AWARD NUMBER: W81XWH-18-1-0524

TITLE: Noninvasive Risk Stratification of Prostate Cancer Patients Using Radiomic Features Derived from Magnetic Resonance Fingerprinting (MRF) and MRI

PRINCIPAL INVESTIGATOR: Rakesh Shiradkar

CONTRACTING ORGANIZATION: Case Western Reserve University, Cleveland, OH

REPORT DATE: October 2021

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release; distribution is unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
October 2021	Annual	30Sep2020-29Sep2021

4. TITLE AND SUBTITLE

Noninvasive Risk Stratification of Prostate Cancer Patients Using Radiomic Features Derived from Magnetic Resonance Fingerprinting (MRF) and MRI

5a. CONTRACT NUMBER
W81XWH-18-1-0524

5b. GRANT NUMBER
GRANT12511573

5c. PROGRAM ELEMENT NUMBER
W81XWH-17-PCRP-IDA

6. AUTHOR(S)

Rakesh Shiradkar

E-Mail: rxs558@case.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Case Western Reserve University
10900 Euclid Ave, Cleveland OH 44106

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
None

14. ABSTRACT

This project is aimed at non-invasive risk stratification of prostate cancer patients achieved through development of computer assisted tools using Magnetic Resonance Imaging (MRI) and Magnetic Resonance Fingerprinting (MRF). During the current reporting period, we build upon earlier reported work and have established associations between MR parameters including T1, T2 MRF, T2w and ADC measurements, tissue compartment ratios derived from whole mount pathology within prostate cancer, prostatitis and normal prostate. We have also showed differential correlations between MRF measurements and tissue compartments within prostate cancer and prostatitis, both within central and transition zone. In addition, we have also developed novel radiomic methods for risk stratification of prostate cancer using MRI including peri-tumoral radiomics, delta radiomics and deep learning based features which have resulted in 4 journal publications.

15. SUBJECT TERMS
None listed.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRDC
Unclassified	Unclassified	Unclassified	Unclassified	21	19b. TELEPHONE NUMBER *(include area code)*

TABLE OF CONTENTS

	Page
1. Introduction	**4**
2. Keywords	**4**
3. Accomplishments	**4**
4. Impact	**16**
5. Changes/Problems	**17**
6. Products	**18**
7. Participants & Other Collaborating Organizations	**19**
8. Special Reporting Requirements	**21**
9. Appendices	**21**

1. INTRODUCTION:

This project is aimed at non-invasive risk stratification of prostate cancer patients achieved through development of computer assisted tools using Magnetic Resonance Imaging (MRI) and Magnetic Resonance Fingerprinting (MRF). Research involved in this project involves co-registration of ex vivo digitized prostate specimens with MRI and MRF maps to enable accurate transfer of pathologic delineations from histology on to MRF and MRI and provide accurate ground truth labels. This will be followed by identification of computer derived radiomic features from these regions of interest (ROIs) on MRF and MRI that separate cancer lesions into various risk categories. These features will be used to design predictive machine learning models for risk stratification.

2. KEYWORDS:

Prostate cancer, magnetic resonance imaging, magnetic resonance fingerprinting, machine learning, computer aided diagnosis, risk stratification.

3. ACCOMPLISHMENTS:

What were the major goals of the project?

Task	**Months**	**% Completion**
Specific Aim 1: To accurately delineate cancer and benign lesions on imaging.		**In Progress**
Major Task : Co-register whole mount digitized histopathology images with MRF and MRI *(N = 40)*		
Subtask 1: *Identify the ROIs corresponding to cancerous and benign lesions on pathology. Begin with an initial subset of N = 40 studies.*	1-6	100%
Subtask 2: *Methods for automatic segmentation of prostate on MRF and MRI*	6-12	100%
Subtask 3: *Methods for co-registering MRF and MRI*	12-18	100%
Subtask 4: *Co-register histology, MRF, MRI.*	12-18	100%
Milestone(s) Achieved: Present research on registration of digitized ex vivo prostate histology with MRF and MRI at AUA 2019	18	*Conference: ISMRM 2019*
Specific Aim 2: Identify a set of MRF and MRI derived radiomic features that stratify cancerous lesions into high, intermediate, low risk categories,		**In Progress**

Major Task : Design and discover radiomic features from MRF, MRI that differentiate various cancer grades and benign lesions *(N = 90)*		
Subtask 1: *Establish morphologic basis for MRF and mpMRI signatures based on co-registration and design radiomic features. Acquire additional datasets and implement methods designed in Specific Aim 1.*	18-28	100%
Subtask 2: *Extract radiomic features from MRF and mpMRI from the ROIs*	28-30	75%
Subtask 2: *Develop a machine learning classifier using the extracted features*	28-30	50%
Milestone(s) Achieved: Present research on radiomic features derived from MRF and MRI that distinguish benign and cancerous lesions at RSNA 2019, ASCO 2020	30	*<u>Conferences</u>: ISMRM 2020, 2021 AUA 2020, SPIE 2020, 2022* *<u>Journal Publications</u>: (Eur.Rad.(2), Cancers(1), Lancet EBiomedicine(1), Lancet Digital health (1), Nature Precision Oncology (1))*
Specific Aim 3: Validate the signatures identified in Aim 2 using an independent validation dataset.		**TBC**
Subtask 1: *Extract specific radiomic features from MRF and MRI, identified in Aim 2 from the independent validation set (N = 90 training, N = 40 validation).*	30-34	25%
Subtask 2: *Validate these features using the machine learning classifier developed in Aim 2*	34-36	10%
Milestone(s) Achieved: Present research on a computerized decision support system to distinguish various grades of cancer and benign lesions using radiomic features derived from MRF and MRI at RSNA 2020, AUA 2020	36	
**TBC: To Be Completed*		

What was accomplished under these goals?

1) ***Major Activities*:**

During the current reporting period, we focused on developing deep learning framework for multi-scale imaging based prostate cancer risk stratification which is the Specific Aim 2.2. This framework was developed using T2W and DWI MRI sequences and will be extended to include T1 and T2 MRF maps. A nomogram integrating clinical variables along with deep learning identified signatures from MRI was developed which resulted in improved risk stratification compared to extant nomograms. We also were able to obtain de-identified MRF, MRI data from about N=400 patients which will be used for validating the approaches that were developed during the current and previous reporting period. This forms a significant portion of the Specific Aim 3.

In addition, we also analyzed digitized H&E stained prostate specimens post-radical prostatectomy and explored computational signatures that were associated with outcome. These will be used to establish the morphologic basis of MRF signatures which is also a part of Specific Aims 1 and 3.

