SEMARAK ILMU PUBLISHING

Semarak International Journal of Public Health and Primary Care

Journal homepage:
https://semarakilmu.online/index.php/sijphpc/index
ISSN: 3083-8401

Diagnostic Performance of AI in Detecting and Classifying Prostate Cancer in MRI in Comparison to Histopathological Result: A Systematic Review

Amanda Tan Jing Ying[1], Nik Nadia Hazwani Nek Kamal[1*]

[1] Department of Medical Imaging, Faculty of Health Sciences, Nursing and Education, MAHSA University, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 10 December 2024
Received in revised form 2 January 2025
Accepted 7 February 2025
Available online 15 March 2025

Keywords:
Artificial Intelligence; diagnostic performance; magnetic resonance imaging; prostate cancer; systematic review

ABSTRACT

Prostate cancer (PCa), the second leading cause of cancer death in men globally, highlights the need for effective early detection methods. While prostate needle biopsy remains the gold standard, it is invasive and relies on the skill of the practitioner. Magnetic resonance imaging (MRI) is currently the primary method for pre-biopsy detection, and artificial intelligence (AI) models are emerging as promising tools to enhance diagnostic accuracy. This systematic review systematically evaluated the diagnostic performance of MRI-based AI models for detecting and classifying prostate cancer, comparing them to histopathological results. Out of 1153 studies, 30 met the criteria for inclusion. Detection models demonstrated high performance with AUC values ranging from 0.78 to 1.00, while classification models had AUC values between 0.64 and 0.93. Sensitivity varied significantly, with detection models showing 69.6% to 100% and classification models showing 46.81% to 100%. Comparisons between AI models and radiologists' interpretations showed similar performance levels in ten studies. Overall, AI models were more effective in detecting prostate cancer than in classifying it, suggesting their potential to improve diagnostic accuracy. However, the variability in performance highlights the need for careful integration of AI into clinical practice and radiological workflows.

1. Introduction

PCa is a critical global health issue, ranking as the second leading cause of cancer mortality among men and the fourth most commonly diagnosed cancer worldwide with 1.41 million new cases reported in 2020 [28,29]. Despite the generally high survival rate associated with PCa due to its typically slow progression, the prognosis for advanced stages is markedly poorer, with a five-year survival rate plummeting to 34% [24]. Early detection plays a pivotal role in improving outcomes by allowing for timely intervention, thereby potentially halting the disease's progression and enhancing survival rates. The current gold standard for diagnosing PCa involves prostate needle biopsy [4,35]. However, this invasive procedure is highly dependent on the skill of the practitioner and can suffer from issues of underdiagnosis or misdiagnosis due to inadequate sample acquisition. The

* *Corresponding author.*
E-mail address: niknadiahazwani@mahsa.edu.my

https://doi.org/10.37934/sijphpc.3.1.109120

development of MRI has significantly improved pre-biopsy detection of PCa, with multiparametric (mp) MRI being favoured for its comprehensive imaging capabilities. Nonetheless, biparametric (bp) MRI has demonstrated comparable performance in certain contexts, highlighting the evolving landscape of MRI diagnostics [18]. In recent years, the integration of AI into medical imaging has shown promise in enhancing diagnostic accuracy and efficiency [16]. AI systems, trained on MRI images and validated against histopathological results, offer the potential to improve both detection and classification of PCa. Despite these advancements, the diagnostic performance of AI tools remains a subject of ongoing research and debate. This study aims to systematically review the diagnostic performance of AI in detecting and classifying PCa in MRI image in comparison to histopathological result. The primary goal is to assess any significant differences in diagnostic performance of AI in detecting and classifying PCa in MRI image in comparison to histopathological result. By addressing this gap, the study seeks to provide updated insights into the potential of AI to enhance PCa diagnosis and guide future integration into clinical practice.

