



Exploring the Use of Artificial Intelligence in the Management of Prostate Cancer

Timothy N. Chu¹ · Elyssa Y. Wong¹ · Runzhuo Ma¹ · Cherine H. Yang¹ · Istabraq S. Dalieh¹ · Andrew J. Hung¹ 

Accepted: 30 January 2023 / Published online: 18 February 2023
© The Author(s) 2023

Abstract

Purpose of Review This review aims to explore the current state of research on the use of artificial intelligence (AI) in the management of prostate cancer. We examine the various applications of AI in prostate cancer, including image analysis, prediction of treatment outcomes, and patient stratification. Additionally, the review will evaluate the current limitations and challenges faced in the implementation of AI in prostate cancer management.

Recent Findings Recent literature has focused particularly on the use of AI in radiomics, pathomics, the evaluation of surgical skills, and patient outcomes.

Summary AI has the potential to revolutionize the future of prostate cancer management by improving diagnostic accuracy, treatment planning, and patient outcomes. Studies have shown improved accuracy and efficiency of AI models in the detection and treatment of prostate cancer, but further research is needed to understand its full potential as well as limitations.

Keywords Artificial intelligence · Machine learning · Prostate cancer · Radiomics · Pathomics

Introduction

Artificial intelligence (AI) and machine learning (ML) are rapidly advancing fields that have the potential to revolutionize many industries, including medicine. AI involves the development of intelligent systems to perform tasks that typically require human intelligence, such as recognizing patterns and making decisions. AI performance is driven by ML, which harnesses algorithms and statistical models to automatically improve system performance on a specific task through experience. Over the last decade, AI has started to become increasingly integrated into medicine. Specifically, in the field of urology, AI is being tested and implemented as a tool to aid in the diagnosis and treatment of prostate cancer. AI-driven techniques are highly appealing because they can quickly analyze large amounts of data, such

as medical images and tissue samples, to identify patterns and make predictions about the likelihood of cancer [1]. In addition, AI-based techniques show the potential to increase the accuracy of prostate cancer diagnosis and improve treatment plans for patients. In this review, we explore recent AI advances in the diagnosis, prognosis, and treatment of localized prostate cancer. We share key studies and the impact they bring to each of these areas and highlight potential avenues for future research.

Methods

A comprehensive review of the current literature was performed using the PubMed-Medline database up to 2023 using the term “urology” combined with the following terms: “prostate cancer” and “artificial intelligence.” To capture recent trends in ML and DL applications, our search was focused on articles published within the last 4 years and originally published in English. Review articles and editorials were excluded. Publications relevant to the subject and their cited references were retrieved and appraised independently by 2 authors for inclusion in the final manuscript.

This article is part of the Topical Collection on *Prostate Cancer*

✉ Andrew J. Hung
andrew.hung@med.usc.edu

¹ Center for Robotic Simulation & Education, Department of Urology, USC Institute of Urology, University of Southern California, Catherine & Joseph Aresty 1441 Eastlake Avenue Suite 7416, Los Angeles, CA 90089, USA

Diagnosics

Radiomics

Analysis of cross-sectional radiographic images, like those produced by CT scans or MRI, is used to identify complex patterns, a task that AI can be trained to perform quickly, accurately, and consistently. The extraction of quantitative features from medical imaging, otherwise known as radiomics, has been studied for use in clinical settings [2]. Specifically, in the context of prostate cancer, ML algorithms can automatically extract qualitative information regarding tumor features such as size, shape, texture, and intensity from medical imaging data. This quantitative data can be delivered back to a urologist to provide objective, data-driven insights into the characteristics and behavior of prostate tumors and offer insight into monitoring response to treatment and prediction of outcomes. This allows clinicians to adjust treatment plans as needed with less delay (Table 1).

Radiomics and Gleason Score

The Gleason grading system is based on the microscopic appearance of cancer cells and remains the gold standard for grading prostate cancer. Traditionally, tissue samples are obtained from the prostate gland during a biopsy or after surgery which can have associated complications (such as bleeding, infection, and urosepsis). Multiple studies have explored whether radiomic features extracted from MRI scans can be used to predict the Gleason score as an alternative to traditional methods. Antonelli et al. performed a study in which machine learning classifiers were trained on various MRI features to classify and compare transitional and peripheral zone prostate tumors. The sensitivity of their peripheral zone model was 0.93, compared to an average sensitivity of 0.72 for three radiologists [3]. In another study, Fehr et al. used a support vector machine classifier trained on apparent diffusion coefficients and T2-weighted MRI-based texture features from a cohort of 217 men to accurately distinguish between Gleason scores 6 and 7 or higher, with an area under the curve (AUC) of 0.93 in both the peripheral and transition zones [4]. The ability to use MRI images to predict prostate tumor Gleason scores may provide information that is useful in guiding treatment while also reducing the risk of complications associated with more invasive means of tissue sampling.