2) ***Specific Objectives*:**

A. Deep leaning based risk score from MRI to identify clinically significant prostate cancer on bi parametric MRI

Summary: The objective of Specific Aim 2.2 is to develop a machine learning classifier that integrates features from MRF and MRI to distinguish clinically significant and insignificant prostate cancer (PCa) as determined on pathology. While radiomic approaches have been proposed, deep learning based methods are becoming popular for their relative ease of training and non-dependence on hand engineered approaches. Towards this end, we developed a multi-scale, deep learning framework using T2W and ADC (derived from DWI) sequences to determine a radiomic risk score. This framework will be extended along with MRF data (which is at a lower resolution) to determine an integrated MRI + MRF radiomic risk score. This study was published in *Lancet Digital Health 2021*[1].

Dataset: A multi-institutional dataset (D_1 –D_5) of N=592 PCa patients with N=823 lesions was used in this study that was approved by the IRB at University Hospitals Cleveland Medical Center. Inclusion criteria included; (a) availability of histopathology or biopsy reports (b) presence of a screening or diagnostic MRI scan in the axial view. Exclusion criteria included (a) Incomplete sequences (b) MRI scans with severe motion and/or susceptibility artifacts. In order to balance the training and the test set and make sure we validate on cases with ground-truth reference to RP specimens, we chose D1, D2 and D3 as our training cohort, Dtrain, and D4, D5 as our testing cohort, Dtest.

Methods: In order to be able to provide both T2W MRI and ADC maps as input to the network simultaneously, we co-registered the ADC maps to their corresponding T2W MRI. Deep learning based imaging predictor (DIP) was designed to exploit the dynamic nature of prostate MRI (Figure 1). An AlexNet architecture was used to identify csPCa lesions on bpMRI. The network comprised of three distinct input channels with T2W MRI, ADC maps being the first two inputs and the corresponding binary manual lesion delineation map being the third input. The network was trained

at the lesion level. The highest GGG and the maximum predicted value obtained by the network among the lesions were used to evaluate the performance at the patient level. AlexNet model available in pytorch (0.4.1) was used for training. The network weights were randomly initialized with a manual seed. Binary entropy loss function was used for training the network and an early stopping criterion was used to stop the network training with respective to the leave one set out cross validation loss.

To evaluate the stability of DIP, QIN-PROSTATE-Repeatability dataset was used. The dataset consisted of mpMRI baseline and repeat prostate MRI exams for 15 subjects with a time interval of two weeks between the scans. Among the 15 patients, suspected tumor was identified in 11 patients. One patient was excluded due to poor quality of the ADC maps. Repeatability of DIP was assessed using the remaining 10 patients. All the baseline repeat scans and manual annotations were co-registered and transformed with respect to the baseline T2W MRI. Patches were extracted from the co-registered scans and provided as input to DIP. The repeatability was calculated between the output network predictions on baseline and repeat scans in terms of intra class correlation coefficient (ICC).

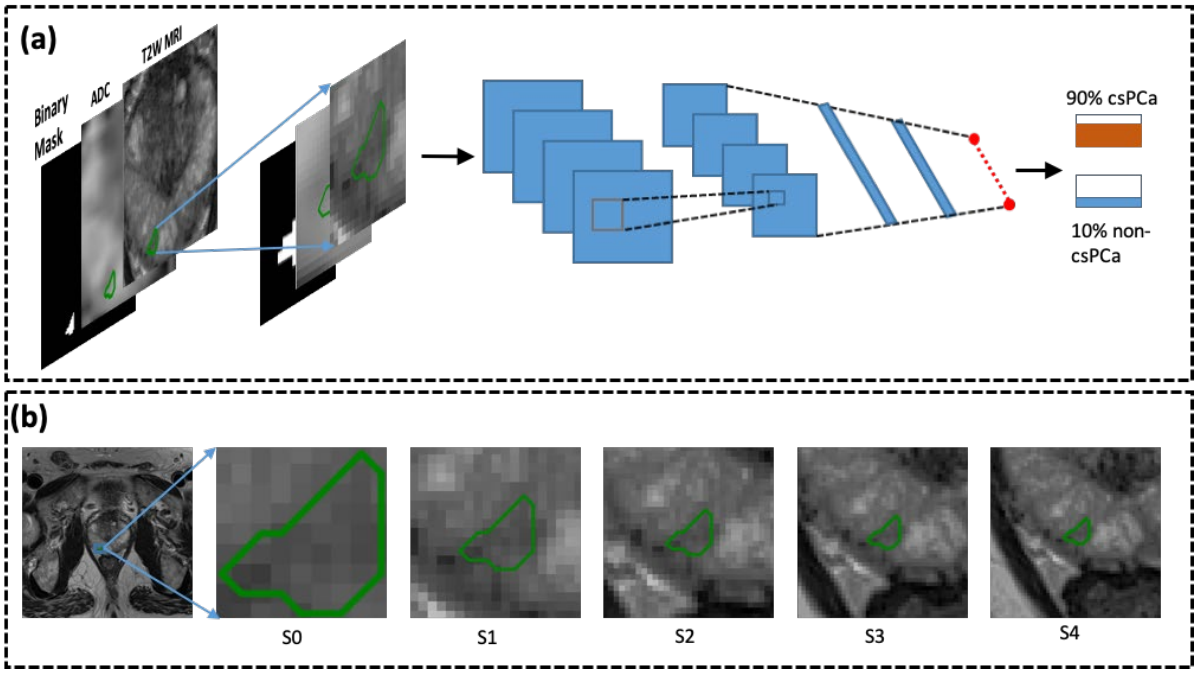

Figure. 1: Architectural diagram of deep learning based imaging predictor (DIP: CNN) and the input configuration of DIP to identify clinically significant prostate cancer lesions on bi-parametric MRI. (a) Architectural diagram of DIP depicting the three different input channels of the network (T2W MRI, ADC maps and binary lesion delineations). (b) Input configurations to DIP. The patches with respect to the first scale (S0) were extracted by drawing a bounding box around the segmented lesion and the patches with subsequent scales (S1-S4) were extracted by extending the bounding box from periphery of the delineated lesion up to 3 mm, 6 mm, 9 mm and 12 mm respectively.

Since clinical information and PI-RADS assessment scores were not available for one of the datasets (D1), a sub-cohort of the training set Dsub (D2, D3) was used to construct the nomograms. Lesion volume and prostate volume were estimated by multiplying the number of voxels from their corresponding manual annotations with their corresponding image spacing. A univariate analysis was conducted including clinical parameters (prostate volume, lesion volume and PSA), PI-RADS and

output probabilities from DIP. Additionally, a logistic regression classifier was trained in the multivariable setting to build an integrated deep learning, PI-RADS and clinical nomogram (ClaD). ClaD was further compared with integrated deep learning and clinical nomogram (DIN) constructed using a multivariable logistic regression classifier trained on DIP's prediction and clinical variables and integrated PI-RADS and clinical nomogram (PIN) constructed using a multivariable logistic regression classifier trained on PI-RADS score and clinical variables.

