738 2D TRUS images and a side-fire training dataset of 4023 2D TRUS images (Table 3.1). This allowed for an assessment of generalizability by training two sets of parameters and testing on the opposite 3D TRUS acquisition type, which was unseen during training.

3.2.2.2 Smaller end-fire, side-fire, and mixed datasets

Smaller datasets were generated by reducing the split end-fire and side-fire datasets to assess how segmentation performance depends on the size of the dataset used for training, with an aim to find the smallest dataset which still maintains high segmentation performance. Using the split end-fire and side-fire datasets, images were removed at random to create training datasets with 1000, 500, 250, and 100 2D TRUS images of each acquisition type (Table 3.1). These smaller datasets were generated by reslicing from 36, 18, 9, and 4 3D TRUS volumes, respectively. Thus, variation in image quality and anatomical features, as determined by the 3D TRUS volume, were similarly reduced. This resulted in eight reduced-size datasets (four end-fire and four side-fire).

Table 3.1. Breakdown of the clinical 3D TRUS training dataset of 206 volumes resliced into 2D images for training 2D neural networks. Rows from top to bottom show the complete and reduced-size side-fire, end-fire datasets, and mixed datasets.

Complete Dataset	6761 2D TRUS Images		
Split Dataset [Images]	4023 Side-fire	2738 End-fire	4023 Mixed 2738 Mixed
Smaller Datasets [Images]	1000 Side-fire 500 Side-fire 250 Side-fire 100 Side-fire	1000 End-fire 500 End-fire 250 End-fire 100 End-fire	1000 Mixed 500 Mixed 250 Mixed 100 Mixed

Similarly, smaller mixed datasets were generated by reducing the complete 2D TRUS dataset. Images were removed at random to create training datasets with 4023, 2738, 1000, 500, 250, and 100 mixed 2D TRUS images (Table 3.1), resliced from 119, 86, 36, 18, 9, and 4 3D TRUS volumes, respectively. The segmentation performance of a network trained using 4023 mixed acquisition images was compared to a network trained using 4023 side-fire images; similarly, the segmentation performance using a training dataset of 2738 mixed images and 2738 end-fire images were compared. In all reduced-size mixed

datasets, the ratio between end-fire and side-fire images matched the complete dataset, with 40.5% end-fire images and 59.5% side-fire images. This resulted in six reduced-size mixed datasets.

3.2.3 Image quality assessment

3D TRUS image quality varies across patients and 3D acquisition methods, and so it is expected to impact segmentation performance. To explore this effect, an experienced interventional and genitourinary radiologist (DC) developed a 3D TRUS image quality grading scale, provided in Table 3.2. To ensure the scale was not biased and was generalizable, it was developed before the physician viewed our 3D TRUS dataset. Image quality was graded using three factors: acquisition quality, artifact severity, and prostate boundary visibility. Acquisition quality rated the quality of the 3D TRUS acquisition itself, ignoring the anatomy artifacts and visibility, ranging from 1 (poor) to 5 (ideal). Examples of poor acquisition quality included image shadowing due to inadequate transducer contact, transducer translation during 3D TRUS acquisition causing anatomy distortion, and issues with ultrasound gain or depth. Artifact severity estimated the degree of image degradation caused by artifact-generating items within the prostate gland, such as calcifications, gas, urinary catheters, and brachytherapy seeds, ranging from 1 (major artifacts) to 5 (no artifacts at all). Prostate boundary graded the visibility or clarity of the prostate boundary with the adjacent periprostatic soft tissue, a key factor in the prostate segmentation task, ranging from 1 (more than 75% of the boundary is indistinguishable) to 3 (40% of the boundary is indistinguishable) to 5 (the entire boundary is clearly visible). The test dataset of 20 end-fire and 20 side-fire 3D TRUS volumes was graded by the same radiologist who was blinded to the qualitative and quantitative segmentation performance. Only the test dataset was graded; as the test dataset was randomly selected from the complete dataset, its images quality distribution was representative of the complete dataset. Five-point numerical grading allowed for a quantitative comparison between end-fire and side-fire 3D TRUS volumes, including the calculation of means and statistical testing.

Table 3.2. Image quality grading scale for 3D TRUS images of the prostate.

Image Quality Factor	Description	Scale
Acquisition Quality	Quality of the 3D TRUS image acquisition regardless of anatomy	1 (poor) - 5 (ideal)
Anatomy Artifacts	Severity of anatomy artifacts (calcification, gas, catheter, etc.)	1 (major artifacts) – 5 (no artifacts)
Prostate Boundary	Visibility/clarity of the prostate boundary	1 (> 75% of boundary indistinguishable) - 3 (40% of boundary indistinguishable) - 5 (entire boundary visible)

3.2.4 3D segmentation algorithm

Our radial prostate segmentation algorithm was first described in Orlando and Gillies *et al.* and will be briefly summarized here (Fig. 3.1). This method utilized a radial segmentation approach, first proposed by Qiu *et al.* for a prostate segmentation algorithm based on convex optimization with shape priors.[27] In this approach, a 3D TRUS volume is resliced radially about the approximate center of the prostate gland at 15° intervals, generating 12 2D TRUS images. The extracted 2D TRUS images appear very similar, as each plane passes through the mid-gland of the prostate, resulting in similar prostate size and shape regardless of the 3D TRUS acquisition method. This radial approach has been shown to improve segmentation performance in the apex and base of the prostate compared to alternative approaches such as transverse reslicing.[27] The 12 radial 2D TRUS images were automatically segmented using neural networks trained with the 2D datasets described in sections 3.2.1 and 3.2.2 to generate 12 segmented prostate boundaries, which were used to reconstruct the 3D surface of the prostate (Fig. 3.1).

3.2.5 2D neural networks

Two neural network architectures were used in this work, which were trained with identical 2D TRUS datasets (See sections 3.2.1 and 3.2.2). Detailed network diagrams are provided in Fig. 3.A1 and 3.A2 in Supplement A for the modified U-Net and U-Net++, respectively. Data augmentation using random combinations of horizontal flips, shifts up to 20%, rotations up to 20°, and zooms up to 20% were applied to double the training

datasets. A personal computer with an i7-9700K central processing unit (CPU) at 3.60 GHz (Intel Corporation, Santa Clara, CA, USA), 64 GB of RAM, and a 24 GB NVIDIA TITAN RTX (NVIDIA Corporation, Santa Clara, CA, USA) graphics processing units (GPU) was used for training all 2D neural networks and for subsequent prediction on unseen testing data.

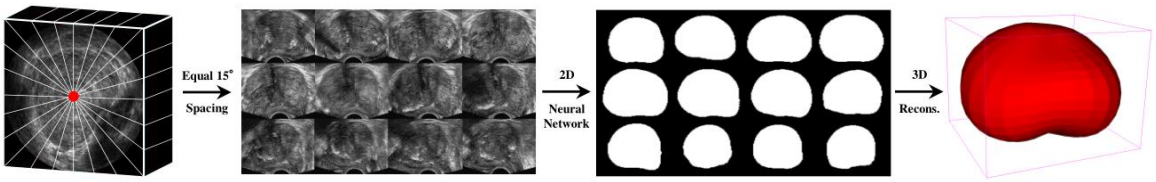

Figure 3.1. 3D prostate segmentation workflow using an example end-fire 3D TRUS volume. The input 3D TRUS volume was resliced radially at 15° spacing to generate 12 2D TRUS images with similar size and shape. A trained 2D neural network was used to predict the prostate boundary locations in 2D binary masks, which were used to reconstruct the 3D prostate surface.

3.2.5.1 Modified U-Net

A five-layer deep modified version of the widely prevalent U-Net[28] was implemented using Keras[29] with TensorFlow.[30] First, 50% dropouts were applied at the last block on the contracting section of the network and at every block on the expansion section of the network to increase regularization and prevent overfitting.[19] In addition, transpose convolutions were applied in the expansion section of the network instead of the standard upsampling followed by convolution (upconvolution), as this allowed for improved performance [19]. Padding and ReLU activation were applied in each (3×3) convolution operation, with sigmoid activation used in the final (1×1) convolution operation. Additional hyperparameter selection based on preliminary experiments included the use of an Adam optimizer, a learning rate of 0.0001, a Dice-coefficient loss function, 100 epochs, and 200 steps per epoch.

3.2.5.2 U-Net++

A state-of-the-art U-Net++ architecture[20,21] was also implemented using Keras[29] with TensorFlow.[30] We used a standard ResNet-50 architecture[31] with batch normalization and a batch size of 10 as our CNN backbone, as it balanced the number of parameters and overfitting risk for the scale of our training datasets. As described in Section 3.2.5.1, the convolution operations and hyperparameters matched the modified U-Net implementation, including the use of transpose convolutions, Adam optimizer, 0.0001 learning rate, Dice-coefficient loss function, and number of epochs.

3.2.6 Evaluation and comparison

All trained models were evaluated using a testing dataset which consisted of 20 end-fire plus 20 side-fire 3D TRUS volumes unseen by the networks during training. The evaluation metrics included Dice Similarity Coefficient (DSC), recall, precision, absolute volume percent differences (VPD), mean surface distances (MSD), and Hausdorff distances (HD), computed for both 2D radial slice and reconstructed 3D segmentations for each prostate. Computation times were recorded for 2D slice segmentation, 3D reconstruction, and overall 3D segmentation time. We have previously demonstrated significantly improved performance with a 2D radial deep learning plus 3D reconstruction approach compared to fully 3D CNNs; consequently, no 3D CNNs were used for comparison in this work. A detailed list of comparisons and corresponding statistical tests is provided in Table 3.3.

Table 3.3. List of comparisons and corresponding statistical tests. Parametric tests are shown with corresponding nonparametric alternative tests presented in parentheses.

Comparison	Statistical Test
Complete Dataset (6,761 2D images)	
Modified U-Net vs U-Net++ with complete training dataset	Paired t-test (Wilcoxon matched-pairs signed-rank test)
Segmentation performance vs prostate volume	Pearson (Spearman) correlation coefficient
Split end-fire and side-fire datasets (Table 3.1)	
End-fire training dataset: U-Net vs U-Net++ tested on both end-fire and side-fire images	Paired t-test (Wilcoxon matched-pairs signed-rank test)
Side-fire training dataset: U-Net vs U-Net++ tested on both side-fire and end-fire images	| |
Split end-fire or side-fire network vs complete training dataset	| |
Smaller end-fire, side-fire, and mixed datasets (Section 3.2.2.2) *	
Performance with reduced-size datasets vs complete dataset	Repeated measured one-way ANOVA with Tukey's correction (Friedman Test with Dunn's correction)
Mixed training dataset vs end-fire or side-fire dataset of equal size	Paired t-test (Wilcoxon matched-pairs signed-rank test)
Image quality (Section 3.2.3)	
End-fire vs side-fire image quality grades for each factor	Unpaired t-test (Mann-Whitney U test)
Segmentation performance for U-Net and U-Net++ vs image quality grade in each factor and overall	One-way ANOVA with Tukey's correction (Kruskal-Wallis test with Dunn's correction)
Correlation between segmentation performance and image quality	Pearson (Spearman) correlation coefficient

* Due to superior performance of the U-Net++ when trained using split datasets, only the U-Net++ was used for reduced-size dataset experiments

Statistical calculations were performed in GraphPad Prism 9.2 (Graphpad Software, Inc., San Diego, CA, USA). The Shapiro-Wilk test was used to evaluate the normality of distributions. Failure of the Shapiro-Wilk test led to the use of nonparametric statistical tests and the reporting of median [quartile 1, quartile 3] results. The significance level for statistical analysis was chosen such that the probability of making a type I error was less than 5% ($p < 0.05$), with statistically significant differences denoted simply as "significant" for the remainder of this manuscript.

3.3 Results

3.3.1 Complete dataset

Example U-Net, U-Net++, and manual segmentations for median end-fire and side-fire cases are shown in Fig. 3.2. The evaluation metric results comparing the modified U-Net to the U-Net++ when trained using the full dataset of 6761 images are shown in Table 3.4. No significant differences were observed between the U-Net and U-Net++ for any metric for the full testing dataset. When considering the side-fire and end-fire test datasets separately, no significant differences were observed for the end-fire testing images, while only the precision and recall metrics were significantly different for side-fire testing images, with the U-Net demonstrating higher precision and the U-Net++ demonstrating higher recall. The mean computation time per 2D segmentation was 0.028 s for the modified U-Net and 0.088 s for the U-Net++. The mean 3D reconstruction time was 0.27s, resulting in a total 3D segmentation time of 0.61 s for the modified U-Net and 1.33 s for the U-Net++.

Of note, a comparison of segmentation performance relative to prostate volume for the U-Net and U-Net++ demonstrated significant correlations between prostate size and the DSC and VPD metrics. The DSC metric showed a Spearman r coefficient of 0.58 and 0.61 for the U-Net and U-Net++, respectively, while the VPD metric showed a Spearman r coefficient of -0.44 and -0.51 for the U-Net and U-Net++, respectively.

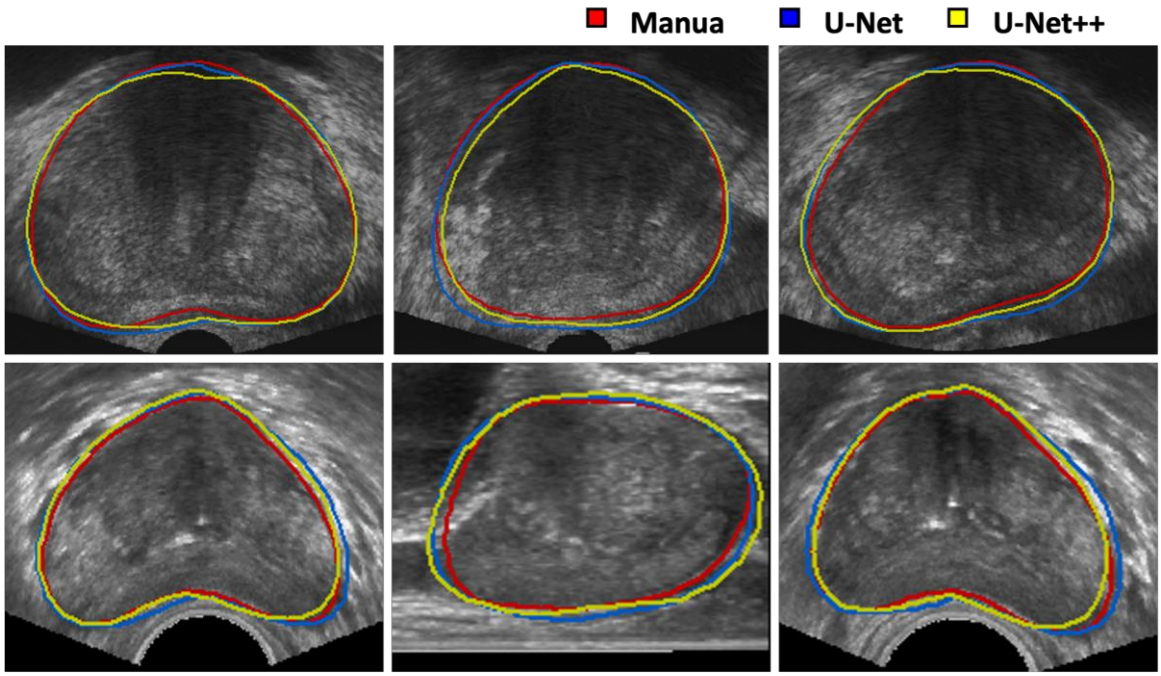

Figure 3.2. Example end-fire (top row) and side-fire (bottom row) prostate segmentation results comparing manual (red), modified U-Net (blue), and U-Net++ (yellow) 3D surfaces for median cases based on DSC. The columns from left to right show the prostate surface in the axial plane, sagittal plane, and an oblique radial plane, respectively.

Table 3.4. Median [Q1, Q3] 3D results showing the modified U-Net and U-Net++ trained using the complete dataset and tested on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS volumes.

Test Dataset	Seg. Alg.	DSC (%)	Recall (%)	Precision (%)	VPD (%)	MSD (mm)	HD (mm)
End-fire	U-Net	94.3 [93.1, 95.2]	96.0 [93.2, 98.7]	94.6 [88.8, 95.8]	5.18 [1.62, 11.2]	1.01 [0.80, 1.21]	3.44 [2.50, 4.65]
	U-Net++	94.5 [92.8, 95.5]	96.0 [93.3, 98.1]	94.1 [90.4, 96.5]	3.54 [1.80, 9.28]	0.93 [0.77, 1.20]	3.39 [2.81, 4.97]
	p-value	0.2904*	0.9932*	0.4749	0.0696	0.4284*	0.8695
Side-fire	U-Net	93.5 [91.1, 94.6]	96.2 [92.5, 98.4]	**91.6 [87.8, 94.8]**	5.89 [3.17, 11.9]	0.86 [0.71, 0.97]	2.73 [2.42, 4.28]
	U-Net++	93.3 [91.7, 95.0]	**97.7 [95.5, 98.4]**	89.8 [86.0, 94.1]	6.27 [3.89, 12.8]	0.84 [0.69, 1.03]	3.19 [2.69, 4.33]
	p-value	0.9563	0.0153	0.0495*	0.3884	0.7562	0.1536
Overall	U-Net	94.1 [92.6, 94.9]	96.0 [93.1, 98.5]	93.2 [88.8, 95.4]	5.78 [2.49, 11.5]	0.89 [0.77, 1.10]	3.14 [2.49, 4.47]
	U-Net++	94.0 [92.2, 95.1]	96.7 [94.1, 98.3]	92.9 [87.1, 95.1]	4.79 [2.93, 10.9]	0.90 [0.73, 1.14]	3.27 [2.71, 4.33]
	p-value	0.6179	0.0641	0.2265	0.6656	0.5657*	0.2317

DSC, Dice similarity coefficient; VPD, volume percent difference; MSD, mean surface dist.; HD, Hausdorff dist.
*Normal distribution = paired t-test; Bolded metrics highlight the method with significantly reduced relative error.

3.3.2 Split end-fire and side-fire datasets

Qualitative segmentation results comparing the modified U-Net and U-Net++ to manual segmentations for networks trained with only end-fire and only side-fire images are shown in Figs. 3.3 and 3.4, respectively, and the corresponding quantitative comparisons are shown in Tables 3.5 and 3.6. Plots showing DSC for the modified U-Net and U-Net++ trained using only end-fire and only side-fire datasets are shown in Fig. 3.5. For both the end-fire and side-fire networks evaluated on the complete testing dataset, which included images from both acquisition methods, the U-Net++ significantly outperformed the modified U-Net for all metrics except VPD for the side-fire networks. When evaluated on the end-fire and side-fire testing datasets separately, the U-Net++ also significantly outperformed the U-Net for all metrics aside from VPD when tested on the same image type it was trained on.

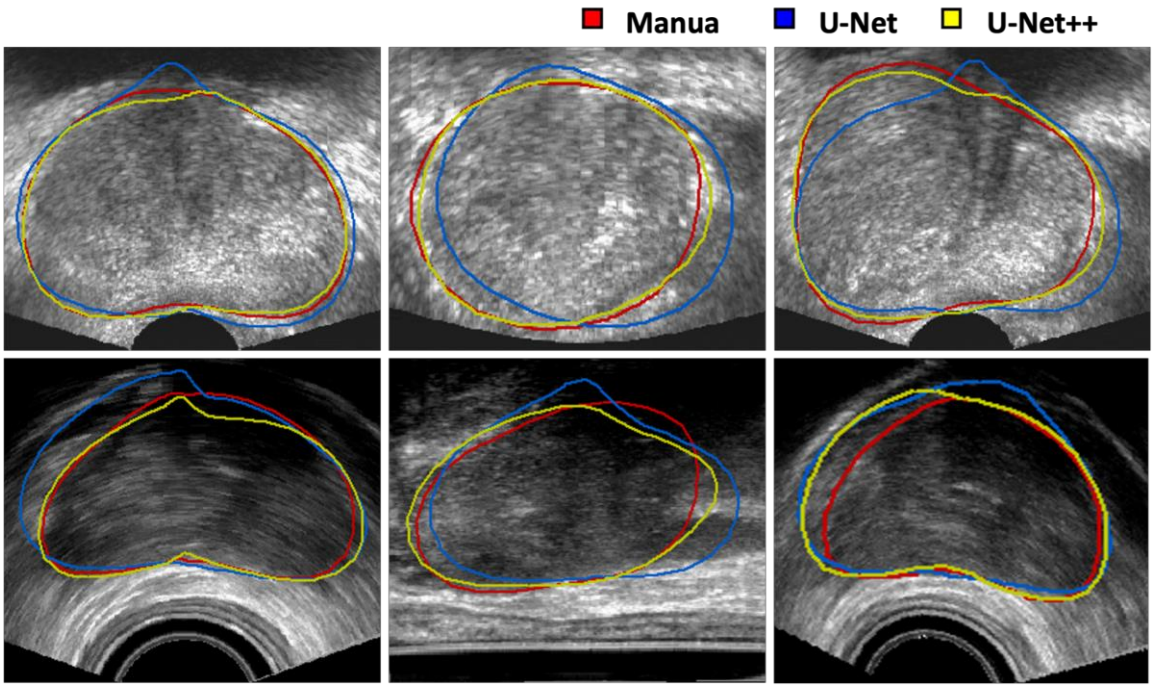

Figure 3.3. Example end-fire (top row) and side-fire (bottom row) median DSC prostate segmentation results comparing manual (red), modified U-Net (blue), and U-Net++ (yellow) 3D surfaces for networks trained only using end-fire images. The columns from left to right show the prostate surface in the axial plane, sagittal plane, and an oblique radial plane, respectively.

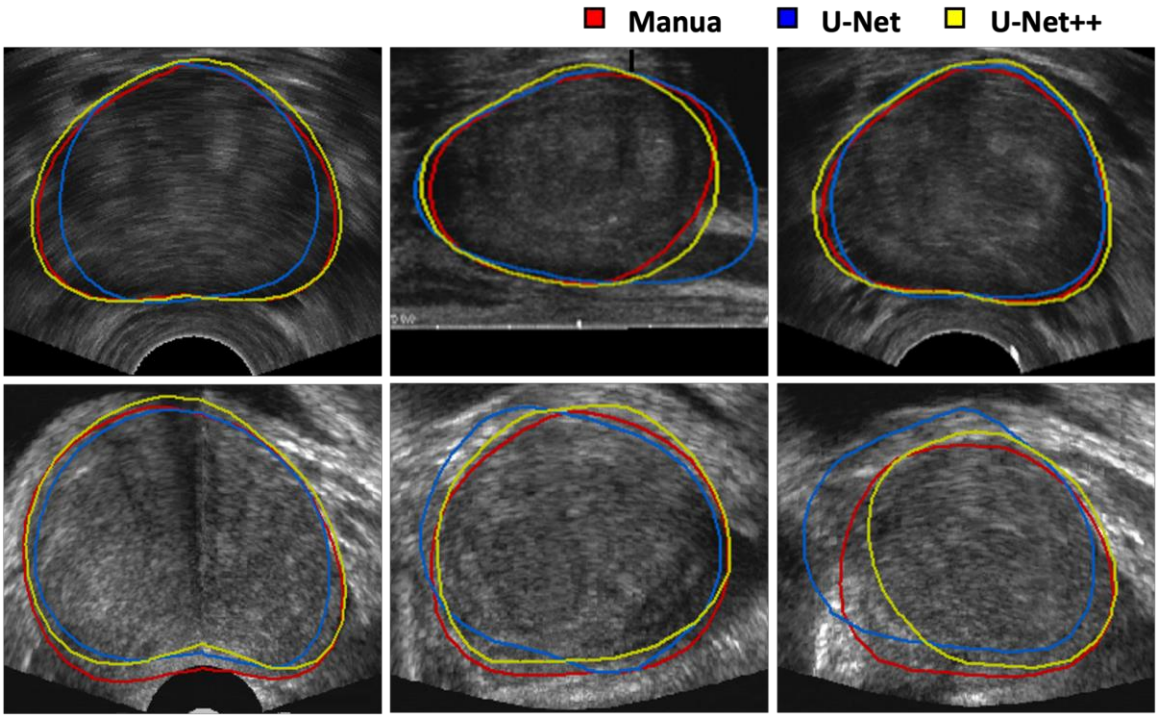

Figure 3.4. Example side-fire (top row) and end-fire (bottom row) median DSC prostate segmentation results comparing manual (red), modified U-Net (blue), and U-Net++ (yellow) 3D surfaces for networks trained only using side-fire images. The columns from left to right show the prostate surface in the axial plane, sagittal plane, and an oblique radial plane, respectively.

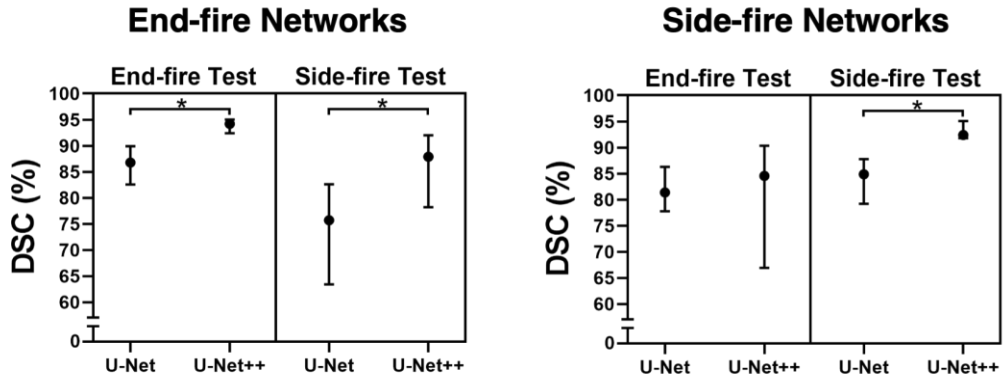

Figure 3.5. Plot of median DSC for the modified U-Net and U-Net++ trained using only end-fire (left) and only side-fire (right) images and tested on both end-fire and side-fire images shown on the left and right half of each graph, respectively.

Table 3.5. Median [Q1, Q3] 3D results showing the modified U-Net and U-Net++ trained using only end-fire images and tested on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS volumes.

Test Dataset	Seg. Alg.	DSC (%)	Recall (%)	Precision (%)	VPD (%)	MSD (mm)	HD (mm)
End-fire	U-Net	86.8 [82.6, 89.9]	89.3 [82.7, 90.8]	84.5 [80.5, 89.4]	7.80 [2.30, 12.2]	1.96 [1.63, 2.39]	6.43 [5.25, 9.11]
	U-Net++	**94.2 [92.4, 95.0]**	**95.8 [92.1, 98.3]**	**93.0 [89.5, 96.2]**	5.31 [3.56, 8.48]	**1.00 [0.85, 1.21]**	**3.78 [2.95, 5.20]**
	p-value	<0.0001	<0.0001*	<0.0001*	0.3118	<0.0001*	<0.0001*
Side-fire	U-Net	75.7 [63.4, 82.6]	91.7 [83.7, 96.0]	64.8 [50.0, 78.1]	35.9 [12.1, 91.9]	3.36 [1.97, 4.49]	12.4 [8.35, 16.1]
	U-Net++	**87.9 [78.2, 92.0]**	93.1 [89.8, 96.6]	**84.0 [71.6, 90.5]**	**18.2 [5.40, 34.9]**	**1.75 [1.17, 2.89]**	**10.0 [5.89, 11.4]**
	p-value	<0.0001	0.2611	<0.0001*	0.0005	0.0003	0.0363*
Overall	U-Net	82.6 [75.5, 87.4]	89.5 [83.7, 94.6]	78.9 [64.2, 86.4]	12.0 [4.24, 37.3]	2.19 [1.69, 3.43]	8.92 [6.15, 12.7]
	U-Net++	**92.3 [87.7, 94.3]**	**93.6 [91.2, 97.4]**	**89.8 [82.9, 95.0]**	**6.84 [4.35, 18.2]**	**1.18 [0.90, 1.78]**	**5.53 [3.61, 10.1]**
	p-value	<0.0001	0.0002	<0.0001	0.0006	<0.0001	<0.0001

DSC, Dice similarity coefficient; VPD, volume percent difference; MSD, mean surface dist.; HD, Hausdorff dist.
*Normal distribution = paired t-test; Bolded metrics highlight the method with significantly reduced relative error.

Comparing the results shown in Tables 3.4 and 3.5 of the U-Net++ trained with the full 6761 image dataset to the U-Net++ trained using only end-fire images, use of the full dataset only demonstrated a significant improvement for the HD metric (0.4 mm) when tested on end-fire images. Similarly, comparing the U-Net++ trained with the full dataset to one trained using only side-fire images, only the precision metric was significantly different (1.8%) when tested on side-fire images. When tested on the unseen acquisition type, the use of the full dataset demonstrated improved performance for every metric in both cases.

3.3.3 Smaller end-fire, side-fire, and mixed datasets

Example segmentation results for the U-Net++ trained with varying-sized end-fire, side-fire, and mixed datasets are shown in Fig. 3.6. Plots of DSC as a function of training dataset size are shown in Fig. 3.7, highlighting the high performance of the U-Net++ when trained with small datasets. Corresponding quantitative comparisons are provided in Supplement B Tables 3.B1, 3.B2, and 3.B3.

Table 3.6. Median [Q1, Q3] 3D results showing the modified U-Net and U-Net++ trained using only side-fire images and tested on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS volumes.

Test Dataset	Seg. Alg.	DSC (%)	Recall (%)	Precision (%)	VPD (%)	MSD (mm)	HD (mm)
End-fire	U-Net	81.4 [77.8, 86.3]	74.4 [68.3, 85.3]	90.2 [86.3, 93.9]	**17.4 [9.87, 26.3]**	2.48 [1.79, 3.38]	8.57 [6.80, 11.8]
	U-Net++	84.5 [66.9, 90.3]	75.3 [51.0, 86.9]	**96.5 [95.7, 99.2]**	23.3 [13.0, 48.3]	2.44 [1.50, 4.72]	8.90 [7.18, 13.8]
	p-value	0.9854	0.2679*	<0.0001	0.0121	0.5459	0.2436*
Side-fire	U-Net	84.9 [79.2, 87.8]	85.9 [78.7, 92.3]	82.2 [77.5, 85.9]	5.16 [4.13, 13.0]	1.81 [1.43, 2.37]	7.77 [6.29, 8.68]
	U-Net++	**92.4 [91.8, 95.1]**	**95.2 [91.8, 97.9]**	**91.6 [87.2, 95.5]**	7.11 [3.34, 12.4]	**0.95 [0.66, 1.05]**	**3.21 [2.58, 4.75]**
	p-value	<0.0001	<0.0001*	<0.0001*	0.4091	<0.0001	<0.0001
Overall	U-Net	82.8 [78.4, 87.4]	80.4 [71.4, 87.6]	86.3 [80.7, 90.6]	11.9 [5.03, 20.4]	2.26 [1.54, 2.83]	7.98 [6.62, 9.60]
	U-Net++	**91.5 [84.3, 92.7]**	**89.3 [75.3, 96.0]**	**95.6 [88.0, 96.9]**	13.0 [6.78, 23.4]	**1.25 [0.95, 2.51]**	**5.46 [3.18, 9.48]**
	p-value	0.0035	0.0394	<0.0001	0.0972	0.0192	0.0436

DSC, Dice similarity coefficient; VPD, volume percent difference; MSD, mean surface dist.; HD, Hausdorff dist.
*Normal distribution = paired t-test; Bolded metrics highlight the method with significantly reduced relative error.

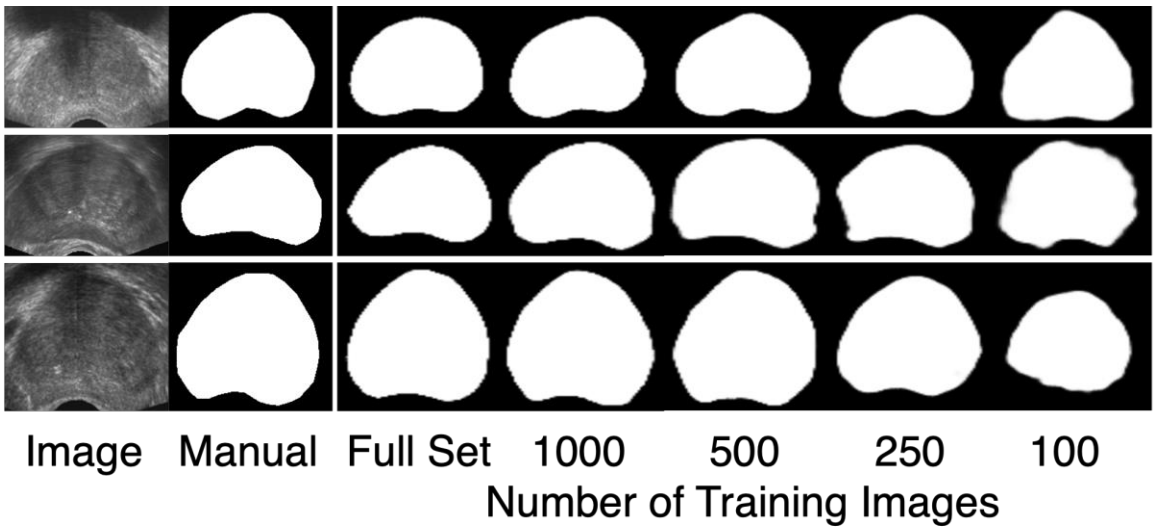

Figure 3.6. Example results for the U-Net++ trained with varying-size end-fire (top row), side-fire (middle row), and mixed (bottom row) datasets. The full dataset corresponds to 2738 images for the end-fire network, 4023 images for the side-fire network, and 6761 for the mixed network.