Clinically Significant Prostate Cancer vs Clinically Insignificant Prostate Cancer

Multiparametric prostate MRI (mpMRI) has become more widely used in the diagnosis of clinically significant

prostate cancer. The Prostate Imaging Reporting and Data system (PI-RADS) is an international standard for acquiring, interpreting, and reporting MRI images of the prostate [5]. PI-RADS 1 and 2 lesions usually indicate lower chances of clinically significant cancer, while PI-RADS 4 and 5 lesions usually indicate higher chances of clinically significant prostate cancer. Category 3 lesions pose challenges to clinicians and radiologists alike due to the ambiguity of this designation in terms of clinically significant or insignificant prostate cancer. There is growing evidence that ML models can rival the performance of clinical radiologists in the assessment of PI-RADS lesions [6]. These models may offer the means to clearly distinguish between clinically significant prostate cancer (csPCa) and clinically insignificant prostate cancer (ciPCa). Hectors et al. used machine learning to construct and cross-validate a model using radiomic features from T2-weighted imaging of PI-RADS 3 lesions to identify clinically significant prostate cancer [7]. Using a training set of 188 subjects and a test set of 52 subjects, they were able to train a random forest classifier with an AUC of 0.76 for predicting csPCa in the test set. In another study, Min et al. used nine radiomic features to train a LASSO algorithm that accurately distinguished between csPCa and ciPCa, with an AUC of 0.82, sensitivity of 0.883, and specificity of 0.753 in the training set, and an AUC of 0.823, sensitivity of 0.841, and specificity of 0.727 in the test cohort [8]. Woznicki et al. developed predictive machine learning models and compared them to PI-RADS-based assessments by radiologists to determine their ability to identify malignant versus benign prostate cancer and csPCa versus ciPCa [9]. In their test cohort, they achieved an AUC of 0.889 with their ensemble model in differentiating between malignant and benign prostate lesions. Their model was also able to achieve an AUC of 0.844 in distinguishing csPCa and ciPCa.

Risk Stratification

Risk stratification is used to assess the likelihood of cancer progression and helps determine the appropriate course of treatment based on this risk. Using existing tools, patients can be classified as having low-, intermediate-, or high-risk prostate cancer with different clinical implications for each. Appropriate treatment depending on risk level may include active surveillance, surgery, radiation therapy, and/or hormone therapy, given the differences in the importance of an accurate risk stratification method. One application of ML in risk stratification was demonstrated by Winkel et al., who investigated whether machine learning algorithms in combination with biparametric imaging could accurately detect and classify prostate lesions in asymptomatic men [10]. They trained a model using a cohort of 48 men, 38 of whom had high-risk lesions while 10 were lesion-free. The model was able to identify and classify 100% of the highest-risk

Table 1 Radiomics

Reference number	Reference	Year	Population	Modality	Outcome/prediction	AUC/performance
[3]	Antonelli et al.	2019	164 male patients (119 PZ, 45 TZ) training, 30 test set	mpMRI	Gleason 4 component	Peripheral zone model: AUC 0.83 Transitional zone model: AUC 0.75
[4]	Fehr et al.	2015	147 male patients	mpMRI	Gleason score	Distinguished GS 6 (3 + 3) and GS ≥ 7 cancers with 93% accuracy Distinguished GS 7 (3 + 4) from GS 7 (4 + 3) with 92% accuracy
[6•]	Schelb et al.	2019	250 male patients training set, 62 male patients test set	mpMRI	csPCa	Sensitivity 88% and specificity 50% for PI-RADS ≥ 4 vs U-Net
[7]	Hectors et al.	2021	188 male patients training set, 52 male patients test set	mpMRI	csPCa	AUC 0.76 ($p = 0.022$) for prediction of csPCa in test set using T2WI radiomics features
[8]	Min et al.	2019	187 male patients training set, 93 male patients test set	mpMRI	csPCa vs ciPCa	Training set: AUC 0.872 Test set: AUC 0.823
[9]	Wozniacki et al.	2020	191 male patients	mpMRI	csPCa vs ciPCa	Differentiation of malignant from benign prostatic lesions: AUC 0.889 csPCa from ciPCa: AUC 0.844
[10]	Winkel et al.	2020	48 male patients	bpMRI	PCa screening	Sensitivity 87%, specificity 50%, $k = 0.42$ PI-RADS 3: 43% detection PI-RADS 4: 73% detection PI-RADS 5: 100% detection
[11]	Varghese et al.	2019	68 male patients	mpMRI	To identify the best performing classifier for PCa risk stratification	AUC 0.92
[12]	Cysouw et al.	2021	76 male patients	PSMA PET-CT	Metastatic disease or high-risk features (LNI, distant metastasis, and Gleason score)	LNI AUC 0.86, nodal or distant metastasis AUC 0.86, Gleason score AUC 0.81, and ECE 0.76
[13]	Papp et al.	2021	52 male patients	PET/MRI	Prostate lesion-specific low-vs-high risk	Low-vs-high risk lesion prediction model AUC: 0.86
[14]	Akatsuka et al.	2022	772 male patients, 2899 ultrasound images	Ultrasound	High-grade PCa	Biochemical recurrence model AUC: 0.9 Overall patient risk model AUC: 0.94 AUC using clinical data 0.691 AUC using clinical data and ultrasound imaging data 0.835
[15]	Wildeboer et al.	2020	48 male patients	Ultrasound	Multiparametric classification of PCa	The multiparametric classifier reached a ROC-AUC of 0.75 and 0.90 for PCa and Gleason > 3 + 4 significant PCa, respectively
[16]	Khosravi et al.	2021	400 male patients	MRI	Benign vs cancerous, high-risk vs low-risk PCa	Cancer vs benign: 0.89 (95% confidence interval [CI]: [0.86–0.92]) High- vs low-risk: 0.78 (95% CI: [0.74–0.82])

bpMRI biparametric MRI, *mpMRI* multi-parametric MRI, *PSMA* quantitative prostate-specific membrane antigen, *ECE* extracapsular extension, *LNI* lymph node invasion, *SWE* shear-wave elastography, *DCE-US* dynamic contrast-enhanced ultrasound, *ROC-AUC* region-wise area under the receiver operating characteristics curve