For identifying csPCa lesions on bpMRI, the models were trained by labeling csPCa lesions as '1' and ciPCa lesions as '0'. Area under the receiver operating characteristic curve (AUC) and other performance metrics such as accuracy, sensitivity and specificity were used to compare the performance between the models (DIP, DIN, PIN and ClaD). The difference in AUCs between the models were tested for statistical significance using DeLong's test. The cross validation AUCs were reported as mean ± standard deviation and 95% confidence intervals were indicated wherever necessary, calculated by bootstrapping the ROC curve over 2000 times. The optimal cutoff points were chosen by maximizing the accuracy in the training ROC curves. Test-retest analysis of the DIP was conducted by analyzing repeatability of the output predictions using bpMRI scans of patients taken two weeks apart. ICC(3,1) scores were reported as a measure of repeatability. Univariate and multivariable analysis was conducted to construct the nomograms. AUCs, log of odds ratio and p-values of individual variables were computed for univariate analysis, while log of odds ratio and p-values were reported for multivariable analysis. Decision curve analysis was performed to illustrate overall net-benefit of using a model with respect to the other. Additional details are provided in the manuscript [1].

3) **Significant Results**:

a) A direct attention based convolutional neural network to identify clinically significant prostate cancer:

Among the different input configurations to the network, patches with scale S1 best identified csPCa lesions with a 3-fold cross validation AUC of 0.752±0.01 using training set. The patient level analysis using the scale S1 resulted in an AUC of 0.719±0.03 in identifying patients with the presence of csPCa lesions. Furthermore, the optimal DIP resulted in an AUC=0.76; 95% confidence interval (CI) [0.71, 0.81] in identifying patients with csPCa lesions on Dtest.
In order to provide a better interpretability of DIP, gradient class activation maps (Grad-CAM) (Figure 2) were generated to identify specific regions in the image patches that saliently contributed to successful predictions. Even though the networks were forced to focus on the tumoral region using binary segmentation mask as one of the channels, intriguingly, the activation maps seemed to emphasize even on the PT regions. This finding is consistent with previous studies where deep features and radiomic features extracted from PT region added value in characterizing disease.

DIP was analyzed for test-retest repeatability using the QIN-PROSTATE-Repeatability dataset (21), that comprised two mpMRI scans (baseline and repeat) taken two weeks apart for each patient. A moderate repeatability with ICC=0.71 (N=10) was observed between the DIP's predictions on baseline and repeat scans.

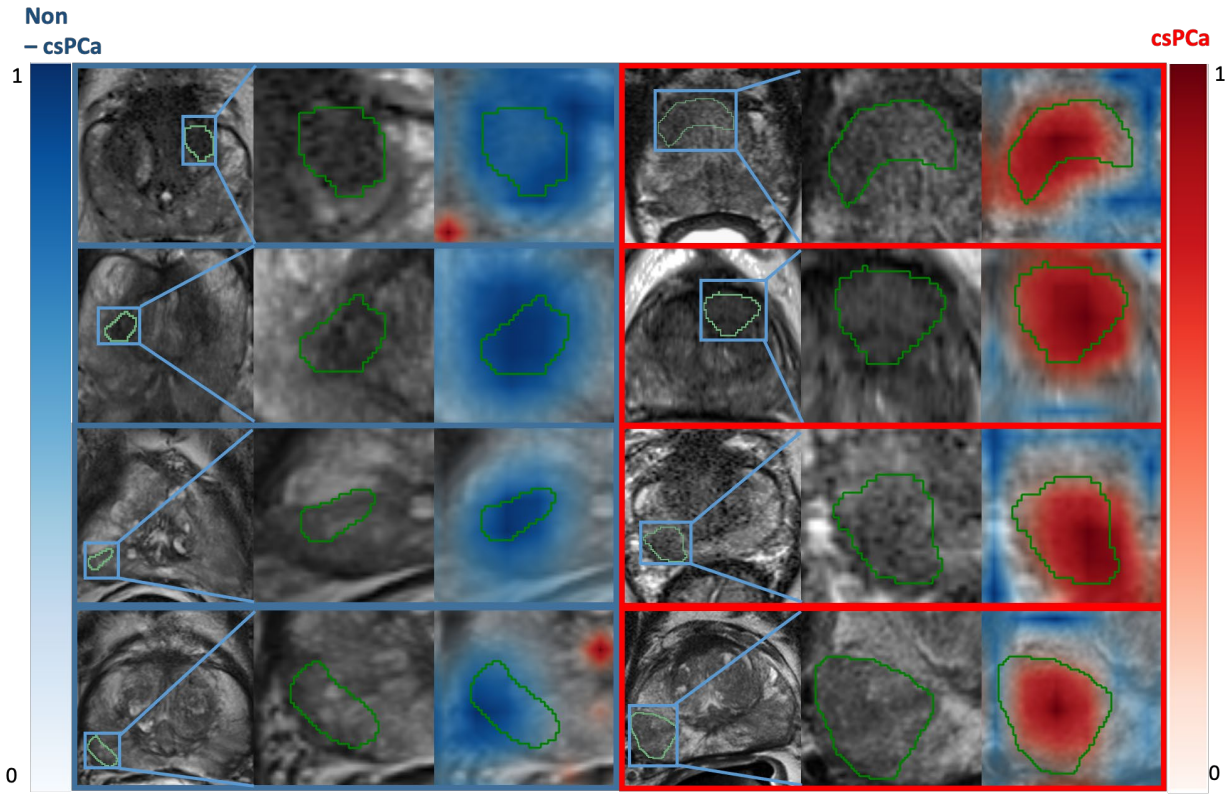

Figure. 2: Model interpretability results leveraging guided Grad-CAM to identify regions contributing to predictions of clinically insignificant / benign prostate cancer lesions (left, blue) and clinically significant prostate cancer lesions (right, red) by deep learning based imaging predictor (DIP: CNN) on scale S1 (patches extracted with bounding box extended 3mm from the tumor periphery). Activation maps emphasize that along with the tumoral region, the peri-tumoral (region immediately surrounding the tumoral region) contribute to the predictions.

Table 1: Univariate analysis of DIP, PI-RADS scoring and clinical variables (lesion volume, prostate volume and PSA).

Variable	AUC	Log (Odds Ratio)	p-value
DIP (0-1)	0.761	3.98	<0.0001
PI-RADS (1-5)	0.719	0.523	0.001
PSA (ng/ml)	0.624	0.015	0.01
Prostate Volume (mm^3)	0.760	-0.06	<0.0001
Lesion Volume (mm^3)	0.751	3.8	0.003

b) An integrated deep learning, PI-RADS and clinical nomogram to identify clinically significant prostate cancer

Univariate analysis (Table. 1) indicated that all the variables (DIP, PI-RADS, prostate volume, lesion volume and PSA) were predictive ($p<0.01$) of csPCa with AUC in the range 0.62-0.76. Multivariable analysis using DIN indicated that only DIP's prediction and prostate volume were independently predictive of csPCa irrespective of other variables. Similarly, PIN suggested that only PI-RADS and prostate volume were independently predictive of csPCa irrespective of other variables. While DIN yielded an AUC=0.79; 95% CI [0.74-0.84], PIN resulted in AUC=0.80; 95% CI [0.75-0.84] on Dtest. DeLong's test indicated that there was no statistically significant difference between predictions of DIN and PIN ($p=0.6$).

