

Review

Harnessing artificial intelligence for prostate cancer management

Lingxuan Zhu,^{1,2,3,16} Jiahua Pan,^{1,16} Weiming Mou,^{4,16} Longxin Deng,⁵ Yinjie Zhu,¹ Yanqing Wang,¹ Gyan Pareek,^{6,7} Elias Hyams,^{6,7} Benedito A. Carneiro,⁸ Matthew J. Hadfield,⁸ Wafik S. El-Deiry,⁹ Tao Yang,¹⁰ Tao Tan,¹¹ Tong Tong,¹² Na Ta,¹³ Yan Zhu,¹³ Yisha Gao,¹³ Yancheng Lai,^{1,15} Liang Cheng,^{6,14,17,*} Rui Chen,^{1,17,*} and Wei Xue^{1,17,*}

¹Department of Urology, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200127, China

²Department of Etiology and Carcinogenesis, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

³Changping Laboratory, Beijing, China

⁴Department of Urology, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

⁵Department of Urology, Shanghai Changhai Hospital, Second Military Medical University, Shanghai 200433, China

⁶Department of Surgery (Urology), Brown University Warren Alpert Medical School, Providence, RI, USA

⁷Minimally Invasive Urology Institute, Providence, RI, USA

⁸The Legorreta Cancer Center at Brown University, Lifespan Cancer Institute, Providence, RI, USA

⁹The Legorreta Cancer Center at Brown University, Laboratory of Translational Oncology and Experimental Cancer Therapeutics, Department of Pathology & Laboratory Medicine, The Warren Alpert Medical School of Brown University, The Joint Program in Cancer Biology, Brown University and Lifespan Health System, Division of Hematology/Oncology, The Warren Alpert Medical School of Brown University, Providence, RI, USA

¹⁰Department of Medical Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

¹¹Faculty of Applied Sciences, Macao Polytechnic University, Address: R. de Luís Gonzaga Gomes, Macao, China

¹²College of Physics and Information Engineering, Fuzhou University, Fujian 350108, China

¹³Department of Pathology, Shanghai Changhai Hospital, Second Military Medical University, Shanghai 200433, China

¹⁴Department of Pathology and Laboratory Medicine, Department of Surgery (Urology), Brown University Warren Alpert Medical School, Lifespan Health, and the Legorreta Cancer Center at Brown University, Providence, RI, USA

¹⁵The First School of Clinical Medicine, Southern Medical University, Guangzhou, China

¹⁶These authors contributed equally

¹⁷Senior author

*Correspondence: liang_cheng@yahoo.com (L.C.), drchenrui@foxmail.com (R.C.), uorxuewei@163.com (W.X.)

<https://doi.org/10.1016/j.xcrm.2024.101506>

SUMMARY

Prostate cancer (PCa) is a common malignancy in males. The pathology review of PCa is crucial for clinical decision-making, but traditional pathology review is labor intensive and subjective to some extent. Digital pathology and whole-slide imaging enable the application of artificial intelligence (AI) in pathology. This review highlights the success of AI in detecting and grading PCa, predicting patient outcomes, and identifying molecular subtypes. We propose that AI-based methods could collaborate with pathologists to reduce workload and assist clinicians in formulating treatment recommendations. We also introduce the general process and challenges in developing AI pathology models for PCa. Importantly, we summarize publicly available datasets and open-source codes to facilitate the utilization of existing data and the comparison of the performance of different models to improve future studies.

INTRODUCTION

Prostate cancer (PCa) is the leading cause of cancer-related deaths in males, with an annual global incidence of approximately 1,414,259 new cases.¹ Diagnosis typically involves conducting a biopsy after an initial suspicion of PCa based on an elevated prostate-specific antigen (PSA) level. Treatment plans for localized PCa are primarily established based on pathological examination of the biopsy. Pathologists determine whether the patient has PCa and assign a Gleason grade group (GG) by evaluating hematoxylin and eosin (H&E)-stained sections of the tissue.^{2,3} A multi-disciplinary team of urologists, medical oncologists, and radiation

oncologists then determine in a collaborative manner based on the biopsy pathology report and radiographic and clinical features whether the patient should pursue active treatment, such as radical prostatectomy (RP) or radiation therapy, or opt for active surveillance (AS). The pathology report following RP is crucial for determining the patient's postoperative prognosis, predicting the likelihood of biochemical recurrence (BCR), and guiding the selection of adjuvant therapy or other treatment options.^{2,3}