lesions (PI-RADS category 5) and 73% of the intermediate-risk lesions (PI-RADS category 4). Varghese et al. developed a quadratic kernel-based support vector machine (SVM) algorithm that used 110 radiomic features to distinguish between high-risk and low-risk prostate cancer [11]. The algorithm achieved a positive predictive value of 0.57 and a negative predictive value of 0.84 when tested on a study cohort of 68 patients who were divided into high-risk and low-risk groups. Cysouw et al. used pre-operative PET-CT scans from 76 patients with intermediate- to high-risk PCa, to train random forest models to predict lymph node metastasis, Gleason score > 8, and extracapsular extension [12]. This ML model was capable of predicting lymph node invasion (AUC 0.86 ± 0.15 , $p < 0.01$), lymph node/distant metastasis (AUC 0.86 ± 0.14 , $p < 0.01$), Gleason score > 8 (AUC 0.81 ± 0.16 , $p < 0.01$), and extracapsular extension (AUC 0.76 ± 0.12 , $p < 0.01$) with high accuracy [12]. Papp et al. used a combination of PET-MRI to predict low versus high-risk lesions, biochemical recurrence, and overall patient risk [13].

Ultrasound

Ultrasound is a cost-effective and efficient tool that offers insightful information in the context of prostate cancer. There have been recent attempts to apply ML analysis to ultrasound imaging data in high-grade prostate cancer [14, 15]. The model demonstrated by Akatsuka et al. used a combination of ultrasound images and clinical data to achieve an AUC of 0.835 in the detection of high-grade cancer (Gleason grade group ≥ 4) [14]. Wildeboer et al. leveraged ultrasound imaging and a random forest-based classifier to improve the localization of Gleason > 3 + 4 prostate [15]. The application of US imaging is still an emerging area of research within radiomics that has the potential to supplement more established imaging modalities such as MRI and CT.

Mixed Modality

Multimodal ML-based approaches blend the strengths of different imaging modalities to optimize the visualization of anatomic structures (i.e., CT or MRI) with other modalities that emphasize function (i.e., PET, US). Khosravi et al. utilized an AI-driven approach that combined MRI and histopathologic data from biopsy reports to increase the accuracy of PI-RADS scoring [16]. Another study utilized mpMRI as a prescreening test before TRUS-biopsy among men with clinical suspicion of prostate cancer.

Pathomics

Pathomics involves using AI to analyze tissue samples, such as biopsy samples, to identify prostate cancer at the

molecular level [17]. The Gleason grading scale remains the strongest predictor of prostate cancer prognosis. ML systems provide an opportunity to reduce inter-observer variability, improve diagnostic accuracy, and streamline the process of grading prostate biopsies. Automated Gleason grading has the potential to produce more objective and reproducible score assignments [18], while performing at a level similar to pathologists, proving a reliable tool for screening or an additional layer of verification [19]. Kott et al. tested an AI-based system for detecting prostate cancer which yielded 91.5% accuracy in classifying slides as either benign or malignant, and 85.4% accuracy in finer classifications of benign vs Gleason 3 vs 4 vs 5. The model experienced the greatest difficulty in differentiating between Gleason 3 and 4, and Gleason 4 and 5 [20]. Gleason 4 pathology posed a challenge for automated detection methods when it presented as small or fused glands without lumina. Automatic detection of Gleason pattern and grade groups classification using a convolutional neural network (CNN) resulted in a 90% accuracy in differentiating between Gleason scores 3 and 4 [21]. Another algorithm based on CNN and ML showed accuracy in detecting Gleason 3 and 4, but to a lesser degree for Gleason 5 [22]. AI computing power has also been harnessed to transform two-dimensional histopathology slides into 3D computational models in an effort to improve risk stratification for patients with prostate cancer [23]. In a study conducted by da Silva et al., the implementation of AI-based systems in histopathology was shown to reduce analysis and diagnostic time by approximately 65.5% and aid in the identification of prostate cancer in patients who were not previously diagnosed by 3 histopathologists [24]. A population-based diagnostic study trained an AI system to reliably detect and grade prostate cancer in needle core biopsies comparable to expert pathologists. AI systems with clinically acceptable accuracy could alleviate the demand on pathologists by screening out benign biopsies and automating the measurement of cancer length in malignant biopsies [25]. One important caveat of AI-based techniques is the potential for bias in classification performance due to patch-wise comparison and training on a single expert data. Nir et al. found that this can be ameliorated by using patient-based cross-validation and training on multiple expert data [26]. Efforts to understand and reduce bias are valuable to improving AI algorithms.

Treatment

Surgical Skill Assessment

Conventionally, surgical skill evaluation is performed manually by human graders, which is time-consuming and prone to observer biases [27]. AI provides an ideal

solution to both issues. Utilizing enriching data (i.e., surgical videos and instrument kinematics) derived from surgery, AI is starting to show promise in surgical assessment [28]. AI can be combined with surgical metrics to assess surgeons. Hung et al. used kinematic metrics derived from surgical robots (e.g., path length and velocity of instruments) to distinguish surgeons' skill levels and predict surgical outcomes after robotic assisted radical prostatectomy (RARP) [29]. Empowered by AI, such models were able to predict surgeons' experience level, short-term outcomes such as length of hospital stay, and long-term outcomes such as continence recovery [29–32]. Interestingly, AI models using surgical performance represented by automated performance metrics (APMs) better predict surgical outcomes than only using a surgeon's prior experience. This brings into question the view of a surgeon's experience as the presumed gold standard for performance proxy [30]. AI-aided vision recognition has also been used to assess surgical performance directly from surgical videos. Khalid et al. developed a machine learning model using the JIGSAWS video footage to accurately detect surgical actions (needle passing, suturing, knot tying) and predict performance levels (novice, intermediate, expert) [33]. Baghdadi et al. described machine learning analysis of color and texture to recognize anatomical structures during pelvic lymph node dissection and predict dissection quality. The automated skill assessment output from their model compares favorably with manually scored expert ratings of lymph node dissection quality (83.3% accuracy), setting the stage for further evaluation of these training tools [34]. Hung et al. trained a deep learning model to give robotic suturing assessments in four domains—needle positioning, needle entry angle, needle driving, and needle withdrawing [35•].