However, a multivariable analysis on ClaD indicated that DIP ($p=0.03$), PI-RADS ($p=0.02$) and prostate volume ($p=0.004$) were independently predictive of other clinical variables (lesion volume, PSA), and ClaD (Fig. 3) improved the performance over DIN ($p=0.01$) and PIN ($p=0.0006$) yielding an AUC=0.83; 95% CI [0.78-0.87] on Dtest. Furthermore, when ClaD was investigated for performance on patients with reference ground-truth from RP specimens (D4), it was found that ClaD resulted in AUC=0.85; 95% CI [0.79-0.92]. Additionally, the decision curve analysis on Dtest indicated an added net-benefit in using the integrated model ClaD over DIP and PIN.

Table 1 depicts other performance metrics of the classifiers such as accuracy, sensitivity and specificity at given cut-off points on the ROC curves. The resulting accuracies of DIP, DIN, PIN and ClaD was found to be 69.36%, 77.92%, 76.57% and 78.82% respectively.

Table 2: Performance metrics such as area under the receiver operating characteristic curve (AUC), accuracy, sensitivity and specificity for the models; deep learning imaging predictor (DIP: 2D CNN), integrated deep learning and clinical nomogram (DIN), integrated PI-RADS and clinical nomogram (PIN); and integrated deep learning, PI-RADS and clinical nomogram (ClaD).

Metric	DIP	DIN	PIN	ClaD
Threshold	0.361	0.598	0.631	0.585
AUC 95% CI	0.76 (0.71-0.81)	0.79 (0.74-0.84)	0.80 (0.75-0.84)	0.83* (0.78-0.87)
p-value	0.004	0.01	0.0006	-
Accuracy (%)	69.36	77.92	76.57	78.82
Sensitivity (%)	69.94	90.17	87.86	87.28
Specificity (%)	67.34	34.69	36.73	48.97

c) Comparison of accuracy of ClaD predictions with biopsy and extant nomograms

A subset of patients (N=36) from D4 who underwent RP along with prior 12-core systematic biopsy and MR/Ultrasound fusion biopsies was used to compare the under grading of the 12- core biopsies, MR/targeted fusion biopsies and ClaD predictions with respect to the ground-truth from RP specimens. From Fig. 4, we can observe that, although, 12- core systematic biopsies and MR/Ultrasound fusion biopsies missed indicating the presence of csPCa lesions from 10, 13 patients respectively, ClaD was found to miss csPCa lesions only from 6 patients, resulting with a sensitivity of 82.4%.

A subset of N=158 patients (49 patients with GGG=1 and 109 patients with GGG=2) from Dtest was chosen to evaluate the performance of the models DIP, DIN, PIN and ClaD in distinguishing PCa patients with GGG=1 and GGG=2 lesions. The models resulted in an AUC in the range of 0.72-0.76 in distinguishing PCa patients with GGG=1 and GGG=2 lesions with ClaD obtaining the highest AUC of 0.76; 95% CI [0.70-0.81] ($p<0.01$). The corresponding accuracy, sensitivity and specificity of ClaD in distinguishing patients with GGG=1 and GGG=2 was found to be 70.25%, 79.81% and 48.97% respectively (Figure 3A)

Prostate Biopsy Collaborative Group (PBCG) which has replaced the earlier Prostate Cancer Prevention Trial (PCPT) PCa risk calculator is one of the widely used prediction tools to estimate risk of high grade PCa (GGG>1). Among the validation sets D3 and D4, since only D4 (N=73) had all the information to evaluate the PBCG nomogram, we compared DIN, PIN and ClaD with PBCG risk calculator. All the proposed models DIN, PIN and ClaD outperformed the PBCG risk calculator ($p<0.03$) (Figure 3B) with an AUC in the range of 0.83-0.89 with ClaD resulting in the highest AUC of 0.89; 95% CI [0.83-0.94].

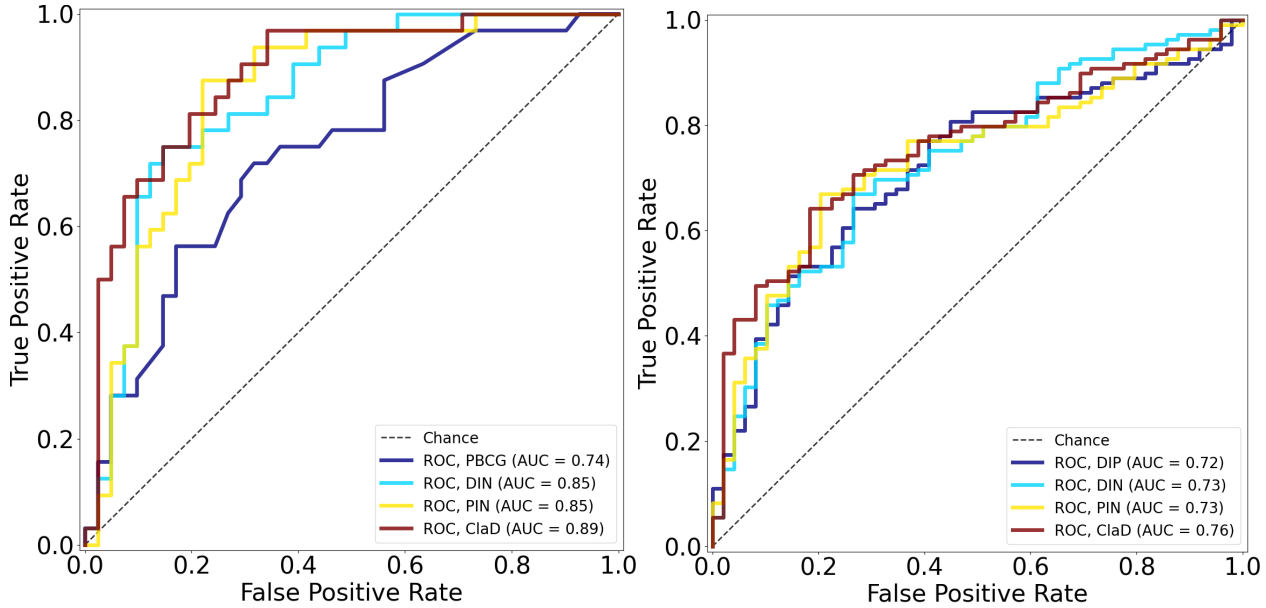

Figure 3: (A) Comparison of receiver operating characteristic curves of (a) deep learning imaging predictor (DIP), (b) integrated deep learning and clinical nomogram (c) integrated PI-RADS and clinical nomogram(PIN), (c) integrated deep learning, PI-RADS and clinical nomogram (ClaD) in distinguishing patients with GGG=1 and GGG=2 lesions (N=158 patients). (B) Comparison of receiver operating characteristic curves of (a) integrated deep learning and clinical nomogram (DIN), (b) integrated PI-RADS and clinical nomogram (PIN), (c) integrated deep learning, PI-RADS and clinical nomogram (ClaD) with Prostate Biopsy Collaborative Group (PBCG) risk calculator.