2. Methodology

In this systematic review, the content generated was followed the guidelines and checklists stated in the PRISMA 2020, and Meta-analysis was not carried out. An English language literature search from 2013 to 2023 was carried out using the PubMed and Google Scholar database with the keywords with their variations: “Artificial intelligence”, “Magnetic Resonance Imaging”, “Prostate cancer”, and “Diagnostic performance”. A total 1153 articles were obtained in the beginning. 917 articles with irrelevant tittle and abstract, and 59 duplicated articles were excluded. Articles that appropriate for inclusion and exclusion criteria were retrieved for full-text. The inclusion criteria for the eligible studies as following: (1) articles concerning AI tool or computer-aided system in detecting and/or classifying PCa on MRI; (2) histopathological result such as biopsy and prostatectomy specimen served as the reference standard; (3) articles consist of measurable data or performance metrics such as sensitivity, specificity, accuracy, and Area Under the Curve (AUC); (4) articles 16 were in full text. The excluded studies were those met the exclusion criteria of: (1) non-English written articles; (2) article without AI-based tool or model; (3) article without diagnostic performance data; (4) article focus on other events rather than detection and classification of PCa; (5) animal studies; (6) review articles and (7) guidelines. In the end, only 30 eligible full-text studies were included in this systematic review, with 16 articles related to PCa detection and 14 articles related to PCa classification using AI models.

In this systematic review, “detection” refers to those AI tools that intentionally aimed to distinguish specific type of lesion, such as detect clinically significant (cs) and non-cs PCa, or malignant and benign lesions. “Classification” refers to those AI tools aimed to classify the prostate lesions into specific categories such as PIRADS score, CAD score, or other valid categorization methods. Figure 1 shows the flow of study selection in this study.