Another innovation in recent years is that AI models can automatically recognize basic phases in surgery such as different surgical steps. This can significantly reduce the time associated with surgical video review and help maintain a good surgical library for educational purposes. For example, Zia et al. applied a machine learning model to automate the segmentation of RARP into 12 surgical steps. Compared with expert annotations, the model correctly annotated most RARP steps with less than 200 s of error [36].

AI has even demonstrated the ability to recognize single instrument movements in a surgical procedure and classify them into different categories, namely, surgical gestures. Luongo et al. trained deep-learning-based computer vision algorithms to identify different suturing gestures during vesicourethral anastomosis of RARP with an AUC of 0.87 [37]. Kiyasseh et al. trained a multi-purpose model which could not only recognize surgical gestures (both dissection and suturing), but could also evaluate surgical quality for multiple different steps of RARP [38]. Furthermore, by

breaking down surgery into individual surgical gestures, differences in gesture usage have been found between experienced surgeons and trainees, providing new insights for both surgical assessment and training. By giving surgical gesture sequences to AI models, a study has been able to predict patients' erectile function recovery in the long term [39•]. This opens up a new avenue for surgical assessment and training (Table 2).

Brachytherapy/Radiation

AI also has been studied in the context of nonsurgical treatment planning for prostate cancer. In a study done by McIntosh et al., a ML model was trained to make plans for external radiation therapy (RT) [40]. When evaluated by a third blinded clinician, 89% of the ML-generated RT plans were deemed clinically acceptable, and 72% were chosen in head-to-head comparisons over human-generated RT plans. The median amount of time needed for the full RT planning procedure decreased by 60.1% (118 to 47 h). Interestingly, though the ML-generated plan performed well in simulation, the treating physician's choice of the consensus-reviewed, quantitatively superior ML RT plans at the deployment phase decreased by 21%. These results highlight the fact that even in the presence of expert blinded review, retrospective or simulated evaluation of ML approaches may not be an accurate representation of algorithm acceptance in a real-world clinical situation when patient care is at risk.

Low-dose-rate prostate brachytherapy treatment takes place by implantation of small radioactive seeds in—and sometimes adjacent to the prostate gland—under the guidance of transrectal ultrasound images. In the planning process, it is standard practice to draw a line that closely resembles the genuine prostate boundary. To obtain a planned goal volume, the border is then dilated in relation to the clinical recommendations. This manual contouring is a laborious task with a significant amount of observer variability. To combat this, Nouranian et al. proposed an efficient learning-based multi-label segmentation algorithm to achieve clinically acceptable instantaneous segmentation results for seed implantation planning [41].

Patient-Informed Treatment Decision-Making

AI has also been used in the informed decision-making process. Auffenberg et al. developed a web-based system that allows patients to input their own specific information to generate treatment options [42]. It was trained using random forest model processing data from 7543 prostate cancer patients covering a variety of different therapies (i.e., active surveillance, radical prostatectomy, radiation therapy, and androgen-deprivation therapy) for newly diagnosed prostate

Table 2 Skill assessment

Reference number	Reference	Application	Population	Performance
[29]	Hung et al.	Automated assessment of surgical performance and predicting clinical outcomes after RARP	78 RARP cases	Using the kinematic data from the surgical robot to represent surgeon performance, random forest-50 had 87.2% accuracy in predicting LOS (≤ 2 days vs > 2 days)
[30]	Hung et al.	Automated prediction of urinary continence recovery after RARP using kinematic data, followed by assessment of surgeon's historical patient outcomes	100 contemporary + 493 historical RARPs cases	DL model obtained a C-index of 0.6 and an MAE of 85.9. Surgeon group with better kinematic data performance had superior continence rates at 3 and 6 months postoperatively (47.5 vs 36.7%, $p = 0.034$; 68.3 vs 59.2%, $p = 0.047$)
[31]	Trinh et al.	Automated prediction of urinary continence recovery after RARP evaluated with patient and treatment factors, kinematic data, and technical skills of surgeon	115 RARP cases	Best models were produced by technical skills (C index: CoxPH – 0.695, DeepSurv: 0.708). Model performance for posterior/anterior VUA outperformed other steps (C index 0.543–0.592). Needle driving achieved top-performing models (C index 0.614–0.655)
[32]	Lee et al.	Using ML to predict positive surgical margins with surgeon performance (represented by kinematic data) and clinical factors	236 RARP cases	Model achieved AUC of 0.74. When assessing only clinical factors or kinematics, the model obtained AUC 0.72 and 0.64, respectively
[33]	Khalid et al.	Using ML models to assess surgical characteristics and performance	103 video clips of surgical procedures performed by 8 surgeons from JIGSAW dataset (benchmark models)	For surgical actions and surgical skill level of operators, deep machine learning yielded a mean precision of 0.97 and 0.77, and a mean recall of 0.98 and 0.78, respectively
[34]	Baghdadi et al.	Automated assessment of surgical performance in PLND	20 PLND video recordings	Model achieved an accuracy of 83.3% in predicting the expert-based PLACE scores
[36]	Zia et al.	Automating recognition of individual surgical tasks by their effect on task-based efficiency metrics	100 RARP cases	Model obtained a Jaccard index of 0.85. Median difference between human rater and ML algorithm < 200 s for most steps
[37]	Luongo et al.	Automating the recognition and categorization of suturing gestures for needle driving attempts	2395 live suturing videos for identification and 511 live suturing videos for classification	Models identified a gesture with AUC 0.88 and classified the gesture with AUC 0.87
[38]	Kiyasseh et al.	Using a unified deep learning framework to recognize surgical phase, classify surgical gestures, and assess surgical skills	86 surgical videos of the NS step of RARP by 15 surgeons, 60 videos of RARP from 8 surgeons, 27 videos of RAPN from 16 surgeons, 78 videos of VUA from 19 surgeons	Models achieved AUCs between 0.857 and 0.898 for phase recognition, between 0.680 and 0.899 for gesture classification, and between 0.797 and 0.880 for surgical skills assessment in multi-center data validation
[39•]	Ma et al.	Automated prediction of 1-year erectile function with surgical gesture sequences	80 nerve-sparing RARP across 2 international medical centers	Models performed better than traditional clinical features in predicting 1-year erectile function (team 1: AUC 0.77, 95% CI 0.73–0.81; team 2: AUC 0.68, 95% CI 0.66–0.70)

