Chinese Medical Journal®

Artificial intelligence in prostate cancer

Wei Li[1], Ruoyu Hu[1], Quan Zhang[1], Zhangsheng Yu[2,3], Longxin Deng[4], Xinhao Zhu[1], Yujia Xia[2], Zijian Song[1], Alessia Cimadamore[5], Fei Chen[6], Antonio Lopez-Beltran[7], Rodolfo Montironi[8], Liang Cheng[9], Rui Chen[1]

[1]Department of Urology, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200127, China;
[2]Department of Bioinformatics and Biostatistics, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China;
[3]Clinical Research Institute, Shanghai Jiao Tong University School of Medicine, Shanghai 200025 China;
[4]Department of Urology, The First Affiliated Hospital of Naval Medical University (Second Military Medical University), Shanghai 200433, China;
[5]Institute of Pathological Anatomy, Department of Medicine, University of Udine, Udine, Italy;
[6]Department of Pathology and Laboratory Medicine, New York University Grossman School of Medicine and NYU Langone Health, New York, NY, USA;
[7]Department of Surgery, Cordoba University Medical School, Cordoba, Spain;
[8]Molecular Medicine and Cell Therapy Foundation, Polytechnic University of the Marche Region, Ancona, Italy;
[9]Department of Pathology and Laboratory Medicine, Department of Surgery (Urology), Brown University Warren Alpert Medical School, the Legorreta Cancer Center at Brown University, and Brown University Health, Providence, RI, USA.

Abstract
Prostate cancer (PCa) ranks as the second most prevalent malignancy among men worldwide. Early diagnosis, personalized treatment, and prognosis prediction of PCa play a crucial role in improving patients' survival rates. The advancement of artificial intelligence (AI), particularly the utilization of deep learning (DL) algorithms, has brought about substantial progress in assisting the diagnosis, treatment, and prognosis prediction of PCa. The introduction of the foundation model has revolutionized the application of AI in medical treatment and facilitated its integration into clinical practice. This review emphasizes the clinical application of AI in PCa by discussing recent advancements from both pathological and imaging perspectives. Furthermore, it explores the current challenges faced by AI in clinical applications while also considering future developments, aiming to provide a valuable point of reference for the integration of AI and clinical applications.
Keywords: Prostate cancer; Artificial intelligence; Whole-slide image; Pathology; Foundation model; Imaging

Introduction

Prostate cancer (PCa) is the second most prevalent malignant tumor among men worldwide.[1] The latest data show that PCa is the fifth leading cause of cancer death in men. In 2020, there were nearly 1.5 million (7.3% of all sites) new cases of PCa and 397,000 (4.1% of all sites) deaths globally.[2] The incidence of PCa is nearly three times higher in transitioned countries than in transitioning countries; however, the difference in mortality rates is not large, suggesting differences in early screening and treatment across regions.[2,3]

Patients with PCa typically exhibit the absence of characteristic clinical symptoms during the early stages of the disease. Current strategies for PCa screening and diagnosis rely on prostate-specific antigen (PSA), digital rectal examination (DRE), ultrasound, magnetic resonance imaging (MRI), and biopsy. Studies have shown that PSA-based screening significantly reduces PCa-specific mortality (relative risk [RR] = 0.73) and metastatic risk (RR = 0.67) in men aged 55–69 years. However, overdiagnosis rates due to PSA-only screening may reach 20–50%,[4] necessitating the integration of clinical parameters such as PSA density, free-to-total PSA ratio, and multiparametric MRI (mpMRI) to minimize unnecessary biopsies.[5] The objective and efficient identification of PCa lesions, assessment of the malignancy of PCa, prediction of patient prognosis, and guidance of personalized treatment are essential for reducing mortality rates and improving quality of life.[1] Pathology and imaging play a pivotal role in achieving these objectives. Due to the demanding clinical tasks and varying levels of expertise among doctors, subjectivity and inconsistency are prevalent, posing significant challenges for precision medicine in PCa patients.[6,7]

Access this article online

Quick Response Code:

Website: www.cmj.org

DOI: 10.1097/CM9.0000000000003689

Wei Li, Ruoyu Hu, and Quan Zhang contributed equally to this work.

Correspondence to: Liang Cheng, Department of Pathology and Laboratory Medicine, Brown University Warren Alpert Medical School, Providence, RI, USA
E-Mail: liang_cheng@yahoo.com;
Rui Chen, Department of Urology, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China. No. 160 Pujian Road, Shanghai 200433, China
E-Mail: drchenrui@foxmail.com

Copyright © 2025 The Chinese Medical Association, produced by Wolters Kluwer, Inc. under the CC-BY-NC-ND license. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Chinese Medical Journal 2025;138(15)
Received: 16-12-2024; **Online:** 09-07-2025 **Edited by:** Yuanyuan Ji

The application of AI in the medical field is gradually becoming widespread and mature.[8] The Asian Prostate Cancer Artificial intelligence (APCA) model, developed by Song *et al*[9] through the integration of multimodal clinical parameters, targets high-grade PCa (HGPCa) and optimizes screening protocols. Notably, it reduces unnecessary biopsies by 38.4% while maintaining a 10% missed diagnosis rate for HGPCa cases, as validated in a multicenter Asian cohort spanning 8 referral centers (area under the curve [AUC] = 0.76). Moreover, based on whole-slide image (WSI), MRI, and positron emission tomography/computed tomography (PET/CT), AI can assist the management of PCa with high consistency and repeatability, which reduces the clinical workload and improves the efficiency of clinicians [Figure 1]. Due to advancements in algorithms and the continuous accumulation of data, the foundation model has emerged. Deep learning (DL) algorithm provides the foundation model with strong learning ability and expression ability. It can better adapt to new tasks through self-supervised learning pre-training on large multimodel unannotated data and fine-tuning on small, annotated data sets of downstream tasks. With superior performance and generalization abilities, it holds significant potential for clinical applications in PCa.[10]

Previously, we have summarized the application of AI in the pathology of PCa.[7] In this review, we focus on AI-assisted personalized management and precision medicine for PCa patients from the aspects of pathology and imaging; summarize the cutting-edge advances of AI in PCa diagnosis, treatment, patient prognosis prediction, and molecular subtyping; and discuss the application of foundation model in PCa. At last, we discuss the challenges and future of its clinical application.

Application of AI in PCa Pathology

Application of AI in PCa pathological diagnosis: Enhancing accuracy and efficiency

The Gleason grading system has long been used to assess the severity of PCa, but the subjectivity of different pathologists can lead to varying results and a lack of consistency. To clarify the impact caused by subjectivity, Ozkan *et al*[11] asked two pathologists to score 407 pathology slides of prostate needle biopsies from 34 consecutive patients with

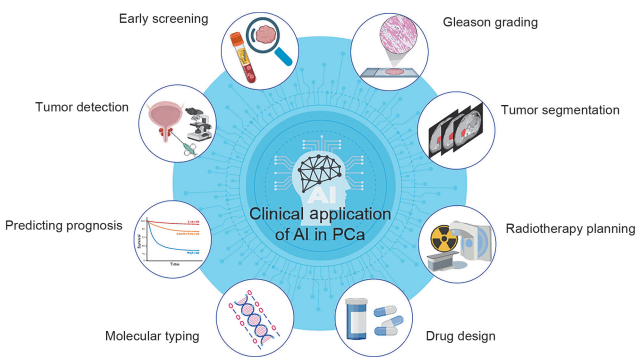

Figure 1: Clinical application of AI in PCa. AI: Artificial intelligence; PCa: Prostate cancer.

PCa by using the Gleason grading system. Then they used the kappa coefficient test, which revealed a consistency of only 51.7% (k = 0.39), to confirm the existence of errors due to subjectivity. Subsequently, they did further analysis of the lack of scoring consistency and found that in addition to subjectivity, differences in education and training levels, difficulty in distinguishing low-grade tumors, and experience levels are all potential factors. The inconsistency of Gleason grading highlights the need for a more objective diagnostic tool.

The emergence of AI has provided an opportunity to improve the accuracy and efficiency of pathological diagnosis.[12–14] Currently, many AI applications have been developed for clinical use in PCa [Table 1]. The Galen Prostate,[15] an auxiliary diagnostic tool based on multilayer convolutional neural networks (CNNs), was trained on 1,357,480 image patches from 549 hematoxylin and eosin (H&E) stained slides and achieved AUCs of 0.997 and 0.991 for internal test set and external validation set of PCa detection, respectively. In the external validation set, the AUC for distinguishing between low-grade (Gleason score 6 or atypical small acinar proliferation) and high-grade (Gleason score 7–10) cancer diagnosis was 0.941 and the AUC for detecting Gleason pattern 5 was 0.971 (0.943–0.998). There was also consistency between the percentage of cancer in core needle biopsy material calculated by pathologists and the algorithm (r = 0.882).

Furthermore, a comparison between a DL model for Gleason grading of prostate with 19 general pathologists and experts demonstrated that AI has the ability to eliminate errors caused by subjectivity.[16] In the external validation set for cancer detection, the model's consistency with experts was 94.3% and that of general pathologists with experts was 94.7%. In the external validation set for grading tumor-containing biopsy specimens, the model's consistency with experts was 71.7% and that of general pathologists with experts was 58.0%. This model showed higher accuracy in Gleason grading than general pathologists, avoiding the omission of minor lesions, proving the potential utility of AI as a decision support tool, and the prospect of improving the quality of PCa grading.

Moreover, AI is continuously evolving in the field of diagnosis. The emergence of semiautomatic labeling technology has eliminated the need for pathologists to manually annotate.[17] The construction of weakly supervised DL models has avoided the high cost and time-consuming pixel-level manual annotation while ensuring diagnostic accuracy, which is currently a popular algorithm.[18] Accurate computational H&E staining of original unstained prostate biopsy images can shorten diagnostic time and achieve immediate diagnosis without the large amount of time traditionally spent on staining.[19]

AI in PCa treatment: Personalized medicine and therapeutic management

Currently, in addition to traditional surgical treatment, targeted drugs, endocrine therapy, and radiotherapy[20] have been proven valuable in the treatment of PCa and

Table 1: Studies assessing AI in pathology applications of PCa.