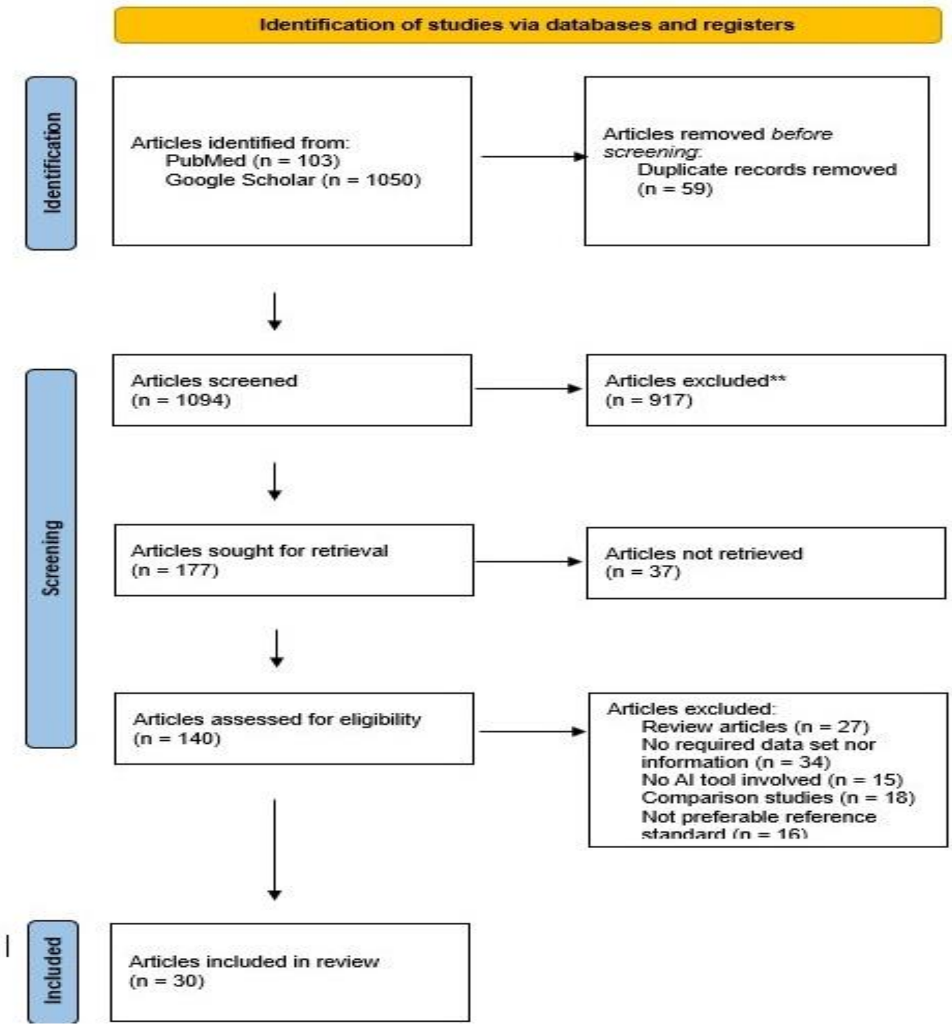

Fig. 1. Flow of study selection

3. Results

Table 1 is the summary of modelling characteristics of studies reporting on AI models in detecting prostate cancer in MRI, while Table 2 is the summary of modelling characteristics of studies reporting on ai models in classifying prostate cancer in MRI. In the evaluation of AI models for prostate cancer (PCa) detection across 16 studies, a diverse array of approaches and technologies were utilized. One study employed a commercially available system, Quantib Prostate, while others leveraged DL (7 studies), ML (5 studies), and CAD-based models (3 studies). A majority of these studies (10/16) applied automatic image segmentation, with 5 studies using manual segmentation by radiologists, and one study did not specify its segmentation approach. The diagnostic performance metrics reveal a wide range of results. For whole prostate (WP) lesion detection, the area under the curve (AUC) varied from 0.70 to 1.00. For studies focusing on both WP and specific zones: peripheral zone (PZ) and transitional zone (TZ), AUC ranged from 0.775 to 0.890. Zone-specific detection yielded an AUC of 0.88 in one study. Sensitivity ranged broadly from 69.6% to 100%, and specificity from 30.0% to 100%. Accuracy was reported in five studies, ranging from 81.4% to 92.3%, while positive predictive value (PPV) was presented in five studies with values from 76.5% to 90.1%. Negative predictive value (NPV) was reported in only two studies, showing 81.6% and 86.4%. Cohen's Kappa and precision were less commonly reported, with only one study providing these metrics (Kappa = 0.467; Precision = 83.5%). Regarding the comparison between AI models and radiologists, ten studies assessed standalone AI models' performance 3 against radiologists and reference standards. Some studies indicated that AI models improved diagnostic performance compared to radiologists, while others showed comparable or slightly inferior results. AI models demonstrated high sensitivity but varied specificity across studies.

Table 1
Summary of modelling characteristics of studies reporting on AI models in detecting prostate cancer in MRI

Author	AI model and algorithm	MRI input	Dataset (n)	Test set (n)	Validation	Image segmentation	Zone	AUC	Comparison	Outcome
Bonekamp, *et al.,* [2]	ML radiomic (RF)	T2WI, DWI-b1500, ADC	316	133	NR	Manual (MITK)	PZ, TZ	Global: .88 PZ: .84 TZ: .89	Radiologist	csPCa from benign
Chen *et al.,* [3]	ML radiomic (RF, LR)	T2WI, ADC	381	NR	30%	Manual (Artificial Intelligence Kit)	WP	(1): .999 (2): .930	Reference Standard	PCa and non-PCa (1), high- and low-grade PCa (2)
Ellmann *et al.,* [5]	ML-CAD (XGBoost, RF)	T2WI, DWI, ADC, DCE	124	24	10-fold CV	Manual	WP	.913	Radiologist	Malignant and benign
Faiella *et al.,* [6]	Quantib Prostate-DL (CNN)	T2WI, DWI, DCE	108	A: 73 B: 14 C: 21	NR	Automated	WP, PZ, TZ	NR	Radiologist	PCa
Greer *et al.,* [10]	CAD (RF)	T2WI, DWI-b2000, ADC	163	NR	NR	Automated (iCAD)	WP, TZ, PZ	.849	Radiologist	PCa
Khosravi *et al.,* [10]	AI-biopsy DL (CNN)	Axial T2WI	400	28	Five-fold CV	Automated	WP	(1): .89 (2): .78	Reference Standard	Malignant and benign lesion (1), High and low risk (2)
Li *et al.,* [11]	DL (CNN, V-Net, DenseNet)	T2WI, DWI, ADC	739	200	80	Automated (V-Net)	WP	NR	Radiologist	csPCa and non-PCa
Mehralivand *et al.,* [12]	ML (RF)	T2WI, ADC, DWI-b1500	236	236	NR	Automated	WP, TZ, PZ	Patient level: .78 Lesion level: .775	Radiologist	PCa