As shown in Table 3.B1, for the U-Net++ trained with reduced-size end-fire datasets and tested on end-fire images, significant differences were observed between the full (2738 images) end-fire training dataset and the 250 and 100 image sets for the DSC and MSD metric, and all reduced-size image sets for the HD metric.

When the U-Net++ was trained with reduced-size side-fire datasets and tested on side-fire images (Table 3.B2), multiple comparisons tests showed significant differences for the DSC, MSD, and HD metrics between the full (4023 images) side-fire training dataset and the 500 and 100 image sets.

As shown in Table 3.B3, for the U-Net++ trained with reduced-size mixed datasets and tested on end-fire images, multiple comparisons tests showed significant differences between the full (6761 images) mixed training dataset and the 500, 250, and 100 image sets for the DSC, MSD, and HD metrics. When tested on side-fire images significant differences were observed between the full mixed training dataset and the 1000 through 100 image sets for the DSC, MSD, and HD metrics.

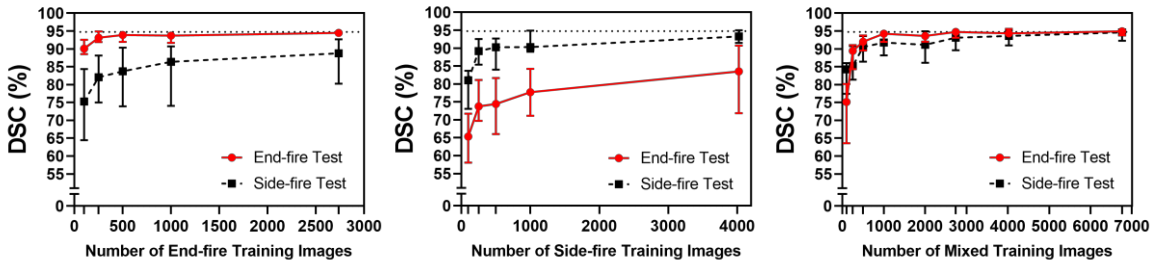

Figure 3.7. Median DSC for the U-Net++ trained using varying-sized end-fire (left), side-fire (middle), and mixed (right) datasets, and tested on both end-fire (solid red) and side-fire (dashed line) images. The number of training images are provided along the x-axis. The dotted line denotes the maximum performance achieved by the U-Net++ trained with the full 6761 image dataset.

Comparing the U-Net++ trained with 2738 mixed images to the U-Net++ trained with 2738 end-fire images, no significant differences were observed when tested on end-fire images, but when tested on side-fire images, use of the mixed training dataset demonstrated significantly improved performance for all metrics. Similarly, for the U-

Net++ trained with 4023 mixed images compared to the network trained with 4023 side-fire images, only the precision and recall metrics were significantly different when tested on side-fire images, with all metrics except precision significantly improved with use of the mixed training dataset when tested on end-fire images.

3.3.4 Image quality

A comparison of average image quality grading results for side-fire and end-fire 3D TRUS images of the prostate is shown in Table 3.7. There were no significant differences between end-fire and side-fire image quality for any image quality factor or for the total averaged image quality.

Table 3.7. Mean plus/minus standard deviation image quality grading results for side-fire and end-fire 3D TRUS images of the prostate. Mean grades are provided out of five for each individual image quality factor and for the total image quality.

Image Quality Factor	Description	End-fire 3D TRUS	Side-fire 3D TRUS
Acquisition Quality	Quality of the 3D TRUS image acquisition regardless of anatomy	4.4 ± 0.7	4.2 ± 1.2
Anatomy Artifacts	Severity of anatomy artifacts (calcification, gas, catheter, etc.)	4.0 ± 0.7	4.1 ± 1.0
Prostate Boundary	Visibility/clarity of the prostate boundary	4.2 ± 0.7	4.2 ± 1.0
Total	Mean of three individual image quality factors	4.2 ± 0.5	4.1 ± 0.9

A graph of DSC as a factor of grade for each individual image quality factor is shown in Fig. 3.8. For end-fire testing images, image quality grade did not have a significant effect on segmentation performance in any metric. For side-fire testing images, only the boundary visibility grade had a significant effect for the modified U-Net, while all image quality factors except anatomy artifact grade had a significant effect on the DSC metric for the U-Net++.

Analysis of plots of DSC as a function of total image quality grade for the U-Net and U-Net++ (Fig. 3.9) showed no significant correlation for the end-fire testing dataset for any metric, with Spearman r coefficients less than 0.4. For the side-fire testing images, the modified U-Net showed a significant correlation between total image quality grade and DSC, recall, and HD metrics, with Spearman r coefficients of 0.60, 0.61, and -0.56, respectively, while the U-Net++ showed a significant correlation for the DSC and recall metrics with Spearman r coefficients of 0.46 and 0.55, respectively.

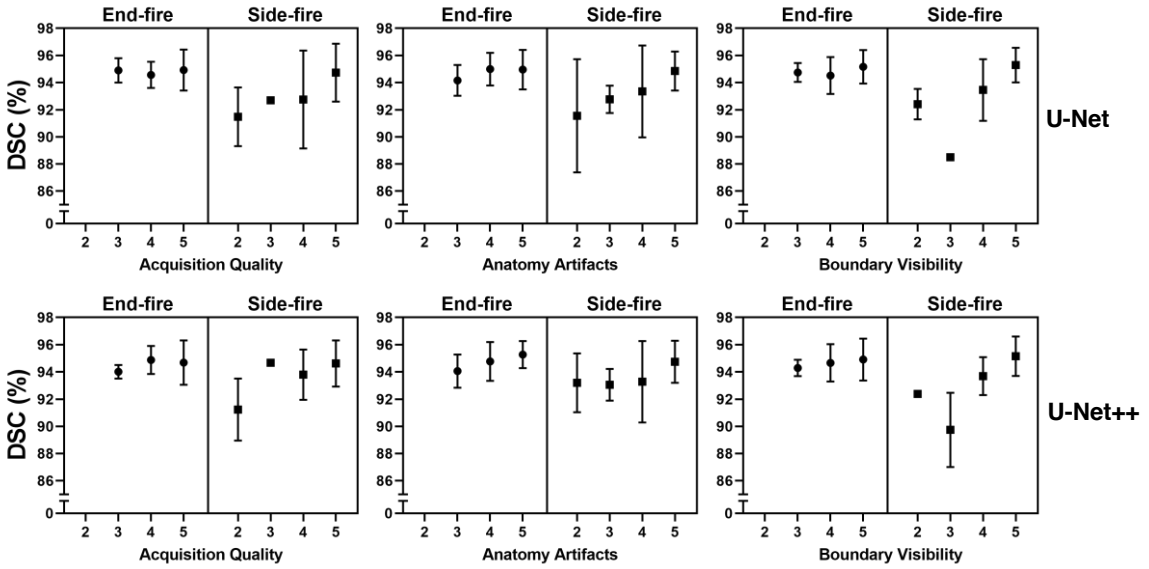

Figure 3.8. Mean DSC as a factor of acquisition quality (left), anatomy artifacts severity (middle), and boundary visibility (right) for the modified U-Net (top row) and U-Net++ (bottom row) trained on the full dataset and tested on end-fire and side-fire images shown on the left and right side of each graph, respectively.

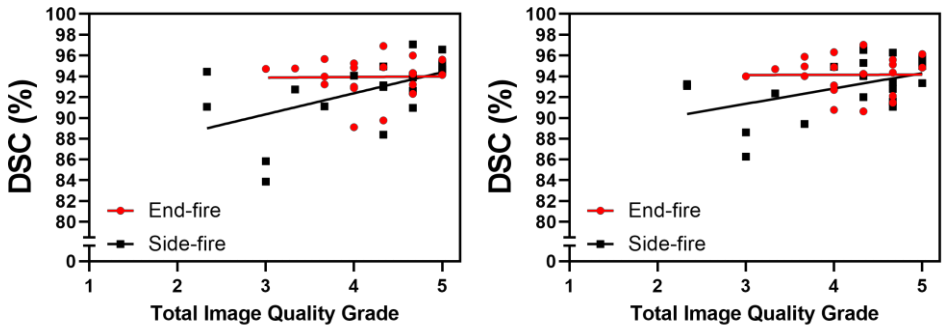

Figure 3.9. DSC as a function of total averaged image quality for the modified U-Net (left) and U-Net++ (right) trained on the full dataset and tested on end-fire (red) and side-fire (black) images. A linear regression is provided for each dataset.

3.4 Discussion

3.4.1 Complete dataset

To provide a baseline maximum performance level, we first compared the segmentation accuracy of the modified U-Net to the U-Net++ for both networks trained on the complete dataset. The results shown in Table 3.4 demonstrate the nearly equivalent performance of the networks. This highlights that with a large training dataset of nearly 7000 2D images, the more advanced U-Net++ network with significantly more parameters did not offer any improvement in performance, motivating the experiments described in sections 3.3.2 and 3.3.3 focused on reduced-size datasets. Using the same 24 GB NVIDIA TITAN RTX GPU, the modified U-Net demonstrated a segmentation time that was three times faster, with speeds of 0.028 s per 2D slice compared to 0.088 s per slice for the U-Net++. After reconstruction of the 2D predictions into a 3D prostate surface, the total segmentation time was 0.61 s for the modified U-Net, which was half of the 1.33 s for the U-Net++. While this is a large relative difference, in a clinical setting the difference is inconsequential, as both present a significant reduction in segmentation time relative to manual segmentations, which can take 10-20 minutes.

Correlations between segmentation performance and prostate size were only significant for the DSC and VPD metrics. This is an expected result due to the nature of these metrics, as absolute differences that would be readily apparent for smaller prostate volumes would be reduced for large volumes when considering these overlap and volume-based metrics. As expected, boundary-based metrics showed no correlation with prostate size. The correlations we did observe were still weak, however, with Spearman coefficients of roughly r = 0.6 for DSC and r = -0.4 to -0.5 for VPD, highlighting the general robustness of our approach to prostate size differences.

Recent work by Lei *et al.* and van Sloun *et al.* report state-of-the-art methods for automatic prostate segmentation in 3D TRUS volumes. Lei *et al.* report best results for DSC, MSD, and HD_{95} of 93.0 ± 3.0, 0.57 ± 0.20 mm, and 2.27 ± 0.79 mm, respectively.[17] Our proposed 3D segmentation method using the modified U-Net and the U-Net++ demonstrated median [Q1, Q3] DSC of 94.1 [92.6, 94.9] % and 94.0 [92.2, 95.1] %, showing improved performance. Our proposed networks demonstrated median MSD [Q1, Q3] of 0.89 [0.77, 1.10] mm and 0.90 [0.73, 1.14] mm, with Lei *et al.* showing improved performance for this metric. We do not report an HD_{95} metric, but our networks demonstrated a median HD [Q1, Q3] of 3.14 [2.49, 4.47] mm and 3.27 [2.71, 4.33], which while higher, is a more stringent metric showing the full HD, not excluding the worst five percent. van Sloun *et al.* report a median (95% confidence interval) accuracy of 98 (95-99)%, Jaccard index of 93.0 (80.0-96.0), and HD of 3.0 (1.3-8.7) mm.[18] While not reported, our modified U-Net and U-Net++ demonstrated a median [Q1, Q3] accuracy of 98.8 [98.4, 99.2] % and 98.9 [98.4, 99.2] %, and a median HD [Q1, Q3] of 3.14 [2.49, 4.47] mm and 3.27 [2.71, 4.33], both of which are similar, with our approach reporting higher accuracy and van Sloun *et al.* reporting lower HD. Our results were obtained with considerably less data, however, as van Sloun *et al.* utilized a training dataset of 436 3D TRUS volumes compared to our complete training dataset of 206 3D TRUS volumes.

3.4.2 Split end-fire and side-fire datasets

Segmentation performance of the modified U-Net and U-Net++ trained with only end-fire or only side-fire images (Fig. 3.5 and Tables 3.5 and 3.6) showed that the U-Net++ significantly outperformed the modified U-Net in nearly all cases. When trained using side-fire images and tested on end-fire images, no difference was observed, but the U-Net++ did have higher median performance, countered by a larger variation. These differences highlight the generalizability and efficiency of the U-Net++ in utilizing small training datasets. The modified U-Net had boundary errors due to shadowing artifacts, even when tested on the same image type as seen in the top row of Fig. 3.3. When tested on the image type not seen during training of the network, the U-Net++ still performed better, although it also had difficulties with shadowing artifacts (e.g. the bottom row of Fig. 3.3, with the heavily shadowed region seen near the top of the prostate). The modified U-Net had a depth of five layers compared to 50 for the U-Net++. This reduction in depth and number of parameters for the U-Net compared to the U-Net++ may alleviate the overfitting problem, which is important as training dataset size is reduced.

When assessing how the U-Net++ trained with only end-fire or only side-fire images compared to one trained with the full dataset, we found that there was little difference when tested on the same TRUS acquisition type the networks were trained with. This highlights a potentially practical finding that the presence of other image types in the training dataset do not add a significant benefit to the segmentation performance when only one image type is required to be segmented. However, when the U-Net++ trained with only end-fire or only side-fire images were tested on the TRUS acquisition type they had never seen before, use of the full dataset significantly improved performance. This demonstrates the necessity of including all image types in the training dataset, especially when generalizability and widespread application is important. DSC performance in these cases was still in the range of 85-89% for the U-Net++, however, demonstrating the generalizability of our approach.

Differences between end-fire and side-fire images, including image quality and artifact prevalence, may explain the observed segmentation performance differences between TRUS image types. The differences in acquisition method between end-fire and

side-fire 3D TRUS may result in artifacts in side-fire images such as air gaps due to lack of transducer contact or distal shadowing due to transducer distance from the prostate. Due to the nature of end-fire image acquisition, the radial plane used for deep learning segmentation matches closely the acquisition plane, resulting in improved segmentation accuracy. For side-fire images, only one of the twelve radial planes is the acquisition plane and the other eleven are interpolated slices resulting in reduced resolution, potentially explaining some of the observed differences in segmentation performance. In HDR brachytherapy procedures where side-fire 3D TRUS is utilized, urinary catheters are commonly used, which create artifacts that are not seen in end-fire images used for prostate biopsy. The appearance of other organs such as the rectum and bladder also differ between end-fire and side-fire leading to increased prostate segmentation error where the algorithm included parts of the rectum or bladder when tested on the 3D TRUS type unseen by the network. Furthermore, due to differences in patient selection and the prevalence of hormone therapy prior to HDR brachytherapy treatment, the prostate sizes in patients presenting for end-fire TRUS-guided biopsy are typically larger than the prostate sizes of patients undergoing side-fire TRUS-guided HDR brachytherapy. This led to underpredictions for side-fire networks tested on end-fire images and overpredictions for end-fire networks tested on side-fire images, limiting generalizability and necessitating the presence of both 3D TRUS types in the training dataset so the network can learn differences in size and shape.

3.4.3 Smaller end-fire, side-fire, and mixed datasets

For as small as 500 end-fire images used in the training dataset, which is just over 7% of the full dataset, DSC performance was within 1% of the U-Net++ trained with the full dataset of 6761 images. Results were similar for the U-Net++ trained with reduced-size side-fire datasets and tested on side-fire images. Networks trained with end-fire images performed better when tested on side-fire images compared to networks trained with side-fire images and tested on end-fire images, suggesting features the network learns from end-fire images are more generalizable to side-fire images. As expected, mixed training datasets had high segmentation performance when tested on both image types even as the dataset

size was reduced. This improved performance and generalizability is apparent in Fig. 3.7, highlighting the benefit of including all image types in the training dataset.

For a segmentation task involving only one image type, performance plateaus at a training dataset size of 1000 2D training images of that type, which were obtained from approximately 36 3D volumes. A dataset of this size is achievable at even small hospitals or research centers, showing that deep learning segmentation in 3D may be possible even with limited data by utilizing organ symmetry and a radial approach. The reduced training data requirement reduces the amount of manual segmentation required, a key benefit as accurate manual segmentation is a difficult and time-consuming process that is often a bottleneck in supervised machine learning. These results also show that for a segmentation task involving multiple image types, the presence of all image types in the training dataset is critical. Segmentation performance for mixed training datasets also plateaus at approximately 1000 training images, suggesting that deep learning segmentation in two image types is possible even if data is scarce.

3.4.4 Image quality

We developed a 5-point image quality grading scale based on three factors specifically for 3D TRUS prostate images. This grading scale helps provide transparency regarding the image quality of our clinical dataset, helping to contextualize our results. A numerical scale with clearly defined image quality factors rated from one to five may enable an easier comparison of segmentation performance between networks trained using different datasets. Designing the image quality grading scale independently of our dataset should allow it to be successfully applied to 3D TRUS datasets of varying quality.

Mean image quality grades for each individual factor provided in Table 3.7 highlight the overall high quality of our dataset and the general similarity in image quality between end-fire and side-fire images, with no statistically significant differences observed and a maximum difference in mean of only 0.2. Side-fire images did have an increased standard deviation for each individual factor, highlighting the larger range of image qualities, including the presence of grades of 2 in each factor, which was not seen in the end-fire images. Our dataset contained no images with a grade of 1.

For end-fire images in our testing dataset, image quality had no significant effect on segmentation performance. In contrast, for side-fire images in our testing dataset, the boundary visibility grade and the acquisition quality, boundary visibility, and total averaged image quality grades significantly impacted segmentation performance for the U-Net and U-Net++, respectively. Boundary visibility showed to be a key factor in the algorithm's ability to accurately segment the prostate boundary for both networks, as expected. These results were further confirmed with the correlation analysis shown in Fig. 3.9, highlighting the significant effect of image quality on segmentation performance for side-fire images, but not for end-fire images. Correlations were not strong with Spearman r coefficients in the range of 0.46-0.6 for the DSC metric for both the U-Net and U-Net++.

The lack of significant differences observed when comparing how segmentation performance varies with image quality, especially considering the end-fire images, may be attributed to the high mean image quality and subtle variation between the poorest quality image and the highest quality image. A dataset with more variation in image quality may better demonstrate the dependence of segmentation performance on image quality. In addition, due to the testing set size of 20 end-fire and 20 side-fire 3D TRUS volumes, some individual image quality grades had a very small sample size, which likely factored into the lack of significant differences observed for some of the image quality factors.

The differences in image quality and its effect on segmentation performance for end-fire compared to side-fire images may be explained in part due to the nature of image acquisition. Ultrasound transducer orientation is one critical component; during end-fire image acquisition the transducer contacts the rectal wall at the end of the transducer pointing towards the prostate. During side-fire acquisition, however, the transducer is positioned horizontally inside the rectum, requiring a much larger contact area, which can result in increased prevalence of air gaps due to lost contact, reducing image quality. Furthermore, due to differences in transducer position based on the intended application the side-fire transducer is further away from the prostate, leading to hypoechoic regions away from the transducer due to issues with time-gain compensation.

3.4.5 Limitations and future work

Only one observer provided manual gold standard segmentations, thus inter- and intra-observer variability were not directly assessed; however, these considerations were addressed in Orlando *et al*.[19] In addition, only one observer defined the image quality grading scale and graded the testing dataset, which did not assess the impact of inter- or intra-observer variability. Future work will include validation of our image quality grading scale and its reliability, including an assessment of inter- and intra-observer variability. Image quality of the training dataset may play a critical role in segmentation performance, and although image quality of the testing dataset should have been representative of the training dataset, direct grading of the training images would allow for confirmation of this assumption. As shown in Table 3.7, our 3D TRUS dataset was of high quality on average. A wide range in image quality is important for algorithm generalizability. Future work should investigate our segmentation approach when trained and tested with a lower quality dataset, ideally from a different center.

Patient clinical information, such as age, stage of prostate cancer, and Gleason score was not recorded for our dataset, and thus an assessment of how segmentation performance is impacted by these measures could not be completed. While this has not been assessed in previous work to our knowledge, differences in such measures could manifest as differences in image quality, potentially captured by our image quality grading scale as anatomy artifact severity for example. Future work could explicitly investigate the influence of patient clinical information on segmentation quality.

For our U-Net++ implementation, only one type of CNN backbone (ResNet) was used. Future work will utilize a U-Net++ ensemble network with results from multiple CNN backbones combined into one segmentation result using a method such as averaging, majority vote, or the STAPLE algorithm.[32] Finally, a leave-one-out vendor study examining the impact of ultrasound machine vendor on segmentation performance would offer a strong assessment of generalizability, which is critical for widespread clinical translation.

3.5 Conclusions

This study investigated the effect of training dataset size, image quality, and image type on prostate segmentation in 3D TRUS volumes using a 2D radial plus 3D reconstruction approach, comparing a modified U-Net to a U-Net++ architecture. Beginning with a large, clinically diverse dataset of TRUS images, smaller training datasets were generated by splitting and reducing the dataset. Segmentation performance for the U-Net++ plateaued at end-fire, side-fire, or mixed training dataset sizes of 1000 2D images, resliced from approximately 36 3D volumes. This high performance with small datasets highlights the potential for widespread use of our approach or similar methods, even if data is scarce, demonstrating the possibility for increased access to automated segmentation methods. The development of an image quality grading scale specifically for 3D TRUS imaging provides a quantitative tool for assessing segmentation performance, with an aim to increase transparency regarding dataset quality and aid in comparison between segmentation methods trained using different datasets.

3.6 References

1. Leite KRM, Camara-Lopes LHA, Dall'Oglio MF, et al. Upgrading the Gleason Score in Extended Prostate Biopsy: Implications for Treatment Choice. *Int J Radiat Oncol Biol Phys*. 2009;73(2):353-356. doi:10.1016/j.ijrobp.2008.04.039

2. Cool DW, Romagnoli C, Izawa JI, et al. Comparison of prostate MRI-3D transrectal ultrasound fusion biopsy for first-time and repeat biopsy patients with previous atypical small acinar proliferation. *Can Urol Assoc J*. 2016;10(9-10):342-348. doi:10.5489/cuaj.3831

3. Yamada Y, Rogers L, Demanes DJ, et al. American Brachytherapy Society consensus guidelines for high-dose-rate prostate brachytherapy. *Brachytherapy*. 2012;11(1):20-32. doi:10.1016/J.BRACHY.2011.09.008

4. Yoshioka Y, Suzuki O, Isohashi F, et al. High-Dose-Rate Brachytherapy as Monotherapy for Intermediate- and High-Risk Prostate Cancer: Clinical Results for a Median 8-Year Follow-Up. *Int J Radiat Oncol*. 2016;94(4):675-682. doi:10.1016/J.IJROBP.2015.05.044

5. Hrinivich WT, Hoover DA, Surry K, et al. Three-dimensional transrectal ultrasound

guided high-dose-rate prostate brachytherapy: A comparison of needle segmentation accuracy with two-dimensional image guidance. *Brachytherapy*. 2016;15(2):231-239. doi:10.1016/j.brachy.2015.12.005

6. Reich DL, Hossain S, Krol M, et al. Predictors of Hypotension After Induction of General Anesthesia. *Anesth Analg*. 2005;101(3):622-628. doi:10.1213/01.ANE.0000175214.38450.91

7. Kasivisvanathan V, Rannikko AS, Borghi M, et al. MRI-Targeted or Standard Biopsy for Prostate-Cancer Diagnosis. *N Engl J Med*. 2018;378:1767-1777. doi: 10.1056/NEJMoa1801993

8. Eklund M, Jäderling F, Discacciati A, et al. MRI-Targeted or Standard Biopsy in Prostate Cancer Screening. *N Engl J Med*. 2021;385(10):908-920. doi:10.1056/NEJMOA2100852

9. Alayed Y, D'Alimonte L, Helou J, et al. MRI assisted focal boost integrated with HDR monotherapy study in low and intermediate risk prostate cancer (MARS): Results from a phase II clinical trial. *Radiother Oncol*. 2019;141:144-148. doi:10.1016/j.radonc.2019.09.011

10. Holly R, Morton GC, Sankreacha R, et al. Use of cone-beam imaging to correct for catheter displacement in high dose-rate prostate brachytherapy. *Brachytherapy*. 2011;10(4):299-305. doi:10.1016/J.BRACHY.2010.11.007

11. Piccialli F, Somma V Di, Giampaolo F, Cuomo S, Fortino G. A survey on deep learning in medicine: Why, how and when? *Inf Fusion*. 2021;66:111-137. doi:10.1016/j.inffus.2020.09.006

12. Anas EMA, Mousavi P, Abolmaesumi P. A deep learning approach for real time prostate segmentation in freehand ultrasound guided biopsy. *Med Image Anal*. 2018;48:107-116. doi:10.1016/j.media.2018.05.010

13. Ghavami N, Hu Y, Bonmati E, et al. Integration of spatial information in convolutional neural networks for automatic segmentation of intraoperative transrectal ultrasound images "Integration of spatial information in convolutional neural networks for automatic segmentation of intraoperative transrectal ultrasound images Integration of spatial information in convolutional neural networks for automatic segmentation of intraoperative transrectal ultrasound images. *J Med Imag*. 2018;6(1):11003. doi:10.1117/1.JMI.6.1.011003

14. Lei Y, Tian S, He X, et al. Ultrasound Prostate Segmentation Based on Multi-Directional Deeply Supervised V-Net. *Med Phys*. 2019;46(7):3194-3206. doi:10.1002/mp.13577

15. Karimi D, Zeng Q, Mathur P, et al. Accurate and robust deep learning-based segmentation of the prostate clinical target volume in ultrasound images. *Med Image Anal*. 2019;57:186-196. doi:10.1016/j.media.2019.07.005

16. Wang Y, Ni D, Dou H, et al. Deep Attentive Features for Prostate Segmentation in 3D Transrectal Ultrasound. *IEEE Trans Med Imaging*. 2019;38(12):2768-2778. doi:10.1109/TMI.2019.2913184

17. Lei Y, Wang T, Roper J, et al. Male pelvic multi-organ segmentation on transrectal ultrasound using anchor-free mask CNN. *Med Phys*. 2021;48(6):3055-3064. doi:10.1002/mp.14895

18. van Sloun RJG, Wildeboer RR, Mannaerts CK, et al. Deep Learning for Real-time, Automatic, and Scanner-adapted Prostate (Zone) Segmentation of Transrectal Ultrasound, for Example, Magnetic Resonance Imaging–transrectal Ultrasound Fusion Prostate Biopsy. *Eur Urol Focus*. 2021;7(1):78-85. doi:10.1016/j.euf.2019.04.009

19. Orlando N, Gillies DJ, Gyacskov I, Romagnoli C, D'Souza D, Fenster A. Automatic prostate segmentation using deep learning on clinically diverse 3D transrectal ultrasound images. *Med Phys*. 2020;47(6):2413-2426. doi:10.1002/mp.14134

20. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2018;11045:3-11. http://arxiv.org/abs/1807.10165

21. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation. *IEEE Trans Med Imaging*. 2020;39(6):1856-1867. doi:10.1109/TMI.2019.2959609

22. Fenster A, Parraga G, Bax J. Three-dimensional ultrasound scanning. *Interface Focus*. 2011;1(4):503-519. doi:10.1098/rsfs.2011.0019

23. Tong S, Downey DB, Cardinal HN, Fenster A. A three-dimensional ultrasound prostate imaging system. *Ultrasound Med Biol*. 1996;22(6):735-746. doi:10.1016/0301-5629(96)00079-8

24. Bax J, Smith D, Bartha L, et al. A compact mechatronic system for 3D ultrasound guided prostate interventions. *Med Phys*. 2011;38(2):1055-1069. doi:10.1118/1.3531540

25. Bax J, Cool D, Gardi L, et al. Mechanically assisted 3D ultrasound guided prostate biopsy system. *Med Phys*. 2008;35(12):5397. doi:10.1118/1.3002415

26. Milletari F, Navab N, Ahmadi S-A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. *International Conference on 3D Vision*. 2016:565-571. doi:10.1109/3DV.2016.79

27. Qiu W, Yuan J, Ukwatta E, Fenster A. Rotationally resliced 3D prostate TRUS segmentation using convex optimization with shape priors. *Med Phys*. 2015;42(2):877-891. doi:10.1118/1.4906129

28. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2015;9351:234-241. doi:10.1007/978-3-319-24574-4_28

29. Chollet F and others. Keras. 2015. http://keras.io/.

30. Abadi M, Agarwal A, Barham P, et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. 2016; arXiv Prepr. arXiv1603.04467.

31. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Vol 2016-December. IEEE Computer Society; 2016:770-778. doi:10.1109/CVPR.2016.90

32. Warfield SK, Zou KH, Wells WM. Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation. *IEEE Trans Med Imaging*. 2004;23(7):903-921. doi:10.1109/TMI.2004.828354

3.7 Supplement A

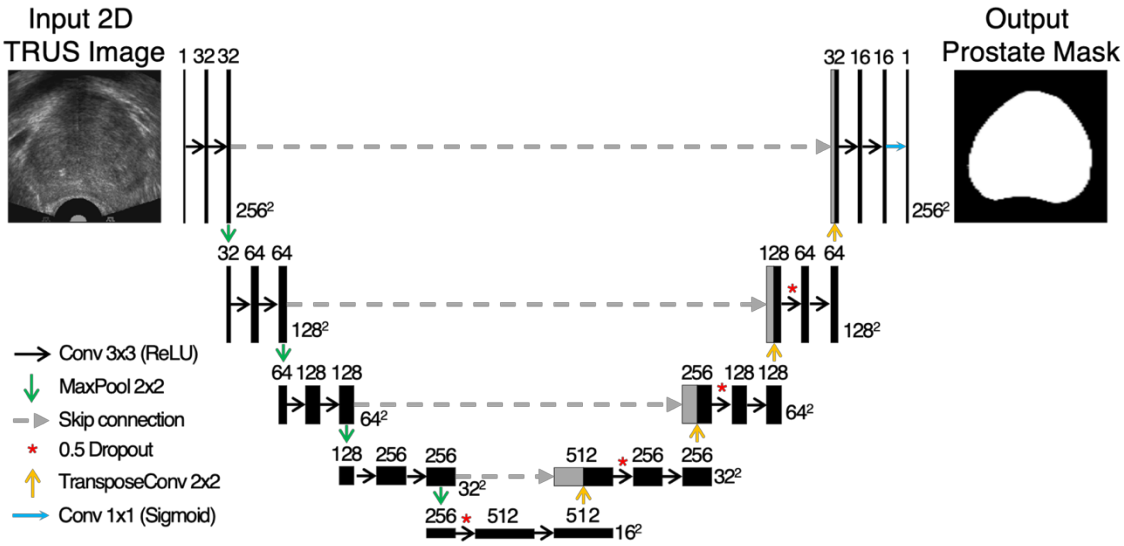

Figure 3.A1. Network diagram for the modified U-Net.

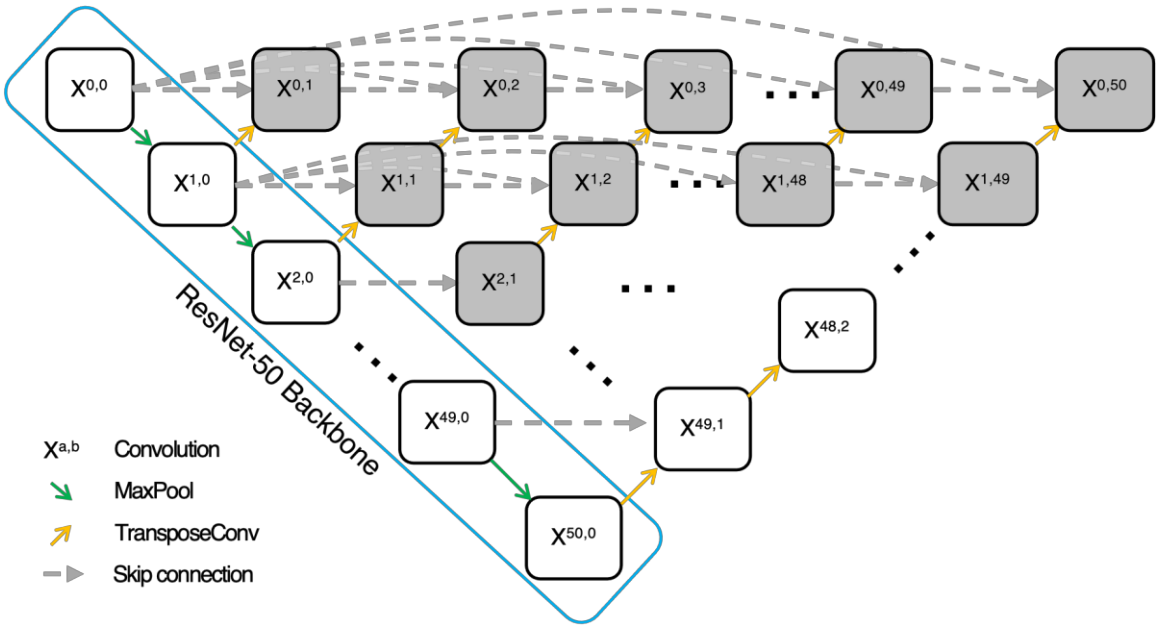

Figure 3.A2. Network diagram for the U-Net++ based on Zhou et al. 2018.

3.8 Supplement B

Table 3.B1. Median [Q1, Q3] 2D results for the U-Net++ trained using end-fire datasets of varying size, from 2738 (full end-fire set) to 100 images. The networks were evaluated on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate.