4) **Key Findings:**

In this study, we constructed an integrated DL, PI-RADS and clinical nomogram (ClaD) to identify the presence of csPCa lesions and compared the nomogram with other integrated nomograms; DIN and PIN. ClaD outperformed other nomograms, DIN and PIN in identifying patients with presence of csPCa lesions. ClaD improved performance by 4% over DIN ($p=0.01$) and 3% over PIN ($p=0.0006$) in terms of AUC. The performance improvement was found to be statistically significant by DeLong's

test. The ClaD model was also used to individualize risk assessments which showed benefit over using only the DL approach or clinical factors. Nomogram predicted score of greater than 0.585 (optimal cut off point on ROC curve) suggested the presence of csPCa lesion in a PCa patient while scores equal to 0.585 and less indicated ciPCa and benign lesions. While sensitivity of PIN and ClaD remain the same, integrating DL predictions into the clinical nomogram led to an increase in specificity from 36.73% for PIN to 48.97% for ClaD. Choosing a cut-off of 0.585 on ClaD could help in avoiding 51.03% of un-necessary biopsies among the low-risk PCa patients with ciPCa lesions or benign lesions, at the cost of missing the clinically significant disease in 12.72% of intermediate/high risk PCa patients.

Although a few previous studies have demonstrated the association of radiomic features with BCR free survival, none of the studies we are aware of has evaluated the association of DL predictions and representations with BCR free survival. Shiradkar et al trained a ML classifier on radiomic signatures to predict BCR and showed that the signatures can be predictive of BCR. However, in our work, we showed that the integrated nomogram, ClaD trained on identifying csPCa lesions on bpMRI could also predict BCR free recurrence which might help identify men who would benefit from adjuvant chemotherapy.

While a number of studies have explored MRI based PCa diagnosis and disease characterization, most of them involve data from a single site or evaluate their approach only on cross-validation set. In this work, we cross-validated our approach using data from three different sites and independently validated on data collected from two different sites. We found that our model generalized well across the external sites. All the three models (ClaD, DIP and PIN) were trained and evaluated on a large multi-institutional dataset with N=592 patients.

5) **Other achievements:**

a) Computer extracted gland features from H&E predicts prostate cancer recurrence comparably to a genomic companion diagnostic test: a large multi-site study

Existing tools for post-radical prostatectomy (RP) prostate cancer biochemical recurrence (BCR) prognosis rely on human pathologist-derived parameters such as tumor grade, with the resulting inter-reviewer variability. Genomic companion diagnostic tests such as Decipher tend to be tissue destructive, expensive, and not routinely available in most centers. We present a tissue non-destructive method for automated BCR prognosis, termed "Histotyping", that employs computational image analysis of morphologic patterns of prostate tissue from a single, routinely acquired hematoxylin and eosin slide. Patients from two institutions (n = 214) were used to train Histotyping for identifying high-risk patients based on six features of glandular morphology extracted from RP specimens. Histotyping was validated for post-RP BCR prognosis on a separate set of n = 675 patients from five institutions and compared against Decipher on n = 167 patients. Histotyping was prognostic of BCR in the validation set ($p < 0.001$, univariable hazard ratio [HR] = 2.83, 95% confidence interval [CI]: 2.03–3.93, concordance index [c-index] = 0.68, median years-to-BCR: 1.7).

Histotyping was also prognostic in clinically stratified subsets, such as patients with Gleason grade group 3 (HR = 4.09) and negative surgical margins (HR = 3.26). Histotyping was prognostic independent of grade group, margin status, pathological stage, and preoperative prostate-specific antigen (PSA) (multivariable $p < 0.001$, HR = 2.09, 95% CI: 1.40–3.10, n = 648). The combination of Histotyping, grade group, and preoperative PSA outperformed Decipher (c-index = 0.75 vs. 0.70, n = 167). These results suggest that a prognostic classifier for prostate cancer based on digital images

could serve as an alternative or complement to molecular based companion diagnostic tests. The method is illustrated in Figure 4 and this study was published in *Nature Precision Oncology 2021*[2].

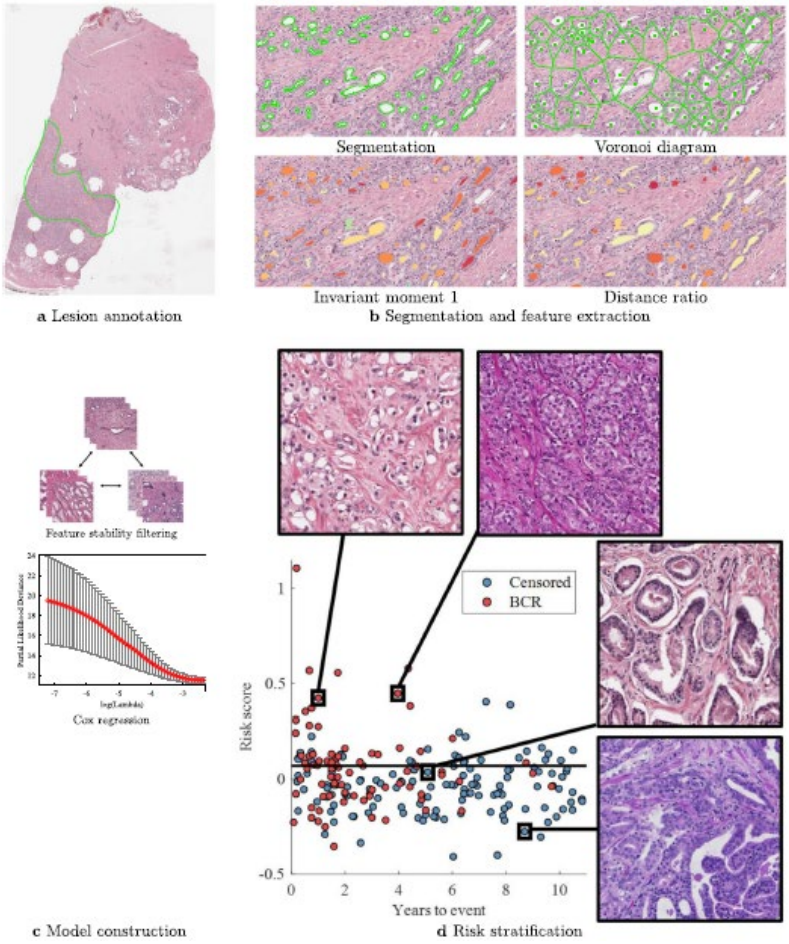

Figure 4: Pipeline of histotyping development process. a) Pathologist annotation of a representative tumor region on a whole-slide image. b) Result of automated lumen segmentation and feature visualizations from a region of interest in the pathologist-annotated tumor region. C) Steps of model training, where features are filtered for stability using the three cohorts of the training set. D) Results of model training

b) Sensitivity of radiomics to inter-reader variations in prostate cancer delineation on MRI should be considered to improve generalizability