Mehralivand *et al.,* [13]	DL (3D UNet, AH-Net)	T2WI, DWI, ADC	525	78	79	Automated	WP	NR	Reference Standard	PCa
Min *et al.,* [14]	ML (mRMR, LASSO, radiomic signature)	T2WI, ADC, DWI-b1500	280	93	10-fold CV	Manual (ITK-SNAP)	WP	.823	Reference Standard	csPCa and ciPCa
Reda *et al.,* [20]	DL-CAD (SNCAE)	Axial DWI	53	53	Four-fold CV	Automated (NMF-based level sets)	WP	≈1.00	Reference Standard	Benign and malignant lesion
Sun *et al.,* [23]	DL (UNet)	DWI, ADC, T2WI, FS-T2WI	480	NR	NR	Automated (ITK-SNAP)	WP	NR	Radiologist	csPCa from ciPCa
Wang *et al.,* [26]	DL (DCNN) non-DL (SVM, BoW)	T2WI	172	172	10-fold CV	Automated	WP	DL: .84 Non-DL: .70	Reference Standard	PCa from benign
Woznicki *et al.,* [30]	ML (mRMR, LR),	T2WI, ADC	191	40	Five-fold CV	Manual (MITK)	WP, TZ, PZ	(1) WP: .889 PZ: .824 TZ: .683 (2) WP: .844 PZ: .894 TZ: .587	Radiologist	Malignant from benign lesions (1), csPCa from ciPCa (2)
Zhu *et al.,* [36]	CAD (ANN)	T2WI, DWI, ADC, DCE	153	153	NR	NR	WP, TZ, PZ	.89	Radiologist	csPCa from ciPCa
Zhu *et al.,* [37]	DL (CNN, Res-UNet)	T2Wi, ADC	347	140	21 csPCa cases	Automated (three segmentation Res-UNet)	WP, PZ, TZ	NR	Radiologist	csPCa from ciPCa

Table 2
Summary of modelling characteristics of studies reporting on AI models in classifying prostate cancer in MRI

Author	AI model and algorithm	MRI input	Dataset (n)	Test set (n)	Validation	Image segmentation	Zone	AUC	Comparison	Outcome
Arif *et al.,* [1]	DL-CAD (CNN)	T2WI, DWI-b800, ADC	356	36	Three-fold CV	Automated	WP	.78	Reference Standard	csPCa in low-risk patient
Gaudiano *et al.,* [7]	ML (LASSO, SVM)	T2WI, ADC	102	50	Three-fold CV	Manual	WP	.88	Reference Standard	csPCa (GG ≥ 3)
Jaouen *et al.,* [9]	CAD (binomial LR)	ADC, DCE	639	Internal test: 158 External test: 104	100 stratified CV	Manual	PZ, TZ	Internal: .82 - .84 External: .82 - .86	Radiologist	PCa (PIRADS)
Niaf *et al.,* [17]	CAD (SVM)	T2WI, ADC, DCE	30	NR	Leave-one-ROI-out CV	Manual	PZ	.872	Radiologist	PCa and benign focal lesion
Prata *et al.,* [19]	ML radiomic (Wrapper, RF)	T2WI, ADC	91	91	10-fold CV	NR	WP- PZ	.804	Reference Standard	csPCa and non-csPCa
Schelb *et al.,* [22]	DL (UNet)	T2WI, DWI-b1500, ADC	259	NR	NR	Automated	WP	NR	Radiologist	csPCa
Schelb *et al.,* [21]	DL (2D UNet)	T2WI, DWI	312	62	CV	Automated	WP	NR	Radiologist	csPCa
Thon *et al.,* [25]	Watson Elementary™ - CAD	T2WI, ADC, DCE	79	NR	NR	Manual	WP	.64	Reference Standard	PCa and benign lesion
Winkel *et al.,* [27]	ProstateAI DL (DNN)	DWI-2000, T2WI, ADC	49	NR	NR	Manually (Annotator Tool, V03_B41)	WP-PZ	NR	Reference Standard	PCa (PIRADS)
Youn *et al.,* [31]	Prostate AI DLA	T2WI, DWI	121	121	NR	NR	WP	All PCa: .808 csPCa: .828	Radiologist	PCa (PIRADS)

Zhang *et al.,* [32]	ML (nomogram, mRMR, LASSO, LR)	T2WI, DWI, ADC	159	NR	Internal: 22 External: 83	Manual (ITK-SNAP)	WP	Internal: .93 External: .84	Reference Standard	csPCa from ciPCa
Zhao *et al.,* [33]	CAD (ANN, SFS, BP, LM)	T2WI	71	NR	Leave-one-ROI-out CV	Manual	PZ, CG	PZ: .849 CG: .821	Radiologist	PCa and non-PCa
Zhong *et al.,* [34]	DTL (ResNet)	T2 SPACE, ADC	140	30	One random splitting	Manual	WP	DTL: .726 DL: .702	Radiologist	csPCa and indolent PCa
Zhong *et al.,* [35]	ML (mRMR, LR, GBDT)	T2WI, DWI, ADC, DCE	171	52	Five-fold CV	Manual (ITK-SNAP)	WP	Test set: .922 Entire set: .927	Radiologist	csPCa and non-PCa