Studies	Data	Training methods	Results
Higher accuracy of PCa detection			
Pantanowitz *et al*[14] 2020	Training set: 549 slides Validation set: 391 slides Test set: 1627 slides	Convolutional neural network	AUC = 0.991
Ström *et al*[25] 2020	Training set: 6682 slides Validation set: 330 biopsies Test set: 1631 biopsies	Deep neural networks	AUC = 0.986
Campanella *et al*[17] 2019	Training set: 17,401 slides Validation set: 3729 biopsies Test set: 3729 biopsies	Multiple instance learning	AUC = 0.994
Rana *et al*[18] 2020	Training set: 74,000 patches Test set: 13,500 patches	Generative adversarial neural networks	SSIM = 0.902 PCC = 0.962 PSNR = 22.821dB
Individualized treatment of PCa			
Spratt *et al*[21] 2023	Training set: 5727 patients Test set: 1594 patients	Multimodal AI	The absolute benefit of short-term ADT was 10.5%
McIntosh *et al*[22] 2021	Training set: 99 plans Test set: 17/50/50 patients	Random forest	89% of ML-generated RT plans are considered clinically acceptable 72% of ML-generated RT plans were selected for treatment
PCa prognosis based on Gleason grading or personalized predictive models			
Lee *et al*[24] 2021	Training set: 110,043 patients Validation set: 27,511 patients Test set: 34,388 patients	Survival quilts	c-index = 0.829
Ström *et al*[26] 2020	Training set: 6682 slides Validation set: 330 biopsies Test set: 1631 biopsies	Deep neural networks	AUC = 0.986
Xie *et al*[27] 2022	Test set: 50 patients	Generative adversarial networks	AUC = 0.800
Wei *et al*[29] 2021	Training set: 2018 patients Test set: 866 patients	XGBoost	AUC = 0.883
Molecular subtypes prediction of PCa			
Harmon *et al*[36] 2021	Training set: 2048 cores Validation set: 224 cores Test set: 357 cores	Convolutional neural networks	AUC = 0.964
Dadhania *et al*[37] 2022	Training set: 261 cases Test set: 131 cases	Convolutional neural network	AUC >0.82

AI: Artificial intelligence; ADT: Androgen deprivation therapy; AUC: Area under the curve; ML: Machine learning; PCa: Prostate cancer; PCC: Pearson correlation coefficient; PSNR: Peak signal-to-noise ratio; RT: Radiotherapy; SSIM: Structural similarity; TCGA-PRAD: The cancer genome atlas prostate adenocarcinoma.

are widely used in clinical practice. To improve the effectiveness of treatment, clinicians have been committed to verifying the treatment effects of various therapies on different patients. Recent research has shown the guiding role of AI in PCa treatment, which can assist clinicians in personalized medicine for patients and also lay the foundation for the entire patient management process.

For patients with local PCa, benefits can be obtained through radiotherapy combined with androgen deprivation therapy (ADT), but ADT may have a negative impact on their quality of life, such as hot flashes, decreased libido, loss of muscle mass, erectile dysfunction, osteoporosis and so on.[21] Therefore, doctors and patients need to carefully weigh the potential benefits of ADT against these possible decreases in quality of life. To assist clinical doctors in making better decisions, there is a model[22] based on DL that can identify intermediate-risk PCa patients who may benefit from short-term ADT, thus guiding the use of ADT. The Multimodal Artificial Intelligence (MMAI) Prostate Prognostic Model outperforms existing prediction methods based on the National Comprehensive Cancer Network (NCCN) risk grouping, indicating its greater accuracy in identifying patients who may benefit from ADT and helping clinicians make more personalized treatment decisions for local PCa patients.

Radiotherapy is also one of the current research directions. A model[23] based on the random forest algorithm was developed for radiation therapy planning with treatment intent and conducted prospective deployment and evaluation in the clinical workflow. The results showed that 89% of machine learning (ML)-generated radiotherapy plans were considered clinically acceptable, and 72% were chosen for treatment, demonstrating the potential value of AI in guiding radiotherapy. Subsequently, Gao *et al*[24] developed an AI called the Virtual Treatment Planner (VTP), which can automatically adjust the relevant parameters in the treatment planning system to generate high-quality PCa stereotactic body radiotherapy treatment plans. It has now been used in the commercial treatment planning system of Varian Medical Systems.

AI in prognostic assessment of PCa: From Gleason grading to personalized predictive models

In terms of prognosis for PCa, Gleason grading has long been one of the recognized predictive indicators, but AI can now develop new predictive models using its unique algorithms. For example, Survival Quilts,[25] which is based on an ensemble of traditional and ML modeling techniques, served as a prognostic model for 10-year PCa-specific mortality in non-metastatic PCa patients and showed good discrimination, with a consistency index (c-index) of 0.829, and this discrimination can still be maintained in different age and racial stratifications.

Biochemical recurrence (BCR), as a major sign of PCa recurrence, is also a key indicator that needs to be monitored. Extracting 2D features is currently the mainstream research method; an AI model[26] based on it achieved an average pairwise kappa of 0.62 in Gleason grading, which is similar to that of pathologists. However, extracting 3D features has shown its potential in predicting BCR. A DL strategy[27] based on interpretable glandular features of image translation assistance segmentation (ITAS3D), extracting 3D rather than 2D glandular features for prediction. Comparing the two, the results showed AUC values of 0.80 and 0.65, respectively, proving better BCR prediction effects.

In addition, lymph node metastasis is also an important indicator affecting disease prognosis. A newly developed AI model for PCa lymph node metastasis detection, known as Prostate Cancer Lymph Node Metastasis Detector (ProCaLNMD),[28] has demonstrated remarkably high accuracy, with area under the receiver operating characteristic curve (AUROC) values ranging from 0.975 to 0.992 across multiple datasets. In addition, the XGBoost (XGB) model[29] showed the best predictive performance, with the highest net benefit compared to other algorithms, and an internal validation test AUC of 0.883. Besides, lymphatic micrometastasis poses difficulties in predicting the prognosis of patients due to the high rate of artificial misdiagnosis. However, The Lymph Node Metastases Diagnostic Model (LNMDM)[30] achieved an AUC of 0.922 in detecting micrometastasis in PCa lymph node metastasis, which demonstrates the excellent performance of the model in detecting lymph node metastasis in PCa, while also demonstrating the potential of AI in this field.

Currently, in terms of slice selection, most studies select multiple tissue slices for each patient,[15–19,22–27,29,31] which can ensure data volume while providing a more comprehensive tissue view, helping to detect small changes and rare events, and increasing the credibility of the results. However, this consumes a lot of time and money; a few studies choose a single tissue slice for each patient, which has the advantage of reducing costs and simplifying analysis, but the information from a single slice is limited, making it difficult to perform reliable statistical analysis and validation.

AI in molecular typing of PCa: New frontier of precision medicine and gene detection

Molecular typing of PCa is crucial for its precise treatment, prognosis assessment, and new drug development, aiding in personalized medicine and patient management. However, traditional gene analysis requires a lot of time and money and is not the preferred plan. Pathological characteristics are a major direction for predicting the progression of PCa, and humans can only extract a small amount of information from it. However, with technological advancements, AI is now able to judge molecular markers related to PCa from it.

Some genes can become targets for detection. The loss of PTEN (phosphatase and tensin homologue deleted on chromosome ten), as a poor prognostic factor,[32–35] can be predicted by a model based on DL and CNNs[36] with a sensitivity of 95.0% and a specificity of 98.4% in the internal cohort. In the external cohort, after fine-tuning, the sensitivity was 90.6%, and the specificity was 95.7%. In addition, for the most common E26 transformation-specific (ETS) gene fusion in PCa – TMPRSS2-ERG gene rearrangement, Dadhania *et al*[37] developed a model based on the CNNs with an AUC of 0.82–0.85 for predicting gene fusion. These successes have proven the potential of AI in gene detection.

In addition to known related genes, artificially developed molecular markers can also be the target of AI detection. A lncRNA marker related to mitochondrial autophagy has been developed by integrated ML technology,[38] which is significantly related to the RIG-I-like receptor signaling pathway, and the results showed an AUC greater than 0.6, which can be used to predict the progression of PCa.

Foundation models in PCa diagnosis and treatment: New breakthroughs

Traditional AI for specific tasks has achieved many results in the diagnosis, treatment, and prognosis of PCa,[39] but with the emergence of foundation models, it has provided new opportunities for the application of AI in the diagnosis and treatment of PCa [Figure 2]. The application of many large models in PCa also demonstrates their potential [Table 2].

Foundation models, as unique models in the field of AI, can handle more complex tasks. For example, the visual-language foundation model CONtrastive learning from Captions for Histopathology (CONCH) developed

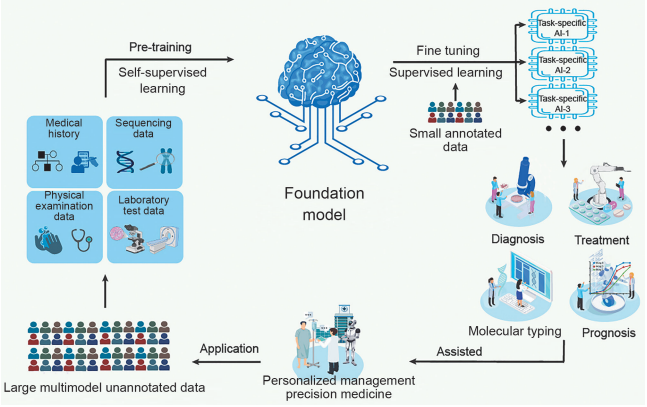

Figure 2: A prospective schematic layout of using foundation models for various downstream tasks. AI: Artificial intelligence.

by Lu *et al*[40] not only performed well in image classification tasks but also showed strong capabilities in image retrieval and few-shot learning tasks. At the same time, as one of the largest pre-trained datasets used in the field of histopathology to date, it can be directly applied to various downstream tasks, including image classification and segmentation, title generation, and image-to-text and text-to-image retrieval, without further supervision or fine-tuning.