AUC area under the curve, DL deep learning, MAE mean absolute error, ML machine learning, NS nerve spare, PLND pelvic lymph node dissection, RAPN robot-assisted radical nephrectomy, RARP robot-assisted radical prostatectomy, VUA vesicourethral anastomosis

cancer patients. Both patients and doctors may benefit from using this tool to help them make informed decisions about the appropriate therapy that is best tailored for each individual patient.

Prognostics

Survival/Mortality Prediction

ML models are capable of rapidly processing large sums of clinical data which allows for increased prognostic capabilities in prostate cancer. Bibault et al. used an ML model that could predict the risk of prostate cancer mortality within 10 years of diagnosis and consisted of 30 clinical features with an accuracy of 0.98. Gleason score, PSA at diagnosis, and age had the largest impact on the model's prediction [43].

With advances in precision oncology, there is an increasing trend toward personalizing the management of diseases to better fit the individual patient. Koo et al. created an online support tool using a long short-term memory ANN model to predict survival outcomes based on initial treatment modalities using data from 7267 patients which provided accurate, individualized survival outcomes at 5 and 10 years [44]. To better understand disparities among prostate cancer patients of different racial backgrounds, efforts are being made to investigate the importance of race and other nonbiological factors on prostate cancer-specific mortality. Hanson et al. used the SEER database and applied a random forest model to analyze different variables and interactions across 4 major categories of factors crucial to prostate cancer mortality, tumor characteristics, race, health care, and social factors [45]. Ultimately tumor characteristics at diagnosis were found to be the most important factor for PCa mortality. While race was also found to be a significant predictor of PCa mortality, health care and social factors had just as important implications on PCa mortality. Zhang et al. used a prognostic ML model to screen for DNA methylation of gene targets and identified FOXD1 as a therapeutic target for patients that have a poor prognosis. Due to the heterogeneity among patients, it is unlikely a singular model will encompass every aspect that contributes to mortality prediction [46]. Lee et al. developed a Survival Quilts model that predicts and compares predictions to other leading models to help improve its accuracy for personalized prognostics [47]. Similar calibration efforts may offer improvements in personalized prognostics.

Recurrence Prediction

In addition to mortality prognosis, recurrence risk prediction is investigated following radical prostatectomies. The recurrence in PCa patients after a radical prostatectomy

often couples with a higher mortality rate. ML algorithms can demonstrate accuracy at higher rates than the previously used predictive nomograms in predicting the recurrence of prostate cancer. Tan et al. trained 3 ML models that outperformed traditional nomograms in predicting biochemical recurrence at 1, 3, and 5 years with the best model reaching an AUC of 0.894 for 5-year recurrence. This provides an alternative to tailored care in multimodal therapy [48]. As seen in the ML involved in mortality prognosis, population-based variations suggest more aggressive tumors in African American prostate cancer patients. Bhargava et al. trained a random forest model (AAstro ML) with an African American-specific stromal signature that outperformed clinical standard Kattan and CAPRA-S nomograms with AUCs of 0.87 and 0.77 [49]. Biochemical factors have a considerable effect in causing recurrence in PCa patients. Machine learning has also shown that it can accurately predict the biochemical recurrence of PCa after robot-assisted prostatectomy using magnetic resonance imaging (MRI) quantitative features. This has significant implications in optimizing treatment such as neoadjuvant or adjuvant therapies for patients.

Conclusions

In conclusion, there is widespread potential of the implementation of AI in the field of urology, particularly in the diagnosis and treatment of prostate cancer. Many studies have shown that AI-powered systems can accurately detect prostate cancer and help predict patient outcomes, leading to higher potential for improved patient care. However, there are several limitations to the use of AI in medicine that must be considered. For example, AI systems rely on the quality and quantity of data that they are trained on and may not perform as well when applied to real-world situations that differ from the data that was used to initially develop and train these algorithms. In addition, there are concerns about the ethical implications of using AI in medical decision-making, as well as the potential for bias in the algorithms that drive these systems. While there are still limitations and challenges to the widespread adoption of AI in medicine, the available evidence suggests that AI has the potential to revolutionize the field of urology and improve patient outcomes in relation to prostate cancer. Further research is needed to fully understand the potential and limitations of AI in this field, and to develop strategies for implementing AI in a way that maximizes its benefits while minimizing potential risks. Ultimately, AI remains as a potential tool to be used by urologists or other specialists to help guide their clinical decision-making and has not yet reached a point that it can or should supplant trained clinical professionals.