Test Dataset	End-fire Training Images	DSC (%)	Recall (%)	Precision (%)	MSD (mm)	HD (mm)
End-fire	2738	94.5 [93.5, 95.1]	93.7 [91.0, 95.9]	95.8 [93.6, 97.4]	1.18 [0.98, 1.42]	3.79 [2.97, 4.58]
	1000	93.7 [91.7, 94.8]	93.8 [89.8, 95.3]	95.9 [94.0, 97.2]	1.35 [1.20, 1.97]	**4.76 [3.89, 6.19]**
	500	93.9 [92.0, 94.8]	95.1 [90.3, 96.2]	**94.5 [91.8, 97.1]**	1.38 [1.06, 1.96]	**4.39 [3.54, 5.60]**
	250	**93.2 [91.9, 94.9]**	93.1 [90.1, 95.8]	95.4 [93.1, 97.2]	**1.52 [1.13, 1.70]**	**4.66 [3.86, 5.67]**
	100	**90.1 [88.5, 92.6]**	**87.5 [84.0, 92.2]**	95.1 [92.0, 96.8]	**2.04 [1.46, 2.91]**	**6.24 [4.55, 7.60]**
Side-fire	2738	88.8 [80.3, 92.7]	89.7 [84.7, 93.4]	90.8 [82.4, 94.5]	2.00 [1.46, 3.64]	6.38 [5.01, 11.1]
	1000	86.4 [74.1, 90.7]	92.6 [89.2, 95.2]	**83.2 [63.2, 91.2]**	**3.81 [1.96, 6.17]**	**12.4 [6.72, 17.5]**
	500	**83.8 [73.9, 90.4]**	**82.1 [67.1, 90.9]**	88.9 [81.8, 94.8]	**2.89 [1.89, 4.89]**	9.83 [6.14, 13.0]
	250	**82.1 [75.0, 88.2]**	**85.1 [70.8, 91.1]**	85.9 [73.7, 92.9]	**3.62 [2.64, 4.99]**	**12.0 [8.83, 14.5]**
	100	**75.3 [64.4, 84.4]**	86.3 [78.6, 92.5]	**71.3 [59.3, 82.9]**	**5.29 [3.57, 6.84]**	**16.1 [11.7, 18.6]**
Overall	2738	93.3 [88.5, 94.7]	92.0 [88.5, 95.1]	94.0 [90.6, 96.2]	1.44 [1.07, 2.03]	4.78 [3.47, 6.94]
	1000	**91.3 [86.2, 94.2]**	93.2 [89.5, 95.3]	**93.3 [82.9, 96.0]**	**1.95 [1.34, 4.55]**	**6.40 [4.26, 14.1]**
	500	**91.5 [83.3, 93.9]**	90.5 [81.8, 95.3]	92.7 [86.9, 95.7]	**1.94 [1.24, 2.93]**	**5.83 [4.08, 9.83]**
	250	**89.2 [81.7, 93.2]**	**90.6 [82.2, 93.8]**	93.1 [83.8, 96.5]	**1.98 [1.48, 3.75]**	**6.60 [4.63, 12.0]**
	100	**87.7 [74.8, 90.3]**	**87.0 [82.5, 92.3]**	**87.3 [70.8, 95.2]**	**3.01 [2.03, 5.32]**	**9.13 [5.88, 16.7]**

DSC, Dice similarity coefficient; MSD, mean surface dist.; HD, Hausdorff dist.

Bolded metrics denotes significant differences compared to the full (2738 images) dataset.

Table 3.B2. Median [Q1, Q3] 2D results for the U-Net++ trained using side-fire datasets of varying size, from 4023 (full end-fire set) to 100 images. The networks were evaluated on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images of the prostate.

Test Dataset	Side-fire Training Images	DSC (%)	Recall (%)	Precision (%)	MSD (mm)	HD (mm)
End-fire	4023	83.5 [71.8, 90.7]	74.8 [58.3, 85.6]	97.6 [96.5, 98.8]	3.07 [2.15, 4.89]	8.56 [7.32, 12.4]
	1000	77.7 [71.1, 84.2]	64.5 [57.2, 74.2]	**98.8 [97.8, 99.7]**	3.83 [3.13, 5.44]	10.9 [8.47, 14.2]
	500	**74.4 [66.0, 81.7]**	**62.5 [50.0, 71.3]**	98.0 [95.8, 99.4]	**4.05 [3.33, 6.15]**	**12.4 [9.31, 16.5]**
	250	**73.8 [69.7, 81.1]**	60.2 [57.1, 69.8]	98.6 [95.2, 99.6]	4.59 [3.76, 5.57]	12.3 [10.2, 14.6]
	100	**65.3 [58.0, 71.7]**	**49.2 [52.6, 59.1]**	97.0 [94.8, 99.6]	**6.28 [4.19, 7.28]**	**14.8 [10.8, 16.3]**
Side-fire	4023	93.3 [91.5, 95.0]	92.1 [88.5, 94.6]	95.1 [93.7, 97.7]	1.13 [0.90, 1.48]	3.50 [2.96, 4.67]
	1000	90.3 [89.0, 94.9]	93.8 [87.4, 95.5]	**93.7 [87.9, 95.1]**	1.59 [0.95, 2.13]	4.95 [3.59, 6.89]
	500	**90.3 [84.0, 92.7]**	**86.6 [78.9, 90.5]**	95.7 [91.1, 98.0]	**1.68 [1.20, 2.79]**	**5.84 [3.88, 8.99]**
	250	**89.2 [85.4, 92.6]**	87.2 [84.7, 91.8]	**91.8 [89.0, 95.9]**	**2.21 [1.20, 2.98]**	**8.44 [4.17, 10.1]**
	100	**81.0 [73.1, 83.7]**	**75.9 [67.0, 85.2]**	**85.4 [81.9, 95.9]**	**3.46 [2.88, 4.27]**	**9.76 [8.80, 11.6]**
Overall	4023	91.0 [83.3, 93.6]	86.7 [74.5, 92.8]	97.0 [94.1, 98.1]	1.65 [1.09, 3.08]	5.30 [3.48, 8.60]
	1000	85.9 [76.2, 92.9]	84.6 [63.7, 94.4]	95.9 [92.6, 99.0]	2.35 [1.26, 4.59]	7.43 [4.12, 11.8]
	500	**82.4 [70.9, 90.5]**	**74.4 [58.3, 87.2]**	97.1 [94.3, 98.3]	**2.85 [1.59, 4.79]**	**9.32 [5.54, 14.3]**
	250	**83.6 [73.4, 89.5]**	**80.4 [60.1, 87.5]**	95.5 [91.0, 98.6]	**3.13 [1.85, 4.70]**	**10.2 [6.89, 12.7]**
	100	**72.1 [62.3, 81.3]**	**63.9 [48.8, 77.2]**	**93.5 [84.7, 97.1]**	**4.25 [3.26, 6.76]**	**11.5 [9.45, 15.5]**

DSC, Dice similarity coefficient; MSD, mean surface dist.; HD, Hausdorff dist.

Bolded metrics denotes significant differences compared to the full (4023 images) dataset.

Table 3.B3. Median [Q1, Q3] 2D results for the U-Net++ trained using mixed (end-fire and side-fire) datasets of varying size, from 6761 (full dataset) to 100 images. The networks were evaluated on an unseen test dataset of 20 end-fire and 20 side-fire 3D TRUS images.

Test Dataset	Mixed Training Images	DSC (%)	Recall (%)	Precision (%)	MSD (mm)	HD (mm)
End-fire	6761	94.9 [93.8, 95.5]	94.6 [92.7, 95.6]	96.3 [94.0, 97.2]	1.10 [0.95, 1.38]	3.68 [2.87, 4.54]
	4023	94.4 [93.4, 95.6]	96.2 [93.3, 97.6]	**94.0 [90.7, 96.3]**	1.31 [1.04, 1.50]	4.34 [3.60, 5.33]
	2738	94.8 [93.2, 95.5]	94.1 [90.4, 95.6]	96.4 [95.5, 98.1]	1.17 [0.99, 1.41]	3.71 [3.30, 4.68]
	1000	94.3 [92.0, 95.0]	95.2 [91.4, 96.6]	94.4 [91.8, 97.3]	**1.29 [1.05, 1.95]**	4.05 [3.45, 5.87]
	500	**92.0 [89.6, 93.8]**	**90.0 [84.5, 94.6]**	96.4 [93.4, 98.0]	**1.58 [1.38, 2.07]**	**5.03 [4.14, 6.65]**
	250	**89.5 [84.3, 91.1]**	**84.3 [74.2, 88.9]**	96.5 [93.3, 98.2]	**2.11 [1.82, 3.42]**	**6.16 [5.10, 9.08]**
	100	**75.2 [63.6, 80.2]**	**60.5 [48.9, 69.0]**	**99.6 [97.1, 99.9]**	**4.90 [3.88, 6.99]**	**11.8 [8.51, 14.9]**
Side-fire	6761	94.6 [92.3, 95.1]	95.7 [91.3, 96.5]	94.2 [92.2, 96.4]	1.00 [0.83, 1.32]	3.27 [2.85, 4.19]
	4023	93.6 [91.0, 95.2]	94.4 [90.8, 96.6]	93.7 [90.1, 95.9]	1.09 [0.91, 1.80]	3.84 [3.10, 5.72]
	2738	93.1 [89.6, 94.9]	**92.5 [89.3, 93.6]**	95.5 [91.3, 96.7]	1.23 [0.88, 1.69]	4.29 [3.12, 5.69]
	1000	**91.8 [88.2, 94.6]**	92.5 [88.8, 95.6]	94.2 [88.0, 96.4]	**1.44 [1.00, 2.33]**	**4.96 [3.63, 7.14]**
	500	**90.5 [86.4, 93.6]**	93.8 [88.3, 95.5]	**90.9 [83.8, 95.2]**	**1.73 [1.21, 2.74]**	**5.61 [3.78, 8.69]**
	250	**85.2 [81.4, 90.8]**	92.2 [88.6, 94.9]	**80.9 [74.1, 89.1]**	**3.05 [1.73, 4.56]**	**8.85 [5.85, 15.6]**
	100	**84.3 [77.4, 86.0]**	**80.9 [76.4, 87.1]**	**85.6 [80.2, 92.5]**	**2.74 [2.52, 3.88]**	**8.48 [6.66, 11.6]**
Overall	6761	94.7 [92.6, 95.4]	95.1 [92.7, 96.4]	95.7 [92.9, 96.8]	1.07 [0.90, 1.33]	3.60 [2.87, 4.45]
	4023	94.2 [92.6, 95.4]	95.2 [92.8, 97.1]	93.9 [90.6, 96.1]	1.21 [1.00, 1.52]	4.00 [3.50, 5.33]
	2738	94.1 [92.1, 95.2]	**92.9 [90.2, 94.8]**	96.1 [92.4, 97.4]	1.19 [0.98, 1.56]	3.98 [3.30, 4.76]
	1000	**93.2 [89.7, 94.9]**	94.6 [89.4, 95.9]	94.3 [88.3, 96.5]	**1.32 [1.01, 2.03]**	**4.45 [3.52, 6.15]**
	500	**91.8 [88.4, 93.8]**	**92.0 [87.4, 94.9]**	95.1 [87.5, 96.7]	**1.63 [1.35, 2.40]**	**5.14 [4.10, 7.99]**
	250	**87.0 [82.9, 91.0]**	**88.9 [82.9, 92.5]**	**91.4 [80.8, 96.5]**	**2.41 [1.82, 3.56]**	**7.84 [5.46, 10.6]**
	100	**79.6 [67.7, 85.8]**	**73.1 [58.0, 81.8]**	93.2 [83.4, 99.7]	**3.88 [2.61, 5.81]**	**9.86 [7.28, 13.4]**

DSC, Dice similarity coefficient; MSD, mean surface dist.; HD, Hausdorff dist.

Bolded metrics denotes significant differences compared to the full (6761 images) dataset.

Chapter 4

4 A power Doppler ultrasound method for improving intraoperative needle tip localization in interstitial prostate brachytherapy

Intraoperative needle tip localization is critical for treatment planning in prostate brachytherapy. However, tip visibility can be limited with conventional B-mode US due to image artifacts. A method to improve needle tip visibility intraoperatively may improve tip identification accuracy for difficult-to-see needles, thus improving treatment accuracy. The purpose of Chapter 4 is to describe the development and validation of a PD US-based needle tip localization method designed specifically for interstitial prostate brachytherapy.

The contents of this chapter have been submitted for publication in *Medical Physics* and are currently under peer review: Orlando N, Snir J, Barker K, D'Souza D, Velker V, Mendez LC, Fenster A, and Hoover DA.

4.1 Introduction

High-dose-rate (HDR) brachytherapy (BT) is a common treatment technique for localized intermediate to high-risk prostate cancer.[1–3] In this treatment a single high-activity radioactive source is passed through needles inserted into the prostate through the perineum, guided by a rigid grid of evenly spaced holes known as a template. Needles are typically inserted under transrectal ultrasound (TRUS) guidance, as it is a widely available and low-cost modality, offering real-time imaging and comparable needle tip localization accuracy to computed tomography (CT),[4] making it the preferred method for prostate HDR-BT guidance at many centers.[5] In addition, TRUS-guided procedures allow for all aspects of the treatment workflow to occur in the operating room, including intraoperative imaging and treatment planning, allowing the patient to remain stationary.

Identification of needle tip position directly influences the radiation treatment plan, making accurate tip localization critical for a safe and effective HDR-BT treatment.[6–8] Error

in needle tip localization is the primary component of uncertainty in HDR-BT setup, with errors greater than 3 mm potentially leading to adverse outcomes including overexposed organs-at-risk or cancer recurrence due to an under-dosed tumor.[7] For TRUS-guided HDR-BT procedures, needle tip positions are typically identified using live two-dimensional (2D) brightness (B)-mode sagittal ultrasound (US).[6] Siebert *et al.*[6] investigated needle tip localization accuracy in water phantoms, demonstrating errors between 0.8 and 2.8 mm for 2D sagittal US. Due to the homogeneity of water, images obtained using this phantom are of higher quality than typical clinical images, making this a best-case scenario. Image artifacts such as shadowing or reverberation from nearby anatomy and needles can make it difficult to clearly visualize the needle tip in clinical ultrasound images, limiting the accuracy of needle tip localization.[9]

The use of Doppler US has been suggested to improve the visualization of surgical instruments in minimally invasive procedures.[10–21] Doppler US images display the velocity component in the plane of the B-mode US image of tissues or instruments as a colored overlay. Doppler-based visualization techniques apply some form of perturbation to the surgical instrument, generating motion that can be detected and displayed by Doppler US, typically in one of two modes, Colour Doppler (CD) or Power Doppler (PD), to improve visibility. While CD can display the direction of motion, PD is typically preferred in these applications as it has a higher sensitivity, and the direction of motion is unimportant. Doppler-based techniques for improving surgical instrument detection have been applied in a variety of procedure types and disease sites.[10–21] This includes various tissue biopsies,[10,11,13] pericardiocentesis,[14] and radiofrequency ablation,[12,13,15,20] with none focusing on brachytherapy to our knowledge. The motion was generated in numerous ways including vibrating devices clipped onto the needle,[10,11,14,20] piezoelectric crystals,[13,15,17] and electromagnetic actuation of small permanent magnets inside the needle.[12] These methods often required modifications to the clinical equipment or physical attachment to the needle or surgical tool. This is impractical in the BT workflow where 12 or more needles are typically used. There is also a clear lack of clinical validation in the literature of the use of PD methods to identify BT needles in human tissues. Thus, a PD-based method designed specifically for the BT workflow is required, in addition to an exploration of clinical feasibility.

We have previously proposed a PD needle visualization method utilizing a simple mechanical device capable of oscillating at a fixed and controllable frequency, requiring no modifications to the clinical workflow.[22] The device was both powered and controlled by a laptop computer, requiring two operators to use the device. The additional computer and required cables were cumbersome in the busy operating room environment. In addition, consistent Doppler signal generation was difficult with this end-piece design, as it required precise positioning of the device. These limitations made the device not feasible for long-term and efficient clinical use. To overcome these limitations and improve clinical translation potential, we presented the initial feasibility design of a wireless mechanical oscillator designed specifically for BT, tested in a small-scale proof-of-concept phantom study with six needles.[23]

In this paper, we report on the development and validation of our clinically-ready novel wireless mechanical oscillator and PD US needle tip localization method through comprehensive tissue-equivalent phantom experiments simulating prostate HDR-BT procedures and scenarios with extensive shadowing artifacts limiting needle visibility. In addition, we validated our PD US method in five prostate cancer patients undergoing standard HDR-BT as part of a prospective feasibility clinical trial, the first of its kind to our knowledge. Our cost-effective and easy to implement method, requiring no modifications to the clinical equipment or workflow, may allow for improved intraoperative needle tip localization accuracy, especially in cases with reduced needle visibility in standard US images. This has the potential to improve both patient safety and treatment accuracy in HDR-BT and more broadly in any minimally invasive needle-based procedure.

4.2 Materials and Methods

4.2.1 Wireless mechanical oscillator design

Photographs of the wireless mechanical oscillator are shown in Fig. 4.1. The oscillator contains a Faulhaber 1331T006SR brushed DC motor and MCDC 3002 series motion controller (Faulhaber MicroMo LLC, Clearwater, FL, USA) powered by a 1200

milliamp hour Lithium-ion polymer battery, rechargeable via a micro-USB cable. The ergonomic cylindrical case, 3D printed using a nylon-aluminum material, had a diameter of 4.2 cm and length of 17.8 cm (including the end-piece). The oscillation frequency is controlled on the device using a 16-position rotary switch with resistances tuned to specific oscillation frequencies ranging from 22.0 Hz to 44.5 Hz. Frequency values are printed directly on the device, as shown in Fig. 4.1C, allowing the user to easily change to the desired frequency while still in the operating room. As the oscillator is battery-powered with frequency-control directly on the device, it is completely wireless and easily operated by one user, requiring no additional equipment, computer, or personnel in the operating room.

The end-piece, machined from 400 series stainless steel, features a cylindrical cup shape, custom-designed for BT applications to fit over top of the commonly used cylindrical mandrins. The end-piece and example mandrins are shown in Fig 4.1B. This design provides two points of contact for needle vibration, while also removing ambiguity surrounding oscillator placement, as the cylindrical cup can be placed directly over the mandrin. Thus, our method requires no modifications or attachment to the standard clinical equipment. Finally, the end-piece design ensures needle oscillation is perpendicular to the insertion direction, minimizing the possibility of vibration pushing the needle deeper into the patient. When powered, the end-piece oscillates with an amplitude of 1 mm. As clinical needle mandrins vary in shape, the end-pieces can be machined to any size. In our application, two end-piece sizes were required as shown in Fig. 4.1B. For easy removal, the end-piece is attached to the oscillator magnetically. This allows for easy bagging of the device in the sterile operating room environment, as a plastic covering can be placed over the 3D-printed case, with the magnetic end-piece attached through the plastic. The end-piece and plastic covering can then be disposed of after the procedure, maintaining a sterile device.

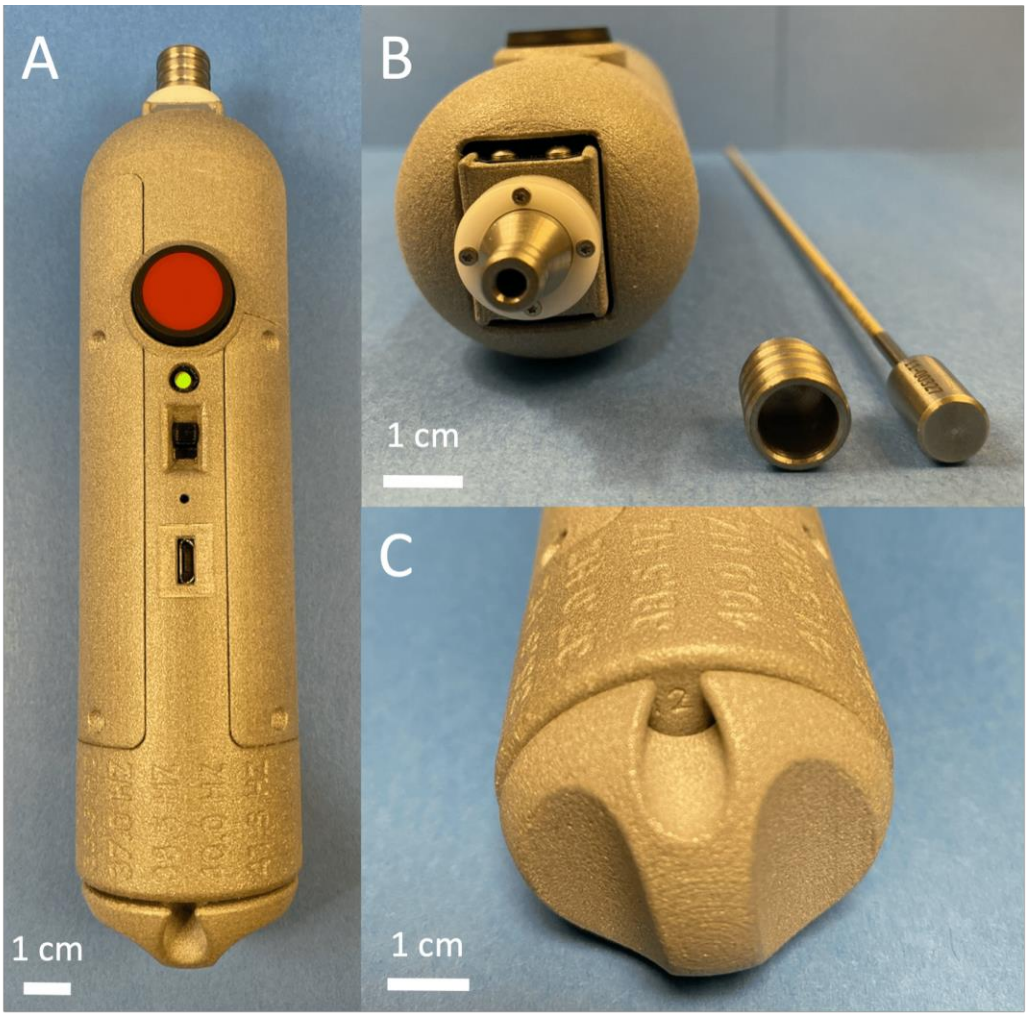

Figure 4.1. (A) Photograph of our wireless mechanical oscillator. (B) Front view showing the magnetically attached cylindrical end-pieces and example plastic needle with the corresponding mandrin. Two end-piece sizes were used corresponding to plastic and metal Varian needles. (C) The oscillation frequency is controlled using a 16-position rotary dial at the bottom of the device. Frequency values are printed on the device for ease of use.

4.2.2 Phantom validation

Prior to clinical testing, phantom validation of our PD needle tip localization method involved three experiments, with details outlined in sections 4.2.2.1 and 4.2.2.2. For each experiment, a tissue-mimicking agar phantom was made containing 35 g agar powder (Sigma Aldrich Co., St. Louis, MO, USA) and 80 mL glycerol (Sigma Aldrich Co.,

St. Louis, MO, USA) per 1 L of distilled water, simulating the speed of sound in soft tissue, approximately 1540 m/s.[24] 10 g of SigmaCell cellulose powder (Sigma Aldrich Co., St. Louis, MO, USA) per 1 L distilled water was also added to simulate soft tissue scatter. A prostate mold with a volume of 50.0 cm^3 was 3D printed based on a prostate segmented from a patient MR image. An identical agar mixture was used to fill the mold; however, only 1 g of cellulose powder per 1 L distilled water was used to provide contrast between the prostate and the background in the US image and 10 g of Tungsten powder (Sigma Aldrich Co., St. Louis, MO, USA) per 1 L distilled water was added to ensure visibility of the prostate in CT images. In addition to the embedded agar prostate, multiple landmarks were embedded in three layers 1 cm away from the prostate to facilitate US-to-CT rigid image registration. Each layer contained two 1.5 mm steel ball bearings (McMaster-Carr, Elmhurst, IL, USA) and three 9.5 mm agar spheres. To provide contrast with the phantom background and allow visibility in CT, the agar spheres contained 10 g Tungsten powder per 1 L distilled water but no cellulose powder.

All US imaging was completed using a BK3000 system with an E14CL4b biplane endocavity transducer (BK Medical, Peabody, MA, USA), supported by a CIVCO EX^3 Stepper (CIVCO Medical Solutions, Coralville, IA, USA). B-mode gain and frequency were 60% and 9 MHz, respectively, with a depth of 6.5 cm. For all experiments, the pulse repetition frequency (PRF) was 0.8 kHz and Res/Hz was 4/22 Hz. Doppler gain and oscillation frequency were the two key parameters used to optimize the power Doppler signal. To best optimize the Doppler signal, Doppler gain and oscillation frequency were increased until flash artifacts were observed, then the gain was reduced. The optimal Doppler gain and oscillation frequency varied depending on needle type, distance from the transducer, and artifact presence. Specific values are provided in sections 4.2.2.1 and 4.2.2.2. All PD US imaging and needle perturbations were done with the mandrins inserted and the needle position locked by the template. This ensured that needle vibration did not change the position of the needle.

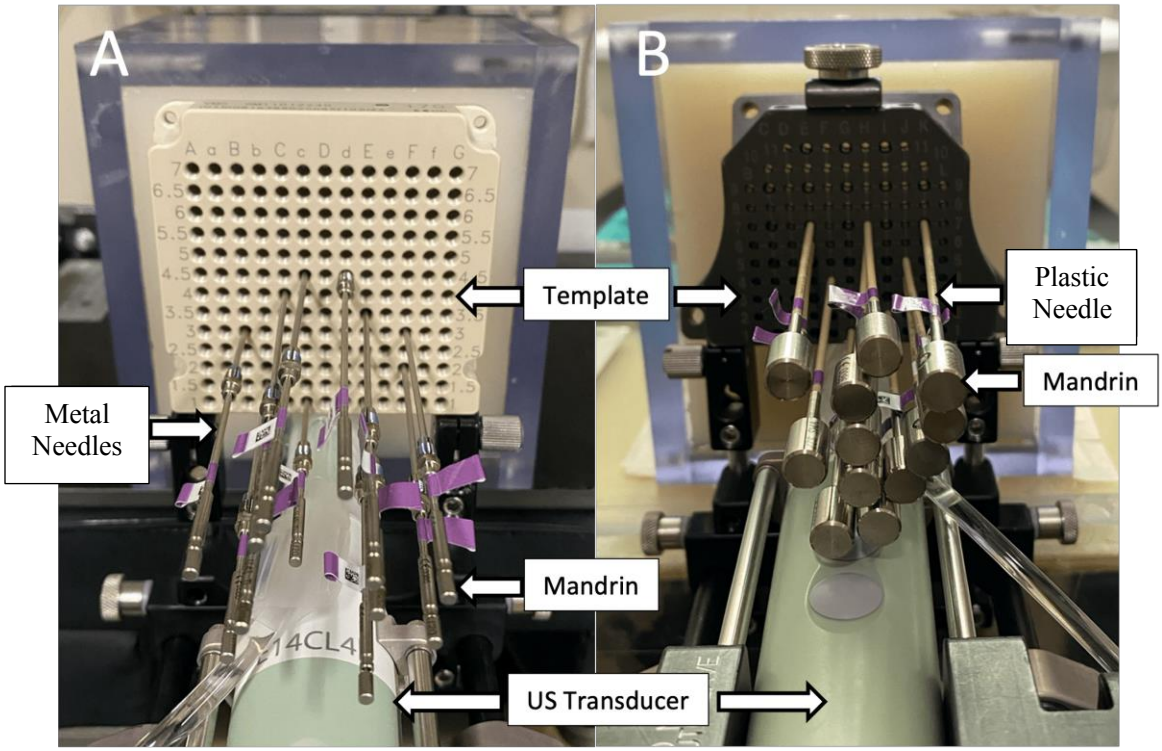

Figure 4.2. Photograph of our experimental set-up showing needle implant patterns for a simulated HDR-BT procedure with metal needles (A), and a simulated case with extensive needle shadowing with plastic needles (B). The US transducer, needle template, and varying size needle mandrins can be seen in both experiments.

4.2.2.1 Mock HDR-BT procedure

In a phantom experiment simulating standard prostate HDR-BT procedures at our center, 12 stainless steel needles (17G) with corresponding mandrins (Varian Medical Systems, Palo Alto, CA, USA) were inserted into our tissue-equivalent phantom through the Crook prostate template (Varian Medical Systems, Palo Alto, CA, USA) shown in Fig. 4.2A. Following the standard clinical procedure, needles were inserted to provide the best possible radiation coverage for the embedded agar prostate, as shown in Fig. 4.2A. All 12 needles were first inserted into the prostate mid-gland, where they were identified in the axial US plane in Vitesse (Varian Medical Systems, Palo Alto, CA, USA). Needles were then advanced, one at a time, to the base of the prostate and verified in the sagittal US plane.

Needle tip locations were first identified using standard B-mode US in the sagittal plane by a board-certified medical physicist experienced with prostate brachytherapy (D.H.). To maintain the highest image resolution, needle tip positions were identified directly on the US screen and subsequently transferred to Vitesse. Next, needle tip locations were identified using PD US with perturbation from our wireless mechanical oscillator. To test needle tip localization accuracy without B-mode US information, the observer was blinded by reducing the B-mode gain to 0, as shown in Fig. 4.3C. The needle tip position was then identified using only the PD signal after mechanical perturbation. This also simulated the clinical case of a needle that is not visible in the standard B-mode US image due to shadowing. Finally, needle tip positions were identified using B-mode US and PD US together with mechanical perturbation. The needle tip location was selected directly on the US screen as the furthest point in the Doppler signal. The distance from the template to the end of the needle, the so-called end length, was recorded for each needle, verified by a second observer.

For PD US imaging, the optimal Doppler gain and oscillation frequency varied depending on the needle position and the presence or absence of needle shadowing obstructions. For unobstructed needles close to the US transducer, Doppler gain was 10-15% and oscillation frequency was 32.5-40 Hz. For the needles far from the US transducer or obstructed by needle shadowing artifacts the Doppler gain was 32-40% and oscillation frequency was 43 Hz.

4.2.2.2 Explicit needle shadowing

The second set of phantom experiments examined needle tip localization accuracy in the presence of needle shadowing artifacts. 15 stainless steel needles with corresponding mandrins were inserted into our tissue-equivalent agar phantom. Needles were inserted in three columns extending from the transducer. The left and right columns formed a V-shape, while the middle column extended upwards away from the transducer, as shown in Fig. 4.2B. This needle insertion pattern ensured every needle, apart from the three closest to the US transducer, was shadowed by one or multiple needles below it, simulating the clinical scenario where the needle tip location in the B-mode US image is obscured by shadowing

artifacts. This implant pattern was repeated in a second experiment utilizing 13 plastic needles (2 mm diameter) with corresponding metal mandrins (Varian Medical Systems, Palo Alto, CA, USA). In this implant, the left and right columns had four needles each, while the central column had five needles.

As described in section 4.2.2.1, needle tip locations were identified in an identical manner, starting with standard B-mode US alone, then PD US with mechanical perturbation alone by reducing B-mode gain to 0, and finally B-mode and PD US together. Once again, the distance from the template to the end of the needle, the end length, was recorded for each needle.

For PD US imaging with the stainless-steel needles, the Doppler gain and oscillation frequency ranged from 35% and 32.5 Hz for needles close to the transducer to 45-55% and 40-44.5 Hz for shadowed needles far from the transducer. Similarly, for the plastic needles, the Doppler gain and oscillation frequency ranged from 35-40% and 32.5 Hz for close needles to 40-50% and 43 Hz for obstructed needles far from the transducer.

4.2.2.3 Evaluation of tip localization accuracy

Needle tip localization error was computed using a clinical method based on ideal reference needles. For each phantom experiment, 1-2 clearly visible, unobstructed needles were selected as reference needles, which were typically along the row closest to the US transducer. Using the measured exposed needle end lengths and the known total needle length, the difference between the identified needle tip location and the expected tip location based on the reference needles was computed in Vitesse, providing a metric for tip localization error. In addition, the absolute distance in the insertion (z) direction between needle tip positions identified using PD US only and PD plus B-mode US were computed for each needle using MATLAB (MathWorks, Natick, MA, USA) to examine the impact of B-mode context on PD US tip localizations.

Needle tip localizations using B-mode and PD US were also compared to tip localizations using CT as a gold standard. To enable registration between US and CT, axially reconstructed step-back 3D US images were acquired with 1 mm step spacing using Vitesse after needle insertion was completed. 3D US images had a size of [1100×700×70]

voxels with a voxel size of [0.083×0.110×1.00] mm^3. CT imaging was completed using a Philips Brilliance Big Bore CT Scanner (Philips Healthcare, Andover, MA, USA) with 120 KV potential, 325 mA x-ray tube current, and 2 mm slice thickness, generating images with a size ranging from [512×512×299] to [512×512×339] voxels and voxel size ranging from [0.355×0.355×1.00] to [0.684×0.684×1.00] mm^3. All CT scans were performed with the mandrins removed to limit metal artifacts. The needle tip positions were manually identified in the CT images by an experienced medical physicist (D.H.). Manual rigid landmark registration between 3D US and CT was performed using the embedded landmarks described in section 4.2.2. With a total of six 1.5 mm steel ball bearings and eight 9.5 mm agar spheres per phantom, three ball bearings, and four agar spheres were used as fiducials for registration, while the remaining landmarks were used to compute target registration error (TRE).[25] After registration to the same coordinate system, the distance in the insertion (z) direction between needle tip positions identified in B-mode or PD US and in CT could be calculated.

4.2.3 Clinical validation

Our PD US needle localization method was validated in a prospective feasibility clinical trial (NCT03861507) approved by the Research Ethics Board at Western University (London, ON, Canada). This feasibility trial had a planned accrual of five patients with a maximum of 20 patients, all of whom were scheduled to undergo standard interstitial HDR prostate brachytherapy at the London Regional Cancer Program. To be considered for the study, patients had to be male, aged 18 years or older, willing to provide informed consent, have pathologically confirmed prostate cancer on a previous biopsy, and are suitable for and consenting to HDR-BT for treatment as standard of care. Any patient who previously received radiotherapy to the prostate was excluded from the study. Treatments were completed by three radiation oncologists specializing in brachytherapy. Our study protocol was completed during the time required to create and validate the radiation treatment plan and did not influence treatment decision-making or the total procedure time. 12-15 stainless steel needles were used per patient, inserted following the standard clinical protocol (63 needles total).