In terms of pathological image data collection, most foundation models currently resample every part of one WSI, which may result in the loss of information.[41] To solve this problem, Prov-GigaPath[42] was developed, which provides a new approach. They proposed the GigaPath architecture, combined with the LongNet method, using multimodal pre-training can effectively handle whole gigapixel WSIs, achieving state-of-the-art performance in 25 out of 26 different pathological tasks.

In cancer prediction, foundation models have made breakthroughs in the diagnosis of cancer mutations and rare cancers. The Clinical Histopathology Imaging Evaluation Foundation (CHIEF)[43] performed well in multiple pathological assessment tasks such as cancer detection, tumor origin recognition, molecular profiling, and prognosis prediction. Through pre-learning, it also can predict the mutation status of 16 types of cancers, including PCa, with an average c-index of 0.74 in various cancer types. The Virchow developed by Vorontsov *et al*[10] showed its generalization ability to rare data in the detection of PCa rare cancers, including neuroendocrine tumors, atrophic tumors, small lymphocytic lymphoma, foam cell carcinoma, and follicular lymphoma, with an AUC of 0.937.

Although foundation models have made breakthroughs in various aspects, data acquisition and the interpretability of foundation models have always been challenges for all pathological large models. However, recent research has found potential solutions to these two problems. The Pathology Language-Image Pretraining (PLIP) model[44] learned from a large amount of public data from Twitter and achieved excellent performance on many external data with an AUC of 0.856–0.972. This concept of knowledge sharing may help with the current shortage of pathological data. In addition, the P-NET,[45] based on bioinformatics knowledge, used sparse connections instead of fully connected dense networks. With researchers able to view the internal structure and node activation of the model

Table 2: Studies assessing foundation model in pathology applications of PCa.

Studies	Pretraining dataset	Name	Innovations
Lu *et al*[40] 2024	1,170,000 image–caption pairs	CONCH	The pretraining data for CONCH is one of the largest pre-training datasets used in the field of histopathology to date. It has good zero-shot transfer capability, efficient few-shot learning, and wide downstream task generalization.
Xu *et al*[42] 2024	1.3 billion 256 × 256 pathology image tiles	Prov-GigaPath	The Prov-GigaPath model uses its innovative GigaPath architecture and LongNet method to achieve in-depth learning of entire slides, capturing local and global pathological patterns, highlighting its transformative potential in the field of digital pathology.
Wang *et al* [43] 2024	44 terabytes of high-resolution pathology imaging datasets+ 60,530 whole-slide images spanning 19 anatomical sites	CHIEF	The CHIEF model is used to extract features from pathological images for systematic cancer assessment. By combining unsupervised pretraining with weakly supervised pretraining, it is capable of predicting the mutational status of 16 types of cancers, including PCa.
Vorontsov *et al*[10] 2024	1,488,550 H&E-stained WSIs	Virchow	Virchow has been evaluated for its ability to predict biomarkers and identify cells and has demonstrated the capability of large models in pan-cancer detection, including both common and rare cancers, particularly in detecting rare cancer variants.
Huang *et al*[44] 2023	608,562 image data + 425,828 text data	PLIP	PLIP leveraged publicly available medical Twitter data to create the OpenPath dataset and developed a multimodal AI model capable of understanding pathological images and their natural language descriptions, which greatly facilitated knowledge sharing and education.

AI: Artificial intelligence; H&E: Hematoxylin and eosin; PCa: Prostate cancer; WSI: Whole-slide images.

through visualization technology, full model interpretability was achieved, which means it can reveal the molecular drivers affecting treatment resistance, which is crucial for the selection of clinical treatment targets.

The foundation model research currently underway has shown its potential in PCa pathology, first by improving diagnostic efficiency and accuracy, providing preliminary judgments to help clinicians make efficient diagnoses; second, in areas lacking medical resources, foundation models can serve as an alternative to obtain more reliable diagnoses; in addition, the unique multimodal applications of foundation models allow them to assist in disease research in addition to diagnosis, prognosis, and treatment. With further technological development, foundation models are expected to play an increasingly important role in PCa pathology.

Application of AI in PCa Imaging

The mpMRI relies on three sequences: T2-weighted imaging (T2WI), diffusion-weighted imaging (DWI), and dynamic contrast-enhanced imaging (DCE). This complexity poses challenges for practitioners to master and generalize.[46,47] Similar limitations persist in biparametric MRI (bpMRI), despite its reduced number of sequences. With the assistance of AI, clinical practitioners can have easier access to interpret the images. For the time being, Prostate Imaging Reporting and Data System (PI-RADS) is still the most popular standard to evaluate patients' MRI and predict the possibility of malignant tumor. According to a cohort study,[48] only 67% of 715 enrolled participants (modified PI-RADS ≥3) turned out to be real PCa while 28% of low-risk ones are diagnosed PCa through biopsy. Obviously, the accuracy of current mainstream diagnostic criteria fails to meet clinical demands for diagnostic precision, potentially leading to substantial misdiagnosis rates and delays in therapeutic intervention and new standards must be established as soon as possible.

Apart from MRI, prostate-specific membrane antigen PET/CT (PSMA PET/CT) is much more accurate in identifying metastases in the prostate gland as a whole. The maximum standardized uptake value (SUVmax) is a critical semi-quantitative parameter in PSMA PET/CT imaging. It serves to assess the tracer uptake intensity of neoplastic lesions for the administered radiopharmaceutical, reflecting the metabolic activity of tumor cells through quantification of radioligand accumulation. However, the cut-off value of SUVmax to diagnose PCa varies a lot from 3.2 to 9.1,[49–51] which presents relatively poorer performance (accuracy = 87.7% in PI-RADS grade 1 patients) compared with the DL models.[52] With the use of DL, the diagnosis is needless to rely on current clinical benchmarks, but on the features extracted by models themselves. A study[53] has already demonstrated that conventional neural network-based trained models perform at the same level as specialized nuclear medicine physicians in determining whether lymph nodes are accumulated on PSMA PET-CT. A meta-analysis[54] evaluating the performance of AI methods in identifying tumor metastases from radiological imaging showed that the combined AUC value of the AI methods was 0.90 in different cancers including PCa.

Given the reliability and great clinical potential of imaging technologies, corresponding surgical guidelines, objective imaging assessment criteria, and case reports have been established. In today's healthcare environment, it is a challenge for professional radiologists to detect and locate all cancer lesions accurately, and the task of the tumor contour is also time-consuming. The combination of such technologies is becoming more and more necessary. For example, computer vision algorithms can assist radiologists in detecting and accurately quantifying metastases, reducing manual labor and increasing accuracy, thus streamlining the radiologist's clinical workflow and improving work efficiency. More applications of AI will be discussed later [Table 3].

Application of AI in PCa radiological diagnosis: Useful assistance to clinical workflow

Imaging is an important part of the diagnosis, not only to non-invasively screen out a subset of low-risk patients but also to guide subsequent biopsy procedures, providing more accurate localization and avoiding unnecessary injuries caused by excessive needle biopsy. According to the recommendations of the European Association of Urology (EAU), mpMRI should be performed before prostate biopsy in first-treatment men, and if positive (PI-RADS ≥3), a targeted biopsy (index lesion) and systematic sampling of the biopsy should follow.[55,56]

Using mpMRI for PCa lesion region delineation can greatly help doctors distinguish benign prostatic hyperplasia and malignant lesions and also clarify whether there is extracapsular extension (ECE), which is of positive significance for prognosis prediction such as BCR. Currently, ML can differentiate lesions between benign and malignant prostate by unsupervised learning through Fuzzy C-Means clustering algorithm. Rundo *et al*[57] used this algorithm to segment the prostate on T1-weighted imaging (T1WI) and T2WI and evaluated 21 patients, yielding a mean Dice score of 0.91, demonstrating that the technique is capable of segmenting and evaluating the volume of the prostate gland with a high degree of precision. Tian *et al*[58] developed a model based on CNN trained on T2WI of 140 patients and achieved a Dice score of 0.85. Another CNN-based model developed by Karimi *et al*[59] on a limited number of T2WI were trained and supplemented with data augmentation, achieving a Dice score of 0.88. The high Dice scores of these two studies suggest that AI can achieve high precision in the segmentation of the prostate with commonly used algorithms simply based on MRI images.

PSMA PET/CT also can provide an accurate diagnosis of PCa at first diagnosis and more information than conventional imaging for clinical practice, reducing patients' cost.[60] An automated PCa lesion characterization model[61] used ^{68}Ga-PSMA-11 PET/CT of 193 metastatic castration-resistant PCa (mCRPC) patients based on a triple combination 2.5D U-Net framework. When compared with manual segmentation, the model had high sensitivity

Table 3: Examples of applications of AI in PCa imaging.

