# UCLA
## UCLA Electronic Theses and Dissertations

**Title**

Prostate Cancer Detection with Micro-Ultrasound: Evaluation of Expert Readers and Deep Learning Classifiers Through Accurate Co-registration with Final Pathology

**Permalink**

https://escholarship.org/uc/item/57x309wn

**Author**

Pensa, Jake

**Publication Date**

2024

Peer reviewed|Thesis/dissertation

UNIVERSITY Of CALIFORNIA

Los Angeles

Prostate Cancer Detection with Micro-Ultrasound:

Evaluation of Expert Readers and Deep Learning Classifiers

Through Accurate Co-registration with Final Pathology

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Bioengineering

by

Jake William Pensa

2024

# ABSTRACT OF THE DISSERTATION

Prostate Cancer Detection with Micro-Ultrasound:

Evaluation of Expert Readers and Deep Learning Classifiers

Through Accurate Co-registration with Final Pathology

by

Jake William Pensa

Doctor of Philosophy in Biomedical Engineering

University of California, Los Angeles, 2024

Professor Leonard S. Marks, Co-Chair

Professor Corey Wells Arnold, Co-Chair

Prostate cancer (PCa) is a leading cause of cancer related mortality in men in the United States. Imaging is crucial for PCa staging and treatment decisions. Currently, the gold standard is a multi-parametric magnetic resonance imaging (mpMRI) scan followed by confirmatory ultrasound (US) guided biopsy. While mpMRI has good sensitivity and specificity for index lesion detection, the largest and most aggressive lesion, it is resource intensive, expensive, and generally limited to larger centers.

Micro-ultrasound (microUS) is a high-resolution US imaging system that has recently been introduced as a low-cost alternative to mpMRI for imaging PCa. While

preliminary results are encouraging, microUS lacks direct comparison with mpMRI and ground-truth whole-mount (WM) pathology. It also requires substantial hands-on experience for accurate interpretation. Ultimately, skepticism relative to established mpMRI imaging and barriers to entry for new users among other reasons have limited microUS adoption.

This thesis details a comprehensive investigation into microUS imaging and its role in PCa diagnosis. To this end we performed a clinical trial collecting microUS images of patients undergoing radical prostatectomy for PCa with corresponding mpMRI imaging and WM pathology. Using this dataset, we were able to quantitatively explain the diagnostic ability of microUS, however its utility relative to mpMRI required further investigation.

To quantify the diagnostic ability of microUS, expert and novice readers reviewed the collected microUS and mpMRI images. For correlation of the delineated regions of interest (ROIs), we developed and validated a novel co-registration process to allow for direct registration of microUS and mpMRI with WM pathology with a registration error of 3.90±0.11mm. Comparing the confirmed regions of cancer on WM pathology with the identified regions on microUS and mpMRI, we found microUS to have similar detection of index lesions to equivalent MRI review (91.7% vs 80%). In an attempt to alleviate inter-reader variability and the reduced performance of novice users, we trained deep learning classifiers to identify PCa on microUS imaging using our accurately labeled data. These models exhibited expert level performance for image classification and targeted biopsy guidance. Further work is necessary for clinical adoption of deep learning models for microUS imaging of PCa.

The dissertation of Jake William Pensa is approved.

Rory Geoghegan

Shyam Natarajan

Robert N. Candler

Holden H. Wu

Leonard S. Marks, Committee Co-Chair

Corey Wells Arnold, Committee Co-Chair

University of California, Los Angeles

2024

*Dedicated to my family, especially Karl von Emster*

# Table of Contents

# LIST OF FIGURES

# LIST OF TABLES

# ACKNOWLEDGMENTS

This thesis would not have been possible without the unwavering support of my advisors and mentors. Thank you, Drs. Leonard Marks, Rory Geoghegan, and Wayne Brisbane. Your continual guidance, mentorship, and encouragement over the years has been instrumental in my development as a scientist and engineer. I am immensely grateful for the countless hours you have selflessly dedicated to my success. I would also like to thank my other committee members, Drs. Corey Arnold, Shyam Natarajan, Robert Candler, and Holden Wu. I would especially like to extend my appreciation to Dr. Corey Arnold for serving as the co-chairman of my doctoral committee.

I would also like to posthumously thank the late Dr. Warren Grundfest. As my first mentor at UCLA, you opened my eyes to biomedical research, taught me to think critically, and inspired my career as an academic and researcher. I will forever be grateful for your sponsorship of my doctoral pursuits at UCLA.

Thank you to my colleagues and collaborators Drs. Adam Kinnaird, David Kuppermann, Alan Priester, Anthony Sisk, Ely Felker, Erik Dutson, and Brian Wodlinger. Your invaluable expertise and contributions to this work ensured its success and I am extremely grateful for your help.

I would also like to recognize the help and support over the years from my fellow CASIT engineers and clinical colleagues: Catherina, Griffith, Derrick, Kyle, Merdie, Sean, Samantha, Eli, Lorna, and Mark.

Lastly, I would like to thank my family and friends for their never-ending support throughout this long and arduous process. Thank you, Karyn, Dave, Tyler, Arley, and Cassidy, your unwavering support and encouragement has been instrumental in helping me navigate this journey and accomplish beyond what I originally thought possible.

This thesis contains prior published work as part of a collaborative effort by myself and others. Permissions for use have been obtained from the proper party. The published works in the order they appear in this thesis are included below.

Pensa J, Geoghegan R, Natarajan S. 3D Ultrasound for Biopsy of the Prostate. In *3D Ultrasound*: *Devices, Applications, and Algorithms* 154-175. CRC Press. 2023

Pensa J, Brisbane W, Priester A, Sisk A, Marks L, Geoghegan R. A System for Co-Registration of High-Resolution Ultrasound, Magnetic Resonance Imaging, and Whole-Mount Pathology for Prostate Cancer. 2021 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2021.

Pensa J, Brisbane W, Kinnaird A, Kuppermann D, Hughes G, Ushko D, Priester A, Gonzalez S, Reiter R, Chin A, Sisk A, Felker E, Marks L, Geoghegan R. Evaluation of Prostate Cancer Detection using Micro-Ultrasound versus MRI through Co-registration to Whole-Mount Pathology. Scientific Reports, 2024

Pensa J, Brisbane W, Kinnaird A, Kuppermann D, Hughes G, Ushko D, Sisk A, Marks L, Geoghegan R. Deep Learning Classification of Prostate Cancer on Confidently Labeled Micro-Ultrasound Images. 2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2024. (In press)

Pensa J, Brisbane W, Kinnaird A, Kuppermann D, Hughes G, Sisk A, Arnold C, Marks L, Geoghegan R. Deep Learning Classification of Prostate Cancer on Micro-Ultrasound Imaging; Validated Against Whole-Mount Pathology and Expert Readers. (In progress)

# VITA

| | |
|---|---|
| 2014-2018 | **B.S. Bioengineering**<br>University of California, Los Angeles |
| 2018-2019 | **M.S. Bioengineering**<br>University of California, Los Angeles |
| 2019 | **Research and Development Intern**<br>Edwards Lifesciences, Irvine |
| 2018-present | **Graduate Student Researcher**<br>UCLA Health, Department of Urology<br>University of California, Los Angeles |

# PUBLICATIONS & PRESENTATIONS

**Pensa J**, Brisbane W, Kinnaird A, Kuppermann D, Hughes G, Ushko D, Priester A, Gonzalez S, Reiter R, Chin A, Sisk A, Felker E, Marks L, Geoghegan R. Evaluation of Prostate Cancer Detection using Micro-Ultrasound versus MRI through Co-registration to Whole-Mount Pathology. Scientific Reports, 2024

**Pensa J**, Brisbane W, Kinnaird A, Kuppermann D, Hughes G, Ushko D, Sisk A, Marks L, Geoghegan R. Deep Learning Classification of Prostate Cancer on Confidently Labeled Micro-Ultrasound Images. 2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2024. (In press)

**Pensa J**, Brisbane W, Kinnaird A, Kuppermann D, Hughes G, Sisk A, Arnold C, Marks L, Geoghegan R. Deep Learning Classification of Prostate Cancer on Micro-Ultrasound Imaging; Validated Against Whole-Mount Pathology and Expert Readers. (In progress)

**Pensa J**, Geoghegan R, Natarajan S. 3D Ultrasound for Biopsy of the Prostate. In *3D Ultrasound: Devices, Applications, and Algorithms* 154-175. CRC Press. 2023

**Pensa J**, Brisbane W, Priester A, Sisk A, Felker E, Marks L, Geoghegan R. Prostate Cancer Diagnosis with Micro-Ultrasound – A Pilot Clinical Study. American Urological Association; Engineering and Urology Society 2023.

**Pensa J**, Ushko D, Hughes G, Brisbane W, Kinnaird A, Kuppermann D, Sisk A, Marks L, Geoghegan R. Prostate Ductal Anatomy as a Cancer Contrast Mechanism for Ultrasound. 2022 44th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2022.

**Pensa J**, Brisbane W, Priester A, Hughes G, Marks L, Geoghegan R. Co-Registration of Micro-Ultrasound, Magnetic Resonance Imaging, and Whole-Mount Pathology for Prostate Cancer. American Urological Association; Engineering and Urology Society 2022.

**Pensa J**, Brisbane W, Priester A, Sisk A, Marks L, Geoghegan R. A System for Co-Registration of High-Resolution Ultrasound, Magnetic Resonance Imaging, and Whole-Mount Pathology for Prostate Cancer. 2021 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2021.

Imran M, Nguyen B, **Pensa J**, Falzarano S, Sisk A, Liang M, DiBianco J, Su L, Zhou Y, Joseph J, Brisbane W, Shao W. Image registration of in vivo micro-ultrasound and ex vivo pseudo-whole mount histopathology images of the prostate: A proof-of-concept study. Biomedical Signal Processing and Control. 2024.

Jiang H, Imran M, Muralidharan P, Patel A, **Pensa J**, Liang M, Benidir T, Grajo J, Joseph J, Terry R, DiBianco J, Su L, Zhou Y, Brisbane W, Shao W. MicroSegNet: A deep learning approach for prostate segmentation on micro-ultrasound images. Computerized Medical Imaging and Graphics. 2024.

Schaer S, Rakauskas, Dagher J, La Rosa S, **Pensa J**, Brisbane W, Marks L, Kinnaird A, Abouassaly R, Klein E, Thomas L, Meuwly J, Parker P, Roth B, Valerio M. Assessing cancer risk in the anterior part of the prostate using micro-ultrasound: validation of a novel distinct protocol. World Journal of Urology. 2023.

Juo Y, **Pensa J**, Sanaiha Y, Abiri A, Sun S, Tao A, Vogel S, Kazanjian K, Dutson E, Grundfest W, Lin A. Reducing retraction forces with tactile feedback during robotic total mesorectal excision in a porcine model. Journal of Robotic Surgery. 2022.

Ouyang Q, Wu J, Sun S, **Pensa J**, Abiri A, Dutson E, Bisley J. Bio-Inspired Haptic Feedback for Artificial Palpation in Robotic Surgery. IEEE Transactions on Biomedical Engineering. 2021.

Juo Y, Abiri A, **Pensa J**, Sun S, Tao A, Bisley J, Grundfest W, Dutson E. Center for Advanced Surgical and Interventional Technology Multimodal Haptic Feedback for Robotic Surgery. In *Handbook of Robotic and Image-Guided Surgery*. 2020.

# CHAPTER 1
## Introduction

Prostate cancer (PCa) is a leading cause of cancer related morbidity and mortality for men in the United States with a projected 35,000 deaths and 300,000 new cases in 2024 [1]. While high-grade, clinically significant PCa can be deadly, low-grade, clinically insignificant PCa has less than a 2% chance of cancer specific mortality 10 years after diagnosis [2], [3]. As such accurate diagnosis is vital for proper treatment of PCa, since overtreatment of low-grade cancer can severely impact the quality of life of patients and increase healthcare expenditures, while undertreatment of high-grade cancer can be fatal [3].

The conventional diagnostic pathway utilizes an ultrasound (US) guided systematic biopsy following an elevated value on a prostate-specific antigen (PSA) screening test [4]. While US allows for general gland visualization and biopsy guidance, it has poor sensitivity and specificity for tumor detection [5]. In contrast, multi-parametric magnetic resonance imaging (mpMRI) has high sensitivity (86-92%) and specificity (60-83%) for identification of prostate cancer index lesions [6], [7], [8]. Index lesions are defined as the largest and most aggressive tumor in a prostate, the size and grade of which are often used to decide the appropriate treatment pathway [9]. Consequently, current recommendations include an MRI-guided biopsy or more commonly: MR-US fused biopsy [10], [11]. For fusion biopsies, a pre-biopsy mpMRI identified region-of-interest (ROI) is fused to a live US scan allowing for US guided targeted biopsy of an MRI ROI [12]. This targeted approach improves cancer detection,

1

however, mpMRI is resource intensive, has substantial inter-observer variability, can still fail to detect 10-15% of clinically significant lesions, and tumor volume estimates can be low as 1/3 of the true volume [7], [13], [14], [15]. In an effort to resolve these issues, a low-cost high resolution ultrasound technology known as micro-ultrasound (microUS) (ExactVu, Exact Imaging, Ontario Canada) has recently been introduced. This technology may enable direct identification of cancerous tissue on live US imaging, potentially mitigating the need for a pre-biopsy MRI.

MicroUS is a promising new modality, however it has not been robustly evaluated against mpMRI in clinical studies validated by ground truth histology. Previous studies validated with tracked biopsy have demonstrated similar cancer detection rates between microUS (35%) and mpMRI (39%) with comparable sensitivities (93% vs 94%) and specificities (21% vs 17%) [16], [17], [18]. Although these studies have demonstrated promise for microUS imaging, the gold standard evaluation for prostate imaging is validation against ground truth whole-mount (WM) pathology slides. Furthermore, the system lacks a scientific justification for why increased resolution allows for PCa to be visualized. Additionally, as a new imaging modality, there is a considerable training requirement for new users to correctly interpret microUS images and considerable inter-reader variability among expert reviewers [19]. Ultimately, skepticism relative to established mpMRI imaging and barriers to entry for new users among other reasons have limited microUS adoption.

This thesis details our efforts to investigate and improve the diagnostic ability of microUS. Our central hypotheses are: 1) microUS enables cancer detection similar to mpMRI and 2) accuracy and ease of use can be improved through implementation of deep learning image classification algorithms.

2

## 1.1.   Overview of Chapters

Chapter 2 provides background information on prostate cancer and related imaging as well as an extensive review of microUS. Chapter 3 outlines a clinical trial in which microUS scans, mpMRI scans, and WM pathology slides were collected from patients undergoing radical prostatectomy for prostate cancer. Chapter 4 describes the development and validation of an ex vivo microUS scanning system and related co-registration with WM pathology. Chapter 5 details the development, evaluation, and implementation of in vivo microUS registration with final histology. A reader study implementing this registration technique is subsequently performed to evaluate the diagnostic accuracy of microUS and MRI. Lastly, the implementation of deep learning classifiers for prostate cancer detection on microUS is explored in Chapter 6. Below is a brief summary of each chapter.

### 1.1.1.   Chapter 2

This chapter serves to introduce the reader to prostate anatomy and associated cancer, with a focus on the morphology of cancer and the index lesion theory. From there it describes prostate cancer screening and diagnosis. Beyond that there is an overview of imaging, primarily ultrasound and mpMRI due to their prevalence in PCa diagnosis as well as the related fusion of the two modalities for targeted biopsy. The last section of this chapter covers microUS specifically, as well as deep learning approaches for prostate cancer imaging. The intersection of PCa morphology and imaging physics is vital for understanding mechanisms behind PCa visualization. Additionally, while US and mpMRI are both imaging modalities for visualizing the prostate, their specific roles

in the diagnostic pathway are different, as mpMRI is typically a noninvasive triage for biopsy. Conversely, US is transrectal, often requires local anesthesia, and imaging is performed during biopsy. MicroUS presents a new method of PCa visualization and diagnosis and the specifications of the system and preliminary studies are discussed. In addition, prior application of deep learning techniques for PCa on both MRI and microUS are highlighted. Ultimately, this chapter serves to frame the work presented in this thesis in context for the reader.

## 1.1.2. Chapter 3

In order to thoroughly investigate microUS imaging of prostate cancer, we needed to collect a high-quality dataset. Patients with biopsy confirmed prostate cancer and a recent mpMRI scan set to undergo radical prostatectomy were enrolled in a clinical trial. In the operating room on the day of surgery, patients received a comprehensive microUS scan of their prostate. Following radical prostatectomy and gland removal, the prostate received another microUS scan using a custom-built ex vivo scanning gantry developed and validated for this trial. After scanning, the prostate was grossed and sectioned following routine clinical practice. Furthermore, as a part of routine clinical practice, the prostate is sectioned in a patient specific 3D printed sectioning mold, a process implemented by our group. Following this process, we obtained spatially tracked, in vivo microUS, ex vivo microUS, mpMRI, and WM pathology for 86 patients with prostate cancer. To our knowledge this is the highest quality microUS imaging dataset collected to date, allowing for comprehensive examination of both mpMRI and microUS imaging.

As we acquired WM pathology and microUS imaging data, we used this data to investigate our hypothesis of why microUS imaging is diagnostic for PCa when

conventional US is not. In summary, our hypothesis was: there are morphological features that are 1) smaller than the resolution limit on conventional US (200µm) and larger than the resolution limit on microUS (70µm) and 2) these features correlate with the presence of cancer. Using computer vision, prostatic features were measured on WM pathology in regions of cancer and regions of benign tissue. Aligning with our hypothesis, we found prostatic ducts are generally too small to be resolved by conventional US but big enough to be resolved on microUS, and there is a significantly reduced amount of ductal space present in areas of cancer relative to benign tissue. While this is not a comprehensive explanation for why microUS is diagnostic for PCa, it does provide confidence that there is a measurable explanation for why microUS can visualize PCa when conventional US is non-diagnostic.

### 1.1.3. Chapter 4

In this chapter we cover the development and validation of our ex vivo scanning gantry introduced in the prior chapter. The system was designed to enable spatially tracked microUS imaging in the axial orientation, which is impossible to obtain natively with an in vivo scan. Using a translation stage, 3D printed flexible alignment molds, and US absorbent material, we were able to collect images that can be matched directly to mpMRI images and WM pathology slides. Furthermore, we also used the system to scan patients with conventional US probes, allowing for direct comparison between microUS and conventional US for visualization of prostate cancer.

The goal of this scanning system is to facilitate co-registration of microUS imaging, mpMRI, and WM pathology. The crux of the issue lies in in vivo microUS imaging being acquired in a freehand rotational fan shaped, para-sagittal sweep in the rectum,

while mpMRI and WM pathology are both acquired axially. Furthermore, microUS probes only have an integrated rotational sensor, leaving the pitch and yaw untracked in addition to any translation. As such, accurate registration with mpMRI and WM pathology is challenging. At the time of starting this study it was not clear if this would be a surmountable challenge, as a prior study by Palmeri et al had to utilize zonal based registration rather than direct registration of para-sagittal US to final histology even with a mechanical stepper for scanning [20].

The primary purpose of ex vivo microUS imaging was to facilitate straightforward registration of microUS with mpMRI and WM pathology, due to all imaging being in the axial orientation if in vivo registration proved too difficult. Furthermore, alternatively ex vivo microUS imaging could serve to inform the transformation of para-sagittal in vivo microUS to axial orientation due to improved spatial tracking. Briefly, ex vivo microUS imaging and mpMRI were registered to WM pathology through slice alignment, then a two-step capsule based registration. The first step involved a rigid alignment of capsule orientation, scaling, and rotation. The second step involved a non-rigid thin-plate-spline (TPS) registration to account for tissue deformation and differences in the capsule. To evaluate target registration error (TRE), common internal fiducials across all three methodologies were selected and the distance between each selected fiducial was measured. Registration of ex vivo microUS and mpMRI to WM pathology had a TRE of 2.16mm ± 1.02mm and 2.34mm ± 1.35mm respectively, similar to other WM registration studies.

We then performed a pilot reader study. An expert microUS reader delineated regions of suspected cancer on ex vivo microUS imaging in 3 patients. These ROIs were subsequently registered to WM pathology and evaluated against the gold standard

mpMRI ROIs. The expert microUS reader was able to correctly identify 3/3 index lesions and 13/20 tumor cross-sections. Comparatively mpMRI review identified 2/3 index lesions and 4/20 tumor cross-sections. Ultimately these results provide confidence that if in vivo microUS registration with WM pathology is not possible, ex vivo imaging can be used as a proxy for tumor identification on WM pathology and data labeling for deep learning. While promising, the results presented in this chapter will need to be confirmed in a larger study. Additionally, ex vivo microUS imaging is not necessarily indicative of in vivo microUS imaging and its use limits the generalizability of this study. As such, implementation of in vivo imaging is preferred, the focus of the next chapter.

## 1.1.4.  Chapter 5

In this chapter, we develop and validate a methodology for direct registration of in vivo microUS imaging with mpMRI and WM pathology. As mentioned in the prior chapter, this is challenging due to the scanning orientation and tissue deformation of native microUS imaging. More specifically, the para-sagittal orientation collected with limited probe positional information and the compression required to sufficiently interrogate the prostate makes registration difficult. In our preliminary efforts we found capsule-based information alone is not informative enough to compensate for these registration issues, while the high amount of texture from diffuse US reflections provides excessive noise for mutual information-based registration.

To overcome the challenges of in vivo microUS registration with WM pathology, we developed and validated a three-step registration methodology. The process consists of: 1) reconstruction of microUS imaging in the true sagittal orientation to allow for 3D positioning and approximation of the overall prostate gland shape 2) 3D capsule and

urethra based TPS registration of reconstructed microUS with MRI to account for tissue deformation from microUS scanning and reconstruction errors due to a lack of tracked positional information during scanning, and 3) 2D capsule based registration of MRI imaging with WM pathology sections. After validation of this process, the overall registration error was found to be 3.90±0.11mm, on par with commercial MR-US fusion systems and prior published MRI to WM registration studies [21], [22].

Following registration, we performed a reader study to evaluate the diagnostic accuracy of microUS relative to MRI using direct evaluation against WM pathology. Utilizing the developed registration methodology we accurately registered microUS and MRI delineations to WM pathology sections and measured the degree of overlap between reviewer identified ROIs and the actual extent of cancer.

Expert and novice readers delineated regions of interest on microUS scans from 15 patients. In addition to the original MRI review prior to surgery, an expert radiologist also performed a retrospective read of the microUS cases for equal comparison to the retrospective microUS review. Utilizing the developed registration process, the reviewer delineations were co-registered to WM pathology and ROI agreement with WM confirmed cancer locations and extent was quantitatively measured. This direct registration allows for an especially granular analysis of both mpMRI and microUS providing a rigorous investigation into the two modalities. Expert microUS readers were found to have equivalent performance to mpMRI for prostate cancer detection with similar rates of index lesion identification (91.7% vs 80% respectively), slightly improved identification of tumor extent (52.5% vs 36.7% respectively), and similar false positive rates (38.3% vs 40.8%). We also found a significant difference in microUS results based on reviewer expertise, as the urologist with limited microUS experience identified only 50% of the

index lesions and had significantly worse lesion overlap than the other three expert microUS readers. As such, we concluded training beyond baseline competency and overall biopsy experience is vital for correct interpretation of microUS imaging highlighting the need for deep learning solutions to alleviate these barriers to adoption.

## 1.1.5. Chapter 6

This chapter describes the development and evaluation of a deep learning solution for prostate cancer identification on microUS. As noted earlier, utilization of accurate deep learning models may eradicate much of the inter-reader variability and training requirement surrounding microUS implementation. In this chapter, we train and validate microUS classifiers on high quality labeled data with direct comparison to the classification ability of expert microUS and MRI reviewers.

Inverting the order of the registration process developed in chapter 5, we labeled reconstructed microUS imaging with the location of WM pathology confirmed prostate cancer. If an image contained confirmed cancer, it was labeled as cancerous, and if it did not contain any WM confirmed lesions it was labeled as benign. With the data labeled, pretrained image net models were fine-tuned on a subset of the dataset. Withheld data from training was used to quantify the accuracy of the microUS models. For a pilot study we evaluated the potential of deep learning classifiers in a dataset of 15 patients (12 training, 3 test) and found similar results to expert reviewers (78.9% vs 60.6% respective sensitivity, 72.7% vs 80.5% respective specificity) with an AUC of 0.80. We then confirmed these results in a larger study with 100 patients and three different methods of training. In this study the deep learning classifier achieved an AUC of 0.82 and equal sensitivity to the expert reviewers (0.72 vs 0.71 respectively) at similar specificity (0.8 vs 0.83

respectively). Furthermore, when confined to only predicting one image per patient as prostate cancer, to approximate identifying a targeted biopsy location, the microUS classifier identified clinically significant PCa 82.6% of the time. For reference, the center of the MRI identified ROI (gold standard) identified clinically significant cancer 80% of the time. Ultimately, these results indicate deep learning models can be used to identify prostate cancer on microUS and have the ability to alleviate much of the challenges associated with microUS adoption.

# CHAPTER 2

# Background

## 2.1.   Prostate Anatomy and Cancer

### 2.1.1.  Prostate Anatomy

The prostate is a "walnut-shaped" gland located in subperitoneal compartment anterior to the rectum [23]. The gland is comprised of 3 main zones: the peripheral zone (70% of prostate volume), central zone (25% of prostate volume), and transition zone (5% of prostate volume) (Figure 1)  [23], [24].

The peripheral zone is primarily located in the posterior of the prostate at midline and also comprises the apical horn and lateral edges. The peripheral zone is also

**Zones of the Prostate**
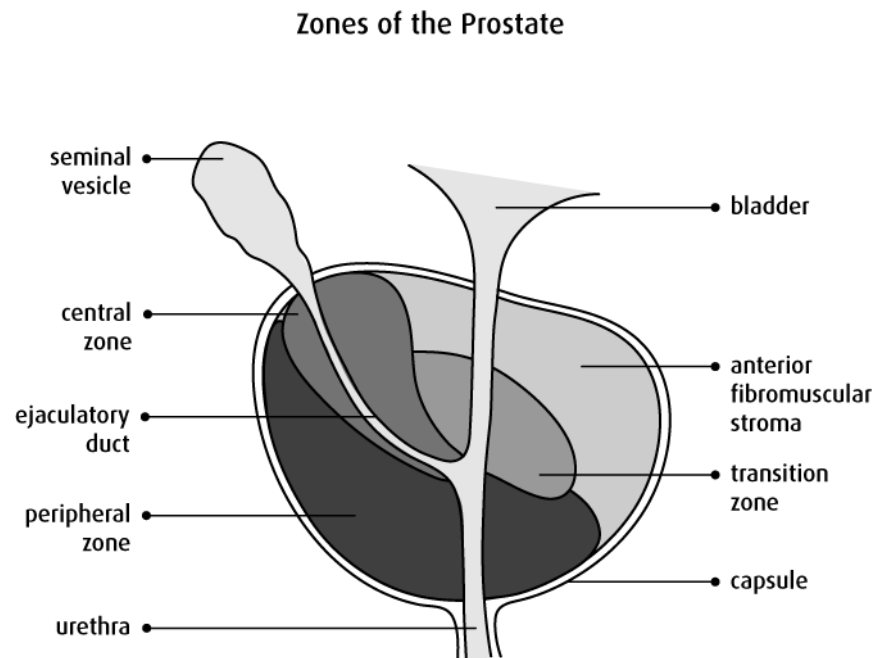


Figure 1: Sagittal view of Prostate Zonal Anatomy [226].

filled with prostate ducts with a width of approximately 135µm [25].The base of the prostate is connected to the base of the bladder and seminal vesicles, the urethra runs through the anterior base of the prostate at midline and exits the prostate at the apical horn in a curved fashion. The neurovascular bundles responsible for controlling urinary and sexual function are located at both lateral sides outside the apex of the prostate. The ejaculatory ducts run parallel to the posterior capsule at the base of the prostate at midline, and separate the peripheral zone from the central zone. Approximately 70% of cases of prostate cancer arise in the peripheral zone, 25% originates in the transition zone, and very few cancers originate in the central zone [23], [24].

## 2.1.2. Prostate Cancer

Prostate cancer (PCa) is a leading cause of cancer related morbidity and mortality for men in the United States with a projected 35,000 deaths and 300,000 new cases in 2024 [1]. While clinically significant, Gleason score 3+4 or higher, prostate cancer can be deadly for the patient, patients with low grade clinically insignificant prostate cancer, Gleason score 3+3 or lower, have less than a 2% risk of cancer specific mortality 10 years after diagnosis when properly managed [2], [3], [26]. Autopsy studies have shown that up to 40% of men older than age 50 will develop some level of microscopic prostate cancer but less than 20% will develop clinically significant prostate cancer in their lifetime [27]. Also referred to as Grade Group 1, early detection of clinically insignificant prostate cancer can lead to overtreatment of the disease, side effects of which are decreased quality of life for the patient as well as increased health care expenditures [3]. Common complications arise from damaged nerves during treatment, most notably incontinence and impotence which greatly diminish patient quality of life. As a result,

12

patients diagnosed with low risk prostate cancer are recommended to be enrolled in active surveillance of their disease, where the disease is monitored over a span of several years, and only treated once the cancer becomes clinically significant [2], [28].

## 2.1.3.   Histopathology of Prostate Cancer

The Gleason grade and type of prostate cancer are best determined through histopathological analysis from either biopsy cores or tissue specimen (Figure 2) [29]. Prostate cancer is defined as a malignant neoplasm or uncontrolled abnormal growth of cells in the prostate. The vast majority of these neoplasms are carcinomas and of epithelial origin [29].

The two main forms of prostate cancer are intraductal carcinoma (17-40% of radical prostatectomies) and adenocarcinomas (72% of radical prostatectomies) which fall under the World Health Organizations glandular neoplasms classification of prostate tumor morphology [29], [30], [31]. Intraductal carcinoma is characterized by epithelial cell proliferation that fills acini and prostate ducts and is often associated with solid or dense cribriform pattern. Linked to aggressive prostate cancer, it is often found alongside high-grade, high-stage adenocarcinoma [29], [32]. Adenocarcinoma is
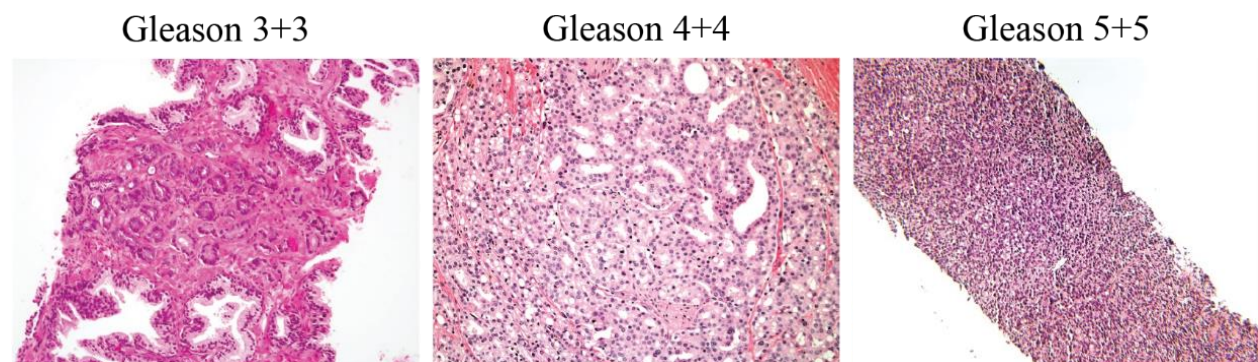


Figure 2: Gleason Grades of adenocarcinoma from biopsy cores. Gleason 3+3 (left), 4+4 (middle) and 5+5 (right). Gleason grades of 4 or higher are considered clinically significant cancer [29].

characterized by the disturbance of benign epithelial-stromal relationships. Gleason grade 1 and 2 glands are single, separate, and well formed, at Gleason grade 3 the glands become crowded and begin to infiltrate the stroma, however, the glands are still discrete and well-formed [29], [33]. High-grade Gleason grade 4 has poorly formed fused cribriform glands and at Gleason 5 cells form sheets of tumors, linear arrays, and solid nests of malignant cells with an overall lack of gland formation [29], [33]. As a result of malignant neoplasm many high grade clinically significant prostate cancers are characterized by dense tumors, a predominant reason why T2 weighted MRI imaging is a useful tool for delineation of tumors due to decreased signal from a lack of free-water content [34], [35]. Gleason score is often used interchangeably with grade group (GG), the latter released in 2014 by the International Society of Urological Pathology (ISUP) [33]. Grade group 1 is equivalent to Gleason score 3+3, as such GG ≥ 2 is often used to denote clinically significant cancer.

## 2.1.4. Index Lesion Theory

More than 60% of prostate cancers at the point of radical prostatectomy are multi-focal with multiple distinct genetically distinct cancer clones [36], [37], [38], [39], [40], [41]. However, metastatic cancer often responsible for patient death originates from a single precursor parent cell from a single precursor lesion [42]. This information is the basis behind the index lesion theory which states the largest and most aggressive tumor is responsible for disease progression [43], [44], [45]. Based on this theory, the primary focus for prostate cancer management should be to accurately diagnose and base treatment on the size, severity, and location of the index lesion [43]. Identifying and

grading the index lesion in patients is crucial for patient care, and heavily dependent on prostate imaging and biopsy.

## 2.1.5. Prostate Cancer Screening

The strongest risk factors for prostate cancer are old age, black ethnicity, and a positive family history [46]. In the United States, 90% of prostate cancer is detected by screening [46], [47]. From 1985 to 2007 the lifetime risk of diagnosis of prostate cancer nearly doubled from 9% to 16% due to the advent of prostate-specific antigen (PSA) screening [46], [48], [49]. Prior to the use of PSA testing a digital rectal exam (DRE) was the primary screening test for prostate cancer. However, this exam has considerable inter-examiner variability and the majority of DRE positive detections are advanced stage cancers [46], [50], [51]. A PSA test measures the level of PSA in the bloodstream, a protein produced by normal as well as malignant prostate cells [52]. However, most abnormal PSA values are false positives and approximately 15% of normal PSA values are false negatives [46]. Furthermore, it is estimated that 23-42% of PSA-detected cancers are over-diagnosed, leading to potential over-treatment of insignificant cancer [46], [53]. Results from 1,226 men have shown that the PSA values are more insightful when normalized for the patient's gland size also known as PSA density (PSAD), and in particular a low PSAD below $0.09ng/mL^2$ is protective against the patient having clinically significant prostate cancer [54]. Due to the uncertainty of PSA and DRE testing, it is difficult to accurately diagnose clinically significant prostate cancer [46]. As a result, positive screening results often warrant a follow up biopsy to allow for histopathological analysis [55].

## 2.1.6. Prostate Diagnosis

Following an elevated result during cancer screening, the traditional diagnostic pathway will typically employ a confirmatory ultrasound guided biopsy. When available, larger centers will also utilize multiparametric MRI to identify any suspicious regions for targeted biopsy.

Since the development of transrectal ultrasound (TRUS) in the 1970's, US has had a vital role in prostate gland visualization and guiding of systematic 12-core biopsies (Figure 3) [56], [57], [58]. Prior to the advent of PSA testing in the late 1980's, prostate cancer screening relied heavily on DRE's for diagnosis, resulting in cancers that were later stage and larger at time of biopsy than they are today, many of which were visible
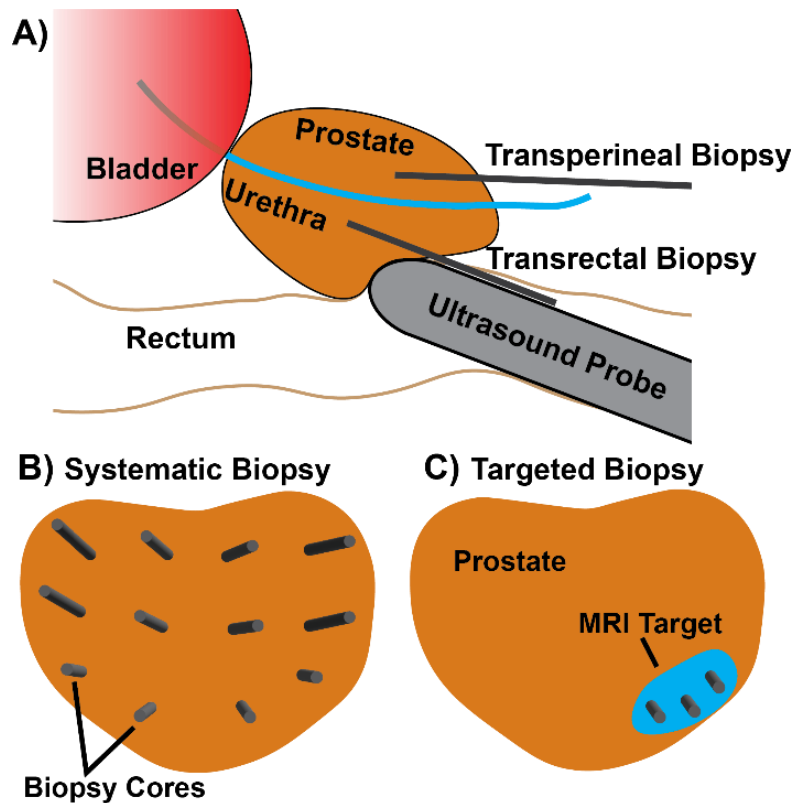


Figure 3: A) - Anatomic position of the prostate and the relative orientation of transrectal and transperineal biopsies; both are guided via transrectal ultrasound. B) - Illustration of a systematic biopsy. C) – Illustration of a targeted biopsy.

on US [5], [59]. Following the United States Food and Drug Administration (FDA) approval of PSA testing in 1986 prostate cancer is generally identified at earlier stages and smaller than it was in prior decades. Consequently, ultrasound has poor sensitivity and specificity for prostate cancer with biopsy results often identifying benign tissue in hypoechoic regions while cancer is found within otherwise inconspicuous isoechoic regions [5]. However, US scans are still vital for prostate cancer screening as they help guide biopsy cores for systematic sampling, and allow physicians to accurately measure the volume of the gland to calculate PSA density, a metric that is reported to have similar sensitivity, but superior specificity to PSA for predicting clinically significant prostate cancer [60], [61], [62], [63].

Due to US having limited diagnostic utility for visualizing PCa, if alternative imaging is not available, TRUS biopsies are typically performed following a 12-core template, known as a systematic biopsy (Figure 3B). A systematic biopsy samples the prostate at standardized locations to minimize the number of cores taken while retaining confidence for prostate cancer identification [4].

However, it has been shown targeted biopsies improve clinically significant cancer detection rates and as such are recommended for biopsy [10], [64]. Targeted biopsy seeks to maximize biopsy yield by acquiring samples from suspicious regions identified on diagnostic imaging (Figure 3C). A combination of these strategies is often used to identify prostate cancer during biopsy. The current recommended diagnostic strategy is a pre-biopsy, multi-parametric magnetic resonance imaging (mpMRI) scan which has demonstrated improved sensitivity (89%) and specificity (73%) over standard US for imaging prostate cancer [6], [8], [10], [64]. To fully utilize mpMRI imaging for prostate cancer detection, radiologist-identified suspicious lesions are fused to 3D ultrasound

17

scans to allow for real-time visualization during biopsy [12], [65]. While it is possible to do in-bore MRI-guided biopsy, these procedures are resource intensive, time consuming, complex, and often cost prohibitive, and do not offer a substantial diagnostic benefit over less expensive, and more time efficient, MRI-US fusion biopsies [11]. In general, use of US-guided MRI targeted prostate biopsies are superior to traditional systematic US-guided biopsy, with higher rates of identification of clinically significant prostate cancer and fewer diagnosis of clinically insignificant cancer [64].

## 2.2. Prostate Cancer Imaging

With advances in PSA screening early detection of prostate cancer has dramatically increased [46]. As a result, the importance of tumor grading and patient risk stratification has become vital to patient care to prioritize treatment for patients with high grade PCa and avoid treatment for low-grade PCa. Biopsy cores are the gold standard for grading prostate cancer [5], [29]. Imaging when used before or during prostate biopsy plays a critical role in prostate cancer diagnosis as well as treatment planning [5], [12], [66]. There are a number of different imaging modalities used for prostate cancer imaging including: ultrasound (US), magnetic resonance imaging, (MRI), computed tomography (CT) and positron emission tomography (PET)-CT, and radionuclide bone scanning [67]. In order to provide optimal care and avoid overtreatment, an ideal imaging modality should allow physicians to identify clinically significant cancer from clinically insignificant cancer [2], [3], [67].

## 2.2.1. Ultrasound

Conventional US is widely used clinically for visualizing the prostate during biopsy procedures and measuring the size of the prostate [67]. Performed either transrectally or transperineally, US without biopsy is about as accurate as DRE, with retrospective studies showing upwards of 40% of tumors are isoechoic under US [59], [67], [68], [69]. With earlier detection of prostate cancer due to PSA testing, the majority of lesions are smaller and earlier-stage than before [67]. In the 1980s, cancer was often diagnosed at a later stage resulting in US having higher sensitivity and specificity for prostate cancer, but today with many biopsy detected cancers are not visible on US and many hypoechoic areas do not appear to be malignant at biopsy [67]. As a result, US alone has poor sensitivity and specificity for the detection of prostate cancer without biopsy [67].

**Ultrasound Imaging**

US utilizes a transducer to generate images. An US transducer converts electrical energy into mechanical pressure waves based on the piezoelectric effect [70], [71]. The piezoelectric effect is the ability of certain materials to generate electrical charge in response to applied mechanical stress, or inversely the generation of stress when an electric field is applied, and this phenomenon is the basis for artificial sound production and detection [71], [72]. A material's piezoelectric properties are based on atomic scale polarization, and to be piezoelectric a material must be non-centrosymmetric: does not have points of inversion symmetry throughout its volume [71], [73]. US machines commonly utilize lead zirconate titanate (PZT) crystal elements either as a single element or as a broadband transducer with multiple elements [70]. Applying an electric field to these crystals in a controlled manner allows for the generation of pressure waves. The peak of the wave is where the pressure is higher than ambient pressure and, and the

valley of the wave is where the pressure is lower than ambient pressure [74]. Conventional US machines generate longitudinal mechanical pressure waves at a frequency of 2-15 MHz, much higher than the upper human auditory limit of 20 kHz [69], [74], [75]. Mechanical waves must pass through a medium. Longitudinal waves oscillate in the same direction of wave propagation. Transverse or shear waves are perpendicular to the longitudinal wave and are rapidly attenuated in tissue, providing no role in standard brightness-mode (B-mode) medical imaging [74].

US machines primarily employ a pulse-echo imaging with a B-mode display [74]. In pulse-echo imaging an initial US pressure wave, generated by the transducer, is transmitted through the imaging specimen. The power per unit cross-sectional area of the pulse is defined as the acoustic intensity [74]. The more concentrated or focused the US pulse, the higher the intensity. Acoustic intensity is directly correlated with bioeffects such as tissue destruction from heating due to exposure to US energy and is the basis for tissue death in High-Intensity Focused Ultrasound (HIFU) [76]. In standard imaging as the pulse (often referred to as the US beam) propagates through the tissue a small fraction reflects off of heterogeneous structures and part of this echo travels back to the transducer which detects the echo during the listening phase. The rest of the US beam travels deeper into tissue and reflections continue with each encountered heterogeneous structure (Figure 4). These echoes are detected and processed to generate a pixel map of the intensity of reflections detected by the transducer, giving a B-mode image. In B-mode imaging the stronger the reflection, the brighter the structure appears [74]. A complete image is obtained by many co-planar pulse-echo cycles, with the next US beam pulse beginning after all the all the echoes have been detected and processed
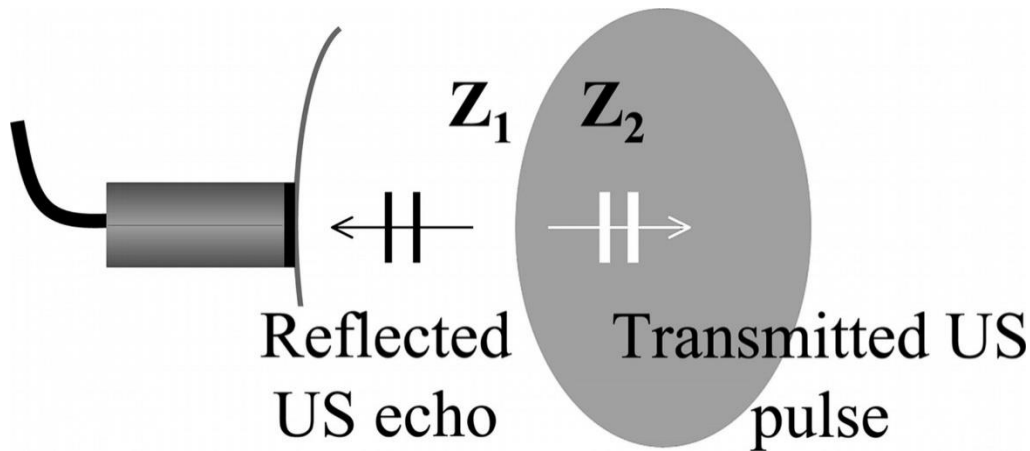
Figure 4: US beam imaging a heterogenous specimen. Part of the beam reflects off of the interface of $Z_1$ and $Z_2$ and part of the beam continues to travel into the specimen [74].

from the first pulse. This process of successive pulse-echoes is then repeated for the next frame, and imaging can generate up to 40 frames per second [74].

**Ultrasound Properties**

The intensity of the reflected US echo is heavily dependent on an intrinsic material property known as acoustic impedance (Z). The specific acoustic impedance of a medium ($Z$) is defined as:

$$Z = \rho c$$

$\rho$ is the tissue density and $c$ is the speed of sound in the medium [74]. A table of speeds of sound for biological tissue can be seen in below (Table 1). The average speed of sound in biological tissue is 1,540 m/sec and US scanners assume this speed to determine the depth at which echoes are produced [74]. The formula for depth ($D$) is:

$$D = \frac{t * c}{2}$$

$t$ is the time it takes from pulse generation to echo detection. The wavelength of the US beam ($\lambda$) is defined as a function of the speed of sound divided by the frequency ($f$)

Table 1: Speed of sound in various biologic mediums [74].

| Medium | Speed of Sound [m/sec] |
|---|---|
| Fat | 1450 |
| Amniotic Fluid | 1540 |
| Kidney | 1565 |
| Muscle | 1600 |
| Bone | 4080 |

$$\lambda = \frac{c}{f}$$

Ultrasound wavelengths in tissue range from 0.77mm at 2MHz to 0.10 mm at 15MHz [74]. The echo pressure amplitude of an US beam can vary by a factor of $10^5$ or greater so relative pressure and intensity are expressed in decibels (dB) [74]. The formulas for relative pressure and relative intensity are:

$$20 \log \frac{P_2}{P_1} \text{ and } 10 \log \frac{I_2}{I_1}$$

When an incident US pulse hits a large smooth interface between two different homogeneous mediums a partially reflected echo travels back to the transducer, and a partially transmitted pulse travels deeper into the imaging specimen (Figure 3). Known as specular reflection, the intensities of the reflected and transmitted pulses sum to the intensity of the original pulse, and for a normal (90°) incidence the intensity of the reflected pulse ($I_r$) is defined as:

$$I_r = I_0 \frac{(Z_2 - Z_1)^2}{(Z_2 + Z_1)^2}$$

Where $I_0$ is the initial beam intensity, $Z_1$ is the acoustic impedance of the first medium, and $Z_2$ is the acoustic impedance of the second medium. The ratio of $I_r/I_0$ is the intensity reflection coefficient between two mediums. As the acoustic impedance difference ($\Delta Z$) between two mediums increases the intensity of the reflected echo increases as well [74]. For example, the $\Delta Z$ between air and tissue is extremely high and can result in a lack of beam penetration, so to remedy this US gel is used which decreases the acoustic impedance difference allowing for better beam penetration.

The angel of incidence of a specular boundary is not always normal, and in this case the echo will not travel directly back to the transducer, and if the boundary is rough there will be a diffuse reflection through a wide variety of angles and only a portion of the beam will reflect back to the transducer resulting in a dimmer echo on the B-mode image. For nonperpendicular incidence of an US beam on two different media, the behavior of the transmitted beam behaves in accordance with Snell's law:

$$\frac{\sin \theta_2}{\sin \theta_1} = \frac{v_2}{v_1} = \frac{n_1}{n_2}$$

Where $\theta$ is the angle from the normal of the boundary, $v$ is the velocity of the US beam in the respective medium, and $n$ is the refractive index of the medium [74].