Radiomics based predictive models using prostate MRI have been previously shown to enable better characterization of prostate cancer (PCa) and improved risk stratification. A large majority of these methods rely on manual delineation of prostate cancer (PCa) region of interest (ROI) on MRI by radiologist which may be influenced by inter-reader variations. The effect of these variations on performance and reliability of PCa predictive models has not been explored before. The purpose of our study is to investigate if radiomics derived from PCa ROIs delineated by multiple radiologists (N=3) on bi-parametric MRI (bpMRI: including T2W and ADC) significantly affect the performance of machine learning classifiers in identifying biopsy proven clinically significant PCa. A publicly available dataset (D_1) consisting of 99 patients with access to prostate 3T MRI scans, centroid location of PCa lesions and corresponding Gleason Grade Group (GGG) from targeted biopsy were included

in the study. Patients with GGG=1 were considered to be with clinically insignificant PCa (ciPCa) and those with GGG>1 were with clinically significant PCa (csPCa). T2W MRI, ADC maps and the lesion location were provided to 3 experienced GU radiologists (R1, R2 & R3) for PCa delineation using 3D Slicer software. We observed that inter-reader variations in delineating PCa ROIs affected radiomic features that showed significant differences between csPCa and ciPCa. On T2W MRI, most texture features were not consistent across readers, which could potentially be due to its relatively higher resolution compared to ADC and texture features capturing underlying heterogeneity were dependent on the extent of ROI considered. The performance on D_1 was good and relatively consistent in terms of AUC however when validated on D_2, there were significant differences. The dice similarity coefficient (DSC) measurements between R1-R2 was higher compared to the other pairs suggesting that R1 and R2 had much more consistent delineations. This was presented at the *ISMRM 2021*[3].

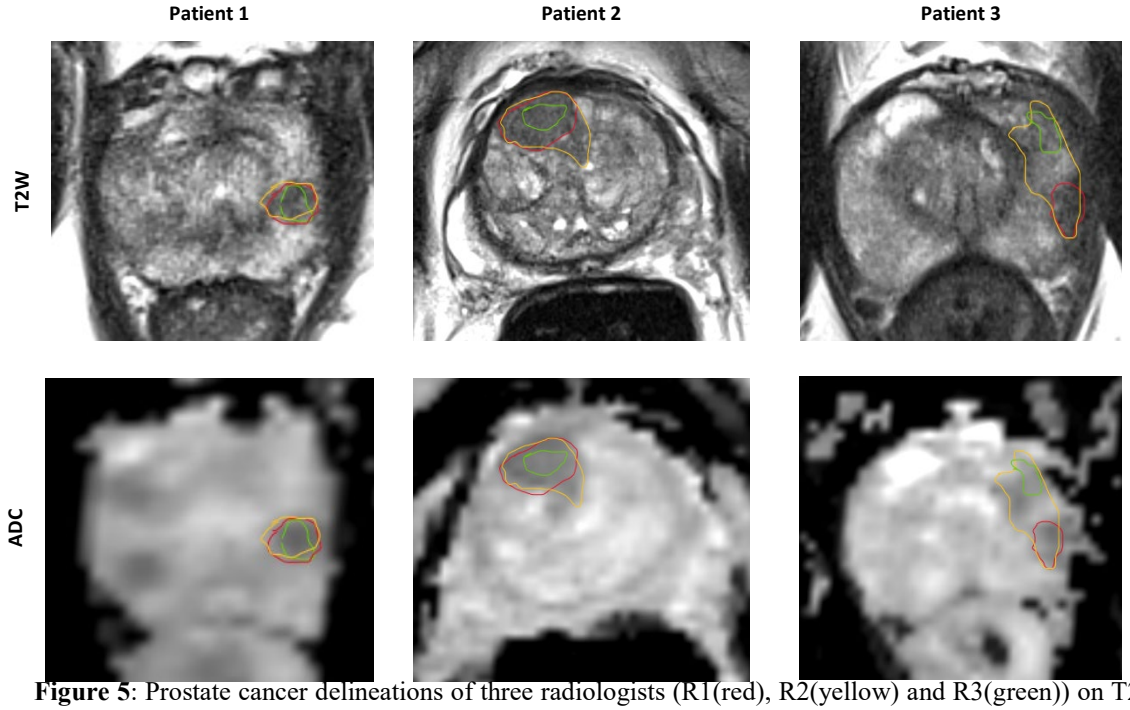

Figure 5: Prostate cancer delineations of three radiologists (R1(red), R2(yellow) and R3(green)) on T2W and ADC. For smaller lesions (patient 1), all 3 readers had a good overlap. For larger (patient 2) and multi-focal lesions (patient 3), there was considerable variation in delineations that affect radiomics and robustness of classifiers.

c) Evaluating the sensitivity of deep learning to inter-reader variations in lesion delineations on bi-parametric MRI in identifying clinically significant prostate cancer

Deep learning (DL) based convolutional neural networks (CNNs) for prostate cancer risk stratification using MRI typically use radiologist delineation of regions of interest (ROI) for analysis, which can be affected by inter-reader variations. In this study, we sought to analyze the sensitivity of classification performance of CNNs to radiologist delineations of PCa ROIs in the context of distinguishing clinically significant (csPCa) and insignificant PCa (ciPCa). We employed n=180 patient studies from 3 institutions who underwent 3T multi-parametric MRI followed by targeted

biopsies and/or radical prostatectomy. ISUP grade groups (GGG) were obtained from targeted biopsy to determine csPCa (GGG□2) and ciPCa (GGG=1). 5 radiologists delineated PCa ROIs on T2W MRI on the training set (n1=160 lesions) which were then used to train individual CNNs (N1-N5) using a SqueezeNet architecture to distinguish csPCa and ciPCa. These models were evaluated on independent test sets (n2=85 lesions, n3=29 lesions) which resulted in no significant difference in AUCs through the DeLong's Test ($p<0.05$). When CNNs were evaluated for sensitivity with respect to reader annotations, we found that all exhibited a degree of sensitivity, specifically to datasets from other institutions, with ICC(2,1) scores across D1, D2, and D3 being 0.98, 0.74, and 0.54, respectively. These results suggest that CNNs are influenced by variations in inter-reader differences in ROI delineations. This study will be presented at the *SPIE 2022*[4].