Studies	Data	Framework	Result	Objective
Diagnosis				
Zhao *et al*[61] 2019*	Training set: 130 patients Test set: 63 patients Validation set: fivefold cross validation	2.5D U-net	Bone lesion detection precision: 0.99 Lymph node lesion detection precision: 0.94 Local lesion detection precision: 0.79	Automated detection and segmentation of PCa lesions on ^{68}Ga-PSMA-11 PET/CT images to improve treatment planning for PSMA-targeted radioligand therapy.
Trägårdh *et al*[62] 2022*	Training set: 125 patients Validation set: 36 patients Test set: 50 patients	CNN	Sensitivity: 82%	Detection of pelvic lymph node metastases in patients with high-risk PCa on ^{18}F PSMA PET/CT.
Rundo *et al*[57] 2017	Training set: 21 patients	Fuzzy C-Means clustering algorithm	Volume similarity: 97.14% DSC: 90.77 ± 1.75	Automatically segmenting the prostate gland effectively.
Tian *et al*[58] 2018	Training set: 112 patients Test set: 28 patients	CNN	Mean DSC: 85.0 ± 3.8%	Automated segmentation of prostate MRI images to improve efficiency and accuracy in PCa diagnosis and treatment.
Karimi *et al*[59] 2018	Training set: 49 pieces T2WI images Test set: 26 images	CNN	Dice score: 0.88	Improving automatic segmentation accuracy of prostate MRI images by combining statistical shape modeling and DL.
Soerensen *et al*[63] 2021	Training set: 805 patients Retrospective testing set: 100 patients Prospective test set: 11 patients	U-Net	Mean DSC: 0.90	Reducing prostate segmentation time for radiologists and urologists in targeted biopsies and potentially improving biopsy accuracy through automated DL models.
Balagopal *et al*[64] 2021	Training set: 255 patients Validation set: 35 patients Test set: 50 patients	2D U-Net	Mean DSC: 0.87	Presenting a deep learning-based framework for segmenting invisible CTV with estimated uncertainties for post-operative PCa radiotherapy.
Lay *et al*[65] 2017	Training set: 112 patients Test set: 112 patients Internal validation: twofold cross-validation	Random Forest	AUC: 0.93	Quickly interpret large amounts of image data through mpMRI images and assist less experienced radiologists in biopsy by detecting the lesion.
Mehralivand *et al*[66] 2022	Training set: 943 patients Validation set: fivefold cross-validation Test set: 150 patients	U-Net, AH-Net	Sensitivity: 56.1% PPV: 62.7%	Developing and testing a cascaded DL detection and classification system trained on biparametric prostate MRI using PI-RADS for assisting radiologists during prostate MRI read out.
Hou *et al*[67] 2021	Training: 596 External validation: 103 Internal validation: 150	ResNeXt	AUC (training): 0.857 AUC (internal): 0.807 AUC (external): 0.728	Detection extracapsular extension of PCa using mpMRI.
Assistance on treatment				
Largent *et al*[70] 2019	Training set:25 patients Validation set:14/14/11 patients	U-Net+GAN	MAE ≤34.4 HU	Generating pseudo-CT images directly from MRI to improve the workflow of radiotherapy without the need to transfer MRI to CT.
Holzschuh *et al*[71] 2023*	Training set:128 patients Internal test set:19 patients External test set:33 patients	3D U-Net	Specificity: 0.83 (IQR: 0.57–0.97) Sensitivity: 0.88 (IQR: 0.68–0.97)	Create a DL model for the accurate delineation of the intraprostatic GTV in PSMA-PET for focal radiation dose escalation and other treatment approaches in PCa.
Prognosis				
Lee *et al*[72] 2020*	Training set: 233 patients Validation set: 10% of training set	2D CNN(ResNet50)/3D CNN(ResNet-14)	2D CNN (better)–Sensitivity: 0.907 Specificity: 0.951, AUC: 0.971	Evaluate the performance of DL classifiers in discriminating normal and abnormal ^{18}F-FACBC PET scans based on the presence of tumor recurrence and/or metastases in patients with PCa and biochemical recurrence.
Zhang *et al*[73] 2016	Training set: 205 patients Validation set: fivefold cross validation	Support vector machine	AUC: 0.959 Sensitivity: 0.933 Specificity: 0.917	Develop an imaging-based approach in the prediction of 3-years biochemical recurrence.
Kendrick *et al*[74] 2022*	Training set: 209 scans from 121 patients Test set: 75 scans from 72 patients	nnU-net	Lesion detection–PPV: 0.972 Sensitivity: 0.730	Develop and assess an automated segmentation framework based on DL for metastatic PCa lesions in whole-body ^{68}Ga-PSMA-11 PET/CT images for the purpose of extracting patient-level prognostic biomarkers.

*Studies are based on PSMA PET/CT, others are based on MRI. AI: Artificial intelligence; AUC: Area under curve; CNN: Conventional neural network; CTV: Clinical target volume; DL: Deep learning; DSC: Dice score coefficient; GAN: Generative adversarial network; GTV: Gross tumor volume; MAE: Mean absolute error; mpMRI: Multiparametric magnetic resonance imaging; PBM: Patch-based method; PCa: Prostate cancer; PET/CT: Positron emission tomography/computed tomography; PL: Perceptual loss; PPV: Positive predictive value; PSMA: Prostate-specific membrane antigen.

for detecting both bone lesions and lymph nodes but low sensitivity for detecting local lesions (sensitivity 79%). Besides the function mentioned above, a model[62] applied another way of imaging,^{18}F-PSMA-1007 PET/CT, achieving a mean sensitivity of 82% for detecting suspected pelvic lymph node metastasis(compared to 77% for the nuclear medicine physician). Thus, the performance of AI in detecting cancer lesions may be comparable to that of nuclear medicine physicians. Although these studies demonstrate the great performance of AI in identifying and classifying metastatic lesions, there are still some challenges in accurately quantifying tumors using PSMA PET/CT. It is an evolving field, with many of the models yet to be tested for their generalizability with a larger external test set. In addition, the limited availability of large, well-designed datasets in PSMA PET/CT, variable PSMA expression, and the lack of well-defined anatomical boundaries pose challenges for lesion depiction and quantification. Rigorous testing and algorithmic refinement are required before integrating these models into clinical workflows.[12]

AI can be even better at guiding biopsy sampling based on its precise segmentation. Prospectively, the ProGNet[63] structure was integrated into fusion biopsies of 11 patients which outperformed the overall holistically nested edge detector, U-Net, and radiology technicians in a retrospective internal test set, with a Dice similarity coefficient (DSC) of 0.92 (P <0.0001). ProGNet required 35 seconds for each segmentation task, whereas radiology technicians take 10 minutes. Instead of focusing on the primary lesion, a DL model[64] segments organs near a lesion and then uses the relationship of these organs to the clinical target volume (CTV) to assist in CTV segmentation. The model achieved an average Dice similarity coefficient DSC of 0.87 on a preserved dataset of 50 patients. The study also estimated the uncertainty of the automatically segmented CTV contours and showed a level comparable to that of a physician. The segmentation function is the most basic function of the AI model for lesion detection, and its improvement in accuracy and efficiency not only provides a more reliable basis for diagnosis, but also lays a solid foundation for subsequent treatment, prognosis, and risk assessment.

However, AI does not always give a satisfying result. For prostate lesion detection, a computer-aided diagnosis (CAD)[65] for prostate lesion detection based on random forest, showed great performance (AUC = 0.93). At the same time, the study also pointed out that there were serious false positives in Gleason's 7-point judgment, which could be attributed to the fact that the CAD system may have mistaken some transition zone lesions for fibromuscular stroma. A model[66] based on the U-Net and AH-Net architectures focused on PCa lesion detection and segmentation by bpMRI. In the test validation, the U-Net model had 148 false positives, of which 21 corresponded to regions detected by radiologists but with negative biopsy results; the AH-Net model had 109 false positives, of which 19 corresponded to regions were detected by radiologists but with negative biopsy results. Although a PAGNet model[67] performed well in all training test sets (AUC = 0.728–0.857), surpassing the performance of two human experts in detecting extraperitoneal invasion of PCa, when the AI-assisted the experts in their prediction work, the performance of both experts was improved.

Obviously, the current studies are still dominated by supervised training, where the quality of data is of great importance. However, the sample size for the internal training dataset is roughly in the triple-digit order of magnitude, lacking a sufficiently large number of external training dataset samples to ensure the generalizability of the models. Even with supervised training, some models had more false positives, and stability could not be guaranteed in the external validation cohort. However, a larger dataset volume implies a greater workload for labeling. The quality of the dataset may not be as good as that of smaller datasets. Despite some flaws, it has to be acknowledged that the results shown by AI under training with a limited sample size are undoubtedly encouraging. Compared to completely replacing the decision-making of imaging physicians, future research may further explore how to better integrate these models into the clinical workflow as an aid to better help physicians make diagnoses.

Applications of AI imaging in PCa treatment: New angles beyond primary lesions

In addition to the fact that radiotherapy is one of the effective methods of PCa treatment, its efficacy is affected by the imprecision of segmentation, which has great side effects on intestinal function.[68] Precise delineation of the malignant lesion contour improves the accuracy of the localization of targeted radiotherapy against the lesion and the relative radiation therapy dose, thus reducing the recurrence rate[69] and side effects. Therefore, more powerful imaging-aiding tools are needed to tailor to patients' conditions in case of overtreatment.

To improve the workflow of radiotherapy, a study[70] tried to use AI to eliminate the step of transferring MRI to CT when obtaining electron density information. Two methods—a DL-based method (including U-Net and generative adversarial network [GAN] to generate DL models for adversarial networks) and a patch-based method (PBM)—were used in generating pseudo-CT (pCT) images which were then evaluated the accuracy of these pCTs in prostate radiotherapy dose planning. It turned out that the U-Net and GAN model exhibited extremely low measurement uncertainty (lowest mean absolute error (MAE) [≤34.4 HU]) and high efficiency (the average computation time for generating a pCT is 15 seconds for DLM and 62 minutes for PBM).

Gross tumor volume (GTV) within the prostate refers to the tumor tissue visible on imaging, including the tumor within the prostate as well as areas of possible metastatic lymph nodes. GTV accurate delineation is critical for determining the target area and dose of radiation therapy. To automate the delineation of GTV, a study[71] combining 3D U-Net framework and three PSMA PET tracers can greatly boost GTV segmentation with high accuracy, which is comparable to experts' manual segmentation.

It can be said that AI can provide more accurate lesion area segmentation and delineation, which can provide a

new reference for targeted treatment and dose calculation of radiotherapy, and help to improve the effectiveness of radiation therapy, which will undoubtedly provide a new reference for targeted treatment and radiation dose of radiotherapy in the future.