If the reflector with dimensions ($d$) is such that $d \ll \lambda$ the US beam will scatter and the beam will be reflected in a wide range of angles (Figure 5). Similar to diffuse reflection the echo intensity will be greatly reduced, but the echo intensity will be detected regardless off the incident angle. This speckle signal provides the visible texture in most organs due to the inherent heterogenous nature of tissue and the presence of many small scattering structures. The appearance of the speckle is dependent on the axial and lateral dimensions of the US pulse as well as the orientation
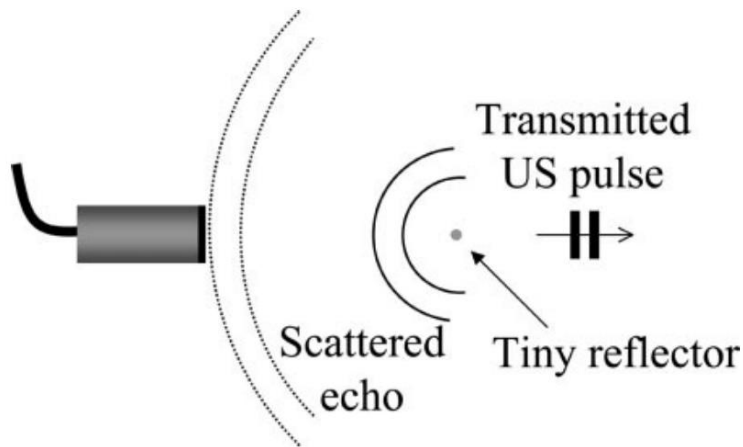
Figure 5: Specular reflections depend on the size of the reflector relative to the US wavelength resulting in scattering of the US signal [74].

of the of the speckle cells. This appearance will not change with time, but rather with changes of the transducer position and orientation [74]. As US resolution increases larger specular reflections will transition to diffuse reflections.

Another tenant of US imaging is that as US pulses travel through tissue their intensity is attenuated due to reflection and scattering, beam divergence, and friction-like losses (heat absorption). Absorption is the most important factor of US attenuation. Attenuation is measured in loss of intensity decibels per centimeter of tissue traveled per megahertz (dB/cm/MHz), as it is approximately proportional to the path length and US frequency. For soft tissue attenuation is generally between 0.3-0.8 dB/cm/MHz, and thus to image deeper into tissue lower US frequencies are required at the cost of imaging resolution [74].

**Ultrasound Pulses**

US pulses generated by a transducer are generally have a short spatial pulse length, the number of cycles in a pulse multiplied by the wavelength. Short pulse lengths generally have the best axial resolution. This is accomplished by electronically exciting the

piezoelectric elements for a very short time (1μsec or less). Axial resolution is constant as the pulse propagates to greater depths. Similarly, the narrower a pulse is in the lateral direction, the better the lateral resolution. This is achieved by focusing the transducer, which can be accomplished for single crystal elements by mechanical focusing. Examples of this include curving the transducer or using an acoustic lens. Regardless of mechanical focusing, all US beams have widths that (lateral resolution) that vary with distance from the transducer. The focal distance is the point of narrowest width, with the near field being closer than the focal distance, and the far field being further. Focusing can only occur in the near field. Higher US frequencies have narrower widths at the focal distance, giving improved lateral resolution, but this produces a trade-off since higher frequencies have increased attenuation leading to weaker signals at greater depths. Lastly side lobes are unwanted regions of US intensity that lie off the axis of the main beam. Side lobes are much weaker than the main US beam but are present for all transducers and have the ability to produce image artifacts[74].

**US Beam Scanning**

Most modern US probes consist of an array of transducers that are activated in subsets to scan the US beam across the imaging plane. These arrays often consist of around 128 to 196 elements. By activating a subset of elements, the US beam is centered over the activated elements, and changing the activated elements laterally the beam line also shifts laterally (Figure 6). Generally, for readout a larger number of elements are used to receive the returning echoes than those used to produce the US beam. This results in image acquisition across the entire field of view (FOV) in less than 1/10 of a second [74].
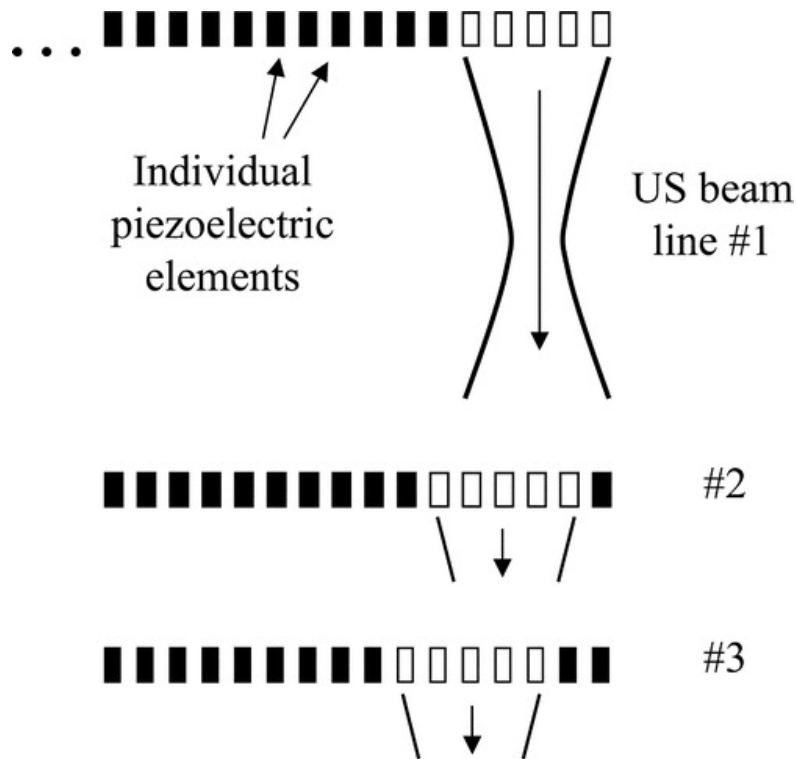
Figure 6: : Diagram of US scanning in a linear-array transducer [74].

By altering the shape of the array, such as a convex arc, also known as a curvilinear-array, FOV and beam shape can be changed. Furthermore, the timing of element activation to produce non perpendicular US beams. Timing offsets, at both beam formation and readout, can also be used to focus beams laterally for improved lateral resolution. Elevational resolution, also known as beam thickness, is commonly controlled through mechanical focusing, and like lateral resolution, varies with depth [69].

**US Echo Detection and Signal Processing**

After receiving an echo, it is amplified in several stages. The transducer immediately applies a uniform preamplification to the detected signal, and then a user-controlled gain is applied in addition. To account for attenuation in the image, a depth dependent

26

gain, also known as timer-gain compensation (TGC), is applied to the echoes as well, with deeper echoes receiving greater amplification [74].

A logarithmic transformation is then applied to the signal to decrease the dynamic range and allow for representation on an 8-bit display. Furthermore, oscillations at the US frequency are filtered out along with very small signals in an effort to improve the signal to noise ratio [74].

**Advanced Ultrasound Techniques**

Beyond standard B-mode imaging, there has been considerable research in contrast enhanced ultrasound (CEUS), elastography, and quantitative ultrasound, for visualization of prostate cancer [77], [78], [79], [80], [81]. Furthermore, these techniques can be applied to microUS imaging which can allow for further improvement in ultrasound imaging of prostate cancer as techniques are developed [78], [82].

CEUS utilizes gas filled microbubbles to increase US reflections for greater signal to noise ratio in vascular architecture through subharmonic imaging (signal at half the transmit frequency) isolation of the micro-bubble signature from surrounding tissue [78], [83]. Advances in CEUS imaging have allowed for detection of individual microbubbles allowing for microvascular visualization [78]. While CEUS is promising, subharmonic imaging relies on low amplitude signals and tissue perfusion, which is often low for prostate relative to other organs such as the liver. Beyond that it requires an injection of contrast during the biopsy for visualization increasing the complexity of the procedure. Meta-analysis of six CEUS studies report only a minor increase of 2-8% in cancer detection rates for CEUS targeted biopsies over systematic template biopsies [77]. For reference the same meta-analysis of 15 studies reported up to a 43% increase in cancer detection for mpMRI targeted biopsies over systematic biopsies [77].

Elastography has two modes: strain and shear wave elastography. Utilizing the increased density of PCa relative to surrounding tissue, elastography attempts to quantify tissue stiffness, overlaying elastography information on B-mode imaging [78]. Strain elastography requires tissue compression, then the transducer performs compression and decompression cycles to measure tissue stiffness [78]. Shear wave elastography measures shear wave speed related to elasticity or the Young's modulus, and requires the transducer to not compress the prostatic tissue to avoid incorrect measurements [78]. Both types of elastography are highly dependent on operator skill, furthermore, acoustic forces reaching the anterior of the prostate are often limited and compression from prior B-mode imaging can diminish shear wave elastography results [78]. Meta-analysis from five elastography studies found elastography increased cancer detection rates by 7-15% relative to systematic biopsy [77].

Quantitative ultrasound applies power-spectrum analysis on US radiofrequency (RF) signals to quantify the bio-acoustic properties of tissue. Backscatter normalization allows for elimination of instrument-setting variables on results [80]. Resulting parametric maps have demonstrated potential for prostate cancer visualization when evaluated against ground truth WM pathology [80]. However, quantitative US requires substantial processing to create parametric maps limiting real-time application.

## 2.2.2. Multi-parametric Magnetic Resonance Imaging

MRI, specifically, multiparametric MRI (mpMRI) (Figure 7), has been in use for prostate cancer imaging since the 1980's [6]. In contrast to US, mpMRI has high sensitivity (86-92%) and specificity (60-83%) for identification of prostate cancer index lesions [6], [7], [8]. However, mpMRI is resource intensive, has high inter-reader variability, can still fail
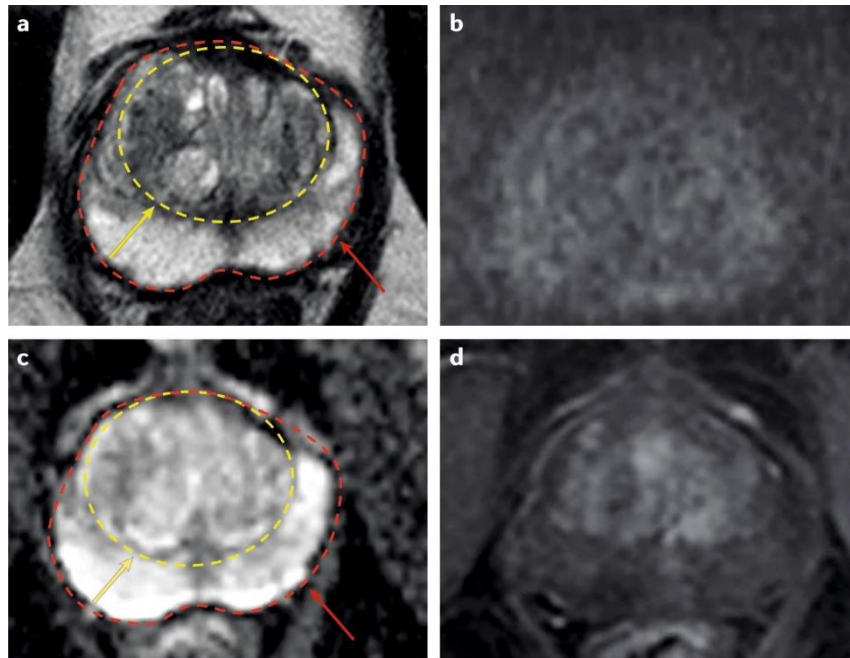
Figure 7: Example subset of sequences comprising an mpMRI prostate scan. The images are all from the same location in the prostate and the specific sequences are: a) T2 b) DWI c) ADC d) DCEI [6].

to detect 10-15% of tumors, and routinely underestimates tumor volume by as much as 1/3 [7], [13], [14], [15]. Originally focusing on only T1-weighted and T2-weighted sequences, mpMRI imaging for prostate cancer have expanded to include diffusion-weighted imaging (DWI), dynamic-contrast enhanced imaging (DCEI), and, when available, magnetic resonance spectroscopy imaging (MRSI) providing a variety of different contrast mechanisms to improve diagnostic accuracy of prostate cancer [6].

MRI is based on manipulating the alignment of atomic nuclei of water molecules with an external magnetic field to produce an external radio frequency (RF) signal. By aligning the nuclei and measuring the RF signal upon relaxation at various points various emitted signals can be measured [84].

**T1-Weighted Imaging**

T1-weighted imaging uses longitudinal relaxation of protons for contrast and is obtained from short repetition time (TR) and short time from RF pulse to echo signal (TE) [84].

29

T1-weighed imaging provides good contrast for identifying lymph nodes and bone structures, and in the context of prostate cancer, is useful for detecting hemorrhaging that can obscure or mimic cancers [6], [85], [86]. However, prostate cancer does not have any noticeable changes on T1-weighted imaging [6], [87].

**T2-Weighted Imaging**

T2-weighted imaging uses transverse relaxation time for contrast and is obtained from longer TE and TR times relative to T1 imaging [84]. T2-weighted imaging highlights the water content of tissue which is directly related to cellularity [6], [87]. For prostate imaging T2 imaging provides a highly defined anatomical image with distinguishable zonal architecture and excellent soft-tissue contrast [6], [88]. In healthy prostate tissue, the peripheral zone appears homogenous and hyperintense on T2 imaging due to high glandular ductal tissue content [6], [87]. Prostate cancer is characterized by high cellularity and low water constant due to a disruption of the normal ductal pattern, resulting in hypointense regions on T2-weighted imaging [6], [87]. The transition zone, however, is heterogeneous and given every prostate is morphologically unique, cancer detection in this zone on T2 imaging is challenging [6], [86]. Furthermore, other changes may mimic prostate cancer on T2 imaging, such as prostatitis, scars, irradiation, hormone treatment, and post-biopsy hemorrhaging [6], [85]. As a result, T2 imaging is great at identifying suspicious areas, and the sensitivity of detection is improved with the use of DWI [6].

**Diffusion-Weighted Imaging**

DWI imaging quantifies the random movement of water molecules within tissue, which is generally high in nonmalignant tissue, but lower in cancerous tissue due to increased epithelium and higher cellularity [6], [85]. The apparent diffusion coefficient (ADC)

reflects the ability of water to move and is mapped by performing DWI at multiple magnetic field strengths [6]. On DWI suspicious areas appear as a bright spot surrounded by low signal tissue, and conversely on an ADC map prostate cancer appears as a low signal area with a lower signal for increasing Gleason scores [6], [89]. DWI with T2 imaging has been proven to improve sensitivity (0.76) and specificity (0.82) over T2 imaging alone as well as improve characterization of transition zone tumors [6], [90], [91].

**Dynamic Contrast-Enhanced Imaging**

DCEI as the name suggests relies on gadolinium-based contrast, and is generated by rapid acquisition of a series of T1-weighted images following intravenous injection. The goal of DCEI is to evaluate tumor angiogenesis based on the velocity and intensity of contrast agent uptake of tissue[6], [92]. Areas of increased contrast intensity are associated with prostate cancer, however other benign conditions can also be hyperintense, resulting in a sensitivity range of 46-90% and specificity range of 74-96% for DCEI alone [6], [92]. DCEI is particularly valuable when T2 imaging and DWI are uncertain or degraded by artifacts, and in instances of local recurrence after intervention such as resection and focal therapy where prostate morphology is altered [6], [86], [93], [94].

**Magnetic Resonance Spectroscopy Imaging**

MRSI sequences highlight the expression of specific metabolites such as citrate and choline [6], [95]. Citrate is produced by nonmalignant prostatic tissue but its production is reduced in cancerous tissue, while conversely choline, a cell membrane component, levels are high in cancerous tissue but low in non-malignant tissue [6], [95]. MRSI alone has a sensitivity of 75-89% and a specificity of 77-91%, but is currently limited to use in

large academic centers due to related costs, availability, and limited research supporting extensive use [96], [97].

**MRI Interpretation**

Interpreting a mpMRI scan can be subject to significant interobserver variability, and a big challenge has been standardizing scoring between observers [6], [15]. Initially, a subjective score was assigned based on the radiologist's subjective opinion, known as the Likert score [6], [98]. Due to the high rate of inter-observer variability, the Prostate Imaging Reporting and Data System (PI-RADS) v1 was developed in 2012, applying rigid criteria to specific scores [6], [87]. This system assigned a score for each sequence (T2, DWI, DCEI, and MRSI when available) and a total score based on these scores, however it was a complex system, time-consuming, and had poor reducibility [6]. To fix these issues PI-RADS v2 was published in 2014 [86]. The scoring system is continually being optimized with new guidelines being established with the most current version being PI-RADS v2.1. As a pre-biopsy scan PI-RADS is designed to prioritize specificity to potentially avoid unnecessary biopsies while also maintaining high sensitivity. Consequently, mpMRI identified suspicious regions can miss 10-15% of tumors and routinely underestimate tumor volumes by as much as a 1/3 [7], [13], [14].

## 2.2.3. PSMA PET-CT

Prostate-specific membrane antigen (PSMA) is overly expressed in many PCa cells allowing for PET imaging ligands [99]. The FDA has approved a gallium based radiotracer for prostate cancer imaging [99]. In PET-CT imaging a radioactive tracer is created in a cyclotron and injected into a patient prior to scanning to circulate in the body. The unstable radioactive material will decay after some time and emit a positron. The

positron scatters in tissue losing energy before colliding with an electron and annihilating in two antiparallel photons [100]. These paired photons are subsequently detected by the PET scanner and used to localize the radiotracer in a CT image. The amount of distance the positron travels before annihilating is based on the energy of the positron and directly affects the spatial resolution of the signal relative to the true location of the tracer decay [100]. The gallium-based radiotracer for PSMA PET has a maximum path length of approximately 9mm, and with scattering the misposition error is roughly 1.6mm on average [100], [101]. Through fusion with MRI or region-based mapping of PSMA PET targets it is possible biopsy identified hotspots for PCa.

PSMA PET-CT has demonstrated similar performance to mpMRI for prostate cancer detection when evaluated against ground-truth whole-mount (WM) pathology [102]. Furthermore, the use of both PSMA PET-CT and mpMRI allows for improved detection and delineation of tumors [102], [103]. While promising, PSMA PET-CT is similarly resource intensive as a cyclotron is required to manufacture radiotracers with relatively short half-lives, ultimately limited it to larger centers. Furthermore, the radioactive tracer causes potentially harmful radiation resulting in longer stays for patients before and after scans.

## 2.3.  MRI-US Fusion

In the last decade, 3D ultrasound fusion devices have revolutionized prostate biopsies, allowing for precision targeting of suspicious lesions and improved detection rates of clinically significant cancer [104]. The following section covers in detail the various aspects of MRI-US fusion biopsies. This is particularly relevant to the work in this thesis as MRI-US fusion biopsies are the gold standard for diagnosis, and if microUS eventually

replaces MRI as a diagnostic, the need for MRI-US fusion and the associated drawbacks would ultimately be alleviated. Furthermore, this section also introduces topics of fusion or co-registration and target registration error, two topics that are vital for the work presented in this thesis.

## 2.3.1.  Ultrasound Tracking

US volume acquisition is typically necessary to perform fusion with MRI. Generally, two types of US probes are used for acquiring US volumes (Figure 8):

1) One-dimensional array transducers with position tracking: these are traditional US probes that create standard two-dimensional images. The depth of the reflection of the US beam is recorded along the dimension of the transducer, providing one-dimensional positional data with a time component, resulting in a beam-formed 2D image. The location and orientation of the probe in 3D space is tracked while the 2D images are collected to allow for a 3D reconstruction of the imaging sweep. Multiple arrays can be embedded within a probe to image different planes simultaneously. Embedded motors may robotically move the array or probe to generate a smooth 3D reconstruction.

2) Two-dimensional array (matrix) transducers: these are more advanced US transducers that create native three-dimensional images. The depth of the reflection of US beam is recorded along the transducer matrix, resulting in two-dimensional positional data with a time component, which provides a 3D image.

The primary advantage of array transducers is the ability to image the entire 3D volume without moving the probe. In contrast to 1D transducers, this enables biopsy needle insertion under direct visualization without having to navigate the probe to the
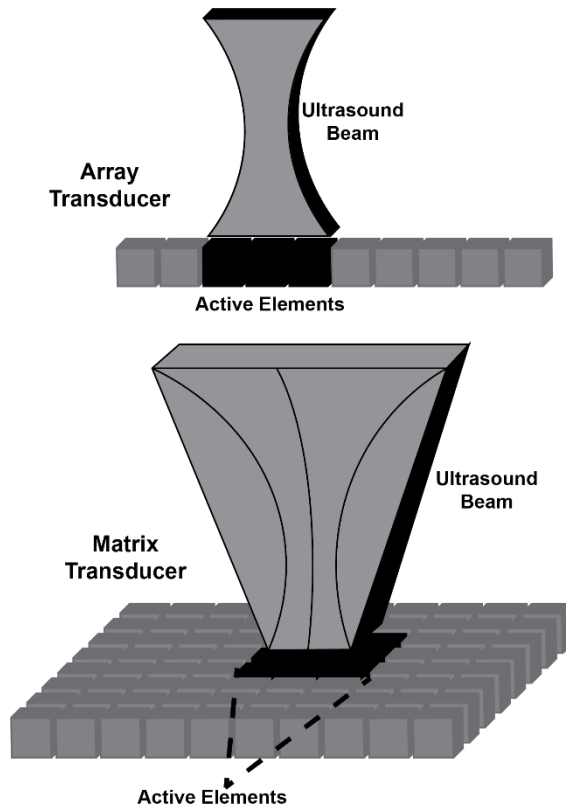
Figure 8: One-dimensional array transducers (top) create US beams from a subset of elements in the array, and the resulting echo is read by the full array to create a 2D planar image. Two-dimensional array (matrix) transducers (bottom) also create US beams from a subset of elements and read the echo from elements in a two-dimensional area to create a 3D volumetric image.

desired biopsy site. This simplifies user interaction and also improves registration with MRI as each time the probe is moved, the prostate undergoes new deformation which needs to be accounted for. Imaging is acquired transrectally, therefore, minimizing probe size is of paramount importance for patient comfort. However, 2D transducer arrays have not been commercialized for prostate biopsy. The cost and complexity of 2D probe design is prohibitive given that one-dimensional probes already achieve adequate levels of co-registration with MRI [65], [105], [106], [107], [108]. Today, all commercial 3D ultrasound systems for prostate biopsy rely on 1D array transducers that employ various tracking techniques during the interrogation of the prostate (Table 2). The first 3D ultrasound device for prostate biopsy to receive regulatory approval

35

incorporated a 1D array with a mechanical fixed arm, receiving clearance in 2008 (Artemis, Table 2). Many other platforms were subsequently brought to market using a variety of approaches to track probe movement as described below.

**Integrated Encoder**

Integrated encoders (Figure 9A) are one of the simplest methods for monitoring US probe position. Encoders built into the casing of the probe monitor the roll, pitch, and yaw while scanning the prostate [109]. Probe translation can also be monitored with an inertial measurement unit. While user-friendly and able to provide general correlation with MRI planes, this tracking mechanism is not as rigorous other approaches and is not commonly used for fusion biopsies.

Table 2: List of FDA approved commercial 3D Ultrasound machines for prostate biopsy and their respective tracking mechanism.

| Device Name | Manufacturer | Approval | Tracking Method |
|---|---|---|---|
| Artemis | Eigen | 2008 - FDA | Mechanical Arm |
| Trinity/Urostation | Koelis | 2010 - FDA | Mechanical Scanning |
| Navigo | UC-Care | 2011 - FDA | Electromagnetic |
| iSR'obot Mona Lisa | BioBot | 2011 - FDA | Mechanical Arm |
| BioJet | Medical Targeting Technologies | 2012 - FDA | Mechanical Arm |
| UroNav | Philips | 2012 - FDA | Electromagnetic |
| Virtual Navigator | Esaote | 2014 - FDA | Electromagnetic |
| Bk3000 | BK Ultrasound/MIM | 2016 - FDA | Electromagnetic |
| Fusion Bx | Focal Healthcare | 2016 - FDA | Mechanical Arm |
| ExactVu | Exact Imaging | 2016 - FDA | Integrated Encoder |
| BiopSee | MedCom | 2017 - FDA | Mechanical Arm |
| SmartTarget | SmartTarget Ltd. | 2017 - FDA | Mechanical Arm |
| RS85 Prestige | Samsung | 2018 - FDA | Electromagnetic |
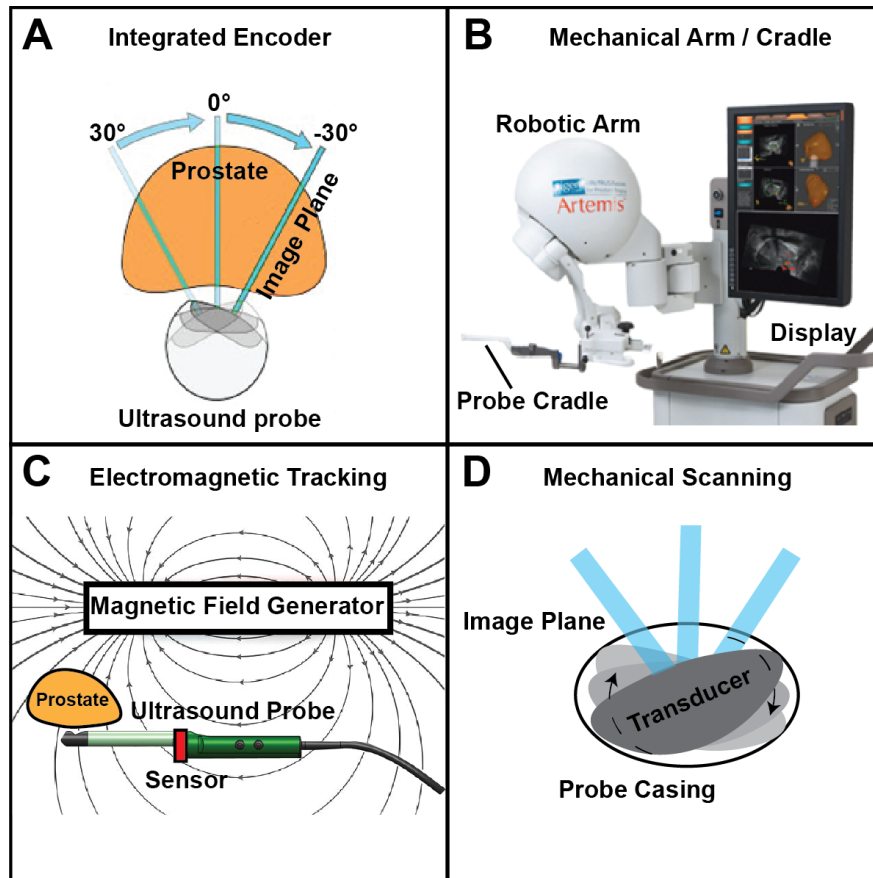
Figure 9: Tracking mechanisms for creating 3D US volumes. A - Integrated encoders use built in encoders in the probe casing to track probe movement. B - Mechanical arm or cradle systems stabilize and record the location of an US probe in a cradle using a jointed mechanical arm containing positional encoders. C - Electromagnetic tracking records the position of a US probe in a time-varying magnetic field. D - Mechanical scanning: the transducer array is automatically rotated inside the probe inside a casing to perform the sweep of the prostate while recording orientation.

## Mechanical Arm or Cradle

In contrast, to integrated encoders, mechanical arms or cradles offer improved tracking and allow for the full 6 degrees of freedom during scanning. Here, the US probe is mounted in a cradle connected to a jointed mechanical arm (Figure 9B) containing positional encoders which provide precise, real-time tracking in 3D space [107], [108]. This positional information along with the US image data is used to create highly

accurate 3D US volumes. This ensures that existing MRI defined targets can be accurately fused with the real-time US scan. Additionally, the exact biopsy site can be stored allowing re-biopsy in a subsequent session with an error of a few millimeters [110]. Although the mechanical arm removes movement error and improves tracking accuracy, the large form factor of these machines and unconventional fixed axis for scanning increase the learning curve and space required for TRUS prostate biopsies [107].

**Electromagnetic Tracking**

Electromagnetic tracking (Figure 9C), based on Faraday's Law, generates current in a solenoid sensor based on its location in a time-varying magnetic field [107], [108]. Three sensors, oriented in three orthogonal directions, are mounted on an US probe and measure the strength of a time-varying magnetic field via the generated voltage in each sensor [108]. From the sensor values it is possible to track the position and orientation of the probe relative to the magnetic field generator [107], [108]. With this tracked positional information, it is possible to generate 3D reconstructions of the prostate and track biopsy cores, while allowing the operator to maneuver the probe unrestricted [107]. However, this approach can be prone to distortion and reduced tracking accuracy due to electromagnetic interference and metallic objects [108]. Additionally, without stabilization, errors may arise during biopsy needle deployment that reduce targeting accuracy relative to mechanical arm approaches [107]. Electromagnetic tracking has also been used to track biopsy needles independent of probe tracking, removing the need for a mounted guide for tracking core location [111].

**Mechanical Scanning**

Mechanical scanning (Figure 9D) probes utilize mechanized transducers housed within the probe casing [108]. The location of the transducer is electronically rotated inside the

probe casing to interrogate the volume of the prostate [108]. Based on the angular position of the transducer, these systems may recreate a 3D volume, though are influenced by precession during rotation. These systems do not require the operator to move the probe after positioning, as the mechanized transducer is able to rotate the image plane to the desired location within the prostate [112]. The accuracy of this approach can be degraded by the motion of the patient or the US probe, however, it is possible to update the 3D scan following deformation or substantial misalignment through image-based tracking [107], [108]. Furthermore, through sophisticated registration methods and organ tracking (image based tracking of capsule deformation during biopsy) it is possible to improve prediction accuracy of the location of a biopsy core prior to biopsy [113].

**Alternative Tracking Approaches**

Apart from the approaches in clinical use listed in Table 2, two other tracking techniques are in development – optical tracking and sensor-less tracking. In optical tracking, a visual target is mounted to the handle of the US probe and tracked by a camera [114], which calculates the location of the US transducer. This technique is limited in clinical use due to difficulties maintaining line-of-sight [30]. In sensor-less tracking, 3D US volumes are generated from array transducers using image data alone through software. Several groups have been able to achieve a drift error of ~5 mm using speckle decorrelation algorithms combined with Gaussian filtering to estimate the elevational distance between neighboring US frames [115], [116], [117], though in *ex vivo* setting. *In vivo*, drift errors of 10 mm were achieved using a type of deep learning network, a deep contextual-contrastive network (DC$^2$-Net) with a 3D ResNeXt backbone structure, that focused on areas of strong speckle patterns. This model also managed to achieve dice

similarity scores of 0.89 [115]. With continued active research into deep learning, sensor-less 3D US reconstruction is a promising approach to obviate errors associated with motion and deformation.

## 2.3.2. Volume Reconstruction

Following acquisition of a series of 2D US images using one of the various tracking techniques, the volume of the scan needs to be reconstructed before segmentation and fusion with MRI. There are a variety of techniques that can be employed to reconstruct the 3D volume, but all methods use the tracked location of the US plane for each image and the B-mode intensity data as inputs for reconstruction. Prostate volumetric correlation of MRI and US volumes, against excised whole gland measurements can vary (US: 0.70-0.90, MRI: 0.8-0.96) [62], [63], but in general it is possible to achieve accurate targeting with a ±20% variation [110]. There are three general subcategories used for reconstruction of 3D US volumes: voxel-based algorithms, pixel-based algorithms, and function-based algorithms (Figure 10) [108], [118].

**Voxel Grid Selection**

In all three techniques, prior to reconstruction, a 3D grid is established across the data volume to determine sampling density, called the voxel grid [118]. An appropriate voxel grid has two main components: a sufficient amount of information for the defined volume to preserve information present in the original images; and a homogenous information density throughout the whole volume to reduce the amount of interpolated voxels [118], [119]. If voxel size and spacing in the 3D grid is too large, there will be a loss of information and resolution (aliasing) from the original 2D images, and if the voxel size and spacing is too small, there will be a large number of voxels that need to be interpolated in the 3D sweep inducing unnecessary error into the 3D reconstruction.

40

One approach for selecting an appropriate voxel grid is to first perform a principal component analysis (PCA) to select an appropriate coordinate system based on the pixel distribution in 3D space, and then calculating the appropriate voxel size based on probe resolution and the rotation of 2D image planes in 3D space as a trigonometric minimization function [119]. Another approach involves subdivision of a tetrahedral grid that is specifically adapted to match the complexity of the dataset, resulting in efficient and compressed volume grids [118], [120].

**Voxel-Based Algorithms**

Voxel-based reconstruction algorithms assign every voxel in a pre-defined volume a value based on nearby pixel information. Depending on the specific technique a voxel can be assigned an intensity value based on one pixel or many [108]. A general approach for voxel-based reconstruction is as follows. First, a voxel grid is established for a desired volume within the US sweep. Next, for each voxel, relevant nearby US image frames are selected based on their proximity to the voxel in 3D space. Within those frames relevant pixel information is selected along with associated weights depending on distance from the voxel. Then, an intensity value is assigned to the voxel based on the selected surrounding pixels and their corresponding weights. For one-pixel approaches, each voxel is generally assigned the value of the nearest pixel, while for multi-pixel approaches, the value of a voxel is normally set to an interpolated value depending on the nearby pixels from multiple frames [121], [122]. For multi-pixel approaches a weight factor is typically applied to the pixel values prior to interpolation that is inversely proportional to the pixels distance from the specific voxel [108], [122]. As a result, the closer pixel values have stronger influence on the voxel's intensity value. Since voxel-

based algorithms loop through each voxel and assign an intensity value; there are generally no gaps in the resulting reconstruction.

**Pixel-Based Algorithms**

Pixel-based reconstruction algorithms assign the value of each pixel in the series of images to one or more voxels [108]. A general pixel-based reconstruction approach would be as follows: a voxel grid is created around the imaging sweep. The algorithm then loops through each frame and assigns the pixel values from each frame to the voxels that intersect the corresponding pixel. Pixel distribution to various voxels can be based on a single pixel, an average of pixels, or an interpolated vale from a surrounding group of pixels. A common approach is to take a local kernel of pixels and apply various weights to each pixel value to populate the surrounding voxels [108], [123]. Depending on the scan distribution and voxel size, there may be unfilled voxels after pixel distribution due to voxels located away from any nearby frames. As such, a gap-filling or interpolation step is used to fill the remaining gaps in the data. Following pixel distribution and assignment to various voxels, a follow up interpolation step based on nearby pixel values is applied to fill in the empty voxels, however sometimes it is beneficial to leave empty voxels unfilled to let the reviewer know there is missing data [108].
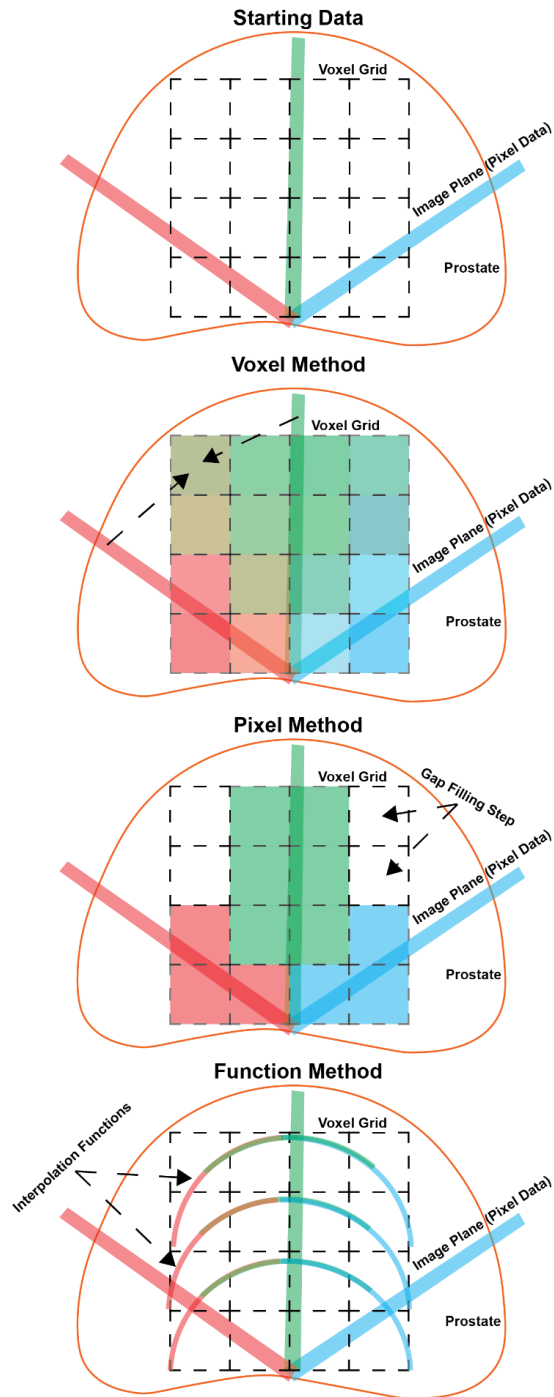
Figure 10: Example of various volume reconstruction techniques using three images at different orientations (red, green, blue). Voxel techniques assign a value to each voxel often based on weighted interpolation from nearby pixels. Pixel methods assign pixel values to nearby voxels and can have empty voxels based on scan distribution. Function based methods fit a curve to the pixel data to populate voxels.

**Function-Based Algorithms**

Function-based reconstruction algorithms utilize function derived interpolation to populate voxel values. By taking small windows of nearby frames at known distances apart, a function can be fit to pixel values within these frames to interpolate to voxel values between frames. Commonly used functions are polynomials, splines, radial basis functions, or Bézier curves [108], [118]. By overlapping the series of neighboring scans for each section of interpolation, it is possible to further improve reconstruction accuracy through distance weighted averaging of shared voxels [108].

**Computation time vs Reconstruction Quality**

In order to choose an adequate reconstruction algorithm, balance between reconstruction speed and quality must be considered [108]. The more sophisticated algorithms, such as spline-based approaches, or those that create local pixel kernels to approximate voxel values, generally allow for improved reconstruction accuracy at the cost of longer computation times. However, some approaches such as Bézier curve interpolation manages to achieve extremely fast reconstruction times (0.008 sec/image) while maintaining a high level of reconstruction accuracy [108], [124]. In general, advances in computer computational power and GPU performance have improved the overall speed of reconstruction, allowing for near real-time volume reconstruction, allowing for the use of 3D US in the clinic setting during biopsy [108].

## 2.3.3. Segmentation

In addition to 3D reconstruction, before fusion with MRI can occur, a segmentation step is typically implemented. The specific process varies from device to device, but in general the operator is asked to mark points around the prostate capsule and some fiducial points visible on both MR and US [65]. Examples of fiducials include: BPH

nodules, calcifications, the urethra, seminal vesicles, or the boundary between the peripheral and transition zones. Typically, the user is asked to mark the prostate capsule on a subset of slides and an interpolative algorithm fills in the remaining slides to create the prostate capsule volume [125]. Prior to biopsy the MRI is also segmented by a radiologist or experienced reviewer, these segmentations of the prostate capsule and any included fiducials are used as the basis for registration between MRI and US [65], [110]. There are a number of different prostate segmentation methods, as well as research into semi and fully-automated segmentation algorithms to speed up and improve registration accuracy. Manual segmentation tends to be time consuming and prone to variability, while automated segmentation is difficult due to inherently low signal-to-noise and contrast-to-noise ratios of US imaging [125], [126]. A comprehensive review paper, published in 2022 by *Jiang et. al*, describes the popular and developing techniques for prostate segmentation; a brief summary is included below [125].

**Edge-Based Algorithms**

Edge-based algorithms employ an edge identification step followed by a point ordering step to create an outline of the prostate capsule [125]. For edge detection, a filtering process is applied to identify the boundaries of the prostate. This can be done via conventional filtering, probabilistic filtering, or through more complicated filtering processes such as super-imposing the original image over an inverted multi-resolution filtered image, a method known as Radial Basi-Relief [127]. In general, edge-based filtering algorithms need to be finely tuned to allow for adequate edge detection, and tend to be very sensitive to image noise and artifacts [125]. As such they are commonly used alongside other filtering algorithms to achieve adequate segmentation, in both fully and semi-automated approaches.

**Initialization Points**
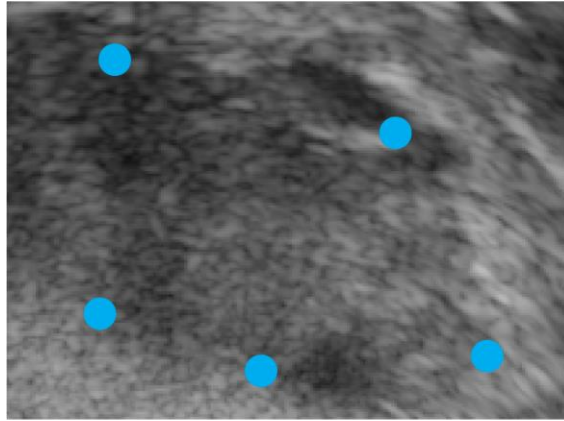


**Region Expansion**
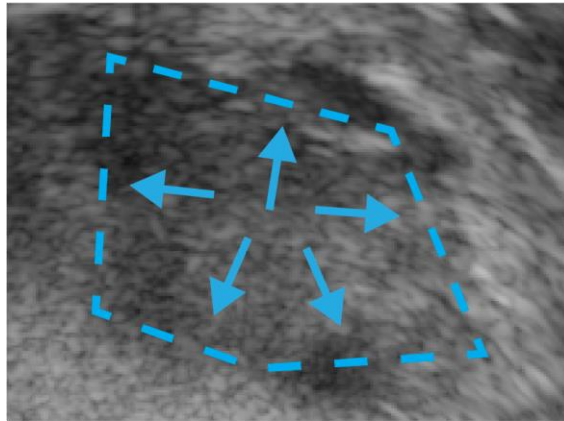


**Final Segmentation**
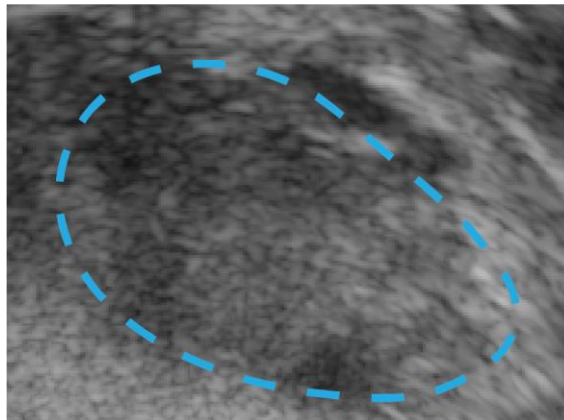


Figure 11: Region based segmentation approach. After a urologist or assistant places initialization points a region is established that is then expanded to find the boundaries of the prostate.

46

**Region-Based Algorithms**

Region-based algorithms utilize both regional and edge information to segment the prostate based on common areas (Figure 11). This can be achieved through regional clustering and statistical analysis [125]. Regional clustering invokes a parameter to divide an image into regions to find the boundaries of the prostate. Common parameters include the greyscale value of pixels, or more robustly, a parameter known as the Local Binary Pattern, a measure of a pixels value relative to its neighborhood and a metric for localized image contrast [125], [128]. By applying adaptive thresholding, optimizing for inter-class variance, and accounting for density and Euclidean distance of points (Density-based spatial clustering of applications with noise, DBSCAN) it is possible to identify the boundaries of the prostate [125], [129]. Statistical analysis region identification uses region-based statistics, such as intensities, averages, and standard deviations, to identify similar areas within images through minimization techniques. This approach can be assisted through the use of *a priori* knowledge of greyscale distributions and pixel intensity gradients of various regions [125], [130].

**Deformable Model-Based Algorithms**

Deformable model-based algorithms implement a template prostate model that is modified based on image data and energy minimization functions while balancing internal and external forces on the model that maintain contour smoothness [125], [131]. Three different frameworks are generally used: the active contour model framework (also known as snake or energy minimization curve framework), edge level set framework, and curve fitting framework. The active contour model is a spline curve created from input points that is deformed based on image gradient data in a specified region of interest, characterized by an energy minimization function that balances

preservation of the original shape's smooth contours with deformation to match the input image data [132]. This approach is dependent on proper assignment of internal and external forces to drive and control the deformation [125]. Edge level set framework utilizes level set implicit functions as desired boundaries, applying finite differences to approximate the solution of a partial differential equation [133]. The level set propagates from the seed point, perpendicular to the curve until it reaches the maximum of the intensity gradient. This approach is characterized by good signal to noise ratio and well defined edges, thus is commonly used [125], [130]. Curve fitting framework approximates the prostate boundary through fitting ellipsoids or splines. These ellipsoids or splines are initialized from a series of user selected points, and are then transformed using varying combinations of thin plate splines, Bayesian frameworks, *a priori* knowledge, and a variety of other optimization techniques [125]. Without using *a priori* knowledge, deformable model segmentations may differ significantly from the actual prostate boundaries reducing segmentation accuracy [134].

**Statistical Model-Based Algorithms**

Statistical model algorithms utilize big data and computational power to delineate the prostate. Utilizing a database of prostate shape features and the associated B-mode intensity data an optimal segmentation is selected [125]. One such approach is a Point Distribution Model (PDM) selection of an Active Shape Model (ASM) framework is performed, followed by a Principal Component Analysis (PCA) of the PDM with a generalized Procrustes analysis to allow for model variation. Building on this approach, the Active Appearance Model (AAM) algorithm framework was developed, utilizing the shape information of the ASM framework in combination with greyscale information to generate texture information for improved segmentation results.

**Learning Model-Based Algorithms**

Learning Model algorithms generally utilize either support vector machines (SVM) or neural networks to train a model to delineate the prostate boundary [125]. SVM are supervised learning models that attempts to create a hyper-plane between different types of data, such as the prostate capsule and the surrounding tissue, with an optimal compromise between model complexity and training results to avoid overfitting and promote generalization. A general SVM approach utilizes Gabor filters applied to the image data to extract texture features which are then input into a SVM for classification. This classification allows for a statistical shape model which automatically delineates the prostate [135], [136]. SVM approaches tend to work well with limited nonlinear analytical data [125]. Neural networks, designed to mimic brain neurons, are learning algorithms capable of identifying complex multivariate relationships, and generally outperform linear regression techniques when working with patient data. In particular, U-Net convolutional neural networks (CNN) perform particularly well at medical image segmentation tasks, and this holds true for prostate delineation [125], [137], [138]. Learning based approaches often experience degraded performance in the far apex and base of the prostate due to poor contrast to noise ratios and increased variability in prostate shape. Recent research has managed to overcome this hurdle through radially re-slicing a diverse set of 3D prostate volumes about the center axis into 2D images. In doing so, each image slice passes through the approximate center of the prostate, avoiding images only containing the far apex or base. These images are then subsequently segmented by a modified 2D U-Net model and reconstructed back into a 3D US sweep with a dice similarity coefficient of around 0.95 and a processing speed of

less than 0.7 seconds per prostate, outperforming 3D V-Net segmentation networks [139], [140].

## 2.3.4. Rendering

Following reconstruction and typically after segmentation, rendering is performed to relay 3D US information to the physician performing the biopsy. Given the large volume of data acquired in a 3D US scan and the time constraints associated with use in the clinic, it is important to consider the optimal approach to render the data [108]. Displaying too much information can increase rendering times and consume resources, while on the other hand, rendering not enough information can diminish the benefit of 3D US leaving out information that could be potentially helpful to the physician. In general there are three main rendering approaches: slice projection, surface rendering, and volume rendering [108], [141].

Slice projection (Figure 12A) displays a 2D slice of the 3D US data, either in the native imaging plane or in a reconstructed plane. The operator can typically scroll through the rendered image plane to visualize various regions of the prostate. However, there can be an increased cognitive load for the physician in this approach as they are required to mentally correlate the location of the image plane in 3D space [108], [142].

Surface rendering (Figure 12B) is a fast-rendering approach that only displays a 3D model of the prostate capsule. This approach is dependent on a segmentation step prior to rendering in order to identify the capsule boundaries [108]. Visualizing the shape of the prostate during biopsy allows for simple real time visualization of systematic and targeted biopsy locations as well an easy-to-read overlay model for any MRI targets.

## Slice Projection
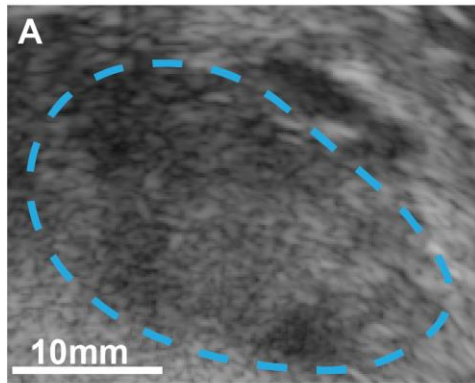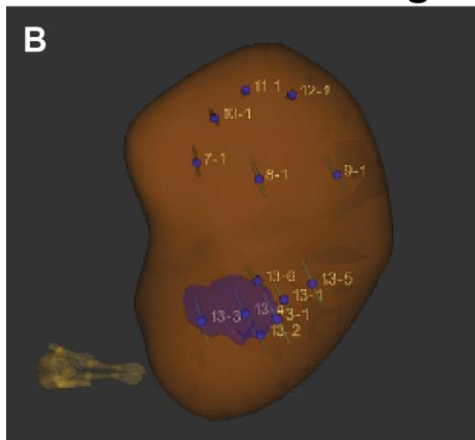
## Surface Rendering

## Volume Rendering
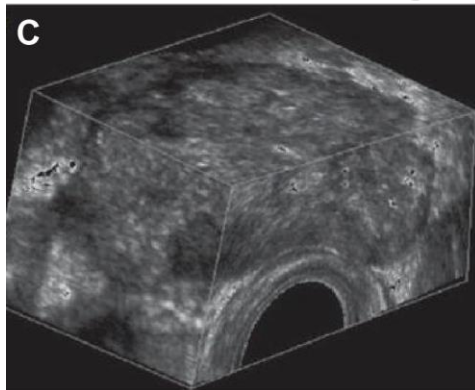
Figure 12: Various methods for rendering 3D US data of the prostate. Slice projection (A) displays a 2D planar image of the US data in native and non-native orientations. Surface rendering (B) displays a model of the prostate capsule for localization, core placement tracking, and MRI target visualization. Volume rendering (C) presents the full 3D sweep of the prostate allowing for thorough review by the physician [143].