Among the 14 studies focused on PCa classification using AI models, there was a mix of commercially available systems, non-commercial DL software, and CAD-based models. Most studies (9/14) used manual image segmentation, with three employing automated segmentation and two not specifying the method. The classification performance was assessed with AUC values ranging from 0.64 to 0.93 for WP-level classification and from 0.82 to 0.872 for zone-specific models. Sensitivity ranged from 46.81% to 100%, and specificity varied between 24.0% and 88.4%. Accuracy was reported in six studies, ranging from 50.0% to 86.4%, while PPV ranged from 48.0% to 90.5%, and NPV from 50.0% to 97.0%. One study reported Kappa and precision values of 0.2 and 84.4%, respectively. Only one study examined AI-assisted classification of lesions, showing a trend towards improved AUC, specificity, and sensitivity, though these improvements were not statistically significant. The majority of studies focused on standalone AI models, comparing them to radiologists or reference standards. Some studies found AI models to have superior specificity but inferior sensitivity compared to radiologists, while others reported improvements in sensitivity or comparable performance.

The comparison of standalone AI models with radiologists showed variable results. Only three studies demonstrated significant improvements in specificity with AI models, while others reported either declines or no significant changes in sensitivity and specificity. This variability likely arises from differences in study design, model implementation, or dataset characteristics, underscoring the need to contextualize AI results with clinical judgment and radiologist expertise. Seven studies reported on AI-assisted diagnosis, generally showing improvements in diagnostic performance, although one study [8] noted reduced specificity with AI assistance. This suggests that AI models are best used as decision support tools rather than replacements, offering valuable second opinions to radiologists. The review also highlights that AI models could potentially reduce unnecessary biopsies. Analysis of PPV and NPV from half of the studies showed PPV ranging from 57.0% to 88.3% and NPV from 50.0% to 97.0%. These results indicate that AI models can enhance diagnostic accuracy and reliability, potentially reducing the number of unnecessary biopsies. The inclusion of DCE sequences was less common but showed promise for improving diagnostic performance. Although DCE was only used in seven studies, with mixed results, integrating DCE into AI models should be further explored to determine its impact on diagnostic accuracy. However, by looking at the diagnostic performance of other AI models that did not involve DCE, the findings were aligned with the study of Monti, *et al.,* [15] that AI models involved DCE did not outperform the others.

The review assessed AI models' effectiveness based on the AUC, with varying results between detection and classification tasks. For detection, 16 studies revealed that three AI models had AUCs greater than 0.9, seven had AUCs above 0.8, and one had an AUC below 0.8. Classification models, on the other hand, showed two studies with AUCs greater than 0.9, six with AUCs above 0.8, and three with AUCs between 0.6 and 0.7. Generally, detection models outperformed classification models in diagnosing PCa. Different AI model types were analysed, including DL, ML, and CAD models. For detection, DL models had a mean AUC of 0.89, ML models had a mean AUC of 0.8895, and CAD models had a mean AUC of 0.8695, demonstrating comparable performance. In classification, ML models achieved the highest mean AUC of 0.867, followed by DL models (mean AUC of 0.755) and CAD models (mean AUC of 0.756). This indicates that ML models generally performed better than DL and CAD models in classification tasks, although the variability in AUCs among DL models complicates direct comparisons.

4. Conclusions

This systematic review evaluates the effectiveness of AI models in detecting and classifying PCa using MRI images. It finds that detection models generally outperform classification models, though both standalone and assistive AI tools hold promise for enhancing PCa diagnosis. The variability in performance metrics like sensitivity, specificity, and AUC underscores the need to integrate AI results thoughtfully into clinical workflows and contextualize them within real-world settings. To improve AI model efficacy, future research should focus on refining models, incorporating diverse datasets, and addressing inconsistencies in diagnostic processes. Prospective studies are recommended over retrospective ones for a more realistic evaluation of AI performance. Consistent performance metrics should be pre-specified to enable meaningful comparisons, and data from multiple centers should be included to reduce overfitting and enhance model reliability. Additionally, incorporating biopsy data alongside radical prostatectomy specimens in studies could improve AI models' ability to detect PCa at various stages, including early detection. Finally, involving multiple authors in systematic reviews is advised to minimize personal bias and ensure comprehensive feedback.

Acknowledgement
This research was not funded by any grant.