Needle tip locations were first identified using standard B-mode US alone. The treating radiation oncologist rotated the US transducer to the optimal plane. The needle tip location was identified by consensus between the radiation oncologist and board-certified medical physicist (D.H.) on the live sagittal US view directly on the US machine as described in section 4.2.2.1. The tip location was then transferred to Vitesse. This B-mode localization is distinct from the clinical version as the needles are already inserted to the base of the prostate and are locked in position by the needle template, meaning they cannot be rotated. The visibility of each needle in the live B-mode US image was labeled on a three-tiered scale as clearly visible, partially visible, or mostly/completely obscured. Needle tip locations were then identified using PD US with mechanical perturbation using our wireless oscillator, which was controlled by the radiation oncologist. The optimal Doppler gain and oscillation frequency varied between patients and was dependent on needle position and the presence or absence of obstructions, leading to a wide range of values. Doppler gain ranged from 8% to 50%, with an average for unobstructed needles of 15-25%. The oscillation frequency was consistently high, ranging from 38.5 Hz to 44.5 Hz. All needle perturbations were performed with the mandrins inserted and the needles locked in place individually using the Crook prostate template. Exposed needle end length measurements were repeated after the study protocol was complete to ensure that needle position remained the same.

As described in section 4.2.2.3, needle tip localization accuracy was computed using the clinical method based on the ideal reference needles selected during the clinical procedure. Tip localization error was computed in Vitesse using the measured exposed needle end lengths and the known total needle length to compare the identified needle tip location and the expected tip location based on the reference needles. Our center does not use a post-implant CT scan for HDR prostate brachytherapy, so the comparison to CT as ground truth was not possible for the clinical study.

4.2.4 Statistical analysis

All statistical analyses were performed in GraphPad Prism 9.3.1 (Graphpad Software, Inc., San Diego, CA, USA). The normality of distributions was assessed using

the Shapiro-Wilk test. Failure of this test led to the use of alternative nonparametric statistical tests, presented in parentheses in the following paragraph. The significance level for statistical analysis was chosen such that the probability of making a type I error was less than 5% ($p < 0.05$). Statistically significant differences are denoted simply as significant for the remainder of this manuscript.

Comparisons between needle tip localization accuracy using B-mode US, PD US only, and PD plus B-mode US in phantom experiments, including comparisons to CT as a gold standard, were completed using repeated measures one-way ANOVA with Tukey's multiple comparisons tests (Friedman test with Dunn's multiple comparisons tests). Differences in tip localization accuracy in the mock HDR-BT procedure and the explicit shadowing experiment for each US visualization method were compared using two-tailed unpaired t-tests (Mann-Whitney U Tests). Similarly, the accuracy for identifying plastic and metal needle tips in the explicit shadowing experiments were directly compared using the same tests. For clinical validation, comparisons of needle tip localization accuracy using B-mode US and PD US for all needles and specifically for unobstructed, partially obstructed, and unobstructed needles were compared using two-tailed paired t-tests (Wilcoxon matched-pairs signed-rank tests) for each individual patient and overall. A comparison of needle tip localization accuracy between the three visibility grades for B-mode and PD US was completed using a two-way ANOVA with Tukey's multiple comparisons test. Needle localization accuracy between physicians for both B-mode and PD US were compared using one-way ANOVA with Tukey's correction for multiple comparisons (Kruskal-Wallis tests with Dunn's correction for multiple comparisons).

4.3 Results

4.3.1 Phantom validation

Examples of needle visualization using standard B-mode US, PD plus B-mode US, and PD US only are shown in Fig. 4.3.

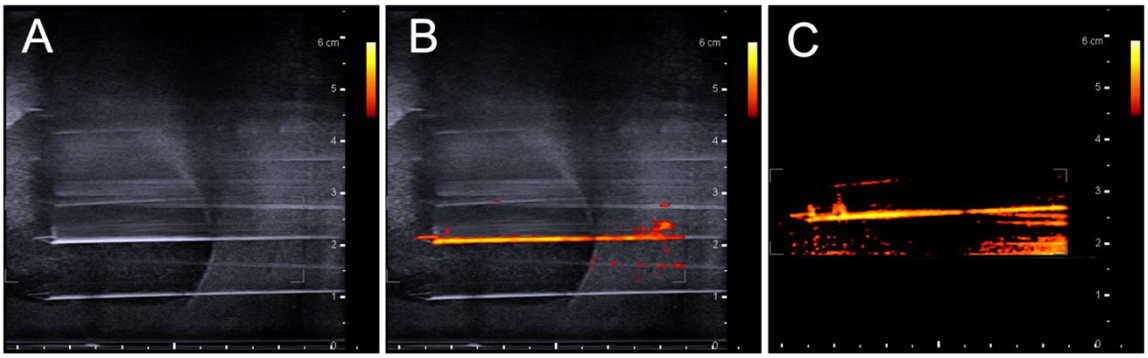

Figure 4.3. Example needle visualization results in phantom, showing (A) a standard B-mode US image, (B) PD US image with perturbation from our mechanical oscillator, and (C) PD US only with B-mode context blinded to the observer.

Quantitative comparisons of absolute needle tip localization error, computed based on clearly visible reference needles, are shown in Fig. 4.4 for three phantom experiments, with complete absolute and signed results provided in Table 4.1. For the metal needles inserted to mimic a standard HDR-BT procedure, only PD US alone had significantly higher signed tip error compared to PD plus B-mode US, with no significant difference observed for absolute tip error (Fig. 4.4A). With metal needles implanted to maximize shadowing artifacts, no significant differences in tip error were observed for either absolute or signed metrics (Fig. 4.4B). Similarly, no significant differences were observed for the identical shadowing experiment with plastic needles, although the mean tip error and variance we variance were much higher for standard B-mode US compared to both PD US approaches, as shown in Fig. 4.4C.

The impact of B-mode information on needle tip identification was investigated by computing the difference in the insertion (z) direction between needle tip locations identified using PD US alone and PD plus B-mode US. The absolute difference was 0.7 ± 0.8 mm, 0.5 ± 0.5 mm, and 0.7 ± 0.9 mm while the signed difference was 0.6 ± 0.9 mm, 0.5 ± 0.5 mm, and -0.6 ± 0.9 mm for the mock HDR-BT procedure, shadowing experiment with metal needles, and shadowing experiment with plastic needles, respectively.

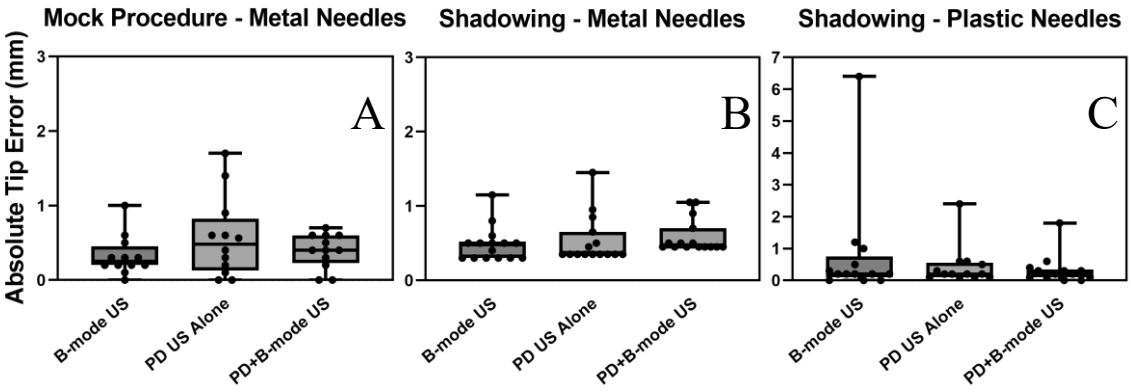

Figure 4.4. Box and whisker plots of absolute needle tip localization errors, in mm, comparing needle visualization with standard B-mode US, PD US alone, and PD plus B-mode US. Plots from left to right show results for the mock HDR-BT procedure with (A) metal needles (N=12) and the explicit needle shadowing implant patterns with (B) metal (N=15) and (C) plastic needles (N=13). Boxes denote interquartile range from the 25th to 75th percentile while whiskers show the minimum and maximum values. Each individual tip error is shown with a bar representing the median.

Table 4.1. Mean ± standard deviation absolute and signed tip localization errors based on ideal reference needles for visualization with standard B-mode US, PD US alone, and PD plus B-mode US in three phantom experiments with two needle types. Negative values represent an underprediction relative to the expected tip location.

		Tip Error (mm)					
		Absolute			Signed		
Experiment	N	B-mode US	PD US Alone	PD+B-mode	B-mode US	PD US Alone	PD+B-mode
Mock HDR-BT Procedure	12	0.3 ± 0.3	0.6 ± 0.5	0.4 ± 0.2	0.1 ± 0.4	0.6 ± 0.5	-0.2 ± 0.4
Shadowing – Metal Needles	15	0.5 ± 0.2	0.5 ± 0.3	0.6 ± 0.2	0.3 ± 0.5	0.3 ± 0.5	0.3 ± 0.5
Shadowing – Plastic Needles	13	0.8 ± 1.7	0.4 ± 0.6	0.3 ± 0.5	0.3 ± 1.9	-0.1 ± 0.8	-0.3 ± 0.5

Needle tip localization error for B-mode US, PD US alone, and PD plus B-mode US compared to CT imaging as a gold standard is shown in Table 4.2. The mean ± standard deviation TRE using a total of 18 landmarks for the rigid registrations between 3D US and CT was 0.8 ± 0.3 mm. For metal needles inserted in a standard HDR-BT configuration, tip locations identified using B-mode US had significantly higher agreement to the CT gold standard compared to PD plus B-mode US. When metal needles were inserted to maximize shadowing artifacts, PD plus B-mode US demonstrated significantly higher error compared to PD US alone and B-mode US. For plastic needles inserted to maximize needle shadowing artifacts, the difference compared to needle tip locations identified using CT imaging was significantly lower for PD plus B-mode US compared to PD US alone. Although B-mode US and PD plus B-mode US compared to the CT gold standard were not significantly different in this case, we did observe reduced mean tip error and variance compared to CT for PD plus B-mode US.

Table 4.2. Mean ± standard deviation absolute and signed tip localization errors compared to the CT gold standard for needle visualization with standard B-mode US, PD US alone, and PD plus B-mode US in three phantom experiments with two needle types.

		Tip Error vs CT (mm)					
		Absolute			Signed		
Experiment	N	B-mode US	PD US Alone	PD+B-mode	B-mode US	PD US Alone	PD+B-mode
Mock HDR-BT Procedure	12	0.9 ± 0.5	1.6 ± 0.7	1.9 ± 0.6	0.9 ± 0.6	1.3 ± 1.2	1.9 ± 0.6
Shadowing – Metal Needles	15	0.7 ± 0.4	0.9 ± 0.5	1.4 ± 0.5	0.7 ± 0.4	0.9 ± 0.5	1.4 ± 0.5
Shadowing – Plastic Needles	13	3.7 ± 0.5	4.1 ± 0.5	3.7 ± 0.4	3.3 ± 1.9	4.1 ± 0.5	3.7 ± 0.4

4.3.2 Clinical validation

Six patients consented to our prospective feasibility clinical trial. Due to treatment complications unrelated to our study, we had insufficient time to complete the study protocol for one patient (PD-04), meaning five patients completed the study. 12 needles were inserted for all patients apart from PD-05, where 15 needles were used. Example B-mode US and PD US needle visualizations in clinical prostate HDR-BT patients are shown in Fig. 4.5 for varying needle visibilities.

Figure 4.6 shows clinical results comparing absolute tip localization error computed using the reference needles chosen during the clinical procedure for standard B-mode US and PD US with mechanical perturbation. Complete numerical results are provided in Table 4.3, with the mean absolute tip error computed for each individual patient and averaged for each of the three physicians who participated in the trial. As shown in Table 4.3, there were no significant differences in tip localization error between B-mode US and PD US with mechanical perturbation for any individual patient or overall, although PD US did show reduced mean tip error and variance. A scatter plot of PD US tip error as a function of B-mode US tip error for all needles in clinical cases is shown in Fig. 4.7.

When tip errors were averaged for each physician, there was once again no significant difference between the US visualization methods, as seen in the right side of Table 4.3. Physician 1, who completed three cases, was observed to have the largest improvement when utilizing PD US, although B-mode tip error was also the highest with a mean of 1.0 mm. Comparing standard B-mode US tip localization between physicians, significant differences were observed, specifically between physician 1 and 3. When PD US tip localizations were compared between physicians, no significant differences were observed.

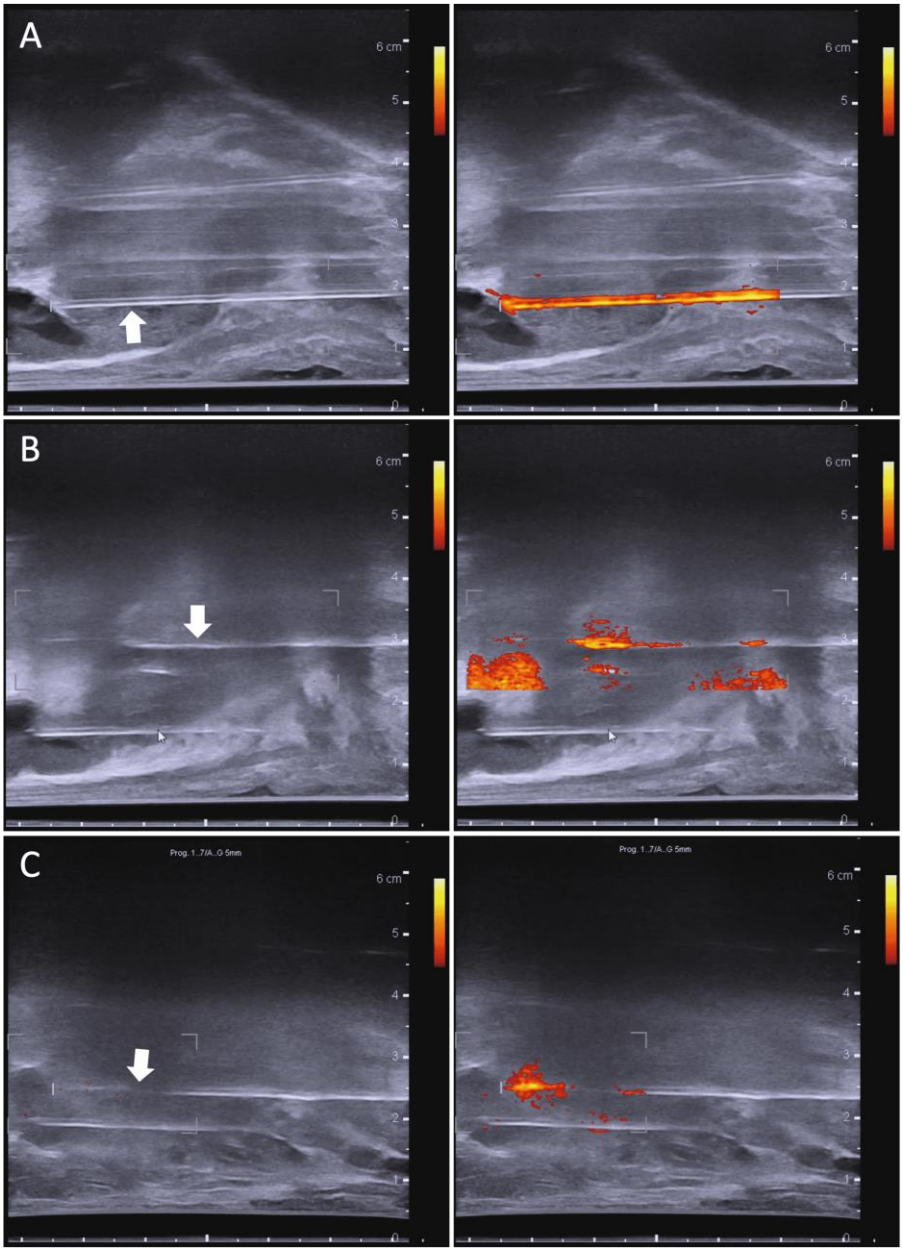

Figure 4.5. Example images from the clinical trial showing needle visualization in patients using standard B-mode US (left) and PD US with perturbation (right). The rows show an example (A) unobstructed needle, (B) partially obstructed needle, and (C) obstructed needle. Arrows point to the needle in question.

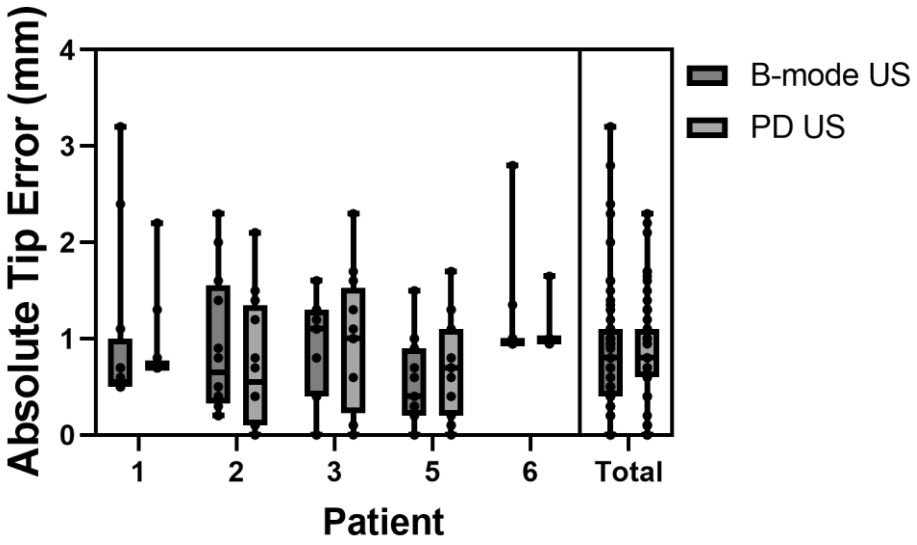

Figure 4.6. Box and whisker plot of absolute needle tip localization errors comparing needle visualization with standard B-mode US to PD US with perturbation in clinical cases. Tip error for each individual patient and the total tip error are shown from left to right. Boxes denote interquartile range from the 25th to 75th percentile while whiskers show the minimum and maximum values. Each individual tip error is shown with a bar representing the median.

Table 4.3. Mean ± standard deviation absolute tip localization errors comparing needle visualization with standard B-mode US to PD US with perturbation for clinical cases. Tip errors for each individual patient and averaged for each physician are shown in the left and right halves of the table, respectively. Overall tip error is shown in the bottom row.

		Tip Error (mm)			Tip Error per Physician (mm)		
Physician	Patient	B-mode US	PD US	*p*-value	B-mode US	PD US	*p*-value
1	1	1.0 ± 0.9	0.9 ± 0.5	—	1.0 ± 0.7	0.9 ± 0.5	0.6036[b]
	2	0.9 ± 0.7	0.7 ± 0.7	—			
	6	1.2 ± 0.5	1.0 ± 0.2	—			
2	3	0.9 ± 0.6	1.0 ± 0.7	—	0.9 ± 0.6	1.0 ± 0.7	0.7681[a]
3	5	0.5 ± 0.4	0.7 ± 0.5	—	0.5 ± 0.4	0.7 ± 0.5	0.1843[a]
Total		0.9 ± 0.7	0.8 ± 0.5	0.9563[b]	—	—	—

p-values correspond to [a]paired t-test or [b]Wilcoxon matched-pairs signed-rank test between B-mode and PD US

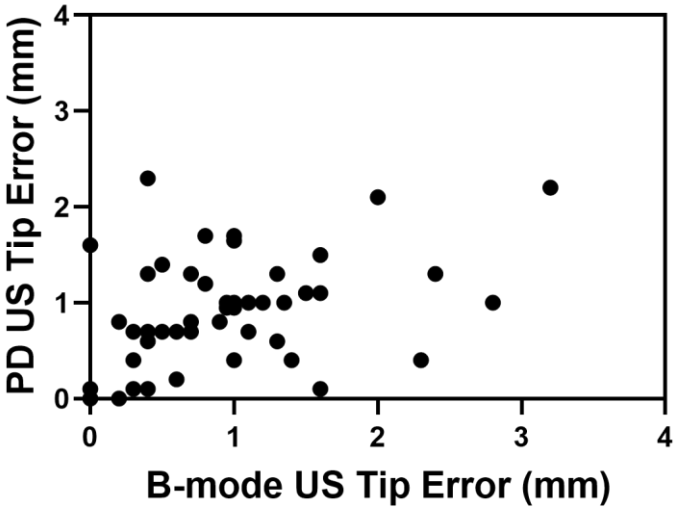

Figure 4.7. Scatter plot of PD US absolute tip error as a function of B-mode US absolute tip error for all 63 needles in the feasibility clinical trial.

For a total of 63 needles inserted across our five clinical cases, 27 were identified as unobstructed needles, 25 were identified as partially obstructed needles, and 11 were identified as obstructed needles. Absolute mean tip errors using B-mode US and PD US needle tip localization methods for each needle visibility tier are shown in Fig. 4.8. A two-way ANOVA showed a significant main effect for needle visibility grade, with the multiple comparison tests showing significant differences between unobstructed and obstructed needles. The US visualization method showed no significant main effect, although PD US corresponded to lower variance as visibility worsened, also reducing mean tip error for partially obstructed needles, as shown in Fig. 4.8.

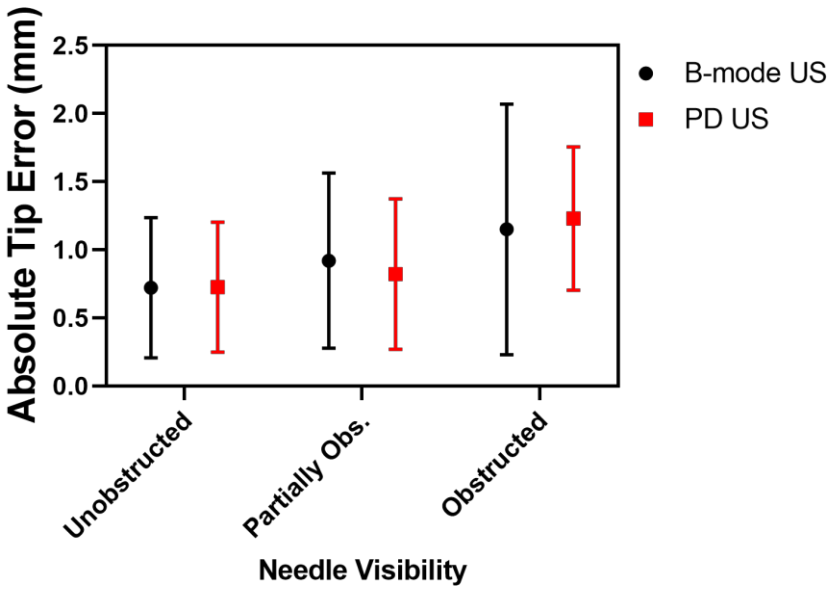

Figure 4.8. Plot of the mean ± standard deviation absolute tip localization errors comparing B-mode US to PD US with perturbation in clinical cases with varying needle visibility. Mean tip error for unobstructed, partially obstructed, and obstructed needles are shown from left to right.

4.4 Discussion

4.4.1 Phantom validation

For the first experiment, an implant with metal needles meant to simulate a clinical prostate HDR-BT procedure showed nearly identical performance between B-mode US alone and PD plus B-mode US, with our PD approach reducing the maximum tip error (Fig. 4.4A). Tip error for our PD US method alone had a higher mean and variation, likely due to the inherent limitations of our blinding approach. By reducing the B-mode gain to zero for the entire PD procedure, it became impossible to tell whether transducer contact with the phantom was maintained, or whether the transducer was in proper alignment with the needle plane. Doppler signal is dependent on transducer alignment and contact, potentially resulting in increased mean tip error. For phantom experiments where needle implant was determined to maximize needle shadowing artifacts, a second observer

ensured the transducer alignment to the needle plane and contact with the phantom was maintained before reducing B-mode gain to zero, effectively eliminating this source of uncertainty. For the explicit shadowing experiment with metal needles, tip error was nearly identical regardless of visualization method, as shown in Fig. 4.4B, but for the same experiment with plastic needles, large differences can be seen in Fig. 4.4C. Maximum tip error and variation for B-mode US alone is more than double that of PD US alone and PD plus B-mode US. The high mean and standard deviation are due to difficulty visualizing the needle tip as a result of excessive needle shadowing, including two needles where the tip was not at all visible using only B-mode US, resulting in the observer needing to guess the tip location. These same needles were clearly visible in PD US both with and without B-mode context to less than 1 mm tip error, demonstrating a clear use case for the PD method. The ability to clearly visualize needles using PD US with mechanical perturbation that were otherwise not visible in standard B-mode US was also shown in Orlando *et al.*[23] The similarity in tip error between PD US alone and PD plus B-mode US for both explicit shadowing experiments demonstrates the reliability of our PD US method, as it does not depend on B-mode information.

The impact of needle implant pattern and shadowing artifacts on tip error was investigated by comparing the mock HDR-BT implant with metal needles and the implant designed to maximize needle shadowing with metal needles. For B-mode US, the tip error was significantly higher when needles were implanted to maximize shadowing, showing that artifacts significantly influenced needle visibility in B-mode images. Contrary to this, no significant differences were observed for PD US visualization, demonstrating increased robustness to shadowing artifacts, with the PD signal helping to improve visibility.

Calculation of the difference in the insertion (z) direction between needle tip positions identified using PD US alone and PD plus B-mode US allowed for examination of the impact of B-mode information on needle tip localization. With mean absolute difference ranging from 0.5-0.7 mm, needle tip localization with PD US was very similar regardless of the presence or absence of B-mode information. This demonstrates the robustness of our method, as we can accurately identify the needle tip without seeing the B-mode image at all.

As shown in Table 4.2, B-mode US demonstrated higher agreement with gold standard tip locations identified using CT imaging compared to PD+B-mode US for metal needles. Although tip error vs CT was higher, the standard deviation for PD+B-mode US was approximately equal to B-mode US. This demonstrated that the PD method was consistently identifying the same point on the needle, thus this systematic offset could be corrected with commissioning experiments to adjust the needle digitization technique. For plastic needles, both B-mode and PD had identical absolute tip error compared to CT; however, B-mode US had considerably higher variation when considering signed tip error. Deviation from the CT tip location for all US methods was higher for plastic needles with mean values between 3.7 and 4 mm. Standard deviation was less than or equal to 0.5 mm, so this large difference may be explained by a systematic shift. Tip locations identified in the US images rely on the step-back 3D US image coordinate system, which is defined before needle insertion begins. The needle tip location identified in the live sagittal US image is then mapped to the 3D US coordinate system based on this initial registration. Any phantom movement during needle insertion could result in a global shift in needle position relative to the 3D US coordinate system, and thus potential large differences even if tip localization in the live US image was accurate. In addition, the needle tip does not appear as clearly in the B-mode US image for plastic needles compared to metal needles. The London Regional Cancer Program also transitioned from plastic needles to metal needles as their standard of care during this work, so recent practice was heavily weighted towards metal needles. Registration error was low with an average TRE of only 0.8 mm. As fiducial landmarks included large agar spheres, we believe fiducial localization error (FLE) may account for some of the observed TRE due to difficulty identifying the centroid. Future work could include a formal evaluation of FLE to assess the impact on registration accuracy.

Limitations of our phantom experiments included the use of only one observer to identify the needle tips and landmarks in both US and CT, preventing the evaluation of inter-observer variability. The use of live US meant needle tip positions were identified only once, thus intra-observer variability could not be computed either. Although the reference needle-based error metrics are utilized clinically, there are limitations to address. Importantly, the accuracy of reference-based approaches relied on accurately identifying

the needle tip of the reference needle(s). This error is mitigated by selecting clear, easy to see, needle tips. In addition, the calculation assumed no needle bending or deflection, which for a homogenous phantom is a good assumption. Any bending that does occur could influence the calculated tip error. The reference needle calculations also rely on end-length measurements, which are only accurate to 0.5 mm as they were measured using a ruler. The experiments were also limited by their small sample size, with only 12-15 needles inserted per experiment. This may have been a factor in the lack of observed statistically significant differences.

4.4.2 Clinical validation

Our PD US-based needle tip localization method was validated in five patients who underwent HDR-BT to treat prostate cancer. To our knowledge, this was the first Doppler US-based needle localization technique applied in a clinical BT procedure. The wireless design of the oscillator, with no cables or computer required to provide power to or control of the device, allowed for easy operation by one user, which will help increase the ease of further clinical translation. The easy-to-use design also allowed for the application of the oscillator in quick bursts as envisioned for difficult-to-see needles. The magnetically attached end-piece allowed for easy bagging of the device for use in the sterile operating room environment, while not negatively impacting Doppler signal generation. During the clinical HDR-BT procedures, the oscillator was controlled by the treating radiation oncologist. Three physicians participated in the trial, and in all cases, the physicians were able to competently use the oscillator after only a brief two-minute demonstration, highlighting the ease-of-use.

For clearly visible needles as shown in Fig. 4.5A, the addition of our PD approach did not offer much benefit for these cases. As needle visibility decreased (Fig. 4.5B and 4.5C), the PD US signal helped to clearly define the needle tip, increasing physician confidence in its location. This example also shows the PD “flash” artifacts that occur more commonly in clinical cases due to the tissue softness compared to the firm agar phantom. As seen in the PD image, these artifacts typically have no negative impact on needle tip localization.

Although there were no statistically significant differences in tip error observed between B-mode and PD US for both individual patients and overall (Fig. 4.6 and Table 4.3), some interesting trends were observed. In three of five patients, the mean and standard deviation tip error were reduced with the use of PD US. This improvement in performance was especially clear in patient 1 and patient 6, with a clear decrease in mean tip error and a reduction in standard deviation by a factor of two to three. Patient 5 represented an interesting case, as both B-mode US and PD US tip errors were smaller than for any other patient due to very clear needle visibility for almost all needles, and thus B-mode US alone allowed for good visualization of the needle tips. Although the PD US tip error was also small, it did not offer an improvement compared to B-mode US, further highlighting the use case of difficult-to-see needles. Considering the total tip error across all patients, PD US showed a reduction in mean tip error and variance. Fig. 4.6 also highlights the reduction in outliers seen with PD US compared to B-mode US, demonstrating our method's ability to improve needle visibility for challenging needles. This is clear when examining the tip error of all patients together, with an approximately equal median but greatly reduced variation and reduction in maximum tip error. This is critical clinically, as 1-2 needles with tip error greater than 2 mm could result in deviation from the planned radiation dose. This point is further reinforced in Fig. 4.7 where the tip error distribution is skewed towards the B-mode US axis, including several needles where tip error was greater than 2mm with B-mode US and less than 1.5 mm with PD US. This demonstrated ability to reduce the tip error of outlier needles with PD US, offering the potential for improved patient safety and treatment effectiveness.

Needle tip localization error separated by physician was shown in Table 4.3. Interestingly, the three patients where PD US offered the largest improvement compared to B-mode US corresponded to physician 1 who had substantially more time and experience using the device compared to physician 2 and 3 who only treated one patient each. Difference in the level of needle tip localization improvement based on physician could be explained in several ways. Although all physicians were able to successfully use the device after a brief tutorial, there was a learning curve for generating a strong, consistent Doppler signal. As physician 1 had substantially more time using the oscillator, the Doppler signal was more consistent suggesting that with further practice Doppler signal

generation would improve. For clinical cases, there was more difficulty generating a useful PD image compared to phantom experiments, necessitating trial and error with oscillator position to achieve the best signal which contributed to the learning curve for physicians. In addition, alignment of the transducer with the needle plane is critical for Doppler signal generation, and can be very difficult, so any deviation in this angular alignment could result in increased tip localization error. Differences between physicians may also be explained in part by the inherent differences in needle visibility between patients as opposed to differences in skill identifying the needle tip. In addition, the metal needles have a trocar tip, so alignment of the "face" of the needle tip with the US plane results in better needle visibility. As our experiments were completed with the needle position locked, they could not be rotated to take advantage of this in all cases.

Needle visibility was approximately equally split between unobstructed and partially obstructed needles, with only 11 fully obstructed needles, indicating in general good needle visibility. There was a clear increase in the mean and variance as needle visibility worsened for B-mode US, but PD US showed approximately no change in variation, demonstrating some level of robustness to needle visibility. This was most evident in the obstructed case where PD US showed approximately half the standard deviation of B-mode US. These results clearly demonstrate the utility of PD US in improving needle tip localization accuracy when needle visibility is poor.

There are several limitations to consider in our prospective clinical trial. Although no inter-observer calculations were possible, this is mitigated by the fact that needle tip locations in the clinical trial were the result of a consensus between the physician and expert medical physicist. Many of the clinical results utilize a reference needle-based metric, so concerns outlined in section 4.4.1 also apply here. Since we did not have access to alternative imaging such as CT or MRI, we did not have a proper ground truth to compare to. As we could not lengthen the standard procedure time, there was a clear time constraint imposed by the planning period in which we completed our study protocol resulting in different lengths of time case-to-case, often leading to rushed PD needle localizations that could have a negative impact on the obtained results. The relatively small sample size of 63 total needles may have played a role in the lack of statistically significant differences observed. As needle tip localization was completed after all needles were inserted to the

base of the prostate, there may be increased needle shadowing artifacts compared to the clinical scenario where needle tips are typically localized individually starting anteriorly and moving posteriorly to limit shadowing.