Volume rendering (Figure 12C) provides all of the 3D US data together, typically utilizing translucent anatomy to allow for visualization of interior feature information [143]. While computationally intense, this approach provides physicians with the most comprehensive view of the prostate.

The primary roles of US in prostate biopsy are for gland visualization, biopsy needle targeting, and fusion with MRI, rather than for cancer identification. Consequently, it is often better to mitigate the processing times by hiding extraneous information that could potentially distract the user while providing no utility. As such, one common approach that 3D US fusion devices utilize is a hybrid surface rendering and slice projection approach. In this approach, a 3D model of the prostate is displayed with associated slice projections, while tracking probe locations on the model aid with probe navigation. As such, the surface model allows the physician to quickly reference MRI targets and identify biopsy sites; while slice projection allows the physician to review the 3D data as desired. The US probe itself can still be used to interrogate the prostate for biopsy core targeting, with some systems overlaying 3D targeting data onto the real-time US.

## 2.3.5. Fusion Techniques

Following image acquisition and segmentation of both 3D US and T2-weighted MRI scans, a fusion step is performed to co-register MRI image data and target information to the US scan [104], [107]. The goal of the fusion transformation is to map the MRI space and coordinates to the US space to allow for real-time visualization of MRI information on the US scan, thus, improving clinically significant prostate cancer detection rates [104]. This fusion step can also be performed multiple times throughout the biopsy to account for any changes to the prostate shape which often occurs as the US probe

deforms the prostate when it is moved across the gland. One study found that the majority of missed lesions (1.9% of patients) during prostate biopsy were due to errors in targeting (51.2% of misses), rather than the lesion being MRI invisible (40.5% of misses) or missed by the radiologist (7.1% of misses), highlighting the importance of accurate co-registration and targeting [144]. It is estimated in order to have a 90% biopsy hit rate on a MRI identified target, the target registration error (TRE) between MRI and US needs to be less than 2.87 mm not accounting for needle deflection, frequently substantial in clinical practice [145]. Existing MRI-US fusion systems have a validated TRE of approximately 3 mm [113], [146], [147], [148], [149], which may be sufficient for diagnosis but may not be accurate for focal therapy guidance [150]. MRI-US fusion is typically performed using one of the following approaches: rigid registration, elastic (non-rigid) registration, or learning-based registration [104], [107], [151]. A 2018 meta-analysis found no significant difference in the odds ratio for identifying clinically significant prostate cancer between the two most common approaches: elastic (1.45) and rigid (1.40) registration [104]. However, some studies have found improved TRE for elastic registration (2-3 mm) over rigid registration alone (5-6 mm) [113], [146], [147], [148], [149]. Regardless, software-based fusion outperforms cognitive based fusion for identifying clinically significant prostate cancer at time of biopsy [152].

**Rigid Registration**

Rigid or affine registration preserves the shape and relative positioning between pixels of the transformed data [104], [107]. Typical transformations encompassed by rigid registration algorithms include: translation, scaling, shearing, and rotating of unregistered data to match a registration target (Figure 13A,B) [107]. A general rigid registration approach for prostate biopsy starts with US and MRI surface segmentations
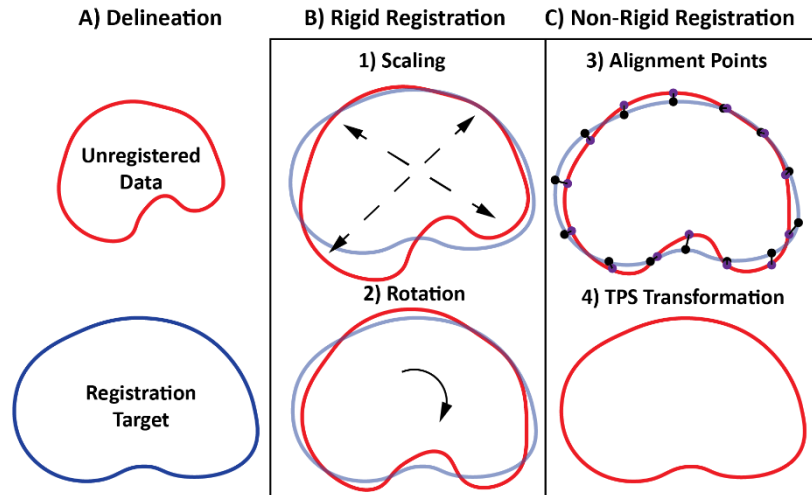
Figure 13: Simple example of a rigid and non-rigid capsule-based registration process in 2D. Using delineated capsules (A) of the unregistered MRI data (red) and the US registration target (blue), the MRI data is scaled (1) and rotated (2) to maximize alignment with the US capsule for rigid registration (B). The subsequent elastic registration step (C) identifies alignment points (3) between the two capsules and performs a thin-plate spline (TPS) transformation (4) to match the MRI data with the US data.

(capsule-based registration) or image data (image-based registration). For capsule or surface-based registration, the MRI data is scaled to match the US data based on the capsule volume differences, and the MRI is translated to align the centroids of the prostate capsule on MRI and US. Following scaling and translation, the MRI is rotated to maximize alignment between the two gland segmentations. At this point the image data between MRI and US is relatively aligned, but there will still be some differences based on tissue deformation from the US probe, an issue that can be accounted for with subsequent elastic registration [107].

An alternative to gland-based registration uses anatomic landmarks (features) and other sources of mutual information in the image data to drive the registration. While image-based registration is more computationally intensive, as it compares raw image data, improved registration accuracy has been reported for image-based mutual information registration over surface-based registration with the added benefit of not

requiring preliminary segmentation [150]. Image based registration often use a technique known as block matching consisting of several iterative steps of increasingly finer adjustments [153]. First, the source image data are divided into rectangular regions called blocks. These blocks are then matched to a similar block within the target image data based on image-to-image similarity metrics, in this case mutual information between image sets [154]. Next, a vector field is generated that maps each source image block to its best matching target image block. A regularization step is applied to the vector fields to generate the transformation map for the source image set to the target image set [150].

**Elastic (Non-Rigid) Registration**

Elastic or non-rigid registration is typically performed after rigid registration and attempts to account for tissue deformation through a transformation map based on alignment points (Figure 13C) [104], [107]. As the US probe pushes against the prostate gland, the shape can significantly deviate from the image of the gland on MRI. This deformation is also operator dependent and changes throughout the procedure as the probe is navigated to different locations of the prostate. Additionally, needle biopsy causes bleed and localized swelling, further affecting registration [104]. Alignment or control points may include a combination of points along prostate capsule or internal landmarks, (urethra, BPH nodules, zonal boundary points, calcifications, etc.) selected by the operator on both the MRI and US scan. The transformation algorithm, typically a spline based transform (b-spline or thin-plate-spline), will warp the MRI data to align the selected control points on MRI and US [22], [107], [155]. While warping the MRI alignment points to match the US control points, the transform will also warp the surrounding voxels proportionally depending on relative distance from the alignment points. For

image based registration, instead of using the capsule segmentation, a moving image is compared to a fixed image and an optimization step compares image-to-image metrics (mutual information) through changing the parameters of a grid of b-splines to yield the transformation with the best match between the two images [150].

**Learning-Based Registration**

Boundary and feature-based registration generally rely on the operator to identify corresponding locations or gland boundaries on MRI and US scans, but are tedious for the operator, challenging to accurately find corresponding points, and are prone to error [22], [151]. To improve fusion quality, consistency, and ease there is considerable research into automated learning-based MRI-US fusion approaches. Conventional machine learning fusion approaches utilize a deformable model and are generally employed as an elastic registration step following rigid registration, making them dependent on operator segmentation for initialization [151]. Utilizing patient specific deformable models and a hybrid surface point matching method, it is possible to achieve a TRE of 1.44 mm [151], [156]. Recent research efforts have focused on developing deep learning models for performing MRI-US fusion, with the goal of predicting the necessary transformation matrix. Due to a lack of a ground truth co-registration transformation matrices for training, research primarily focuses on unsupervised or weakly supervised models [151]. Various approaches include combinations of fully convolutional networks (FCN), 2D and 3D CNNs, and multi-modality generative adversarial networks (GAN) [151], [157], [158], [159]. With the addition of biomechanical constraints to deformation one group was able to achieve a TRE of 1.57 mm and a dice similarity coefficient of 0.94. The framework is capable of first segmenting both MRI and US, then applying a point-cloud-

based deformation for accurate registration in a manner that could potentially be translated to the clinic [151], [160].

## 2.3.6. Validation Techniques

Validation of biopsy accuracy is a key step in demonstrating the clinical utility of a 3D US system for prostate biopsy. While systems may differ in image acquisition, tracking, reconstruction, and registration, the same fundamental sources of error can be independently characterized (Figure 14):

1. **a - Initial registration error**: Imperfect registration between the MRI defined lesion and the initial 3D ultrasound scan.

   **b - Dynamic registration error**: Further error is introduced throughout the biopsy procedure due to tissue deformation from movement of the probe and tissue trauma induced by the needle

2. **Targeting error** – Inaccuracy in positioning needle over the intended target which occurs for many reasons including user inexperience and the spacing of the holes in the biopsy template (if used).

3. **Needle deflection** - Caused by tissue mechanical properties and needle design

4. **Stored location**: In many systems, the coordinates of the biopsy needle are digitally recorded, to enable re-biopsy at a future date. However, the capture of this information may involve a user-confirmed action to mark the biopsy location or needle depth, which introduces perceptual error in the needle localization. This error is also introduced when registering this location with a previously obtained 3D US scan during a follow-up biopsy, typically 6 months after the initial biopsy.
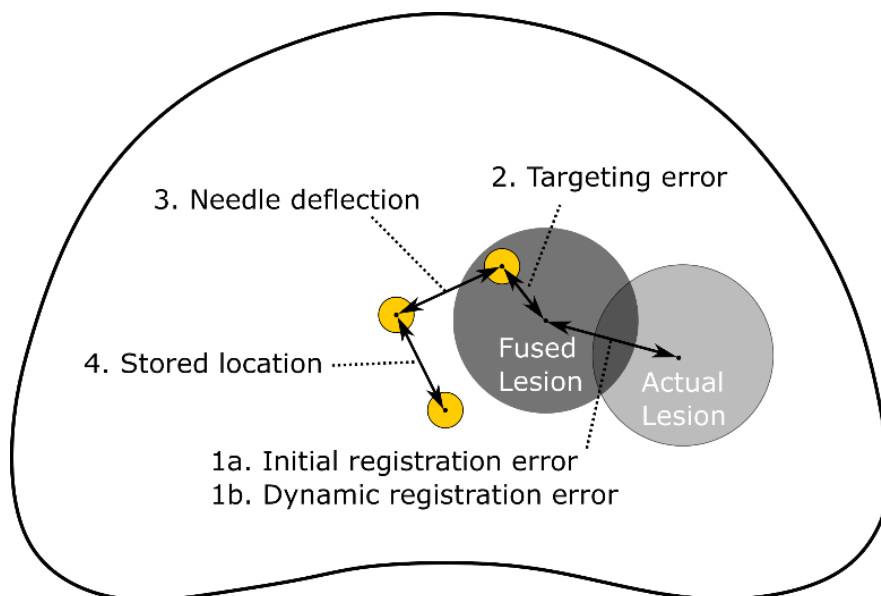
Figure 14: Sources of error when performing MRI-US fusion biopsy. Here, the goal is to biopsy the center of the true lesion.

It is important to note that these errors do not sum together linearly, and the overall error will be less than the sum. It should also be noted that many of the preclinical validation studies were performed on early system versions, sometimes before commercialization, thus, existing literature may not correctly reflect biopsy accuracy. In existing validation studies of commercially available systems, the metric for biopsy accuracy can be defined in numerous ways thus preventing direct comparison between studies. Additionally, due to the nature of regulation of 3D ultrasound systems for prostate biopsy, many commercially available systems report benchtop validation, with unknown accuracy in a real-world clinical setting. Moreover, clinical studies often report on cancer detection rate in comparison to conventional systematic or cognitive fusion biopsy. This further complicates analysis as it is dependent on the reliability of the MRI defined lesion which is essentially a function of radiologist skill and cancer grade. Consequently, when comparing clinical studies, it is important to account for the Prostate Imaging Reporting & Data System (PI-RADS) score and radiologist experience.

58

**Pre-clinical**

Pre-clinical studies typically employ prostate tissue mimicking phantoms composed of agar or hydrogels to assess the targeting error of fusion guidance platforms [149]. In order to determine TRE biopsy samples are taken of MR- or CT-visible targets so that the distance between the needle tracks and lesions can be calculated. Using this approach, Uronav, an electromagnetic-tracking system, reported a TRE of 2.4 ± 1.2 mm [66].

The Trinity system, which uses an image-based tracking approach, was validated using a similar study with a hydrogel phantoms except the needle tracks were identified by injecting a mixture of dye and gadolinium though the outer sheath of the biopsy needle immediately after deployment [148]. The phantoms were then scanned on MRI to reconstruct the 3D volume and also sectioned and photographed to identify each needle track based on dye color coding. Here, procedural targeting error is defined as the shortest distance between the needle track and lesion center, was found to be 2.09 ± 1.28 mm. Additionally, the system registration error was 0.83 ± 0.54 mm. It was then assumed that these errors are independent and additive giving an overall error of 2.92 mm.

The BiopSee system was validated using the same approach and found a targeting error of 0.83 ± 0.48 mm [161] which was defined as the distance between virtually planned cores and the confirmed needle tracks. The needle track was also derived from the US images during biopsy. Importantly, procedural targeting error (PTE) is defined differently than above, corresponding to the distance between this tracked core and the virtually planned core. This measure of PTE was found to be 0.26 ± 0.46 mm. There was no statistical difference between the two metrics which was used to support the use of this automatically calculated error in later *in vivo* studies [162]. While this study

supports the use of PTE, it is not clear if it provides an accurate assessment of overall error (distance from lesion center to true needle track). Indeed in a similar a follow-up phantom study [163], the PTE (0.39 mm) was found to be only a modest contributor to overall error (2.33 mm). Moreover, errors are likely to be even greater *in vivo* due to needle deflection from inhomogeneous tissue resistance and reduced precision of contouring.

A similar study was also performed during the initial validation of the Artemis platform [164]. Here, the authors constructed an agar phantom using an MRI scan of a patient's prostate and previously developed phantom recipe [165]. The total needle guidance error was found to be 2.13 ± 1.28 mm which largely arises due to needle deflection which accounts for 2.08 ± 1.59 mm. Additionally, a commercially available hydrogel phantom (Zerdine, CIRS) was used for assessing the accuracy of segmentation and volume calculation which exhibited a 5% error. Agar phantoms are advantageous as they are inexpensive, can be easily cast in patient-specific molds and also permit placement of fiducials of any desired shape and size [166], [167]. Similar phantoms can be produced for extending the study to include analysis of ultrasound guided focal therapy modalities such as laser ablation [168], [169], high-intensity focused ultrasound [170] and radiofrequency ablation [171].

## Clinical – Target Registration Error

Clinical studies typically focus on cancer detection rates rather than attempting to assess biopsy accuracy because of the inability to assess ground truth needle tracks. However, there are some notable exceptions such as the a study by Moldovan et al which rigorously quantifies the registration accuracy of the Trinity fusion platform [172]. Three intraprostatic fiducials were inserted into patients prior to radiotherapy. The fiducials

were inserted through a needle under ultrasound guidance using the same procedure that would have been used for biopsy. The patient then received an MRI on the same day. Using the ultrasound images, a virtual needle was positioned with the tip defined by the fiducial and the angle defined by that of the ultrasound probe. Virtual lesions were defined as a sphere overlaying the fiducial on the MRI image set. The ultrasound and MR images were then fused together using the software on the Trinity system. Registration error was defined as follows: 1) $TRE_{3D}$ is the distance between the tip of the virtual needle and the virtual lesion, and 2) $TRE_{2D}$ is the orthogonal distance between the virtual needle tract and the virtual lesion. $TRE_{2D}$ is the better metric of registration accuracy as small errors along the needle trajectory are less critical given that most biopsy cores are 15-20 mm. Here, $TRE_{3D}$ and $TRE_{2D}$ ranged from 3.8-5.6mm and 2.5-3.6 mm respectively with much of the variation attributed to location and user experience. Importantly, the design of the study removes contributions from needle deflection thus providing an accurate assessment of registration error alone.

Studies such as this suffer from three intrinsic sources of error: 1) fiducial localization error, 2) fiducial migration and 3) segmentation variability. When determining registration accuracy with fiducials, it is necessary to define the center of the fiducial on each image set. This is generally a manual process and as such it is user dependent. In an earlier study, the same user identified fiducials on the same image set for 5 consecutive days resulting in a fiducial localization error (FLE) of $0.21 \pm 0.11$ mm [173]. This suggests that FLE is a relatively minor contributor to overall error. In contrast, the same study found that segmentation variability was responsible for ~50% of the calculated TRE. Segmentation error arises when using semi-automated algorithms in which the user is required to manually outline the prostate on a subset of the images.

Finally, implanted fiducials are susceptible to migration with one study showing mean migration of 1.2 ± 0.2 mm in patients receiving radiotherapy [174]. Fiducial localization error and migration can be mitigated by using anatomical landmarks (e.g., calcifications, capsule, peripheral zone – transition zone boundary) instead of implants. Accurately identifying these landmarks on different imaging modalities is quite challenging but can be achieved as demonstrated in multiple studies including ultrasound, MRI and wholemount registration [157], [167], [175], [176].

In one clinical study on 106 patients, TRE was determined to 1.7 mm [162] using a stepper-stabilizer, while image-based TRE in another system was found to be 2.8 mm in a clinical study on 88 patients [113]. In the same study, the accuracy of cognitive fusion was 7.1 mm. It is important to note that the methodology and definitions of error between the two systems are very different, thus, they cannot be directly compared. Moreover, these automatically calculated errors may not accurately reflect the true TRE.

### Clinical – Cancer Detection Rates

In general, clinical studies have continuously demonstrated that targeted biopsy has a greater detection rate than conventional systematic or cognitive fusion biopsy, though taking both targeted and systematic biopsy cores yields the highest rate. In an early clinical trial on 101 patients, the per core cancer detection rate was significantly higher for targeted biopsy than systematic biopsy at 20.6% vs 11.7% [177]. The difference was even more pronounced for targets deemed to be highly suspicious for cancer on MRI at 53.8% vs 29.9%. The same group subsequently reported on targeted biopsy performed on patients with a prior negative systematic biopsy and found cancer in 37.5% of cases and high grade cancer in 11% [178]. Additionally, targeted biopsy was responsible for pathological upgrading in 38.9% compared to systematic biopsy alone. Similar results

were reported in a later trial where 23.5% of cancers diagnosed as clinically insignificant on systematic biopsy were upgraded after targeted biopsy [179]. In a study targeting high grade regions of interest, fusion biopsy exhibited a 33% positivity rate in comparison to the 7% achieved via systematic biopsy [110]. In another clinical trial, 91% of men with prostate cancer had significant cancer when diagnosed via targeted biopsy vs 54% with systematic biopsy [180]. Comparable results have been reported for numerous biopsy platform [113], [162], [181], [182], [183], [184], [185].

**Ground-Truth Validation Studies**

Co-registration and targeting error validation studies are vital for fusion devices. Without these studies, it is impossible to be confident that targeting is accurate and biopsy needles are placed in the right location. Improper targeting can result in false negatives, or poor mapping of positive cores to the right location, resulting in possible changes to clinical treatment of a patient's disease. Furthermore, many studies rely on phantom studies because they are highly controllable and do not require human or animal subjects, however it is worth considering if these phantoms are adequate for assessing true targeting error. These phantoms are generally homogenous, and designed to mimic prostate shape and acoustic properties, for similar US imaging patterns to human tissue. However, this does not necessarily mean these phantoms necessarily share similar heterogenous mechanical properties of human tissue, and sources of error such as tissue deformation and needle deflection may not be properly represented. In general, it is found targeting error increases when moving from phantoms to real prostates. At the same time, it can be difficult to assess the true accuracy with prostate tissue. For even highly controlled radical prostatectomy studies, the associated error of matching a pathology slide with an associated MRI image can be on the order of 1-3mm,

which could potentially stack with any targeting errors potentially inflating registration errors [175]. Additionally, landmark studies are dependent on a reviewer's perception of the tissue, which can be subconsciously biased if there is an inherent understanding of the underlying registration algorithm, or over-report the registration error if a reviewer incorrectly pairs two points between MRI and US. Furthermore, targeting error is not ubiquitous throughout the whole gland, and superior results are typically obtained in the mid-gland where tissue deformation is mitigated compares to the far apex, base, or lateral zones of the prostate. As such, it is possible to improve targeting error by increasing the sampling density at mid-glad and decreasing the sampling density in areas prone to tissue deformation. For this reason, it is important to note not only overall targeting accuracy, but also sampling distributions and targeting accuracy in various zone of the prostate. In summary, there are inherent weaknesses to all validation techniques making it difficult to assess the true targeting error for prostate biopsy fusion systems, however, by performing multiple types of validation studies with detailed reporting and large sample sizes it is possible to mitigate random error and more confidently quantify the targeting accuracy of these systems.

## 2.4.　Micro-Ultrasound

Micro-ultrasound (microUS) was originally developed as a pre-clinical ultrasound system for imaging mice models of human diseases at frequencies between 15-50MHz [186]. The system allowed for impressive visualization of mice cardiac activity and embryonic development (Figure 15). The system was later translated for use in prostate cancer imaging as the ExactVu system (Exact Imaging, Ontario, Canada) (Figure 16). Approved by the FDA for prostate biopsy in 2016, the system has demonstrated promising results
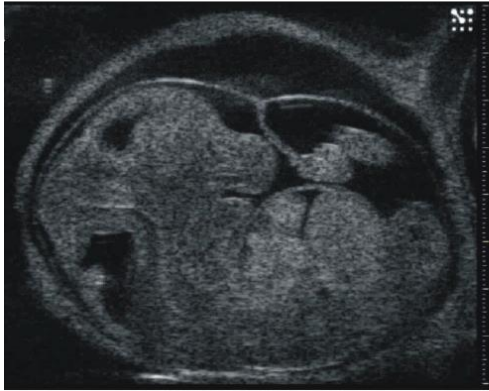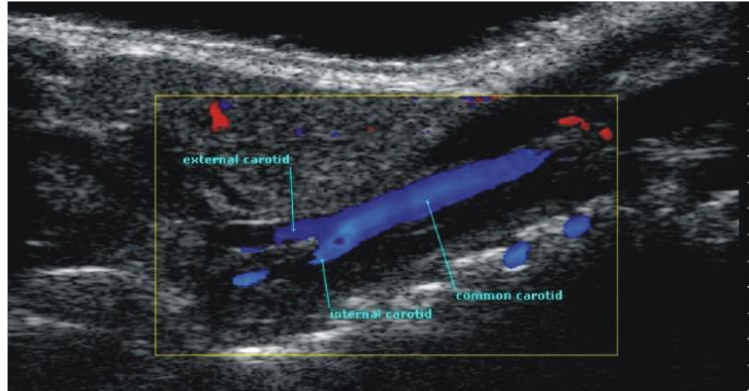
Figure 15: Mouse imaging using microUS allowing for visualization of embryos and the carotid artery [186].

for prostate cancer visualization in tracked biopsy studies following development of a microUS suspicion scoring [16], [17], [18]. The high-resolution ultrasound images allow for higher quality prostate imaging over conventional US imaging for direct visualization and biopsy of prostate cancer. As such, microUS could represent a low-cost alternative or supplement to mpMRI imaging.

## 2.4.1. System Design

MicroUS uses a 512-element linear array 29MHz transducer with inter-element spacing of 90μm based off of the work of Foster *et al* [186]. This allows for an axial resolution of 70μm in proximal regions of the prostate, encompassing much of the peripheral zone, which also accounts for approximately 70% of all prostate cancers [186], [187]. To image deeper in the prostate the microUS system uses a distance-based beamforming design that images more distal regions of the prostate at lower frequencies creating image bands of differing axial resolutions (70μm [proximal] - 100μm [distal]) that comprise the full B-mode image. This approach allows for an imaging depth of 6cm at the cost of an increased axial resolution. In contrast, conventional US generally has an imaging

Figure 16: ExactVu micro-ultrasound system created by Exact Imaging.

frequency in the 2-15MHz range and an axial resolution as fine as 200μm [16], [186]. The superior microUS resolution allows for visualization of prostate features and tumors that are not resolvable on conventional US which may facilitate tumor visualization.

## 2.4.2. PRI-MUS System

The first prospective randomized clinical trial with microUS against conventional US for prostate cancer detection initially had no improvement in cancer detection rates for
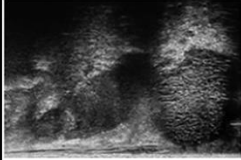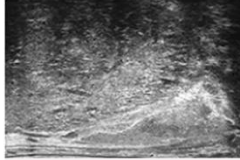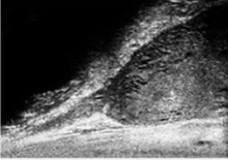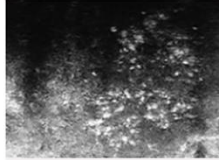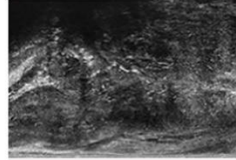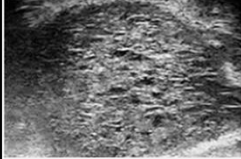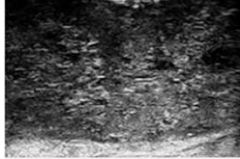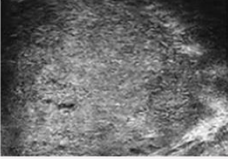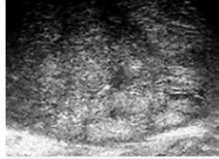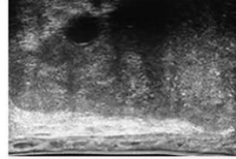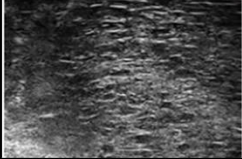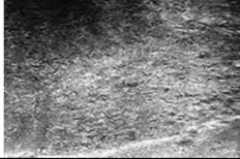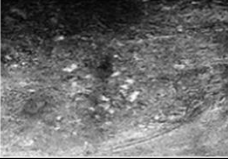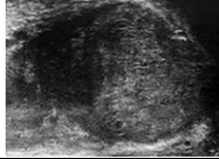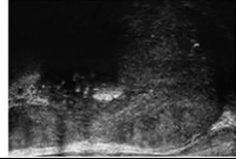
Figure 17: The PRI-MUS scoring system used to highlight suspicious features of prostate cancer on microUS imaging.

microUS over conventional US guided systematic biopsy for 1113 patients [188]. Using retrospective analysis of microUS with 400 tracked biopsies, the first iteration of the prostate risk identification using micro-ultrasound (PRI-MUS) suspicion scoring system was developed (Figure 17) [18]. This risk identification system was designed to highlight indications of PCa on microUS similar to the PI-RADS system for mpMRI. Following initial development, this system was used for brief physician training (< 1 hour) and the prospective clinical trial was resumed, in which there was an increase in microUS sensitivity from 24.6% to 60.8% [18], [188]. However, there was an associated decrease in specificity and no improvement in cancer detection rates over conventional US guided biopsy [188]. Regardless, the increase in sensitivity from brief training was promising and warranted further investigation into microUS. Since initial inception, there has been

67

effort to further refine the PRI-MUS system, particularly for detection of anterior tumors [189].

## 2.4.3. Existing Studies

Following development of the PRI-MUS system, follow up studies were performed to investigate cancer detection rates, sensitivity, and specificity of microUS related to gold standard mpMRI. A multi-center prospective biopsy study by Klotz *et al.* with 1040 patients found microUS targeted biopsies had a sensitivity of 94% for PCa compared to 90% for mpMRI targeted biopsies [16]. Furthermore, both modalities exhibited similar specificity for PCa at 22%. This study also highlights the importance of mpMRI as a benchmark for evaluation of microUS. Without the mpMRI benchmark a reported specificity of 22% for microUS would be concerning. However, since gold standard mpMRI also had a similar specificity for PCa (that is often reported between 60-83% [6], [7], [8]), it is likely a specificity of 22% is not abnormally low and rather likely due to the type of analysis performed. Sensitivity and specificity are highly dependent on the number of possible events in a study, so results can fluctuate between studies. This is highlighted further in this thesis in chapter 5, but in general as the granularity of an analysis increases sensitivity will decrease and specificity will increase. Furthermore, it is much easier to record a positive event in a biopsy study (taking a biopsy core) compared to a negative event (not taking a core).

A blinded prospective biopsy analysis by Ghai *et al.* found similar cancer detection rates between microUS and mpMRI (35% vs 39% respectively) [17]. Between the two modalities 67% of mpMRI identified ROIs were also identified on microUS, however clinically significant PCa was only present in 58% of the concordant lesions. Interestingly

the study found there was no added benefit from systematic biopsy over the combination of mpMRI and microUS guided biopsy. This highlights a potential synergistic benefit of implementation of the combination of microUS and mpMRI for centers that are able to utilize both modalities.

After the start of our efforts to investigate microUS there have been additional studies investigating the potential of microUS. Callejas *et al* found microUS was able to correctly identify the index lesion cancer in 87% of patients undergoing radical prostatectomy for PCa using cognitively aligned quarter-mount pathology for validation [190]. Similarly using cognitive alignment of microUS and quarter-mount pathology, Pedraza *et al* found similar sensitivity (68.7% vs 68.6%) and specificity (96.3% vs 97.2%) between microUS and mpMRI for index lesion identification in 363 radical prostatectomy patients [191].

There is a growing body of evidence that microUS is able to identify PCa with similar accuracy to mpMRI. In this thesis we sought to build upon the existing literature through rigorous validation of microUS through computational registration to ground truth WM pathology with benchmark comparison to mpMRI.

## 2.5. Deep Learning for Prostate Cancer

With the explosion of artificial intelligence (AI), there have been numerous applications and commercialization of machine learning in prostate cancer imaging, ranging from segmentation of prostate capsules, to MRI-US fusion, treatment planning, and prostate cancer detection [151], [192]. Due to the prevalence of radical prostatectomy for prostate cancer treatment, it is possible to obtain high-quality ground truth for model training from surgically removed prostates. Furthermore, MRI-US fusion is common for prostate

cancer biopsy, requiring segmentation of prostate capsules on both MRI and US and co-registration between the two modalities. Additionally with the development of focal therapy there has been considerable effort to develop AI derived treatment margins for optimizing PCa treatment to minimize treatment impact on impact on patient quality of life [192].

## 2.5.1.  Applications for Cancer Detection on MRI

There have been a multitude of studies for prostate cancer detection on MRI using machine learning [151], [193]. General approaches often focus on image classification and follow with lesion localization and delineation [151], [194]. Specific approaches have varied but models have obtained impressive AUC values often ranging between 0.8-0.9 [151]. There are a few publicly available prostate MRI databases such as the PI-CAI dataset containing over 10,000 MRI scans from 4 centers. Using this dataset, a multi-reader study of 62 radiologists from 45 centers was used for benchmark comparison in a test set of 400 patients [195]. The AI model was found to have a statistically superior AUC compared to the radiologists (0.91 vs 0.86) [195]. At the same specificity of the readers, the AI model detected 6.8% more clinically significant cancer, and had 50.4% less false positives than the radiologist cohort [195].

Typical approaches for detection and delineation on mpMRI typically employ nnU-net as a backbone due to its flexible nature and strong performance for a wide variety of medical image segmentation, often outperforming specialized deep-learning approaches [196], [197]. Attention network approaches have also demonstrated promising results, utilizing whole image classification in conjunction with patch based classification to create localization maps of suspected regions of cancer [198]. There has

70

been numerous demonstrating strong performance of AI and deep learning for PCa identification on MRI. Implementation of these models in future work could aid radiologists by alleviating much of the inter-reader variability associated with MRI while improve PCa detection and patient care.

## 2.5.2.   Applications for Cancer Detection on Micro-Ultrasound

There have been several studies investigating the utility of machine learning for PCa detection on microUS with biopsy validation [82], [199], [200], [201], [202]. The specific approaches range from linear discriminant classifiers and support vector machines to self-supervised learning with promising AUC values (0.80-0.93) for classification along the biopsy track in datasets between 160 to 700 patients. Approaches have predominantly relied on weakly labeled microUS patches labeled from biopsy data for training and validation. Implementation of confidently labeled data validated from WM pathology should allow for a higher quality dataset that could increase training results. Additionally, studies have primarily used broadband RF US signals to extract useful features for training. Broadband RF data provides substantially more data per image, but requires clever approaches for extracting "useful" information and greatly increases data storage requirements.

Deep learning approaches for cancer detection on microUS are promising. Ultimately, there is a lack of labeled datasets for microUS training relative to approaches with MRI but efforts are ongoing to address this. In this thesis we work towards that end through the development of a high-quality microUS dataset and application of deep learning techniques.

# CHAPTER 3

# Multi-modal Imaging of Prostate Cancer and Ex Vivo Ultrasound Scanning

## 3.1. Introduction

Micro-ultrasound has recently been introduced as a potential low-cost alternative to mpMRI for PCa visualization during biopsy. Early studies have demonstrated similar performance between microUS and mpMRI for sensitivity (93% vs 94%), specificity (21% vs 17%), and cancer detection rates (35% vs 39%) when utilized by expert readers [16], [17], [18]. While these studies are promising, they are validated via prostate biopsy which provide limited information regarding cancer extent in the prostate. As such, a more rigorous investigation of microUS would implement validation against ground truth WM pathology, the most comprehensive form of final pathology, with mpMRI as a benchmark for the clinical gold standard.

To perform this investigation, a dataset would need to be collected at an institution capable of capturing WM pathology as well as pre-operative mpMRI and microUS scans for patients undergoing radical prostatectomy for prostate cancer. Unfortunately, at the time of starting this thesis, no such dataset is available. However, UCLA is one of the few medical centers in the United States capable of capturing all three modalities (microUS, mpMRI, and WM pathology). In this chapter, we detail the creation of a clinical trial to allow for rigorous investigation of microUS imaging. The data

collected in this trial will be used in the subsequent chapters to quantify the diagnostic capability of microUS (chapters 4 & 5) as well as create confidently labeled data for deep learning model training (chapter 6). To collect a dataset that could allow for accurate registration between the three modalities, there needed to be emphasis on controlled data collection. More specifically, the positional information of images needs to be tracked at all times. Fortunately, this is routine practice for mpMRI, with the DICOM standard, and UCLA routinely processes WM pathology to be aligned with mpMRI. However, microUS is recorded as a cine (video) with limited positional information, and WM pathology is recorded as a series of physical slides. To address this special care is taken in this dataset to capture as much data as possible while preserving positional information. Beyond that data also needed to be represented in a way that could also be easily reviewed and analyzed retrospectively. To address these issues, microUS and WM pathology data were converted into custom DICOM series using custom developed MATLAB scripts allowing for retrospective review similar to mpMRI and preserved positional information.

While microUS has demonstrated promise in earlier clinical studies, there is a lack of an identifiable contrast mechanism for its improved diagnostic capability over conventional US. The PRI-MUS system used to guide suspicion scoring of cancer on microUS is validated from tracked biopsies, but there is limited published investigation into what these features represent morphologically and why they are visible on microUS but not conventional US machines. Due to the costs associated with a clinical trial and collecting the dataset needed for this thesis, we wanted to provide a reason for the increased diagnostic capability of microUS using quantitative evidence to justify continuation of the study. To achieve this, in the last part of this chapter we utilize

computer vision to investigate differences in cancer morphology in regions of cancer and benign tissue on WM pathology.

## 3.2. Enrollment

Eighty-six patients were enrolled into a registered clinical trial (NCT04299620) and UCLA institutional review board (IRB) approved study (IRB#19-001136) after providing informed consent. To be eligible for enrollment patients had to satisfy the following eligibility requirements:

1. Scheduled to undergo **robotic-assisted radical prostatectomy** (RALP) at UCLA Ronald Reagan hospital with the prostate removed as a whole gland

2. **Prostate biopsy** in the last 18 months with at least one core with confirmed prostate cancer

3. An **mpMRI** scan within a year prior to surgery with adequate scan quality for radiologist review

4. **No prior prostate cancer treatment**, including hormone therapy, radiation, and focal therapy

5. A **prostate size** capable of being sectioned in a pathology sectioning mold (approximately **110g** or less)

6. A **prostate anterior to posterior axis length** of no more than **6cm** (the maximum window size of a microUS image)

Potential patients were screened weeks in advance of their surgery. If patients were deemed to be eligible, they were contacted by a physician who provided the details of the study and how it would affect their procedure specifically. Patients were then given the opportunity to participate or opt out of the study. If the patient consented, a

physician also discussed the procedure with them during pre-operative care on the day of surgery. If the patient still consented to participate in the study, necessary enrollment forms were signed.

## 3.3. Data Collection

Following enrollment into the study, the following groups of data were collected for each patient: clinical variables, mpMRI scans, microUS scans (both in vivo and ex vivo), and WM pathology data (Figure 18). Each of these are described in detail in the following sections. To our knowledge, this was the most comprehensive and controlled dataset involving microUS imaging at the time this data was collected. This data would ultimately allow us to assess 1) the true accuracy of mpMRI and microUS delineated lesions through computational registration to WM pathology with tracked registration
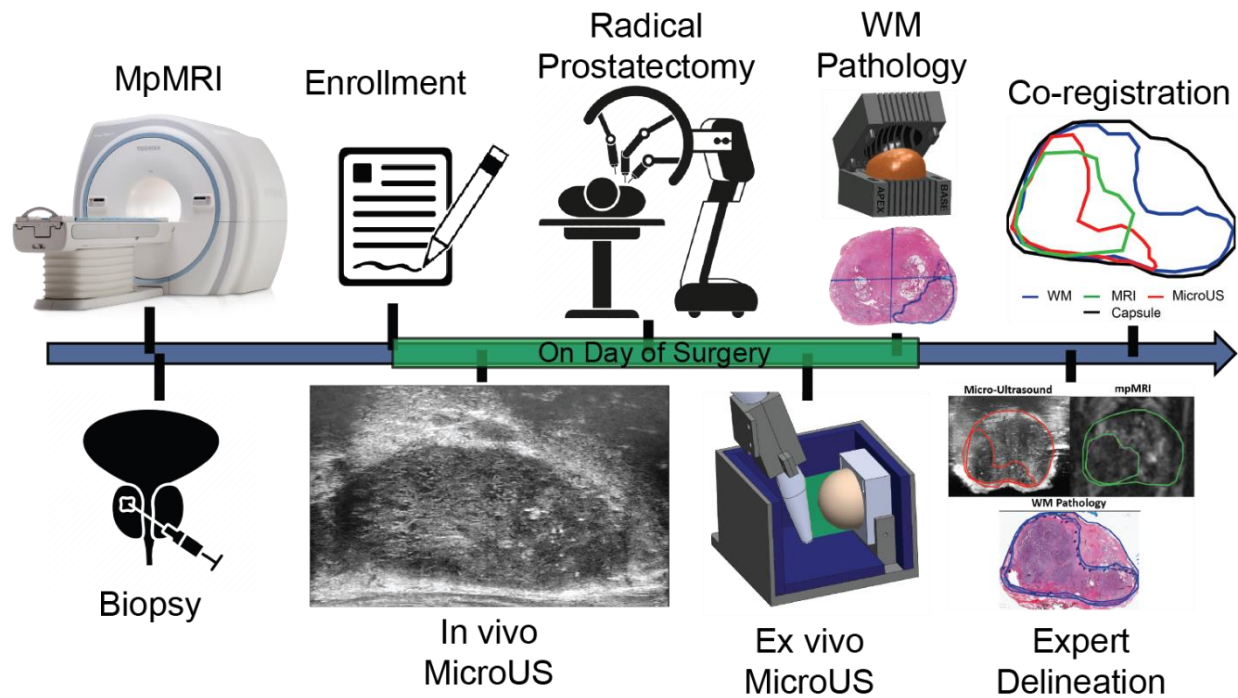


Figure 18: Clinical trial workflow.

error, and 2) create confidently labeled microUS images allowing for deep learning classification of prostate cancer.

### 3.3.1. Clinical Variables

To frame the results of this study, it is necessary to collect clinical data from every patient enrolled. More specifically, prostatectomy patients tend to have more aggressive, higher volume tumors than general biopsy cohorts. While we attempted to overcome this limitation through paired evaluation using gold-standard mpMRI as a benchmark, it is an inherent limitation of the work presented in this thesis. As such, the following clinical parameters help frame the results of this work in context, particularly for physicians, and are vital for the quality of the dataset. The specific parameters collected include the following:

1. **PSA** – prostate specific antigen score, higher values generally correlated with larger prostates or more higher grade/volume cancer.

2. **PSAd** – prostate specific antigen density, equal to PSA divided by prostate volume removing the volume dependence of PSA scores.

3. **DRE** – digital rectal exam at the time of surgery, positive results are indicative of aggressive disease

4. **Age** – Older patients are more likely to have later stage prostate cancer

5. **Prostate Volume** – Smaller prostates are more easily interrogated by US imaging due to attenuation decreasing with pulse length, especially at higher frequencies.

6. **Cancer stage** after resection – The extent of tumor involvement in the prostate, another indication of the aggressiveness and size of the cancer.

## 3.3.2. Multi-parametric MRI

As mentioned above, to be eligible for enrollment each patient needed to have a mpMRI within a year prior to surgery. As prostate cancer is generally a slow growing cancer, it is not expected the tumor would grow significantly over the span of one year. Each patient had a T1, T2, DWI, ADC, and DCEI scans comprising their mpMRI in a 3 Tesla scanner. An experienced radiologist reviewed each case pre-operatively in accordance with standard clinical practice and identified any PI-RADS targets with associated suspicion scores. Furthermore, the quality of the MRI scan was also assessed by an experience radiologist to ensure adequate image quality prior to enrollment in the study. If the MRI occurred at another institution it was re-read by an experienced UCLA radiologist. While there is variability in scanning parameters across institutions, the UCLA MRIs were generally acquired using consistent scan parameters. The resolution of the T2 sequence, the sequence most often used for delineation, is 0.66mm in-plane with a through plane spacing of 1.5mm.

## 3.3.3. Micro-Ultrasound

**In vivo Micro-Ultrasound**

On the day of surgery, each patient received multiple transrectal b-mode microUS scans of their prostate following anesthesia but before the start of surgery. Each scan was performed by an experienced urologic fellow trained on microUS imaging to ensure consistency across scans. First, the urologist positioned the probe at mid-gland in the prostate, indicated by the appearance of the urethra in the prostate. The overall gain of the images as well as the time gain compensation sliders were adjusted to provide a quality image as deemed by a urologist and imaging scientist, and the scan angle was
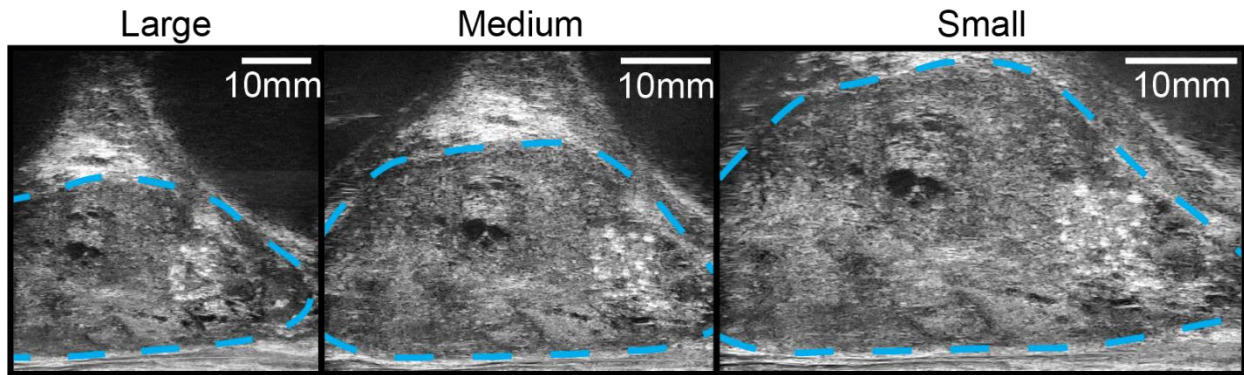
Figure 19: Large, medium, and small imaging presets for microUS imaging of the same location in a prostate. The prostate capsule is marked by a blue dashed line. By using a smaller imaging preset it is easier for the operator to visualize smaller anatomical features due to the larger appearance on the screen.

calibrated to align with the urethra. The microUS probe records the roll angle for each frame in a scan. Next, the physician oriented the probe to just beyond the anatomical right edge of the prostate. Following this, the physician rotated the prostate to the anatomical left edge of the prostate over the course of 30 seconds under three different imaging presets: small, medium, and large. A cine video was then saved to the US machine, at a framerate of 10 frames per second for a maximum of 300 frames, the reasoning for the 30 second scan. After a scan was completed, it was reviewed by the scanning urologist for quality, if necessary, the scan was repeated. As discussed in the background chapter, microUS utilizes three different beam frequencies to image at deeper depths of the prostate resulting in bands of frequency windows ranging from 70μm-100μm for proximal to distal areas for a single image. As such the smaller the scan window preset, ranging from 3cm-6cm for small to large presets, the better the overall image quality and view of the prostate (Figure 19). The pixel spacing for the three imaging presets are Small: 0.03mm/pixel, Medium: 0.04mm/pixel, and Large: 0.05mm/pixel. It is recommended by the manufacturer to use the smallest possible imaging preset that still encompasses the suspected region of cancer. Since the microUS images were not interpreted at the time of the scan and to avoid bias at the time of

review, each prostate was scanned under all three imaging presets. Later, when the scans are interpreted, the reader was provided the option of all three scans. Following image collection, the images were exported to a database under a study assigned identifier for the patient.

To allow for ease of interpretation, we developed custom pre-processing scripts to convert the microUS cine videos to DICOM image series with embedded pixel spacing and angle tracking. In doing so, readers were able to segment, measure, and localize various locations in the prostate which would not be possible using a video. Any reviewer segmentations were also able to be saved and shared for analysis purposes. Since DICOM is generally designed for cartesian coordinates, opposed to the cylindrical coordinates of the microUS sweeps, the rotation angle was recorded as the slice location for the DICOM metadata. While this allows for visualization during review of a sweep in a DICOM viewer, it does not allow for accurate 3D visualization of the prostate in alternative imaging orientations as is possible with MRI, an issue that will be discussed in detail in chapter 5.
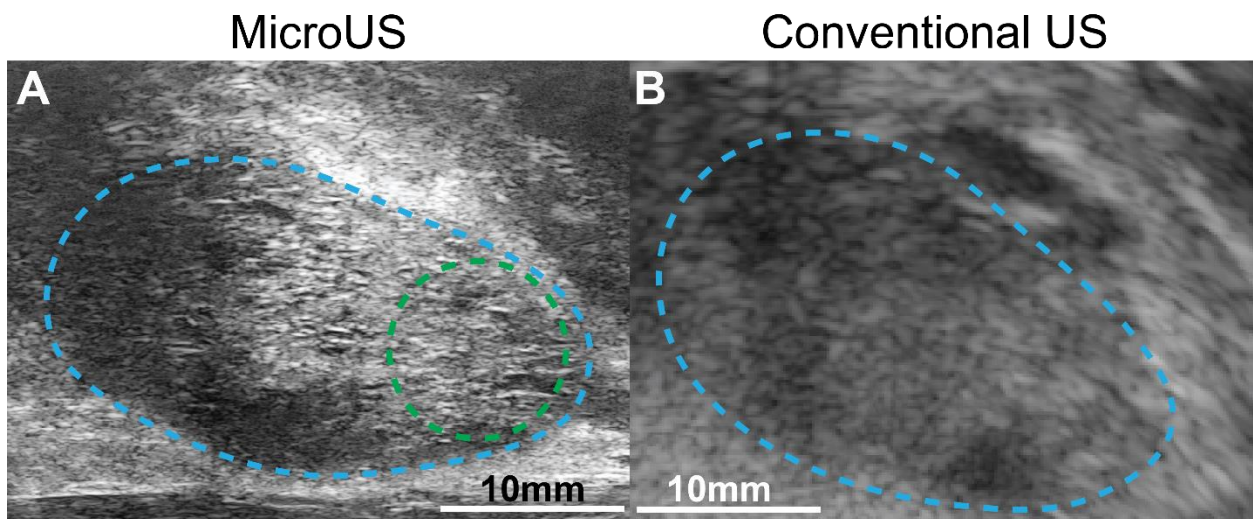


Figure 20: A) MicroUS image of a prostate (outlined in blue) with visible prostate ducts (outlined in green). B) Conventional US image of the same prostate tissue (outlined in blue) with no visible prostate ducts

For a very limited number of patients (N=2), it was also possible to acquire conventional US scans of the prostate. Following the same process as with the microUS system, a Hitachi Noblus ultrasound system and a C41L47RP probe was used to scan the prostate at 9MHz just prior to surgery. The prostate was interrogated from anatomical right to left and the scans were saved and exported for comparison with microUS. Since the probe roll is not tracked on this system, any comparison between conventional US and microUS for in vivo imaging relied on visual correlation. Due to operating room time constraints and availability of an additional US system for scanning, it was not possible to acquire conventional US imaging for more patients. Despite the limited patients with conventional US imaging, the direct comparison between microUS and conventional US allowed for appreciation of the improved image quality microUS provides and the visualization of previously non-resolvable structures.