Application of AI in prognosis prediction of PCa: Integration and selection of diverse information and data

AI can predict BCR, distant metastasis and other prognosis of PCa patients based on preoperative imaging data. In view of its powerful analysis and learning ability, it can more accurately distinguish the risk stratification of PCa patients and provide a reliable basis for the subsequent clinical personalized treatment and full management, to achieve a better prognosis. For automatically identifying patients with no recurrence and patients with local or distant recurrence, two models based on CNN were developed. Instead of the assumed result that more complicated data came out a more powerful model, CNN with 2D slices achieved the best prediction results (AUC = 0.971 in the training set), outperforming 3D CNN.[72] in detecting pelvic tumor recurrence and metastasis. Similarly, a traditional model structure was not doomed to be inferior in prediction ability. A support vector machine (SVM) model,[73] a much older structure, that combines clinicopathological and MRI imaging variables also effectively facilitated the prediction of BCR within three years in patients with localized PCa before surgical treatment (AUC = 0.970). SVM-constructed nomograms significantly outperformed the logistic regression analyses. Simply based on the images, another study[74] constructed an AI based on a CNN framework with PET-CT and MRI as inputs, which was able to detect and accurately delineate lesions using a cascade network architecture, as well as extract the patient's total lesion volume (TLV) and total lesion uptake (TLU) for survival analysis assessment. The new model also succeeded in extracting prognostic biomarkers.

In the other fields, cancer imaging biomarkers were built with the help of foundation models giving the model much more powerful abilities.[75] Such foundation model implementations showed superiority over compared baseline methods in lesion anatomical site classification (mean average precision [mAP] = 0.85, accuracy = 0.804). When training data was limited, the performance of the model kept improving. Moreover, a large amount of data enables the model to figure out features at the level of the gene. The absolute correlation coefficients between gene expression profiles and model predictions were significantly higher ($P = 0.008$) for the foundation model, indicating a stronger association with underlying tumor biology.

AI model in prognostic prediction using imaging data is different from diagnosis and treatment, in that imaging information alone cannot directly suggest a good or bad prognosis. More diverse types should be included for the model to learn. Therefore, almost all of today's AI models incorporate more information types and more complicated information. However, such an upgrade does not necessarily mean higher predictive power, and data that contains higher-dimension information is often constrained by data quality and size and cannot guarantee reliable output. For large models trained by a great amount of data, it can still behave well even on scarce data. There is a strong indication that large models will thrive in the near future.

Application for large visual model in PCa: A promising future

The application of large visual models in the field of PCa imaging remains scarce at present, and most of the models are not able to test the generalizability and stability of the models under more data tests due to the lack of sufficient data volume. In addition, the parameters are not uniform among different centers, and there is a lack of an established standard to generalize the parameters in the study. Currently, open PCa image databases are extremely scarce, limiting many researchers, especially independent research teams, to obtaining enough samples to conduct studies that require more multicenter collaboration to obtain large amounts of data and professional support. In the future, with the continuous updating of the AI framework and datasets, the research on large models in PCa imaging will become a hot spot.

Application and Prospective of Large Language Model in PCa: Powerful Tools of Great Scalablity and Stability

The Transformer model architecture (Transformer model) has revolutionized the field of AI, especially in the field of natural language processing (NLP), where the most famous Chat Generative Pre-trained Transformer (ChatGPT) is built based on this model. The Transformer model introduces a new multi-self attention mechanism, which allows modeling of dependencies regardless of their distance in the input or output sequences which is important for understanding and generating natural language.

Another advantage of the Transformer model over traditional recurrent neural networks (RNN) is that it allows for parallel computation. To be specific, all the words are processed at the same time instead of one after the other as in RNN which makes the Transformer model more efficient in processing large amounts of data at the same time.

Transformer-based pre-trained models such as Bidirectional Encoder Representations from Transformer (BERT), GPT-2, GPT-3, etc. have achieved state-of-the-art results in various NLP tasks. These models are first pre-trained on a large amount of textual data to learn the general properties of the language and then fine-tuned on specific downstream tasks. Such workflow has become standard practice in the field of NLP. Nowadays, more and more institutions have laid emphasis on large language models (LLMs) and several commercial models have been available to the public already [Table 4].

Due to the current establishment of the Transformer architecture, a large language model that can be applied to PCa aspects is gradually being developed. Early identification of metastatic PCa (mPCa) is crucial for patient prognosis, so an NLP algorithm[76] was developed and validated, which can also be deployed on oncology, urology, and

Table 4: Examples of commercial large language models.

Large language models	Developing institution	Function
Med-PaLM	Google	Clinical decision support, medical Q&A, medical record analysis, personalized medicine, literature search, health education
BioMedLM	Stanford University	Biomedical literature search, clinical decision support, disease and symptom analysis, give personalized recommendations, medical education, research assistance
GatorTron	University of Florida	Medical Q&A system, electronic medical record analysis, personalized medical recommendations, medical literature processing, health management and education
Shuzhi Qihuang	East China Normal University, Shanghai University of Traditional Chinese Medicine, Naval Medical University, East China University of Science and Technology	Intelligent Q&A of knowledge in the field of Chinese medicine, health consultation, dynamic interaction of Chinese medicine knowledge map
HuatuoGPT	Shenzhen Research Institute of Big Data, The Chinese University of Hong Kong	Medical counseling patient training, health counseling, triage for medical care
Zhongjin	Fudan University, Tongji University	Traditional Chinese medicine Q&A, disease diagnosis, personalized plan, case analysis, literature and data search
Pangu	Zhejiang Jiuwei Health Technology Co., Ltd, Huawei Cloud Computing Technology Co., Ltd	Drug discovery, intelligent medicine
TCMLLM	Beijing Jiaotong University	Chinese medicine clinical aids, disease diagnosis, prescription recommendations, Chinese medicine knowledge quiz

Q&A: Question and answer.

radiology clinical records, to identify 1,144,610 veterans diagnosed PCa in the Department of Veterans Affairs (VA) (specificity = 0.979, sensitivity = 0.919). It also could be used to enhance cancer registry data, facilitate research queries and identify patients who may benefit from mPCa treatment innovations. Another NLP-based model[77] with a similar function found electronic health records (EHRs) as a new resource for input. The study used 139,057 radiology reports from 6211 patients, and after training the NLP model showed high sensitivity (85%) and high specificity (96%) with an AUC of 0.911, which was better than ICD9/10 judgment. One of the most notable advantages is that NLP is compatible with various data types. However, both studies are constricted to one dataset meaning the lack of proof and training in models' generalization. Upon further validation, these models can be effective in identifying mPCa patients in the population.

Apart from the prediction of mPCa, due to the variability of the large language model, it can also be trained by certain commands. Our team[78] explored the potential of the new generalized multimodal AI model, ChatGPT-4V, for pathology slice review and democratization of medical knowledge on urological malignancies, and found that ChatGPT-4V performed excellently in distinguishing renal cell carcinoma (RCC) from normal renal tissue, and could perform medical image analysis and medical basic knowledge explanation for simulated patients only by setting a certain situation and commands. However, ChatGPT-4V performed poorly in distinguishing PCa from benign prostate tissue. The reason for this may be that ChatGPT-4V can recognize intracellular changes in RCC, such as nuclear heterogeneity, but cannot detect the tissue structural changes present in PCa, and there is a scarcity of training volume.

In addition to LLMs that specialize in a particular disease area,[79] there are many LLMs that are used throughout the healthcare system to improve efficiency in the underlying architecture, assist in diagnosis, etc. The novel model Foresight,[80] a Transformer-based architecture, was pretrained on EHRs of 811,336 patients from three institutions: King's College Hospital (KCH), South London and Maudsley NHS Foundation Trust (SLaM), and the Medical Information Mart for Intensive Care III (MIMIC-III) dataset. It demonstrated high predictive precision across these cohorts: 0.68 (KCH), 0.76 (SLaM), and 0.88 (MIMIC-III). Foresight is able to handle a large amount of EHR data, including unstructured free text, and predicts a wide range of future medical outcomes, is scalable, and its performance gets better as the amount of data increases.

Overall, the models generated by domain-specific pretraining on the medical corpus perform well on biomedical tasks and surpass the general LLM (with some exceptions such as GPT-4). However, fine-tuning methods for adapting generic models such as BERT and GPT-3 achieve good results on medical tasks in a more computationally efficient manner. The models themselves are learning and getting better, which outperform those models when trained by less data.[81] Although LLMs such as ChatGPT show great potential in the medical field, the answers they provide may contain errors and inconsistent responses can occur. When it comes to humanistic care, LLMs are far from human beings and can only show this

with specific questions.[82] In the future, such flaws should be paid more attention to when developing LLM.

Challenges and Future Development of the Clinical Application of AI

Lack of interpretability

The application of AI in the medical field encompasses crucial considerations such as legal implications, ethical concerns, and privacy issues. Despite the significant potential demonstrated by AI in various aspects of PCa clinical applications, such as diagnosis, treatment, prognosis, etc., it is crucial to acknowledge that the current decision-making process of AI is still a "black box".[83] The interpretability of AI is crucial for establishing trust, ensuring fairness in decision making, and adherence to ethical standards, which is fundamental to clinical applications. Currently, the interpretation and comprehension of DL methods for analyzing pathological and imaging images of PCa remain challenging, thereby impeding their acceptance and implementation in clinical applications. Since 2010, a series of technologies have gradually emerged to improve the interpretability of AI.[84] A DL model called Lesion-Aware Generative Adversarial Networks (LAGAN)[85] incorporates a model visualization tool to generate an object heat map. This heatmap can visually show the areas that the model is focusing on, providing a visual explanation of the decision-making process of the model. Another DL model called P-NET[45] achieves its interpretability through visualization techniques such as node activation distribution and feature importance plots. Despite some progress in the field of AI interpretability, it remains at an exploratory stage. To ensure the widespread development of AI in clinical applications, addressing AI interpretability becomes a crucial concern for future advancements.