References

[1] Arif, Muhammad, Ivo G. Schoots, Monique J. Roobol, Wiro Niessen, and Jifke F. Veenland. "Computer aided diagnosis of clinically significant prostate cancer in low-risk patients on multi-parametric MR images using deep learning." In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pp. 1482-1485. IEEE, 2020.

[2] Bonekamp, David, Simon Kohl, Manuel Wiesenfarth, Patrick Schelb, Jan Philipp Radtke, Michael Götz, Philipp Kickingereder et al. "Radiomic machine learning for characterization of prostate lesions with MRI: comparison to ADC values." *Radiology* 289, no. 1 (2018): 128-137.

[3] Chen, Tong, Mengjuan Li, Yuefan Gu, Yueyue Zhang, Shuo Yang, Chaogang Wei, Jiangfen Wu, Xin Li, Wenlu Zhao, and Junkang Shen. "Prostate cancer differentiation and aggressiveness: assessment with a radiomic-based model vs. PI-RADS v2." *Journal of Magnetic Resonance Imaging* 49, no. 3 (2019): 875-884.

[4] Cuocolo, Renato, Maria Brunella Cipullo, Arnaldo Stanzione, Lorenzo Ugga, Valeria Romeo, Leonardo Radice, Arturo Brunetti, and Massimo Imbriaco. "Machine learning applications in prostate cancer magnetic resonance imaging." *European radiology experimental* 3 (2019): 1-8.

[5] Ellmann, Stephan, Michael Schlicht, Matthias Dietzel, Rolf Janka, Matthias Hammon, Marc Saake, Thomas Ganslandt et al. "Computer-aided diagnosis in multiparametric MRI of the prostate: An open-access online tool for lesion classification with high accuracy." *Cancers* 12, no. 9 (2020): 2366.

[6] Faiella, Eliodoro, Daniele Vertulli, Francesco Esperto, Ermanno Cordelli, Paolo Soda, Rosa Maria Muraca, Lorenzo Paolo Moramarco, Rosario Francesco Grasso, Bruno Beomonte Zobel, and Domiziana Santucci. "Quantib prostate compared to an expert radiologist for the diagnosis of prostate cancer on mpMRI: a single-center preliminary study." *Tomography* 8, no. 4 (2022): 2010-2019.

[7] Gaudiano, Caterina, Margherita Mottola, Lorenzo Bianchi, Beniamino Corcioni, Arrigo Cattabriga, Maria Adriana Cocozza, Antonino Palmeri et al. "Beyond multiparametric MRI and towards radiomics to detect prostate cancer: a machine learning model to predict clinically significant lesions." *Cancers* 14, no. 24 (2022): 6156.

[8] Greer, Matthew D., Nathan Lay, Joanna H. Shih, Tristan Barrett, Leonardo Kayat Bittencourt, Samuel Borofsky, Ismail Kabakus et al. "Computer-aided diagnosis prior to conventional interpretation of prostate mpMRI: an international multi-reader study." *European radiology* 28 (2018): 4407-4417.

[9] Jaouen, Tristan, Rémi Souchon, Paul C. Moldovan, Flavie Bratan, Audrey Duran, Au Hoang-Dinh, Florian Di Franco et al. "Characterization of high-grade prostate cancer at multiparametric MRI using a radiomic-based computer-aided diagnosis system as standalone and second reader." *Diagnostic and Interventional Imaging* 104, no. 10 (2023): 465-476.

[10] Khosravi, Pegah, Maria Lysandrou, Mahmoud Eljalby, Qianzi Li, Ehsan Kazemi, Pantelis Zisimopoulos, Alexandros Sigaras et al. "A deep learning approach to diagnostic classification of prostate cancer using pathology–radiology fusion." *Journal of Magnetic Resonance Imaging* 54, no. 2 (2021): 462-471.

[11] Li, Danyan, Xiaowei Han, Jie Gao, Qing Zhang, Haibo Yang, Shu Liao, Hongqian Guo, and Bing Zhang. "Deep learning in prostate cancer diagnosis using multiparametric magnetic resonance imaging with whole-mount histopathology referenced delineations." *Frontiers in medicine* 8 (2022): 810995.

[12] Mehralivand, Sherif, Stephanie A. Harmon, Joanna H. Shih, Clayton P. Smith, Nathan Lay, Burak Argun, Sandra Bednarova et al. "Multicenter multireader evaluation of an artificial intelligence–based attention mapping system for the detection of prostate cancer with multiparametric MRI." *American Journal of Roentgenology* 215, no. 4 (2020): 903-912.