**Ex vivo Micro-Ultrasound**

Following prostate resection but prior to grossing and sectioning, the prostate was imaged in a custom ex vivo microUS scanning system to assist with registration provide an aligned imaging orientation relative to WM pathology and mpMRI. The development and validation of this system will be covered in detail in the following chapter. Briefly, the prostate was placed in a patient specific alignment mold, positioned in the scanning tank and imaged in the axial orientation (the same orientation as MRI) allowing for straightforward registration between modalities. Using an electronic translation stage, a mounted microUS probe scanned the prostate from apex to base. The prostate was interrogated in detail using 5 scans per imaging preset (small, medium, large) for a total of 15 ex vivo scans. Confined by the 30 second scan limit, the prostate was interrogated in one full scan from apex to base at a scan speed of 2.75mm/sec and an imaging spacing

| Ex Vivo Micro-Ultrasound | Ex Vivo Conventional Ultrasound |

**Pathology Confirmed Prostate Cancer**
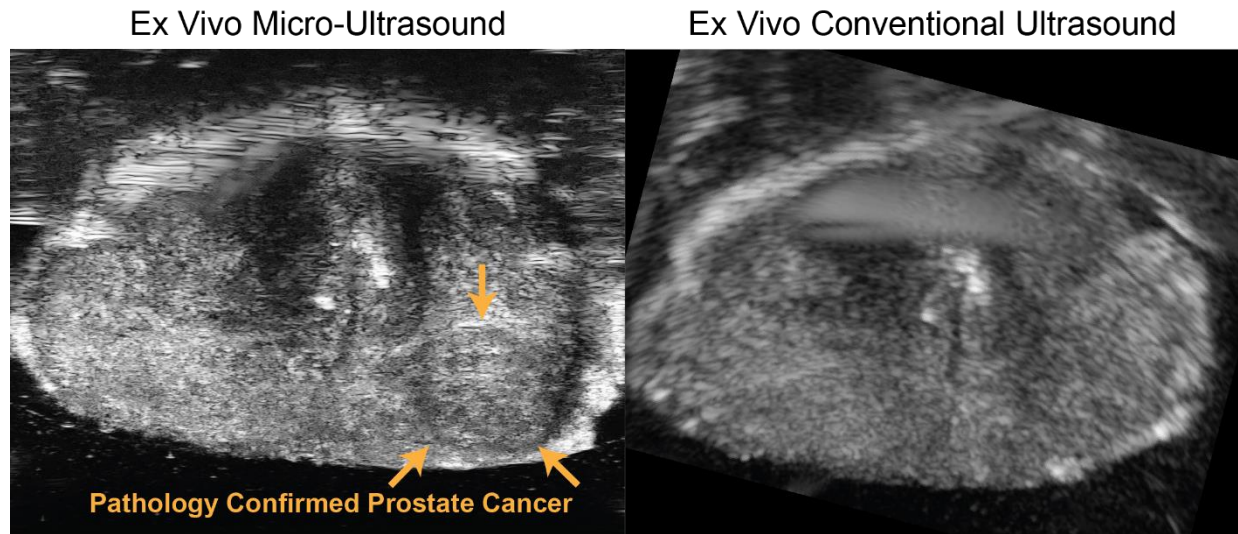
Figure 21: Comparative images between microUS (left) and conventional US (right) acquired in a custom-built ex vivo scanning system. A WM confirmed lesion is visible on microUS only (orange arrows).

of 0.275mm/frame. Additionally, the prostate was interrogated in 4 scans traversing approximately the quarter of the distance from apex to base each. These scans were translated at a slower speed of 0.65mm/sec and an imaging spacing of 0.065mm/frame. These 5 scans were repeated for each imaging preset. Prior to imaging, the prostate was aligned to be centrally located in the frame and the overall gain and time gain compensation sliders were adjusted to provide a quality image as deemed by a urologist and imaging scientist. Following scanning, the images were converted to DICOM series as with in vivo microUS based on the controlled location of the translation stage during scanning. The 4 quarter scans were also stitched into one complete scan.

As with in vivo imaging, ex vivo conventional US imaging was acquired for a limited number of patients (N=4). As with microUS, the conventional US system and probe were a Hitachi Noblus ultrasound system and a C41L47RP probe. The conventional probe was mounted to the translation stage and scanned the prostate at 9MHz from apex to base. Cine video sweeps were exported and used for comparison with microUS. Since ex vivo imaging is axial, it is possible to compare to WM pathology. An example of

cancer visualization of WM pathology confirmed cancer on microUS that is not visible on conventional US can be seen in Figure 21.

### 3.3.4. Whole-Mount Pathology

Following ex vivo imaging, each prostate was processed to create WM pathology slides in accordance with routine clinical practice at UCLA. WM pathology is not widely available, and there are a limited number of centers in the United States capable of capturing this information. Fortunately for this work, UCLA has a robust WM pathology processing pipeline that not only allows for it to be created for every patient, but also takes special care to section the prostates in the same orientation as MRI using patient specific 3D printed sectioning molds. The overall process used at UCLA to create WM pathology slides is detailed below.

**Pathology Sectioning Molds**

A few days before surgery, the preoperative T2 MRI of each prostate was delineated by an experienced radiologist. Using a MATLAB script, the delineations are converted into a 3D model. This virtual model of the patient's prostate was then subtracted from a surgical processing mold template to create patient specific surgical processing mold. The mold template has slits running from the apex to base every 4.5mm to provide a track for sectioning. The processing mold was then loaded into a slicing application to create instructions for 3D printing (gcode). Using fused deposition modeling (FDM) based 3D printing, polylactic acid (PLA) was extruded from a 3D printer to create the surgical processing mold (Figure 22A). This mold is designed to house the prostate to allow for it to be cut in the axial orientation with less error than when sliced by hand [167] and is used routinely at UCLA for radical prostatectomy patients. A similar mold

creation process, discussed in the following chapter, was used to create alignment molds for ex vivo microUS imaging.

**Whole-mount Processing**

Following surgery and ex vivo microUS imaging, the prostate and pathology sectioning mold were delivered to a surgical processing laboratory. A brief overview of the WM pathology processing provided below.

To allow for slide orientation, the prostate capsule is stained in the following manner: right: black, left: blue, posterior: red, anterior: yellow. The prostate is subsequently grossed, removing any fat and seminal vesicles on the prostate. After grossing and capsule staining, the prostate is placed into the pathology sectioning mold and sliced to create axial cross-sections of the prostate with a thickness of 4.5mm.
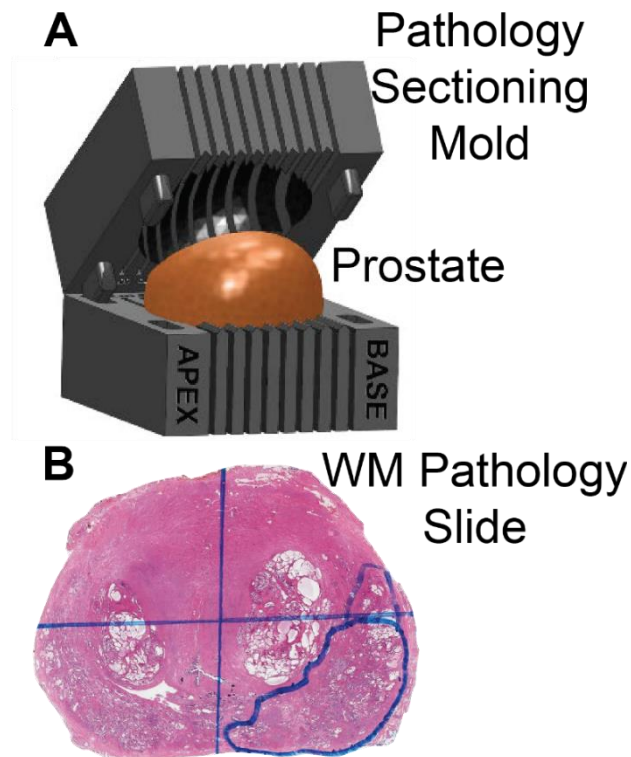


Figure 22: Example of a patient specific pathology sectioning mold (A) used to create WM pathology slides (B) aligned in the axial orientation.

Generally, prostates will have approximately 7 sections. Two of those sections are typically used for genetic processing, as such they are not available for WM processing. The remaining sections are subsequently placed in a microtome to generate thin slides. The slides are stained using hematoxylin and eosin (H&E) for cell and cancer visualization (Figure 22B). The slide generation process typically decreases the overall cross-sectional area by approximately 15-20% due to fixation, staining, and water loss [203]. Beyond that, there is often changes in the prostate shape or loss of tissue during the process of generating histology slides. As a result, any registration between WM pathology and other imaging modalities needs to account for this decrease in area and tissue deformation. Occasionally a microtome of a prostatic cross-section was too large to be visualized on one slide. In these instances, the slide was cut in half from posterior to anterior and represented as two WM slides.

**WM Pathology Review**

Using a microscope, an experienced genitourinary (GU) pathologist delineated the margins of cancer on each WM pathology slide (Figure 22B). The specific Gleason grades of each cancer were also recorded. These slides were then oriented anatomically from apex to base and captured as a reference image for later orientation. Each slide was then scanned digitally at 20X resolution and stored in the imaging database. Similar to microUS, all the WM pathology slides for each patient were converted to a DICOM series with embedded pixel spacing using a custom MATLAB script. This allows for easy visualization of a complete pathology series. For the WM slides that had to be represented by two halves, another custom MATLAB script was used to allow for the images to be stitched together while also embedding the pixel spacing.

## 3.4.   Computer Vision for Prostate Cancer Morphology

While enrollment and data collection were ongoing, we investigated why microUS could be diagnostic for prostate cancer when conventional US is generally non-diagnostic [5]. As discussed in the background section, while early studies with microUS have demonstrated similar cancer detection rates, sensitivity, and specificity as MRI, there is limited explanation or investigation into a rationale for the reported increase in diagnostic capability over conventional ultrasound. Drawing from ultrasound physics, cancer morphology, and the PRI-MUS scoring system for microUS, we created the following hypothesis. The improved diagnostic capability of microUS over conventional US can be explained, in part, by the ability of microUS to resolve prostate ducts, which are 1) large enough to be resolved on microUS but too small to be resolved on conventional US and 2) the presence and size of these prostate ducts is indicative of healthy tissue and the absence of ducts is indicative of cancerous tissue.

As mentioned, the rationale for this hypothesis comes from multiple areas. First the fundamentals of US describe that if a scatter is sufficiently smaller than the wavelength of the US pulse ($d \ll \lambda$), it will scatter in a wide variety of angles (speckle pattern) and appear as noise or texture in the US image [74]. Increasing the frequency of an US probe, decreases the wavelength of the US beam, resulting in small scatters that would be represented as a speckle pattern to generate resolvable reflections that can be visualized on US imaging. As such, conventional US typically has a resolution of 200μm while microUS has a reported resolution as small as 70μm [16]. Next, prostate cancer morphology describes that as Gleason Grade increases there is a general disturbance of benign epithelial-stromal relationships. More specifically, intraductal carcinoma infiltrates the prostate ducts and adenocarcinoma results in gland crowding resulting in
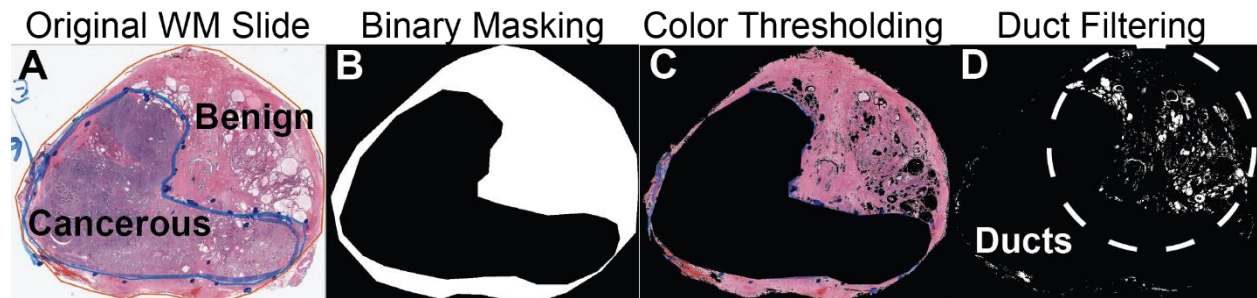
Figure 23: Annotated prostate WM slides (A) are masked into benign and cancerous regions (B). A color thresholding step is then used to separate tissue from ducts (C). The ducts are then extracted and measured (D).

gland compression or, at higher grade, disappearance of prostate ducts [29]. Lastly the PRI-MUS scoring system describes the presence of a ductal pattern in a microUS image (Figure 20) as indicative of benign tissue [18]. While prostate cancer is multi-faceted and its appearance on imaging likely cannot be reduced to the presence or absence of prostate ducts, more complex features of prostate cancer are challenging to measure and correlate between US imaging and pathology, with reduced prevalence across patients compared to the inherent ductal architecture of prostatic tissue. Ultimately, this hypothesis sets forth one possible explanation for why microUS may be diagnostic for prostate cancer and is an anatomical feature that is quantifiable using WM pathology data.

## 3.4.1. Methods

Ten prostatectomy prostates were grossed, sectioned, stained (H&E), and all clinically significant PCa lesions were annotated by an experienced pathologist following routine clinical practice. The scanned slides were then analyzed using a custom computer vision script (MATLAB) to identify and quantify ductal features (Figure 23). Briefly, using the annotated lesions, binary masks were made to delineate benign and PCa groups. A color thresholding step was then applied to the tissue within each mask to select the white space where the prostate ducts are located. Two main features were extracted: the duct
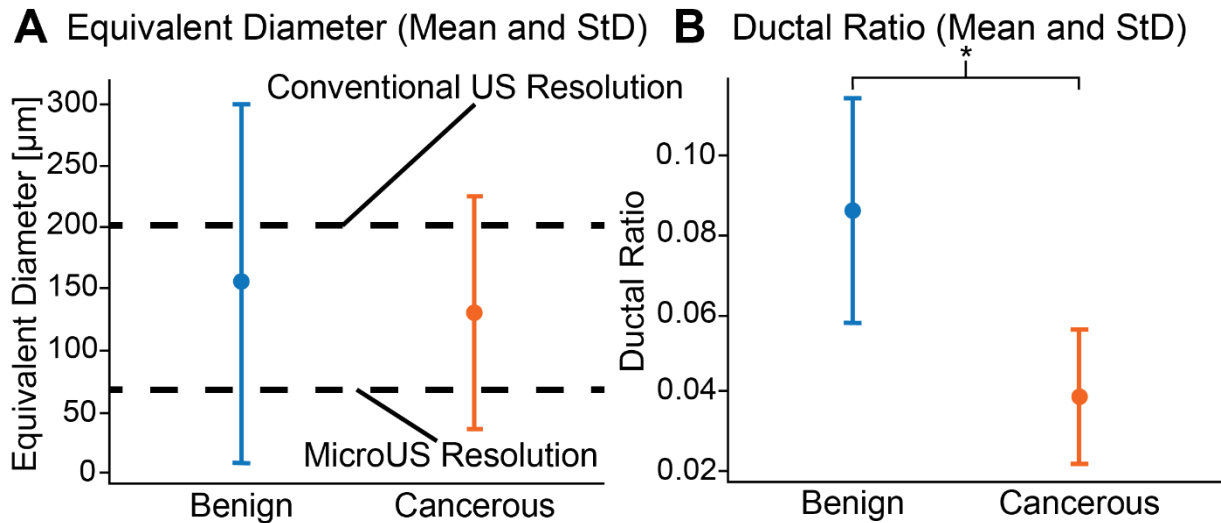
Figure 24: A) The equivalent diameter for benign and cancerous prostate ducts. Note, the resolution of conventional US is insufficient to visualize most of the ducts. B) The ductal ratio for cancerous tissue is approximately half that of benign tissue suggesting utility of ductal ratio for cancer identification (*p-value << 0.001 for two-sample t-test)

equivalent diameter (the diameter of circular region with the same area as the duct), and the area of all the ducts in a region divided by the area of the region (ductal ratio).

### 3.4.2. Results

A total of 54 slides were processed with 145,150 benign ducts identified. The mean equivalent diameter of benign ducts was 155.7 ± 147.8µm (Figure 24A). Additionally, 9,245 PCa ducts had a mean equivalent diameter of 129.4 ± 94.3µm. The average ductal ratio for benign tissue was 0.0863 ± 0.0289 and 0.0392 ± 0.0169 for PCa tissue (Figure 24B).

The average equivalent diameter for both groups (benign: 155.7µm, PCa: 131.8µm) is smaller than the minimum resolution of conventional US (200µm), but large enough to be visualized under microUS (resolution: 70µm). Only 18.3% of benign ducts and 12.4% of PCa ducts were larger than 200µm. Furthermore, PCa tissue exhibits an

approximate 50% reduction in ductal ratio indicating a reduced presence of ducts in regions of cancer relative to benign tissue.

## 3.5.  Discussion

Eighty-six patients were enrolled in a clinical trial allowing for the creation of a high-quality PCa dataset for microUS evaluation against mpMRI and ground truth WM pathology. This dataset included a multitude of clinical variables, an mpMRI, in vivo microUS, ex vivo microUS, and mold-sliced WM pathology. Ultimately, this dataset is distinguishable from others due to effort for preserving positional information in the related modalities. Specifically, all modalities, apart from in vivo microUS, are acquired in the same orientation as mpMRI, the gold standard diagnostic. In vivo microUS is supplemented by ex vivo microUS providing an alternative in the instance in vivo cannot be accurately registered to WM pathology. Furthermore, this dataset is validated through mold-sliced WM pathology ground truth, the most comprehensive final histology available for prostate cancer. The collected data provides the information needed to objectively evaluate microUS imaging, as well as accurate data labeling of microUS images for model training.

The results from the computer vision study provide evidence in support of our proposed hypothesis, demonstrating microUS has the potential to stratify regions of benign tissue from regions of cancerous tissue through the appearance of prostate ducts. As mentioned, this is by no means an all-encompassing explanation for prostate cancer visualization on microUS but it does identify a microUS visible contrast mechanism possibly allowing for the increased diagnostic ability of microUS relative to conventional US. It would follow that there are additional morphological markers of

prostate cancer that are unidentifiable on conventional US and resolvable on microUS, however quantification of these features may require hundreds of patients to identify and quantify due to inconsistent presentation across patients. As a result, such an investigation is outside the bounds of this thesis, but represents a possible avenue for future microUS-histology research.

## 3.6.    Conclusion

MicroUS has the potential to be a low-cost alternative to mpMRI imaging of prostate cancer. However, the majority of clinical studies evaluating the diagnostic ability of microUS have relied on biopsy data for validation which do not provide the true margins of cancer extent. In this chapter we detail data collection from a clinical trial that will provide the necessary information to robustly evaluate the diagnostic ability of microUS. With spatially tracked mpMRI, microUS, and WM pathology data, this high-quality dataset provides the information necessary to accurately evaluate the diagnostic accuracy of both microUS and mpMRI. To support the rationale for this investigation, using computer vision and WM pathology data, we identified quantifiable evidence for the diagnostic capability of microUS through its visualization of prostate ducts, which inversely correlate to cancer. In this chapter, we have collected the necessary data for the rest of this thesis as well as provided supporting evidence justifying continued investigation into microUS imaging of prostate cancer.

# Chapter 4

# Ex Vivo Micro-Ultrasound Imaging of Prostate Cancer

## 4.1.    Introduction

Recently, microUS has been introduced for prostate cancer diagnosis. Commercially available machines operate at a much higher frequency (29MHz) than conventional US probes, while maintaining an imaging depth of 6cm, allowing for a superior spatial resolution of approximately 70μm versus 200μm for conventional ultrasound [16]. The matching of microUS readings to 400 tracked biopsies led to the development of the Prostate Risk Identification using Micro-Ultrasound (PRI-MUS) grading scale for detecting and grading prostate cancer [18]. In preliminary studies microUS has performed similarly to MRI and could potentially provide a low cost alternative to MRI for targeted biopsies [16].

MicroUS has demonstrated potential for prostate cancer diagnosis, however, it lacks validation to WM pathology, the gold standard. MicroUS utilizes a side fire probe, therefore, in vivo imaging is achieved by sweeping the ultrasound plane across the prostate in a para-sagittal, fan-shaped pattern. Since WM pathology is acquired in the axial orientation, computational registration is challenging even with highly-controlled scanning setups [20]. In this chapter we detail the development and validation of a scanning system that facilitates accurate in-plane co-registration of ex vivo microUS with

WM pathology and MRI. We then use this scanning system to compare the diagnostic ability of ex vivo microUS and MRI against WM pathology in a pilot clinical study.

## 4.2.    System Design

The ex vivo scanning system was designed to interrogate a surgically removed prostate under US imaging in the axial orientation from apex to base (Figure 25A). The system is mounted to an optical board and designed to scan at a controlled speed allowing for position tracking. The three main components of the scanning system are: the probe holder, the scanning tank, and the imaging alignment molds. There is also an image processing component following data collection to convert the ex vivo microUS scans into single DICOM series. Each of these components are covered in more detail in the following sub-sections. The ultimate goal of the scanning system is to acquire microUS images in the same orientation as WM pathology and mpMRI are acquired allowing for co-registration between the modalities. An ideal system would provide clear images of a
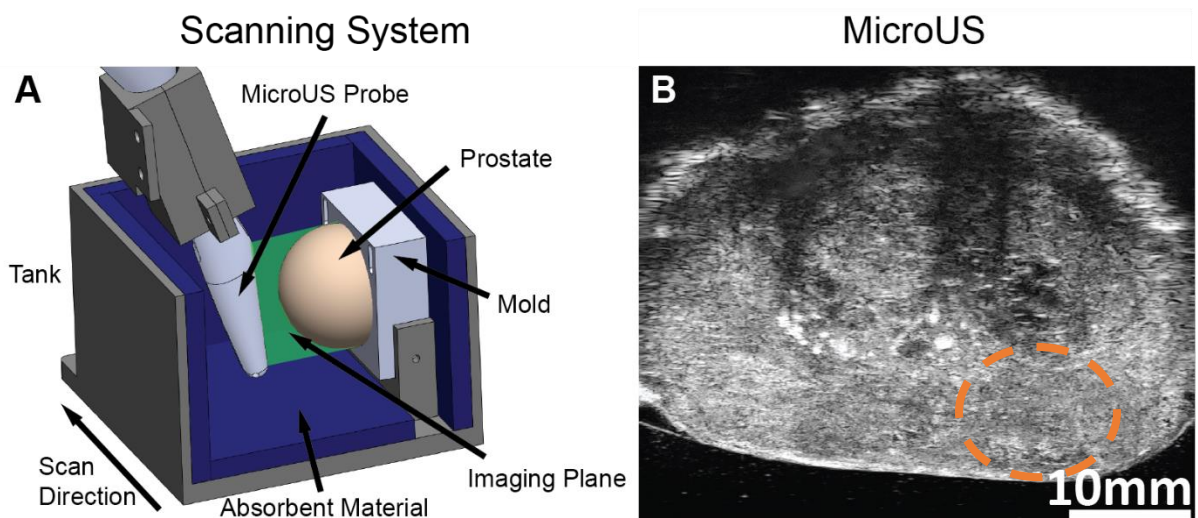


Figure 25: A) Ex vivo scanning system design used to scan the prostate axially from apex to base and B) an example image from the ex vivo scanning system. There is pathology confirmed prostate cancer (orange circle) in the left peripheral zone (lower right region of the image).

prostate with minimal noise artifacts. Through iterative development, we were able to refine the US scanning system to provide quality images that ultimately allow for evaluation of microUS imaging of PCa (Figure 25B).

## 4.2.1. Probe Mounting and Translation

The microUS probe needed to be translated at a controlled speed from the apex to the base of the prostate. To achieve this a mechanical translation stage (X-LSQ300B, Zaber Technologies) was mounted to an optical board. The translation stage used is a linear stage with a built-in controller and supported Python application programming interface (API). The stage allows for 300mm of travel, a step size of 50μm, and a maximum load of 500N allowing for adequate scan distance and resolution while also being able to support a microUS probe, the related mounting components, and any induced forces during scanning. Using the API a custom Python script was written to control the translation stage. The script provides functionality to move to the mid-gland of the prostate prior to scanning to allow for image optimization of probe position and image gain. Beyond that it allows for repeated scans from the apex to the base at a translation speed of 2.75mm/second or 0.65mm/second giving a respective imaging spacing of 0.275mm/frame or 0.065mm/second due to the microUS system having a framerate of 10 frames/second. Due to limitations of the microUS system acquiring a maximum of 300 frames per cine video, the faster translation speed completes the full travel scan in just under 30 seconds allowing the operator time to save the sweep after the scan is completed. The slower scan speed follows the same pattern splits the full travel distance into 4 periods allowing for scan acquisition in between each scan. For conventional US
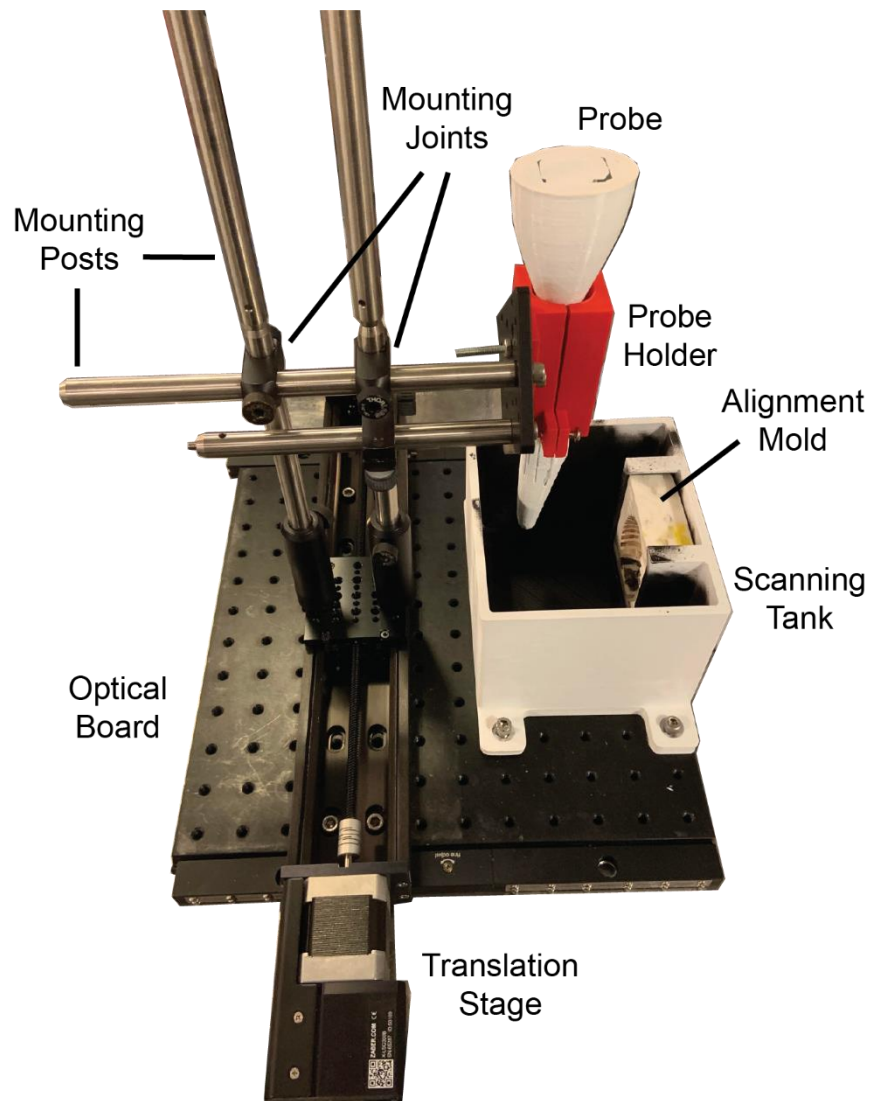
Figure 26: A built version of the ex vivo scanning gantry used for preliminary testing. The alignment mold, probe holders, and scanning tank were updated in later iterations.

scanning, the Hitachi system allows for custom scan lengths so the entire scan was performed as a continuous 0.65mm/sec translation and saved.

In order to align the probe for scanning it needed to be appropriately mounted to the translation stage (Figure 26). Two vertical mounting posts were attached to the translation stage attached to two horizontal mounting posts via mounting joints. The raised horizontal mounting posts were connected to a small optical board allowing for

93

custom designed US probe holders to be attached. The US probe holders were designed to align the US probe in the axial orientation and utilizes unique features on the neck of the probe to prevent rotation or misalignment. To create these models a 3D computer-aided design (CAD) model of the microUS probe is used to extract the probe volume from a holder template file. The CAD model is then split into two components with screw connection points to secure them to each other with the probe in the middle. The CAD models were then 3D printed and screwed into the vertical optical board on the mounting system. This process was followed to create a custom holder for the conventional US probe. The initial microUS holder was designed without an angle offset, however the microUS probe has a canted design at the location of the transducer to ease
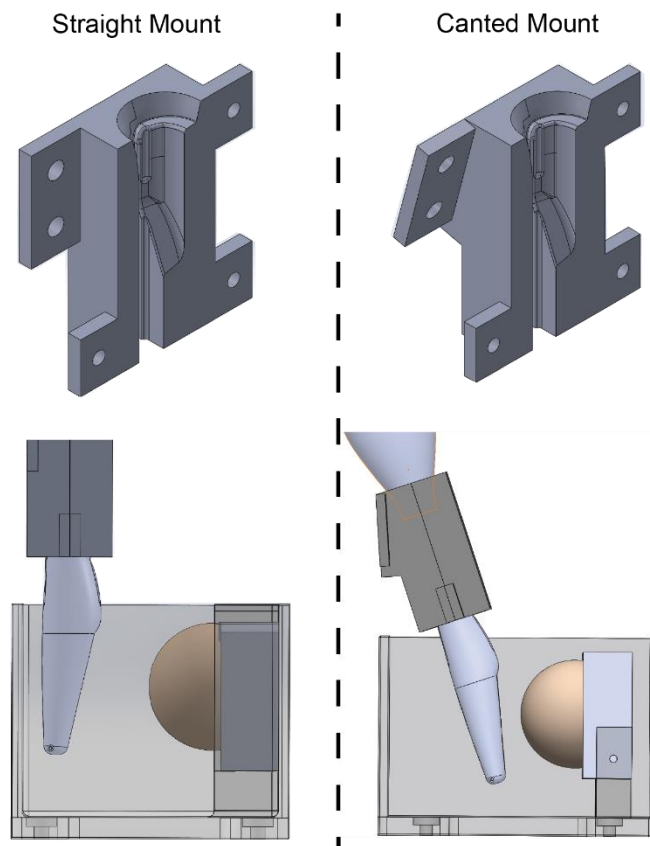


Figure 27: Initial straight mount design (left) resulting in a transducer angle mismatch with the prostate. Adding an offset of 30 degrees in the canted mount (right) allows for straight on imaging from the transducer.

transrectal insertion. As such to allow for the microUS probe to sit parallel to the prostate and scan tank wall, it was necessary to add an angled offset to the microUS probe holder of approximately 30 degrees (Figure 27).

## 4.2.2. Scanning Tank

To allow for adequate interrogation of prostate tissue, we needed to create a conducive ultrasound imaging environment. In the preliminary proof of concept efforts, we designed and printed a 15x13x11cm PLA tank with a housing for the imaging alignment mold. The tank was coated with a silicon rubber spray to increase the absorption of the US beam. For agar phantom validation studies, the tank was filled with water to allow for US interrogation, but for prostate studies the tank was filled with 0.9% saline for an isotonic solution to avoid osmotic tissue degradation during scanning.

While these scans allowed for adequate interrogation of the prostate, we continued iterative development over the course of the clinical trial to further improve scan quality. Specifically, PLA allows for easy prototyping due to its ability to be 3D printed, but creates considerable US reflections as a rigid plastic material. To decrease ultrasonic reflections, the box was redesigned to allow for room for US absorbent material. The absorbent material (AptFlex F28P, Precision Acoustics) is a micro-bubble filled sheet made of a polyurethane with small pyramidal structures that decrease specular reflections and increase the number of reflections required to transmit back to the transducer. The material has a density and speed of sound similar to water to promote absorption. All of this results in an echo reduction of approximately 40dB. All of the surfaces in the US scanning tank were lined with this material (Figure 25A). To further reduce the amount of PLA exposed to the US probe, the size of the housing used

to secure the alignment mold was reduced and instead was secured in place by a small screw.

### 4.2.3. Patient Specific Scanning Molds

To correctly orient the imaging molds during scanning we utilized patient-specific imaging alignment mold. Similar to the molds used for WM pathology sectioning, the molds are designed to orient the prostate to allow for axial scanning and are constructed following the same process. As detailed in chapter 3, to create the mold the prostate shape is delineated on MRI and extracted from a template file that is subsequently 3D printed. For imaging alignment, only the anterior half of the mold is needed. This orients the prostate with the peripheral side facing the transducer allowing for the correct imaging plane.

To further improve the ex vivo scan quality, two different types of imaging alignment molds were tested. One mold was a traditional pathology slicing mold made of PLA and the other was made of thermoplastic polyurethane (TPU). An agar phantom (4%w/v) was scanned in the HRUS scanning gantry using both molds. The subsequent image stacks were compared qualitatively and quantitatively to assess the impact of the material properties on the imaging artifacts, full width at half maximum (FWHM) at the edge of the mold, pixel intensities, signal to noise ratios (SNR), and contrast to noise ratios (CNR). A superior mold generates less noise and has sharper edges, generally characterized by a short FWHM, high SNR and CNR, and low background pixel intensities.
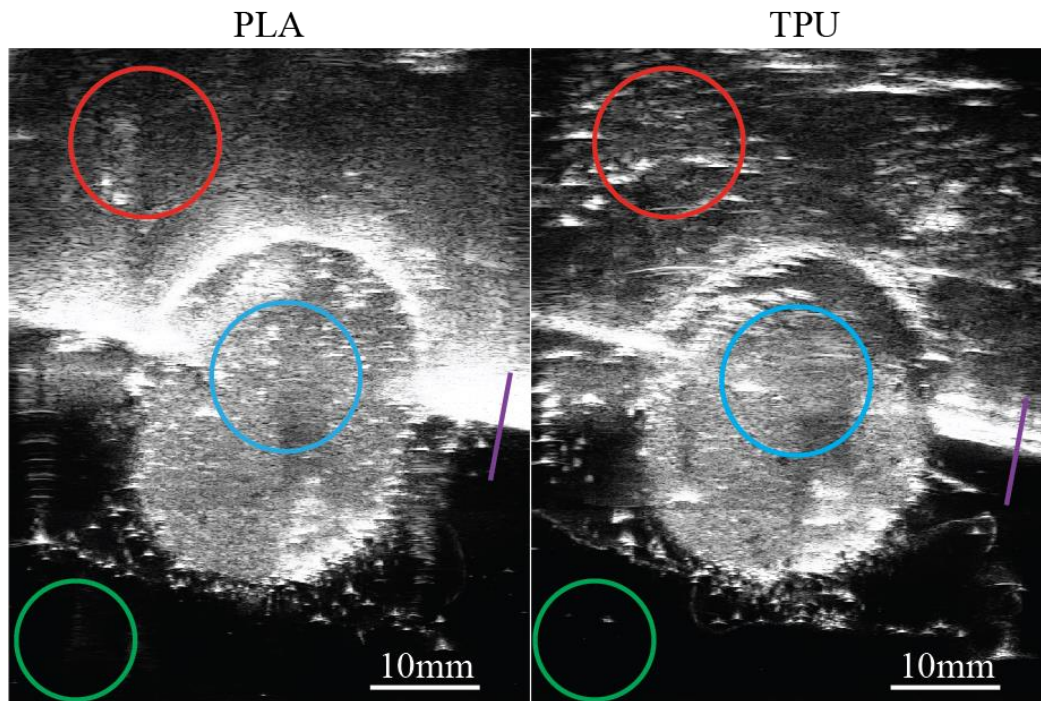
Figure 28: Comparison of PLA and TPU phantom images. (Red) Background noise within the mold. (Blue) Signal from the phantom. (Green) Background noise in the water. (Purple) FWHM across the interface between the mold and the water.

A total of 477 matched images were compared between a phantom imaged in both PLA and TPU molds (Figure 28). Matched background pixels (N = 96,000) were compared between PLA and TPU with the former exhibiting a higher average pixel intensity by 27.5±54.8. From a sample of 50 images for each mold, the FWHM was found to be 4.6mm±1.1mm for PLA and 1.8mm±0.4mm for TPU, indicating the TPU molds have sharper edges under HRUS. For a sample of 100 images, the SNR was found to be 84.2 and 97.8 for PLA and TPU respectively. The CNR for PLA was 43.3 and 72.5 for TPU. These data indicate that TPU molds result in fewer imaging artifacts, sharper edges, better signal and contrast to noise ratios. Due to the superior imaging quality, the TPU mold was chosen for imaging ex vivo human prostates. Beyond this, the flexible nature of TPU molds allows for the prostate to be sutured to the imaging mold to help secure it for imaging.

97

PLA Mold

TPU Mold
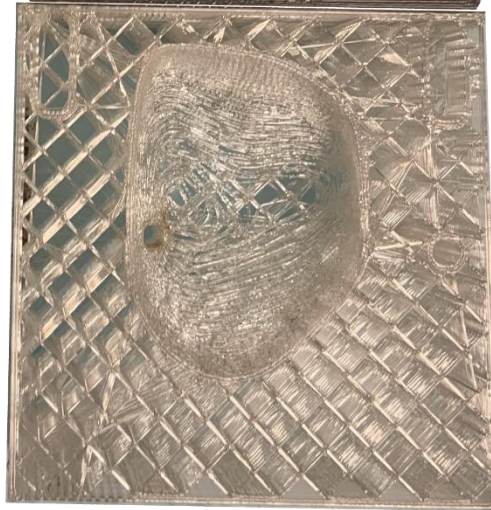
Figure 29: Example of the anterior half of the PLA pathology sectioning mold and TPU imaging alignment mold. The flexible TPU provides a superior acoustic environment to the more rigid PLA.

To allow for equal comparison between materials in this test, the TPU mold was designed and printed the same way as the PLA pathology sectioning molds. After confirming the more flexible TPU molds allow for higher quality imaging, the TPU imaging alignment molds were further altered to use less material resulting in less opportunities for reflection (Figure 29). The pathology sectioning mold needs to have slits to allow for the cutting blade, but this is not required for ex vivo imaging alignment. As such the mold template was redesigned to not have these slits. The reduced volume

of walls within the print as well as utilizing a smaller infill percentage (10%), helped diminish US reflections further.

The last iterative development to improve ex vivo US imaging was the implementation of an US stand-off (Aquaflex, Parker Laboratories). An US stand-off was implemented to help with near-field visualization of the prostate, specifically the peripheral zone. Using a custom designed stand-off slicer mold, the stand-off was cut in
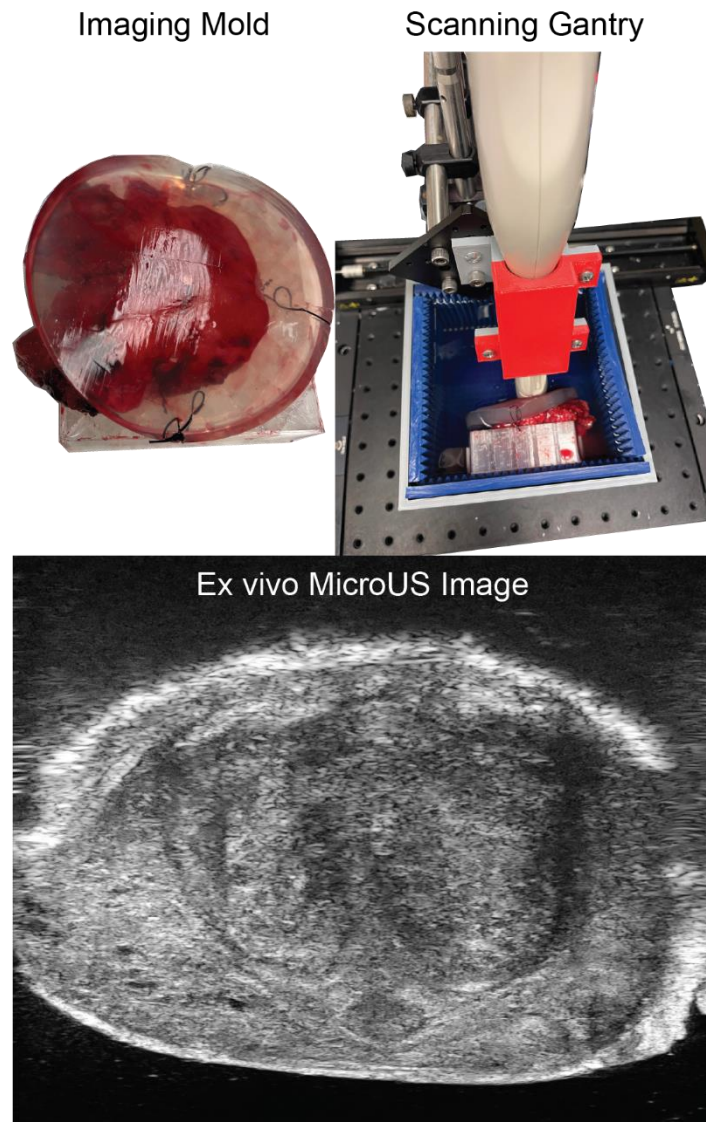


Figure 30: Ex vivo imaging mold, scanning gantry, and resulting image. The high-quality scans allow for visualization of intraprostatic features

half to reduce its thickness. It was then sutured to the imaging alignment mold on top of the peripheral zone of the prostate. Ultimately all of the iterative develops allowed for high quality ex vivo imaging of human prostates following radical prostatectomy (Figure 30).

## 4.2.4. Image Processing

Following ex vivo scan collection, the files needed to be extracted from the microUS machine and converted into a usable file format. Each microUS sweep is saved as a video (.bmp) file on the US machine. We created a custom MATLAB script to convert the video file to individual image files. Each frame is turned into a DICOM image with embedded pixel spacing. Each DICOM image is assigned a position based on the translation speed of the scan and the number of images since the start of the scan. However, since the microUS video is saved as the prior 30 seconds and is not linked to the movement of the translation stage, there is a period at the beginning of each scan where the probe is not moving. To account for this, the first image in the series is compared to consecutive images following it using image similarity metrics. Once a substantial difference between frames is detected, indicating the transducer is moving, the preceding frames are trimmed from the series. A script also takes the time a scan was taken and uses it to stitch the four consecutive scans together automatically. The resulting DICOM series are subsequently available for review, delineation, and analysis.

# 4.3.  Co-Registration

Following development of the scanning system, it was necessary to validate its correspondence with WM pathology and MRI. The purpose of the ex vivo scanning system is to compare microUS with WM pathology and MRI, and in order to do this a co-registration process needs to be developed and validated. To this end, the target registration error (TRE) of the system was evaluated in two studies: a phantom study with embedded fiducials, and a prostate fiducial study using internal prostate landmarks.

## 4.3.1.  Phantom Study

First, the system and co-registration were validated in a phantom study. To achieve this 5 prostate mimicking agar phantoms (4%w/v) were created following a process from a prior study [167]. Briefly, warm agar solution was poured into a prostate mold with needles oriented in specific positions throughout the prostate and placed in a refrigerator to cool. Once the agar set, the needles were removed leaving behind tracks that serve as fiducial markers. The prostate phantoms were then removed and placed in imaging alignment molds and scanned in the gantry with the microUS probe. The phantoms were subsequently sliced axially in the full pathology mold and scanned on a conventional desktop scanner (MFC-9340CDW, Brother Industries, Ltd. Japan) at 600dpi to mimic WM pathology slides. Using a DICOM viewer (Horos, Horos Project), both the microUS and desktop scan of the phantom were segmented and the fiducials were marked (Figure 31). With the prostate capsule as the basis for co-registration, the segmentations were then co-registered via a custom MATLAB script. This was performed through a rigid transformation of the sliced phantoms to account for scaling and in-
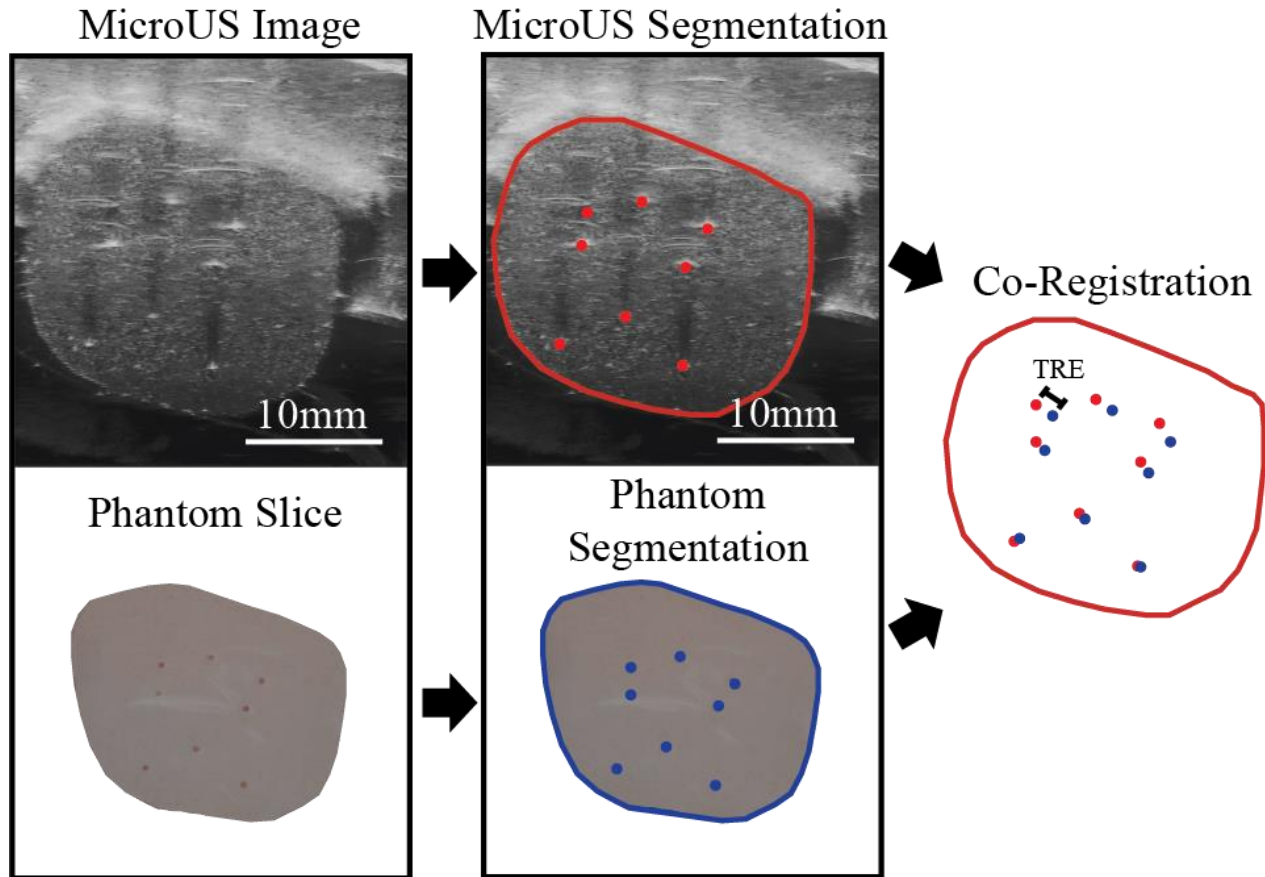
Figure 31: Phantom study overview. MicroUS phantom images were matched to phantom slices based on recorded scan and slice locations. The distance between marked needle track fiducial pairs were measured following co-registration.

plane rotational misalignment, followed by a non-rigid thin-plate spline transformation using control points on the prostate capsule to account for non-rigid specimen deformation as described by Fei *et al.* [155]. The in-plane distance between matched fiducials on microUS and desktop scans was recorded for each fiducial pair and used for determining in-plane TRE.

From the 5 agar phantoms, a total of 294 matched fiducial pairs were compared from 40 phantom slices (Figure 31). The average in-plane TRE was 1.9mm±1.4mm. This is similar to the TRE observed when co-registering axial MRI to WM pathology (1.9mm±0.6mm) [167].