Future work includes redesigning the end-piece to improve signal generation and decrease operator dependence. As well, identification of multiple needles at once using PD US may be beneficial. Using an axial US image as shown in Fig. 4.9A, the needles can be seen as bright spots in the prostate. By vibrating the needle template directly, motion is induced in all needles, allowing for visualization of all needles in the PD US image (Fig. 4.9B). By obtaining a 3D US scan during PD US with mechanical perturbation, all needles could be simultaneously identified in 3D, with potential use as an initialization for an automatic needle segmentation approach[26] or as a confirmation for the needle position before the radiation treatment plan is generated. Furthermore, although our PD US-based method was first tested in prostate HDR-BT, it could be extended to any template-based brachytherapy procedure, such as gynecological brachytherapy, and more broadly to other minimally invasive procedures such as radiofrequency ablation. Future work will investigate these applications.

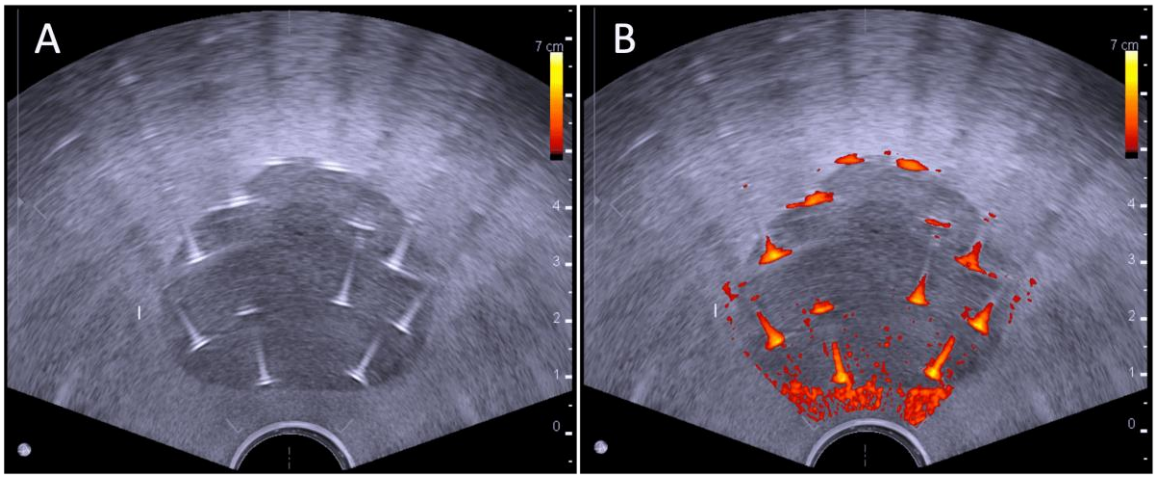

Figure 4.9. (A) Example axial B-mode US image of a tissue-mimicking agar phantom with an embedded agar prostate showing 12 implanted needles. (B) Example PD US image after vibration of the needle template, showing visualization of all needles.

4.5 Conclusions

This study investigated the development and validation of a PD US-based needle tip localization method utilizing a novel wireless mechanical oscillator in both phantom experiments and clinical HDR-BT treatments. The oscillator design is inexpensive and easy to manufacture, and the PD US method is easy to implement requiring no modifications to the clinical equipment or workflow. Phantom validation demonstrated improved tip localization accuracy relative to B-mode US for heavily shadowed needles, including the ability to accurately visualize needles previously not visible using B-mode US alone. Results of our prospective feasibility clinical trial demonstrated, for the first time to our knowledge, the promising utility of a PD US tip localization method in clinical HDR-BT procedures. PD US offered clear visualization of needle tips with a demonstrated ability to reduce tip error for outlier needles that were difficult to see in B-mode US, increasing confidence in tip localization, especially as needle visibility worsens. The proposed PD US method may offer improved needle tip localization accuracy and reduced variability, potentially improving HDR-BT treatment accuracy.

4.6 References

1. Yamada Y, Rogers L, Demanes DJ, et al. American Brachytherapy Society consensus guidelines for high-dose-rate prostate brachytherapy. *Brachytherapy*. 2012;11(1):20-32. doi:10.1016/J.BRACHY.2011.09.008

2. Yoshioka Y, Suzuki O, Isohashi F, et al. High-Dose-Rate Brachytherapy as Monotherapy for Intermediate- and High-Risk Prostate Cancer: Clinical Results for a Median 8-Year Follow-Up. *Int J Radiat Oncol*. 2016;94(4):675-682. doi:10.1016/J.IJROBP.2015.05.044

3. Hauswald H, Kamrava MR, Fallon JM, et al. Clinical Investigation High-Dose-Rate Monotherapy for Localized Prostate Cancer: 10-Year Results Radiation Oncology. *Int J Radiat Oncol Biol Phys*. 2016;94(4):667-674. doi:10.1016/j.ijrobp.2015.07.2290

4. Batchelar D, Gaztañaga M, Schmid M, Araujo C, Bachand F, Crook J. Validation study of ultrasound-based high-dose-rate prostate brachytherapy planning compared with CT-based planning. *Brachytherapy*. 2014;13(1):75-79.

doi:10.1016/J.BRACHY.2013.08.004

5. Morton GC. Prostate high-dose-rate brachytherapy: Transrectal ultrasound based planning, a technical note. *Pract Radiat Oncol*. 2015;5(4):238-240. doi:10.1016/J.PRRO.2014.12.009

6. Siebert F-A, Hirt M, Niehoff P, Kovács G. Imaging of implant needles for real-time HDR-brachytherapy prostate treatment using biplane ultrasound transducers. *Med Phys*. 2009;36(8):3406-3412. doi:10.1118/1.3157107

7. Tiong A, Bydder S, Ebert M, et al. A Small Tolerance for Catheter Displacement in High–Dose Rate Prostate Brachytherapy is Necessary and Feasible. *Int J Radiat Oncol*. 2010;76(4):1066-1072. doi:10.1016/J.IJROBP.2009.03.052

8. Mason J, Al-Qaisieh B, Bownes P, Thwaites D, Henry A. Dosimetry modeling for focal high-dose-rate prostate brachytherapy. *Brachytherapy*. 2014;13(6):611-617. doi:10.1016/j.brachy.2014.06.007

9. Hamper UM, Savader BL, Sheth S. Improved needle-tip visualization by color Doppler sonography. *Am J Roentgenol*. 1991;156(2):401-402. doi:10.2214/ajr.156.2.1898823

10. Feld R, Needleman L, Goldberg BB. Use of a needle-vibrating device and color Doppler imaging for sonographically guided invasive procedures. *Am J Roentgenol*. 1997;168(1):255-256. doi:10.2214/ajr.168.1.8976955

11. Jones CD, McGahan JP, Clark KJ. Color Doppler ultrasonographic detection of a vibrating needle system. *J Ultrasound Med*. 1997;16(4):269-274. doi:10.7863/jum.1997.16.4.269

12. Cabreros SS, Jimenez NM, Greer JD, Adebar TK, Okamura AM. Remote electromagnetic vibration of steerable needles for imaging in power Doppler ultrasound. *Robot Autom (ICRA), 2015 IEEE Int Conf*. 2015:2244-2249. doi:10.1109/ICRA.2015.7139496

13. Kuang Y, Hilgers A, Sadiq M, Cochran S, Corner G, Huang Z. Modelling and characterisation of a ultrasound-actuated needle for improved visibility in ultrasound-guided regional anaesthesia and tissue biopsy. *Ultrasonics*. 2016;69:38-46. doi:10.1016/J.ULTRAS.2016.02.018

14. Armstrong G, Cardon L, Vilkomerson D, et al. Localization of needle tip with color Doppler during pericardiocentesis: In vitro validation and initial clinical application. *J Am Soc Echocardiogr*. 2001;14(1):29-37. doi:10.1067/mje.2001.106680

15. Fronheiser MP, Wolf PD, Idriss SF, Nelson RC, Lee W, Smith SW. Real-time 3D color flow doppler for guidance of vibrating interventional devices. *Ultrason Imaging*. 2004;26(3):173-184. doi:10.1177/016173460402600304

16. Harmat A, Rohling RN, Salcudean SE. Needle tip localization using stylet vibration. *Ultrasound Med Biol*. 2006;32(9):1339-1348. doi:10.1016/j.ultrasmedbio.2006.05.019

17. Klein SM, Fronheiser MP, Reach J, Nielsen KC, Smith SW. Piezoelectric vibrating needle and catheter for enhancing ultrasound-guided peripheral nerve blocks. *Anesth Analg*. 2007;105(6):1858-1860. doi:10.1213/01.ane.0000286814.79988.0a

18. Fronheiser MP, Idriss SF, Wolf PD, Smith SW. Vibrating interventional device detection using real-time 3-D color Doppler. *IEEE Trans Ultrason Ferroelectr Freq Control*. 2008;55(6):1355-1362. doi:10.1109/TUFFC.2008.798

19. Reddy KE, Light ED, Rivera DJ, Kisslo JA, Smith SW. Color Doppler imaging of cardiac catheters using vibrating motors. *Ultrason Imaging*. 2008;30(4):247-250. doi:10.1177/016173460803000408

20. Adebar TK, Fletcher AE, Okamura AM. 3-D ultrasound-guided robotic needle steering in biological tissue. *IEEE Trans Biomed Eng*. 2014;61(12):2899-2910. doi:10.1109/TBME.2014.2334309

21. Greer JD, Adebar TK, Hwang GL, Okamura AM. Real-Time 3D Curved Needle Segmentation Using Combined B-Mode and Power Doppler Ultrasound. *Med Image Comput Comput Assist Interv*. 2014;17(2):381-388. doi:10.1007/978-3-319-10470-6_48

22. Orlando N, Snir J, Barker K, Hoover D, Fenster A. Power Doppler ultrasound imaging with mechanical perturbation for improved intraoperative needle tip identification during prostate brachytherapy: a phantom study. *Proc. SPIE 10951, Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling*. 2019;10951:111. doi:10.1117/12.2513082

23. Orlando N, Snir J, Barker K, Hoover D, Fenster A. Wireless oscillating device for power Doppler-based interstitial needle tip identification. *Proc. SPIE 11602, Medical Imaging 2021: Ultrasonic Imaging and Tomography*. 2021;116020J. doi:10.1117/12.2581103

24. Rickey DW, Picot PA, Christopher DA, Fenster A. A wall-less vessel phantom for Doppler ultrasound studies. *Ultrasound Med Biol*. 1995;21(9):1163-1176. doi:10.1016/0301-5629(95)00044-5

25. Fitzpatrick JM, West JB, Maurer CR. Predicting Error in Rigid-Body Point-Based Registration. *IEEE Trans Med* Imaging. 1998;17(5):694-702. doi:10.1109/42.736021

26. Gillies DJ, Rodgers JR, Gyacskov I, Roy P, Kakani N, Cool DW, Fenster A. Deep learning segmentation of general interventional tools in two-dimensional ultrasound images. *Med Phys*. 2020;47(10):4956-4970. doi:10.1002/mp.14427

Chapter 5

5 Validation of a surface-based deformable MRI-3D ultrasound image registration algorithm towards clinical implementation for interstitial prostate brachytherapy

In tumour-targeted brachytherapy procedures, the integration of MRI information into the intraoperative 3D US image is commonly done using cognitive fusion, resulting in variation between physicians. An automated registration algorithm could standardize the registration procedure, potentially reducing operator dependence and facilitating dose escalation to the MR-defined tumour. The purpose of Chapter 5 is to describe the development and clinical implementation of a surface-based deformable MRI-3D US image registration algorithm designed for prostate brachytherapy.

The contents of this chapter have been submitted for publication in *Brachytherapy* and are currently under peer review: Orlando N, Edirisinghe C, Gyacskov I, Vickress J, Sachdeva R, Gomez JA, D'Souza D, Velker V, Mendez LC, Bauman G, Fenster A, and Hoover DA.

5.1 Introduction

Conventional prostate brachytherapy (BT) aims to deliver a uniform dose to the whole prostate, with maximum dose limited by the surrounding organs-at-risk.[1] Studies have demonstrated that the site of local recurrence following radiation therapy is often the dominant intraprostatic lesion (DIL).[2–4] Tumour-targeted brachytherapy has been suggested as an alternative treatment technique, proposing the escalation of dose to the DIL while maintaining the conventional whole-gland dose, potentially offering improved tumour control while limiting toxicity.[5,6] Unfortunately, US imaging lacks the sensitivity to effectively detect and localize the DIL and other sensitive organs-at-risk such as the neurovascular bundles (NVB).[7] Magnetic resonance imaging (MRI) provides higher soft tissue contrast compared to US, allowing for effective DIL identification and localization.[8–

[10] Intraoperative MRI needle guidance is costly and challenging, requiring incorporation of MRI information into the operating room environment where US is the primary imaging modality. In many centres, the current standard of care involves rigid or cognitive fusion of MR contours into the US image, done intraoperatively by the physician. This results in inter-physician variability and high operator dependence. In addition, this manual approach is cumbersome if performed in the operating room, adding up to 15 minutes to complete both the registration and contour transfer. An automated MR-US registration approach is required to standardize the registration and reduce procedure time.

Conventional automated and semi-automated MR-US registration approaches have been proposed for use in minimally invasive prostate procedures.[11–15] The majority of these image registration methods have focused on MR-US fusion-guided prostate biopsy, with both rigid[16–18] and non-rigid approaches.[11,12,19–22] The presence of the US transducer or endorectal coil in the rectum during US or MR imaging, respectively can deform the prostate, suggesting deformable registration approaches may offer improved registration accuracy.[15] Shaaer *et al*. proposed a deformable registration approach for prostate BT based on a modified rigid registration algorithm using 6 degrees-of-freedom, demonstrating increased registration accuracy compared to rigid registration in 10 patients.[14] More recently, deep learning-based registration approaches have been proposed,[23–28] utilizing methods including segmentation-based algorithms[24–26] and adversarial networks.[27] These approaches require large, curated image datasets, which can be rare in BT settings, presenting a barrier for widespread clinical translation.

Prostate biopsy-based registration approaches typically use different 3D US image acquisition methods than those used in BT. However, no study focused on BT applications for the purpose of dose sparing has described the registration and translation of MR-defined organ-at-risk contours outside of the prostate, such as the NVBs. Furthermore, prospective validation and clinical implementation for these approaches is lacking. An automated algorithm capable of accurate MR-US deformable registration of both target and organ-at-risk contours, designed for BT, including demonstrated clinical feasibility, may improve the clinical translation potential for MR-US fusion-guided BT procedures.

In this paper, we report on the development and validation of a surface-based MR-US deformable image registration (DIR) algorithm for prostate BT, including registration

and translation of MR-defined targets and organs-at-risk into the intraoperative US image. Our DIR method was validated in tissue-mimicking phantom experiments with varying deformation levels as well as in retrospective high-dose-rate brachytherapy (HDR-BT) clinical cases. Clinical implementation of our method demonstrated the feasibility of intraoperative registration, facilitating tumour-targeted dose escalation. This has the potential to reduce procedure time, standardize the registration procedure between physicians, thus reducing operator dependence, and could be extended to any minimally invasive prostate procedure employing MR and US imaging.

5.2 Methods and materials

5.2.1 Deformable surface-based image registration

The workflow of our DIR algorithm is summarized in Fig. 5.1. The algorithm takes as input an MR image, 3D US image, and corresponding 3D prostate contours. The algorithm is fully automated, requiring no manual initialization, but does require a pre-operative preprocessing step if the 3D US and MR images are not in the same orientation. Surface contours can be equidistant parallel 2D contours or radial 2D contours with equal angular separation about an arbitrary axis. Using these 2D contours, 3D surfaces are formed using Delaunay Triangulation, available in the Visualization Toolkit (VTK).[29]

The registration workflow has two fully automated steps: rigid translation, scaling, and alignment, and non-rigid warping. First, 3D cuboids are determined that enclose all the points in each of the US-and MR-defined prostate surfaces. The center points of each cuboid are then matched, overlapping the surfaces. The MR-defined cuboid is then scaled in the x, y, and z-directions to align its edges with the US-defined cuboid and the triangulated surfaces are scaled accordingly. Then, iterative closest point (ICP) transformation is used to align the MR and US prostate surfaces through a series of affine transformations.[29] Finally, thin plate spline (TPS) deformation is used for non-rigid warping of the MR prostate surface to match the US prostate surface. This is accomplished by creating a sphere with a 5 mm margin around the aligned surfaces. From the center point of the sphere, 2304 equidistant lines are drawn outwards to the sphere surface. The MR

and US prostate surface points that intersect with each line are labelled as corresponding point pairs. Using the VTK TPS module, a transformation is then computed to warp the MR prostate image such that corresponding points on the MR prostate surface match the points on the US prostate surface.[29] Using the computed transformation, other organs contoured in the MR image are warped and translated to the US image space.

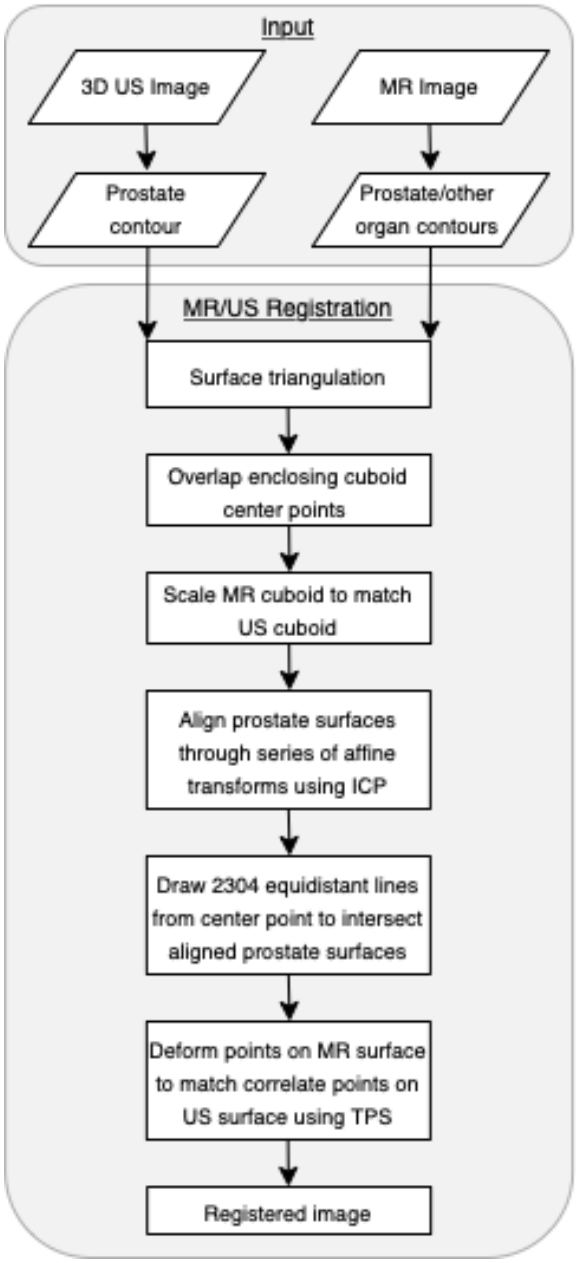

Figure 5.1. Summary of proposed deformable image registration algorithm workflow.

The DIR algorithm was implemented on a personal computer using a Windows 10 operating system (Microsoft, Redmond, WA, USA) with an Intel Core i7-4770 central processing unit (Intel Corporation, Santa Clara, CA, USA), 32.0 GB of memory, and a 6 GB Ge-Force GTX TITAN (NVIDIA Corporation, Santa Clara, CA, USA) graphics processing unit.

5.2.2 Phantom validation

The proposed DIR algorithm was first validated in phantom using the 053L tissue-equivalent US deformable prostate phantom (CIRS Inc., Norfolk, VA, USA) for use with an endocavity US transducer. A simulated 53 cm^3 prostate, urethra, and three stiff 1 cm diameter spherical simulated lesions distributed uniformly throughout the prostate are contained within the $11.5 \times 7.0 \times 9.5$ cm^3 acrylic container.

3D US images of the phantom were acquired at three different deformation levels to simulate potential clinical scenarios. Deformation levels were no deformation, moderate deformation, and large deformation, controlled by the amount of upward force applied to the phantom by the transducer. The US images were acquired using a BK3000 system with an E14CL4b biplane endocavity transducer (BK Medical, Peabody, MA, USA). The transducer was supported by an encoded CIVCO EX^3 Stepper (CIVCO Medical Solutions, Coralville, IA, USA), allowing for axially reconstructed 3D US step-back acquisition with a 1-mm step spacing using Vitesse v4.03 (Varian Medical Systems, Palo Alto, CA, USA). 3D US images had a size of $[1100 \times 700 \times 73]$ voxels with a voxel size of $[0.083 \times 0.110 \times 1.00]$ mm^3. B-mode US frequency was 9 MHz with a depth of 6.5 cm for all image acquisitions.

The phantom was imaged using 1.5T and 3T MR scanners to simulate different clinical scenarios and examine the impact of MR image quality on registration accuracy. The 1.5T MR image was acquired using a GE SIGNA HDxt scanner (GE Healthcare Systems, Chicago, IL, USA) with an 8-channel body coil. An axial T2-weighted fast spin echo (FSE) sequence was used, acquiring a 3D image with a size of $[512 \times 512 \times 30]$ voxels, voxel size of $[0.430 \times 0.430 \times 5.00]$ mm^3, and slice thickness of 2 mm. The 3T MR image was acquired using a GE Discovery MR750 scanner (GE Healthcare Systems, Chicago,

IL, USA) with a 32-channel cardiac coil. For consistency, an axial T2-weighted FSE sequence was once again used for the 3T MR image acquisition, resulting in a 3D image with a size of [512×512×65] voxels, voxel size of [0.352×0.352×2.00] mm^3, and slice thickness of 2 mm.

The prostate, urethra, and spherical landmarks were manually contoured by a trained observer (NJO) using custom-made image visualization software.[30] To overcome the limitation of large slice spacing for the MR images, manual prostate and landmark contours were generated using a radial approach. Contours were drawn on 18 radial slices extracted every 10 degrees about an arbitrary axis of rotation defined by the observer at the object's centroid. The boundary points of neighbouring slices were then connected with a smoothing filter, generating a 3D surface. For the 3D US images, the parallel manual contouring method was used on the axial/acquisition planes, which were separated by 1mm.

The 1.5T and 3T MR images were each registered to the 3D US images with three levels of deformation, resulting in a total of 6 registrations. Each registration used our surface-based DIR algorithm and the manually defined prostate surfaces in the MR and 3D US images. For comparison, manual rigid registrations of the 1.5T and 3T MR images to the 3D US images based on the prostate surfaces were completed using MIM 7.0.5 (MIM Software Inc., Cleveland, OH, USA).

5.2.3 Retrospective clinical validation

Our DIR algorithm was retrospectively validated using three patients who were imaged with 3D US and MRI and underwent standard whole-gland HDR-BT treatment. All 3D US images were acquired as described in section 5.2.2. 3D US image sizes were [1100×700×52], [1100×700×57], and [1100×700×46] voxels for patients 1, 2, and 3, respectively. All MR images were acquired with 1.5T scanners. Images of patients 1 and 2 were acquired using a Philips Achieva 1.5T MRI system (Philips Healthcare, Andover, MA, USA) with an axial T2-weighted pulse sequence generating images with a voxel size of [0.3125×0.3125×3.00] mm^3 and image sizes of [576×576×32] and [576×576×30], respectively. The image of patient 3 was acquired using a GE Optima MR450w 1.5T MRI

System (GE Healthcare Systems, Chicago, IL, USA) with an axial T2-weighted FSE pulse sequence to produce a [512×512×76] voxel image with voxel size of [0.3906×0.3906×3.00] mm^3.

The prostate in each 3D US image and the prostate, urethra, clinical target volume (CTV), and left/right NVBs in each MR image were manually contoured by a radiation oncologist specializing in brachytherapy (physician 1) in parallel axial slices using contouring tools within ARIA 15.6 (Varian Medical Systems, Palo Alto, CA, USA). Using these prostate contours, the surface-based DIR algorithm was used to register the MR image to the 3D US image. For comparison, a manual rigid registration from MRI to 3D US was completed for each clinical case using MIM. The proposed DIR algorithm was then compared to the standard approach at our cancer centre, which involves rigid MR-to-US registration and transfer of MR contours to the US space followed by manual physician editing using a cognitive fusion approach. Three physicians with 5, 7, and 21 years of experience performing prostate brachytherapy procedures completed these cognitive fusions using the contours from physician 1.

5.2.4 Evaluation of registration accuracy

Phantom validation

As the spherical simulated lesions and urethra were clearly visible in both US and MR images, direct comparisons between the contour locations in the 3D US image and the locations of the MR contours registered into the 3D US image could be completed for all registrations. The target registration error (TRE)[31] was computed for each registration by comparing the centroid locations of each of the three spherical landmarks in the 3D US image to the registered 1.5T and 3T MR-defined contour centroid locations. The centroid locations were selected by a trained observer (NJO) using custom-made image visualization software.[30] In addition, the fiducial localization error (FLE)[31] was computed for both the 3D US contours and the registered MRI contours to assess the variability in the centroid localization. Centroid localization was repeated three times by the same observer with a one-week washout period between each localization.

Using tools within MIM, the Dice similarity coefficient (DSC) metric was computed to assess the overlap between the US-identified contours and the registered MR contours of the three spherical landmarks and the urethra.

Retrospective clinical validation

Due to the lack of reproducible identification of anatomical landmarks visible in both US and MRI, evaluation of registration accuracy utilized the DSC metric. The DSC between the registered contours generated using our proposed DIR algorithm and the manual rigid registration were compared to the manual MR-to-US cognitive registrations performed by the three physicians. Inter-physician variability was computed as the average DSC when comparing between physician cognitive fusion outputs. This allowed for comparison of the algorithm accuracy and variability to the variability between different physicians. The DSC metric was computed for the CTV and the left and right NVBs in MIM.

5.2.5 Clinical implementation

Our proposed DIR algorithm was implemented for the first patient of a prospective randomized clinical trial at the London Regional Cancer Program approved by the Research Ethics Board at Western University. The trial involved a two-part clinical procedure including an initial targeted biopsy based on MRI and prostate specific membrane antigen-positron emission tomography (PSMA-PET) imaging followed by standard HDR-BT treatment with boosted dose to the lesion identified in MRI and PSMA-PET. Pre-procedural MRI and PSMA-PET images were acquired simultaneously using a Siemens Biograph mMR PET-MR system (Siemens Healthineers, Erlangen, Germany). The prostate, urethra, rectum, MR-defined gross tumour volume (GTV), PSMA-PET-defined GTV, and CTV were manually contoured by a radiation oncologist.

In the operating room, a pre-biopsy 3D US image was first acquired using the BK3000 system with the same image parameters described in section 5.2.2. The prostate was contoured in the pre-biopsy 3D US image by the treating radiation oncologist using Vitesse v4.03. MR-to-US registration was completed in the operating room using our DIR algorithm and the physician-defined prostate contours, thus registering the PET-MR

contours to the 3D US image space. The registered CTV and combined GTV contours were imported into Vitesse using existing rigid registration tools and were subsequently used to guide 3 targeted biopsy needles. The standard HDR-BT procedure was initiated following the targeted biopsy. With all needles implanted, a second 3D US image was acquired, and the prostate was contoured by the treating radiation oncologist. MR-to-US registration was once again completed using our DIR algorithm in the operating room. The registered combined GTVs and CTV were edited by the radiation oncologist and were subsequently used during the treatment planning process to ensure dose coverage and facilitate accurate dose escalation to the CTV. To quantify the registration accuracy of our DIR algorithm, one of the targeted biopsy cores was analyzed pathologically to determine the presence of cancer. Due to the lack of ground truth or reproducible anatomical landmarks visible in both US and MRI, further evaluation of registration accuracy was limited to qualitative comparison.

5.2.6 Statistical analysis

All statistical analyses were performed in GraphPad Prism 9.3.1 (Graphpad Software, Inc., San Diego, CA, USA). Shapiro-Wilk tests were used to assess the normality of distributions. Failed normality tests led to the use of nonparametric statistical tests, presented in parentheses in the following paragraph. The significance level for statistical analysis was set such that the probability of making a type I error is less than 5% ($p < 0.05$). Statistically significant differences are denoted simply as significant for the remainder of this manuscript.

A comparison of registration accuracy in phantom in terms of DSC and TRE between the three deformation levels for our DIR algorithm and manual rigid registration was completed using a two-way ANOVA with Tukey's correction for multiple comparisons. This test was repeated for comparisons of DSC and TRE between 1.5T and 3T MR images registered to 3D US images with varying deformation levels. 1.5T and 3T MRI FLE, and US and total MRI FLE for the varying deformation levels were compared using the same two-way ANOVA. For the retrospective clinical validation, registration accuracy in terms of DSC for our DIR algorithm, rigid registration, and physician cognitive

fusion were compared using one-way ANOVA with Tukey's multiple comparisons tests (Kruskal-Wallis tests with Dunn's multiple comparisons tests).

5.3 Results

5.3.1 Phantom validation

Examples of the registration results obtained using our DIR algorithm and manual rigid registration for three deformation levels of the phantom are shown in Fig. 5.2.

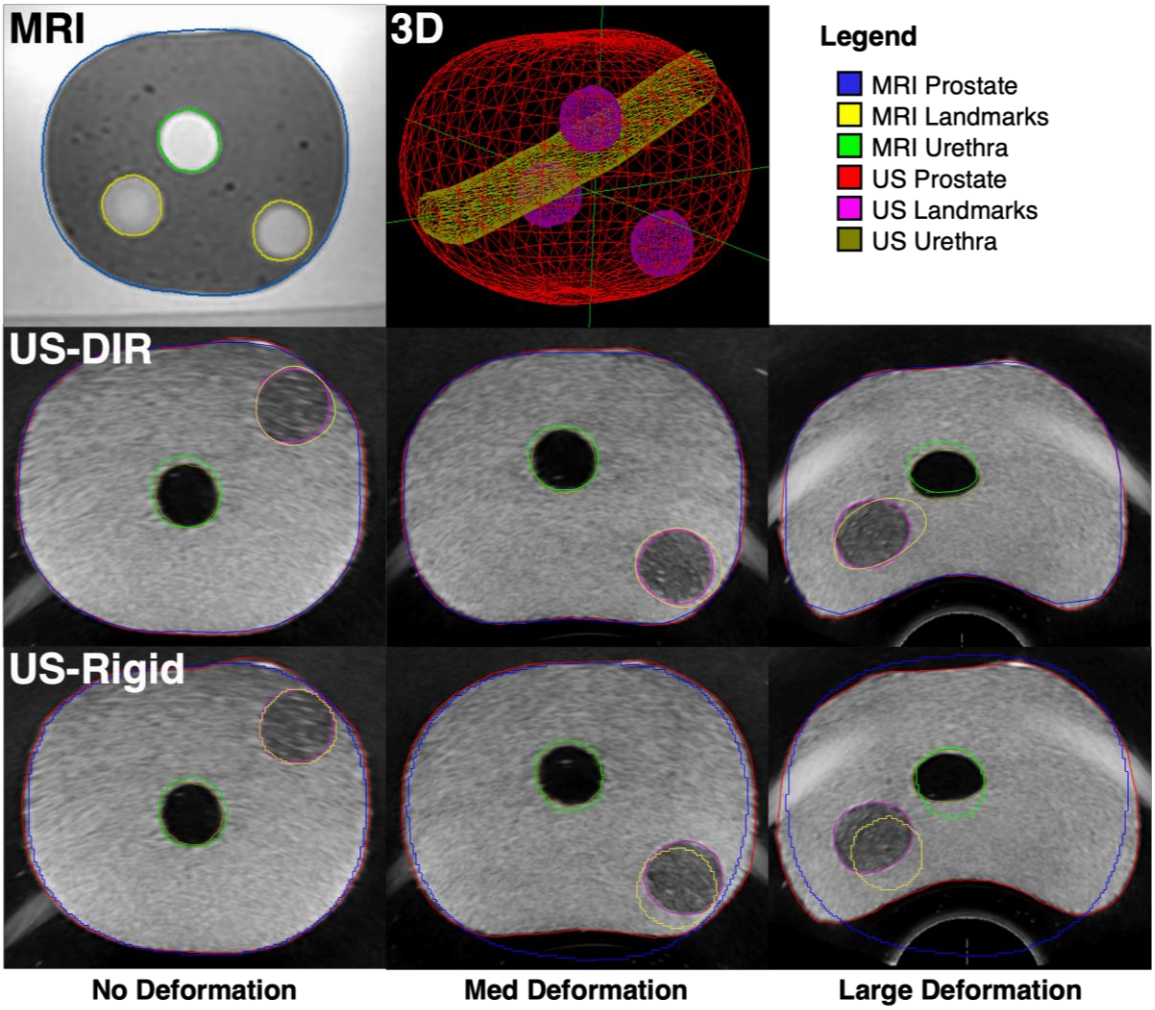

Figure 5.2. Example phantom registration results for three deformation levels ranging from no deformation to large deformation shown in columns from left to right, respectively. The

top row shows an example 3T MR image, 3D surface view of the phantom highlighting landmark distribution, and contour legend. The middle and bottom rows show MR-to-US registration results generated using our DIR algorithm and manual rigid registration, respectively. Each deformation level examines a different landmark location.

Quantitative comparisons of the registration accuracy using our DIR algorithm and rigid registration in terms of landmark DSC and TRE are shown in Figs. 5.3A and 5.3B, respectively, with complete results provided in Table 5.1. Mean DSC across all deformation levels and MR images were 0.80±0.12 and 0.72±0.17 for the DIR algorithm and rigid registration, respectively. Mean TRE for the embedded spherical landmarks using the DIR algorithm and rigid registration was 0.94±0.49 mm and 1.88±1.24 mm, respectively. Although registration type and deformation level had no significant effect on DSC for the prostate and urethra, the mean prostate DSC was lower for the rigid registrations. For the spherical landmarks, registration type, deformation level, and the interaction between them had significant effect on DSC score, with the DSC increasing when using the DIR algorithm and decreasing as deformation increases. These results were mirrored for the TRE measurements based on landmark centroid location, with registration type, deformation level, and the interaction between them significantly impacting TRE values. MR field strength showed no significant main effect on DSC and TRE for algorithmic or rigid registrations. Landmark location within the prostate had no significant effect on DSC and TRE for either registration method. Total registration computation time ranged from 85s to 150s depending on the number of registered contours.