Funding Open access funding provided by SCEL, Statewide California Electronic Library Consortium.

Compliance with Ethical Standards

Conflict of Interest Andrew J. Hung has financial disclosures with Intuitive Surgical, Inc. The other authors declare that they have no conflict of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

Papers of particular interest, published recently, have been highlighted as:

- Of importance

1. Cui M, Zhang DY. Artificial intelligence and computational pathology. *Lab Invest.* 2021;101(4):412–22. <https://doi.org/10.1038/s41374-020-00514-0>.
2. Stoyanova R, Takhar M, Tschudi Y, Ford JC, Solórzano G, Erho N, et al. Prostate cancer radiomics and the promise of radiogenomics. *Transl Cancer Res.* 2016;5(4):432–47. <https://doi.org/10.21037/tcr.2016.06.20>.
3. Antonelli M, Johnston EW, Dikaios N, Cheung KK, Sidhu HS, Appayya MB, et al. Machine learning classifiers can predict Gleason pattern 4 prostate cancer with greater accuracy than experienced radiologists. *Eur Radiol.* 2019;29(9):4754–64. <https://doi.org/10.1007/s00330-019-06244-2>.
4. Fehr D, Veeraraghavan H, Wibmer A, Gondo T, Matsumoto K, Vargas HA, et al. Automatic classification of prostate cancer Gleason scores from multiparametric magnetic resonance images. *Proc Natl Acad Sci USA.* 2015;112(46):E6265–73. <https://doi.org/10.1073/pnas.1505935112>.
5. Turkbey B, Rosenkrantz AB, Haider MA, Padhani AR, Villeirs G, Macura KJ, et al. Prostate imaging reporting and data system version 2.1: 2019 update of prostate imaging reporting and data system version 2. *Eur Urol.* 2019;76(3):340–51. <https://doi.org/10.1016/j.eururo.2019.02.033>.
6. • Schelb P, Kohl S, Radtke JP, Wiesenfarth M, Kickingereder P, Bickelhaupt S, et al. Classification of cancer at prostate MRI: deep learning versus clinical PI-RADS assessment. *Radiology.* 2019;293(3):607–17. <https://doi.org/10.1148/radiol.2019190938>. **AI algorithm (UNET) achieved similar performance to PI-RADS assessment system in assessment of prostate cancer.**
7. Hectors SJ, Chen C, Chen J, Wang J, Gordon S, Yu M, et al. Magnetic resonance imaging radiomics-based machine learning prediction of clinically significant prostate cancer in equivocal PI-RADS 3 lesions. *J Magn Reson Imaging.* 2021;54(5):1466–73. <https://doi.org/10.1002/jmri.27692>.
8. Min X, Li M, Dong D, Feng Z, Zhang P, Ke Z, et al. Multiparametric MRI-based radiomics signature for discriminating between clinically significant and insignificant prostate cancer: cross-validation of a machine learning method. *Eur J Radiol.* 2019;115:16–21. <https://doi.org/10.1016/j.ejrad.2019.03.010>.
9. Woźnicki P, Westhoff N, Huber T, Riffel P, Froelich MF, Gresser E, et al. Multiparametric MRI for prostate cancer characterization: combined use of radiomics model with PI-RADS and clinical parameters. *Cancers (Basel).* 2020;12(7). <https://doi.org/10.3390/cancers12071767>.
10. Winkel DJ, Wetterauer C, Matthias MO, Lou B, Shi B, Kamen A, et al. Autonomous detection and classification of PI-RADS lesions in an MRI screening population incorporating multi-center-labeled deep learning and biparametric imaging: proof of concept. *Diagnostics (Basel).* 2020;10(11). <https://doi.org/10.3390/diagnostics10110951>.
11. Varghese B, Chen F, Hwang D, Palmer SL, De Castro Abreu AL, Ukimura O, et al. Objective risk stratification of prostate cancer using machine learning and radiomics applied to multiparametric magnetic resonance images. *Sci Rep.* 2019;9(1):1570. <https://doi.org/10.1038/s41598-018-38381-x>.
12. Cysouw MCF, Jansen BHE, van de Brug T, Oprea-Lager DE, Pfaehler E, de Vries BM, et al. Machine learning-based analysis of [(18)F]DCFPyL PET radiomics for risk stratification in primary prostate cancer. *Eur J Nucl Med Mol Imaging.* 2021;48(2):340–9. <https://doi.org/10.1007/s00259-020-04971-z>.
13. Papp L, Spielvogel CP, Grubmüller B, Grahovac M, Krajnc D, Ecsedi B, et al. Supervised machine learning enables non-invasive lesion characterization in primary prostate cancer with [(68)Ga]Ga-PSMA-11 PET/MRI. *Eur J Nucl Med Mol Imaging.* 2021;48(6):1795–805. <https://doi.org/10.1007/s00259-020-05140-y>.
14. Akatsuka J, Numata Y, Morikawa H, Sekine T, Kayama S, Mikami H, et al. A data-driven ultrasound approach discriminates pathological high grade prostate cancer. *Sci Rep.* 2022;12(1):860. <https://doi.org/10.1038/s41598-022-04951-3>.
15. Wildeboer RR, Mannaerts CK, van Sloun RJG, Budäus L, Tilki D, Wijkstra H, et al. Automated multiparametric localization of prostate cancer based on B-mode, shear-wave elastography, and contrast-enhanced ultrasound radiomics. *Eur Radiol.* 2020;30(2):806–15. <https://doi.org/10.1007/s00330-019-06436-w>.
16. Khosravi P, Lysandrou M, Eljalby M, Li Q, Kazemi E, Zisimopoulos P, et al. A deep learning approach to diagnostic classification of prostate cancer using pathology-radiology fusion. *J Magn Reson Imaging.* 2021;54(2):462–71. <https://doi.org/10.1002/jmri.27599>.
17. Schuettfort VM, Pradere B, Rink M, Comperat E, Shariat SF. Pathomics in urology. *Curr Opin Urol.* 2020;30(6):823–31. <https://doi.org/10.1097/mou.0000000000000813>.
18. Arvaniti E, Fricker KS, Moret M, Rupp N, Hermanns T, Fankhauser C, et al. Automated Gleason grading of prostate cancer tissue microarrays via deep learning. *Sci Rep.* 2018;8(1):12054. <https://doi.org/10.1038/s41598-018-30535-1>.
19. • Bulten W, Pinckaers H, van Boven H, Vink R, de Bel T, van Ginneken B, et al. Automated deep-learning system for Gleason grading of prostate cancer using biopsies: a diagnostic study. *Lancet Oncol.* 2020;21(2):233–41. [https://doi.org/10.1016/s1470-2045\(19\)30739-9](https://doi.org/10.1016/s1470-2045(19)30739-9). **AI/ML algorithm achieved similar performance to pathologists in determination of Gleason grades for prostate cancer biopsies.**
20. Kott O, Linsley D, Amin A, Karagounis A, Jeffers C, Golijanin D, et al. Development of a deep learning algorithm for the histopathologic diagnosis and gleason grading of prostate cancer biopsies: a pilot study. *Eur Urol Focus.* 2021;7(2):347–51. <https://doi.org/10.1016/j.euf.2019.11.003>.