## 4.3.2.  Prostatic Fiducial Study

Following a similar process as the phantom study, imaging data from four prostates from the dataset detailed in chapter 3 were used for this evaluation. Specifically, following surgery, the prostate was placed in a patient-specific mold, submerged in saline in a custom imaging tank, and imaged in the axial plane with a microUS probe from apex to base using a translation stage. The prostate was then placed in a 3D printed pathology mold and sectioned to produce slides in the same orientation as the MRI and microUS scan, as described in chapter 3. Digitally scanned histological slides at 20X magnification were created from each specimen. Matched corresponding images from
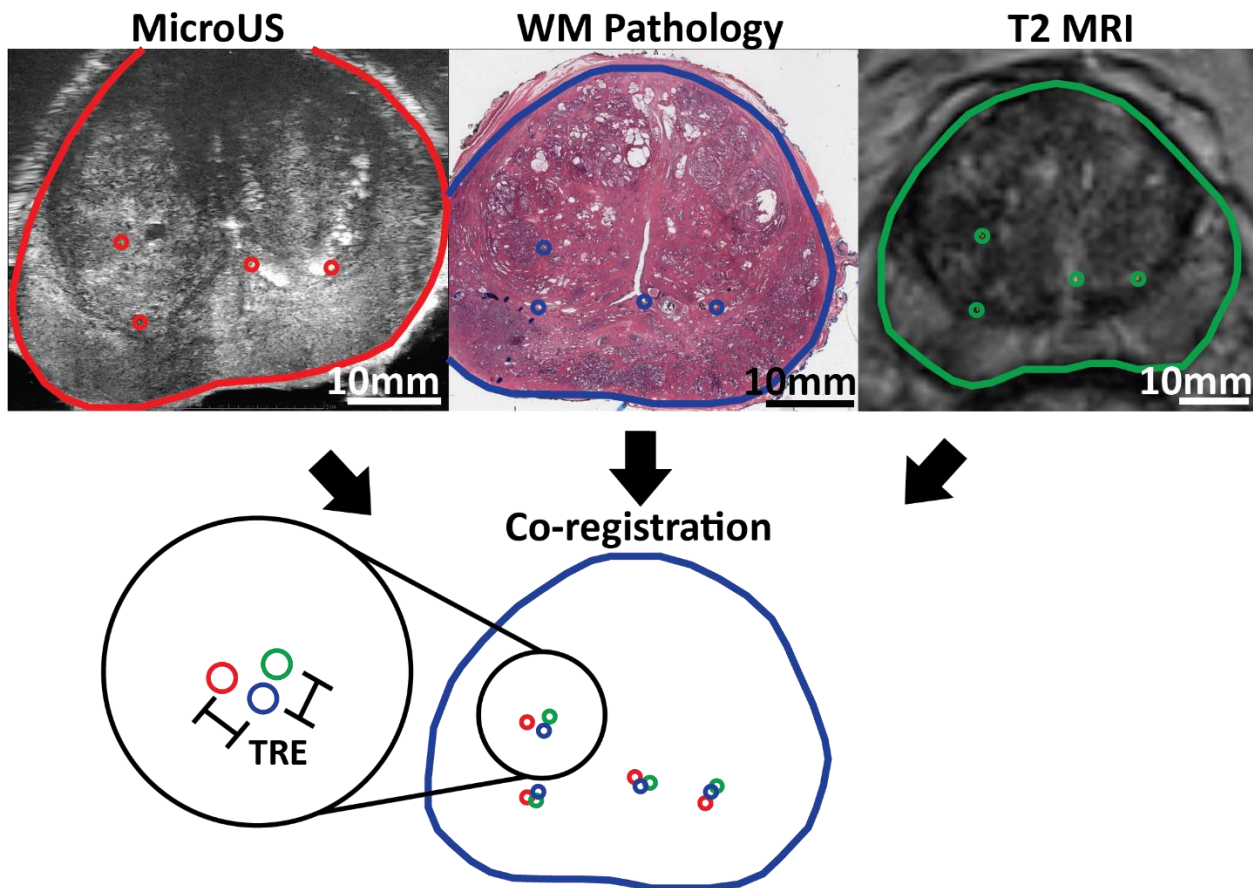


Figure 32: Co-registration of microUS (red), WM pathology (blue), and T2 MRI (green). The four corresponding landmarks are marked as circles and the distance between microUS and MRI from the corresponding WM landmark is used to quantify the target registration error.

both T2 MRI and ex vivo microUS were chosen for each histology slide. For each slide the prostate capsule was segmented and at least four anatomical landmarks (urethra, BPH nodules, cysts, ejaculatory ducts) were marked that were visible in all 3 image sets. Using only the prostate capsule segmentations, a rigid registration step followed by a thin-plate spline (TPS) non-rigid registration step was performed to co-register each set of image triplicates (Figure 32). The in-plane TRE was calculated for each landmark.

Matched microUS and MRI images were co-registered to a total of 22 pathology slides. A total of 88 anatomical landmarks were evaluated. Based on previously published results, the out of plane error between MRI and microUS matched with WM pathology slides is expected to be around 1.5-3mm [175]. The TRE following TPS registration was 2.16mm ± 1.02mm for microUS and 2.34mm ± 1.35mm for MRI demonstrating the developed co-registration process facilitates accurate co-registration between microUS, MRI, and WM pathology.

## 4.4. Pilot Clinical Study

While the ex vivo scanning system has been validated for registration accuracy in the prior sections, ex vivo scanning is in a different imaging environment than the in vivo prostatic environment. As such, it is crucial to determine if it is possible to identify PCa on ex vivo microUS imaging. If an expert microUS reader is able to identify PCa on ex vivo microUS with similar performance to mpMRI, the ex vivo scanning system and imaging environment can be deemed sufficient quality for microUS evaluation. Furthermore, this will also provide promising evidence for the overall diagnostic potential of microUS. However, it is worth stating that the ex vivo imaging environment is not indicative of an in vivo imaging environment, and as such there is the possibility

it is more favorable for microUS visualization of PCa than an in vivo environment. For this reason, an eventual in vivo evaluation of microUS is preferrable. If it is not possible to achieve accurate registration, however, ex vivo imaging could potentially provide an adequate alterative. To confirm this, we performed a pilot clinical study to assess the diagnostic accuracy of ex vivo microUS imaging for PCa and the quality of the ex vivo imaging environment.

All research involving human subjects was approved by the Institutional Review Board. Three patients set to undergo radical prostatectomy were consented and enrolled in the study. Using a pre-operative MRI scan with annotated suspicious lesions, custom PLA and TPU molds were printed for prostate slicing and imaging respectively. Following radical prostatectomy and prior to fixation, the prostate was placed in the anterior half of a TPU mold, secured via a suture, and imaged under microUS in the ex vivo scanning gantry in the same manner as the phantom scans. After scanning, the prostate was placed in the PLA pathology mold and sectioned axially following routine clinical protocol. Histology slides were made from each slice and cancerous lesions were annotated by an experienced pathologist. The annotated slides were subsequently scanned at high resolution (20x). Following a similar process to the phantom experiment, the histology slides were then cognitively compared to the microUS scan and pre-operative MRI to find matched image triplicates. The prostate and visible lesions were identified and segmented by separate expert blinded observers for the microUS scan, MRI scan, and pathology slides. Following metrics used in similar studies [204], [205], the overlap percentage, DICE similarity coefficient (DSC), and centroid distance were calculated for each co-registered ROI triplicate. This allows for preliminary analysis of the ability of microUS to identify PCa, using MRI imaging as a benchmark.

From 3 subjects, a total of 14 microUS, MRI and pathology image triplicates were matched and computationally registered together (Figure 33). Lesions were present on 11 of the pathology slides with a total of 20 lesions. Of these 20 lesions 13 were identified by the microUS observer with only 4 identified on MRI. Overall, 26 lesions were delineated on microUS with 18 true positives and 8 false positives. For MRI, 4 true positive lesions were delineated. All 3 index lesions were identified on microUS and 2 out of 3 were identified on MRI. The average distance between the centroids of the true positive microUS ROIs and the actual lesions was 10.8mm±7.2mm with 11 out of 26 of the ROIs having a centroid distance from the WM ROI of less than 1cm. For MRI the average centroid distance was 5.8mm±2.5mm with all 4 ROIs having a centroid distance
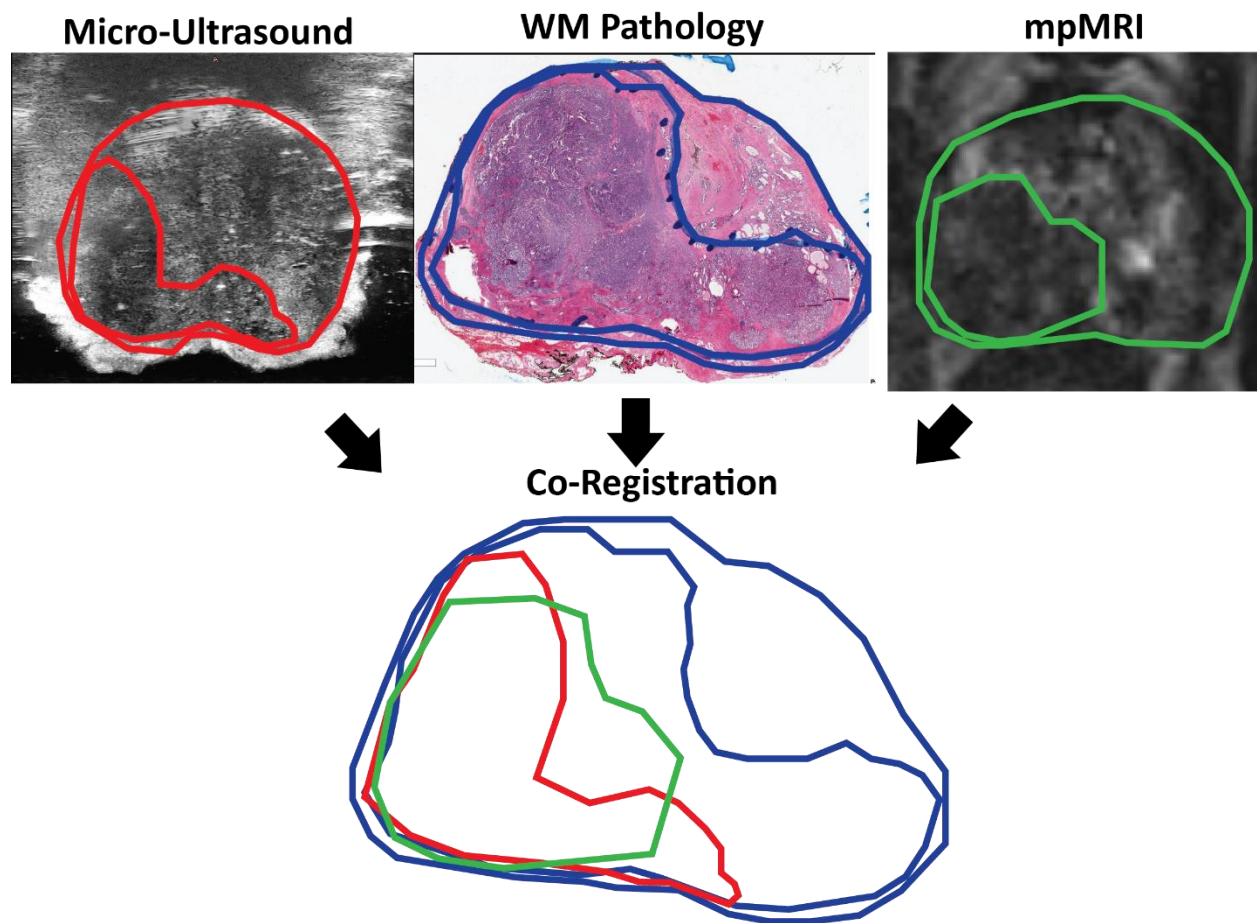


Figure 33: Co-registration of blinded expert delineated regions on ex vivo microUS, WM pathology, and mpMRI.

of less than 1cm from the WM ROI. The average overlap percentage and average DSC of the true positive microUS and WM ROIs was 54.3%±33.1% and 0.33±0.21 respectively. While for MRI the average overlap percentage was 85.1%±17.6% and DSC was 0.61±0.05.

## 4.5.　Discussion

The goal of this chapter was to develop and validate an ex vivo scanning system to allow for in-plane co-registration of microUS and MRI images with WM pathology, for eventual comparison of the diagnostic accuracy of microUS versus MRI imaging.

The ex vivo scanning system was specifically designed to interrogate surgically excised prostates in the axial orientation, allowing for comparison with WM pathology and MRI. The system was developed and iteratively improved to create high quality microUS images. Implementation of TPU molds offered superior imaging quality to PLA. This is likely due to the reduced material stiffness of TPU resulting in a lower acoustic impedance difference at both the water-mold interface and the phantom-mold interface. Furthermore, use of US absorbent material reduced scattering noise and US stand-offs allowed for better interrogation of near field imaging.

Phantom experiments demonstrated that the developed scanning system facilitates in-plane co-registration similar to published values for MRI with a TRE of 1.9mm±1.4mm. This error is acceptable given that the average biopsy needle deflection is 0.9mm and the average registration error for MR-US fusion phantom studies typically report a TRE of 2-3mm [21]. However, the phantoms are more rigid than prostate tissue, so there may be a larger TRE than was found with the phantoms from increased deformation when sectioning. To account for this, we performed a similar study using prostatic fiducials after confirming the system had adequate registration with agar

phantoms. In which microUS and MRI had average TRE values of 2.16mm ± 1.02mm and 2.34mm ± 1.35mm respectively. This represents a slight increase in TRE over the phantom studies, as expected, but is still within the typical reported TRE range of commercial fusion systems.

Overall, the ex vivo scanning system allowed for detection of prostate cancer with 65% of distinct lesions and all 3 index lesions correctly identified under microUS. The index lesion is the largest and most aggressive tumor and believed to be responsible for the progression of prostate cancer making its detection vital for proper disease management [43], [44]. Additionally, 6 out of the 7 lesions that microUS failed to detect, were small lesions with a cross-sectional area less than 40mm2. The only large lesion that was missed by microUS was located in the far anterior of the prostate, indicative of a weakness of microUS as there is a distinct loss of resolution in the far anterior of the prostate due to shadowing, however this lesion was also missed under MRI.

The pilot clinical study demonstrated the scanning system and registration software presented here are allow for visualization of PCa and co-registration of ex vivo microUS with MRI and WM pathology. Ultimately, this facilitates comparison of the clinical utility of microUS and MRI using WM pathology as ground truth. However, it should be noted that microUS scans were performed ex vivo and is not indicative of a clinical scan, while MRI was performed in vivo. In an in vivo setting there are anatomical structures which can potentially degrade image quality; however, ex vivo also has structures that could also degrade image quality such as the TPU mold and tank.

The clinical data present here highlight the promise of microUS as an alternative to MRI for prostate cancer detection as all 3 index lesions were correctly identified. However, this was only an exploratory study with a small sample size (N=3), therefore,

accurate metrics for the sensitivity and specificity of microUS cannot be discerned. Further work is planned to increase the number of patients and ultimately demonstrate the clinical utility of microUS.

## 4.6.    Conclusion

The scanning system presented in this chapter facilitates high quality ex vivo microUS images and allows for accurate in-plane co-registration of with WM pathology slides. As demonstrated in the pilot clinical study, this process enables rigorous assessment of the utility of microUS for prostate cancer detection. The ability of an expert reviewer to identify PCa on ex vivo microUS is a testament to the quality of ex vivo images as well as the promise for microUS as a potential alternative to MRI. The results from this chapter provides added confidence microUS has potential to identify PCa and this investigation is worth pursuing further.

# CHAPTER 5

# Evaluation of Native Micro-Ultrasound and MRI Imaging

## 5.1. Introduction

MicroUS is a promising new modality, however it has not been robustly evaluated against mpMRI in clinical studies validated by ground truth histology. Previous studies validated with tracked biopsy have demonstrated similar cancer detection rates between microUS (35%) and mpMRI (39%) with comparable sensitivities (94% vs 90%) and specificities (22% for both) [16], [17], [18]. Although these studies have demonstrated promise for microUS imaging, the gold standard evaluation for prostate imaging is validation against ground truth whole-mount (WM) pathology slides. Unfortunately, such a study is challenging because it is difficult to co-register imaging planes with pathology slides, because the former are acquired in a para-sagittal sweep and the latter are axial. Moreover, probe rotation is performed manually and thus prone to inconsistent tissue deformation and untracked imaging plane translations. Consequently, most studies comparing microUS and MRI have relied on cognitive registration of microUS and pathology through placement of an ROI into one of 12 corresponding template zones within the prostate [191], [206]. Although this approach is sufficient for assessing general consensus between modalities, cognitive registration is limiting for the following reasons: 1) Zonal placement is subject to reviewer bias and expertise as correlation between non-

orthogonal, and non-parallel microUS and WM image planes require significant cognitive effort and is prone to error 2) Registration error and overall quality of cognitive registration is not necessarily consistent across cases and reviewers, thus, specific ROI-lesion agreement metrics such as dice scores and overlap percentages cannot be accurately evaluated 3) Image masks created from cognitive registration are weak labels and can diminish the accuracy of predictive AI models for microUS, which would help facilitate widespread adoption of microUS.

Here, we demonstrate computational registration of microUS and MRI to WM pathology, with measurable registration accuracy and minimal bias, thus enabling rigorous assessment of both imaging modalities. In addition, this process facilitates generation of high quality, strongly labelled datasets which will pave the way for the development of deep learning models capable of automatically identifying cancerous lesions on microUS. We believe such models are a necessary prerequisite to widespread adoption of microUS as cognitive analysis of the images is challenging particularly in the biopsy setting where time is often limited. In earlier work, our group has validated an ex vivo microUS scanning setup and found comparable results between MRI and ex vivo microUS for prostate cancer identification through computational co-registration with WM pathology [166]. In this work, we build upon our approach with a validated, multi-step registration process for accurate co-registration of in vivo microUS and mpMRI with WM pathology. This approach is then utilized for quantitative analysis and comparison of microUS and mpMRI for blinded cancer detection in a clinical study with 15 patients. The goal of this study is two parts: 1) validate co-registration of microUS and MRI with ground truth WM pathology and 2) determine the diagnostic accuracy of microUS and mpMRI in a pilot clinical study.

## 5.2.  Methods

A general overview of the three-step registration process is as follows (Figure 34): After para-sagittal microUS images, axial MRI images, and axial WM pathology slides are obtained 1) The para-sagittal microUS images are reconstructed and interpolated to align with the MRI axial planes 2) The reconstructed microUS images are co-registered to the MRI via a 3D registration process based on common landmarks 3) The newly co-registered microUS and MRI datasets are registered to WM pathology via 2D capsule registration. The end result is a series of 3 transformation matrices that transform a point in para-sagittal microUS space to "true" sagittal microUS space, then to MRI space, and lastly to WM pathology space. Later in the pilot study, reviewers delineate regions suspicious for cancer on native microUS and MRI and the aforementioned matrices are retrospectively applied to register both datasets to ground truth WM pathology.

### 5.2.1.  Data Collection

Fifteen patients set to undergo radical prostatectomy for cancer treatment gave informed consent and were enrolled into a registered clinical trial (NCT04299620) and UCLA Institutional Review Board approved study (IRB#19-001136). To be eligible for enrollment, each patient needed to have an mpMRI of their prostate with an associated UCLA radiologist report within a year prior to surgery, biopsy confirmed prostate cancer, and no prior treatment. Any radiologist identified PI-RADSv2.1 (Prostate Imaging Reporting and Data System) score 3 or higher suspicious regions were exported from the mpMRI and stored.

In the operating room on the day of surgery, each patient received at least 3 transrectal microUS sweeps of their prostate at large, medium, and small imaging presets prior to the procedure. All microUS scans were performed by the same experienced urologist to ensure consistency across patients. Immediately after a scan was obtained it was reviewed and, if necessary, repeated to confirm sufficient quality and that the entire prostate was captured. After prostatectomy the prostate was placed in a patient-specific 3D printed histology alignment mold and sectioned in accordance with routine clinical practice at UCLA [207]. These molds are specifically designed to section the prostate in the axial orientation to allow for direct comparison between MRI
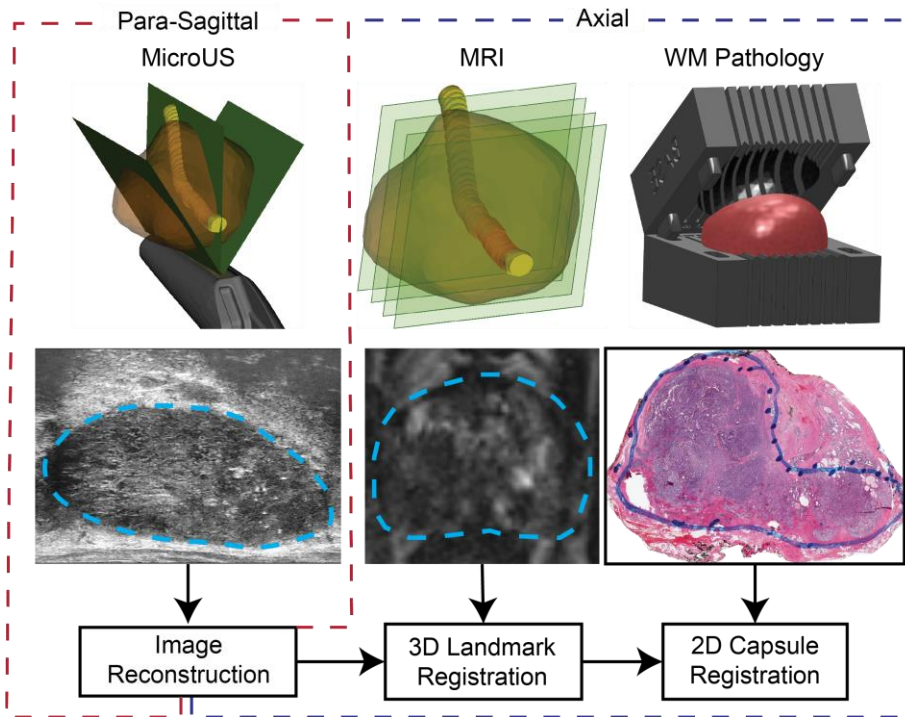


Figure 34: Overview of methodology of alignment of microUS images with MRI and WM pathology. MicroUS images are acquired in a para-sagittal fan sweep, while MRI is acquired in the axial plane. Following prostatectomy, the gland is sectioned in a pathology mold to align the WM sections with the MRI axial orientation. The para-sagittal microUS frames are reconstructed into a "true" sagittal orientation, then registered to MRI via a 3D landmark registration. Finally, MRI and microUS are registered to WM pathology via a 2D capsule registration.

and final pathology. Using a microtome, WM pathology slides were made from each section and subsequently scanned at 20X magnification and saved digitally. Cancerous lesions were delineated on the WM slides by an experienced genitourinary pathologist. In summary, an axial mpMRI scan, a para-sagittal microUS sweep, and axial WM pathology slides were acquired from 15 patients (Figure 34).

## 5.2.2.  MicroUS Image Plane Reconstruction

The first step to register the para-sagittal microUS sweep with the axial MRI and WM datasets required the microUS image data to be reconstructed into parallel frames. Since every frame in an imaging sweep was acquired at a different angle in a para-sagittal sweep, the data cannot be loaded into a standard DICOM viewer without losing geometric positioning of the frames (Figure 35A). A custom MATLAB (MathWorks) script was used to reconstruct the microUS images in the sagittal orientation through positioning in 3D space, resampling, and interpolation. The microUS images were obtained as a 300-frame sweep from the right to left side of the prostate with each image at a varying angle of transducer rotation. The native image frames can be thought of as 2D planes in a cylindrical coordinate system rotated around the rectum (x-axis) at a specific angle and offset from the center of rotation. To mimic this, a rotation around the x-axis was applied to the pixel coordinates of each frame using the stored angle metadata. Due to the microUS probe design, there is a 2mm offset from the edge of the US transducer and the first row of pixels acquired in the image. To account for this, each image was shifted above the x-axis by 2mm prior to rotation. After rotating all the images, empty parallel planes with 1mm spacing along the y-axis (sagittal orientation) were populated with any image data that lies within y±0.5mm of the plane. A gridded interpolation step was
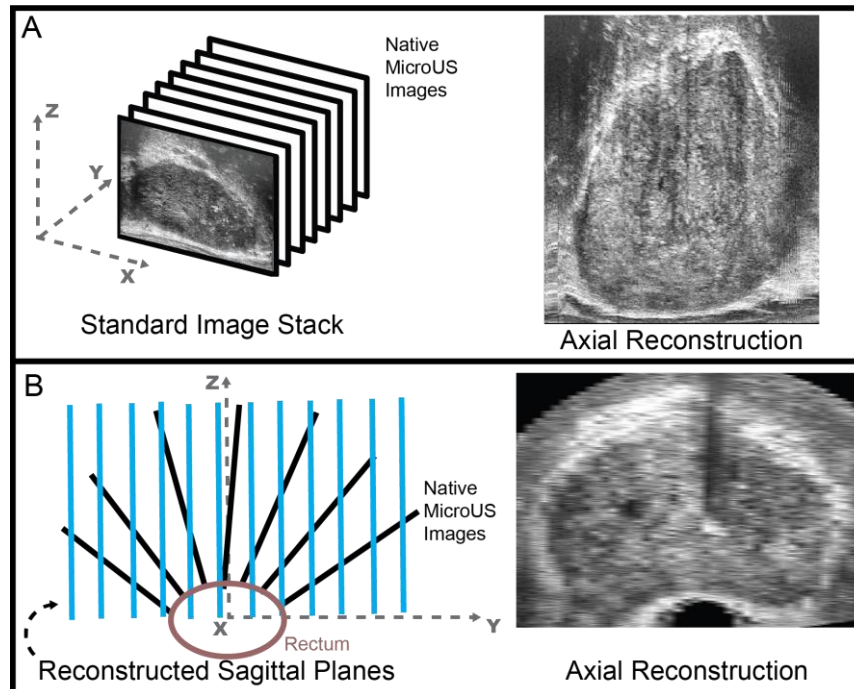
Figure 35: A) Loading consecutive images into a DICOM viewer without accounting for scan angle results in a warped prostate shape that is not representative of the actual gland. B) By accounting for the scan angle and imaging offset it is possible to reconstruct sagittal images (blue), that when loaded into a DICOM viewer preserve the geometric shape of the gland to facilitate axial viewing and 3D reconstruction.

performed to fill any holes in the new image planes. These new images were then saved as a DICOM series with embedded pixel spacing and coordinate locations. When loaded in a standard DICOM viewer, these reconstructed scans allow for correct visualization of the prostate in traditional orientations (axial, sagittal, and coronal) (Figure 35B).

### 5.2.3.  MicroUS to MRI Registration

**Registration Process**

Following microUS image plane reconstruction, the next step is to account for tissue deformation and any misalignment from the reconstruction step. During microUS interrogation of the prostate, the probe deforms the prostate causing compression of

the gland in-plane. Furthermore, only the roll (rotation around the x-axis) of the US probe was tracked and recorded during scanning, so any pitch, yaw, or translation has not been accounted for at this stage. The reconstruction process described in the prior step is designed to approximate the shape of the prostate, but a corrective registration step is required. To account for this, a landmark registration with MRI was performed. MRI was chosen over direct registration to WM pathology due to the superior sampling density of MRI and to avoid potentially confounding factors for tissue deformation from WM slide shrinkage during fixation and processing. Additionally, WM pathology is generally limited to at most 6-7 slides per case with approximately 4.5mm spacing between slides, whereas the majority of MRI data was acquired with 1.5mm spacing providing 3 times the sampling density registration.

For microUS to MRI registration, the reconstructed microUS and MR images were loaded into 3D Slicer[208] and the microUS sweep was visually aligned to the MRI via translation and rotation. Following visual alignment, a landmark based thin-plate spline (TPS) registration was performed. Landmark selection was limited to selection of 3 common points along the urethra at the apex, mid-gland, and base, as well as points along the capsule in 3 axial frames in the apex, mid-gland, and base (Figure 36A).

**Registration Accuracy Assessment**

The accuracy of the registration was then assessed by selecting 5 internal landmarks (excluding the urethra) and measuring the distance between them to get the target registration error (TRE). Landmarks were generally distributed throughout the prostate rather than clustered in the same location (e.g., apex, base, mid-gland, left, right).
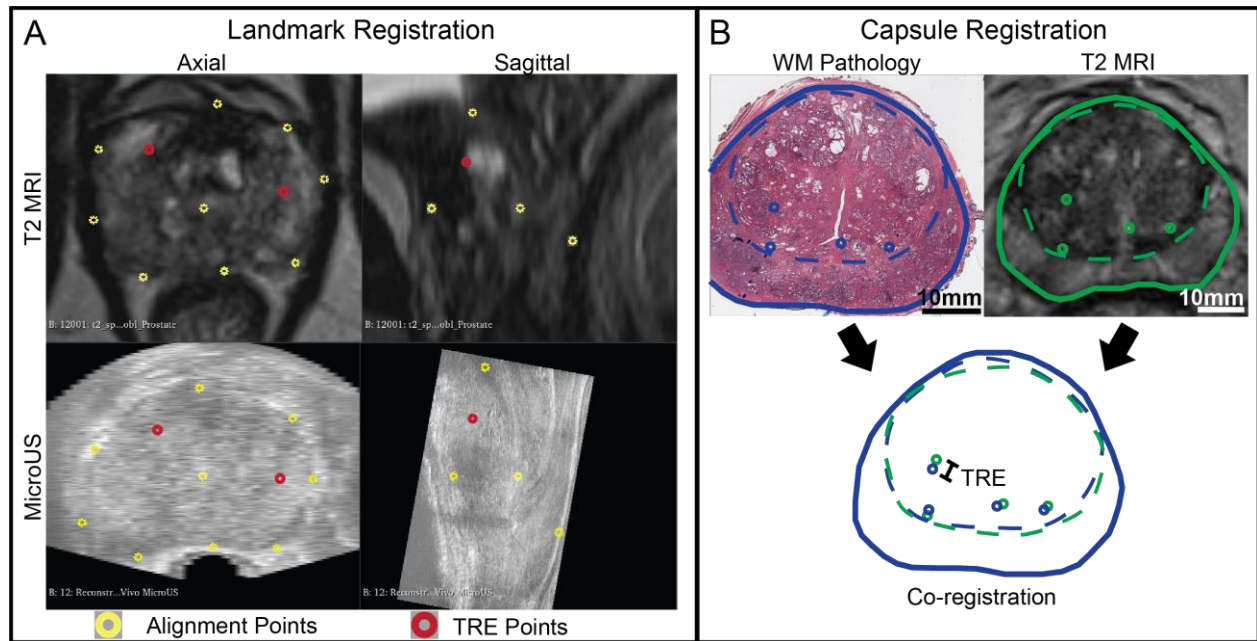
Figure 36: A) Landmark registration between microUS and MRI. Points along the urethra and the capsule were selected to inform the TPS transformation (yellow). The distance between five internal anatomical fiducials (red) were used to evaluate TRE. B) Capsule registration between MRI and WM pathology. The distance between fiducials (circles) were used to evaluate TRE and the overlap of the transition zone (dashed lines) was used to determine DICE scores.

## 5.2.4. MRI to WM Pathology Registration

**Registration Process**

After microUS registration with MRI, a second registration step is performed to register both microUS and MRI with WM pathology. Before this can be done, the WM slides must be matched to a specific MRI image. The general location and the order of each histology slide is known. However, since the thickness of each histology section is 4.5mm and the MRI thickness in our dataset is usually 1.5mm, there are at least 3 possible MRI images that could correspond to a given WM slide. To address this, an experienced reviewer matched each WM slide to a corresponding MRI slide using the knowledge of the slide location, size and shape of the capsule on MRI and WM, anatomical features (e.g., urethra, seminal vesicles, BPH nodules, and calcifications), and the relative spacing from

117

other MRI matched WM slides. After matching all slides, it was confirmed that exactly one WM slide was contained within every 3 MRI frames (when MRI slice thickness was 1.5mm). Following this process, WM slides are able to be matched with an error of approximately 1.5mm (or one MRI frame) as reported by an earlier study from our group [175].

Using MRI and WM capsule segmentations, the MRI was co-registered to the WM slides through a custom MATLAB script. The registration involved a rigid transformation to account for scaling and in-plane rotational misalignment, followed by a non-rigid TPS transformation using control points on the prostate capsule to account for specimen deformation as performed in earlier studies [155], [209].

**Registration Accuracy Assessment**

In order to assess registration accuracy, matched image data were assessed by four reviewers (3 urologists, 1 imaging scientist). Using a DICOM viewer (Weasis Medical Viewer), each of the 15 subjects was reviewed by at least two reviewers and each reviewer was randomly assigned a subset of eight cases to review. For each case, reviewers were asked to mark up to four common anatomical landmarks as well as outline the transition zone and capsule on both MRI and WM for each matched image pair (Figure 36B). The MRI was then co-registered to WM for each case and each reviewer. The distance in millimeters between identified landmarks was recorded as the TRE, and the dice similarity coefficient was calculated from the overlap of the identified transition zones.

## 5.2.5. Pilot Clinical Study

**Whole Mount Pathology**

As described earlier, in accordance with normal clinical practice, an experienced genitourinary pathologist reviewed the WM pathology slides and identified the prostate cancer boundaries as well as the International Society of Urological Pathology (ISUP) grade group of the tumors.

**MRI**

For MRI, any delineated PI-RADSv2.1 target from the original mpMRI prior to surgery was imported. In addition, using the mpMRI scan, an experienced radiologist with over ten years of experience with PI-RADS was asked to delineate any suspicious regions on T2 MR images with the knowledge the patient received a prostatectomy for prostate cancer (retrospective MRI). This was necessary because the PI-RADSv2.1 protocol prioritizes specificity and identification of the center of the index lesion as a triage for biopsy. Consequently, this methodology underestimates tumor extent in an effort to minimize unnecessary biopsy procedures [13]. However, we are interested in analyzing the ability of microUS and MRI to identify the margins of prostate cancer, which is critical for treatment decisions and focal therapy planning. Furthermore, given microUS is a relatively new imaging modality, there are limited numbers of readers familiar with the technology and even fewer expert readers. As a result, the experienced readers available for this study had knowledge of the study design and were aware it is validated via WM pathology, which is only available for prostatectomy patients. For the above reasons, both the MRI and microUS reviewers were aware the patient had a prostatectomy prior to review, but were blinded to the WM pathology and other clinical history of the patient (e.g., cancer grade, PSA, biopsy results).

**MicroUS**

For microUS, the same four reviewers in the landmark study reviewed the native microUS sweeps and were asked to identify up to two suspicious lesions per subject to ensure only the most suspicious regions were annotated. Reviewers had varying levels of microUS experience but they all completed an online training module from the microUS manufacturer to ensure baseline competency. However, the clinical experience of each reviewer varied (number of microUS biopsies performed or observed). Reviewers 1-3 had performed or observed over 50 microUS targeted biopsies at the time of review and frequently interpret microUS scans, while reviewer 4 had knowledge of the system but had not attempted interpretation outside of the training module. Similar to the landmark study, each subject was reviewed by at least two reviewers with each reviewer assigned to the opposite subset of cases they reviewed in the prior landmark study to avoid any previous knowledge bias. As an added precaution, there was also a minimum 90-day washout period between reviews to avoid bias. Reviewers were blinded to all clinical parameters but were aware the patient had a prostatectomy for prostate cancer, similar to prior final histology based studies for microUS [191].

**Registration**

Following the described registration process, the microUS annotations (Figure 37A) were first transformed into a point-cloud and oriented in 3D space in the same manner as the reconstruction of the microUS images (Figure 37B). The annotations were then registered to MRI via the stored transformation matrices created from the 3D landmark registration between microUS and MRI (Figure 37C). After registering microUS annotations with MRI, both the MRI and microUS annotations were registered to the WM pathology following the rigid and non-rigid MRI to WM capsule registration (Figure 37D). Using density-based spatial clustering of applications with noise (DBSCAN), the microUS point-clouds were
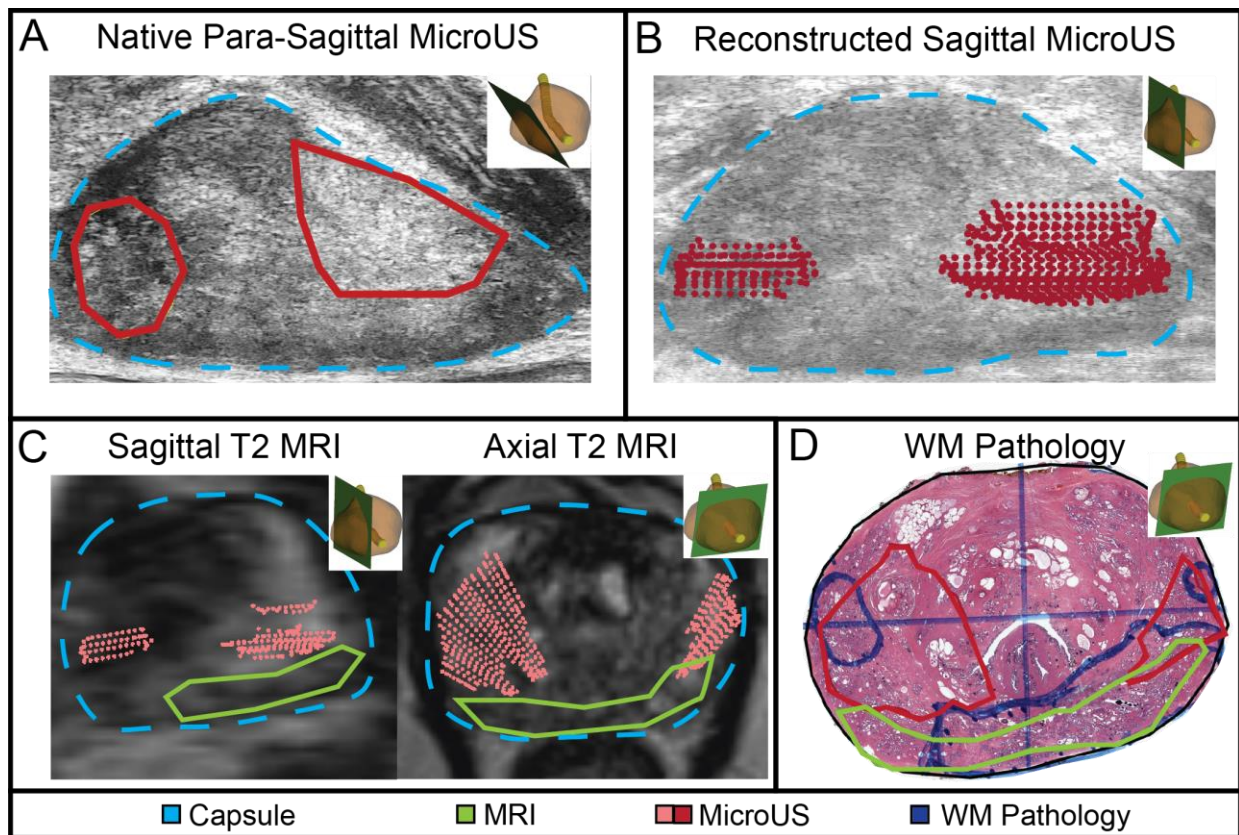
Figure 37: Complete registration process from microUS to WM pathology. A) MicroUS reviewer delineations (red) are made on native microUS imaging. B) These markups are turned into a point-cloud and positioned in 3D space similar to the image reconstruction. C) The transformation matrices for microUS to MRI registration are applied to the markup. D) Lastly, the microUS and MRI (green) delineations are co-registered with WM pathology.

transformed into outlined regions-of-interest (ROIs). For each review of the 15 cases, pixel level detection metrics were quantified, and figures were created for each of the WM slides overlaid with the annotations from prospective PI-RADSv2.1 MRI (pMRI), retrospective MRI (rMRI), and microUS reads.

## 5.3.   Results

### 5.3.1.  Data Collection

Each surgery was performed by one of two high volume surgeons using a robotic assisted laparoscopic prostatectomy approach.  The median patient age was 63 [58-68]

(median [IQR]). Median preoperative PSA was 8.8 [6.3-11.3]. The final pathology revealed 5 patients with organ confined disease (T2), 7 with extra prostatic extension (T3a), and 3 with seminal vesicle invasion (T3b). There was a total of 19 significant lesions in the patient population. The median ISUP Grade Group was assigned to each lesion. The average number of lesions per prostate was 1.6±0.6 (mean ± standard deviation). For the index lesion (highest grade and/or largest tumor) the average largest axial dimension was 29.4±11.9mm with the following ISUP grade group (GG) distribution: GG2: 8, GG3: 5, GG4: 1, and GG5: 1. There were 4 significant secondary lesions, all of which were GG2 and 5 clinically insignificant GG1 secondary lesions.

## 5.3.2. MicroUS Image Plane Reconstruction

A total of 26 different microUS scans were reconstructed (1-2 per patient) with a slice thickness of 1mm. On average the reconstructed scan contained 68±8 (mean ± standard deviation) images created from the original 300 frame scan. The reconstructed images are saved as DICOM files and contain embedded pixel spacing, slice position, and slice thickness information allowing for clear visualization of the prostate in the axial sagittal
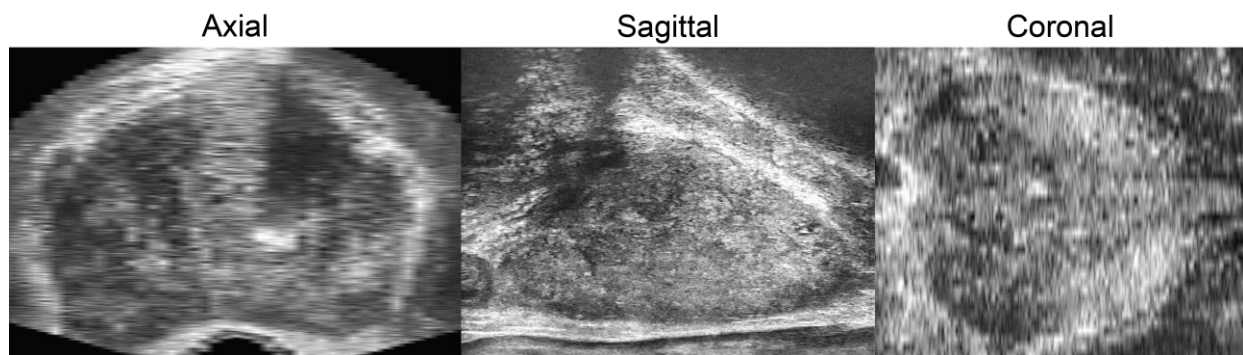


Figure 38: Example of a set of microUS images reconstructed in axial, sagittal, and coronal orientations. Sagittal images were reconstructed via a MATLAB script allowing for visualization of the axial and coronal planes in a DICOM viewer. Sagittal images are of sufficient quality to enable accurate visualization of anatomical features.

and coronal planes in any standard DICOM viewer (Figure 38). Images correctly represented the shape of the prostate while preserving image quality for visualization of anatomical features.

### 5.3.3.  MicroUS to MRI Registration

Following the registration of the reconstructed microUS scan with T2 MRI, five common anatomical landmarks (excluding the urethra) were selected from varying regions of the prostate for each of the 26 scans. A total of 130 landmarks were identified and the distance between them measured. The resulting average TRE was 2.23±0.08mm (mean ± standard error) (Figure 39A).

### 5.3.4.  MRI to WM Pathology Registration

MRI was co-registered to WM pathology on a subset of 8 patients for each of the 4 reviewers, for a total of 32 registrations comprised of 159 slides and 472 TRE landmarks. Landmarks that were greater than 2 standard deviations outside the mean TRE were
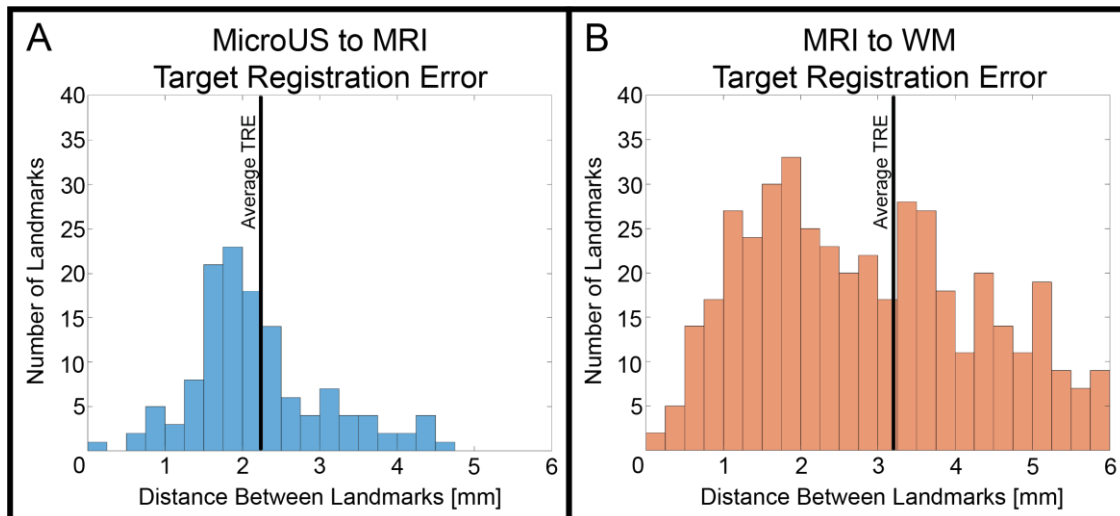


Figure 39: Distance between anatomical landmarks following microUS to MRI (left), and MRI to WM pathology registration (right).

removed as they were determined to be a result of incorrect landmark selection. The average TRE between landmarks on MRI and WM was 3.20±0.08mm (mean ± standard error) (Figure 39B). In total only 26 landmarks from an original 498 (5%) were excluded. The average dice similarity coefficient of the transition zone was 0.84±0.01 (mean ± standard error). When comparing reviewers, landmarks were not excluded in order to preserve reviewer variability. There was no significant difference between reviewers for TRE (p= 0.15) and transition zone dice scores (p = 0.62, Kruskal-Wallis). The combined TRE (summation in quadrature) of microUS to WM pathology from the 3 registration steps was 3.90±0.11mm.

## 5.3.5. Pilot Clinical Study

A total of 32 microUS reviews (8 per reviewer) were registered to the associated MRI and subsequently registered to WM pathology. Two examples of registration of microUS and
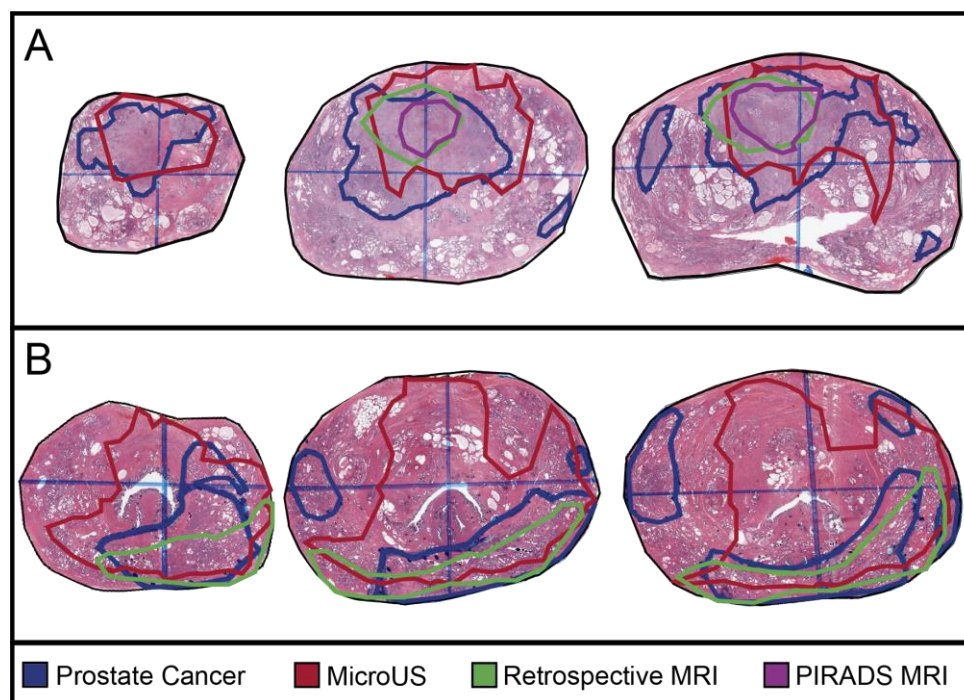


Figure 40: Example co-registration outputs of microUS and MRI with WM pathology for two cases (A and B).

124

MRI annotations to WM pathology are shown below (Figure 40). Only clinically significant, grade group 2 or greater, lesions were considered for this analysis. Results are reported as mean ± standard error where appropriate.

Due to the rigorous nature of the registration process, analysis can be performed at increasing levels of granularity starting at the patient level, then the cross-section level and finally the pixel level. At the patient level, a total of 11/15 (73.3%) index lesions were identified by pMRI and 12/15 (80%) were identified by rMRI. For each individual microUS reviewer: 7/8 (88%), 8/8 (100%), 7/8 (88%), and 4/8 (50%) index lesions were identified respectively (Figure 41A). In a combined microUS analysis, where reviewer 4 was removed, 91.7±4.2% of index lesions were identified (Figure 41B). MRI did not detect any index lesions that were not detected by microUS. If any part of an ROI overlapped the index lesion it was considered a correct identification. Reviewer 4 was removed from combined microUS analysis due to reduced expertise and poor index lesion identification. Reviewers 1-3 (2 attending urologists and 1 expert imaging scientist) had performed and observed over 50 microUS biopsies at the time of review, whereas reviewer 4 (urologic fellow) had not performed any microUS biopsies at the time of review. Furthermore, there was a significant decrease between reviewer 4 and the other 3 reviewers ($p < 0.0001$, Kruskal-Wallis) when analyzing ROI-tumor overlap.