Limitation of data

Currently, the majority of studies are retrospective single-center investigations with limited sample sizes and a lack of external validation, thereby compromising the generalizability of AI models. One significant contributing factor is the variation in data acquisition and screening criteria across different studies, which hampers the widespread implementation of AI in clinical practice. In the future, the widespread adoption of AI models will heavily rely on advancements in data standards and the enhanced capability of AI algorithms to seamlessly transform data across different centers. The increasing availability of publicly accessible databases poses a significant challenge in safeguarding patient privacy, enhancing legal and regulatory frameworks, and promoting responsible data utilization.[86] A large dataset called OpenPath[44] was constructed by collecting publicly available medical information on medical Twitter. Subsequently, PLIP, a fundamental vision-language model developed based on this dataset, has exhibited exceptional performance across multiple external datasets and can be effectively employed for diverse pathological image analysis tasks. The utilization of multimodal data is increasingly prevalent as the volume of data continues to grow. The Generalist Vision-Language Foundation model, known as Biomedical Generative Pre-trained Transformer (BiomedGPT),[87]effectively performs a diverse range of intricate biomedical tasks by integrating multimodal data acquired from 14 publicly accessible datasets. This highlights the significance of multimodal AI as a prominent trend in clinical applications.

Fairness and AI bias

The principle of fairness holds significant importance in the clinical application of AI; however, biases may arise during the development and implementation process, thereby impacting the overall fairness of AI systems. The bias of AI stems from algorithmic and data biases, while in clinical applications, bias also exists in the interaction between AI systems, clinicians, and patients.[88] Chen *et al*[89] evaluated whether the treatment recommendations provided by the large language model GPT-3.5-turbo0301 for cancers such as PCa were consistent with the NCCN guidelines in the United States, which found that at least one-third were inconsistent with the NCCN guidelines, and the main cause of inconsistency was bias. Clinicians must have a clear understanding that AI is merely a tool to assist in clinical diagnosis and treatment. It cannot surpass or replace human expertise, and excessive reliance on AI should be avoided during the process of clinical diagnosis and treatment. The fairness of artificial intelligence recommendations in healthcare (FAIR) principles, proposed by Ueda *et al*[90], serve as a set of recommendations aimed at mitigating bias and fostering fairness in the application of AI within the healthcare domain. The future advancement of algorithms, the utilization of diverse and representative datasets in training processes, and the enhancement of education for clinicians and patients regarding AI biases will serve as crucial strategies to facilitate the integration of AI into clinical applications.

Conclusions

The application of AI in the medical field is becoming increasingly extensive and sophisticated, driven by advancements in computer technology and the optimization of DL algorithms. The integration of AI with pathology and imaging has become a cornerstone of precision medicine for PCa, transforming traditional medical paradigms. AI models are progressively being employed across various aspects of PCa care, including diagnosis, treatment planning, prognosis prediction, and molecular profiling. These models enable clinicians to assess PCa with greater accuracy and efficiency, facilitating personalized treatment and comprehensive management throughout the care continuum. This approach ultimately contributes to reducing mortality rates and improving patients' quality of life. With ongoing updates to AI frameworks and datasets, foundation models exhibit remarkable scalability and stability, presenting new opportunities for their application in medicine. However, it is crucial to acknowledge that AI remains in its early stages, with limitations in clinical practice. In the future, as databases become more robust, algorithms are further refined, and supportive laws and regulations are developed,

AI is poised to play an even more transformative role in precision medicine for PCa.

Funding

This review was supported by the National Natural Science Foundation of China (Nos. 82272905 and 82473385) and the Natural Science Foundation of Science and Technology Commission of Shanghai (No. 22ZR1478000).

Conflicts of interest

None.

References

1. Graham LS, Lin JK, Lage DE, Kessler ER, Parikh RB, Morgans AK. Management of prostate cancer in older adults. Am Soc Clin Oncol Educ Book 2023;43:e390396. doi:10.1200/EDBK_390396.
2. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, *et al*. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA A Cancer J Clinicians 2024;74:229–263. doi: 10.3322/caac.21834.
3. He S, Xia C, Li H, Cao M, Yang F, Yan X, *et al*. Cancer profiles in China and comparisons with the USA: A comprehensive analysis in the incidence, mortality, survival, staging, and attribution to risk factors. Sci China Life Sci 2024;67:122–131. doi: 10.1007/s11427-023-2423-1.
4. De Vos II, Meertens A, Hogenhout R, Remmers S, Roobol MJ. A detailed evaluation of the effect of prostate-specific antigen–based screening on morbidity and mortality of prostate cancer: 21-year follow-up results of the Rotterdam section of the European Randomised Study of Screening for Prostate Cancer. Eur Urol 2023;84:426–434. doi: 10.1016/j.eururo.2023.03.016.
5. Aragona F, Pepe P, Motta M, Saita A, Raciti G, Rosa PL, *et al*. Incidence of prostate cancer in Sicily: Results of a Multicenter case-findings protocol. Eur Urol 2005;47:569–574. doi: 10.1016/j.eururo.2004.11.007.
6. Nir G, Hor S, Karimi D, Fazli L, Skinnider BF, Tavassoli P, *et al*. Automatic grading of prostate cancer in digitized histopathology images: Learning from multiple experts. Med Image Anal 2018;50:167–180. doi: 10.1016/j.media.2018.09.005.
7. Zhu L, Pan J, Mou W, Deng L, Zhu Y, Wang Y, *et al*. Harnessing artificial intelligence for prostate cancer management. Cell Rep Med 2024;5:101506. doi: 10.1016/j.xcrm.2024.101506.
8. Howell MD, Corrado GS, DeSalvo KB. Three epochs of artificial intelligence in health care. JAMA 2024;331:242. doi: 10.1001/jama.2023.25057.
9. Song Z, Zhang W, Jiang Q, Deng L, Du L, Mou W, *et al*. Artificial intelligence-aided detection for prostate cancer with multi-modal routine health check-up data: An Asian multi-center study. Int J Surg 2023;109:3848–3860. doi: 10.1097/JS9.0000000000000862.
10. Vorontsov E, Bozkurt A, Casson A, Shaikovski G, Zelechowski M, Severson K, *et al*. A foundation model for clinical-grade computational pathology and rare cancers detection. Nat Med 2024;30:2924–2935. doi: 10.1038/s41591-024-03141-0.
11. Ozkan TA, Eruyar AT, Cebeci OO, Memik O, Ozcan L, Kuskonmaz I. Interobserver variability in Gleason histological grading of prostate cancer. Scand J Urology 2016;50:420–424. doi: 10.1080/21681805.2016.1206619.
12. Riaz IB, Harmon S, Chen Z, Naqvi SAA, Cheng L. Applications of artificial intelligence in prostate cancer care: A path to enhanced efficiency and outcomes. Am Soc Clin Oncol Educ Book 2024;44:e438516. doi: 10.1200/EDBK_438516.
13. Kartasalo K, Bulten W, Delahunt B, Chen P-HC, Pinckaers H, Olsson H, *et al*. Artificial intelligence for diagnosis and Gleason grading of prostate cancer in biopsies—Current status and next steps. Eur Urol Focus 2021;7:687–691. doi: 10.1016/j.euf.2021.07.002.
14. Tolkach Y, Ovtcharov V, Pryalukhin A, Eich M-L, Gaisa NT, Braun M, *et al*. An international multi-institutional validation study of the algorithm for prostate cancer detection and Gleason grading. NPJ Precis Onc 2023;7:77. doi: 10.1038/s41698-023-00424-6.
15. Pantanowitz L, Quiroga-Garza GM, Bien L, Heled R, Laifenfeld D, Linhart C, *et al*. An artificial intelligence algorithm for prostate cancer diagnosis in whole slide images of core needle biopsies: A blinded clinical validation and deployment study. Lancet Digit Health 2020;2:e407–e416. doi: 10.1016/S2589-7500(20)30159-X.
16. Nagpal K, Foote D, Tan F, Liu Y, Chen P-HC, Steiner DF, *et al*. Development and validation of a deep learning algorithm for Gleason grading of prostate cancer from biopsy specimens. JAMA Oncol 2020;6:1372. doi: 10.1001/jamaoncol.2020.2485.
17. 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;21:233–241. doi: 10.1016/S1470-2045(19)30739-9.
18. Campanella G, Hanna MG, Geneslaw L, Miraflor A, Werneck Krauss Silva V, Busam KJ, *et al*. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. Nat Med 2019;25:1301–1309. doi: 10.1038/s41591-019-0508-1.
19. Rana A, Lowe A, Lithgow M, Horback K, Janovitz T, Da Silva A, *et al*. Use of deep learning to develop and analyze computational hematoxylin and eosin staining of prostate core biopsy images for tumor diagnosis. JAMA Netw Open 2020;3:e205111. doi: 10.1001/jamanetworkopen.2020.5111.
20. Sandhu S, Moore CM, Chiong E, Beltran H, Bristow RG, Williams SG. Prostate cancer. Lancet 2021;398:1075–1090. doi: 10.1016/S0140-6736(21)00950-8.
21. Nguyen PL, Alibhai SMH, Basaria S, D'Amico AV, Kantoff PW, Keating NL, *et al*. Adverse effects of androgen deprivation therapy and strategies to mitigate them. Eur Urol 2015;67:825–836. doi: 10.1016/j.eururo.2014.07.010.
22. Spratt DE, Tang S, Sun Y, Huang H-C, Chen E, Mohamad O, *et al*. Artificial intelligence predictive model for hormone therapy use in prostate cancer. NEJM Evid 2023;2:EVIDoa2300023. doi: 10.1056/EVIDoa2300023.
23. 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;27:999–1005. doi: 10.1038/s41591-021-01359-w.
24. Gao Y, Shen C, Jia X, Kyun Park Y. Implementation and evaluation of an intelligent automatic treatment planning robot for prostate cancer stereotactic body radiation therapy. Radiother Oncol 2023;184:109685. doi: 10.1016/j.radonc.2023.109685.
25. 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. Lancet Digit Health 2021;3:e158–e165. doi: 10.1016/S2589-7500(20)30314-9.
26. 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;21:222–232. doi: 10.1016/S1470-2045(19)30738-7.
27. 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;82:334–345. doi: 10.1158/0008-5472.CAN-21-2843.
28. Wu S, Wang Y, Hong G, Luo Y, Lin Z, Shen R, *et al*. An artificial intelligence model for detecting pathological lymph node metastasis in prostate cancer using whole slide images: A retrospective, multicentre, diagnostic study. EClinicalMed 2024;71:102580. doi: 10.1016/j.eclinm.2024.102580.
29. Wei L, Huang Y, Chen Z, Lei H, Qin X, Cui L, *et al*. Artificial intelligence combined with big data to predict lymph node involvement in prostate cancer: A population-based study. Front Oncol 2021;11:763381. doi: 10.3389/fonc.2021.763381.
30. Wu S, Hong G, Xu A, Zeng H, Chen X, Wang Y, *et al*. Artificial intelligence-based model for lymph node metastases detection on whole slide images in bladder cancer: A retrospective, multicentre, diagnostic study. Lancet Oncol 2023;24:360–370. doi: 10.1016/S1470-2045(23)00061-X.
31. Ross AE, Zhang J, Huang H-C, Yamashita R, Keim-Malpass J, Simko JP, *et al*. External validation of a digital pathology-based multimodal artificial intelligence architecture in the NRG/RTOG 9902 phase 3 trial. Eur Urol Oncol 2024;7:1024–1033. doi: 10.1016/j.euo.2024.01.004.