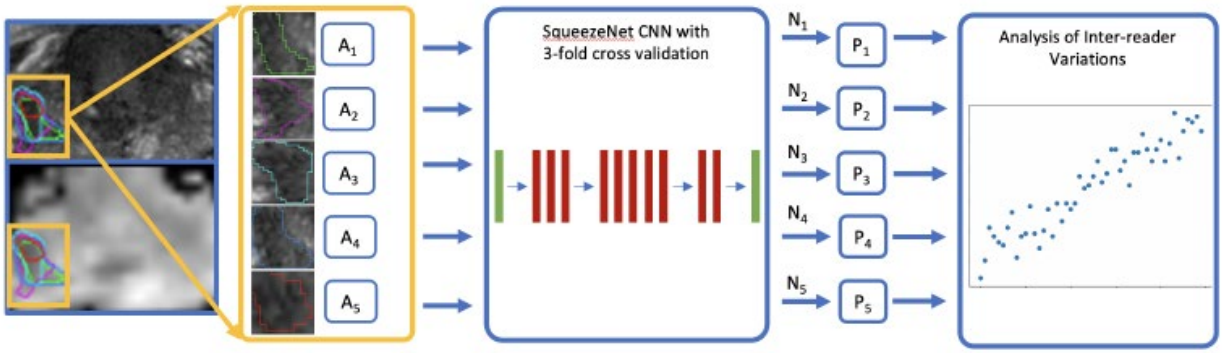

Figure 5. Flowchart for evaluating CNN sensitivity to variabilities in reader annotations on prostate MRI. PCa delineations by five radiologists (R1 (red), R2 (blue), R3 (cyan), R4 (purple), R5 (green)) on T2W MRI and ADC maps.

What opportunities for training and professional development has the project provided?

1) Interactions with my mentor and co-mentor at a higher level of project direction, collaboration has made me more independent, thoughtful in terms of designing experimental strategies, plan for future grant opportunities, publish in higher impact journals and address peer-review concerns.
2) This project thus far has enabled a deeper understanding of the histo-morphometric basis of T1, T2 MRF measurements that will result in designing optimal protocols of MRF acquisition. The ownership of this research project has made me more independent, collaborative and learn better communication skills to a wide variety of audience.
3) I have presented the work resulting out of this project at several conferences including ISMRM 2021, SPIE 2022 which have provided a wide out-reach and discussion.

How were the results disseminated to communities of interest?

The major dissemination has been through journal publications in the *Lancet Digital Health* and *Nature Precision Oncology,* presentation at conferences, including ISMRM 2021 and upcoming SPIE 2022 which have a special focus on MR Fingerprinting research. In addition, I presented my research at an internal talk to the *Program of Advanced Musculoskeletal Imaging (PAMI) at Cleveland Clinic Foundation*.

What do you plan to do during the next reporting period to accomplish the goals?

In this reporting period, we accomplished most activities and goals outlined in Specific Aim 2. The pandemic has affected acquisition of data and this was communicated to the program officer and the DoD. We have requested a 12 month no cost extension to accomplish the goals outlined in Specific Aim 3. Details are provided below:

1) The methods developed thus far for MRI will be adapted to include MRF as well in the context of Specific Aim 2 where radiomics signatures and machine learning will be leveraged to create an imaging based radiomics nomogram for prostate cancer risk stratification. Data for this task has been anonymized and will be available for analysis by the end of December 2021.
2) The goal of Specific Aim 3 would be to evaluate the radiomics based machine learning classifier on an independent hold-out dataset of N=40 patient studies.

These results will be formally compiled as a manuscript and submitted for peer-review at high impact radiology journals (such as Radiology, JMRI, European Radiology, Investigative Radiology).

4. IMPACT:

What was the impact on the development of the principal discipline(s) of the project?

The results thus far have established the pathologic basis of signatures within prostate cancer, benign prostatitis, and normal prostate on MR Fingerprinting maps. This will greatly benefit in optimizing future MRF acquisition while providing a scientific basis for the reason why MRF signatures have been excellent at distinguishing prostate cancer and the normal prostate.

On completion of this project, the radiomic risk score obtained from prostate MRF and MRI sequences will enable development of a robust, non-invasive risk stratification methodology for prostate cancer. This will fill out the current void that has continued use of invasive prostate biopsies to determine the aggressiveness of prostate cancer.

What was the impact on other disciplines?
Nothing to report.

What was the impact on technology transfer?
Nothing to report.

What was the impact on society beyond science and technology?

The results from this project will result in the development of a non-invasive, imaging-based risk score that will inform clinicians of the disease aggressiveness without the need of invasive biopsies. This will have a significant impact on society in the following areas:

1) Make patients more amenable to prostate cancer screening and avoid the stigma associated with prostate biopsies. This can result in reducing over-treatment and over-diagnosis.
2) African American men are more at risk of developing prostate cancer and mortality associated with prostate cancer. One of the reasons for this again is the presence of rectal exam and biopsies that are associated with taboo and stigma within the community. A robust and accurate imaging-based risk score can help potentially reduce the disparity and outcome in African American men.

5. CHANGES/PROBLEMS:

Changes in approach and reasons for change
Nothing to report.

Actual or anticipated problems or delays and actions or plans to resolve them

The COVID-19 pandemic resulted additional delays in our research. Our clinical collaborators had to face a stricter protocol and we were unable to have regular interactions with them needed for our project. In addition, we experienced departure of few research personnel and coordinators, and hiring was difficult due to the pandemic. We have new research personnel now and do not anticipate any further issues until the completion of this project.Our co-mentor Dr. Vikas Gulani moved from University Hospitals Cleveland Medical Center to the University of Michigan in 2019. This move impacted MRF and MRI data acquisition needed for our experiments. However, this is settled now, and we do have access to the required datasets as mentioned in the current progress report.

We have been granted a 12 month no cost extension which will now allow us to complete the proposed deliverables and utilize all remaining project funds.

Changes that had a significant impact on expenditures
Nothing to report.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents
Nothing to report.

Significant changes in use or care of human subjects
Nothing to report.

Significant changes in use or care of vertebrate animals

Nothing to report.

Significant changes in use of biohazards and/or select agents
Nothing to report.

6. PRODUCTS:

- **Publications, conference papers, and presentations**

Journal publications.