At the cross-section level, a total of 90 clinically significant lesions cross-sections (2D slice of a lesion on WM pathology) were present across the 82 WM slides from all 15 patients. The average ± standard error and median [IQR] area of these lesion cross-sections was 89.0±11.7mm$^2$ and 37mm$^2$ [1-130] respectively. Of these 90 lesion cross-sections, pMRI identified 17/90 (18.9%) whereas rMRI exhibited improved performance with 33/90 (36.7%). MicroUS had even better performance identifying 85/162 (52.5±0.1%)

lesion cross-sections between the three reviewers (Figure 41B). Similar to the index lesion analysis, if any ROI overlapped a lesion cross-section it was considered a correct identification. Identification of lesion cross-sections provides a metric to assess the ability of the imaging modalities to identify tumor extent.

The following quantitative metrics were analyzed at the pixel level for increased granularity and improved margin evaluation which is crucial for focal therapy planning. For example, the overlap between an ROI and a cancerous lesion was decided for every pixel contained within a lesion or ROI.

The lesion overlap was defined as:

$$Overlap\ (\%) = \left( \frac{Lesion\ Area\ \cap\ ROI\ Area}{Lesion\ Area} \right) 100$$

The overlap percentage accounts for the amount of a lesion cross-section encapsulated by a modality. The average overlap for pMRI, rMRI, and microUS were 7.4±1.9%, 17.1±2.8% and 22.8±2.4% respectively (Figure 41B). Interestingly, there was a significant difference between both microUS and pMRI, and rMRI and pMRI (p < 0.001, Wilcoxon rank sum) whereas microUS and rMRI exhibited no statistical difference. This suggest that both rMRI and microUS identified significantly more tumor extent than pMRI. The reduction in encapsulation between pMRI and rMRI can likely be attributed to the primary focus of PI-RADSv2.1 on identifying the core of the index lesion while preserving specificity as part of the decision to refer a patient to biopsy. The overlap percentage of each modality was also stratified by cross-sectional area to show how tumor size can affect diagnostic ability (Figure 41C). As expected, overlap generally improved as lesion size increased with pMRI, rMRI, and microUS increasing from 0±0%, 8.9±3.2%, and 15.2±3.1% to 31.5±7.8%, 50.1±8.8%, and 35.5±5.9% respectively. There was also a

126

significant difference between microUS and pMRI for all but the largest groups of lesions (p < 0.05, Wilcoxon rank sum) indicating that microUS may outperform pMRI for tumor extent when detecting all but the largest lesions. However, given that there was no statistical difference between microUS and rMRI, it is likely that the deficiencies of pMRI are not due to limitations of the MRI modality but rather the PI-RADSv2.1 protocol which is not designed for optimizing overlap.

The false positive metric was defined as:

$$False\ Positive\ (\%) = \left(1 - \frac{Lesion\ Area\ \cap\ ROI\ Area}{ROI\ Area}\right) * 100$$

This metric accounts for the amount of an ROI that does not overlap a lesion and is analogous to the specificity of a modality. False positive percentages were reported instead of specificity for the following reasons: 1) To avoid potentially misleading values which arise as the pixel level granularity of the analysis results in inflated specificity values (approximately 95% for all modalities) when compared to conventional zonal based studies [16], [206] 2) To allow for more opportunities for one imaging modality to distinguish itself from the other since the area of the ROI is always substantially smaller than the benign (negative) area. At 36.2±8.6%, the pMRI false positive ratio was significantly better than rMRI (65.4±5.5%) and microUS (73.7±3.0%) (p < 0.01, Wilcoxon rank sum). This reflects the prioritization of specificity in the PI-RADSv2.1 protocol. Furthermore, only 5/22 (22.7%) of pMRI annotations completely missed a significant lesion cross-section, whereas rMRI and microUS had 20/49 (40.8%) and 41/ 107(38.3%) respectively. Collectively, this indicates pMRI has superior specificity to both rMRI and microUS. Interestingly, microUS and rMRI had similar rates of both false positives and completely false ROIs.

Dice scores evaluate the similarity between two ROIs, defined as:

$$DICE = 2\left(\frac{Lesion\ Area\ \cap\ ROI\ Area}{ROI\ Area + Lesion\ Area}\right) * 100$$

Since this metric accounts for both false positive regions and false negative regions, it is a single metric that represents both sensitivity and specificity. For pMRI, rMRI, and microUS the dice scores were 13.9±3.1%, 18.7±3.1%, and 16.6±1.9% respectively (Figure 41B). There was a significant difference between microUS and pMRI dice scores (p < 0.05, Wilcoxon rank sum) indicating microUS allows for better similarity between reviewer identified ROIs and true tumor boundaries than pMRI. Similar to overlap percentage, dice scores were stratified based on cross-sectional area to analyze how tumor size affects diagnostic accuracy. As expected, dice scores generally improved as lesion size increased with pMRI, rMRI, and microUS increasing from 0±0%, 1.9±1.3%, and 4.6±1.5% to 31.7±7.0%, 44.0±6.1%, and 36.3±3.6% respectively (Figure 41D).

The sensitivity of each modality was defined as:

$$Sensitivity = \left(\frac{Lesion\ Area_{Slide}\ \cap\ ROI\ Area_{Slide}}{Lesion\ Area_{Slide}}\right) * 100$$

Although similar to the lesion overlap percentage, sensitivity was calculated per slide rather than per lesion. As such, slides with multiple lesion foci were grouped together. For pMRI, rMRI, and microUS the sensitivity was 11.1±2.7, 22.1±3.6%, and 26.7±3.0% respectively (Figure 41B). A significant difference was found between pMRI and rMRI, and pMRI and microUS (p < 0.01, Wilcoxon rank sum) further indicating the improved ability of microUS and rMRI over pMRI for assessing tumor extent.
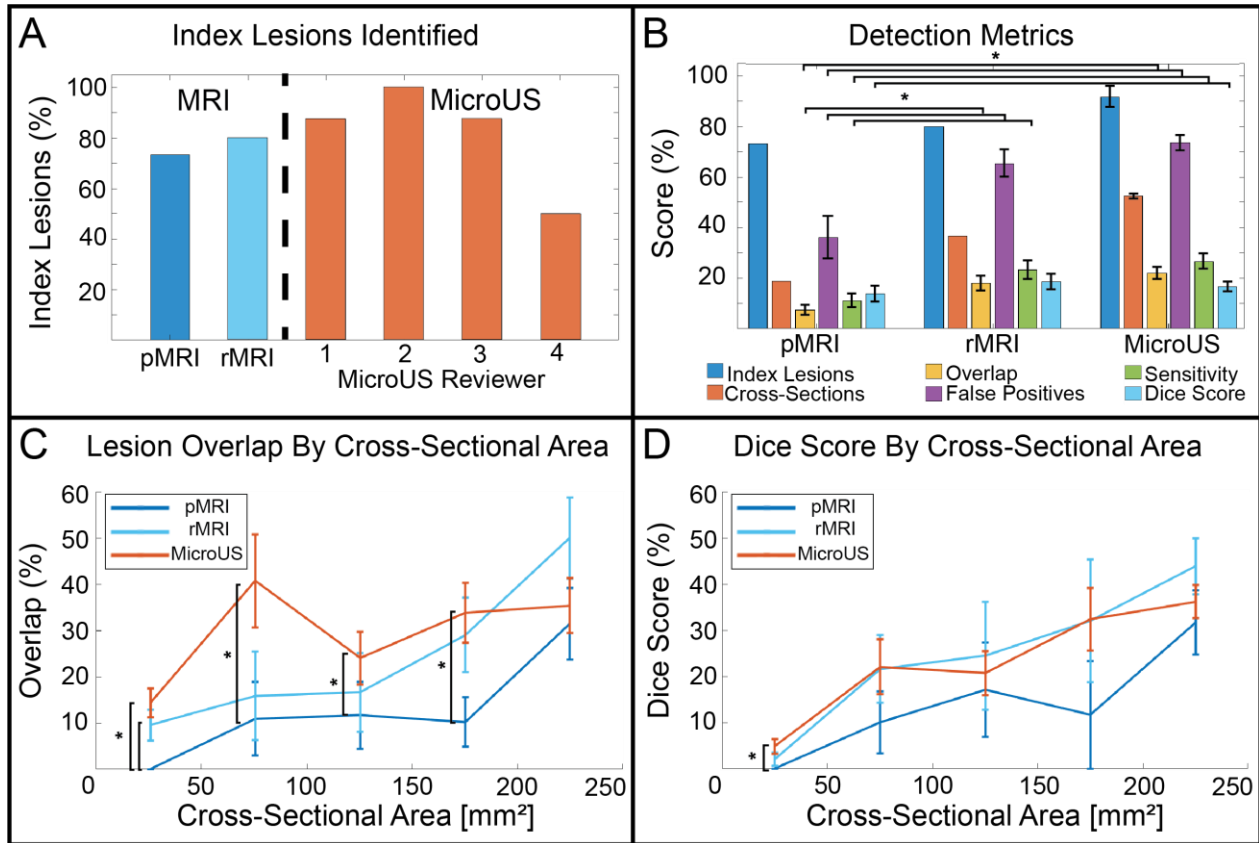
Figure 41: Assessment of microUS and MRI against WM pathology in a 15-patient clinical study. A) Index lesions identified by modality and reviewer. B) Combined reviewer detection metrics for each modality at the patient level (index lesions), cross-section level (cross-sections), and pixel level (overlap, false positives, sensitivity, dice score). Lesion overlap (C) and dice score (D) for each modality stratified into groups by lesion cross-sectional area. (* indicates p<0.05, Wilcoxon rank sum. Error bars are ± one standard error)

## 5.4.    Discussion

Conventional US is non-diagnostic for prostate cancer, as such mpMRI is the recommended imaging modality for prostate cancer detection [5], [10], [11]. However, mpMRI is resource intensive, limiting its use to large centers. High-resolution microUS has been introduced as a potential low-cost alternative with promising preliminary results. Prior studies have relied on biopsies or cognitive zonal registration with histology to compare the diagnostic accuracy of microUS and MRI [191], [206].

129

Computational registration of microUS and MRI with ground truth WM pathology facilitates more rigorous comparison of the two imaging modalities, however, to date that has not been demonstrated as the native orientation of microUS imaging makes direct registration with WM pathology difficult.

In this work we developed and validated the first multi-step process for co-registration of microUS, MRI, and WM pathology, and subsequently assessed the ability of expert microUS and MRI reviewers to identify prostate cancer against ground truth WM pathology. In doing so, we have created a validated methodology that allows for rigorous comparison of microUS and MRI, as well as an approach for labeling microUS images with ground truth histology for predictive model training. The three-step process consists of 1) microUS para-sagittal image reconstruction into the true sagittal orientation, 2) 3D alignment and TPS registration of the reconstructed microUS images with MRI, and 3) 2D alignment and TPS registration of MRI with WM pathology slides.

The first step, microUS image reconstruction, approximates the shape of the prostate capsule at the time of the US scan. This reduces the magnitude of subsequent registration steps, reducing opportunities for misregistration. We have demonstrated it is feasible to reconstruct sagittal microUS images from native images acquired from a handheld transrectal, para-sagittal sweep of the prostate. The image quality from these reconstructions is adequate to identify intraprostatic fiducials such as the urethra, and preserves the anatomical prostate shape. Furthermore, since these images are used as an intermediary in the registration process, reviewers are able to delineate the native microUS images without any loss of image quality. This is an improvement on our prior work which required reviewers to delineate ex vivo imaging in order to align with MRI

130

and WM pathology, which is not necessarily representative of the in vivo imaging environment [166].

In the second step, reconstructed microUS images are co-registered to MRI using a landmark TPS registration. This accounts for tissue deformation during microUS scanning and possible reconstruction error from limited probe position tracking. Intermediate registration to MRI was chosen over direct registration to WM pathology for the following reasons: 1) MRI has superior through plane resolution compared to WM (1.5mm vs 4.5mm) providing more information for registration 2) MRI is in vivo and more representative of the microUS imaging environment than WM slides, and 3) WM pathology processing causes the tissue to shrink introducing additional sources of error [203]. The measured TRE for microUS to MRI co-registration was 2.23±0.08mm which is within the 2-3mm TRE for commercial MRI-US fusion systems [21], [210].

The third and final step, 2D MRI to WM capsule registration, allows for direct comparison with ground truth histology. Since microUS is fused to MRI in step two, the MRI to WM transformation matrix can be applied to both MRI and microUS annotations. The average TRE for MRI to WM was 3.20±0.08 within the range reported in literature of 1-5mm[22] resulting in a total registration error of 3.90±0.11mm for microUS to WM pathology. The registration accuracy is further validated by the pilot clinical study, where MRI and microUS identified 73.3-91.7% of the index lesions, consistent with earlier studies [6], [7], [8], [191], [206], [211].

After validating each step of the registration process, it was applied to 15 patients. Each patient had a pre-operative MRI read, a retrospective MRI read, and at least two independent microUS reads registered to final pathology.

The microUS results emphasize the crucial role of reviewer training and experience beyond baseline competency. Notably, reviewer 4, being the least experienced, achieved significantly inferior results compared to the other reviewers ($p < 0.0001$). This also highlights a potential value of future deep learning models to alleviate both inter-reviewer variability and the need for extensive training.

In the pilot clinical study, microUS demonstrated non-inferiority to both pMRI and rMRI for identifying prostate cancer index lesions and tumor extent. Specifically, microUS identified a higher percentage of index lesions (91.7±4.2%) compared to pMRI (73.3%) and rMRI (80%). It is vital for an imaging modality to identify the index lesion as it often drives treatment decisions. MicroUS also identified a higher percentage of lesion cross-sections with better overlap (52.5±0.1% and 22.8±2.4%) than both pMRI (18.9% and 7.4±1.9%) and rMRI (36.7% and 17.1±2.8%). Lesion cross-sections and overlap percentages are representative of tumor extent throughout the prostate and are related to a modality's ability to identify tumor margins. Patient treatment decisions can be affected by cancer volume and location in the prostate. As such, margin identification is important for cancer staging and deciding treatment options, especially for focal therapy where only part of the gland is treated.

MicroUS also performed comparably to rMRI for both sensitivity (26.7±3.0% vs 22.1±3.6%) and dice scores (16.6±1.9% vs 18.7±3.1%), with both outperforming pMRI (11.1±2.7% and 13.9±3.1%). Conversely, MicroUS and rMRI had similarly high false positive percentages (73.7±3.0% and 65.4±5.5% respectively) and completely false ROIs (38.3% and 40.8% respectively) compared to pMRI (36.2±8.6% and 22.7%). This emphasizes that PI-RADSv2.1 is designed to prioritize specificity of MRI at the cost of sensitivity to avoid unnecessary biopsies. Perhaps the specificity of microUS, often noted

to be worse than pMRI [17], [212], [213] can be increased through the refinement of the microUS diagnosis protocol (PRI-MUS) similar to improved iterations of PI-RADS raising the specificity of MRI [6].

It is worth noting that the recorded results for sensitivity, dice scores, and overlap percentage are lower and the false positive percentages are higher than reported in other related studies [16], [17], [191], [206]. This can be attributed to the granularity of the analysis performed. For these metrics each pixel was treated as a possible true positive, false positive, true negative, or false negative event. Any misalignment between regions reduces the modalities accuracy allowing for more opportunities for one modality to outperform the other, which is important for this study given the limited sample size (N=15). Other studies in the space often rely on zonal based analysis, analogous to the study by Turkbey et al. [214], where the quantity of possible true positive, false positive, true negative, and false negative events are limited to twelve zones. In general, with increased granularity, the overlap, sensitivity, and dice scores decrease and false positive ratios increase. An example of this can be seen in similar work co-registering WM to MRI with a pixel spacing of ~0.6mm, which reported MRI sensitivity and specificity values of 37% and 97.9% respectively, Whereas traditional zonal MRI studies report sensitivities and specificities of 86-92% and 60-83% respectively [6], [7], [8], [192]. To avoid reporting potentially misleading values, specificity values of each modality were intentionally not reported, and instead the false positive ratio was used as a surrogate for specificity. With increased granularity, the specificity for all modalities and all reviewers was approximately 95% since it is related to the number of true negative events, and the majority of WM slides were comprised of mainly benign tissue. Although the specific

results from this study are not directly comparable to other zonal based studies, the general trends can be compared.

The results presented in this study provide good agreement with the larger body of literature in the field. Specifically, microUS demonstrated non-inferiority compared with MRI for index lesion identification and detection of tumor extent [16], [17], [191], [206]. This is particularly important given the reduced cost of microUS compared to mpMRI. However, microUS had higher false positive ratios than MRI, similar to other studies [17], [212], [213]. Additionally, this study demonstrated reviewer expertise is vital for identification of prostate cancer on microUS; a possible reason for earlier studies having lower detection rates compared to later studies [16], [17], [188].

This study also highlights a potentially synergistic relationship with mpMRI and microUS, wherein, MRI is used to triage patients for biopsy and identify the index lesion and microUS is used during biopsy to identify tumor extent throughout the prostate and better inform tumor margins for treatment decisions. MRI is sensitive for index lesion cancer and maintains high specificity, making it an excellent system for triaging patients for biopsy. However, PI-RADSv2.1 severely underestimates the extent of the lesion in patients as reported in this work and in earlier studies[13]. While microUS has reduced specificity, this generally would only result in extra biopsy cores rather than performing a biopsy on a patient who does not need it. Whereas the improved detection of tumor extent under microUS could result in better tumor margin identification. This is particularly valuable in an era where focal therapy modalities have been developed with the ability to create precise ablation zones and are thus capable of achieving oncologic control if tumor margins are evident e.g. focal HIFU [215] or laser ablation with real-time monitoring [169], [216].

This study also demonstrated retrospective MRI reads with knowledge of existing cancer, greatly improved identification of tumor extent over a standard PI-RADSv2.1 read, while conversely lowering the overall specificity for prostate cancer. This suggests tumor margin identification could be improved by re-reading MRI scans after cancer is confirmed via biopsy.

There are a few limitations of this study. The registration process presented here is accurate, however, it is not fully automated, requiring WM to MRI matching and landmark selection for microUS to MRI registration. This can potentially be improved upon with the implementation of algorithms for MRI to WM matching as well as capsule and urethra identification. Another potential limitation of the registration process, inherent to side-fire transrectal US, is larger prostates will have more signal attenuation and larger distances between consecutive images in the anterior regions. This could diminish reconstruction quality and registration. However, this may be a minimal factor as the largest prostate in this cohort was 108g (20g is average [217]) and the average microUS to MRI registration error for this patient was only 0.16mm higher than the total average. Moreover, both microUS reviewers correctly identified all significant lesions for this patient. Additionally, the diagnostic metrics derived here should be interpreted with some caveats. Firstly, the limited sample size (N=15 patients) reduces statistical power and also indicates that a broad of spectrum of tumor presentations (size, location, ISUP grade group) was not examined. Secondly, the patient cohort consists of only prostatectomy patients, an inherent limitation of WM pathology-based studies. As such, the patient population likely has more aggressive disease than an equivalent biopsy or focal therapy cohort. Additionally, due to the relatively limited number of WM pathology slides per prostate and the differences between identifying an ROI and taking an actual

biopsy, a correct detection of cancer on microUS or MRI could still lead to a negative targeted biopsy core, particularly for transrectal biopsies. For this reason, we avoided approximating correct biopsy targets since a core could have been taken from a region of an ROI that did not overlap cancer, especially when accounting for tissue deformation and needle deflection during sampling. Moreover, microUS reviewers were aware the patient had a prostatectomy for prostate cancer due to limited availability of expert microUS readers and their knowledge of the study being validated by WM pathology. As mentioned earlier, this limitation is present in earlier final pathology based studies reported in literature [191]. To account for this limitation, the MRI was also reviewed retrospectively with the same knowledge allowing for an equal comparison between microUS and MRI. It is also worth noting that the microUS images were effectively recorded scans, subsequently the microUS reviewers were not able to utilize potential scanning benefits from live US imaging such as variable compression and probe translation to further interrogate suspicious tissue. It also may not be indicative of a live interpretation in regards to time constraints during an actual biopsy. Furthermore, for non-expert reviewers, reading microUS images is challenging, evident from the performance of the excluded microUS reviewer. This emphasizes a need for automated predictive models for cancer detection to reduce the learning curve and cognitive load on physicians when interpreting microUS images during biopsies.

In contrast to MRI [151], microUS-based AI models for automated tumor detection are much earlier in development and typically rely on weakly labeled biopsy data [201], [202], [218]. This is largely due to the absence of accurately labeled datasets which have been challenging to compile because of the complicated registration process. However, the approach that we have presented here greatly alleviates this problem and thus has

the potential to spark the development of the requisite AI models that would greatly ease the adoption of microUS.

The registration process presented here has great utility in further characterizing the diagnostic capability of microUS. Interest in this area is rapidly growing as evidenced by recent publications [16], [17], [191], [206], [212], [213]. In line with these studies, the pilot study that we have presented demonstrates that microUS has similar performance to MRI for identifying prostate cancer. This growing body of evidence suggests that microUS has the potential to be an inexpensive alternative or adjunct to MRI for prostate cancer visualization.

## 5.5.   Conclusion

MicroUS has the potential to greatly reduce the cost of imaging prostate cancer relative to MRI. However, direct registration with ground-truth WM pathology is challenging. To overcome this, we developed and validated the first methodology for accurate co-registration of microUS and MRI with WM pathology. The methodology was then successfully implemented in a multi-reviewer pilot study which demonstrated microUS is non-inferior to MRI for index lesion identification and potentially more sensitive for tumor extent. At smaller centers, microUS could be an inexpensive alternative to mpMRI for prostate cancer diagnosis. At larger centers, microUS could have a synergistic relationship with mpMRI to capitalize on the specificity of MRI for triaging patients for biopsy and the improved tumor margin identification of microUS during biopsy. However, to preserve diagnostic accuracy it is vital that microUS is used by well trained, expert reviewers. Due to the limited sample size of this study, a larger study is required

to confirm these findings, but the conclusions presented here align closely with other

microUS studies.

# CHAPTER 6

# Deep Learning for Prostate Cancer Classification on Micro-Ultrasound

## 6.1. Introduction

MicroUS has recently been introduced as an inexpensive alternative to mpMRI for imaging prostate cancer. The device operates at a 29MHz frequency, while conventional US is typically around 2-15MHz [186], [219]. Previous studies have demonstrated similar diagnostic accuracy between microUS and mpMRI when utilized by an expert reviewer [17], [18], [219], [220]. Centers that are able to implement both, can utilize MRI as a triage to avoid unnecessary biopsies and microUS to allow for direct visualization and targeting of suspicious regions during biopsy potentially avoiding the registration error associated with MRI-US fusion approaches [21]. Furthermore, lesions that are missed on one modality may be identifiable on the other, potentially reducing any missed cancers.

While microUS has demonstrated similar diagnostic performance to mpMRI, urologists without extensive training and hands-on experience exhibit worse performance than expert users [188], [220]. Additionally, new research has shown that even expert reviewers have a high-degree of inter-reviewer variability with only an 30% agreement between expert reviewers [19]. Furthermore, reading and interpreting microUS images during a biopsy in real-time suffers from a substantial increase in

cognitive load for the operating physician and limited time available to allocate for interpretation compared to a pre-biopsy MRI.

As such, microUS would benefit greatly from the implementation of deep learning models that can localize prostate cancer for the physician. Specifically, accurate deep learning models would alleviate the training requirement and inter-reader variability through standardization of microUS reads, as well as reduce the cognitive load on the physicians through real-time interpretation. Current research in the field is working towards this goal, with development of AI classifiers trained on raw RF microUS images with weakly labeled biopsy data as the ground truth [82], [199], [200], [201], [202]. These studies have exhibited promising results with area under the curve results of 0.80-0.93 for classification along the biopsy track in datasets between 160 to 700 patients.

While biopsy data is relatively easy to acquire and allows for fast dataset accumulation, there are inherent drawbacks resulting in a weak source of data labels. Biopsy data only represents tissue information where it was specifically sampled and while it may correlate to the surrounding tissue, it is impossible to objectively identify the actual tumor extent within the image. Moreover, the core orientation is commonly not preserved or recorded, cores are comprised of combination of benign and cancerous tissue, and cores commonly break into multiple pieces during processing. All of this combined with biopsy needle deflection and tissue deformation increase the challenges associated with cancer localization from biopsy. Due to this, whole-mount (WM) pathology from surgically removed prostates provides is generally considered the gold-standard for cancer localization as it provides much more comprehensive information about tumor location and extent in the prostate, allowing for higher quality ground

truth. However, WM pathology data is challenging to acquire limiting the availability of WM based studies and datasets.

Additionally, while raw RF data provides more information during an US scan, it also increases the file size and complexity of the microUS data, which in turn increases processing and model complexity to extract useful RF information. This added complexity and processing time creates potential delays for utilization of models during a biopsy and creates an additional barrier to eventual adoption by physicians.

In this chapter we develop the first deep learning microUS classifiers validated on ground-truth, mold-sliced, WM pathology and evaluate them against expert readers to allow for added clinical context. Building off our prior work, we employ our recently developed microUS-WM co-registration methodology to confidently label B-mode microUS data for model training and evaluation [220]. The goal of this chapter is to demonstrate the potential of microUS classifiers for identification of prostate cancer with direct comparison against expert readers, as well as demonstrate the utility of confidently labeled data, especially for training models in a limited dataset. To accomplish this, we conduct two studies: 1) a preliminary investigation of deep learning classifiers with 15 patients 2) a more comprehensive study to confirm the results of the preliminary study using 100 patients.

## 6.2.    Data Labeling

As described in chapter 2, eighty patients set to undergo radical prostatectomy provided informed consent and were enrolled into a registered clinical trial (NCT04299620) and UCLA approved study (IRB#19-00113). To be eligible for enrollment patients needed to have a mpMRI with an associate UCLA radiologist report within one year prior to surgery,

and biopsy confirmed prostate cancer. On the day of surgery, a microUS scan of the prostate was collected. Following gland resection, the prostate gland was placed in a patient specific 3D printed alignment mold and sectioned in accordance with routine clinical practice at UCLA. This process is designed to orient prostate sections of 4.5mm thickness in the axial orientation for comparison with axial MRI. Using a microtome, WM slides were created from each section and scanned digitally. These WM sections were subsequently reviewed by an experienced genitourinary pathologist, and the areas of cancer were delineated and assigned a grade group (GG).

Following a similar co-registration process detailed in chapter 4, the native microUS data was reconstructed in the true sagittal orientation to allow for registration with WM pathology (Figure 42). Briefly, the reconstructed microUS images were created
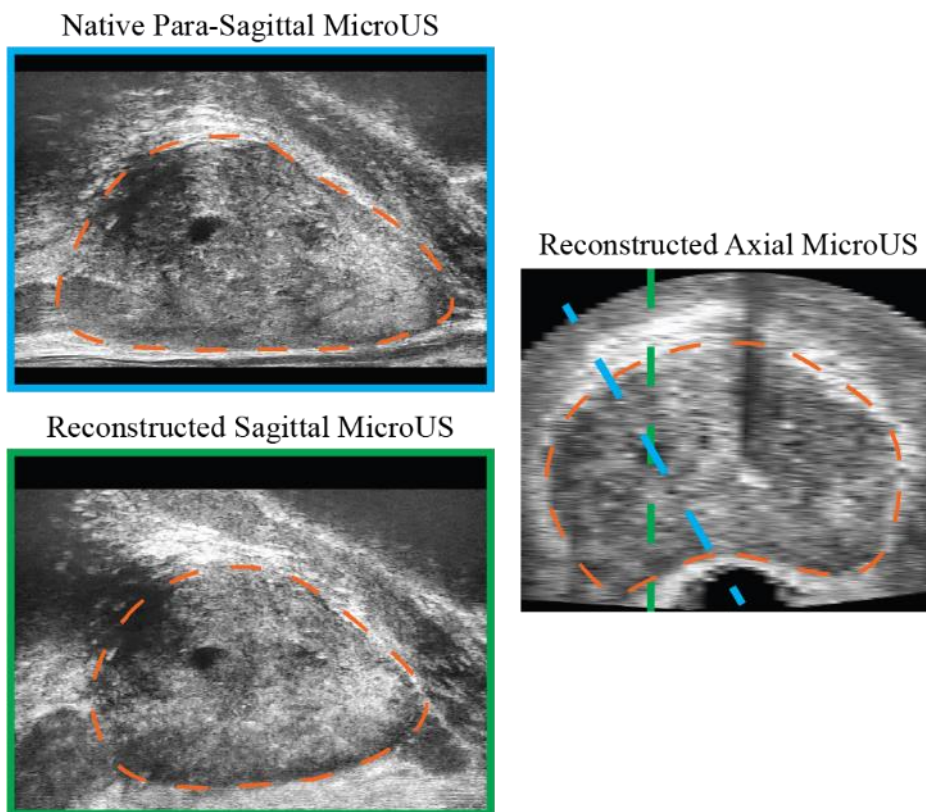


Figure 42: All micro-US images reconstructed from native Para-Sagittal sweep to true sagittal and axial orientations to facilitate registration with axial whole mount slides

through positioning the native microUS images in the approximate orientation they were acquired using the recorded angle associated with each image. Serial, parallel images in the sagittal orientation were then populated with native microUS data that intersects the sagittal image location. The resulting reconstructed microUS images allow approximation of the 3D prostate volume and subsequent co-registration with MRI and WM.

Following data collection and microUS reconstruction, a subset (N=24) of the 80 prostatectomy cases were labeled with the delineated WM lesions via the previously developed registration process (Figure 43) [220]. To accomplish this, the WM slides were
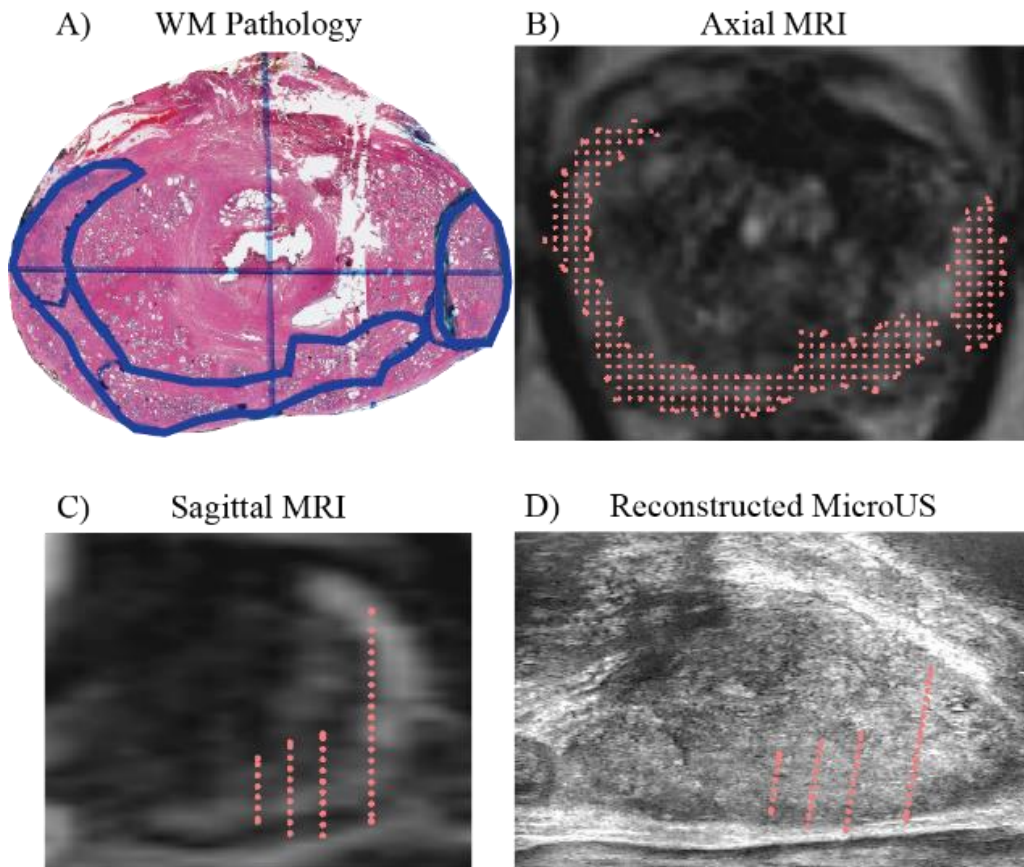


Figure 43: Data labeling process: delineated WM slides (A) are registered to matched MR images (B). This process is performed for each WM slide easily visualized in the sagittal plane (C). The delineations are then registered to sagittal microUS images (D) allowing for confidently labeled images.

first matched to an associated MRI image based on the slice location of the MRI scan and the recorded location of the WM slice from the sectioning mold and intraprostatic fiducials. Next, the WM delineations were registered to the associated MRI through prostate capsule-based registration consisting of 1) rigid alignment (alignment, rotation, and scaling) and 2) non-rigid alignment (thin-plate-spline transformation) to account for tissue deformation. Lastly, the 3D MRI scan was registered to the 3D reconstructed microUS images via a landmark-based registration consisting of 1) visual alignment of the prostate volumes (alignment and rotation), and 2) a thin-plate-spline registration based on common landmarks visible on both MRI and reconstructed microUS. The resulting transformation matrices from each registration step were subsequently used to map the WM delineated lesions from WM pathology to reconstructed microUS. After registration, it was possible to visualize the WM identified lesions on reconstructed microUS. Since WM pathology is orthogonal to reconstructed microUS, and it is also designed to capture the largest and most aggressive lesions to ensure negative margins from surgery, it is possible to represent the extent of the cancerous lesions across images despite the gaps between WM slides. Thus, allowing for accurate identification of any microUS frames containing cancer. All frames that contained clinically significant cancer (GG2 or greater) were labeled as cancerous and images without clinically significant cancer were labeled as benign.

An additional 20 microUS scans of patients with either negative (benign) biopsy results (N=14) or only clinically insignificant (GG1) cancer (N=6) were added to supplement the dataset to make it more representative of a true patient population [221]. Only GG1 and benign biopsy patients were included since an ideal classifier would detect only clinically significant cancer, therefore, these patients can be accurately

labeled as benign without being affected by cancer localization challenges for biopsy data. This data was only used for training because all biopsy data is labeled benign. In theory, if a classifier could identify if data was collected from a biopsy patient rather than a prostatectomy patient, it would be able to classify tissue as benign through confounding means. For this reason, the biopsy patients were used for training purposes only and never included in a test set.

In summary, a total 100 microUS scans of patients were available for model training from the dataset collected in chapter 3. While data was being labeled, there was a concurrent effort to perform a preliminary evaluation of deep learning classifiers using a subset of patients (N=15) detailed in the following section. These results were then ultimately confirmed in a more comprehensive study utilizing the full dataset of 100 patients.

# 6.3.　Preliminary Evaluation: Deep Learning Classifiers

## 6.3.1.　Methods

For each of the 15 patients, the reconstructed microUS images were classified as either benign or cancerous depending on if the image contained any points from the WM annotation point cloud. Out of the 15 patients, 3 patients (20%) were randomly selected as a test set with the remaining 12 saved as the training set. A random sample of 20% of the training data was used as the validation set and all images were resized to 256 x 256 pixels. Utilizing transfer learning to account for limited data, ConvNeXt models were fine-tuned on the reconstructed microUS images to classify images as benign or cancerous using the fastai and PyTorch Image Models (timm) libraries [222], [223]. Each

model utilized Adam optimization and binary cross-entropy loss, but the specific size of the model, the learning rate, and the number of epochs varied. After training a simple ensemble model was made using the mean prediction of the 5 models. The 5 models and the ensemble model were subsequently evaluated against the test set with any cancer prediction greater than or equal to 0.5 considered a cancerous classification. The true positive, false positive, true negative, and false negative rates were recorded for each model. Receiver operating characteristic (ROC) curves were created and the AUC was calculated.

An expert reader (urologist with extensive microUS biopsy experience) and a novice reader (urologist without microUS biopsy experience) were also tasked with classifying the same dataset as benign or cancerous by delineating any suspicious lesions on the native microUS images from the test set. Prior to this, both reviewers completed an online training module to ensure baseline competency. The reader annotations were then resampled to align with the reconstructed microUS images, and any images containing a reviewer delineation were marked as a cancerous classification. If a reconstructed image did not contain any delineation, it was considered a benign classification. The expert and novice classification results were then compared against the fine-tuned ConvNeXt models.

## 6.3.2. Results

The median patient age was 63 [58-68] (median [IQR]), with a median preoperative PSA of 8.8 [6.3-11.3]. There was a total of 24 distinct lesions with the following grade group (GG) distributions from low-grade to high-grade: GG1: 5, GG2: 12, GG3: 5, GG4: 1, GG5: 1. Native microUS images were acquired in a 300-frame sweep of the prostate. A total of

977 reconstructed images with a slice thickness of 1mm were created from all 15 patients with an average of 65 images per patient. A total of 82 WM slides were collected from all 15 patients with approximately 5-6 slides per patient.

Registration of pathologist delineated cancer was performed for all 15 patients. As evaluated in our earlier work, the two-step registration process had a combined registration error of 3.90±0.11mm, similar to commercial MR-US biopsy fusion systems [21], [220]. From the 977 reconstructed microUS images, 603 images were benign and 374 images were cancerous. If any points from the delineated point cloud were within ±0.5mm of the reconstructed slice location, the image was labeled as cancerous. An example of labeled data can be seen in Figure 43D. All prostate locations within the training set contained both cancerous and benign examples across the 12 patients.

After labeling, the dataset was split into training, validation, and test sets. Three patients were randomly selected to comprise the test set which consisted of 128 benign images and 71 cancerous images. The remaining 12 patients comprised the training and validation sets consisting of 475 benign images and 303 cancerous images with a random 20% of these images used as a validation set. The ConvNeXt base 22k model fine-tuned for 20 epochs with a learning rate of 0.01 had the best performance in the validation set, with an error rate of 5.8% and a loss of 0.20 (Table 3).

Table 3: Model parameters for training and validation set performance.

TABLE 3

| Model | Model Parameters and Training Metrics | | | |
| --- | --- | --- | --- | --- |
| | *Epochs* | *Learning Rate* | *Error Rate* | *Validation Loss* |
| 1) ConvNeXt Small 1k | 5 | 1E-3 | 0.142 | 0.304 |
| 2) ConvNeXt Small 1k | 22 | 1E-4 | 0.142 | 0.325 |
| 3) ConvNeXt Base 22k | 5 | 1E-3 | 0.110 | 0.353 |
| 4) ConvNeXt Base 22k | 20 | 2E-3 | 0.103 | 0.294 |
| 5) ConvNeXt Base 22k | 20 | 1E-2 | 0.058 | 0.200 |

Table 4: Model and human performance against the test set.

TABLE II

| Model | Model and Human Reviewer Evaluation on Test Set | | | |
| | *Sensitivity* | *Specificity* | *False Positive Rate* | *False Negative Rate* |
| --- | --- | --- | --- | --- |
| 1 | 0.789 | 0.648 | 0.352 | 0.211 |
| 2 | 0.775 | 0.695 | 0.305 | 0.225 |
| 3 | 0.690 | 0.727 | 0.273 | 0.310 |
| 4 | 0.704 | 0.711 | 0.289 | 0.296 |
| 5 | 0.789 | 0.688 | 0.313 | 0.211 |
| Ensemble | 0.789 | 0.727 | 0.273 | 0.211 |
| Expert | 0.606 | 0.805 | 0.195 | 0.394 |
| Novice | 0.275 | 0.853 | 0.147 | 0.726 |

The five models and an ensemble model were trained and evaluated against the test set and human reviewers (Table 4). All models exhibited similar performance but the best model was the ensemble model which had a sensitivity of 78.9% and specificity of 72.7%. The false positive and false negative rates were 27.3% and 21.1% respectively and the AUC of the model was 0.802. The AUC for the five models was: 0.813, 0.789, 0.788, 0.778, and 0.819 respectively. The receiver operating characteristics of the three models
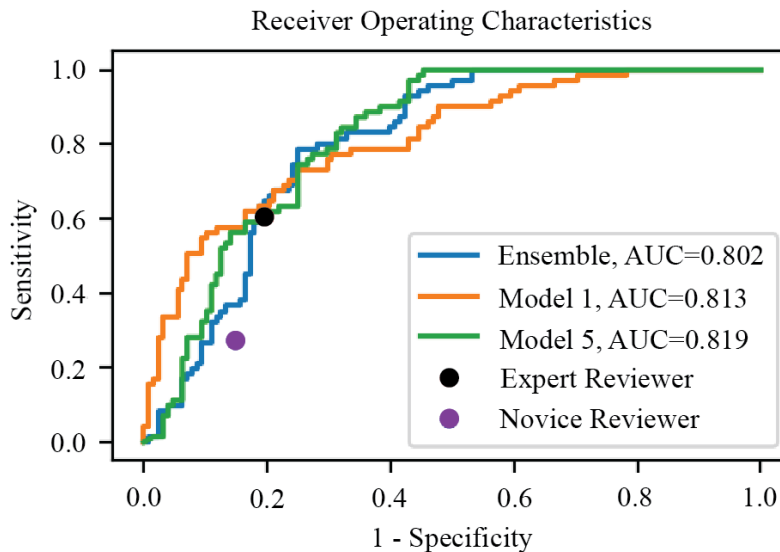


Figure 44: Receiver operating characteristics curve for the three models with the highest AUC. The models exhibited similar performance to the expert reviewer, while outperforming the novice reviewer.

Test Set Predictions
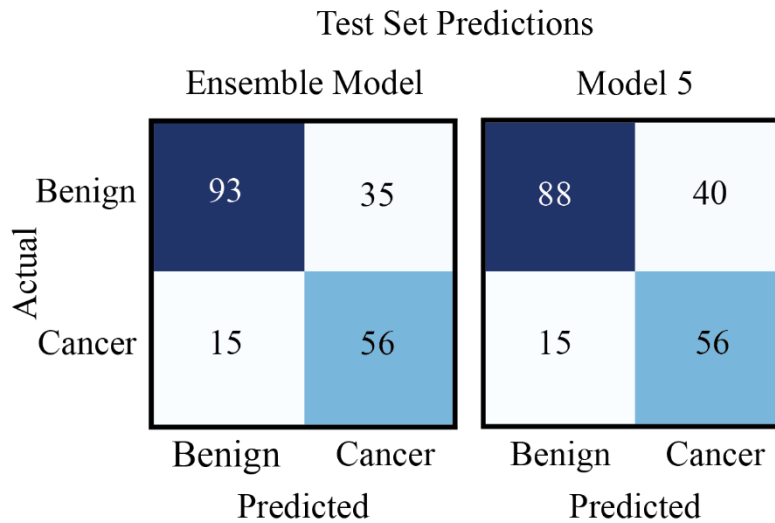


Figure 45: Confusion matrix of the ensemble model and model 5 for the
test set.

with the highest AUC can be seen in Figure 44. The test set confusion matrix for the

ensemble model and model 5 can be seen in Figure 45. These two models are represented

due to having the highest sensitivity and specificity (ensemble) and AUC (model 5).

The novice reader reviewed 2 of the 3 cases in the test set and the expert reader

reviewed all 3 cases in the test set. The expert reviewer had a sensitivity of 60.6% and a

specificity of 80.5%, while for the novice reviewer had a sensitivity of 27.5% and a

specificity of 85.3%. The performance of each reviewer is compared against the highest

performing model in Table 4 and can be visualized on the ROC curve in Figure 44.

### 6.3.3. Discussion

MicroUS has recently been introduced as a potential inexpensive alternative to mpMRI

for identifying prostate cancer, with early studies demonstrating similar performance

when interpreted by expert reviewers [17], [109], [220]. However, extensive training is

required for users to interpret microUS images correctly. Less experienced reviewers

exhibit substantially inferior performance than expert reviewers making it more difficult for new users to adopt the technology. Moreover, as tumor localization is non-trivial, the added cognitive load during biopsies may present a barrier to adoption.

Here, we explore the potential of deep learning to automate prostate cancer detection. Using a limited dataset of 15 patients, we registered ground truth, WM-pathology confirmed lesions to reconstructed microUS images to generate a confidently labeled dataset of 977 images for model training and evaluation. Due to limited availability of WM data and differences in imaging orientations, confidently labeled prostate cancer on microUS imaging is challenging. To our knowledge, this is the first work to create a confidently labeled microUS prostate cancer dataset for deep learning applications. This is important as high-quality labeled data is vital for creating accurate deep learning models.

Following data labeling and utilizing transfer learning, we trained five deep learning models and one ensemble model to classify microUS images as benign or cancerous. Evaluation of these models against expert and novice readers on patients that were excluded from training was encouraging with AUC values of 0.8. All six models had better sensitivity than the novice reviewer when evaluated at a prediction cutoff of 0.5. Moreover, through examining the ROC curves, the models achieved improved specificity at the same sensitivity as the novice reviewer. In regards to the expert reviewer, the models exhibited similar performance, with slightly improved sensitivity and slightly diminished specificity. Furthermore, the ROC curves approximately intersect the expert reviewer, indicative of similar performance. These results are particularly encouraging given these models were trained on a relatively limited dataset indicating that performance should improve with the addition of more training data. This also

highlights the value of confidently labeled microUS images, as the models were able to achieve similar performance to an expert reader despite having limited training data. Furthermore, the deep learning models trained in this work were relatively simple, consisting of ConvNeXt image net architectures fine-tuned for 5 to 20 epochs. Given preliminary models exhibited similar performance to expert reviewers, more complex models will likely yield better results.

It is worth noting the expert reviewer sensitivity (61%) and specificity (81%) are not directly comparable to other microUS studies particularly those focusing on biopsies. Here we evaluated a reviewer on every image in the sweep in order to compare against a classifier model, providing more opportunities for evaluation per patient than biopsy-based studies focusing on targeting cancer foci. As such, biopsy studies have higher sensitivity and lower specificity, such as in the work by Klotz *et al.* (94% and 22% respectively) [109].

There are some inherent limitations of this work. Firstly, these models were tasked with classification and not segmentation. As such it is not possible to confirm that model classification is not due to confounding factors. Another limiting factor is the small dataset of prostatectomy patients used in this study, which may not be generalizable to biopsy patients or patients with types of prostate cancer not represented in this dataset. Additionally, the models were trained on non-native, reconstructed microUS images. Although the reconstructed images preserve much of the native image quality, the models may not be generalizable to native microUS images. We plan to address much of these limitations in future work through increasing the size dataset, training on native microUS images, as well as training segmentation models for cancer delineation.

# 6.4. Expanded Evaluation: Deep Learning vs Experts

## 6.4.1. Methods

Using the 100 patients, three different groups of models were trained: 1) Confident models, using the confidently labeled data from the 24 patients co-registered with WM pathology, 2) Weak models, using the remaining unlabeled microUS data assigned labels by the confident models, and 3) Combined models, using both confident and weakly labeled data. Each group of models was evaluated via 3-fold cross validation and the results were compared against both expert microUS and MRI reviewers.

**Urologist and Radiologist Review**

In order to set an appropriate benchmark for model evaluation, 3 expert urologist readers (extensive microUS biopsy experience) and 1 novice urologist reader (experienced urologist with no microUS biopsy experience) were asked to delineate cancer on native microUS imaging (not reconstructed microUS) for a randomly assigned 8 prostatectomy patients out of a subset of 15 patients. Each patient was reviewed by at least two readers allowing for cross-validation between patients and reviewers. Due to reader availability only 15 out of the 24 prostatectomy patients were reviewed. The microUS readers were aware the patient had a prostatectomy for prostate cancer, but were blinded to all other clinical variables. Additionally, an expert radiologist reviewed each of the same 15 patients associated mpMRI. The MRI was reviewed twice, once prior to biopsy without knowledge of the patient having prostate cancer using the standard PI-RADSv2 scoring system (pMRI) as a part of routine clinical practice, and once retrospectively (rMRI) with the knowledge the patient had a prostatectomy for prostate

cancer, the same knowledge as the microUS readers. Since PI-RADSv2 is designed to prioritize specificity to avoid unnecessary biopsies, the delineated lesions are often much smaller than the true tumor extent. As such, it was important to re-review the MRI with a focus on identifying tumor extent to allow for a more direct comparison with microUS and associated deep learning models.

Following reviewer delineation of suspicious lesions, the marked regions of interest (ROIs) were registered to reconstructed microUS via the stored transformation matrices from data labeling in the previous section. If a reviewer delineation was present on a reconstructed microUS image, the image was counted as a reviewer classification of cancer. If the image did not contain a reviewer delineation, it was counted as a reviewer classification of benign tissue. Evaluation of the reviewer classifications against the ground truth results in directly comparable benchmarks for deep learning-based classification models.

**Preliminary Training and Confident Models**

A total of 44 reconstructed microUS patient scans were used for confident model development, 24 of these patients were confidently labeled prostatectomy patients (924 benign images, 543 cancerous images) and 20 were biopsy patients (1,308 benign images) for a total of 2,775 reconstructed microUS images. All images were converted to PNG file types and normalized between 0 to 255 as greyscale images. For 3-fold cross-validation, the 24 confidently labeled prostatectomy patients were divided into three sets of 8 were withheld from training: test sets A, B, and C. Test set A was specifically chosen to be the same patients reviewed by the highest performing expert microUS reviewer for straightforward comparison. The remaining 16 prostatectomy patients were randomly assigned to test sets B and C. The same 3 test sets were uses across all three

groups of models trained (confident, weak, combined). The data distribution can be seen in Figure 46.