Each patient typically has at least 12 needle biopsies, resulting in over 15 million biopsy samples per year worldwide, highlighting the tremendous workload for pathologists. The



Table 1. AI tool for pathological analysis of prostate cancer approved by the US FDA or that have received the CE mark

Name of the device	Company	Functions	Country	Related work
Paige Prostate Detect ^a	Paige.AI	detecting tumors in prostate needle biopsies	US	Raciti et al., ¹² Perincheri et al., ¹³ da Silva et al., ¹⁴ Campanella et al., ¹⁵ Raciti et al. ¹⁶
Paige Prostate Grade & Quantify	Paige.AI	Gleason grading and quantification, total tumor percentage, and tumor length measurement	US	Eloy et al. ¹⁷
Aiforia Clinical AI Model for Prostate Cancer	Aiforia	cancer detection and Gleason grading	Finland	Sandeman et al. ¹⁸
DeepDx-Prostate Connect	Deep Bio	ROI detection, Gleason grading, and quantification	South Korea	Jung et al. ^{19,20} and Ryu et al. ^{19,20}
Galen Prostate	Ibex Medical Analytics	cancer detection and Gleason grading	Israel	Pantanowitz et al. ²¹
HALO Prostate AI	Indica Labs	prostate cancer detection and Gleason grading	US	Tolkach et al. ²²
INIFY Prostate	Inify Laboratories	cancer detection	Sweden	Vazzano et al. ²³

^aApproved by the US FDA.

shortage of pathologists suggests an even greater workload associated with PCa diagnosis.⁴ In addition, the pathology review of PCa primarily relies on subjective assessments, with reported low inter-observer consistency.^{5–7} This inconsistency may lead to under-treatment of aggressive cancers and over-treatment of indolent cancers. As a result, researchers are attempting to introduce artificial intelligence (AI) into the pathology review of PCa, aiming to attain expert-level and reproducible outcomes.

AI has demonstrated promising capabilities in medicine, especially in the analysis of medical images.^{8–10} In PCa pathology, the US Food and Drug Administration (FDA) has approved the Paige Prostate for determining between benign prostate biopsies and suspicious PCa, and six other AI products have received European Conformity (CE) certification (Table 1). This indicates that AI capabilities in PCa pathology analysis have begun to be recognized by regulatory agencies. Nonetheless, the current repertoire of AI products approved for PCa pathology analysis remains limited, and their capabilities are comparatively rudimentary. With the advent of high-throughput automated pathology slide scanners, complete digitization of pathology laboratories is inevitably going to become a reality in the near future.¹¹ We believe that AI pathology will have a significant impact on the management of PCa.

This review focuses on advanced AI research with potential for translation in PCa pathology analysis and emerging research trends, highlighting how AI can collaborate with pathologists and urologists to improve patient care. In addition to common tasks such as automated malignancy detection and automated Gleason grading, we also review cutting-edge developments in predicting patient outcomes, molecular phenotypes, and uncovering possible molecular mechanisms of the disease. We also introduce the general process and challenges in developing AI models in this field. Importantly, we summarize publicly available prostate pathology image data-

sets and open-source prostate pathology AI model codes, enabling researchers to fully utilize existing data and compare performance against other models, thereby improving the quality of research.