Landmark centroid localization for the three deformation levels had a mean ± standard deviation FLE of 0.16±0.07 mm (N=54), 0.19±0.09 mm (N=27), and 0.18±0.10 mm (N=27) for US, registered 1.5T MR, and registered 3T MR contours, respectively, for a total mean of 0.17±0.08 mm (N=108). There was no significant difference in FLE between 1.5T and 3T MR, between US and MR, or based on landmark position.

Table 5.1. Mean ± standard deviation registration error metrics for our deformable image registration (DIR) algorithm and manual rigid registration evaluated using 1.5T and 3T MRI phantom images with three deformation levels. Dice similarity coefficient (DSC) scores are provided for the prostate, urethra, and three spherical landmarks. Target registration error (TRE), in mm, is provided for the three spherical landmark centroids.

	1.5T MRI			3T MRI		
DIR Algorithm	No Def.	Med Def.	Large Def.	No Def.	Med Def.	Large Def.
DSC						
Prostate	0.98	0.98	0.97	0.98	0.98	0.98
Landmarks	0.79±0.02	0.74±0.03	0.64±0.03	0.82±0.06	0.78±0.04	0.64±0.06
Urethra	0.84	0.80	0.76	0.81	0.79	0.73
TRE (mm)						
Landmarks	0.52±0.36	1.13±0.81	1.34±0.23	0.58±0.01	0.75±0.04	1.32±0.43
Rigid Registration	No Def.	Med Def.	Large Def.	No Def.	Med Def.	Large Def.
DSC						
Prostate	0.94	0.91	0.83	0.95	0.93	0.85
Landmarks	0.76±0.05	0.65±0.03	0.45±0.06	0.81±0.03	0.69±0.05	0.46±0.17
Urethra	0.83	0.84	0.78	0.84	0.80	0.68
TRE (mm)						
Landmarks	0.77±0.07	1.79±0.46	3.22±0.66	0.50±0.09	1.73±0.42	3.24±1.42

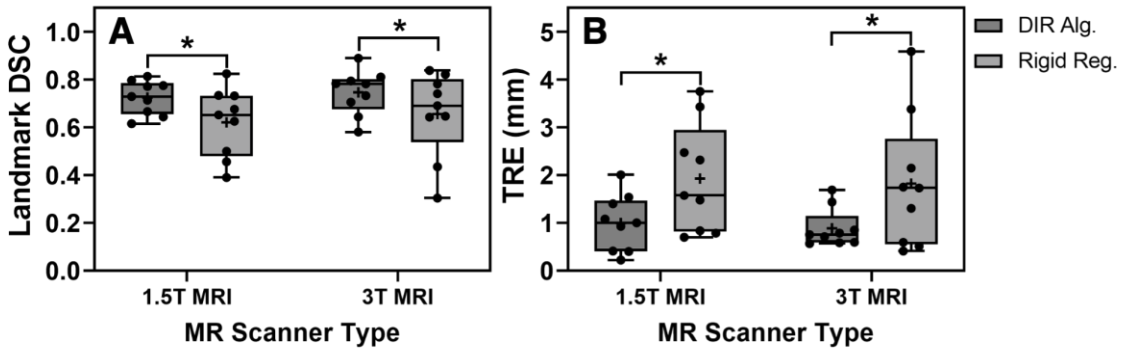

Figure 5.3. Box and whisker plots showing (A) landmark Dice similarity coefficient (DSC) and (B) target registration error (TRE) for 1.5T and 3T MRI images registered to 3D US using our DIR algorithm and manual rigid registration in phantom, averaged across three deformation levels. Boxes denote the interquartile range while whiskers show the

minimum and maximum values. A bar and plus sign represent the median and mean, respectively. Statistically significant differences (p<0.05) are denoted by an asterisk.

5.3.2 Retrospective clinical validation

Example clinical registration results for one of our three retrospective cases are shown in Fig. 5.4, including results obtained using our DIR algorithm, manual rigid registration, and physician cognitive fusion.

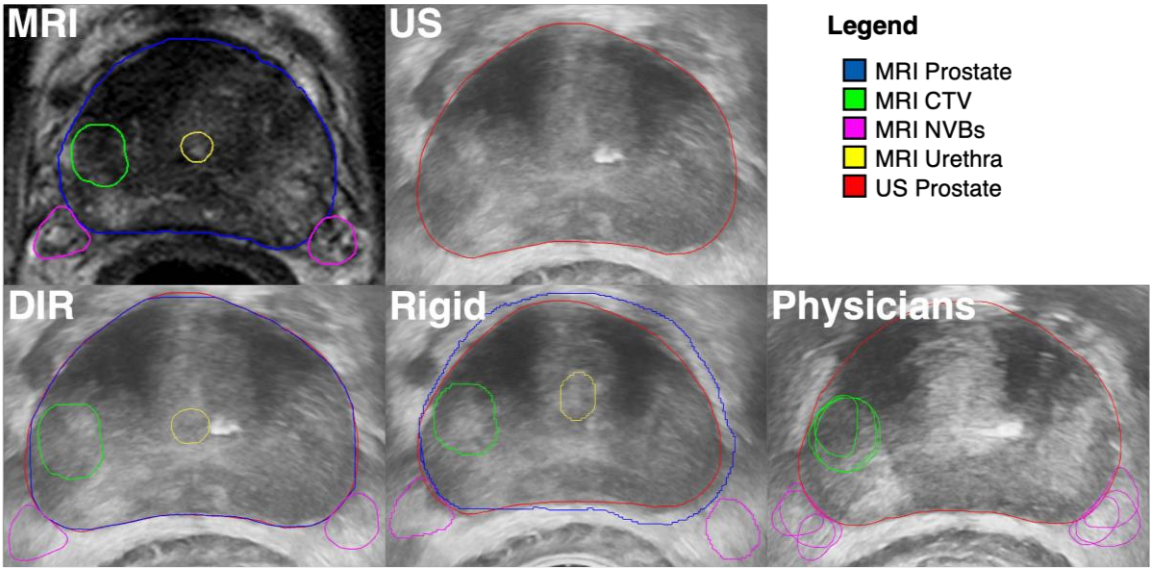

Figure 5.4. Example clinical registration results for one retrospective HDR-BT case. The top row shows the physician-contoured MRI and US as well as the contour legend. The bottom row shows the MR-to-US registration using our DIR algorithm, manual rigid registration, and cognitive fusion completed by three physicians, respectively.

Figure 5.6 shows the DSC score for the MR-defined CTV and NVBs registered to US using our DIR algorithm, rigid registration, and physician cognitive fusion. Complete results are provided in Table 5.2, including comparisons to each physician individually and all together, as well as the inter-physician variability. Mean DSC between the US and registered MR-defined prostate was 0.98±0.01 and 0.89±0.02 for the DIR algorithm and

manual rigid registration, respectively. When comparing the DIR and rigid registration output to the individual physician cognitive fusion results, only comparisons to physician 1 showed significant differences, with the DIR algorithm demonstrating significantly higher DSC score for the right NVB relative to the rigid registration approach. Furthermore, when comparing DSC scores averaged for comparisons to all physicians, there was no significant difference observed between the DIR algorithm and rigid registration results and the inter-physician variability.

Table 5.2. Mean ± standard deviation Dice similarity coefficient (DSC) values averaged from three clinical cases registered using our deformable image registration (DIR) algorithm and manual rigid registration. Columns show DSC values for the clinical target volume (CTV) and the left and right neurovascular bundles (NVB) compared to cognitive fusion results for each physician individually, averaged DSC results compared to all physicians, and inter-physician variability, from left to right respectively.

		Total Mean DSC				
DIR Alg.	Contour	Compared to Physician 1	Compared to Physician 2	Compared to Physician 3	Compared to all Physicians	Inter-physician var.
	CTV	0.66±0.11	0.49±0.21	0.50±0.37	0.55±0.23	0.56±0.29
	NVB-L	0.51±0.18	0.50±0.08	0.59±0.09	0.53±0.12	0.52±0.18
	NVB-R	0.48±0.09	0.50±0.05	0.50±0.08	0.50±0.06	0.41±0.14
	Mean	0.55±0.14	0.49±0.11	0.53±0.20	0.53±0.15	0.50±0.21
Rigid Reg.	Contour	Compared to Physician 1	Compared to Physician 2	Compared to Physician 3	Compared to all Physicians	Inter-physician var.
	CTV	0.67±0.03	0.53±0.19	0.49±0.34	0.56±0.21	0.56±0.29
	NVB-L	0.49±0.13	0.41±0.29	0.51±0.24	0.47±0.20	0.52±0.18
	NVB-R	0.30±0.04	0.50±0.04	0.65±0.17	0.48±0.18	0.41±0.14
	Mean	0.48±0.17	0.48±0.18	0.55±0.24	0.50±0.19	0.50±0.21

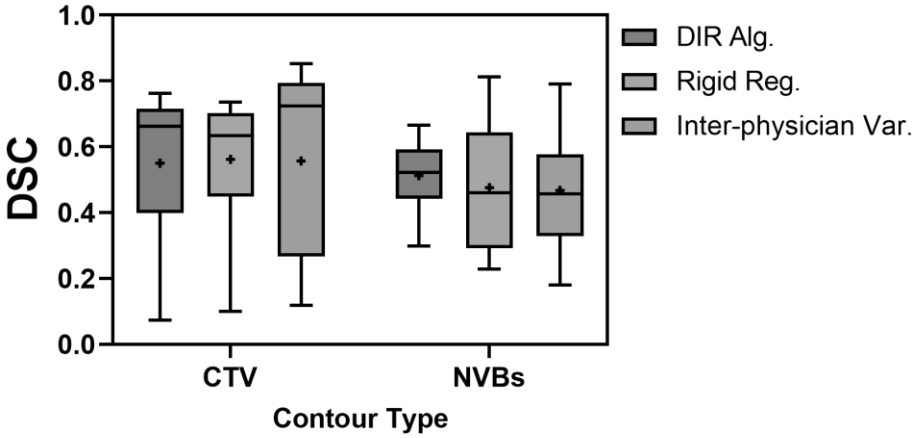

Figure 5.5. Box and whisker plots showing the Dice similarity coefficient (DSC) score for clinical target volume (CTV) and neurovascular bundle (NVB) contours registered using our DIR algorithm, manual rigid registration, and manual cognitive fusion by the physician, presented as inter-physician variability. Boxes denote the interquartile range while whiskers show the minimum and maximum values. A bar and plus sign represent the median and mean, respectively.

5.3.3 Clinical implementation

Our DIR algorithm was implemented intraoperatively for one patient who consented to a prospective clinical trial approved by the Research Ethics Board at Western University. Example MR as well as pre- and post-implant US images and the corresponding registration results are shown in Fig. 5.6. After pathological evaluation, the biopsy core targeted to the combined PET-MRI GTV based on the registration results obtained using our DIR algorithm, was positive for prostate cancer. The corresponding histology slide is shown in Fig. 5.7.

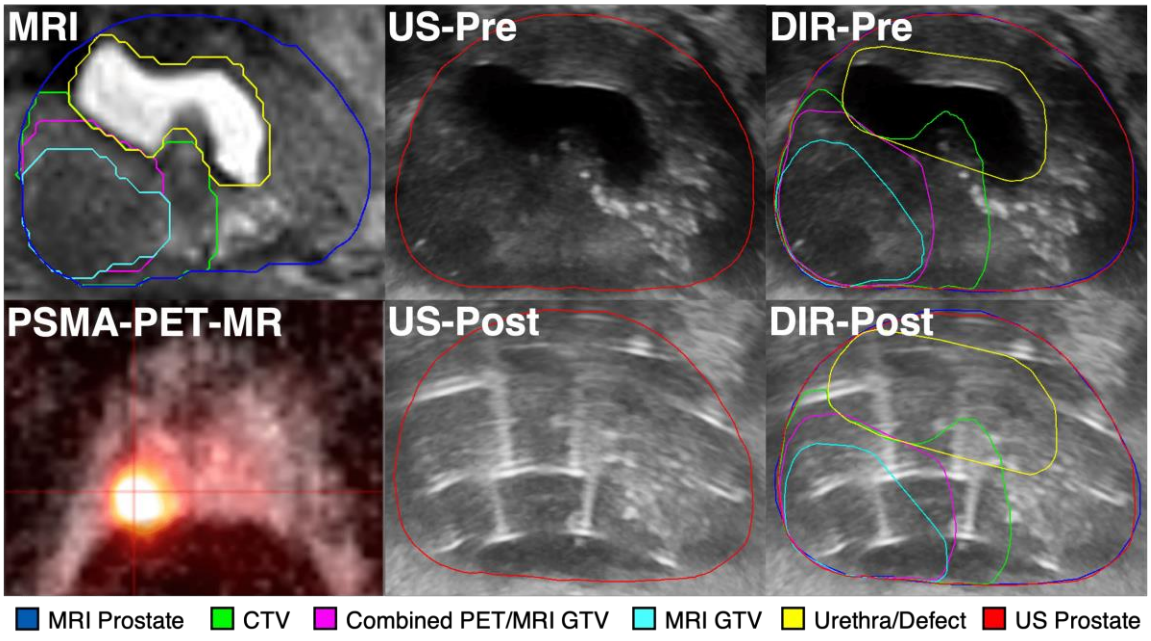

Figure 5.6. Example intraoperative clinical registration results in a prospective clinical case. The physician-contoured MR image is shown in the top left pane. The pre- and post-implant US image and registration results generated using our DIR algorithm are shown in the top and bottom row, respectively. The contour legend is provided in the bottom left pane.

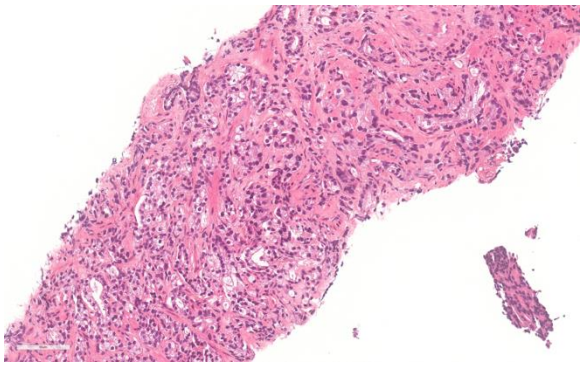

Figure 5.7. Histology slide demonstrating the presence of cancer for the GTV-targeted biopsy result obtained using our DIR algorithm.

5.4 Discussion

Results from the phantom experiments showed that the deformation level and registration type had a significant impact on the segmentation performance, highlighting the need for deformable registration. This is mirrored in the qualitative results shown in Fig. 5.2, which showed similar results between the rigid and deformable methods when no deformation was applied. In the second and third column where deformation was applied, there was significant misalignment between the MR-defined landmarks and US-defined landmarks when using the rigid approach. In addition, the MR prostate contour no longer matches the US prostate contour, which would make accurate contour translation from the MR to the US images and treatment planning difficult. Both DSC and TRE results showed that our DIR algorithm was robust to MR field strength and landmark location within the prostate, which is a critical finding as both these parameters vary between clinical cases. When examining the TRE results, it is important to consider the FLE for centroid identification. A TRE less than 1 mm for all deformation levels with our DIR algorithm and a mean FLE of 0.17 mm suggest that the TRE value is only minimally affected by the observer's ability to reliably identify the landmark centroid. With an average computation time under two minutes, our DIR algorithm may also reduce procedure time, as manual rigid registration and physician cognitive fusion may require up to 15 minutes in the operating room.

The retrospective clinical validation results show similar mean DSC scores between our DIR algorithm, rigid registration, and physician cognitive fusion. This may be due in part to the low levels of deformation in our clinical cases, compounded by the lack of endorectal coil for MR image acquisition. We did observe a greater than 5% reduction in variability for our DIR algorithm compared to the inter-physician variability. This was most prominent for the NVB contours where mean DSC was 5% higher and variation was approximately 7% lower when using the DIR algorithm compared to cognitive fusion, highlighting the ability of our algorithm to accurately translate small contours outside of the prostate volume, which is critical for dose-avoidance during treatment planning. Inter-physician variability is demonstrated in the bottom right pane of Fig. 5.4, where there is inconsistent placement of both the CTV and NVBs between physicians. Deformation may

impact posterior structures such as the NVBs and peripheral zone lesions more than anterior targets, contributing to the DIR algorithm's observed improvement in registration accuracy relative to the rigid approach for the NVB. This reduced variation afforded by the DIR algorithm demonstrates that standardized use of our method would result in more consistent registration results compared to the current inter-physician variability, thus improving treatment planning consistency and highlighting the potentially clinical utility of our method.

Our DIR algorithm was successfully implemented in a prospective clinical case, demonstrating the feasibility of integrating our method into the clinical workflow. The ease of integration highlights the clinical translation potential of our method, as it required only minimal pauses in the workflow and the registration result integrated seamlessly into the clinical software. This allowed for a successful targeted biopsy of the combined PSMA-PET/MR-GTV, confirmed histologically as shown in Fig. 5.7. Qualitative results shown in Fig. 5.6 highlight the large urethra defect resulting from a previous transurethral resection. Our DIR algorithm was shown to be robust to this artifact, successfully registering the prostate contours and translating the CTV and GTVs. Application of the algorithm pre-implant to target biopsy needles and post-implant to facilitate dose escalation demonstrated the versatility of our method, further highlighting translational potential. The method could be easily extended to other minimally invasive prostate procedures employing MR-imaging such as fusion-guided prostate biopsy and tumour-targeted low-dose-rate BT.

The main limitation of our surface-based DIR algorithm is its reliance on accurate and consistent manual contouring, particularly for the prostate in US and MRI. Implementation of defined contouring guidelines for the prostate in both MR and US would reduce contouring variability between physicians, further improving registration consistency and accuracy. The addition of AI-based automatic prostate contouring tools[32,33] could also help to standardize the contouring process and further reduce procedure times. The retrospective clinical dataset used to validate our algorithm is small with only three cases, so further validation with a larger dataset is an important next step in our future work. Furthermore, validation using MR images acquired with endorectal coils would be important, as we expect the increased deformation necessitates a DIR approach. Future work will also include expanding the prospective implementation of our algorithm to

additional patients and additional procedure types, including an approved study assessing the benefits of NVB avoidance in BT.

5.5 Conclusions

This study investigated the development and validation of a surface-based MR-US DIR algorithm for prostate BT. This approach demonstrated improved registration accuracy compared to rigid registration, as well as decreased variation compared to the clinical standard of physician cognitive fusion. Prospective clinical implementation demonstrated the feasibility of applying our DIR algorithm intraoperatively, facilitating targeted biopsies and dose escalation to a PET-MR-defined lesion. This algorithm offers the potential to reduce operator dependence and standardize the registration procedure between physicians.

5.6 References

1. Hoskin P, Rojas A, Ostler P, et al. High-dose-rate brachytherapy alone given as two or one fraction to patients for locally advanced prostate cancer: Acute toxicity. *Radiother Oncol*. 2014;110(2):268-271. doi:10.1016/j.radonc.2013.09.025

2. Pucar D, Hricak H, Shukla-Dave A, et al. Clinically Significant Prostate Cancer Local Recurrence After Radiation Therapy Occurs at the Site of Primary Tumor: Magnetic Resonance Imaging and Step-Section Pathology Evidence. *Int J Radiat Oncol Biol Phys*. 2007;69(1):62-69. doi:10.1016/j.ijrobp.2007.03.065

3. Arrayeh E, Westphalen AC, Kurhanewicz J, et al. Does local recurrence of prostate cancer after radiation therapy occur at the site of primary tumor? Results of a longitudinal MRI and MRSI study. *Int J Radiat Oncol Biol Phys*. 2012;82(5):e787-793. doi:10.1016/j.ijrobp.2011.11.030

4. Cellini N, Morganti AG, Mattiucci GC, et al. Analysis of intraprostatic failures in patients treated with hormonal therapy and radiotherapy: implications for conformal therapy planning. *Int J Radiat Oncol Biol Phys*. 2002;53(3):595-599. doi:10.1016/S0360-3016(02)02795-5

5. Bauman G, Haider M, Van der Heide UA, Ménard C. Boosting imaging defined dominant prostatic tumors: A systematic review. *Radiother Oncol*. 2013;107(3):274-281. doi:10.1016/J.RADONC.2013.04.027

6. Von Eyben FE, Kiljunen T, Kangasmaki A, Kairemo K, Von Eyben R, Joensuu T. Radiotherapy Boost for the Dominant Intraprostatic Cancer Lesion - A Systematic Review and Meta-Analysis. *Clin Genitourin Cancer*. 2016;14(3):189-197. doi:10.1016/j.clgc.2015.12.005

7. Catalona WJ, Smith DS, Ratliff TL, et al. Measurement of Prostate-Specific Antigen in Serum as a Screening Test for Prostate Cancer. *http://dx.doi.org/101056/NEJM199104253241702*. 2010;324(17):1156-1161. doi:10.1056/NEJM199104253241702

8. Thompson J, Lawrentschuk N, Frydenberg M, Thompson L, Stricker P. The role of magnetic resonance imaging in the diagnosis and management of prostate cancer. *BJU Int*. 2013;112(SUPPL. 2):6-20. doi:10.1111/BJU.12381

9. Bauman G, Haider M, Van Der Heide UA, Ménard C. Boosting imaging defined dominant prostatic tumors: A systematic review. *Radiother Oncol*. 2013;107(3):274-281. doi:10.1016/j.radonc.2013.04.027

10. Bouchelouche K, Turkbey B, Choyke P, Capala J. Imaging prostate cancer: An update on positron emission tomography and magnetic resonance imaging. *Curr Urol Rep*. 2010;11(3):180-190. doi:10.1007/S11934-010-0105-9/FIGURES/5

11. Sun Y, Qiu W, Romagnoli C, Fenster A. 3D non-rigid surface-based MR-TRUS registration for image-guided prostate biopsy. *Proc. SPIE 9036, Medical Imaging 2014: Image-Guided Procedures, Robotic Interventions, and Modeling*. 2014;90362J. doi:10.1117/12.2043662

12. Narayanan R, Kurhanewicz J, Shinohara K, Crawford ED, Simoneau A, Suri JS. Mri-ultrasound registration for targeted prostate biopsy. *Proc - 2009 IEEE Int Symp Biomed Imaging From Nano to Macro, ISBI 2009*. 2009:991-994. doi:10.1109/ISBI.2009.5193221

13. Hu Y, Ahmed HU, Taylor Z, et al. MR to ultrasound registration for image-guided prostate interventions. *Med Image Anal*. 2012;16(3):687-703. doi:10.1016/j.media.2010.11.003

14. Shaaer A, Davidson M, Semple M, et al. Clinical evaluation of an MRI-to-ultrasound deformable image registration algorithm for prostate brachytherapy. *Brachytherapy*. 2019;18(1):95-102. doi:10.1016/j.brachy.2018.08.006

15. Venderink W, de Rooij M, Sedelaar J. M, Huisman HJ, Fütterer JJ. Elastic Versus Rigid Image Registration in Magnetic Resonance Imaging–transrectal Ultrasound Fusion Prostate Biopsy: A Systematic Review and Meta-analysis. *Eur Urol Focus*. 2018;4(2):219-227. doi:10.1016/j.euf.2016.07.003

16. Borkowetz A, Platzek I, Toma M, et al. Comparison of systematic transrectal biopsy to transperineal magnetic resonance imaging/ultrasound-fusion biopsy for the diagnosis of prostate cancer. *BJU Int*. 2015;116(6):873-879.

doi:10.1111/BJU.13023

17. Siddiqui MM, Rais-Bahrami S, Turkbey B, et al. Comparison of MR/Ultrasound Fusion–Guided Biopsy With Ultrasound-Guided Biopsy for the Diagnosis of Prostate Cancer. *JAMA*. 2015;313(4):390-397. doi:10.1001/JAMA.2014.17942

18. Radtke JP, Kuru TH, Boxler S, et al. Comparative Analysis of Transperineal Template Saturation Prostate Biopsy Versus Magnetic Resonance Imaging Targeted Biopsy with Magnetic Resonance Imaging-Ultrasound Fusion Guidance. *J Urol*. 2015;193(1):87-94. doi:10.1016/J.JURO.2014.07.098

19. Hu Y, Ahmed HU, Taylor Z, et al. MR to ultrasound registration for image-guided prostate interventions. *Med Image Anal*. 2012;16(3):687-703. doi:10.1016/J.MEDIA.2010.11.003

20. Baco E, Ukimura O, Rud E, et al. Magnetic Resonance Imaging–Transectal Ultrasound Image-fusion Biopsies Accurately Characterize the Index Tumor: Correlation with Step-sectioned Radical Prostatectomy Specimens in 135 Patients. *Eur Urol*. 2015;67(4):787-794. doi:10.1016/J.EURURO.2014.08.077

21. Meng X, Rosenkrantz AB, Mendhiratta N, et al. Relationship Between Prebiopsy Multiparametric Magnetic Resonance Imaging (MRI), Biopsy Indication, and MRI-ultrasound Fusion–targeted Prostate Biopsy Outcomes. *Eur Urol*. 2016;69(3):512-517. doi:10.1016/J.EURURO.2015.06.005

22. Ukimura O, Marien A, Palmer S, et al. Trans-rectal ultrasound visibility of prostate lesions identified by magnetic resonance imaging increases accuracy of image-fusion targeted biopsies. *World J Urol*. 2015;33(11):1669-1676. doi:10.1007/s00345-015-1501-z

23. Guo H, Kruger M, Xu S, Wood BJ, Yan P. Deep adaptive registration of multi-modal prostate images. *Comput Med Imaging Graph*. 2020;84:101769. doi:10.1016/J.COMPMEDIMAG.2020.101769

24. Zeng Q, Fu Y, Tian Z, et al. Label-driven magnetic resonance imaging (MRI)-transrectal ultrasound (TRUS) registration using weakly supervised learning for MRI-guided prostate radiotherapy. *Phys Med Biol*. 2020;65(13):135002. doi:10.1088/1361-6560/AB8CD6

25. Chen Y, Xing L, Yu L, et al. MR to ultrasound image registration with segmentation-based learning for HDR prostate brachytherapy. *Med Phys*. 2021;48(6):3074-3083. doi:10.1002/MP.14901

26. Yang X, Fu Y, Lei Y, et al. Deformable MRI-TRUS Registration Using Biomechanically Constrained Deep Learning Model for Tumor-Targeted Prostate Brachytherapy. *Int J Radiat Oncol Biol Phys*. 2020;108(3):e339. doi:10.1016/J.IJROBP.2020.07.810

27. Yan P, Xu S, Rastinehad AR, Wood BJ. Adversarial image registration with application for MR and TRUS image fusion. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2018;11046:197-204. doi:10.1007/978-3-030-00919-9_23/FIGURES/3

28. Hu Y, Modat M, Gibson E, et al. Weakly-supervised convolutional neural networks for multimodal image registration. *Med Image Anal*. 2018;49:1-13. doi:10.1016/J.MEDIA.2018.07.002

29. Schroeder W, Martin K, Lorensen B. *The Visualization Toolkit*. 4th Ed. Kitware; 2006.

30. Bax J, Smith D, Bartha L, et al. A compact mechatronic system for 3D ultrasound guided prostate interventions. *Med Phys*. 2011;38(2):1055-1069. doi:10.1118/1.3531540

31. Fitzpatrick JM, West JB, Maurer CR. Predicting Error in Rigid-Body Point-Based Registration. *IEEE Trans Med Imaging.* 1998;17(5):694-702. doi:10.1109/42.736021

32. Orlando N, Gillies DJ, Gyacskov I, Romagnoli C, D'Souza D, Fenster A. Automatic prostate segmentation using deep learning on clinically diverse 3D transrectal ultrasound images. *Med Phys*. 2020;47(6):2413-2426. doi:10.1002/mp.14134

33. Orlando N, Gyacskov I, Gillies DJ, et al. Effect of dataset size, image quality, and image type on deep learning-based automatic prostate segmentation in 3D ultrasound. *Phys Med Biol*. 2022;67(7):074002. doi:10.1088/1361-6560/AC5A93

Chapter 6

6 Conclusions, Limitations, and Future Work

Minimally invasive procedures for prostate cancer diagnosis and treatment, including biopsy and brachytherapy, rely on medical imaging for accurate needle guidance, target definition, and treatment planning. Efficient utilization of these imaging modalities, specifically US and MRI, involves challenges such as time-consuming manual segmentation, poor needle tip visualization, and variable MR-US cognitive fusion. The work in this thesis seeks to address these challenges through the development and validation of generalizable and easy-to-implement software- and hardware-based approaches described in Chapters 2-5. This chapter provides a summary and conclusions for the described work including a discussion of limitations and potential future work.

6.1 Summary and Conclusions

While the incorporation of 3D TRUS imaging can improve needle guidance for minimally invasive procedures such as biopsy and brachytherapy, it also necessitates time-consuming and difficult manual prostate segmentation, which must be completed by the physician in the operating room environment while the patient is under sedation or anesthetic. **Chapter 2** described the development and validation of a deep learning-based automatic prostate segmentation algorithm for 3D TRUS images, potentially improving workflow efficiency and reducing procedure time. Our algorithm utilized a modified U-Net architecture[1] to segment 12 radially sampled 2D US images before reconstruction into a 3D surface. The clinically diverse training dataset consisted of 84 end-fire (biopsy) and 122 side-fire (BT) 3D TRUS images acquired using three different US systems, which was subsequently resliced to generate 6773 2D TRUS images for training a 2D neural network. Our proposed algorithm generated 3D surfaces with a median [first quartile (Q1), third quartile (Q3)] DSC, recall, precision, absolute VPD, MSD, and HD of 94.1 [92.6, 94.9] %, 96.0 [93.1, 98.5] %, 93.2 [88.8, 95.4] %, 5.78 [2.49, 11.5] %, 0.89 [0.73, 1.09] mm, and 2.89 [2.37,

4.35] mm, respectively, significantly improving performance compared to fully 3D CNNs including a V-Net, Dense V-Net, and High-resolution 3D-Net trained using the same dataset. Average total 3D segmentation time of 0.62 s with our proposed method demonstrated a significant reduction compared to manual segmentation, which can take up to ten minutes. Our proposed segmentation algorithm demonstrated fast and accurate 3D segmentation of the prostate across variable TRUS acquisition methods, providing a generalizable and robust algorithm that has the potential to reduce procedure time in minimally invasive prostate procedures.

Chapter 2 described a deep learning-based algorithm capable of accurate prostate segmentation, trained using a large, clinically diverse dataset with variable image quality. As large clinical datasets are rare, particularly for US imaging, widespread adoption of deep learning-based segmentation could be facilitated through demonstration of performance with smaller and less diverse datasets. **Chapter 3** examined the impact of the dataset on segmentation performance using our 2D deep learning plus 3D reconstruction approach, specifically the effect of dataset size, image quality, and image type. First, end-fire and side-fire datasets were generated by splitting the complete dataset of 6761 2D images, followed by the creation of smaller end-fire, side-fire, and mixed datasets with 1000, 500, 250, and 100 images. A modified U-Net and U-Net++ algorithm were implemented for this work, as the latter has shown improved performance with small datasets.[2] A 3D TRUS image quality grading was developed based on three factors rated from 1 to 5: acquisition quality, artifact severity, and boundary visibility. While the U-Net and U-Net++ showed no difference in performance when trained with the full dataset, the U-Net++ significantly outperformed the U-Net when trained using only side-fire or only end-fire images. For the U-Net++ trained using smaller datasets, performance plateaued at a training dataset of 1000 mixed images with a median DSC of 93.2 [89.7, 94.9] %. For our testing dataset, image quality showed no significant effect on segmentation performance for end-fire images but did show a significant effect for side-fire images, specifically boundary visibility and acquisition quality. The image quality grading scale specific to 3D TRUS imaging, the first of its kind to our knowledge, may help improve transparency of published results, easing comparison between algorithms trained on datasets of different image quality. The demonstrated performance of our algorithm when

trained with smaller datasets supports the potential for widespread use of deep learning segmentation methods, even when data is scarce.

Needle tip localization in prostate HDR-BT is typically done using standard B-mode TRUS imaging, which can be limited due to image artifacts, potentially resulting in deviation from the planned dose. **Chapter 4** described the development and validation of a PD US-based needle tip localization method designed specifically for HDR-BT. Needle perturbations generated using a custom-built wireless mechanical oscillator could be detected using PD US, presented as a coloured overlay on the B-mode US image, potentially improving needle visualization. The mechanical oscillator featured a cylindrical end piece designed to fit overtop the standard needle mandrins used for HDR-BT, improving Doppler signal generation. Our PD US needle localization method was evaluated in tissue equivalent agar phantoms as well as in 5 HDR-BT patients as part of a prospective feasibility clinical trial. In the phantom, B-mode alone, PD alone, and PD plus B-mode needle visualization methods demonstrated absolute mean ± standard deviation tip error of 0.3 ± 0.3 mm, 0.6 ± 0.5 mm, and 0.4 ± 0.2 mm for the mock HDR-BT needle implant; 0.8 ± 1.7 mm, 0.4 ± 0.6 mm, and 0.3 ± 0.5 mm for the explicit shadowing implant with plastic needles; and 0.5 ± 0.2 mm, 0.5 ± 0.3 mm, and 0.6 ± 0.2 mm for the explicit shadowing implant with metal needles, respectively. The averaged absolute tip error for all five patients in the prospective clinical trial was 0.9 ± 0.7 mm and 0.8 ± 0.5 mm for PD US and B-mode US needle visualization, respectively. PD US was demonstrated a particular advantage for needles classified as difficult to see using B-mode US, including providing the ability to visualize needles previously not visible using standard US alone. Our PD US method demonstrated equivalent needle tip error for unobstructed needles and reduced tip error and variation for obstructed needles, potentially increasing treatment accuracy in HDR-BT. The method is easy-to-implement and requires no modifications to the clinical workflow, while also being generalizable to other minimally invasive procedures, described in detail in Section 6.3.3.