32. Lotan TL, Wei W, Ludkovski O, Morais CL, Guedes LB, Jamaspishvili T, *et al*. Analytic validation of a clinical-grade PTEN immunohistochemistry assay in prostate cancer by comparison with PTEN FISH. Mod Pathol 2016;29:904–914. doi: 10.1038/modpathol.2016.88.
33. Jamaspishvili T, Berman DM, Ross AE, Scher HI, De Marzo AM, Squire JA, *et al*. Clinical implications of PTEN loss in prostate cancer. Nat Rev Urol 2018;15:222–234. doi: 10.1038/nrurol.2018.9.
34. Troyer DA, Jamaspishvili T, Wei W, Feng Z, Good J, Hawley S, *et al*. A multicenter study shows PTEN deletion is strongly associated with seminal vesicle involvement and extracapsular extension in localized prostate cancer. Prostate 2015;75:1206–1215. doi: 10.1002/pros.23003.
35. Krohn A, Diedler T, Burkhardt L, Mayer P-S, De Silva C, Meyer-Kornblum M, *et al*. Genomic deletion of PTEN is associated with tumor progression and early PSA recurrence in ERG fusion-positive and fusion-negative prostate cancer. Am J Pathol 2012;181:401–412. doi: 10.1016/j.ajpath.2012.04.026.
36. Harmon SA, Patel PG, Sanford TH, Caven I, Iseman R, Vidotto T, *et al*. High throughput assessment of biomarkers in tissue microarrays using artificial intelligence: PTEN loss as a proof-of-principle in multi-center prostate cancer cohorts. Mod Pathol 2021;34:478–489. doi: 10.1038/s41379-020-00674-w.
37. Dadhania V, Gonzalez D, Yousif M, Cheng J, Morgan TM, Spratt DE, *et al*. Leveraging artificial intelligence to predict ERG gene fusion status in prostate cancer. BMC Cancer 2022;22:494. doi: 10.1186/s12885-022-09559-4.
38. Dai C, Zeng X, Zhang X, Liu Z, Cheng S. Machine learning-based integration develops a mitophagy-related lncRNA signature for predicting the progression of prostate cancer: A bioinformatic analysis. Discov Onc 2024;15:316. doi: 10.1007/s12672-024-01189-5.
39. George RS, Htoo A, Cheng M, Masterson TM, Huang K, Adra N, *et al*. Artificial intelligence in prostate cancer: Definitions, current research, and future directions. Urol Oncol 2022;40:262–270. doi: 10.1016/j.urolonc.2022.03.003.
40. Lu MY, Chen B, Williamson DFK, Chen RJ, Liang I, Ding T, *et al*. A visual-language foundation model for computational pathology. Nat Med 2024;30:863–874. doi: 10.1038/s41591-024-02856-4.
41. Ilse M, Tomczak JM, Welling M. Attention-based deep multiple instance learning. Proc 35th Int Conf Mach Learn 2018;80:1–9. doi: 10.1145/3208118.
42. Xu H, Usuyama N, Bagga J, Zhang S, Rao R, Naumann T, *et al*. A whole-slide foundation model for digital pathology from real-world data. Nature 2024;630:181–188. doi: 10.1038/s41586-024-07441-w.
43. Wang X, Zhao J, Marostica E, Yuan W, Jin J, Zhang J, *et al*. A pathology foundation model for cancer diagnosis and prognosis prediction. Nature 2024;634:970–978. doi: 10.1038/s41586-024-07894-z.
44. Huang Z, Bianchi F, Yuksekgonul M, Montine TJ, Zou J. A visual–language foundation model for pathology image analysis using medical Twitter. Nat Med 2023;29:2307–2316. doi: 10.1038/s41591-023-02504-3.
45. Elmarakeby HA, Hwang J, Arafeh R, Crowdis J, Gang S, Liu D, *et al*. Biologically informed deep neural network for prostate cancer discovery. Nature 2021;598:348–352. doi: 10.1038/s41586-021-03922-4.
46. Gaziev G, Wadhwa K, Barrett T, Koo BC, Gallagher FA, Serrao E, *et al*. Defining the learning curve for multiparametric magnetic resonance imaging (MRI) of the prostate using MRI-transrectal ultrasonography (TRUS) fusion-guided transperineal prostate biopsies as a validation tool. BJU Int 2016;117:80–86. doi: 10.1111/bju.12892.
47. Rosenkrantz AB, Babb JS, Taneja SS, Ream JM. Proposed adjustments to PI-RADS version 2 decision rules: Impact on Prostate cancer detection. Radiology 2017;283:119–129. doi: 10.1148/radiol.2016161124.
48. Boesen L, Nørgaard N, Løgager V, Balslev I, Bisbjerg R, Thestrup K-C, *et al*. Assessment of the diagnostic accuracy of biparametric magnetic resonance imaging for prostate cancer in biopsy-naive men: The biparametric MRI for detection of prostate cancer (BIDOC) study. JAMA Netw Open 2018;1:e180219. doi: 10.1001/jamanetworkopen.2018.0219.
49. Cytawa W, Kircher S, Kübler H, Werner RA, Weber S, Hartrampf P, *et al*. Diverse PSMA expression in primary prostate cancer: reason for negative [^{68}Ga]Ga-PSMA PET/CT scans? Immunohistochemical validation in 40 surgical specimens. Eur J Nucl Med Mol Imaging 2022;49:3938–3949. doi: 10.1007/s00259-022-05831-8.
50. Zhang L-L, Li W-C, Xu Z, Jiang N, Zang S-M, Xu L-W, *et al*. ^{68}Ga-PSMA PET/CT targeted biopsy for the diagnosis of clinically significant prostate cancer compared with transrectal ultrasound guided biopsy: A prospective randomized single-centre study. Eur J Nucl Med Mol Imaging 2021;48:483–492. doi: 10.1007/s00259-020-04863-2.
51. Rüschoff JH, Ferraro DA, Muehlematter UJ, Laudicella R, Hermanns T, Rodewald A-K, *et al*. What's behind ^{68}Ga-PSMA-11 uptake in primary prostate cancer PET? Investigation of histopathological parameters and immunohistochemical PSMA expression patterns. Eur J Nucl Med Mol Imaging 2021;48:4042–4053. doi: 10.1007/s00259-021-05501-1.
52. Pepe P, Pepe L, Tamburo M, Marletta G, Savoca F, Pennisi M, *et al*. ^{68}Ga-PSMA PET/CT and prostate cancer diagnosis: Which SUVmax value? In Vivo 2023;37:1318–1322. doi: 10.21873/invivo.13211.
53. Hartenstein A, Lübbe F, Baur ADJ, Rudolph MM, Furth C, Brenner W, *et al*. Prostate cancer nodal staging: using deep learning to predict ^{68}Ga-PSMA-positivity from CT imaging alone. Sci Rep 2020;10:3398. doi: 10.1038/s41598-020-60311-z.
54. Zheng Q, Yang L, Zeng B, Li J, Guo K, Liang Y, *et al*. Artificial intelligence performance in detecting tumor metastasis from medical radiology imaging: A systematic review and meta-analysis. EClinicalMed 2021;31:100669. doi: 10.1016/j.eclinm.2020.100669.
55. European Association of Urology. EAU guidelines, 2022 edition, 2022. Available from: https://uroweb.org/guidelines [Last accessed on April 2022].
56. Schaeffer EM, Srinivas S, Adra N, An Y, Barocas D, Bitting R, *et al*. Prostate cancer, version 4.2023, NCCN clinical practice guidelines in oncology. J Natl Compr Canc Netw 2023;21:1067–1096. doi: 10.6004/jnccn.2023.0050.
57. Rundo L, Militello C, Russo G, Garufi A, Vitabile S, Gilardi M, *et al*. Automated prostate gland segmentation based on an unsupervised fuzzy C-means clustering technique using multispectral T1w and T2w MR imaging. Information 2017;8:49. doi: 10.3390/info8020049.
58. Tian Z, Liu L, Zhang Z, Fei B. PSNet: Prostate segmentation on MRI based on a convolutional neural network. J Med Imag 2018;5:1. doi: 10.1117/1.JMI.5.2.021208.
59. Karimi D, Samei G, Kesch C, Nir G, Salcudean SE. Prostate segmentation in MRI using a convolutional neural network architecture and training strategy based on statistical shape models. Int J Comput Assist Radiol Surg 2018;13:1211–1219. doi: 10.1007/s11548-018-1785-8.
60. Hofman MS, Lawrentschuk N, Francis RJ, Tang C, Vela I, Thomas P, *et al*. Prostate-specific membrane antigen PET-CT in patients with high-risk prostate cancer before curative-intent surgery or radiotherapy (proPSMA): A prospective, randomised, multicentre study. Lancet 2020;395:1208–1216. doi: 10.1016/S0140-6736(20)30314-7.
61. Zhao Y, Gafita A, Vollnberg B, Tetteh G, Haupt F, Afshar-Oromieh A, *et al*. Deep neural network for automatic characterization of lesions on ^{68}Ga-PSMA-11 PET/CT. Eur J Nucl Med Mol Imaging 2020;47:603–613. doi: 10.1007/s00259-019-04606-y.
62. Trägårdh E, Enqvist O, Ulén J, Hvittfeldt E, Garpered S, Belal SL, *et al*. Freely available artificial intelligence for pelvic lymph node metastases in PSMA PET-CT that performs on par with nuclear medicine physicians. Eur J Nucl Med Mol Imaging 2022;49:3412–3418. doi: 10.1007/s00259-022-05806-9.
63. Soerensen SJC, Fan RE, Seetharaman A, Chen L, Shao W, Bhattacharya I, *et al*. Deep learning improves speed and accuracy of prostate gland segmentations on magnetic resonance imaging for targeted biopsy. J Urol 2021;206:604–612. doi: 10.1097/JU.0000000000001783.
64. Balagopal A, Nguyen D, Morgan H, Weng Y, Dohopolski M, Lin M-H, *et al*. A deep learning-based framework for segmenting invisible clinical target volumes with estimated uncertainties for post-operative prostate cancer radiotherapy. Med Image Anal 2021;72:102101. doi: 10.1016/j.media.2021.102101.
65. Lay N, Tsehay Y, Greer MD, Turkbey B, Kwak JT, Choyke PL, *et al*. Detection of prostate cancer in multiparametric MRI using random forest with instance weighting. J Med Imag 2017;4:024506. doi: 10.1117/1.JMI.4.2.024506.
66. Mehralivand S, Yang D, Harmon SA, Xu D, Xu Z, Roth H, *et al*. A cascaded deep learning-based artificial intelligence algorithm for