1. Hiremath A, Shiradkar R, Fu P, Mahran A, Rastinehad AR, Tewari A, Tirumani SH, Purysko A, Ponsky L, Madabhushi A. An integrated nomogram combining deep learning, Prostate Imaging–Reporting and Data System (PI-RADS) scoring, and clinical variables for identification of clinically significant prostate cancer on biparametric MRI: a retrospective multicentre study. *The Lancet Digital Health*. 2021 Jul;**3**(7):e445–e454. doi:10.1016/S2589-7500(21)00082-0 *(Acknowledgement of federal support = yes)*

2. Leo P, Janowczyk A, Elliott R, Janaki N, Bera K, Shiradkar R, Farré X, Fu P, El-Fahmawi A, Shahait M, Kim J, Lee D, Yamoah K, Rebbeck TR, Khani F, Robinson BD, Eklund L, Jambor I, Merisaari H, Ettala O, Taimen P, Aronen HJ, Boström PJ, Tewari A, Magi-Galluzzi C, Klein E, Purysko A, NC Shih N, Feldman M, Gupta S, Lal P, Madabhushi A. Computer extracted gland features from H&E predicts prostate cancer recurrence comparably to a genomic companion diagnostic test: a large multi-site study. *npj Precis Onc*. 2021 Dec;**5**(1):35. doi:10.1038/s41698-021-00174-3 *(Acknowledgement of federal support = yes)*

Books or other non-periodical, one-time publications.

Nothing to report

Other publications, conference papers and presentations.

1. Shiradkar R, Sobota M, Madabhushi A. Sensitivity of radiomics to inter-reader variations in prostate cancer delineation on MRI should be considered to improve generalizability. *Proceedings of the Annual Meeting of the ISMRM 2021. (Acknowledgement of federal support = yes)*

2. Roge A, Hiremath A, Sobota M, Tirumani SH, Bittencourt LK, Ream J, Ward J, Verma S, Purysko A, Madabhushi A, Shiradkar R. Evaluating the sensitivity of deep learning to human reader based lesion delineations in identifying clinically significant prostate cancer on MRI.

Proceedings of the SPIE Conference on Medical Imaging. 2022. *(Acknowledgement of federal support = yes)*

- **Website(s) or other Internet site(s)**
  Nothing to report

- **Technologies or techniques**
  Nothing to report

- **Inventions, patent applications, and/or licenses**

  - A Madabhushi, A Algohary, **R Shiradkar**, "Predicting prostate cancer risk of progression with multiparametric magnetic resonance imaging using machine learning and peritumoral radiomics". US Patent 11,011,265, (2021)
  - A Madabhushi, L Li, A Purysko, **R Shiradkar**, "Radiomic features of prostate bi-parametric magnetic resonance imaging (bpmri) associate with decipher score". US Patent 11,017,896, (2021)

- **Other Products**

  Nothing to report

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?

1) Rakesh Shiradkar (PI) – no change
2) Lin Li (graduate student) – no change
3) Anant Madabhushi (mentor) – no change
4) Vikas Gulani (co-mentor) – no change

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

1) Rakesh Shiradkar (PI)

<u>Ongoing Research Support</u>

W81XWH-18-1-0524 Shiradkar (PI) 10/01/2018 – 09/30/2022 **DOD PC170688**

Title: Noninvasive Risk Stratification of Prostate Cancer Patients Using Radiomic Features Derived from Magnetic Resonance Fingerprinting (MRF) and MRI
Objectives: The goal of this project is to study various computational imaging or radiomic tools in the context of prostate MR Fingerprinting (MRF) and MRI to obtain attributes that capture subtle sub-visual differences between benign and cancerous tumors and also of the different grades of cancer. The attributes will then be used in the training of robust machine learning and deep learning models that could be used in clinical settings. This will enable development of computational imaging based methods to determine the risk or aggressiveness of disease of a prostate cancer patient. The current standard of care uses core needle biopsies and blood tests which are both painful and inaccurate. Leveraging radiomics and machine learning methods using MRF and MRI can definitely advance management and care of prostate cancer patients.

Overlap: None

Annual Pilot Grant 2020, Shiradkar (PI) 05/01/2020 – 04/31/2021

CTSC Cleveland, UL1TR002548
Title: Development of a Prostate Cancer Risk Stratification Nomogram Integrating Clinical Variables and MRI derived Radiomic Signatures for Patients on Active Surveillance.
Objectives: Current standard of care for prostate cancer patients on active surveillance (AS) or watchful waiting involves repeat biopsies which are stressful, invasive and cause immense discomfort. In this proposal, we aim develop a machine learning based non-invasive prostate cancer risk stratification nomogram to identify which patients are at higher risk of progression from those at low risk on AS by leveraging computer derived imaging signatures from MRI and routinely acquired clinical variables.

Overlap: None

Completed Research Support

None

What other organizations were involved as partners?

Nothing to report

8. SPECIAL REPORTING REQUIREMENTS

COLLABORATIVE AWARDS:

QUAD CHARTS:

9. APPENDICES:

References:

1. Hiremath A, Shiradkar R, Fu P, Mahran A, Rastinehad AR, Tewari A, Tirumani SH, Purysko A, Ponsky L, Madabhushi A. An integrated nomogram combining deep learning, Prostate Imaging–Reporting and Data System (PI-RADS) scoring, and clinical variables for identification of clinically significant prostate cancer on biparametric MRI: a retrospective multicentre study. *The Lancet Digital Health*. 2021 Jul;**3**(7):e445–e454. doi:10.1016/S2589-7500(21)00082-0

2. Leo P, Janowczyk A, Elliott R, Janaki N, Bera K, Shiradkar R, Farré X, Fu P, El-Fahmawi A, Shahait M, Kim J, Lee D, Yamoah K, Rebbeck TR, Khani F, Robinson BD, Eklund L, Jambor I, Merisaari H, Ettala O, Taimen P, Aronen HJ, Boström PJ, Tewari A, Magi-Galluzzi C, Klein E, Purysko A, NC Shih N, Feldman M, Gupta S, Lal P, Madabhushi A. Computer extracted gland features from H&E predicts prostate cancer recurrence comparably to a genomic companion diagnostic test: a large multi-site study. *npj Precis Onc*. 2021 Dec;**5**(1):35. doi:10.1038/s41698-021-00174-3

3. Shiradkar R, Sobota M, Madabhushi A. Sensitivity of radiomics to inter-reader variations in prostate cancer delineation on MRI should be considered to improve generalizability. *Proceedings of the Annual Meeting of the ISMRM*.

4. Roge A, Hiremath A, Sobota M, Tirumani SH, Bittencourt LK, Ream J, Ward J, Verma S, Purysko A, Madabhushi A, Shiradkar R. Evaluating the sensitivity of deep learning to human reader based lesion delineations in identifying clinically significant prostate cancer on MRI. *Proceedings of the SPIE Conference on Medical Imaging*. 2022.

Copies of journal articles resulting from research in this grant are provided.