Tumour-targeted BT has been proposed to overcome the limitation of conventional whole-gland BT, motivated by studies which have shown the most likely recurrence site following therapy is the DIL.[3–5] The DIL is not typically visible in US, necessitating the use of an additional imaging modality such as MRI, and thus challenging and variable

cognitive fusion of MR-US images is required to utilize the information intraoperatively. **Chapter 5** described the validation and clinical implementation of a surface-based deformable MR-3D TRUS image registration algorithm for prostate BT. Our DIR algorithm has two fully automated steps starting with an affine alignment of the prostate surfaces using the ICP algorithm followed by non-rigid warping using thin plate spline deformation.[6] Validation utilized a deformable prostate phantom and three retrospective HDR-BT clinical cases with comparison to manual rigid registration and physician cognitive fusion. When evaluated using phantom images at three deformation levels, the DIR algorithm significantly improved registration accuracy compared to rigid registration, demonstrating mean DSC and TRE of 0.74±0.08 and 0.94±0.49 mm compared to 0.64±0.16 and 1.88±1.24 mm for the rigid registrations. The mean CTV and NVB DSC compared to physician cognitive fusions, averaged for the three retrospective clinical cases, was 0.53 ± 0.15 for our DIR algorithm, demonstrating reduced variation compared to the inter-physician variability, which had a DSC of 0.50 ± 0.21. Prospective clinical implementation of our DIR algorithm allowed for a successful PSMA-PET-MRI GTV-targeted biopsy, confirmed histologically, and facilitated dose escalation to the registered CTV. Our DIR algorithm allows for the registration and translation of MR-defined target and organ-at-risk contours into the intraoperative US image, allowing for dose escalation and dose avoidance while standardizing the registration procedure between physicians, reducing operator dependence.

6.2 Limitations

6.2.1 General Limitations

A limitation common to the work described in all Chapters of this thesis is the difficulty obtaining gold standard comparisons for assessing the accuracy of our proposed methods. In Chapters 2-3, gold standard manual prostate segmentations were completed by a non-physician observer, limiting the clinical translation potential as physician contours can follow different guidelines and will vary physician-to-physician. In Chapters 4-5, retrospective and prospective clinical validation experiments were conducted, but there was no way to generate gold standard needle tip locations and MR-US registrations,

respectively. This necessitated the use of alternative evaluation metrics such as reference needle-based tip localization error used in Chapter 4 and comparisons to inter-physician variability as the standard of care in Chapter 5. Contrary to this, phantom experiments in Chapters 4-5 provided clear gold standards, specifically through CT imaging for the PD study and through US- and MR-visible landmarks for the registration study.

Another limitation common to each chapter is the difficulty in validating our proposed methods prospectively. Chapters 2-3 rely solely on retrospective data for training and evaluation. Prospective clinical implementation and evaluation of our deep learning-based automatic prostate segmentation algorithm would be critical for successful clinical translation. Among the challenges involved in implementing a deep learning algorithm in the clinic is the difficulty integrating with the clinical software. In particular, the vendor-provided computers typically do not contain a dedicated GPU with enough memory to complete the segmentations with the low computation time described in Chapters 2 and 3. In Chapters 4-5, prospective clinical validation was completed; however, it was limited to small feasibility studies with 5 and 1 clinical cases for the PD and registration studies, respectively. As described for Chapters 2-3, prospective studies with additional clinical cases are critical for demonstrating robustness towards the widespread clinical implementation of our proposed approaches.

6.2.2 Study-Specific Limitations

Chapter 2: Deep learning segmentation of the prostate in 3D TRUS images

The main limitations of the work described in Chapter 2 relate to the TRUS image dataset used in the study. While both side-fire and end-fire images were included, acquired using three different US systems, all images of a certain type were obtained from the same centre. This limited the generalizability and potential performance of our algorithm when implemented in a new centre where different US systems or acquisition geometries may be used. Importantly, all manual gold standard prostate segmentations were completed by one non-physician observer. While this offered consistency in the segmentation approach, there was variability in segmentation technique between physicians clinically, making exposure and robustness to multiple manual segmentation techniques critical for clinical translation. In addition, inter- and intra-observer segmentation variability was not examined for our

dataset, although this has been previously assessed for end-fire[7] and side-fire[8] 3D TRUS images, with our proposed algorithm demonstrating performance comparable to this variability. Finally, the dataset was limited to pre-implant 3D TRUS images with minimal obstructions, disregarding prostate segmentation in post-implant images which is critical to the HDR-BT workflow.

Beyond the dataset, there were also limitations to our algorithm design and network comparison implementation. Our algorithm utilized radial slices extracted with a 15° step angle about the approximate central axis, chosen to balance computation time and segmentation performance. While performance was excellent with this 15° step angle, no experiment with varying step angle was completed, so the optimal angle is not known. When comparing the performance of our proposed method to fully 3D CNNs trained using the same dataset, hyperparameter optimization was completed for the V-Net, with identical hyperparameters used for subsequent CNNs (Dense V-Net and High-resolution 3D-Net). This meant that the performance observed for these two CNNs may not be optimal, as this would require individual optimization for each architecture. While radial segmentation approaches have previously demonstrated improved segmentation performance in the base and apex of the prostate,[7] our analysis was limited to whole-gland segmentation performance. In addition, there was no analysis of the dosimetric impact of segmentation error.

Chapter 3: Effect of dataset size, image quality, and image type on deep learning prostate segmentation in 3D TRUS

As the same TRUS dataset was used for the work described in Chapter 2 and Chapter 3, the same dataset limitations described in the previous section apply to this chapter as well. Specific to Chapter 3, only one physician observer developed the 3D TRUS image quality scale and graded the testing dataset. Thus, inter and intra-observer variability in image quality classification was not assessed. Image quality was only graded for the testing dataset of 40 3D TRUS images, with no grading completed for the larger training dataset, which may play a critical role in algorithm performance. While we expect the randomly sampled testing dataset to be representative of the total dataset, explicit grading of the training dataset would demonstrate any potential biases. Furthermore, our 3D TRUS

dataset was shown to be of high quality on average. To improve algorithm generalizability and further explore the impact of image quality on segmentation performance, an expanded dataset with a wider range of image qualities, including low quality images, could be used, ideally from a different centre. While our dataset included TRUS images from multiple vendors, no experiment was conducted to examine the impact of US machine vendor on segmentation performance. Such a study would offer a strong assessment of generalizability, which is critical for widespread clinical translation. Our U-Net++ implementation was limited to only a ResNet CNN backbone. Higher performance may be achieved through utilization of an ensemble network with results from multiple CNN backbones combined into one segmentation result using averaging, majority vote, or the STAPLE method for example.[9]

Chapter 4: Power Doppler needle localization for interstitial prostate brachytherapy

The main limitation of the work described in Chapter 4 is the evaluation metric used. In both the phantom and clinical validation, needle tip localization error was computed using a reference needle calculation, which relied on accurate identification of the reference needle tip. Although tip error is mitigated by choosing easy to see needles as a reference, any error contributes to the resulting tip localization error. In addition, the reference needle calculation utilizes end-length measurements, which assumes there is no bending or deflection of the needle. While this is a fair assumption in the homogenous phantom, needle deflection can occur in clinical cases. End-length measurements were also made by hand using a ruler, limiting accuracy to within 0.5 mm, which may further impact the computed tip localization error. In the phantom experiments, only one observer identified the needle tip and landmark locations in US and CT, so no inter-observer variability was computed. Live US imaging was used for needle tip localization with both B-mode and PD US, causing needle tips to be identified once, preventing any examination of intra-observer variability. Specific to the clinical validation, no alternative imaging such as CT or MRI was used, so there was no proper ground truth to compare to. Another limitation is the time constraints imposed by the small window in which we completed our study protocol during the HDR-BT clinical workflow, as we could not extend the procedure time. This resulted in rushed needle tip localization, which occurred often for the PD US method localization

as this was completed after the B-mode localizations were done. Finally, both phantom and clinical validation experiments were limited by their sample size, with only 12-15 needles per case and a total of 3 phantom cases and 5 clinical cases. In addition, needle visibility was graded highly on average, with only 11 fully obstructed needles across 5 patients. Further validation in additional phantom and clinical cases with more variable needle visibility would allow for a more robust assessment of needle tip localization error.

Chapter 5: Surface-based deformable MRI-3D TRUS image registration algorithm

The main limitation of the surface-based DIR algorithm described in Chapter 5 is its reliance on accurate and consistent manual prostate contouring. Differences in manual prostate segmentation technique between MR and US images as well as between and within observers[7,8] would directly impact the registration result and subsequent target and organ-at-risk contour translation. Implementation of defined prostate contouring guidelines for both MR and US imaging would help to standardize the contouring process between different physicians and further reduce operator dependence. The datasets used for phantom validation and retrospective clinical validation are both small, with 6 and 3 cases respectively. In addition, clinical implementation was limited to only one prospective HDR-BT case, including one pre-implant registration and one post-implant registration. No MR images utilized in this work were acquired with an endorectal coil, which would result in more deformation than a standard pelvic or body coil. Validation in more retrospective and prospective cases, including for MR images acquired with endorectal coils would strengthen the clinical translation potential. Registration accuracy assessment for the prospective clinical validation case was limited to histological evaluation of the targeted biopsy result. Quantitative evaluation of our DIR algorithm performance in prospective patients in terms of DSC or TRE is another critical step towards widespread implementation.

6.3 Future Work

6.3.1 Multi-Institution and Multi-Physician Validation of Deep Learning Prostate Segmentation Towards Clinical Implementation

A 2D deep learning plus 3D reconstruction automatic prostate segmentation algorithm for 3D US images was described in Chapter 2, evaluated using retrospective images acquired in clinical prostate biopsy and HDR-BT procedures. Future work for this project will be focused on successful clinical implementation through further clinical validation and dataset improvements. The dataset used in Chapter 2 and 3 was limited to 3D TRUS images acquired using three different US systems, all from one centre. Successful clinical implementation of this algorithm, particularly at a different clinical site, is highly dependent on the generalizability and robustness of the trained network. A multi-centre dataset of 3D TRUS images acquired using numerous US systems from different minimally invasive procedures with varied image quality would result in a more generalizable segmentation algorithm, improving performance in prospective clinical scenarios. Manual prostate segmentation also varies between physicians. While our current dataset is limited to one non-physician observer who provided all ground truth segmentations, the inclusion of manual segmentations produced by different observers would strengthen the training dataset and thus algorithm robustness.

Furthermore, in Chapter 2 and 3, algorithm performance evaluation was limited to segmentation accuracy metrics including DSC, boundary distance metrics, and VPD. While the performance of our algorithm was strong according to these metrics, the clinical utility of the segmentation output is related to the physician's confidence in the result. The prostate contour produced by the algorithm would require manual editing by the physician, so a study examining how much editing time is required on average would allow for determination of a clinically relevant total segmentation time. Such a metric may better reflect the clinical utility of the algorithm. Using multiple physician observers, both algorithmic and manual physician-defined prostate contours could be provided blindly to the observer, with the required editing time documented. This would allow for determination of the average editing time for algorithm-generated contours in addition to

an evaluation of how this time compares to the editing time for manual contours, effectively comparing to the current standard-of-care. This multi-center, multi-physician study would serve to further validate and improve our algorithm's robustness and generalizability as a critical step towards clinical implementation.

Finally, widespread clinical implementation of deep learning methods in medical imaging is limited in part by the restricted nature of the proposed models and datasets. Future work could focus on improving public access to deep learning models and the datasets used to train and evaluate them. This would increase transparency in the presented results and would greatly increase access to deep learning methods, accelerating clinical translation. There are significant challenges in widespread data sharing including the protection of patient information, restrictions related to research ethics agreements, and the high cost associated with dataset curation and management.

6.3.2 Automatic Prostate Segmentation in Post-Implant 3D TRUS Images

The training dataset used in Chapter 2 and 3 was limited to end-fire and side-fire 3D TRUS acquisition. As described in Section 1.4.2.2, axially reconstructed step-back 3D TRUS images are commonly used in prostate BT.[10–12] Compared to end-fire and side-fire 3D TRUS acquisitions which utilize transducer rotation, step-back 3D TRUS utilizes transducer translation in inferior steps, resulting in decreased image resolution in the sagittal plane. Evaluation and supplemental training of our segmentation algorithm for use with step-back 3D TRUS images would allow for implementation in any centre where 3D TRUS is used. In addition, the described dataset is limited to pre-insertion 3D TRUS images, which have minimal obstructions, presenting a simpler segmentation task. In HDR-BT procedures, prostate segmentation in post-implant 3D TRUS images is critical for treatment planning. Segmentation of the prostate in post-implant images presents a significant challenge due to the hyperechoic artifacts created by the implanted needles obscuring the prostate boundary. Supplemental training and evaluation of our algorithm for post-implant 3D TRUS images is a necessary step towards clinical implementation for HDR-BT. Successful validation in post-implant 3D TRUS would demonstrate a versatile algorithm that could aid in both pre-implant segmentation for needle guidance and post-

implant segmentation for treatment planning, greatly reducing procedure time and standardizing the segmentation procedure.

6.3.3 Power Doppler Needle Localization Extended to other Minimally Invasive Procedures

Chapter 4 described the development and validation of a PD-based needle tip localization method for interstitial prostate BT. This method was validated in both phantom and prospective clinical cases, focusing on simulated and clinical HDR-BT procedures, respectively. As our PD-based technique demonstrated improved needle tip visibility and reduced variability compared to standard B-mode US for HDR-BT, future work could expand our approach to other minimally invasive procedures where needle tip localization is equally critical. Candidate procedures where PD US may offer potential utility include LDR-BT, gynecological BT, focal ablation procedures, and hyperthermia procedures. As our method requires no additions to the clinical equipment or workflow, translation to new minimally invasive procedures would be smooth, only requiring modification of the end piece design based on the needle-like tools shape. Each procedure does present new challenges, however. Focal ablation and hyperthermia procedures commonly require needle insertion at oblique angles, resulting in needle tip visibility challenges using standard US,[13] and thus a natural application for PD US. In gynecological BT procedures, needle visibility can be limited by the presence of a vaginal cylinder.[14] Validation of PD US for needle tip identification in other minimally invasive procedures could greatly increase the clinical translation and potential impact of our method.

As shown in Fig. 4.9, PD US combined with 3D US image acquisition could allow for the simultaneous identification of all implanted needle tip locations within the 3D US image. This is particularly relevant for interstitial HDR prostate and gynecological BT where multiple needles are inserted, including potential use as an initialization for automatic multi-needle segmentation algorithms.[14,15]

6.3.4 Deep Learning-Based Automatic MRI-3D US Image Registration Algorithm

Chapter 5 described the validation of a deformable surface-based MRI-3D TRUS registration algorithm for tumour-targeted prostate BT. Our proposed surface-based algorithm is limited by its dependence on manual prostate segmentation, which has been shown to vary within and between observers.[7,8] Implementation of deep learning prostate segmentation as input to the DIR algorithm could improve segmentation consistency and accuracy compared to standard manual segmentation as well as further reducing procedure time. 3D TRUS segmentation could utilize the trained networks described in Chapter 2 and 3, requiring no modification.[16,17] Prostate segmentation in MR images could utilize the same algorithm trained with MR data, or a pre-trained MR-specific algorithm could be implemented.[18,19] The underlying registration algorithm could remain unchanged, making this a natural next step with a large potential impact for reducing operator dependence and procedure time.

To further improve registration accuracy, the deep learning segmentation initialization could be augmented with a deep learning-based registration procedure.[20–25] A full deep learning registration procedure would offer significant advantages in terms of computation time compared to our conventional approach which took approximately two minutes on average. There are challenges in implementing a full deep learning registration algorithm, most importantly the difficulty in creating a large dataset of paired 3D TRUS and MR images with high quality prostate contours.

Image-guided minimally invasive procedures are a powerful tool for prostate cancer management, while also presenting numerous challenges. This thesis described several easy-to-implement, generalizable software- and hardware-based approaches designed specifically for prostate biopsy and brachytherapy. These tools offer the potential to reduce procedure time and variability, potentially increasing the effectiveness of minimally invasive procedures towards improved prostate cancer management for men worldwide.

6.4 References

1. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2015;9351:234-241. doi:10.1007/978-3-319-24574-4_28

2. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2018;11045:3-11. http://arxiv.org/abs/1807.10165.

3. Pucar D, Hricak H, Shukla-Dave A, et al. Clinically Significant Prostate Cancer Local Recurrence After Radiation Therapy Occurs at the Site of Primary Tumor: Magnetic Resonance Imaging and Step-Section Pathology Evidence. *Int J Radiat Oncol Biol Phys*. 2007;69(1):62-69. doi:10.1016/j.ijrobp.2007.03.065

4. Arrayeh E, Westphalen AC, Kurhanewicz J, et al. Does local recurrence of prostate cancer after radiation therapy occur at the site of primary tumor? Results of a longitudinal MRI and MRSI study. *Int J Radiat Oncol Biol Phys*. 2012;82(5):e787-793. doi:10.1016/j.ijrobp.2011.11.030

5. Cellini N, Morganti AG, Mattiucci GC, et al. Analysis of intraprostatic failures in patients treated with hormonal therapy and radiotherapy: implications for conformal therapy planning. *Int J Radiat Oncol Biol Phys*. 2002;53(3):595-599. doi:10.1016/S0360-3016(02)02795-5

6. Schroeder W, Martin K, Lorensen B. *The Visualization Toolkit*. 4th Ed. Kitware; 2006.

7. Qiu W, Yuan J, Ukwatta E, Fenster A. Rotationally resliced 3D prostate TRUS segmentation using convex optimization with shape priors. *Med Phys*. 2015;42(2):877-891. doi:10.1118/1.4906129

8. Tong S, Cardinal HN, McLoughlin RF, Downey DB, Fenster A. Intra- and inter-observer variability and reliability of prostate volume measurement via two-dimensional and three-dimensional ultrasound imaging. *Ultrasound Med Biol*. 1998;24(5):673-681. doi:10.1016/S0301-5629(98)00039-8

9. Warfield SK, Zou KH, Wells WM. Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation. *IEEE Trans Med Imaging*. 2004;23(7):903-921. doi:10.1109/TMI.2004.828354

10. Siebert F-A, Hirt M, Niehoff P, Kovács G. Imaging of implant needles for real-time HDR-brachytherapy prostate treatment using biplane ultrasound transducers. *Med Phys*. 2009;36(8):3406-3412. doi:10.1118/1.3157107

11. Schmid M, Crook JM, Batchelar D, et al. A phantom study to assess accuracy of needle identification in real-time planning of ultrasound-guided high-dose-rate prostate implants. *Brachytherapy*. 2013;12:56-64. doi:10.1016/j.brachy.2012.03.002

12. Batchelar D, Gaztañaga M, Schmid M, Araujo C, Bachand F, Crook J. Validation study of ultrasound-based high-dose-rate prostate brachytherapy planning compared with CT-based planning. *Brachytherapy*. 2014;13(1):75-79. doi:10.1016/J.BRACHY.2013.08.004

13. Gillies DJ, Rodgers JR, Gyacskov I, et al. Deep Learning Segmentation of General Interventional Tools in Two-dimensional Ultrasound Images. *Med Phys*. 2020;47(10):4956-4970. doi:10.1002/mp.14427

14. Rodgers JR, Hrinivich WT, Surry K, Velker V, D'Souza D, Fenster A. A semiautomatic segmentation method for interstitial needles in intraoperative 3D transvaginal ultrasound images for high-dose-rate gynecologic brachytherapy of vaginal tumors. *Brachytherapy*. 2020;19(5):659-668. doi:10.1016/j.brachy.2020.05.006

15. Hrinivich WT, Hoover DA, Surry K, et al. Simultaneous automatic segmentation of multiple needles using 3D ultrasound for high-dose-rate prostate brachytherapy. *Med Phys.* 2017;44(4):1234-1245. doi: 10.1002/mp.12148.

16. Orlando N, Gillies DJ, Gyacskov I, Romagnoli C, D'Souza D, Fenster A. Automatic prostate segmentation using deep learning on clinically diverse 3D transrectal ultrasound images. *Med Phys*. 2020;47(6):2413-2426. doi:10.1002/mp.14134

17. Orlando N, Gyacskov I, Gillies DJ, et al. Effect of dataset size, image quality, and image type on deep learning-based automatic prostate segmentation in 3D ultrasound. *Phys Med Biol*. 2022;67(7):074002. doi:10.1088/1361-6560/AC5A93

18. Ghavami N, Hu Y, Gibson E, et al. Automatic segmentation of prostate MRI using convolutional neural networks: Investigating the impact of network architecture on the accuracy of volume measurement and MRI-ultrasound registration. *Med Image Anal*. 2019;58:101558. doi:10.1016/j.media.2019.101558

19. Litjens G, Toth R, van de Ven W, et al. Evaluation of prostate segmentation algorithms for MRI: the PROMISE12 challenge. *Med Image Anal*. 2014;18(2):359-373. doi:10.1016/j.media.2013.12.002

20. Guo H, Kruger M, Xu S, Wood BJ, Yan P. Deep adaptive registration of multi-modal prostate images. *Comput Med Imaging Graph*. 2020;84:101769. doi:10.1016/J.COMPMEDIMAG.2020.101769

21. Zeng Q, Fu Y, Tian Z, et al. Label-driven magnetic resonance imaging (MRI)-transrectal ultrasound (TRUS) registration using weakly supervised learning for MRI-guided prostate radiotherapy. *Phys Med Biol*. 2020;65(13):135002.

doi:10.1088/1361-6560/AB8CD6

22. Chen Y, Xing L, Yu L, et al. MR to ultrasound image registration with segmentation-based learning for HDR prostate brachytherapy. *Med Phys*. 2021;48(6):3074-3083. doi:10.1002/MP.14901

23. Yang X, Fu Y, Lei Y, et al. Deformable MRI-TRUS Registration Using Biomechanically Constrained Deep Learning Model for Tumor-Targeted Prostate Brachytherapy. *Int J Radiat Oncol Biol Phys*. 2020;108(3):e339. doi:10.1016/J.IJROBP.2020.07.810

24. Yan P, Xu S, Rastinehad AR, Wood BJ. Adversarial image registration with application for MR and TRUS image fusion. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2018;11046:197-204. doi:10.1007/978-3-030-00919-9_23/FIGURES/3

25. Hu Y, Modat M, Gibson E, et al. Weakly-supervised convolutional neural networks for multimodal image registration. *Med Image Anal*. 2018;49:1-13. doi:10.1016/J.MEDIA.2018.07.002

APPENDIX A – Copyright Releases

A.1 Permission to reproduce previously published material in Chapter 2

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS

Jun 13, 2022

This Agreement between Mr. Nathan Orlando ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	5327211062693
License date	Jun 13, 2022
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Medical Physics
Licensed Content Title	Automatic prostate segmentation using deep learning on clinically diverse 3D transrectal ultrasound images
Licensed Content Author	Aaron Fenster, David D'Souza, Cesare Romagnoli, et al
Licensed Content Date	Apr 8, 2020
Licensed Content Volume	47
Licensed Content Issue	6

Licensed Content Pages	14
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Title	Software and Hardware-based Tools for Improving Ultrasound Guided Prostate Brachytherapy
Institution name	Western University
Expected presentation date	Jul 2022
Requestor Location	Mr. Nathan Orlando 1151 Richmond Street North Robarts Research Institute Room 5250 London, ON N6A 5B7 Canada Attn: Mr. Nathan Orlando
Publisher Tax ID	EU826007151
Total	0.00 CAD

Terms and Conditions

TERMS AND CONDITIONS

- The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto

- NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.

- WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

- You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

- IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

- Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

- The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

- This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

- Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.

- These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

- In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

- WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

- This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

- This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

WILEY OPEN ACCESS TERMS AND CONDITIONS

Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of Creative Commons Licenses. The license type is clearly identified on the article.

The Creative Commons Attribution License

The Creative Commons Attribution License (CC-BY) allows users to copy, distribute and transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-

Creative Commons Attribution Non-Commercial License

The Creative Commons Attribution Non-Commercial (CC-BY-NC)License permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.(see below)

Creative Commons Attribution-Non-Commercial-NoDerivs License

The Creative Commons Attribution Non-Commercial-NoDerivs License (CC-BY-NC-ND) permits use, distribution and reproduction in any medium, provided the original work is properly cited, is not used for commercial purposes and no modifications or adaptations are made. (see below)

Use by commercial "for-profit" organizations

Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee.

Further details can be found on Wiley Online Library
http://olabout.wiley.com/WileyCDA/Section/id-410895.html

Other Terms and Conditions:

v1.10 Last updated September 2015

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

A.2 Permission to reproduce previously published material in Chapter 3

From: **Permissions** permissions@ioppublishing.org
Subject: Re: PMB Article re-use in Thesis Dissertation
Date: June 15, 2022 at 12:57 AM
To: Nathan Orlando norland2@uwo.ca

Dear Nathan Orlando

Thank you for your request to reproduce content from Physics in Medicine & Biology.

The content you have requested was published under a CC BY licence which permits reuse for any purposes, even commercial, provided the licence terms are adhered to.

Therefore you may reuse the content without permission, so long as you reference it adequately and adhere to the terms of the CC BY licence.

Please note this does not apply to any content/figure which is credited to another source in our publication or has been obtained from a third party, which is not available under a suitable open access licence. Express permission for such content/figures must be obtained from the copyright owner.

Kind regards,

Sophie
Copyright & Permissions Team
Sophie Brittain - Rights & Permissions Assistant
Cameron Wood - Legal & Rights Adviser
Contact Details
E-mail: permissions@ioppublishing.org

For further information about copyright and how to request permission:
https://publishingsupport.iopscience.iop.org/copyright-journals/
See also: https://publishingsupport.iopscience.iop.org/
Please see our Author Rights Policy https://publishingsupport.iopscience.iop.org/author-rights-policies/
Please note: We do not provide signed permission forms as a separate attachment. Please print this email and provide it to your publisher as proof of permission. **Please note**: Any statements made by IOP Publishing to the effect that authors do not need to get permission to use any content where IOP Publishing is not the publisher is not intended to constitute any sort of legal advice. Authors must make their own decisions as to the suitability of the content they are using and whether they require permission for it to be published within their article.

APPENDIX B – Research Ethics and Health Canada Approvals

David Tessier

From: Barras, Eliane (HC/SC) <eliane.barras@canada.ca>
Sent: March 12, 2019 9:44 AM
To: David Tessier
Subject: RE: Classification - Non-Coupling Needle Oscillator
Attachments: RE: Classification Request New ITA device; RE: Classification Request New ITA device; ITA-Med-Dev-Fenster-Robarts-preliminary-v5.docx; RE: Classification Request New ITA device

Hello David,

Based on the information provided, the NCNO-1331T is a Class I medical device per Rule 12 of the *Medical Devices Regulations* (MDR). For your information, Rule 12 states:
Rule 12:
Any other active device is classified as Class I.

The software used to acquire and store 2D ultrasound images and reconstruct the 3D ultrasound image is a Class II medical device per Rule 10(1) of the MDR. For your information, Rule 10(1) states:
Rule 10:

1. ***Subject to subrule (2), an active diagnostic device, including any dedicated software, that supplies energy for the purpose of imaging or monitoring physiological processes is classified as Class II.***
2. *A device described in subrule (1) that is intended to be used to monitor, assess or diagnose a disease, a disorder, an abnormal physical state or a pregnancy, if erroneous readings could result in immediate danger, is classified as Class III.*

If the two devices will be used/sold together as a system, per section 7 of the MDR the highest risk classification applies, and the system would be considered a Class II device.
7 If a medical device can be classified into more than one class, the class representing the higher risk applies.

Please note that this classification decision is based on the information provided at this time. Any changes to the product's representation for use and/or labelling may affect the classification decision.

Please feel free to contact me if you have any questions or need more clarification.

Regards,

Eliane (Ellie) Barras, MSc

Agente principale des affaires réglementaires, Division des services d'homologation
Bureau des matériels médicaux, Santé Canada
eliane.barras@canada.ca | tél: (613) 960-5296

Senior Regulatory Affairs Officer, Device Licensing Services Division
Medical Devices Bureau, Health Canada
eliane.barras@canada.ca | tel: (613) 960-5296

Western Research

Date: 13 September 2019

To: Dr Douglas Hoover

Project ID: 113708

Study Title: Needle Localization during Prostate Brachytherapy using Power Doppler Ultrasound

Application Type: HSREB Initial Application

Review Type: Full Board

Date Approval Issued: 13/Sep/2019

REB Approval Expiry Date: 13/Sep/2020

Dear Dr Douglas Hoover

The Western University Health Science Research Ethics Board (HSREB) has reviewed and approved the above mentioned study as described in the WREM application form, as of the HSREB Initial Approval Date noted above. This research study is to be conducted by the investigator noted above. All other required institutional approvals must also be obtained prior to the conduct of the study.

Documents Approved:

Document Name	Document Type	Document Date	Document Version
Data_collection_form_Feb-27-2019-v1.0	Other Data Collection Instruments	27/Feb/2019	1.0
LOI-Aug-29-2019-v1.2-Clean	Written Consent/Assent	06/Sep/2019	1.2
Prost_Need_Loc_proto FINAL	Protocol	26/Feb/2019	1.0

Documents Acknowledged:

Document Name	Document Type	Document Date	Document Version
Abdollah_Int-J-Urol-2012	References		
Crook-Brachy-2015	References		
Moreira_2018_Phys._Med._Biol._63_20NT02	References		
PowerDoppler-ProceedingsOfSPIE	References	30/Aug/2019	1
TIONG 2010 A small tolerance for catheter displacement in high-dose rate prostate brachytherapy is necessary and feasible	References	13/May/2019	1

No deviations from, or changes to, the protocol or WREM application should be initiated without prior written approval of an appropriate amendment from Western HSREB , except when necessary to eliminate immediate hazard(s) to study participants or when the change(s) involves only administrative or logistical aspects of the trial.

REB members involved in the research project do not participate in the review, discussion or decision.

The Western University HSREB operates in compliance with, and is constituted in accordance with, the requirements of the TriCouncil Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2); the International Conference on Harmonisation Good Clinical Practice Consolidated Guideline (ICH GCP); Part C, Division 5 of the Food and Drug Regulations; Part 4 of the Natural Health Products Regulations; Part 3 of the Medical Devices Regulations and the provisions of the Ontario Personal Health Information Protection Act (PHIPA 2004) and its applicable regulations. The HSREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000940.

Please do not hesitate to contact us if you have any questions.

Sincerely,

Patricia Sargeant, Ethics Officer (ext. 85990) on behalf of Dr. Joseph Gilbert, HSREB Chair

Note: This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).

Western Research

Date: 16 December 2019

To: Dr. Aaron Fenster

Project ID: 114910

Study Title: Using Deep Learning to Automatically Identify Internal Features

Application Type: HSREB Initial Application

Review Type: Delegated

Meeting Date / Full Board Reporting Date: 14/Jan/2020

Date Approval Issued: 16/Dec/2019

REB Approval Expiry Date: 16/Dec/2020

Dear Dr. Aaron Fenster

The Western University Health Science Research Ethics Board (HSREB) has reviewed and approved the above mentioned study as described in the WREM application form, as of the HSREB Initial Approval Date noted above. This research study is to be conducted by the investigator noted above. All other required institutional approvals must also be obtained prior to the conduct of the study.

Documents Approved:

Document Name	Document Type	Document Date	Document Version
DCF-prostate_DL-DB-Oct_23_2019-v1.0	Other Data Collection Instruments	23/Oct/2019	1.0
Protocol-23-Oct-2019-1.0	Protocol	23/Oct/2019	1.0

No deviations from, or changes to, the protocol or WREM application should be initiated without prior written approval of an appropriate amendment from Western HSREB , except when necessary to eliminate immediate hazard(s) to study participants or when the change(s) involves only administrative or logistical aspects of the trial.

REB members involved in the research project do not participate in the review, discussion or decision.

The Western University HSREB operates in compliance with, and is constituted in accordance with, the requirements of the TriCouncil Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2); the International Conference on Harmonisation Good Clinical Practice Consolidated Guideline (ICH GCP); Part C, Division 5 of the Food and Drug Regulations; Part 4 of the Natural Health Products Regulations; Part 3 of the Medical Devices Regulations and the provisions of the Ontario Personal Health Information Protection Act (PHIPA 2004) and its applicable regulations. The HSREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000940.

Please do not hesitate to contact us if you have any questions.

Sincerely,

Daniel Wyzynski, Research Ethics Coordinator, on behalf of Dr. Philip Jones, HSREB Vice-Chair

Note: This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).

Western Research

Date: 6 December 2021

To: Douglas Hoover

Project ID: 115863

Study Title: Development of machine learning tools for prostate brachytherapy

Reference Number/ID: 2

Application Type: HSREB Amendment Form

Review Type: Delegated

Full Board Reporting Date: 07/December/2021

Date Approval Issued: 06/Dec/2021

REB Approval Expiry Date: 07/Aug/2022

Dear Douglas Hoover ,

The Western University Health Sciences Research Ethics Board (HSREB) has reviewed and approved the WREM application form for the amendment, as of the date noted above.

Documents Approved:

Document Name	Document Type	Document Date	Document Version
Protocol-27-Nov-2021-v2.0-Amendment1-CLEAN	Protocol	27/Nov/2021	2.0

REB members involved in the research project do not participate in the review, discussion or decision.