automated lesion detection and classification on biparametric prostate magnetic resonance imaging. Acad Radiol 2022;29:1159–1168. doi: 10.1016/j.acra.2021.08.019.
67. Hou Y, Zhang Y-H, Bao J, Bao M-L, Yang G, Shi H-B, *et al*. Artificial intelligence is a promising prospect for the detection of prostate cancer extracapsular extension with mpMRI: A two-center comparative study. Eur J Nucl Med Mol Imaging 2021;48:3805–3816. doi: 10.1007/s00259-021-05381-5.
68. Donovan JL, Hamdy FC, Lane JA, Mason M, Metcalfe C, Walsh E, *et al*. Patient-reported outcomes after monitoring, surgery, or radiotherapy for prostate cancer. N Engl J Med 2016;375:1425–1437. doi: 10.1056/NEJMoa1606221.
69. Cellini N, Morganti AG, Mattiucci GC, Valentini V, Leone M, Luzi S, *et al*. Analysis of intraprostatic failures in patients treated with hormonal therapy and radiotherapy: implications for conformal therapy planning. Int J Radiat Oncol Biol Phys 2002;53:595–599. doi: 10.1016/S0360-3016(02)02795-5.
70. Largent A, Barateau A, Nunes J-C, Mylona E, Castelli J, Lafond C, *et al*. Comparison of deep learning-based and patch-based methods for pseudo-CT generation in MRI-based prostate dose planning. Int J Radiat Oncol Biol Phys 2019;105:1137–1150. doi: 10.1016/j.ijrobp.2019.08.049.
71. Holzschuh JC, Mix M, Ruf J, Hölscher T, Kotzerke J, Vrachimis A, *et al*. Deep learning based automated delineation of the intraprostatic gross tumour volume in PSMA-PET for patients with primary prostate cancer. Radiother Oncol 2023;188:109774. doi: 10.1016/j.radonc.2023.109774.
72. Lee JJ, Yang H, Franc BL, Iagaru A, Davidzon GA. Deep learning detection of prostate cancer recurrence with ^{18}F-FACBC (fluciclovine, Axumin®) positron emission tomography. Eur J Nucl Med Mol Imaging 2020;47:2992–2997. doi: 10.1007/s00259-020-04912-w.
73. Zhang YD, Wang J, Wu CJ, Bao ML, Li H, Wang XN, *et al*. An imaging-based approach predicts clinical outcomes in prostate cancer through a novel support vector machine classification. Oncotarget 2016;7:78140–78151. doi: 10.18632/oncotarget.11293.
74. Kendrick J, Francis RJ, Hassan GM, Rowshanfarzad P, Ong JSL, Ebert MA. Fully automatic prognostic biomarker extraction from metastatic prostate lesion segmentations in whole-body [^{68}Ga] Ga-PSMA-11 PET/CT images. Eur J Nucl Med Mol Imaging 2022;50:67–79. doi: 10.1007/s00259-022-05927-1.
75. Pai S, Bontempi D, Hadzic I, Prudente V, Sokač M, Chaunzwa TL, *et al*. Foundation model for cancer imaging biomarkers. Nat Mach Intell 2024;6:354–367. doi: 10.1038/s42256-024-00807-9.
76. Alba PR, Gao A, Lee KM, Anglin-Foote T, Robison B, Katsoulakis E, *et al*. Ascertainment of Veterans with metastatic prostate cancer in electronic health records: Demonstrating the case for natural language processing. JCO Clin Cancer Inform 2021:1005–1014. doi: 10.1200/CCI.21.00030.
77. Yang R, Zhu D, Howard LE, De Hoedt A, Williams SB, Freedland SJ, *et al*. Identification of patients with metastatic prostate cancer with natural language processing and machine learning. JCO Clin Cancer Inform 2022:e2100071. doi: 10.1200/CCI.21.00071.
78. Zhu L, Lai Y, Ta N, Cheng L, Chen R. Multimodal approach in the diagnosis of urologic malignancies: Critical assessment of ChatGPT-4V's image-reading capabilities. JCO Clin Cancer Inform 2024:e2300275. doi: 10.1200/CCI.23.00275.
79. Hershenhouse JS, Mokhtar D, Eppler MB, Rodler S, Storino Ramacciotti L, Ganjavi C, *et al*. Accuracy, readability, and understandability of large language models for prostate cancer information to the public. Prostate Cancer Prostatic Dis 2024. doi: 10.1038/s41391-024-00826-y.
80. Kraljevic Z, Bean D, Shek A, Bendayan R, Hemingway H, Yeung JA, *et al*. Foresight—A generative pretrained transformer for modelling of patient timelines using electronic health records: A retrospective modelling study. Lancet Digit Health 2024;6:e281–e290. doi: 10.1016/S2589-7500(24)00025-6.
81. Omiye JA, Gui H, Rezaei SJ, Zou J, Daneshjou R. Large language models in medicine: The potentials and pitfalls: A narrative review. Ann Intern Med 2024;177:210–220. doi: 10.7326/M23-2772.
82. Zhu L, Mou W, Chen R. Can the ChatGPT and other large language models with internet-connected database solve the questions and concerns of patient with prostate cancer and help democratize medical knowledge? J Transl Med 2023;21:269. doi: 10.1186/s12967-023-04123-5.
83. Wang H, Fu T, Du Y, Gao W, Huang K, Liu Z, *et al*. Scientific discovery in the age of artificial intelligence. Nature 2023;620:47–60. doi: 10.1038/s41586-023-06221-2.
84. Singla S, Eslami M, Pollack B, Wallace S, Batmanghelich K. Explaining the black-box smoothly—A counterfactual approach. Med Image Anal 2023;84:102721. doi: 10.1016/j.media.2022.102721.
85. Tao Y, Ma X, Zhang Y, Huang K, Ji Z, Fan W, *et al*. LAGAN: Lesion-Aware generative adversarial networks for edema area segmentation in SD-OCT images. IEEE J Biomed Health Inform 2023;27:2432–2443. doi: 10.1109/JBHI.2023.3252665.
86. Bulten W, Kartasalo K, Chen P-HC, Ström P, Pinckaers H, Nagpal K, *et al*. Artificial intelligence for diagnosis and Gleason grading of prostate cancer: the PANDA challenge. Nat Med 2022;28:154–163. doi: 10.1038/s41591-021-01620-2.
87. Zhang K, Zhou R, Adhikarla E, Yan Z, Liu Y, Yu J, *et al*. A generalist vision–language foundation model for diverse biomedical tasks. Nat Med 2024;30:3129–3141. doi: 10.1038/s41591-024-03185-2.
88. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring fairness in machine learning to advance health equity. Ann Intern Med 2018;169:866. doi: 10.7326/M18-1990.
89. Chen S, Kann BH, Foote MB, Aerts HJWL, Savova GK, Mak RH, *et al*. Use of artificial intelligence chatbots for cancer treatment information. JAMA Oncol 2023;9:1459. doi: 10.1001/jamaoncol.2023.2954.
90. Ueda D, Kakinuma T, Fujita S, Kamagata K, Fushimi Y, Ito R, *et al*. Fairness of artificial intelligence in healthcare: Review and recommendations. Jpn J Radiol 2024;42:3–15. doi: 10.1007/s11604-023-01474-3.

How to cite this article: Li W, Hu RY, Zhang Q, Yu ZS, Deng LX, Zhu XH, Xia YJ, Song ZJ, Cimadamore A, Chen F, Lopez-Beltran A, Montironi R, Cheng L, Chen R. Artificial intelligence in prostate cancer. Chin Med J 2025;138:1769–1782. doi: 10.1097/CM9.0000000000003689