For preliminary training, 3 different types of image net models were imported using the timm library (ConvNext, ViT, and SwinV2) [223], [224], [225]. These models were tasked with classifying reconstructed microUS images as benign or cancerous and fine-tuned on 36 patients of reconstructed microUS images (16 prostatectomy and 20 biopsy) that were compressed to fit to the specifications of each architecture (ConvNext: 256x350, ViT: 224x224, SwinV2: 256x256). Models were trained using PyTorch for 20 epochs, a learning rate of 0.01, and binary cross-entropy. A random selection of 20% of the images from the 36 patients were selected as a validation set. This was decided over withholding entire patients to allow for models to train on more instances of cancer, since prostate cancer can appear differently on microUS for different patients, evident in the Prostate Risk Identification using Micro-Ultrasound (PRI-MUS) scoring system, and the training set being limited to a small dataset of 16 patients containing clinically significant prostate cancer. Batch transformations of brightness and contrast were randomly applied to every image in the training set to increase the generalizability of the data. The models with the lowest validation error and validation loss were selected for use in the remainder of the study in an equal voting ensemble. The best performing models were the ViT Large model and the SwinV2 Small model.
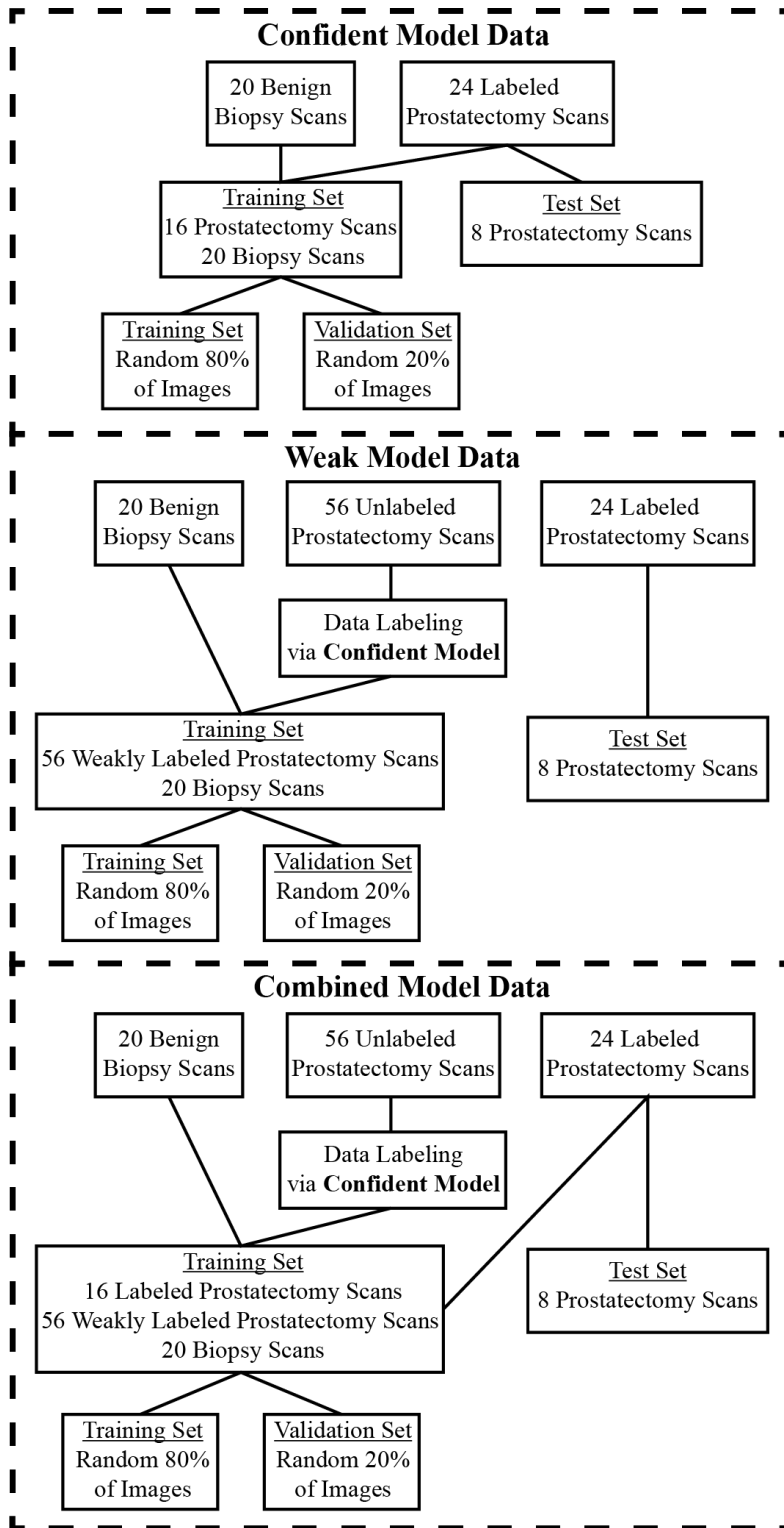
Figure 46: Data distribution for each group of models. The confident models (top) were used to label the unlabeled images for the weak (middle) and combined (bottom) models.

155

For each test set (A, B, and C) an equal voting ensemble was trained using the 36 patients not contained in the 8-patient prostatectomy test set. For each ensemble a ViT Large and SwinV2 Small model were trained using the same parameters from preliminary training, and the mean of the two predications was used as the ensemble prediction. These predictions were then evaluated per image against the associated 8 patient test that was withheld from training. Receiver operating characteristic (ROC) curves were plotted for each of the 3 cross-validation models alongside the results from the human reviewers for benchmark comparison. The AUC was also calculated along with the sensitivity at a specificity of 0.8.

**Weak Models**

The reconstructed microUS images from the 56 unlabeled prostatectomy patients (N = 3,411) were assigned labels by each of the 3 confident ensemble models (A, B, and C) trained in the prior section. The labels were softened by 0.2 due to possible prediction error, resulting in cancer predictions assigned a label of 0.8 and benign predictions assigned 0.2. The 20 biopsy patients were also included and assigned a label of 0 as they are confirmed benign.

Each group of labels (A, B, and C) were used to train an equal voting ensemble model comprised of a ViT Large and a SwinV2 Small model using the same training parameters as the confident models (20 epochs, 0.01 learning rate, brightness and contrast batch transforms). As with the confident models, a random sample of 20% of the training images were withheld and used as a validation set during training. However, these weak models were trained as regression models instead of binary classifiers due to labels ranging on a scale of 0-0.8. As such, root mean squared error (RMSE) loss was used instead of binary cross-entropy loss. The equal voting ensemble predictions were

156

similarly defined as the mean prediction between the ViT and SwinV2 models. The three weak ensemble models were evaluated against their respective 8 patient test set (A, B, and C) that was withheld from training for both the weak model and the confident model used for labeling. The three 8 patient test sets are the same as the three test sets used in the prior section. The 3-fold cross validation results were similarly plotted against the benchmark expert reviewers alongside the AUC and sensitivity at a specificity of 0.8.

**Combined Models**

Three combined models (A, B, and C) were trained using a combination of confidently labeled data and weakly labeled data. More specifically, the confidently labeled data was comprised of the same 16 prostatectomy patients used to train each respective confident model (A, B and C). The images labeled as cancer were assigned a label of 1 and images labeled as benign were assigned a label of 0. As with the weak model training, the reconstructed microUS images from the 56 unlabeled prostatectomy patients were assigned labels using the classification predictions of the 3 previously trained confident models (A, B, and C) to create three sets of labels. Each set of labels was used by its respective combined model. For example, combined model A used the data labels created by confident model A. As before, each predicted label was softened by 0.2 due to possible prediction error, giving images labeled as cancerous a value of 0.8 and images labeled as benign a value of 0.2. The 20 benign biopsy patients were also included and assigned a label of 0. In summary each training set was comprised of 16 confidently labeled prostatectomy patients, 56 weakly labeled prostatectomy patients, and 20 benign biopsy patients. As mentioned earlier, the same 8 patient test sets from the prior sections were used to evaluate the respective combined model and these patients were never included in training or in any training of models used to label data for training.

157

As with the weak and confident models, three equal voting ensembles were created, each comprised of a ViT Large model and a SwinV2 Small model fine-tuned on the respective training data (A, B, and C). The same training parameters as used in the confident and weak models were used here (20 epochs, 0.01 learning rate, brightness and contrast batch transforms). Additionally, a random sample of 20% of the training images were withheld and used as a validation set during training. The combined models were trained as regression models using RMSE loss. The mean predictions of the ViT and Swin models were used to generate the ensemble prediction. Each combined ensemble (A, B, and C), were evaluated against the respective confidently labeled 8 patient test set, which was withheld from training and any training of models used to label training data. The 3-fold cross validation results were similarly plotted against the benchmark expert reviewers alongside the AUC and sensitivity at a specificity of 0.8.

In summary, three groups of classifiers were trained described by the quality of labels used for training: confident, weak, and combined models. Each group was subsequently evaluated via 3-fold cross validation of an 8-patient test set with WM pathology confirmed labels. The three test sets were the same across each group of classifiers and where never included in any training and the related training of any data labeling models for the respective subset of models. For example, test set A was withheld from confident model A, weak model A, and combined model A. This was repeated for test sets B and C. Benchmark classification results were also collected from expert reviewers for both MRI and microUS, allowing for direct comparison with model classification results.

**Clinical Evaluation**

To evaluate the potential usefulness of the trained classifiers for guidance in a biopsy setting, the highest performing group of classifiers were evaluated on cancer detection metrics in the test sets. More specifically, the number of index lesions correctly identified was recorded, defined as the largest and most aggressive tumor responsible for disease progression. The total number of significant lesions correctly identified, as well as the ratio of false positive and false negative classifications per patient were also recorded. These metrics were also calculated for the expert reviewers for both MRI and microUS allowing for direct comparison. Additionally, results were further stratified by scenarios where expert reviewers and the classifiers agreed on a classification and where they disagreed to further examine the value of correspondence between the models and experts. Lastly, we attempted to evaluate the ability of the model predictions as a potential targeted biopsy guide. Using the model predictions, the image with the highest cancer prediction was defined as the location for a targeted biopsy. Assuming ideal placement of the biopsy core within the image, if the target location overlaps the location clinically significant cancer, it was counted as a correct target. If the targeted location does not overlap a clinically significant lesion, it was counted as an incorrect target. While assuming perfect targeting within an image may not be representative of real biopsies, targeted biopsies normally consist of at least 3 cores spread around a region of interest. By only analyzing one image, it is safe to assume that if 3 cores were taken from 3 locations within an image (i.e. apex, mid-gland, and base), that one of those cores would correctly capture the lesion.

## 6.4.2. Results

MicroUS scans of the prostate from 80 prostatectomy patients and 20 biopsy patients were reconstructed from para-sagittal to true sagittal. A total of 6,264 reconstructed microUS images (4,956 prostatectomy and 1,308 biopsy) were created from the patient scans. Out of the 80 prostatectomy scans, 24 scans were labeled as benign or cancerous on a per image basis depending on the co-registered location of WM pathology confirmed clinically significant cancer. From these 24 patients, 924 images were benign and 543 were cancerous. The distribution of cancer and benign images by location across the 24 patients can be seen in Figure 47. It is worth noting that there was no cancer present on the edges of the sweep, this was because these images were outside the prostate, however these images are still worth including because an ideal model would be able to
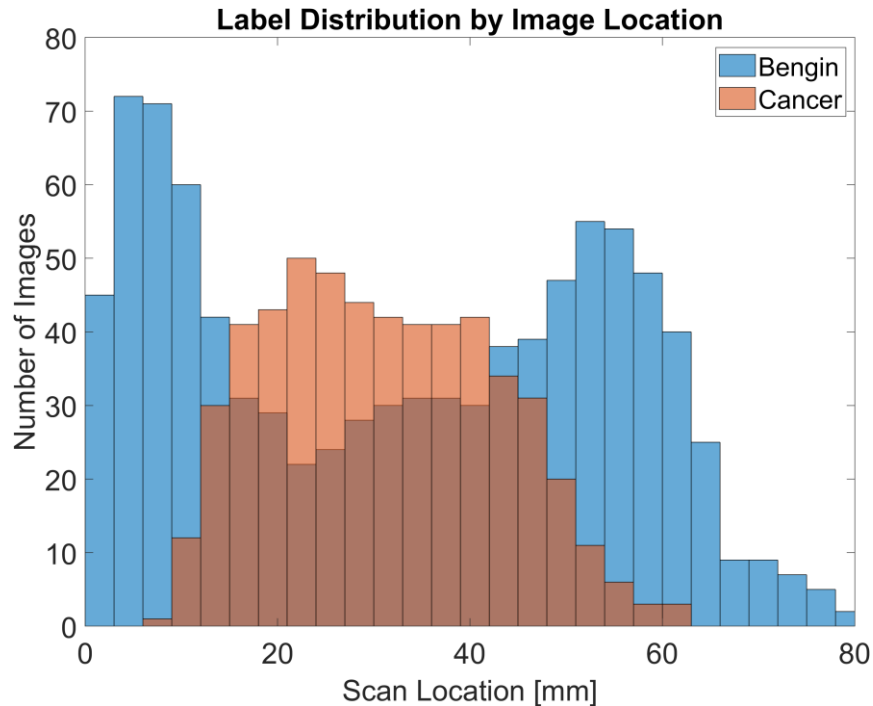


Figure 47: Cancer distribution by position of reconstructed microUS images for 24 confidently labeled prostatectomy patients. Scan location goes from anatomical right to left with the entire prostate fully represented. A few images on the far right and far left are outside of the prostate and always benign.

160

recognize the boundaries of the prostate. There was a total of 32 clinically significant lesions and 24 index lesions across the 24 labeled prostatectomy patients.

**Urologist and Radiologist Review**

For the reviewer classifications the 3 expert microUS reviewers had an average sensitivity of 0.71±0.06 (0.79, 0.69, and 0.64 respectively). The average reviewer specificity was 0.83±0.04 (0.80, 0.89, and 0.80 respectively). The urologist with no microUS biopsy experience had a sensitivity and specificity of 0.27 and 0.88, highlighting the importance of microUS experience for effective review. For MRI the prospective, standard of care, PIRADSv2.1 (pMRI) review had a sensitivity of 0.44 and a specificity of 0.97, while the retrospective MRI (rMRI) review had a sensitivity of 0.73 and specificity of 0.86.

**Preliminary Training and Confident Models**

For preliminary training varying types of models were trained with varied parameters. The top two models from preliminary training by validation set loss and validation error rate (Table 5) were selected to create an equal voting ensemble. The ViT Large model and SwinV2 Small models had the best performance with a validation set loss of 0.13 for both and an error rate of 0.2 and 0.4 respectively.

Table 5: Preliminary model training used to find a final model. The two models with best performance are bolded.

| | | Preliminary Model Training | | | |
|---|---|---|---|---|---|
| *Model* | *Size* | *Image Size* | *Training Set* | *Validation Loss* | *Validation Error Rate* |
| ConvNeXt | Small | 256x350 | All | 0.14 | 0.04 |
| ConvNeXt | Small | 256x350 | Prostatectomy | 0.37 | 0.11 |
| ViT | Small | 224x224 | Prostatectomy | 0.37 | 0.09 |
| ViT | Small | 224x224 | All | 0.22 | 0.05 |
| SwinV2 | Small | 256x256 | Prostatectomy | 0.23 | 0.05 |
| **SwinV2** | **Small** | **256x256** | **All** | **0.13** | **0.04** |
| **ViT** | **Large** | **224x224** | **All** | **0.13** | **0.02** |
| SwinV2 | Large | 192x192 | All | 0.16 | 0.04 |
| SwinV2 | Base | 256x256 | All | 0.16 | 0.04 |
| ConvNeXt | Base | 256x350 | All | 0.20 | 0.03 |

The 24 patients were divided into 3 test sets of 8 patients for cross-validation. Test set A was assigned the same 8 patients reviewed by the highest performing expert reviewer, while the remaining 16 cases were randomly assigned to test sets B and C. One case in test set B was removed for evaluation due to the patient having a prior transurethral resection of the prostate (TURP). This was the only patient in the dataset with a TURP and as expected the trained models exhibited poor performance when tested on this case. While each group of models did correctly identify the region of cancer, the models had poor classification in the regions where the TURP was visible on imaging. This case was left in during training for models evaluated on test sets A and C in an attempt to improve generalization and was not reviewed by any of the human reviewers.

Using 3-fold cross validation, the confidently trained models had an average sensitivity of 0.72±0.07 (0.82, 0.70, and 0.64 respectively) at a specificity of 0.8. The average AUC value was 0.81±0.04 for the three test sets (0.87, 0.76, and 0.81 respectively). These results can be visualized in Figure 48 and Table 6. These results indicate similar performance relative to the expert microUS reviewers.

Table 6: Model and human benchmark performance against the test sets. Each cross-validation group is denoted as A, B, and C.

| | Test Set Performance | | | |
| --- | --- | --- | --- | --- |
| Model Type | Data Group | Sensitivity | Specificity | AUC |
| Confident | A | 0.82 | 0.8 | 0.87 |
| | B | 0.70 | 0.8 | 0.76 |
| | C | 0.64 | 0.8 | 0.81 |
| | **Mean** | **0.72** | **0.8** | **0.81** |
| Weak | A | 0.77 | 0.8 | 0.84 |
| | B | 0.75 | 0.8 | 0.78 |
| | C | 0.62 | 0.8 | 0.77 |
| | **Mean** | **0.71** | **0.8** | **0.80** |
| Combined | A | 0.79 | 0.8 | 0.85 |
| | B | 0.78 | 0.8 | 0.79 |
| | C | 0.58 | 0.8 | 0.81 |
| | **Mean** | **0.72** | **0.8** | **0.82** |
| Expert | 1 | 0.79 | 0.80 | |
| | 2 | 0.69 | 0.89 | |
| | 3 | 0.64 | 0.80 | |
| | **Mean** | **0.71** | **0.83** | |
| | **pMRI** | **0.44** | **0.97** | |
| | **rMRI** | **0.73** | **0.86** | |
| Novice | 1 | 0.27 | 0.88 | |

## Weak Models

For the weak models, the sensitivity was 0.71±0.07 (0.77, 0.75, and 0.62 respectively) at a specificity of 0.8. The average AUC was 0.80±0.03 across the 3 test sets (0.84, 0.78, 0.77 respectively). These results can be visualized in Figure 48 and Table 6. These results relative to the expert reviewers further support the ability of the confident models to correctly classify images, as the training data was labeled by the respective confident models that
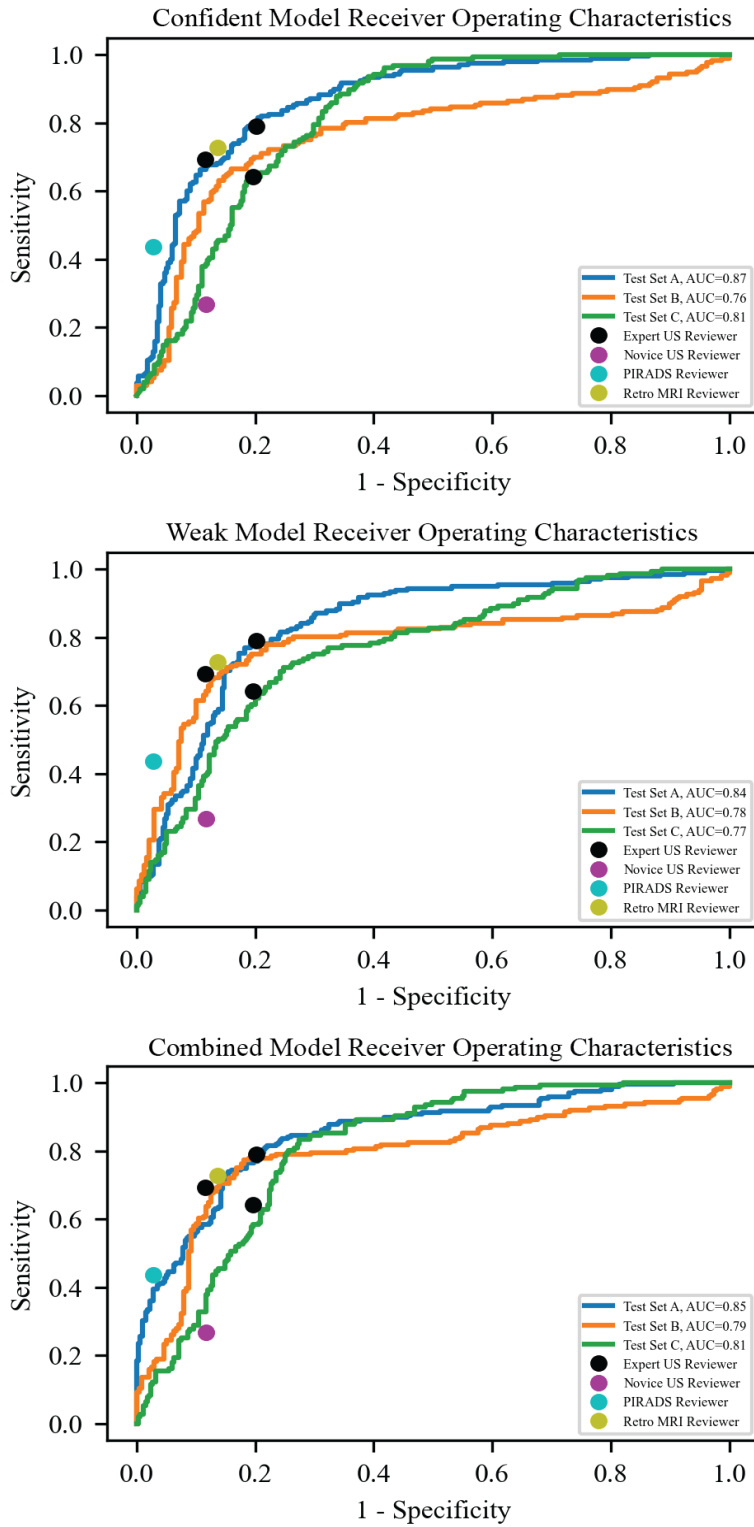
Figure 48: ROC plot of confident, weak, and combined models (top to bottom) with benchmark points of expert reviewers. Each curve represents one cross-validation model

were also blinded to the test set. Models subsequently trained on this model-labeled data approach expert level performance.

**Combined Models**

For the combined models, the sensitivity was 0.72±0.10 (0.79, 0.78, and 0.58 respectively) at a specificity of 0.8. The average AUC was 0.82±0.02 across the 3 test sets (0.85, 0.79, 0.81 respectively). These results can be visualized in Figure 48 and Table 6. While the three groups of models exhibited similar performance to each other and the expert reviewers, the combined classifier had slightly higher performance. Due to this it was used to in the subsequent clinical evaluation.

**Clinical Evaluation**

The combined classifiers had the highest sensitivity and AUC of the classifier groups. These models correctly classified 22/23 (95.7%) of index lesions and 28/31 (90.3%) of the total lesions. Comparatively the benchmark expert microUS reviewers identified 8/8 (100%) of their respective index lesions and 93.0±5.0% (9/10, 8/9, 10/10) of total lesions on microUS. The novice urologist identified 6/8 (75%) of index lesions and 6/9 (66.7%) of total lesions. Expert MRI review identified 13/15 (86.7%) and 15/15 (100%) of index lesions with 13/18 (72.2%) and 15/18 (83.3%) of total lesions for pMRI and rMRI respectively. Furthermore, the classifiers had an average false positive classification of 7.1 frames per patient and a false negative classification of 7.3 frames per patient, while the expert reviewers had 7.0 (8.3, 4.9, 8.0), 1.2, 5.8 false positive classifications per patient and 7.0 (5.1, 7.0, 8.9), 12.8, 6.2 false negative classifications for microUS, pMRI, and rMRI respectively. Examples of patient specific classifications can be seen in Figure 49.

Stratifying by instances where the expert reviewer and the classifier both predicted cancer, there was an average precision of 0.78 (0.81, 0.83, 0.69) for the three microUS reviewers. For pMRI and rMRI the precision when the model and reviewer
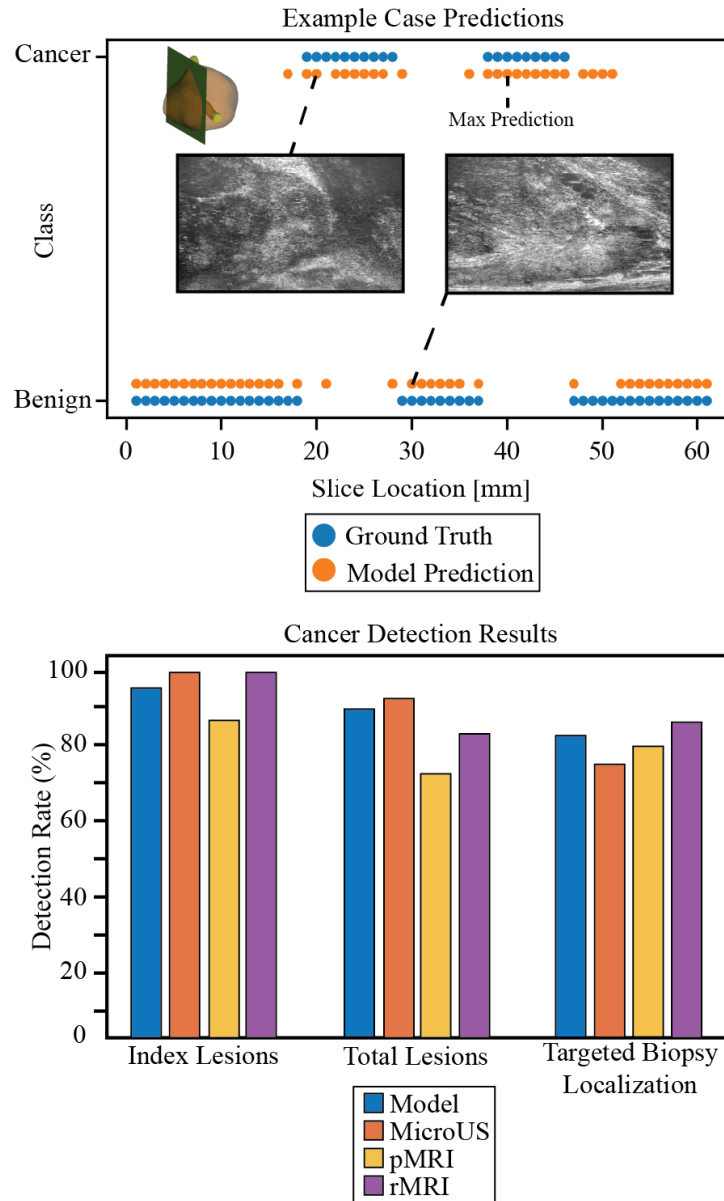


Figure 49: Top: Example patient specific plot with model predictions (orange dots) and ground truth (blue dots) for each image in a scan. X-axis is location within parallel sweep through prostate. Bottom: Cancer detection results for the deep learning model and expert reviewers. Targeted biopsy localization refers to the highest prediction within a scan for the model and the center of prediction for the expert reviewers.

predicted cancer was 0.98 and 0.81 respectively. Comparatively, the individual precision was 0.69 for the classifier, 0.71 (0.70, 0.76, 0.66) for expert microUS reviewers, 0.89 for pMRI, and 0.74 for rMRI. Indicating prediction confidence increases classifier predictions, particularly for pMRI.

Evaluating the ability of the microUS classifier to identify a potential targeted biopsy location, a total of 19/23 (82.6%) targets captured a clinically significant lesion (18 index lesions and 1 non-index lesion) using the maximum of the model predictions. Using the center of the positive predictions of the expert reviewers as a comparison, a total of 75% (6/8, 7/8, 5/8) of targets correctly localized a clinically significant lesion on microUS while 12/15 (80%) and 13/15 (86.7%) targets correctly localized a clinically significant lesion on pMRI and rMRI respectively.

## 6.4.3. Discussion

In this section, we have developed the first whole image classifier for B-mode microUS imaging through co-registration with ground truth WM pathology. We have also evaluated the classifier results in context with direct comparison to MRI with an expert radiologist, the clinical standard, as well as expert microUS readers.

In summary, microUS scans were labeled with ground truth WM pathology via a previously developed registration process [220]. Any images with confirmed clinically significant cancer were labeled as cancerous. Since WM pathology is designed to capture the tumor extent and ensure negative margins from surgery, this form of histology provides the most accurate possible representation of cancer location. The reported registration error from our prior work was 3.90±0.11mm (mean ± standard error), similar to the reported registration accuracy of commercial MRI-US fusion devices used

for biopsy guidance [21]. However, while unlikely, it is possible that the true extent of the cancerous lesions is up to 3-4 images wider or narrower than recorded if the registration error is solely in the left to right orientation.

Following registration, delineations from expert reviewers were registered to the reconstructed microUS images and used to determine reviewer image classifications, allowing for direct comparison with model classifier results. Additionally, the expert microUS reviewers and rMRI review were aware the patients had prostate cancer, due to knowledge the study was validated by WM pathology, a common limitation associated with WM validated studies [191], [220]. Both pMRI and the model classifiers were blinded to the patient having cancer. Additionally, fully benign patient data were added to training to make the patient populations more representative of true patient populations for the classifier. Between the expert reviewers, rMRI and microUS performed similarly as reported in other comparison studies [17], [18], [219], with sensitivities of 0.73 and 0.71±0.06 with cancer detection rates of 83.3% and 93.0±5.0% respectively. For pMRI, the specificity was higher than rMRI and microUS reviewers with a lower overall sensitivity. This can be attributed to PIRADS being optimized for specificity while maintaining good sensitivity for its use as a triage for biopsy. As a result, PIRADS delineations tend to underestimate the true tumor extent while still having high detection rates for index lesions as reported in other studies [6], [13], and evident in this study by the lower sensitivity (0.44) on the image level while still identifying 86.7% of index lesions.

For comparison with expert reviewers, we trained 3 different groups of classifier models with 3-fold cross validation. Each group of classifiers were trained using a different type of data: 1) confident classifiers trained on microUS images that were labeled from co-registration with WM pathology, 2) weak classifiers trained on microUS

images that were labeled from previously developed confident models, and 3) combined models that used a combination of confidently labeled data and weakly labeled data. The test sets used to evaluate each model were a subset of confidently labeled data that was withheld from training, including the training of models used to create weakly labeled data. Each model was comprised of ViT and SwinV2 vison transformer models fine-tuned on the microUS data. All three groups of models performed similarly to the expert microUS reviewers in terms of sensitivity (0.71-0.72 vs 0.71 respectively) and specificity (0.8 vs 0.83 respectively). Furthermore, when analyzed on the patient level, the combined models identified a similar percentage of index lesions and total lesions as the expert reviewers, emphasizing the potential utility of these models for cancer detection. These models could alleviate the training requirement for novice users, allowing for urologists not experienced with microUS to achieve expert level detection.

Perhaps most interestingly, the max prediction of the classifier correctly identified clinically significant cancer in 82.6% of the patients. For perspective, the center of PIRADS ROIs, the gold standard, captured clinically significant cancer in 80% of patients. This is important as the primary role of imaging for prostate biopsy is to guide targeted biopsy to confirm prostate cancer. As such, these classifiers could potentially be used to identify a targeted biopsy location providing biopsy guidance for urologists with a similar accuracy to PIRADS. Furthermore, if the classifier identified images of cancer also contain a PIRADS delineated ROI, there is a 98% chance there is clinically significant cancer within that image based on the results from this study.

It is worth noting that the sensitivities and specificities reported here are not directly comparable to other studies, but they are directly comparable between the reviewers and classifiers within this study. Since sensitivity and specificity are greatly

affected by the amount of positive and negative instances, it is vital for the benchmark comparisons to be evaluated in the same manner as the model classifications. This can further be highlighted by our earlier work which used the exact same reviewer delineations but with a more granular evaluation [220]. As a result, the sensitivities were much lower and the specificities were much higher than the results from this study which were analyzed on the image level and the patient level. However, while the results within this study cannot be directly compared to the sensitivity and specificity of other studies, the general trends between modalities are comparable. Specifically, microUS review having higher sensitivity but lower specificity than PIRADS is commonly reported in literature [78], [220].

Ultimately, the results presented here highlight the value of confidently labeled data and transfer learning for creating quality deep learning models especially for limited datasets. The classifiers trained in this study were able to achieve expert level performance with up to 100 patients and 6,264 images worth of data. The confident models in particular achieved expert level performance with only 36 patients of training data. To this end, future work will focus on registering more WM data to the unlabeled microUS images within this dataset to improve classifier performance. However, it is worth mentioning the weakly labeled data created from the confident models was able to train classifiers with expert level performance with 56 patients of weakly labeled data. As such the trained confident models could potentially be used to generate larger datasets for training and circumvent the time intensive registration process. Additionally, the models fine-tuned in this work are pre-trained models developed for general image classification. It is likely that the results presented in this work can be improved through model optimization in future work. Regardless the models in this

work achieved expert level performance, and application of deep learning classifiers for microUS warrant further investigation.

While these results are promising, there are some limitations of this study. Most notably, the results of this study will need to be confirmed in a larger dataset with larger test sets. Additionally, while reviewers always delineated native imaging on both MRI and microUS, the models were trained on reconstructed microUS to allow for more straightforward registration with ground truth WM pathology. Due to the reconstruction process, the reconstructed images are of slightly lower quality due to interpolation, but the images are compressed to either 224x224 or 256x256 for training, so this may not have a large impact on model performance. Furthermore, a microUS sweep could theoretically be reconstructed during biopsy for model predictions. Regardless, the difference between native and reconstructed images for model performance will be investigated in future work. Additionally, this study relied on prostatectomy patients for WM pathology. As such, the cancer in patients selecting to undergo radical prostatectomy is likely larger and more aggressive than an equivalent biopsy cohort. We attempted to mitigate this issue through the inclusion of benign biopsy patients as well as comparing directly with expert reviewers reviewing the same patient population, but ultimately this is inherent to WM based validation studies. Lastly, the models presented here are performing whole image classification, and any application of them in a clinical setting would depend on the reader to localize cancer within the image. We plan to address this in future work with a patch-based classifier that can be used to delineate the whole image classifier providing delineated regions of interest.

In summary the classifiers trained in this work achieved expert level classification performance for identification of clinically significant prostate cancer on microUS.

Implementation of these classifiers could potentially reduce the training requirement for microUS adoption as well as the inter-reader variability between expert readers.

## 6.5.  Conclusion

Deep learning approaches for prostate cancer detection on microUS have the potential to alleviate the training requirement for novice users as well as reduce the cognitive load during biopsies. We have demonstrated that transfer learning and confidently labeled datasets enable model development despite limited data availability. This is highly advantageous as large microUS datasets are not commonly available at this time.

In this chapter we developed the first whole image microUS cancer classifier trained on ground truth WM pathology. Utilizing both high-quality labeled data and transfer learning, the fine-tuned vision transformers achieved AUC values of 0.80-0.82 in three-fold cross validation. The classifiers performed as well as expert reviewers for both MRI and microUS with similar sensitivity, specificity, and cancer detection rates in an equivalent evaluation. Furthermore, the max prediction of the classifiers demonstrated high accuracy for selecting a targeted biopsy location. The results presented in this study are promising for future development of deep learning based microUS cancer classification, and warrant further investigation using larger datasets.

# CHAPTER 7

# Conclusions and Future Directions

## 7.1.   Summary

Multi-parametric magnetic resonance imaging is the gold standard imaging diagnostic for visualizing prostate cancer. While effective, mpMRI is expensive, resource intensive, requires multiple patient visits, and subject to inter-reader variability. As such it is often limited to larger centers. Recently, micro-ultrasound has been introduced as a low-cost alternative to mpMRI for PCa visualization during biopsy. While early studies have demonstrated similar cancer detection rates, sensitivity, and specificity for prostate cancer visualization, there has been limited validation against ground truth WM pathology especially using computational co-registration. While cancer detection rates and sensitivity for PCa have been promising, there is minimal quantitatively backed evidence for the improved PCa detection over non-diagnostic conventional US. Beyond that, there is a considerable training requirement to properly interpret microUS images and considerable inter-reader variability between expert readers. Ultimately, these reasons have limited microUS adoption.

The purpose of this thesis was a multi-faceted investigation into microUS imaging of prostate cancer. This body of work was focused on 1) identifying quantitative evidence for the reported diagnostic capability of microUS, 2) rigorously investigating the diagnostic potential of microUS imaging through WM pathology validated reader

studies, and 3) evaluating the ability of deep learning classifiers to identify PCa on microUS using confidently labeled data.

In order to investigate microUS imaging, it was first necessary to collect a high-quality dataset. Through a clinical trial, we collected mpMRI, microUS, and WM pathology for patients undergoing radical prostatectomy for prostate cancer. Through careful tracking of spatial orientation for each modality, the data provided the opportunity for accurate correlation between microUS, mpMRI, and WM pathology.

To determine if microUS is capable of identifying prostate cancer, we first analyzed prostate cancer morphology on WM pathology. We found the amount of ductal anatomy of prostate tissue changes in regions of cancerous tissue, and the majority of these ducts are within the resolvable limit of microUS but too small to be resolved on conventional US, providing quantifiable evidence for why microUS could potentially resolve PCa and conventional US generally cannot.

After confirming a reason for the improved diagnostic ability of microUS over conventional US, it was necessary to evaluate the performance of human readers with comparison to radiologists using mpMRI. However, before this was possible it was necessary to develop and validate co-registration techniques to allow for accurate evaluation. Two methods were developed, one utilizing ex vivo microUS allowing for capsule registration of microUS at the cost of a different imaging environment, and one for in vivo microUS utilizing intermediary image reconstruction and three-dimensional registration with MRI to allow for capsule registration to WM pathology. Using these techniques readers delineated areas of suspected prostate cancer on microUS which was registered to ground truth WM along with comparison from radiologist delineated

mpMRI. For both instances expert review demonstrated similar performance to MRI for PCa delineation.

While expert microUS reviewers demonstrated similar performance to radiologists on MRI, novice urologist review demonstrated substantially worse performance, indicative of the high training requirement. Furthermore, microUS is affected by inter-reader variability similar to MRI. For these reasons, microUS would stand to benefit from implementation of deep learning solutions. With the development of accurate co-registration between microUS and WM pathology, we were uniquely positioned to create confidently labeled data for training microUS models. These classification models demonstrated similar performance to expert microUS readers in multiple evaluations with particularly good performance for identifying sites for targeted biopsy. Beyond demonstrating the utility of deep learning for microUS, the performance of these models further proves there is information embedded in microUS images indicative of prostate cancer as models were able to identify images containing cancer with similar accuracy to expert readers.

Conclusively, in this thesis we confirmed the diagnostic capability of microUS to identify prostate cancer. We provided quantitative evidence for microUS having improved diagnostic capability over conventional US, demonstrated expert microUS readers are as diagnostic as expert radiologists on mpMRI in a reader study, and proved there is objectively visible prostate cancer in microUS images through expert level performance of deep learning classifiers. A primary contribution of this thesis is the development of an accurate co-registration methodology between microUS, MRI, and WM pathology. This provides the vehicle for the other work in this thesis. It allows for direct comparison of PCa visualization on microUS and mpMRI validated by ground truth WM

pathology. Additionally, it allows for accurate data labeling of microUS images, allowing for high-quality training data necessary for the development of microUS classifiers. Using the developed co-registration methodology to these ends, this work highlights the utility of this methodology and the potential of microUS and related deep learning applications for identification of prostate cancer.

## 7.1.1. Impact on Prostate Cancer Diagnosis

MicroUS and related application of deep learning models have the potential to greatly improve prostate cancer diagnosis. From this thesis we have demonstrated that microUS continues to perform similarly to gold standard MRI even under rigorous interrogation against WM pathology. We have also demonstrated deep learning classifiers are able to identify microUS images containing prostate cancer with similar accuracy to expert readers. Ultimately this body of work provides objective evidence that microUS is able to identify prostate cancer with similar performance to mpMRI. For smaller centers, microUS could be serve as an alternative to MRI. For larger centers, microUS and mpMRI could potentially have a synergistic relationship, where mpMRI is used as a triage for biopsy and target identification, while microUS is used to directly visualize the MRI target or any targets not visible on MRI, alleviating registration error and potentially providing better tumor margins. Beyond that, deep learning approaches have the potential to alleviate the inherent training requirement for microUS and could reduce inter-reader variability between expert readers allowing for standardized microUS interpretation.

## 7.2. Future work

### 7.2.1. Automated Registration of Micro-Ultrasound and Histology

In chapter 5 we developed a novel registration methodology to co-register microUS, MRI, and WM pathology. This methodology was subsequently validated through TRE studies and used to compare visualization of prostate cancer on microUS and MRI imaging. The methodology was also used to create confidently labeled microUS images for model training. While accurate, a limitation of the registration methodology is the lack of complete automation. In order to evaluate larger microUS datasets and create more microUS training data, it is necessary to develop a fully automated microUS registration pipeline. As such, future work should focus on automation of the registration process allowing for faster data processing and labeling.

To achieve this, a major hurdle will involve deriving transformation matrices for alignment and TPS registration of microUS and MRI automatically. In the current registration pipeline the transformation matrices are created manually, but in theory tissue deformation from scanning on microUS should be similar across patients allowing for similar transformation matrices with only small deviations. Based on this, and the ability to generate actual correct matrices manually, one possible automation approach may be to develop a machine learning based approach to automate registration. The proposed model would utilize 3D image volumes of MRI and microUS to fine tune the transformation matrices. The accuracy of these predictions could be subsequently compared to matrices generated manually as well as by measuring the TRE between microUS and MRI similar to the analysis performed in chapter 5. If achieved, this would improve upon the registration methodology presented in this work.

## 7.2.2. Automated Segmentation of Prostate Cancer

The developed deep learning classifiers detailed in chapter 6 demonstrated expert level performance for identification of prostate cancer on microUS imaging. The developed classifiers also had good performance for correctly selecting a single image for targeted biopsy. The promising results from these classifiers encourages further advancement of machine learning based detection of prostate cancer on microUS. Currently, the models classify images as cancerous or benign. An ideal model would also delineate the cancer boundary on images classified as cancerous removing the need for the operator to identify the location of cancer for biopsy.

Before training a model for delineation it will be necessary to create accurate masks of the regions of cancer. While WM pathology is the most comprehensive form of ground truth available, there are still regions of the prostate between WM slides where the location of cancer will need to be interpolated. Since interpolation will be prone to error, it is important to account for this in our labels. Similar to the label softening used for weakly labeled data in chapter 6, we plan to decrease the strength of a mask as it moves away from WM pathology confirmed regions of cancer. This should diminish the impact of improperly labeled data while still allowing for the models to learn from regions where cancer is likely but not confirmed. Following that, with enough data labeling, it should be possible to train a segmentation model to predict cancer on previously classified images. It is likely that the models will require a substantial amount of data however, so the automated registration process described above would help these efforts. Furthermore, implementation of MRI and PSA with microUS could also help improve artificial intelligence detection of cancer.

# 8.    References

[1]     R. L. Siegel, A. N. Giaquinto, and J. D. Ahmedin, "Cancer statistics, 2024," *CA. Cancer J. Clin.*, vol. 74, no. 1, pp. 12–49, Jan. 2024,

[2]     L. Klotz *et al.*, "Long-term follow-up of a large active surveillance cohort of patients with prostate cancer," *J. Clin. Oncol.*, vol. 33, no. 3, pp. 272–277, Jan. 2015,

[3]     J. L. Mohler *et al.*, "Prostate cancer, version 2.2019," *JNCCN J. Natl. Compr. Cancer Netw.*, vol. 17, no. 5, pp. 479–505, May 2019,

[4]     M. E. Noureldin, M. J. Connor, N. Boxall, S. Miah, T. Shah, and J. Walz, "Current techniques of prostate biopsy: An update from past to present," *Translational Andrology and Urology*, vol. 9, no. 3. AME Publishing Company, pp. 1510–1517, Jun. 01, 2020.

[5]     H. Hricak, P. L. Choyke, S. C. Eberhardt, S. A. Leibel, and P. T. Scardino, "Imaging prostate cancer: A multidisciplinary perspective," *Radiology*, vol. 243, no. 1. pp. 28–53, Apr. 01, 2007.

[6]     A. Stabile *et al.*, "Multiparametric MRI for prostate cancer diagnosis: current status and future directions," *Nat. Rev. Urol. 2019 171*, vol. 17, no. 1, pp. 41–61, Jul. 2019,

[7]     H. U. Ahmed *et al.*, "Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study," *Lancet*, vol. 389, no. 10071, pp. 815–822, Feb. 2017,

[8]     S. Woo, C. H. Suh, S. Y. Kim, J. Y. Cho, and S. H. Kim, "Diagnostic Performance of Prostate Imaging Reporting and Data System Version 2 for Detection of Prostate Cancer: A Systematic Review and Diagnostic Meta-analysis," *Eur. Urol.*, vol. 72, no. 2, pp. 177–188, Aug. 2017,

[9]     H. U. Ahmed, "The Index Lesion and the Origin of Prostate Cancer," *N. Engl. J. Med.*, vol. 361, no. 17, pp. 1704–1706, Oct. 2009,

[10]    M. A. Bjurlin *et al.*, "Update of the Standard Operating Procedure on the Use of Multiparametric Magnetic Resonance Imaging for the Diagnosis, Staging and Management of Prostate Cancer," *J. Urol.*, vol. 203, no. 4, pp. 706–712, Apr. 2020,

[11]    C. Arsov *et al.*, "Prospective Randomized Trial Comparing Magnetic Resonance Imaging (MRI)-guided In-bore Biopsy to MRI-ultrasound Fusion and Transrectal Ultrasound-guided Prostate Biopsy in Patients with Prior Negative Biopsies," *Eur. Urol.*, vol. 68, no. 4, pp. 713–720, Oct. 2015,

[12]    L. Marks, S. Young, and S. Natarajan, "MRI-ultrasound fusion for guidance of targeted prostate biopsy.," *Curr. Opin. Urol.*, vol. 23, no. 1, pp. 43–50, Jan. 2013,

[13]    A. Priester *et al.*, "Magnetic Resonance Imaging Underestimation of Prostate Cancer Geometry: Use of Patient Specific Molds to Correlate Images with Whole Mount Pathology," *J. Urol.*, vol. 197, no. 2, pp. 320–326, Feb. 2017,

[14] D. C. Johnson *et al.*, "Detection of Individual Prostate Cancer Foci via Multiparametric Magnetic Resonance Imaging," *Eur. Urol.*, vol. 75, no. 5, pp. 712–720, May 2019,

[15] A. B. Rosenkrantz *et al.*, "Interobserver reproducibility of the PI-RADS version 2 lexicon: A multicenter study of six experienced prostate radiologists," *Radiology*, vol. 280, no. 3, pp. 793–804, Sep. 2016,

[16] L. Klotz *et al.*, "Comparison of micro-ultrasound and multiparametric magnetic resonance imaging for prostate cancer: A multicenter, prospective analysis," *Can. Urol. Assoc. J.*, vol. 15, no. 1, p. E11, Jul. 2020,

[17] S. Ghai *et al.*, "Comparison of Micro-US and Multiparametric MRI for Prostate Cancer Detection in Biopsy-Naive Men," *https://doi.org/10.1148/radiol.212163*, Jul. 2022,

[18] S. Ghai *et al.*, "Assessing Cancer Risk on Novel 29 MHz Micro-Ultrasound Images of the Prostate: Creation of the Micro-Ultrasound Protocol for Prostate Risk Identification," *J. Urol.*, vol. 196, no. 2, pp. 562–569, Aug. 2016,

[19] S. R. Zhou *et al.*, "Inter-reader Agreement for Prostate Cancer Detection Using Micro-ultrasound: A Multi-institutional Study," *Eur. Urol. Open Sci.*, vol. 66, pp. 93–100, Aug. 2024,

[20] M. L. Palmeri *et al.*, "Identifying Clinically Significant Prostate Cancers using 3-D In Vivo Acoustic Radiation Force Impulse Imaging with Whole-Mount Histology Validation," *Ultrasound Med. Biol.*, vol. 42, no. 6, pp. 1251–1262, Jun. 2016,

[21] J. Pensa, R. Geoghegan, and S. Natarajan, "3D Ultrasound for Biopsy of the Prostate," in *3D Ultrasound*, Boca Raton: CRC Press, 2023, pp. 154–175.

[22] W. Alyami, A. Kyme, and R. Bourne, "Histological Validation of MRI: A Review of Challenges in Registration of Imaging and Whole-Mount Histopathology," *J. Magn. Reson. Imaging*, vol. 55, no. 1, pp. 11–22, Jan. 2022,

[23] C. H. Lee, O. Akin-Olugbade, and A. Kirschenbaum, "Overview of Prostate Anatomy, Histology, and Pathology," *Endocrinology and Metabolism Clinics of North America*, vol. 40, no. 3. Elsevier, pp. 565–575, Sep. 01, 2011.