The Western University HSREB operates in compliance with, and is constituted in accordance with, the requirements of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2); the International Conference on Harmonisation Good Clinical Practice Consolidated Guideline (ICH GCP); Part C, Division 5 of the Food and Drug Regulations; Part 4 of the Natural Health Products Regulations; Part 3 of the Medical Devices Regulations and the provisions of the Ontario Personal Health Information Protection Act (PHIPA 2004) and its applicable regulations. The HSREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000940.

Please do not hesitate to contact us if you have any questions.

Sincerely,

Ms. Jhananiee Subendran, Ethics Coordinator on behalf of Dr. Philip Jones, HSREB Chair

Note: This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).

Western Research

Date: 11 February 2022

To: Dr. Glenn Bauman

Project ID: 120132

Study Title: Radioligand fOr locAl raDiorecurrent proStaTe cancER

Application Type: HSREB Initial Application

Review Type: Full Board

Date Approval Issued: 11/Feb/2022

REB Approval Expiry Date: 11/Feb/2023

Dear Dr. Glenn Bauman

The Western University Health Science Research Ethics Board (HSREB) has reviewed and approved the above mentioned study as described in the WREM application form, as of the HSREB Initial Approval Date noted above. This research study is to be conducted by the investigator noted above. **All other required institutional approvals and mandated training must also be obtained prior to the conduct of the study**.

Documents Approved:

Document Name	Document Type	Document Date	Document Version
211115!_ROADSTER_Female Partner Pregnancy Letter of Information and Consent	Preliminary and Pregnant Partner Info Letter/ICF	15/Nov/2021	
211115!_ROADSTER_Female Partner Preliminary Letter of Information	Preliminary and Pregnant Partner Info Letter/ICF	15/Nov/2021	
211014!_Protocol_ROADSTER	Protocol	14/Oct/2021	
220121!_List of variables_V2	Other Data Collection Instruments	21/Jan/2022	2
220210!_EPIC Questionnaires_V1	Paper Survey	10/Feb/2022	1
220210!_ROADSTER Main Consent Form_V5	Written Consent/Assent	10/Feb/2022	5

Documents Acknowledged:

Document Name	Document Type	Document Date	Document Version
211029!_Study Schema_ROADSTER	Flow Diagram	29/Oct/2021	1
177Lu-PNT2002 IB V1.1-FINAL-21NOV2020_Final	Investigator Brochure	21/Nov/2020	1.1
211007!_ROADSTER Budget_RS	Study budget	07/Oct/2021	1
211112!_ROADSTER trial_nol	NOL/NOA/ITA	12/Nov/2021	1
CPDC_[18F]PSMA-1007 IB v1.0_13Dec2019_FINAL	Investigator Brochure	13/Dec/2019	1
211009!_List of references_V1	References	21/Oct/2021	1
HRSSRC - Bauman - Roadster - Letter of No Objection - REV	Radionuclide Safety and Scientific Review Approval	10/Dec/2021	1
220125!_Data and Safety Monitoring Committee Charter_V2	DSMB/C Charter	25/Jan/2022	2

No deviations from, or changes to, the protocol or WREM application should be initiated without prior written approval of an appropriate amendment from Western HSREB , except when necessary to eliminate immediate hazard(s) to study participants or when the change(s) involves only administrative or logistical aspects of the

trial.

REB members involved in the research project do not participate in the review, discussion or decision.

The Western University HSREB operates in compliance with, and is constituted in accordance with, the requirements of the TriCouncil Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2); the International Conference on Harmonisation Good Clinical Practice Consolidated Guideline (ICH GCP); Part C, Division 5 of the Food and Drug Regulations; Part 4 of the Natural Health Products Regulations; Part 3 of the Medical Devices Regulations and the provisions of the Ontario Personal Health Information Protection Act (PHIPA 2004) and its applicable regulations. The HSREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000940.

Please do not hesitate to contact us if you have any questions. Sincerely,

Patricia Sargeant, Ethics Officer (psargean@uwo.ca) on behalf of Dr. Philip Jones, HSREB Chair

Note: This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).

Curriculum Vitae

Nathan James Orlando, MSc

Education

2017 – 2022	**Doctor of Philosophy in Medical Biophysics** Department of Medical Biophysics Western University, London, ON, Canada Supervisors: Aaron Fenster, Ph.D., FCCPM and Douglas Hoover, Ph.D., FCCPM Thesis: Software and Hardware-based Tools for Improving Ultrasound Guided Prostate Brachytherapy
2017 – 2019	**Master of Science in Clinical Medical Biophysics** CAMPEP accredited Department of Medical Biophysics Western University, London, ON, Canada
2013 – 2017	**Bachelor of Science with Honors: Physics with First Class Honors** Department of Physics University of Alberta, Edmonton, AB, Canada Supervisor: Nicola De Zanche, Ph.D., MCCPM Thesis: SAR Dosimeter for MRI

Publications

Peer-Reviewed Journal Manuscripts (5 First Author, 8 Total)

I. Published (3 First Author, 6 Total)

1. Du Toit C and **Orlando N***, Papernick S, Dima R, Gyacskov I, Fenster A. *Automatic femoral articular cartilage segmentation using deep learning in three-dimensional ultrasound images of the knee*. Osteoarthritis and Cartilage Open **4**(3), 100290 (2022). *Co-first Authors.

2. Park C, Xing S, Papernick S, **Orlando N**, Knull E, Du Toit C, Bax J, Gardi L, Barker K, Tessier D, Fenster A. *Spatially tracked whole-breast 3D ultrasound system toward point-of-care breast cancer screening in high-risk women with dense breasts*. Medical Physics **49**(6), 3944-3962 (2022).

3. **Orlando N**, Gyacskov I, Gillies DJ, Guo F, Romagnoli C, D'Souza D, Cool DW, Hoover DA, Fenster A. *Effect of dataset size, image quality, and image type on deep learning-based automatic prostate segmentation in 3D ultrasound*. Physics in Medicine and Biology **67**(7), 074002 (2022).

4. Zhang C, Hilts M, Batchelar D, **Orlando N**, Gardi L, Fenster A, Crook J. *Characterization and registration of 3D ultrasound for use in permanent breast seed implant brachytherapy treatment planning.* Brachytherapy **20**(1), 248-256 (2021).

5. **Orlando N** and Gillies DJ, Gyacskov I, Romagnoli C, D'Souza D, Fenster A. *Automatic prostate segmentation using deep learning on clinically diverse 3D transrectal ultrasound images*. Medical Physics **47**(6), 2413-2426 (2020). **(Selected as Editor's Choice.)**
6. Morrish W, West P, **Orlando N**, Klantsataya E, Gardner K, Lane S, DeCorby R, Francois A, Meldrum A. *Refractometric micro-sensor using a mirrored capillary resonator*. Optics Express **24**, 24959-24970 (2016).

II. Under Review (2 First Author, 2 Total)

1. **Orlando N**, Edirisinghe C, Gyacskov I, Vickress J, Sachdeva R, Gomez JA, D'Souza D, Velker V, Mendez LC, Bauman G, Fenster A, Hoover DA. *Validation of a surface-based deformable MRI-3D ultrasound image registration algorithm towards clinical implementation for interstitial prostate brachytherapy*. Brachytherapy. Submitted June 15, 2022. Manuscript number BRACHYJOURNAL-D-22-00138.
2. **Orlando N**, Snir J, Barker K, D'Souza D, Velker V, Mendez LC, Fenster A, Hoover DA. *A power Doppler ultrasound method for improving intraoperative needle tip localization in interstitial prostate brachytherapy*. Medical Physics. Submitted March 28, 2022. Manuscript number 22-538.

Peer-Reviewed Conference Proceedings (3 First Author, 5 Total)

1. Park C, Papernick S, **Orlando N**, Jonnalagadda M, Bax J, Gardi L, Barker K, Fenster A. *Toward point-of-care breast cancer diagnosis: validation of a spatially tracked automated 3D ultrasound system*. Proc. SPIE 12038, Medical Imaging 2022: Ultrasonic Imaging and Tomography, 1203804 (2022). **(Robert F. Wagner All Conference Best Student Paper Award runner-up.)**
2. **Orlando N**, Snir J, Barker K, Hoover D, Fenster A. *Wireless oscillating device for power Doppler-based interstitial needle tip identification*. Proc. SPIE 11602, Medical Imaging 2021: Ultrasonic Imaging and Tomography, 11602-19 (2021).
3. Papernick S, **Orlando N**, Park C, Dima R, Gillies DJ, Bax J, Gardi L, Barker K, Appleton T, Fenster A. *Spatially tracked three-dimensional ultrasound imaging for monitoring the synovial membrane in knee arthritis*. Proc. SPIE 11602, Medical Imaging 2021: Ultrasonic Imaging and Tomography, 11602-13 (2021).
4. **Orlando N**, Gillies DJ, Gyacskov I, Fenster A. *Deep learning-based automatic prostate segmentation in 3D transrectal ultrasound images from multiple acquisition geometries and systems.* Proc. SPIE 11315, Medical Imaging 2020: Image-Guided Procedures, Robotic Interventions, and Modeling, 11315-83 (2020).
5. **Orlando N**, Snir J, Barker K, Hoover D, Fenster A. *Power Doppler ultrasound imaging with mechanical perturbation for improved intraoperative needle tip identification during prostate brachytherapy: a phantom study*. Proc. SPIE 10951, Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling, 1095131 (2019).

News Articles

1. Park C and **Orlando N**. *Improving Internal Radiation Treatment of Cervical Cancer with 3D Imaging*. Canadian Cancer Society Research Information Outreach Team (RIOT), News and Views (2021).

Scholarships

2020-2022	**Queen Elizabeth II Graduate Scholarship in Science and Technology** Government of Ontario Value: $15,000 per year ($30,000 total)
2018-2020	**Translational Breast Cancer Research Scholarship** Breast Cancer Society of Canada Value: $18,000 per year ($36,000 total)
2017-2018	**Cancer Research and Technology Transfer (CaRTT) Trainee** CaRTT Strategic Training Program Value: $18,000
2017-2022	**Western Graduate Research Scholarship (WGRS)** Western University Value: $5,000-$7800 per year ($27,800 total)
2014-2016	**Jason Lang Scholarship** Government of Alberta Value: $1,000 per year ($3,000 total)
2013	**University of Alberta Academic Excellence Scholarship** University of Alberta Value: $2,000
2013	**Faculty of Science Academic Excellence Scholarship** University of Alberta Value: $1,000
2013	**Alexander Rutherford Scholarship** Government of Alberta Value: $2,500

Awards and Honours

2022	**First Place Pitch Presentation Award** Imaging Network of Ontario (ImNO) 2022 Imaging for Musculoskeletal Analysis Session Toronto, ON (Virtual Meeting Due to COVID-19) Value: $50

2022 **First Place Pitch Presentation Award**
Imaging Network of Ontario (ImNO) 2022
Hardware, Software, and System Development Session
Toronto, ON (Virtual Meeting Due to COVID-19)
Value: $50

2022 **Robert F. Wagner All Conference Best Student Paper Award runner-up**
SPIE Medical Imaging 2022
San Diego, CA

2021 **New Investigator Award Runner-up in Basic Science**
New Investigator Scientific Session
Annual Integrative Ultrasound Meeting (AIUM) 2021
Orlando, FL (Virtual Meeting Due to COVID-19)

2021 **OICR Rising Stars Poster Award**
OICR Translational Research Conference 2021
Toronto, ON (Virtual Meeting Due to COVID-19)
Value: $100

2021 **First Place Oral Presentation Award**
Imaging Network of Ontario (ImNO) 2021
Machine Learning II Session
Toronto, ON (Virtual Meeting Due to COVID-19)
Value: $250

2020 **Second Place in the John R. Cameron - John R. Cunningham Young Investigators Symposium**
2020 Joint AAPM/COMP Meeting
Vancouver, BC (Virtual Meeting Due to COVID-19)
Value: $230

2020 ***Medical Physics* Editor's Choice**
Medical Physics Volume 47, Issue 6, published in June 2020.

2020 **Dr. Alfred Jay Award for Translational Research**
Department of Medical Biophysics, Western University
Value: $2,000

2019 **Selected to present in "Highlights" section for top scoring abstract submissions**
Canadian Organization of Medical Physicists (COMP) Annual Scientific Meeting
Kelowna, BC

2019 **Third Place Oral Presentation Award**
London Imaging Discovery Day (LIDD) 2019
London, ON

2019	**Top Imaging Abstract** Robarts Research Retreat 2019 London, ON Value: $100
2018	**Third Place Poster Award** London Imaging Discovery Day (LIDD) 2018 London, ON
2018	**Second Place Poster Award in Advances in structural and physiological treatment of disease and therapeutic intervention** London Health Research Day (LHRD) 2018 London, ON
2017	**The Dean's Silver Medal in Science** University of Alberta
2016	**Departmental Undergraduate Student Research Award (USRA)** University of Alberta Value: $4,500
2014-2017	**Dean's Honor Roll** University of Alberta

Research Funding

Agency/Program:	**London Regional Cancer Program (LRCP) Catalyst Grant for Translational Cancer Research**
Project Title:	Development and Implementation of an Accurate and Efficient MRI-to-Ultrasound Registration Framework for Focal Dose Escalation in Prostate Cancer
Project Role:	Co-applicant with Dr. Douglas Hoover (PI), Dr. Lucas Mendez, and Dr. Aaron Fenster
Amount:	Value: $10,574
Funding Period:	January 1, 2021 - December 31, 2022

Presentations

Invited Presentations (5 First Author, 10 Total)

1. **Orlando N**. *Deep learning prostate segmentation in three-dimensional ultrasound.* Invited oral presentation in the inaugural Institute of Physics (IOP) AI in Medical Physics week (Virtual). June 20-24, 2022.
2. Fenster A, Park C, **Orlando N**, Gillies D, Rodgers J. *Applications and trends for use of 3D Ultrasound in image-guided Interventions and point of care diagnostic applications.* Plenary presentation at IEEE International Conference on Imaging Systems & Techniques (IST), New York, NY, USA (Virtual). August 24-26, 2021.

3. Fenster A, **Orlando N**, Gillies D, Rodgers J. *Applications of 3D Ultrasound in Image-guided Prostate and Liver Interventions.* Invited presentation at Xidian University, Xi'an, China. August 23, 2021.
4. **Orlando N**. *A Generalizable and Efficient Deep Learning Algorithm for Automatic Prostate Segmentation in 3D Ultrasound*. 30-minute invited presentation at the Canadian Association of Physicists (CAP) Congress, Virtual. June 6-11, 2021.
5. Fenster A, **Orlando N**, Gillies D, Rodgers J. *Machine Learning for Improving Ultrasound-guided Interventional Cancer Procedures*. Invited oral presentation at 5th International Caparica Conference on Ultrasonic-based applications from analysis to synthesis, Caparica, Portugal. May 31-June 3, 2021.
6. **Orlando N**. *Deep Learning Prostate Segmentation in 3D Ultrasound*. Invited public oral presentation in the Machine Learning for Medical Imaging Consortium (MaLMIC) Forum, Virtual. February 19, 2021.
7. **Orlando N**. *How Artificial Intelligence Can Improve Radiation Therapy for Prostate Cancer.* Invited public oral presentation at RIOT Progress in Cancer Research Symposium, Virtual. November 14, 2020.
8. Fenster A, Rodgers J, **Orlando N**, Gillies D, Park C. *Development of 3D Ultrasound Systems for Image-Guided Interventions*. Invited oral presentation at 4th International Caparica Conference on Ultrasonic-based Applications: from analysis to synthesis, Caparica, Portugal. July 20-23, 2020.
9. **Orlando N**. *Improving Prostate Cancer Radiation Therapy*. Invited public oral presentation at the Kiwanis Club of St. Thomas Golden K, St. Thomas, ON, Canada. November 26, 2019.
10. Fenster A, **Orlando N**, Rodgers J, Park C, Gillies DJ. *2D and 3D Ultrasound Devices for image-guided interventions*. Invited oral presentation at BC Cancer Agency Sindi Ahluwalia Hawkins Centre for the Southern Interior Grand Rounds, Kelowna, BC, Canada. September 30, 2019. **(Group presentation.)**

Oral Presentations (15 First Author, 19 Total)

1. **Orlando N**, Gyacskov I, Gillies DJ, Cool DW, Hoover D, Fenster A. *Deep Learning Prostate Segmentation in 3D Ultrasound and the Impact of Image Quality and Training Dataset Size.* American Association of Physicists in Medicine (AAPM) Annual Meeting, Washington, DC, USA. July 10-14, 2022.
2. Park C, Papernick S, **Orlando N**, Jonnalagadda M, Bax J, Gardi L, Barker K, Fenster A. *Toward point-of-care breast cancer diagnosis: Validation of a spatially tracked automated 3D ultrasound system*. SPIE Medical Imaging: Ultrasonic Imaging and Tomography, San Diego, CA, USA. February 20-24, 2022.
3. **Orlando N**, Snir J, Barker K, Hoover D, Fenster A. *A novel wireless mechanical oscillator for power Doppler-based needle tip identification in brachytherapy*. AAPM Annual Meeting, Virtual. July 25-29, 2021.

4. **Orlando N**, Gillies DJ, Gyacskov I, Romagnoli C, D'Souza D, Hoover D, Fenster A. *Deep learning-based prostate segmentation in 3D ultrasound: a study of dataset size and its effect on performance*. COMP Annual Scientific Meeting, Virtual. June 22-25, 2021.
5. Park C, Papernick S, **Orlando N**, Jonnalagadda M, Bax J, Gardi L, Barker K, Fenster A. *Development of a spatially tracked three-dimensional ultrasound system for whole-breast imaging*. COMP Annual Scientific Meeting, Virtual. June 22-25, 2021.
6. **Orlando N**, Gillies DJ, Gyacskov I, Romagnoli C, D'Souza D, Fenster A. *An Efficient Deep Learning Algorithm for Automatic Prostate Segmentation in Three-Dimensional Ultrasound.* Annual Integrative Ultrasound Meeting (AIUM), Orlando, FL, USA (Virtual). April 10-14, 2021. **(Runner-up award for basic science category in New Investigator Scientific Session.)**
7. **Orlando N**, Gillies DJ, Gyacskov I, Romagnoli C, D'Souza D, Hoover D, Fenster A. *Effect of dataset size and acquisition type on deep learning segmentation of the prostate in 3D ultrasound.* Imaging Network of Ontario (ImNO), Toronto, ON, Canada (Virtual). March 23-24, 2021. **(Received award for oral presentation.)**
8. **Orlando N**, Snir J, Barker K, Hoover D, Fenster A. *Wireless oscillating device for power Doppler-based interstitial needle tip identification*. SPIE Medical Imaging: Ultrasonic Imaging and Tomography, San Diego, CA, USA (Virtual). February 14-18, 2021
9. Papernick S, **Orlando N**, Park C, Dima R, Gillies DJ, Bax J, Gardi L, Barker K, Appleton T, Fenster A. *Spatially tracked three-dimensional ultrasound imaging for monitoring the synovial membrane in knee arthritis*. SPIE Medical Imaging: Ultrasonic Imaging and Tomography, San Diego, CA, USA (Virtual). February 14-18, 2021
10. **Orlando N**, Gillies DJ, Gyacskov I, Romagnoli C, D'Souza D, Fenster A. *rmU-Net: A Generalizable Deep Learning Approach for Automatic Prostate Segmentation in 3D Ultrasound Images*. American Association of Physicists in Medicine (AAPM) and Canadian Organization of Medical Physicists (COMP) Joint Meeting, Vancouver, BC, Canada (Virtual). July 12-16, 2020. **(Second place in the J.R. Cameron - J.R. Cunningham Young Investigator Symposium.)**
11. **Orlando N** and Gillies DJ, Gyacskov I, Romagnoli C, D'Souza D, Fenster A. *A Deep Learning-based Method for Generalized Prostate Segmentation in 3D Ultrasound*. Imaging Network of Ontario (ImNO), Toronto, ON, Canada (Virtual). March 26-27, 2020.
12. **Orlando N**, Gillies DJ, Gyacskov I, Romagnoli C, D'Souza D, Fenster A. *A Deep Learning Approach for Automatic Prostate Segmentation in Clinically Diverse 3D Ultrasound Images*. Western Research Forum (WRF), London, ON, Canada. March 19, 2020 **(Cancelled due to COVID-19.)**
13. **Orlando N**, Gillies DJ, Gyacskov I, Fenster A. *A generalizable deep learning-based method for automatic prostate segmentation on 3D ultrasound*. SPIE Medical Imaging: Ultrasonic Imaging and Tomography, Houston, TX, USA. February 15-20, 2020.
14. **Orlando N**, Snir J, Barker K, Hoover D, Fenster A. *A power Doppler ultrasound-based intraoperative needle tip identification method for interstitial brachytherapy*. Canadian Organization of Medical Physicists (COMP) Annual Scientific Meeting, Kelowna, BC, Canada. September 24-27, 2019.

15. Zhang C, Hilts M, Batchelar D, **Orlando N**, Fenster A, Crook J. *Rigid Registration of Computed Tomography Images and 3D Ultrasound Images for Permanent Breast Seed Implant Planning Procedure*. Canadian Organization of Medical Physicists (COMP) Annual Scientific Meeting, Kelowna, BC, Canada. September 24-27, 2019. **(Selected to present in "Highlights" section for top scoring abstract submissions.)**
16. **Orlando N**, Snir J, Barker K, Hoover D, Fenster A. Development and validation of a power Doppler ultrasound method for improved intraoperative needle tip visualization during high-dose-rate brachytherapy. American Association of Physicists in Medicine (AAPM) Annual Meeting, San Antonio, TX, USA. July 14-18, 2019.
17. **Orlando N**, Snir J, Barker K, Hoover D, Fenster A. *Use of needle-vibrating device and power Doppler ultrasound for improved intraoperative needle tip localization*. London Imaging Discovery Day (LIDD), London, ON, Canada. June 12, 2019. **(Received award for oral presentation.)**
18. **Orlando N**, Snir J, Barker K, Hoover D, Fenster A. *Development of a power Doppler ultrasound method for improved intraoperative needle tip localization during interstitial brachytherapy*. Robarts Research Retreat, London, ON, Canada. June 7, 2019.
19. **Orlando N**, Snir J, Barker K, Hoover D, Fenster A. *Power Doppler ultrasound for improved needle tip visualization during prostate brachytherapy.* Western Canadian Medical Physics Meeting (WesCan), Edmonton, AB, Canada. March 28-30, 2019.

Poster Presentations (19 First Author, 24 Total)

1. **Orlando N**, Snir J, Barker K, D'Souza D, Velker V, Mendez LC, Fenster A, Hoover D. *Clinical validation of a power Doppler-based needle visualization method in high-dose-rate prostate brachytherapy procedures.* Canadian Organization of Medical Physicists (COMP) Annual Scientific Meeting, Quebec City, QC, Canada. June 22-25, 2022.
2. **Orlando N**, Tessier D, Hoover D, Fenster A. *Needle Tip Identification in Clinical Power Doppler Ultrasound Using Induced Vibrations by an Innovative Mechanical Oscillator*. Canadian Association of Physicists (CAP) Congress, Hamilton, ON, Canada. June 5-10, 2022.
3. **Orlando N**, Du Toit C, Papernick S, Dima R, Gyacskov I, Fenster A. *A deep learning algorithm for automatic cartilage segmentation in three-dimensional ultrasound images of healthy participants.* Imaging Network of Ontario (ImNO), Toronto, ON, Canada (Virtual). March 22-24, 2022. **(Received award for pitch presentation.)**
4. Park C, Xing S, Papernick S, **Orlando N**, Bax J, Gardi L, Barker K, Tessier D, Fenster A. *Spatially tracked whole-breast three-dimensional ultrasound system toward point-of-care breast cancer screening in women with dense breasts.* Imaging Network of Ontario (ImNO), Toronto, ON, Canada (Virtual). March 22-24, 2022. **(Received award for pitch presentation.)**
5. **Orlando N**, Gyacskov I, Gillies DJ, Romagnoli C, D'Souza D, Hoover D, Fenster A. *Deep learning prostate segmentation in three-dimensional ultrasound: the role of dataset size*. Image-guided Therapeutics & Diagnostics Symposium, University of British Columbia, Virtual. November 17-18, 2021.

6. Park CK, Xing S, Papernick S, **Orlando N**, Bax J, Gardi L, Barker K, Tessier D, Fenster A. *Whole-breast phantom validation of a spatially tracked 3D ultrasound system for point-of-care breast cancer imaging*. Image-guided Therapeutics & Diagnostics Symposium, University of British Columbia, Virtual. November 17-18, 2021.
7. Park CK, Papernick S, **Orlando N**, Jonnalagadda M, Bax J, Gardi L, Barker K, Fenster A. *Automatic Spatial Registration System for Whole-Breast Three-Dimensional Ultrasound*. ePoster and live discussion session at AAPM Annual Meeting, Virtual. July 25-29, 2021. **(Interactive ePoster Session for high-scoring abstracts.)**
8. **Orlando N**, Gyacskov I, Gillies DJ, Romagnoli C, D'Souza D, Hoover D, Fenster A. *How does dataset size and diversity influence prostate segmentation performance in 3D ultrasound?* Robarts Research Retreat, Virtual. June 17, 2021.
9. Park CK, Papernick S, **Orlando N**, Jonnalagadda M, Bax J, Gardi L, Barker K, Fenster A. *Spatially-tracked 3DUS system for whole-breast imaging*. Robarts Research Retreat, Virtual. June 17, 2021.
10. **Orlando N**, Snir J, Barker K, Hoover D, Fenster A. *A wireless oscillator for power Doppler ultrasound-based needle identification in brachytherapy*. Oncology Research and Education Day, Virtual. June 14-18, 2021.
11. **Orlando N**, Gillies DJ, Gyacskov I, Romagnoli C, D'Souza D, Hoover D, Fenster A. *A Robust and Efficient Deep Learning Prostate Segmentation Algorithm for 3D Transrectal Ultrasound Images*. London Health Research Day 2021, London, ON, Canada (Virtual). May 11, 2021.
12. **Orlando N**, Snir J, Barker K, Hoover D, Fenster A. *Power Doppler Ultrasound: A Method for Improving Needle Tip Identification in Brachytherapy*. OICR Translational Research Conference (TRC), Toronto, ON, Canada (Virtual). March 25-26, 2021. **(Received award for e-poster presentation.)**
13. **Orlando N** and Gillies DJ, Gyacskov I, Fenster A. *Deep learning-based automatic prostate segmentation in 3D transrectal ultrasound images from multiple acquisition geometries and systems*. SPIE Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling, Houston, TX, USA. February 15-20, 2020.
14. **Orlando N**, Snir J, Barker K, Hoover D, Fenster A. *Improved needle tip visualization in interstitial brachytherapy using needle vibration.* Oncology Research and Education Day, London, ON, Canada. June 14, 2019.
15. **Orlando N**, Snir J, Barker K, Hoover D, Fenster A. *Power Doppler ultrasound for improved intraoperative needle tip visualization during prostate brachytherapy*. London Health Research Day (LHRD), London, ON, Canada. April 30, 2019.
16. **Orlando N**, Snir J, Barker K, Hoover D, Fenster A. *Improved Intraoperative Needle Tip Identification Using Power Doppler Ultrasound.* Imaging Network of Ontario (ImNO), London, ON, Canada. March 28-29, 2019.
17. **Orlando N**, Snir J, Barker K, Hoover D, Fenster A. *Improving needle tip localization in interstitial prostate brachytherapy using power Doppler imaging.* OICR Translational Research Conference, Toronto, ON, Canada. March 21-22, 2019.

18. **Orlando N**, Snir J, Barker K, Hoover D, Fenster A. *Power Doppler ultrasound imaging with mechanical perturbation for improved intraoperative needle tip identification during prostate brachytherapy: a phantom study.* SPIE Medical Imaging, San Diego, CA, USA. February 16-21, 2019.
19. **Orlando N**, Hoover D, Edirisinghe C, D'Souza D, Wong E, Fenster A. *An examination of needle deflection in high dose rate prostate cancer brachytherapy.* CARO-COMP-CAMRT Joint Scientific Meeting, Montreal, QC, Canada. September 12-15, 2018.
20. Snir J, **Orlando N**, Patrick J, Surry K, Velker V, Hoover D. *Mechanical perturbation enhanced power Doppler ultrasound imaging for improved intra-operative localization of interstitial brachytherapy needles.* CARO-COMP-CAMRT Joint Scientific Meeting, Montreal, QC, Canada. September 12-15, 2018.
21. **Orlando N**, Hoover D, Edirisinghe C, D'Souza D, Wong E, Fenster A. *Needle deflection in high dose rate prostate cancer brachytherapy.* Eight-minute e-poster presentation at London Imaging Discovery Day (LIDD), London, ON, Canada. June 14, 2018. **(Received award for poster presentation.)**
22. **Orlando N**, Hoover D, Edirisinghe C, D'Souza D, Wong E, Fenster A. *An examination of needle deflection in high dose rate prostate cancer brachytherapy.* Oncology Research and Education Day, London, ON, Canada. June 8, 2018.
23. **Orlando N**, Hoover D, Edirisinghe C, D'Souza D, Wong E, Fenster A. *An examination of needle deflection in high dose rate prostate cancer brachytherapy.* Robarts Research Retreat, London, ON, Canada. June 1, 2018.
24. **Orlando N**, Hoover D, Edirisinghe C, Wong E, Fenster A. *Quantifying needle deflection in high dose rate prostate cancer brachytherapy.* London Health Research Day (LHRD), London, ON, Canada. May 10, 2018. **(Received award for poster presentation.)**

Leadership Experience and Academic Service

09/2021 – 04/2022	**Co-supervisor – Engineering Co-op Student** Summary: Co-supervised 8-month research co-op at the London Regional Cancer Program focusing on deep learning prostate segmentation.
09/2021 – 04/2022	**CAMPEP Mentor – Medical Biophysics Academic Mentorship Program** Summary: Responsible for mentoring a one-year CAMPEP MSc student in the Department of Medical Biophysics.
11/2021 – 12/2021	**Organizer – Robarts Professional Development Miniseries** Summary: Part of the organizing committee for a three-part professional development series featuring Dr. Aaron Fenster.
06/2021	**J.R. Cunningham Young Investigator Symposium Co-moderator** COMP 2021 Summary: Co-moderated this marquee session at the COMP 2021 Annual Scientific Meeting.

03/2021	**Pitch Presentation Co-chair and Judge** ImNO 2021 Summary: Responsible for chairing and judging two pitch presentation sessions at ImNO 2021.
09/2020 – 08/2021	**Mentor - Medical Biophysics Academic Mentorship Program** Summary: Responsible for mentoring a first-year MSc student in the Department of Medical Biophysics.
01/2020 – 07/2022	**Deep Learning Club – Organizing Committee Founding Member** Summary: Founding member of this committee which built and connected a network of students interested in AI from across Western University.
07/2019 – 07/2021	**Robarts Association of Trainees Committee Member** Summary: Elected committee which represents, connects, and advocates for trainees at the Robarts Research Institute.
01/2019 – 04/2019	**Founding member and Academic Coordinator - Medical Biophysics Undergraduate Research Day Planning Committee** Summary: Helped organize the first three annual Medical Biophysics undergraduate research conference at Western University.
12/2019 – 04/2020 **02/2021 – 04/2021**	
09/2018 – 08/2021	**CAMPEP Student Representative** Summary: One of two peer-elected student members on the CAMPEP steering Committee.

Teaching Experience

09/2020 – 04/2021 **09/2021 – 04/2022**	**Teaching Assistant – Practical Radiotherapy Physics - BIOPHYS 9672** Course Instructor: Kathleen Surry, Ph.D., MCCPM Western University
02/2021	**Invited Guest Lecturer - Biomedical Applications of Neural Networks - BIOPHYS 9709** Western University
01/2020 – 05/2020	**Pedagogy in Biophysics – BIOPHYS 9674B** Summary: Contributed to the development of a graduate-level course curriculum titled “Biomedical Applications of Neural Networks”.
10/03/2020	**An Introduction to Deep Learning with Keras Tutorial** Summary: 30-minute tutorial introducing deep learning with Keras.

Clinical Experience

05/2018 – 09/2018 **09/2019 – 01/2020**	**Linear Accelerator Mechanical and Output Quality Assurance** London Regional Cancer Program, London, ON, Canada Supervisor: Rob Barnett, Ph.D., FCCPM

Independently performed routine mechanical and output quality assurance measurements for Varian Clinac iX and TrueBeam linear accelerators.

Undergraduate Research Experience

09/2016 – 12/2016 **Undergraduate Honors Thesis Project**
Title: SAR Dosimeter for MRI
Supervisor: Nicola De Zanche, Ph.D., MCCPM
Summary: Designed and built a dosimeter to independently verify Specific Absorption Rate (SAR) levels of an MRI system.

05/2016 – 08/2016 **Undergraduate Research Assistant**
Department of Physics, University of Alberta, Edmonton, AB, Canada
Supervisor: Alkiviathes Meldrum, Ph.D.
Summary: Assisted with the development of a refractometric liquid and gas sensor using both polymer and silver coated microcapillaries.

Professional Development Experience

09/2017 – 06/2022 **Own Your Future Trainee**
Summary: Curriculum-based professional development program at Western University designed specifically for doctoral students.

09/2017 – 08/2018 **Cancer Research and Technology Transfer (CaRTT) Trainee**
Summary: Training program in cancer biology and treatment which aims to development partnerships between basic and clinical researchers.

Professional Memberships

2019 – Present **American Association of Physicists in Medicine (AAPM)**
Student Member

2018 – Present **Canadian Organization of Medical Physicists (COMP)**
Student Member

2021 – 2022 **Canadian Association of Physicists (CAP)**
Student Member

2019 – 2022 **Society of Photo-Optical Instrumentation Engineers (SPIE)**
Student Member