21. Lucas M, Jansen I, Savci-Heijink CD, Meijer SL, de Boer OJ, van Leeuwen TG, et al. Deep learning for automatic Gleason pattern classification for grade group determination of prostate biopsies. *Virchows Arch.* 2019;475(1):77–83. <https://doi.org/10.1007/s00428-019-02577-x>.
22. Marginean F, Arvidsson I, Simoulis A, Christian Overgaard N, Åström K, Heyden A, et al. An artificial intelligence-based support tool for automation and standardisation of Gleason grading in prostate biopsies. *Eur Urol Focus.* 2021;7(5):995–1001. <https://doi.org/10.1016/j.euf.2020.11.001>.
23. Xie W, Reder NP, Koyuncu C, Leo P, Hawley S, Huang H, et al. Prostate cancer risk stratification via nondestructive 3D pathology with deep learning-assisted gland analysis. *Cancer Res.* 2022;82(2):334–45. <https://doi.org/10.1158/0008-5472.can-21-2843>.
24. da Silva LM, Pereira EM, Salles PG, Godrich R, Ceballos R, Kunz JD, et al. Independent real-world application of a clinical-grade automated prostate cancer detection system. *J Pathol.* 2021;254(2):147–58. <https://doi.org/10.1002/path.5662>.
25. Ström P, Kartasalo K, Olsson H, Solorzano L, Delahunt B, Berney DM, et al. Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. *Lancet Oncol.* 2020;21(2):222–32. [https://doi.org/10.1016/s1470-2045\(19\)30738-7](https://doi.org/10.1016/s1470-2045(19)30738-7).
26. Nir G, Karimi D, Goldenberg SL, Fazli L, Skinnider BF, Tavassoli P, et al. Comparison of artificial intelligence techniques to evaluate performance of a classifier for automatic grading of prostate cancer from digitized histopathologic images. *JAMA Netw Open.* 2019;2(3):e190442. <https://doi.org/10.1001/jamanetworkopen.2019.0442>.
27. Ma R, Reddy S, Vanstrum EB, Hung AJ. Innovations in urologic surgical training. *Curr Urol Rep.* 2021;22(4):26. <https://doi.org/10.1007/s11934-021-01043-z>.
28. Ma R, Collins JW, Hung AJ. The role of artificial intelligence and machine learning in surgery. In: Wiklund P, Mottrie A, Gundeti MS, Patel V, editors. *Robotic urologic surgery*. Cham: Springer International Publishing; 2022. p. 79–89.
29. Hung AJ, Chen J, Che Z, Nilanon T, Jarc A, Titus M, et al. Utilizing machine learning and automated performance metrics to evaluate robot-assisted radical prostatectomy performance and predict outcomes. *J Endourol.* 2018;32(5):438–44. <https://doi.org/10.1089/end.2018.0035>.
30. Hung AJ, Chen J, Ghodoussipour S, Oh PJ, Liu Z, Nguyen J, et al. A deep-learning model using automated performance metrics and clinical features to predict urinary continence recovery after robot-assisted radical prostatectomy. *BJU Int.* 2019;124(3):487–95. <https://doi.org/10.1111/bju.14735>.
31. Trinh L, Mingo S, Vanstrum EB, Sanford DI, Aastha, Ma R, et al. Survival analysis using surgeon skill metrics and patient factors to predict urinary continence recovery after robot-assisted radical prostatectomy. *Eur Urol Focus.* 2022;8(2):623–30. <https://doi.org/10.1016/j.euf.2021.04.001>.
32. Lee RS, Ma R, Pham S, Maya-Silva J, Nguyen JH, Aron M, et al. Machine learning to delineate surgeon and clinical factors that anticipate positive surgical margins after robot-assisted radical prostatectomy. *J Endourol.* 2022;36(9):1192–8. <https://doi.org/10.1089/end.2021.0890>.
33. Khalid S, Goldenberg M, Grantcharov T, Taati B, Rudzicz F. Evaluation of deep learning models for identifying surgical actions and measuring performance. *JAMA Netw Open.* 2020;3(3):e201664. <https://doi.org/10.1001/jamanetworkopen.2020.1664>.
34. Baghdadi A, Hussein AA, Ahmed Y, Cavuoto LA, Guru KA. A computer vision technique for automated assessment of surgical performance using surgeons' console-feed videos. *Int J Comput Assist Radiol Surg.* 2019;14(4):697–707. <https://doi.org/10.1007/s11548-018-1881-9>.
35. Hung AJ, Liu Y, Anandkumar A. Deep learning to automate technical skills assessment in robotic surgery. *JAMA Surg.* 2021;156(11):1059–60. <https://doi.org/10.1001/jamasurg.2021.3651>. **Machine Learning models can be potentially used to assess robotic surgeon technical skill.**