[24] J. E. McNeal, "The zonal anatomy of the prostate," *Prostate*, vol. 2, no. 1, pp. 35–49, Jan. 1981,

[25] R. S. Rittmaster, R. W. Norman, L. N. Thomas, and G. Rowden, "Evidence for atrophy and apoptosis in the prostates of men given finasteride.," *J. Clin. Endocrinol. Metab.*, vol. 81, no. 2, pp. 814–819, Feb. 1996,

[26] D. F. Gleason, G. T. Mellinger, and The Veterans Administration Cooperative Urological Research Group, "Prediction of Prognosis for Prostatic Adenocarcinoma by Combined Histological Grading and Clinical Staging," *J. Urol.*, vol. 111, no. 1, pp. 58–64, 1974,

[27] K. K. Yu and H. Hricak, "Imaging prostate cancer," *Radiol. Clin. North Am.*, vol. 38, no. 1, pp. 59–85, Jan. 2000,

[28] M. R. Cooperberg and P. R. Carroll, "Trends in management for patients with localized prostate cancer, 1990-2013," *JAMA - J. Am. Med. Assoc.*, vol. 314, no. 1, pp. 80–82, Jul. 2015,

[29] P. A. Humphrey, "Histopathology of prostate cancer," *Cold Spring Harb. Perspect. Med.*, vol. 7, no. 10, Oct. 2017,

[30] J. E. McNeal and C. E. M. Yemoto, "Spread of adenocarcinoma within prostatic ducts and acini: Morphologic and clinical correlations," *Am. J. Surg. Pathol.*, vol. 20, no. 7, pp. 802–814, Jul. 1996,

[31] Z. Kaleem, P. E. Swanson, R. T. Vollmer, and P. A. Humphrey, "Prostatic adenocarcinoma with atrophic features: A study of 202 consecutive completely embedded radical prostatectomy specimens," *Am. J. Clin. Pathol.*, vol. 109, no. 6, pp. 695–703, 1998,

[32] B. D. Robinson and J. I. Epstein, "Intraductal carcinoma of the prostate without invasive carcinoma on needle biopsy: Emphasis on radical prostatectomy findings," *J. Urol.*, vol. 184, no. 4, pp. 1328–1333, Oct. 2010,

[33] J. I. Epstein, L. Egevad, M. B. Amin, B. Delahunt, J. R. Srigley, and P. A. Humphrey, "The 2014 international society of urological pathology (ISUP) consensus conference on gleason grading of prostatic carcinoma definition of grading patterns and proposal for a new grading system," *Am. J. Surg. Pathol.*, vol. 40, no. 2, pp. 244–252, 2016,

[34] L. Wang, Y. Mazaheri, J. Zhang, N. M. Ishill, K. Kuroiwa, and H. Hricak, "Assessment of biologic aggressiveness of prostate cancer: Correlation of MR signal intensity with gleason grade after radical prostatectomy," *Radiology*, vol. 246, no. 1, pp. 168–176, Jan. 2008,

[35] C. E. Lovegrove *et al.*, "Prostate imaging features that indicate benign or Malignant pathology on biopsy," *Translational Andrology and Urology*, vol. 7, no. Suppl 4. AME Publishing Company, pp. S420–S435, Sep. 01, 2018.

[36] G. J. Miller and J. M. Cygan, "Morphology of prostate cancer: The effects of multifocality on histological grade, tumor volume and capsule penetration," in *Journal of Urology*, Elsevier Inc., Nov. 1994, pp. 1709–1713.

[37] E. T. Ruijter, C. A. Van De Kaat, J. A. Schalken",', F. M. Debruyne, and D. J. Rulter, "HISTOLOGICAL GRADE HETEROGENEITY IN MULTIFOCAL PROSTATE CANCER. BIOLOGICAL AND CLINICAL IMPLICATIONS," 1996.

[38] M. Aihara, T. M. Wheeler, M. Ohori, and P. T. Scardino, "Heterogeneity of prostate cancer inradical prostatectomy specimens," *Urology*, vol. 43, no. 1, pp. 60–66, Jan. 1994,

[39] C. A. Macintosh, M. Stower, N. Reid, and N. J. Maitland, "Precise Microdissection of Human Prostate Cancers Reveals Genotypic Heterogeneity," *Cancer Res.*, vol. 58, no. 1, 1998.

[40] "Allelic imbalance in the clonal evolution of prostate carcinoma - Cheng - 1999 - Cancer - Wiley Online Library." Accessed: May 25, 2021. [Online]. Available: https://acsjournals.onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291097-

0142%2819990501%2985%3A9%3C2017%3A%3AAID-CNCR20%3E3.0.CO%3B2-V

[41]    L. Cheng *et al.*, "Evidence of independent origin of multiple tumors from patients with prostate cancer," *J. Natl. Cancer Inst.*, vol. 90, no. 3, pp. 233–237, Feb. 1998,

[42]    W. Liu *et al.*, "Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer," *Nat. Med.*, vol. 15, no. 5, pp. 559–565, May 2009,

[43]    A. Stabile, M. Moschini, F. Montorsi, X. Cathelineau, and R. Sanchez-Salas, "Focal therapy for prostate cancer-index lesion treatment vs. hemiablation. A matter of definition," *International Braz J Urol*, vol. 45, no. 5. Brazilian Society of Urology, pp. 873–876, 2019.

[44]    H. U. Ahmed, "The Index Lesion and the Origin of Prostate Cancer," *N. Engl. J. Med.*, vol. 361, no. 17, pp. 1704–1706, Oct. 2009,

[45]    T. A. Stamey, J. M. McNeal, A. M. Wise, and J. L. Clayton, "Secondary cancers in the prostate do not determine PSA biochemical failure in untreated men undergoing radical retropubic prostatectomy," in *European Urology*, S. Karger AG, 2001, pp. 22–23.

[46]    R. M. Hoffman, "Screening for Prostate Cancer," *N. Engl. J. Med.*, vol. 365, no. 21, pp. 2013–2019, Nov. 2011,

[47]    R. M. Hoffman, S. N. Stone, D. Espey, and A. L. Potosky, "Differences between men with screening-detected versus clinically diagnosed prostate cancers in the USA," *BMC Cancer*, vol. 5, no. 1, pp. 1–9, Mar. 2005,

[48]    H. Seidman, M. H. Mushinski, S. K. Gelb, and E. Silverberg, "Probabilities of Eventually Developing or Dying of Cancer--United States, 1985," *CA. Cancer J. Clin.*, vol. 35, no. 1, pp. 36–56, Jan. 1985,

[49]    ALTEKRUSE and S. editors, "SEER cancer statistics review, 1975-2007," *http://seer.cancer.gov/csr/1975_2007/results_merged/sect_13_leukemia.pdf*, 2009, Accessed: Jun. 01, 2021. [Online]. Available: https://ci.nii.ac.jp/naid/10030551832

[50]    D. S. Smith and W. J. Catalona, "Interexaminer variability of digital rectal examination in detecting prostate cancer," *Urology*, vol. 45, no. 1, pp. 70–74, Jan. 1995,

[51]    G. W. Chodak, P. Keller, and H. W. Schoenberg, "Assessment of screening for prostate cancer using the digital rectal examination," *J. Urol.*, vol. 141, no. 5, pp. 1136–1138, May 1989,

[52]    "Prostate-Specific Antigen (PSA) Test - National Cancer Institute." Accessed: Jun. 01, 2021. [Online]. Available: https://www.cancer.gov/types/prostate/psa-fact-sheet

[53]    G. Draisma *et al.*, "Lead time and overdiagnosis in prostate-specific antigen screening: Importance of methods and context," *J. Natl. Cancer Inst.*, vol. 101, no. 6, pp. 374–383, Mar. 2009,

[54]    I. Yusim, M. Krenawi, E. Mazor, V. Novack, and N. J. Mabjeesh, "The use of prostate specific antigen density to predict clinically significant prostate cancer," *Sci. Rep.*, vol. 10, no. 1, pp. 1–6, Dec. 2020,

[55]    B. R. Matlaga, L. A. Eskew, and D. L. McCullough, "Prostate biopsy: Indications and technique," *Journal of Urology*, vol. 169, no. 1. Lippincott Williams and Wilkins, pp. 12–19, Jan. 01, 2003.

[56]    H. Watanabe, D. Igari, Y. Tanahasi, K. Harada, and M. Saitoh, "Development and application of new equipment for transrectal ultrasonography," *J. Clin. Ultrasound*, vol. 2, no. 2, pp. 91–98, Jun. 1974,

[57]    S. W. Hardeman, R. W. Wake, and M. S. Soloway, "Two new techniques for evaluating prostate cancer The role of prostate-specific antigen and transrectal ultrasound," *Postgrad. Med.* , 1989,

[58]    R. Clements, G. J. Griffiths, W. B. Peeling, and P. G. Ryan, "Experience with Ultrasound Guided Transperineal Prostatic Needle Biopsy 1985–1988," *Br. J. Urol.*, vol. 65, no. 4, pp. 362–367, Apr. 1990,

[59]    K. Shinohara, T. M. Wheeler, and P. T. Scardino, "The appearance of prostate cancer on transrectal ultrasonography: Correlation of imagining and pathological examinations," *J. Urol.*, vol. 142, no. 1, pp. 76–82, 1989,

[60]    F. A. Distler *et al.*, "The Value of PSA Density in Combination with PI-RADS$^{TM}$ for the Accuracy of Prostate Cancer Prediction," *J. Urol.*, vol. 198, no. 3, pp. 575–582, Sep. 2017,

[61]    A. Verma, J. St. Onge, K. Dhillon, and A. Chorneyko, "PSA density improves prediction of prostate cancer.," *Can. J. Urol.*, vol. 21, no. 3, pp. 7312–7321, Jun. 2014,

[62]    D. R. H. Christie and C. F. Sharpley, "How Accurately Can Prostate Gland Imaging Measure the Prostate Gland Volume? Results of a Systematic Review," *Prostate Cancer*, vol. 2019, 2019,

[63]    N. R. Paterson *et al.*, "Prostate volume estimations using magnetic resonance imaging and transrectal ultrasound compared to radical prostatectomy specimens," *Can. Urol. Assoc. J.*, vol. 10, no. 7–8, p. 264, Aug. 2016,

[64]    V. Kasivisvanathan *et al.*, "MRI-Targeted or Standard Biopsy for Prostate-Cancer Diagnosis," *N. Engl. J. Med.*, Mar. 2018,

[65]    C. J. Das, A. Razik, A. Netaji, and S. Verma, "Prostate MRI–TRUS fusion biopsy: a review of the state of the art procedure," *Abdom. Radiol.*, vol. 45, no. 7, pp. 2176–2183, Jul. 2020,

[66]    S. Natarajan *et al.*, "Clinical application of a 3D ultrasound-guided prostate biopsy system," *Urologic Oncology: Seminars and Original Investigations*, vol. 29, no. 3. Urol Oncol, pp. 334–342, May 2011.

[67]    H. Hricak, P. L. Choyke, S. C. Eberhardt, S. A. Leibel, and P. T. Scardino, "Imaging prostate cancer: A multidisciplinary perspective," *Radiology*, vol. 243, no. 1.  Radiological Society of North America, pp. 28–53, Apr. 01, 2007.

[68]    F. A. Ganie *et al.*, "Correlation of transrectal ultrasonographic findings with histo pathology in prostatic cancer.," *J. Educ. Health Promot.*, vol. 3, no. 1, pp. 38–38, 2014,

[69]    K. K. Shung, "High Frequency Ultrasonic Imaging," *Journal of Medical Ultrasound*, vol. 17, no. 1.

Elsevier (Singapore) Pte Ltd, pp. 25–30, 2009.

[70]    "Ultrasound transducer | Radiology Reference Article | Radiopaedia.org." Accessed: Jun. 02, 2021. [Online]. Available: https://radiopaedia.org/articles/ultrasound-transducer?lang=us

[71]    Y. J. Lu, Z. F. Shi, C. X. Shan, and D. Z. Shen, "ZnO nanostructures and lasers," in *Nanoscale Semiconductor Lasers*, Elsevier, 2019, pp. 75–108.

[72]    "The Piezoelectric Effect - Piezoelectric Motors & Motion Systems." Accessed: Jun. 02, 2021. [Online]. Available: https://www.nanomotion.com/nanomotion-technology/piezoelectric-effect/

[73]    N. Biosensors and D. Li, "Non-Centrosymmetric," in *Encyclopedia of Microfluidics and Nanofluidics*, Springer US, 2008, pp. 1457–1457.

[74]    N. J. Hangiandreou, "B-mode US: Basic Concepts and New Technology," *Radiographics*, vol. 23, no. 4. Radiological Society of North America Inc., pp. 1019–1033, Jul. 01, 2003.

[75]    F. M. Abu-Zidan, A. F. Hefny, and P. Corr, "Clinical ultrasound physics," *J. Emergencies, Trauma Shock*, vol. 4, no. 4, pp. 501–503, Oct. 2011,

[76]    M. S. Canney, M. R. Bailey, L. A. Crum, V. A. Khokhlova, and O. A. Sapozhnikov, "Acoustic characterization of high intensity focused ultrasound fields: A combined measurement and modeling approach," *J. Acoust. Soc. Am.*, vol. 124, no. 4, pp. 2406–2420, Oct. 2008,

[77]    A. van Hove *et al.*, "Comparison of image-guided targeted biopsies versus systematic randomized biopsies in the detection of prostate cancer: A systematic literature review of well-designed studies," *World J. Urol.*, vol. 32, no. 4, pp. 847–858, Jun. 2014,

[78]    A. B. Dias, C. O'Brien,  jean M. Correas, and S. Ghai, "Multiparametric ultrasound and micro-ultrasound in prostate cancer: A comprehensive review," *Br. J. Radiol.*, vol. 95, no. 1131, Nov. 2022,

[79]    L. Pallwein *et al.*, "Prostate cancer diagnosis: Value of real-time elastography," *Abdom. Imaging*, vol. 33, no. 6, pp. 729–735, Nov. 2008,

[80]    A. Sadeghi-Naini *et al.*, "Quantitative Ultrasound Spectroscopic Imaging for Characterization of Disease Extent in Prostate Cancer Patients," *Transl. Oncol.*, vol. 8, no. 1, pp. 25–34, Feb. 2015,

[81]    F. Sano and H. Uemura, "The Utility and Limitations of Contrast-Enhanced Ultrasound for the Diagnosis and Treatment of Prostate Cancer," *Sensors 2015, Vol. 15, Pages 4947-4957*, vol. 15, no. 3, pp. 4947–4957, Feb. 2015,

[82]    D. Rohrbach, B. Wodlinger, J. Wen, J. Mamou, and E. Feleppa, "High-Frequency Quantitative Ultrasound for Imaging Prostate Cancer Using a Novel Micro-Ultrasound Scanner," *Ultrasound Med. Biol.*, vol. 44, no. 7, pp. 1341–1354, Jul. 2018,

[83]    J. R. Eisenbrey, A. Sridharan, J. Bin Liu, and F. Forsberg, "Recent Experiences and Advances in Contrast-Enhanced Subharmonic Ultrasound," *Biomed Res. Int.*, vol. 2015, 2015,

[84] "MRI Basics." Accessed: Aug. 02, 2021. [Online]. Available: https://case.edu/med/neurology/NR/MRI Basics.htm

[85] A. B. Rosenkrantz and S. S. Taneja, "Radiologist, Be Aware: Ten Pitfalls That Confound the Interpretation of Multiparametric Prostate MRI," *http://dx.doi.org/10.2214/AJR.13.10699*, vol. 202, no. 1, pp. 109–120, Dec. 2013,

[86] W. JC *et al.*, "PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2," *Eur. Urol.*, vol. 69, no. 1, pp. 16–40, Jan. 2016,

[87] B. JO *et al.*, "ESUR prostate MR guidelines 2012," *Eur. Radiol.*, vol. 22, no. 4, pp. 746–757, 2012,

[88] H. Hricak *et al.*, "MR imaging of the prostate gland: normal anatomy," *http://dx.doi.org/10.2214/ajr.148.1.51*, vol. 148, no. 1, pp. 51–58, Nov. 2012,

[89] H. T *et al.*, "Relationship between apparent diffusion coefficients at 3.0-T MR imaging and Gleason grade in peripheral zone prostate cancer," *Radiology*, vol. 259, no. 2, pp. 453–461, May 2011,

[90] J. SI, D. OF, V. HA, G. D, H. H, and A. O, "Transition zone prostate cancer: incremental value of diffusion-weighted endorectal MR imaging in tumor detection and assessment of aggressiveness," *Radiology*, vol. 269, no. 2, pp. 493–503, Nov. 2013,

[91] L.-M. Wu, J.-R. Xu, Y.-Q. Ye, Q. Lu, and J.-N. Hu, "The Clinical Value of Diffusion-Weighted Imaging in Combination With T2-Weighted Imaging in Diagnosing Prostate Carcinoma: A Systematic Review and Meta-Analysis," *http://dx.doi.org/10.2214/AJR.11.7634*, vol. 199, no. 1, pp. 103–110, Nov. 2012,

[92] H. N, O. M, K. H, K. M, and B. V, "Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) is a useful modality for the precise detection and staging of early prostate cancer," *Prostate*, vol. 62, no. 2, pp. 140–147, Feb. 2005,

[93] P. S *et al.*, "Prostatic cancer surveillance following whole-gland high-intensity focused ultrasound: comparison of MRI and prostate-specific antigen for detection of residual or recurrent disease," *Br. J. Radiol.*, vol. 85, no. 1014, pp. 720–728, Jun. 2012,

[94] D. V. R *et al.*, "Dynamic contrast-enhanced MR evaluation of prostate cancer before and after endorectal high-intensity focused ultrasound," *Radiol. Med.*, vol. 118, no. 5, pp. 851–862, Aug. 2013,

[95] H. Hricak, "MR imaging and MR spectroscopic imaging in the pre-treatment evaluation of prostate cancer," *https://doi.org/10.1259/bjr/11253478*, vol. 78, no. SPEC. ISS. 2, Jan. 2014,

[96] K. V, J. NR, T. S, and K. R, "Prebiopsy magnetic resonance spectroscopy and imaging in the diagnosis of prostate cancer," *Int. J. Urol.*, vol. 19, no. 7, pp. 602–613, Jul. 2012,

[97] A. Stabile *et al.*, "Multiparametric MRI for prostate cancer diagnosis: current status and future directions," *Nat. Rev. Urol. 2019 171*, vol. 17, no. 1, pp. 41–61, Jul. 2019,

[98]     F. Giganti *et al.*, "The Evolution of MRI of the Prostate: The Past, the Present, and the Future," *https://doi.org/10.2214/AJR.18.20796*, vol. 213, no. 2, pp. 384–396, Apr. 2019,

[99]     A. Farolfi *et al.*, "Current and Emerging Clinical Applications of PSMA PET Diagnostic Imaging for Prostate Cancer," *J. Nucl. Med.*, vol. 62, no. 5, pp. 596–604, May 2021,

[100]    A. Shukla and U. Kumar, "Positron emission tomography: An overview," *J. Med. Phys.*, vol. 31, no. 1, p. 13, 2006,

[101]    Z. H. Cho *et al.*, "Positron Ranges Obtained from Biomedically Important Positron-Emitting Radionuclides," *J. Nucl. Med.*, vol. 16, no. 12, pp. 1174–1176, Dec. 1975, Accessed: Aug. 09, 2024. [Online]. Available: https://jnm.snmjournals.org/content/16/12/1174

[102]    I. Sonni *et al.*, "Head-to-Head Comparison of 68Ga-PSMA-11 PET/CT and mpMRI with a Histopathology Gold Standard in the Detection, Intraprostatic Localization, and Determination of Local Extension of Primary Prostate Cancer: Results from a Prospective Single-Center Imaging Trial," *J. Nucl. Med.*, vol. 63, no. 6, p. 847, Jun. 2022,

[103]    D. Sitharthan *et al.*, "The Sensitivity and Specificity of Multiparametric Magnetic Resonance Imaging and Prostate-Specific Membrane Antigen Positron Emission Tomography/Computed Tomography for Predicting Seminal Vesicle Invasion in Clinically Significant Prostate Cancer: A Multicenter Retrospective Study," *J. Clin. Med. 2024, Vol. 13, Page 4424*, vol. 13, no. 15, p. 4424, Jul. 2024,

[104]    W. Venderink, M. de Rooij, J. P. M. Sedelaar, H. J. Huisman, and J. J. Fütterer, "Elastic Versus Rigid Image Registration in Magnetic Resonance Imaging–transrectal Ultrasound Fusion Prostate Biopsy: A Systematic Review and Meta-analysis," *Eur. Urol. Focus*, vol. 4, no. 2, pp. 219–227, Mar. 2018,

[105]    D. H. Turnbull and F. S. Foster, "Fabrication and Characterization of Transducer Elements in Two-Dimensional Arrays for Medical Ultrasound Imaging," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 39, no. 4, pp. 464–475, 1992,

[106]    J. K. Logan *et al.*, "Current status of magnetic resonance imaging (MRI) and ultrasonography fusion software platforms for guidance of prostate biopsies," *BJU Int.*, vol. 114, no. 5, pp. 641–652, Nov. 2014,

[107]    M. Kongnyuy, A. K. George, A. R. Rastinehad, and P. A. Pinto, "Magnetic Resonance Imaging-Ultrasound Fusion-Guided Prostate Biopsy: Review of Technology, Techniques, and Outcomes," *Curr. Urol. Rep.*, vol. 17, no. 4, pp. 1–9, Apr. 2016,

[108]    Q. Huang and Z. Zeng, "A Review on Real-Time 3D Ultrasound Imaging Technology," *Biomed Res. Int.*, vol. 2017, 2017,

[109]    L. Klotz *et al.*, "Comparison of micro-ultrasound and multiparametric magnetic resonance imaging for prostate cancer: A multicenter, prospective analysis," *Can. Urol. Assoc. J.*, vol. 15, no. 1, p. E11, Jul. 2021,

[110] S. Natarajan *et al.*, "Clinical application of a 3D ultrasound-guided prostate biopsy system," *Urol. Oncol. Semin. Orig. Investig.*, vol. 29, no. 3, pp. 334–342, May 2011,

[111] P. Fletcher *et al.*, "Vector Prostate Biopsy: A Novel Magnetic Resonance Imaging/Ultrasound Image Fusion Transperineal Biopsy Technique Using Electromagnetic Needle Tracking Under Local Anaesthesia," *Eur. Urol.*, no. xxxx, 2023,

[112] "3D Ultrasound | KOELIS," *https://koelis.com/*, Accessed: Nov. 16, 2022. [Online]. Available: https://koelis.com/3d-ultrasound/

[113] F. Cornud *et al.*, "Precision matters in MR imaging-targeted Prostate biopsies: Evidence from a prospective study of cognitive and elastic fusion registration transrectal biopsies," *Radiology*, vol. 287, no. 2, pp. 534–542, May 2018,

[114] A. Shah *et al.*, "An Open Source Multimodal Image-Guided Prostate Biopsy Framework," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8680, pp. 1–8, 2014,

[115] H. Guo, H. Chao, S. Xu, B. J. Wood, J. Wang, and P. Yan, "Ultrasound Volume Reconstruction From Freehand Scans Without Tracking," *IEEE Trans. Biomed. Eng.*, 2022,

[116] L. Tetrel, H. Chebrek, and C. Laporte, "Learning for graph-based sensorless freehand 3D ultrasound," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10019 LNCS, pp. 205–212, 2016,

[117] T. A. Tuthill, J. F. Krücker, J. B. Fowlkes, and P. L. Carson, "Automated three-dimensional US frame positioning computed from elevational speckle decorrelation," *Radiology*, vol. 209, no. 2, pp. 575–582, 1998,

[118] M. H. Mozaffari and W. S. Lee, "Freehand 3-D Ultrasound Imaging: A Systematic Review," *Ultrasound Med. Biol.*, vol. 43, no. 10, pp. 2099–2124, Oct. 2017,

[119] R. San José-Estépar, M. Martín-Fernández, P. P. Caballero-Martínez, C. Alberola-López, and J. Ruiz-Alzola, "A theoretical framework to three-dimensional ultrasound reconstruction from irregularly sampled data," *Ultrasound Med. Biol.*, vol. 29, no. 2, pp. 255–269, Feb. 2003,

[120] T. Roxborough and G. M. Nielson, "Tetrahedron based, least squares, progressive volume models with application to freehand ultrasound data," *Proc. IEEE Vis. Conf.*, pp. 93–100, 2000,

[121] S. Sherebrin, A. Fenster, R. N. R. M.D., and D. Spence, "Freehand three-dimensional ultrasound: implementation and applications," *Med. Imaging*, vol. 2708, pp. 296–303, Apr. 1996,

[122] J. W. Trobaugh, D. J. Trobaugh, and W. D. Richard, "Three-dimensional imaging with stereotactic ultrasonography," *Comput. Med. Imaging Graph.*, vol. 18, no. 5, pp. 315–323, Sep. 1994,

[123] C. D. Barry *et al.*, "Three-dimensional freehand ultrasound: Image reconstruction and volume analysis," *Ultrasound Med. Biol.*, vol. 23, no. 8, pp. 1209–1224, Jan. 1997,

[124] Q. Huang, Y. Huang, W. Hu, and X. Li, "Bezier interpolation for 3-D freehand ultrasound," *IEEE Trans. Human-Machine Syst.*, vol. 45, no. 3, pp. 385–392, Jun. 2015,

[125] J. Jiang, Y. Guo, Z. Bi, Z. Huang, G. Yu, and J. Wang, "Segmentation of prostate ultrasound images: the state of the art and the future directions of segmentation algorithms," *Artif. Intell. Rev. 2022*, pp. 1–37, Apr. 2022,

[126] P. Wu, Y. Liu, Y. Li, and Y. Shi, "TRUS image segmentation with non-parametric kernel density estimation shape prior," *Biomed. Signal Process. Control*, vol. 8, no. 6, pp. 764–771, Nov. 2013,

[127] Y. J. Liu, W. S. Ng, M. Y. Teo, and H. C. Lim, "Computerised prostate boundary estimation of ultrasound images using radial bas-relief method," *Med. Biol. Eng. Comput. 1997 355*, vol. 35, no. 5, pp. 445–454, 1997,

[128] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002,

[129] R. Manavalan and K. Thangavel, "TRUS image segmentation using morphological operators and DBSCAN clustering," *Proc. 2011 World Congr. Inf. Commun. Technol. WICT 2011*, pp. 898–903, 2011,

[130] A. Sarti, C. Corsi, E. Mazzini, and C. Lamberti, "Maximum likelihood segmentation with rayleigh distribution of ultrasound images," *Comput. Cardiol.*, vol. 31, pp. 329–332, 2004,

[131] F. Shao, K. V. Ling, W. S. Ng, and R. Y. Wu, "Prostate Boundary Detection From Ultrasonographic Images," *J. Ultrasound Med.*, vol. 22, no. 6, pp. 605–623, Jun. 2003,

[132] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis. 1987 14*, vol. 1, no. 4, pp. 321–331, Jan. 1988,

[133] S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations," *J. Comput. Phys.*, vol. 79, no. 1, pp. 12–49, Nov. 1988,

[134] S. Ghose *et al.*, "A survey of prostate segmentation methodologies in ultrasound, magnetic resonance and computed tomography images," *Comput. Methods Programs Biomed.*, vol. 108, no. 1, pp. 262–287, Oct. 2012,

[135] Y. Zhan and D. Shen, "Deformable segmentation of 3-D ultrasound prostate images using statistical texture matching method," *IEEE Trans. Med. Imaging*, vol. 25, no. 3, pp. 256–272, Mar. 2006,

[136] X. Yang and B. Fei, "3D prostate segmentation of ultrasound images combining longitudinal image registration and machine learning," *Med. Imaging*, vol. 8316, pp. 803–811, Feb. 2012,

[137] W. Weng and X. Zhu, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *IEEE Access*, vol. 9, pp. 16591–16603, May 2015,

[138] R. J. G. Van Sloun *et al.*, "Zonal Segmentation in Transrectal Ultrasound Images of the Prostate

Through Deep Learning," *IEEE Int. Ultrason. Symp. IUS*, vol. 2018-October, Dec. 2018,

[139]  N. Orlando, D. J. Gillies, I. Gyacskov, C. Romagnoli, D. D'Souza, and A. Fenster, "Automatic prostate segmentation using deep learning on clinically diverse 3D transrectal ultrasound images," *Med. Phys.*, vol. 47, no. 6, pp. 2413–2426, Jun. 2020,

[140]  N. Orlando *et al.*, "Effect of dataset size, image quality, and image type on deep learning-based automatic prostate segmentation in 3D ultrasound," *Phys. Med. Biol.*, vol. 67, no. 7, p. 074002, Mar. 2022,

[141]  T. R. Nelson and T. T. Elvins, "Visualization of 3D Ultrasound Data," *IEEE Comput. Graph. Appl.*, vol. 13, no. 6, pp. 50–57, 1993,

[142]  J. N. Welch, J. A. Johnson, M. R. Bax, R. Badr, and R. Shahidi, "A real-time freehand 3D ultrasound system for image-guided surgery," *Proc. IEEE Ultrason. Symp.*, vol. 2, pp. 1601–1604, 2000,

[143]  A. Fenster, G. Parraga, and J. Bax, "Three-dimensional ultrasound scanning," *Interface Focus*, vol. 1, no. 4, p. 503, Aug. 2011,

[144]  C. Williams *et al.*, "Why Does Magnetic Resonance Imaging-Targeted Biopsy Miss Clinically Significant Cancer?," *J. Urol.*, vol. 207, no. 1, pp. 95–107, Jan. 2022,

[145]  W. J. M. Van De Ven, G. J. S. Litjens, J. O. Barentsz, T. Hambrock, and H. J. Huisman, "Required accuracy of MR-US registration for prostate biopsies," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 6963 LNCS, pp. 92–99, 2011,

[146]  Y. Guo *et al.*, "Image Registration Accuracy of a 3-Dimensional Transrectal Ultrasound–Guided Prostate Biopsy System," *J. Ultrasound Med.*, vol. 28, no. 11, pp. 1561–1568, Nov. 2009,

[147]  S. Il Hwang *et al.*, "Comparison of Accuracies between Real-Time Nonrigid and Rigid Registration in the MRI–US Fusion Biopsy of the Prostate," *Diagnostics 2021, Vol. 11, Page 1481*, vol. 11, no. 8, p. 1481, Aug. 2021,

[148]  O. Ukimura *et al.*, "3-Dimensional Elastic Registration System of Prostate Biopsy Location By Real-Time 3-Dimensional Transrectal Ultrasound Guidance With Magnetic Resonance/Transrectal Ultrasound Image Fusion," *J. Urol.*, vol. 187, no. 3, pp. 1080–1086, 2012,

[149]  S. Xu *et al.*, "Real-time MRI-TRUS fusion for guidance of targeted prostate biopsies," *Comput. Aided Surg.*, vol. 13, no. 5, pp. 255–264, 2008,

[150]  V. V. Karnik *et al.*, "Assessment of image registration accuracy in three-dimensional transrectal ultrasound guided prostate biopsy," *Med. Phys.*, vol. 37, no. 2, pp. 802–813, 2010,

[151]  H. ; Li *et al.*, "Machine Learning in Prostate MRI for Prostate Cancer: Current Status and Future Opportunities," *Diagnostics 2022, Vol. 12, Page 289*, vol. 12, no. 2, p. 289, Jan. 2022,

[152]  D. W. Cool, X. Zhang, C. Romagnoli, J. I. Izawa, W. M. Romano, and A. Fenster, "Evaluation of MRI-TRUS fusion versus cognitive registration accuracy for MRI-targeted, TRUS-guided prostate

biopsy," *AJR. Am. J. Roentgenol.*, vol. 204, no. 1, pp. 83–91, Jan. 2015,

[153]   S. Ourselin, A. Roche, S. Prima, and N. Ayache, "Block matching: A general framework to improve robustness of rigid registration of medical images," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 1935, pp. 557–566, 2000,

[154]   W. M. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Med. Image Anal.*, vol. 1, no. 1, pp. 35–51, Mar. 1996,

[155]   B. Fei, C. Kemper, and D. L. Wilson, "A comparative study of warping and rigid body registration for the prostate and pelvic MR volumes," *Comput. Med. Imaging Graph.*, vol. 27, no. 4, pp. 267–281, Jul. 2003,

[156]   Y. Wang *et al.*, "Towards personalized statistical deformable model and hybrid point matching for robust MR-TRUS registration," *IEEE Trans. Med. Imaging*, vol. 35, no. 2, pp. 589–604, Feb. 2016,

[157]   Y. Hu *et al.*, "Weakly-supervised convolutional neural networks for multimodal image registration," *Med. Image Anal.*, vol. 49, pp. 1–13, Oct. 2018,

[158]   P. Yan, S. Xu, A. R. Rastinehad, and B. J. Wood, "Adversarial image registration with application for MR and TRUS image fusion," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11046 LNCS, pp. 197–204, 2018,

[159]   Q. Zeng *et al.*, "Label-driven magnetic resonance imaging (MRI)-transrectal ultrasound (TRUS) registration using weakly supervised learning for MRI-guided prostate radiotherapy," *Phys. Med. Biol.*, vol. 65, no. 13, p. 135002, Jun. 2020,

[160]   X. Yang *et al.*, "Deformable MRI-TRUS Registration Using Biomechanically Constrained Deep Learning Model for Tumor-Targeted Prostate Brachytherapy," *Int. J. Radiat. Oncol.*, vol. 108, no. 3, p. e339, Nov. 2020,

[161]   T. Kuru *et al.*, "Phantom Study of a Novel Stereotactic Prostate Biopsy System Integrating Preinterventional Magnetic Resonance Imaging and Live Ultrasonography Fusion," *J. Endourol.*, vol. 26, no. 7, pp. 807–813, 2012,

[162]   B. A. Hadaschik *et al.*, "A novel stereotactic prostate biopsy system integrating pre-interventional magnetic resonance imaging and live ultrasound fusion," *J. Urol.*, vol. 186, no. 6, pp. 2214–2220, 2011,

[163]   O. Wegelin *et al.*, "An Ex Vivo Phantom Validation Study of an MRI-Transrectal Ultrasound Fusion Device for Targeted Prostate Biopsy," *J. Endourol.*, vol. 30, no. 6, pp. 685–691, 2016,

[164]   J. Bax *et al.*, "Mechanically assisted 3D ultrasound guided prostate biopsy system," *Med Phys*, vol. 35, no. 12, pp. 5397–5410, 2008,

[165]   D. W. Rickey, P. A. Picot, D. A. Christopher, and A. Fenster, "A wall-less vessel phantom for Doppler ultrasound studies," *Ultrasound Med. Biol.*, vol. 21, no. 9, pp. 1163–1176, 1995,

[166] J. Pensa, W. Brisbane, A. Priester, A. Sisk, L. Marks, and R. Geoghegan, "A System for Co-Registration of High-Resolution Ultrasound, Magnetic Resonance Imaging, and Whole-Mount Pathology for Prostate Cancer," *2021 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 3890–3893, Nov. 2021,

[167] A. Priester *et al.*, "Registration Accuracy of Patient-Specific, Three-Dimensional-Printed Prostate Molds for Correlating Pathology with Magnetic Resonance Imaging," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 1, pp. 14–22, Jan. 2019,

[168] R. Geoghegan *et al.*, "A tissue-mimicking prostate phantom for 980 nm laser interstitial thermal therapy," *Int. J. Hyperth.*, vol. 36, no. 1, pp. 993–1002, 2019,

[169] R. Geoghegan, L. Zhang, A. M. Priester, H. Wu, L. Marks, and S. Natarajan, "Interstitial Optical Monitoring of Focal Laser Ablation," *IEEE Trans. Biomed. Eng.*, 2022,

[170] A. Eranki, A. S. Mikhail, A. H. Negussie, P. S. Katti, B. J. Wood, and A. Partanen, "Tissue-mimicking thermochromic phantom for characterization of HIFU devices and applications," *Int. J. Hyperth.*, vol. 36, no. 1, pp. 518–529, 2019,

[171] Z. Bu-Lin, H. Bing, K. Sheng-Li, Y. Huang, W. Rong, and L. Jia, "A polyacrylamide gel phantom for radiofrequency ablation." *Int. J. Hyperthermia*, vol. 24, no. 7, pp. 568–576, 2008,

[172] P. Moldovan *et al.*, "Accuracy of elastic fusion of prostate magnetic resonance and transrectal ultrasound images under routine conditions: A prospective multi-operator study," *PLoS One*, vol. 11, no. 12, pp. 1–11, 2016,

[173] V. V. Karnik *et al.*, "Assessment of image registration accuracy in three-dimensional transrectal ultrasound guided prostate biopsy," *Med. Phys.*, vol. 37, no. 2, pp. 802–813, 2010,

[174] M. M. Poggi, D. A. Gant, W. Sewchand, and W. B. Warlick, "Marker seed migration in prostate localization," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 56, no. 5, pp. 1248–1251, 2003,

[175] H. H. Wu *et al.*, "A system using patient-specific 3D-printed molds to spatially align in vivo MRI with ex vivo MRI and whole-mount histopathology for prostate cancer research," *J. Magn. Reson. Imaging*, vol. 49, no. 1, pp. 270–279, Jan. 2019,

[176] Y. Hu *et al.*, "MR to ultrasound registration for image-guided prostate interventions," *Med. Image Anal.*, vol. 16, no. 3, pp. 687–703, Apr. 2012,

[177] P. A. Pinto *et al.*, "Magnetic resonance imaging/ultrasound fusion guided prostate biopsy improves cancer detection following transrectal ultrasound biopsy and correlates with multiparametric magnetic resonance imaging," *J. Urol.*, vol. 186, no. 4, pp. 1281–1285, 2011,

[178] S. Vourganti *et al.*, "Multiparametric magnetic resonance imaging and ultrasound fusion biopsy detect prostate cancer in patients with prior negative transrectal ultrasound biopsies," *J. Urol.*, vol. 188, no. 6, pp. 2152–2157, 2012,

[179] A. R. Rastinehad *et al.*, "Improving detection of clinically significant prostate cancer: Magnetic

resonance imaging/transrectal ultrasound fusion guided prostate biopsy," *J. Urol.*, vol. 191, no. 6, pp. 1749–1754, 2014,

[180] G. A. Sonn *et al.*, "Value of targeted prostate biopsy using magnetic resonance-ultrasound fusion in men with prior negative biopsy and elevated prostate-specific antigen," *Eur. Urol.*, vol. 65, no. 4, pp. 809–815, 2014,

[181] S. Shoji *et al.*, "Accuracy of real-time magnetic resonance imaging-transrectalultrasound fusion image-guided transperineal target biopsy withneedle tracking with a mechanical position-encoded stepper indetecting significant prostate cancer in biopsy-na€ıve men," *Int. J. Urol.*, vol. 78, 2017,

[182] R. Sun, A. Fast, I. Kirkpatrick, P. Cho, and J. Saranchuk, "Assessment of magnetic resonance imaging (MRI)-fusion prostate biopsy with concurrent standard systematic ultrasound-guided biopsy among men requiring repeat biopsy," *Can. Urol. Assoc. J.*, vol. 15, no. 9, pp. 495–500, 2021,

[183] A. Magnier, C. Nedelcu, S. Chelly, M. C. Rousselet-Chapeau, A. R. Azzouzi, and S. Lebdai, "Prostate cancer detection by targeted prostate biopsy using the 3D Navigo system: a prospective study," *Abdom. Radiol.*, vol. 46, no. 9, pp. 4381–4387, 2021,

[184] S. Miah *et al.*, "A prospective analysis of robotic targeted MRI-US fusion prostate biopsy using the centroid targeting approach," *J. Robot. Surg.*, vol. 14, no. 1, pp. 69–74, 2020,

[185] S. Hamid *et al.*, "The SmartTarget Biopsy Trial: A Prospective, Within-person Randomised, Blinded Trial Comparing the Accuracy of Visual-registration and Magnetic Resonance Imaging/Ultrasound Image-fusion Targeted Biopsies for Prostate Cancer Risk Stratification," *Eur. Urol.*, vol. 75, no. 5, pp. 733–740, 2019,

[186] F. S. Foster *et al.*, "A New 15–50 MHz Array-Based Micro-Ultrasound Scanner for Preclinical Imaging," *Ultrasound Med. Biol.*, vol. 35, no. 10, pp. 1700–1708, Oct. 2009,

[187] J. E. McNeal, E. A. Redwine, F. S. Freiha, and T. A. Stamey, "Zonal Distribution of Prostatic Adenocarcinoma: Correlation with Histologic Pattern and Direction of Spread," *Am. J. Surg. Pathol.*, vol. 12, no. 12, 1988, [Online]. Available: https://journals.lww.com/ajsp/fulltext/1988/12000/zonal_distribution_of_prostatic_adenocarcin oma_.1.aspx

[188] C. P. Pavlovich *et al.*, "A multi-institutional randomized controlled trial comparing first-generation transrectal high-resolution micro-ultrasound with conventional frequency transrectal ultrasound for prostate biopsy," *BJUI Compass*, vol. 2, no. 2, p. 126, Mar. 2021,

[189] S. Schaer *et al.*, "Assessing cancer risk in the anterior part of the prostate using micro-ultrasound: validation of a novel distinct protocol," *World J. Urol.*, vol. 41, no. 11, pp. 3325–3331, Nov. 2023,

[190] M. F. Callejas *et al.*, "Detection of clinically significant index prostate cancer using micro-ultrasound: correlation with radical prostatectomy," *Urology*, vol. 0, no. 0, Jul. 2022,

[191] A. M. Pedraza *et al.*, "Microultrasound in the detection of the index lesion in prostate cancer,"

*Prostate*, 2023,

[192] A. Priester *et al.*, "Prediction and Mapping of Intraprostatic Tumor Extent with Artificial Intelligence," *Eur. Urol. Open Sci.*, vol. 54, pp. 20–27, Aug. 2023,

[193] M. Ferro *et al.*, "Radiomics in prostate cancer: an up-to-date review," *Ther. Adv. Urol.*, vol. 14, 2022,

[194] J. J. Twilt, K. G. van Leeuwen, H. J. Huisman, J. J. Fütterer, and M. de Rooij, "Artificial intelligence based algorithms for prostate cancer classification and detection on magnetic resonance imaging: A narrative review," *Diagnostics*, vol. 11, no. 6, p. 959, Jun. 2021,

[195] A. Saha *et al.*, "Artificial intelligence and radiologists in prostate cancer detection on MRI (PI-CAI): an international, paired, non-inferiority, confirmatory study," *Lancet Oncol.*, vol. 25, no. 7, pp. 879–887, Jul. 2024,

[196] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation," *Nat. Methods 2020 182*, vol. 18, no. 2, pp. 203–211, Dec. 2020,

[197] H. Zheng *et al.*, "AtPCa-Net: anatomical-aware prostate cancer detection network on multi-parametric MRI," *Sci. Reports 2024 141*, vol. 14, no. 1, pp. 1–13, Mar. 2024,

[198] E. Redekop *et al.*, "Attention-Guided Prostate Lesion Localization and Grade Group Classification with Multiple Instance Learning," *Proceedings of Machine Learning Research*, vol. 172. PMLR, pp. 975–987, Dec. 04, 2022. Accessed: Aug. 13, 2024. [Online]. Available: https://proceedings.mlr.press/v172/redekop22a.html

[199] Y. Shao, J. Wang, B. Wodlinger, and S. E. Salcudean, "Improving Prostate Cancer (PCA) Classification Performance by Using Three-Player Minimax Game to Reduce Data Source Heterogeneity," *IEEE Trans. Med. Imaging*, vol. 39, no. 10, pp. 3148–3158, Oct. 2020,

[200] M. Gilany *et al.*, "Towards Confident Detection of Prostate Cancer Using High Resolution Micro-ultrasound," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*, Springer Science and Business Media Deutschland GmbH, 2022, pp. 411–420.

[201] M. Gilany *et al.*, "TRUSformer: improving prostate cancer detection from micro-ultrasound using attention and self-supervision," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 18, no. 7, pp. 1193–1200, Jul. 2023,

[202] P. F. R. Wilson *et al.*, "Self-Supervised Learning with Limited Labeled Data for Prostate Cancer Detection in High-Frequency Ultrasound," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 70, no. 9, pp. 1073–1083, Sep. 2023,

[203] A. R. Schned *et al.*, "Tissue-shrinkage correction factor in the calculation of prostate cancer volume," *Am. J. Surg. Pathol.*, vol. 20, no. 12, pp. 1501–1506, Dec. 1996,

[204] K. H. Zou *et al.*, "Statistical Validation of Image Segmentation Quality Based on a Spatial Overlap

Index," *Acad. Radiol.*, vol. 11, no. 2, pp. 178–189, 2004,

[205] M. Piert *et al.*, "Accuracy of tumor segmentation from multi-parametric prostate MRI and 18F-choline PET/CT for focal prostate cancer therapy applications," *EJNMMI Res.*, vol. 8, no. 1, p. 23, Dec. 2018,

[206] V. Lorusso *et al.*, "Comparison Between Micro-Ultrasound and Multiparametric MRI Regarding the Correct Identification of Prostate Cancer Lesions," *Clin. Genitourin. Cancer*, vol. 20, no. 4, pp. e339–e345, Aug. 2022,

[207] A. Priester *et al.*, "Registration Accuracy of Patient-Specific, Three-Dimensional-Printed Prostate Molds for Correlating Pathology with Magnetic Resonance Imaging," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 1, pp. 14–22, Jan. 2019,

[208] A. Fedorov *et al.*, "3D Slicer as an Image Computing Platform for the Quantitative Imaging Network," *Magn Reson Imaging*, vol. 30, no. 9, pp. 1323–1341, 2012,

[209] A. Priester *et al.*, "A system for evaluating magnetic resonance imaging of prostate cancer using patient-specific 3D printed molds," *Am. J. Clin. Exp. Urol.*, vol. 2, no. 2, p. 127, 2014, Accessed: Oct. 27, 2023. [Online]. Available: /pmc/articles/PMC4219304/

[210] S. Xu *et al.*, "Real-time MRI-TRUS fusion for guidance of targeted prostate biopsies," *https://mc.manuscriptcentral.com/tcas*, vol. 13, no. 5, pp. 255–264, Jan. 2010,

[211] M. F. Callejas, E. A. Klein, M. Truong, L. Thomas, J. K. McKenney, and S. Ghai, "Detection of Clinically Significant Index Prostate Cancer Using Micro-ultrasound: Correlation With Radical Prostatectomy," *Urology*, vol. 169, pp. 150–155, Nov. 2022,

[212] L. Wiemer *et al.*, "Evolution of Targeted Prostate Biopsy by Adding Micro-Ultrasound to the Magnetic Resonance Imaging Pathway," *Eur. Urol. Focus*, vol. 7, no. 6, pp. 1292–1299, Nov. 2021,

[213] G. Lughezzani *et al.*, "Comparison of the Diagnostic Accuracy of Micro-ultrasound and Magnetic Resonance Imaging/Ultrasound Fusion Targeted Biopsies for the Diagnosis of Clinically Significant Prostate Cancer," *Eur. Urol. Oncol.*, vol. 2, no. 3, pp. 329–332, May 2019,

[214] B. Turkbey *et al.*, "Prostate Cancer: Value of Multiparametric MR Imaging at 3 T for Detection—Histopathologic Correlation," *Radiology*, vol. 255, no. 1, p. 89, Apr. 2010,

[215] B. Nahar *et al.*, "Prospective Evaluation of Focal High Intensity Focused Ultrasound for Localized Prostate Cancer," *J. Urol.*, vol. 204, no. 3, pp. 483–489, Sep. 2020,

[216] S. Natarajan *et al.*, "Focal Laser Ablation of Prostate Cancer: Feasibility of Magnetic Resonance Imaging-Ultrasound Fusion for Guidance," *J. Urol.*, vol. 198, no. 4, pp. 839–847, Oct. 2017,

[217] M. Ittmann, "Anatomy and Histology of the Human and Murine Prostate," *Cold Spring Harb. Perspect. Med.*, vol. 8, no. 5, May 2018,

[218] M. Gilany *et al.*, "Towards Confident Detection of Prostate Cancer using High Resolution Micro-

ultrasound," Jul. 2022, Accessed: Aug. 10, 2022. [Online]. Available: http://arxiv.org/abs/2207.10485

[219]  L. Klotz *et al.*, "Comparison of micro-ultrasound and multiparametric magnetic resonance imaging for prostate cancer: A multicenter, prospective analysis," *Can. Urol. Assoc. J.*, vol. 15, no. 1, p. E11, Jul. 2021,

[220]  J. Pensa *et al.*, "Evaluation of prostate cancer detection using micro-ultrasound versus MRI through co-registration to whole-mount pathology," *Sci. Reports 2024 141*, vol. 14, no. 1, pp. 1–14, Aug. 2024,

[221]  C. Dariane *et al.*, "Micro-ultrasound-guided biopsies versus systematic biopsies in the detection of prostate cancer: a systematic review and meta-analysis," *World J. Urol.*, vol. 41, no. 3, pp. 641–651, Mar. 2023,

[222]  J. Howard and S. Gugger, "Fastai: A Layered API for Deep Learning," *Inf. 2020, Vol. 11, Page 108*, vol. 11, no. 2, p. 108, Feb. 2020,

[223]  Z. Liu, H. Mao, C. Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2022-June, pp. 11966–11976, 2022,

[224]  Z. Liu *et al.*, "Swin Transformer V2: Scaling Up Capacity and Resolution," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2022-June, pp. 11999–12009, 2022,

[225]  A. Dosovitskiy *et al.*, "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," *ICLR 2021 - 9th Int. Conf. Learn. Represent.*, Oct. 2020, Accessed: Jun. 06, 2024. [Online]. Available: https://arxiv.org/abs/2010.11929v2

[226]  "The prostate - Canadian Cancer Society." Accessed: Jun. 01, 2021. [Online]. Available: https://www.cancer.ca/en/cancer-information/cancer-type/prostate/prostate-cancer/the-prostate/?region=on