DEVELOPMENT OF AI MODELS FOR PROSTATE CANCER MANAGEMENT

AI can be broadly divided into two categories: traditional machine learning (ML) and deep learning (DL). The key difference is that the former relies on hand-crafted features (such as color and texture), while the latter utilizes features automatically extracted by convolutional neural networks (CNNs), thus enabling the detection of information that is difficult for humans to obtain. For example, in the development of a DL model for automated Gleason grading, AI learns from annotations provided by pathologists. Different training strategies require different levels of annotation information.²⁴ Supervised learning requires pathologists to manually draw every gland on the whole-slide images (WSIs) and provide Gleason pattern (GP) information, which is time consuming. Weakly supervised learning typically uses slide- or patient-level Gleason scores recorded in pathology reports as inputs, without the need for additional annotations. Semi-supervised learning leverages a small set of precisely labeled data alongside larger unlabeled datasets to improve performance. Self-supervised learning generates its own supervisory signal from unlabeled data by defining pretext tasks, then fine-tuning the model on downstream tasks, allowing it to exploit large unlabeled datasets. Generally speaking, these strategies, with the exception of supervised learning, require larger datasets. Weakly supervised learning is currently the most popular strategy.

Due to the large size of images, WSIs are preprocessed and split into small patches as model input. After its development, the AI model is validated on independent external validation

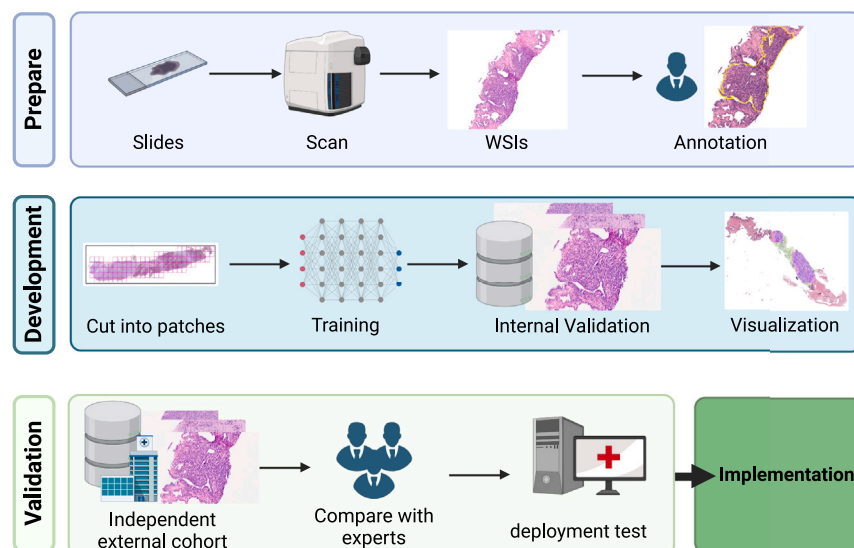


Figure 1. The development process of pathology AI models, using automated Gleason grading as an example

sets to assess its generalization ability and is compared with human pathologists to determine whether it has reached or exceeded the performance of expert pathologists. Furthermore, prospective deployment testing is conducted to validate its real-world performance and impact on clinical practice. Thus, a prostate AI pathology tool is completed from development to validation and eventually moves toward clinical application (Figure 1).

AI for the identification of benign and malignant lesions

AI has made remarkable achievements in identifying benign and malignant WSIs (Table 2). One example is Paige Prostate, which has been approved by the FDA for clinical application. This model used multiple instance learning - recurrent neural network (MIL-RNN) technology to train on 12,132 WSIs with the diagnosis of pathology report as labels.¹⁵ In their study, the model achieved an area under the curve (AUC) of 0.99 on the test set and 0.93 on an independent external validation set of over 12,000 slides.¹⁵ In a user study, the sensitivity of three board-certified pathologists in diagnosing PCa was significantly improved from 74% to 90% with the assistance of the AI.¹² The model also performed excellently on an independent external cohort of 600 slides from 100 consecutive patients¹⁴ and another cohort of 1,876 prostate core biopsy WSIs,¹³ demonstrating its good generalizability.