36. Zia A, Guo L, Zhou L, Essa I, Jarc A. Novel evaluation of surgical activity recognition models using task-based efficiency metrics. *Int J Comput Assist Radiol Surg.* 2019;14(12):2155–63. <https://doi.org/10.1007/s11548-019-02025-w>.
37. Luongo F, Hakim R, Nguyen JH, Anandkumar A, Hung AJ. Deep learning-based computer vision to recognize and classify suturing gestures in robot-assisted surgery. *Surgery.* 2021;169(5):1240–4. <https://doi.org/10.1016/j.surg.2020.08.016>.
38. Kiyasseh D, Ma R, Haque TF, et al. Decoding surgeon activity from surgical videos with a unified artificial intelligence system. *Nature Biomedical Engineering.* In Press.
39. Ma R, Ramaswamy A, Xu J, Trinh L, Kiyasseh D, Chu TN, et al. Surgical gestures as a method to quantify surgical performance and predict patient outcomes. *NPJ Digit Med.* 2022;5(1):187. <https://doi.org/10.1038/s41746-022-00738-y>. **Machine learning models were able to break down surgical gesture sequences during robotic assisted radical prostatectomies and effectively predict postoperative 1 year erectile function recovery.**
40. McIntosh C, Conroy L, Tjong MC, Craig T, Bayley A, Catton C, et al. Clinical integration of machine learning for curative-intent radiation treatment of patients with prostate cancer. *Nat Med.* 2021;27(6):999–1005. <https://doi.org/10.1038/s41591-021-01359-w>.
41. Nouranian S, Ramezani M, Spadinger I, Morris WJ, Salcudean SE, Abolmaesumi P. Learning-based multi-label segmentation of transrectal ultrasound images for prostate brachytherapy. *IEEE Trans Med Imaging.* 2016;35(3):921–32. <https://doi.org/10.1109/tmi.2015.2502540>.
42. Auffenberg GB, Ghani KR, Ramani S, Usoro E, Denton B, Rogers C, et al. askMUSIC: leveraging a clinical registry to develop a new machine learning model to inform patients of prostate cancer treatments chosen by similar men. *Eur Urol.* 2019;75(6):901–7. <https://doi.org/10.1016/j.eururo.2018.09.050>.
43. Bibault JE, Hancock S, Buyyounouski MK, Bagshaw H, Leppert JT, Liao JC, et al. Development and validation of an interpretable artificial intelligence model to predict 10-year prostate cancer mortality. *Cancers (Basel).* 2021;13(12). <https://doi.org/10.3390/cancers13123064>.
44. Koo KC, Lee KS, Kim S, Min C, Min GR, Lee YH, et al. Long short-term memory artificial neural network model for prediction of prostate cancer survival outcomes according to initial treatment strategy: development of an online decision-making support system. *World J Urol.* 2020;38(10):2469–76. <https://doi.org/10.1007/s00345-020-03080-8>.
45. Hanson HA, Martin C, O'Neil B, Leiser CL, Mayer EN, Smith KR, et al. The relative importance of race compared to health care and social factors in predicting prostate cancer mortality: a random forest approach. *J Urol.* 2019;202(6):1209–16. <https://doi.org/10.1097/ju.0000000000000416>.
46. Zhang E, Hou X, Hou B, Zhang M, Song Y. A risk prediction model of DNA methylation improves prognosis evaluation and indicates gene targets in prostate cancer. *Epigenomics.* 2020;12(4):333–52. <https://doi.org/10.2217/epi-2019-0349>.
47. Lee C, Light A, Alaa A, Thurtle D, van der Schaar M, Gnanapragasam VJ. Application of a novel machine learning framework for predicting non-metastatic prostate cancer-specific mortality in men using the Surveillance, Epidemiology, and End Results (SEER) database.e158–e65.
48. Tan YG, Fang AHS, Lim JKS, Khalid F, Chen K, Ho HSS, et al. Incorporating artificial intelligence in urology: supervised machine learning algorithms demonstrate comparative advantage over

- nomograms in predicting biochemical recurrence after prostatectomy. *Prostate*. 2022;82(3):298–305. <https://doi.org/10.1002/pros.24272>.
49. Bhargava HK, Leo P, Elliott R, Janowczyk A, Whitney J, Gupta S, et al. Computationally derived image signature of stromal morphology is prognostic of prostate cancer recurrence following prostatectomy in African American patients. *Clin Cancer Res*. 2020;26(8):1915–23. <https://doi.org/10.1158/1078-0432.ccr-19-2659>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.