[13] Mehralivand, Sherif, Dong Yang, Stephanie A. Harmon, Daguang Xu, Ziyue Xu, Holger Roth, Samira Masoudi et al. "Deep learning-based artificial intelligence for prostate cancer detection at biparametric MRI." *Abdominal Radiology* 47, no. 4 (2022): 1425-1434.

[14] Min, Xiangde, Min Li, Di Dong, Zhaoyan Feng, Peipei Zhang, Zan Ke, Huijuan You et al. "Multi-parametric MRI-based radiomics signature for discriminating between clinically significant and insignificant prostate cancer: Cross-validation of a machine learning method." *European journal of radiology* 115 (2019): 16-21.

[15] Monti, Serena, Valentina Brancato, Giuseppe Di Costanzo, Luca Basso, Marta Puglia, Alfonso Ragozzino, Marco Salvatore, and Carlo Cavaliere. "Multiparametric MRI for prostate cancer detection: new insights into the combined use of a radiomic approach with advanced acquisition protocol." *Cancers* 12, no. 2 (2020): 390.

[16] Nematollahi, Hamide, Masoud Moslehi, Fahimeh Aminolroayaei, Maryam Maleki, and Daryoush Shahbazi-Gahrouei. "Diagnostic performance evaluation of multiparametric magnetic resonance imaging in the detection of prostate cancer with supervised machine learning methods." *Diagnostics* 13, no. 4 (2023): 806.

[17] Niaf, Emilie, Carole Lartizien, Flavie Bratan, Laurent Roche, Muriel Rabilloud, Florence Mège-Lechevallier, and Olivier Rouvière. "Prostate focal peripheral zone lesions: characterization at multiparametric MR imaging—influence of a computer-aided diagnosis system." *Radiology* 271, no. 3 (2014): 761-769.

[18] Pesapane, Filippo, Marzia Acquasanta, Rosario Di Meo, Giorgio Maria Agazzi, Priyan Tantrige, Marina Codari, Simone Schiaffino, Francesca Patella, Anastasia Esseridou, and Francesco Sardanelli. "Comparison of sensitivity and specificity of biparametric versus multiparametric prostate MRI in the detection of prostate cancer in 431 men with elevated prostate-specific antigen levels." *Diagnostics* 11, no. 7 (2021): 1223.

[19] Prata, Francesco, Umberto Anceschi, Ermanno Cordelli, Eliodoro Faiella, Angelo Civitella, Piergiorgio Tuzzolo, Andrea Iannuzzi et al. "Radiomic machine-learning analysis of multiparametric magnetic resonance imaging in the diagnosis of clinically significant prostate cancer: new combination of textural and clinical features." *Current Oncology* 30, no. 2 (2023): 2021-2031.

[20] Reda, Islam, Ahmed Shalaby, Mohammed Elmogy, Ahmed Abou Elfotouh, Fahmi Khalifa, Mohamed Abou El-Ghar, Ehsan Hosseini-Asl, Georgy Gimel'farb, Naoufel Werghi, and Ayman El-Baz. "A comprehensive non-invasive framework for diagnosing prostate cancer." *Computers in biology and medicine* 81 (2017): 148-158.

[21] Schelb, Patrick, Simon Kohl, Jan Philipp Radtke, Manuel Wiesenfarth, Philipp Kickingereder, Sebastian Bickelhaupt, Tristan Anselm Kuder et al. "Classification of cancer at prostate MRI: deep learning versus clinical PI-RADS assessment." *Radiology* 293, no. 3 (2019): 607-617.

[22] Schelb, Patrick, Xianfeng Wang, Jan Philipp Radtke, Manuel Wiesenfarth, Philipp Kickingereder, Albrecht Stenzinger, Markus Hohenfellner, Heinz-Peter Schlemmer, Klaus H. Maier-Hein, and David Bonekamp. "Simulated clinical deployment of fully automatic deep learning for clinical prostate MRI assessment." *European radiology* 31 (2021): 302-313.

[23] Sun, Zhaonan, Kexin Wang, Zixuan Kong, Zhangli Xing, Yuntian Chen, Ning Luo, Yang Yu et al. "A multicenter study of artificial intelligence-aided software for detecting visible clinically significant prostate cancer on mpMRI." *Insights Into Imaging* 14, no. 1 (2023): 72.

[24] Survival Rates for Prostate Cancer. (2024, January 17). Retrieved from American Cancer Society.

[25] Thon, Anika, Ulf Teichgräber, Cornelia Tennstedt-Schenk, Stathis Hadjidemetriou, Sven Winzler, Ansgar Malich, and Ismini Papageorgiou. "Computer aided detection in prostate cancer diagnostics: A promising alternative to biopsy? A retrospective study from 104 lesions with histological ground truth." *PLoS One* 12, no. 10 (2017): e0185995.

[26] Wang, Xinggang, Wei Yang, Jeffrey Weinreb, Juan Han, Qiubai Li, Xiangchuang Kong, Yongluan Yan et al. "Searching for prostate cancer by fully automated magnetic resonance imaging classification: deep learning versus non-deep learning." *Scientific reports* 7, no. 1 (2017): 15415.

[27] Winkel, David J., Christian Wetterauer, Marc Oliver Matthias, Bin Lou, Bibo Shi, Ali Kamen, Dorin Comaniciu, Hans-Helge Seifert, Cyrill A. Rentsch, and Daniel T. Boll. "Autonomous detection and classification of PI-RADS lesions in an MRI screening population incorporating multicenter-labeled deep learning and biparametric imaging: proof of concept." *Diagnostics* 10, no. 11 (2020): 951.

[28] World Cancer Research Fund International. (2022, March 23). Prostate cancer statistics. Retrieved from World Cancer Research Fund International.

[29] World Health Organization. (2022, February 3). Cancer. Retrieved from World Health Organization.

[30] Woźnicki, Piotr, Niklas Westhoff, Thomas Huber, Philipp Riffel, Matthias F. Froelich, Eva Gresser, Jost von Hardenberg et al. "Multiparametric MRI for prostate cancer characterization: combined use of radiomics model with PI-RADS and clinical parameters." *Cancers* 12, no. 7 (2020): 1767.

[31] Youn, Seo Yeon, Moon Hyung Choi, Dong Hwan Kim, Young Joon Lee, Henkjan Huisman, Evan Johnson, Tobias Penzkofer et al. "Detection and PI-RADS classification of focal lesions in prostate MRI: Performance comparison between a deep learning-based algorithm (DLA) and radiologists with various levels of experience." *European Journal of Radiology* 142 (2021): 109894.

[32] Zhang, Yongsheng, Wen Chen, Xianjie Yue, Jianliang Shen, Chen Gao, Peipei Pang, Feng Cui, and Maosheng Xu. "Development of a novel, multi-parametric, MRI-based radiomic nomogram for differentiating between clinically significant and insignificant prostate cancer." *Frontiers in oncology* 10 (2020): 888.

[33] Zhao, Kai, ChengYan Wang, Juan Hu, XueDong Yang, He Wang, FeiYu Li, XiaoDong Zhang, Jue Zhang, and XiaoYing Wang. "Prostate cancer identification: quantitative analysis of T2-weighted MR images based on a back propagation artificial neural network model." *Science China Life Sciences* 58 (2015): 666-673.

[34] Zhong, Xinran, Ruiming Cao, Sepideh Shakeri, Fabien Scalzo, Yeejin Lee, Dieter R. Enzmann, Holden H. Wu, Steven S. Raman, and Kyunghyun Sung. "Deep transfer learning-based prostate cancer classification using 3 Tesla multi-parametric MRI." *Abdominal Radiology* 44 (2019): 2030-2039.

[35] Zhong, Jian-Guo, Lin Shi, Jing Liu, Fang Cao, Yan-Qing Ma, and Yang Zhang. "Predicting prostate cancer in men with PSA levels of 4–10 ng/mL: MRI-based radiomics can help junior radiologists improve the diagnostic performance." *Scientific reports* 13, no. 1 (2023): 4846.

[36] Zhu, Lina, Ge Gao, Yi Liu, Chao Han, Jing Liu, Xiaodong Zhang, and Xiaoying Wang. "Feasibility of integrating computer-aided diagnosis with structured reports of prostate multiparametric MRI." *Clinical Imaging* 60, no. 1 (2020): 123-130.

[37] Zhu, Lina, Ge Gao, Yi Zhu, Chao Han, Xiang Liu, Derun Li, Weipeng Liu et al. "Fully automated detection and localization of clinically significant prostate cancer on MR images using a cascaded convolutional neural network." *Frontiers in Oncology* 12 (2022): 958065.