In contrast to previous studies that used retrospectively collected samples, Ström et al. developed a model based on a prospective, population-based diagnostic study cohort (STHLM3 study) and achieved an AUC of 0.986 in the validation set, even when facing various difficult cases from the real world.³⁷ Unlike using AI to directly provide a diagnosis, Dov et al. designed a hybrid human-machine approach in which the AI identified the top 20 regions of interests (ROIs) with the highest malignancy probability for each biopsy. Pathologists were then able to make negative biopsy diagnoses by examining only these ROIs while reserving potentially malignant biopsies for further examination. This approach yielded an outstanding sensitivity

of 99.2% and filtered out approximately 70% of all benign WSIs.³⁹

Prostate biopsy can sometimes miss cancerous areas, leading to missed diagnosis. Liu et al. hypothesized that there may be tumor-induced morphological changes in benign cores sampled in the vicinity of cancerous regions. They employed CNNs to uncover these subtle morphological variations and successfully detected 70% of GG 5 PCa cases and approximately 30% of low-grade PCa cases by analyzing only the benign biopsy WSIs and incorporating relevant patient clinical information.

This can help identify high-risk PCa cases missed by regular biopsies without increasing overdiagnosis risk for low-grade lesions.⁴⁰

AI for the grading of prostate cancer

Gleason grading is a powerful prognostic factor for PCa and is crucial for treatment decision-making. Automated Gleason grading with AI is practical but challenging. Numerous studies have shown that AI algorithms surpass non-uropathologists and achieve results that are comparable to those of expert-level uropathologists^{19,37,41–43} (Table 2).

Numerous AI challenges have promoted development in this field, with the most influential being the PANDA Challenge.³⁰ Algorithms developed using 10,616 prostate biopsy samples from multiple centers achieved expert-level performance in two independent cross-continental validation sets, reaching concordance of 0.862 and 0.868 with expert uropathologists, demonstrating a robust performance of the AI models that is comparable to that of pathologists.³⁰ Some researchers have proposed fine-grained Gleason grading concepts, such as using the prediction scores of DL system to subdivide the GP into GP 3.5 and GP 4.5, which achieved more accurate risk stratification.⁴³

On the other hand, as current technology cannot perfectly assign each patient to a specific group, many studies opt to divide patients into low-grade or high-grade (GG1 vs. GG2-5 or GP 3 vs. GP \geq 4) based on clinical treatment implications.^{2,21,35,38,42,44–48} representing whether patients are suitable for AS or RP. For example, Pantanowitz et al.'s model achieved an AUC of 0.941 in distinguishing between high- and low-grade PCa in an external validation set.²¹ In another study, Bulten et al.'s model achieved AUCs of 0.878 and 0.869 in distinguishing between benign and GG 1 vs. GG \geq 2 in two external validation sets.³⁸

As the field progresses, weak supervision technology has emerged as a promising solution to free pathologists from time-consuming pixel-level annotation and fully leveraging

Table 2. Studies on prostate cancer detection and automated Gleason grading models with independent external validation

Author	Purpose	Methods	Type	Dataset size	Performance on external dataset
Vazzano et al. ²³	external validation of the INIFY Prostate model (detection and quantification of tumor area)	CNN	biopsy	external: 30 patients, 30 slides from two centers	detection: specificity 0.97 sensitivity 0.994 quantification of tumor area: correlation coefficient 0.76–0.80
Yang et al. ²⁵	improving the recognition of small tumor regions (detection of prostate cancer)	intensive-sampling MIL	biopsy	training: PANDA dataset external: HEBEI (100 patients, 844 slides) and DiagSet-B dataset	for all cases, AUC 0.9711 and 0.9822; F1 score: 0.8410 and 0.9709 for hard cases, AUC 0.7348 and 0.9707; F1 score 0.6669 and 0.7957
Xiang et al. ²⁶	detection and automated GG grading (0–5)	GCN-MIL	biopsy	training: PANDA dataset external: HEBEI (100 patients, 844 slides) and DiagSet-B dataset	detection: AUC 0.985 and 0.986 GG grading [#] : κ , 0.723; κ_{quad} 0.801 (only HEBEI dataset was used for validation in this task)
Oner et al. ²⁷	detection of prostate cancer	multi-resolution R-CNN-ResNet-18, supervised learning	biopsy	training: PANDA dataset external: 280 cores, 46 patients	AUC: 0.992
Jung et al. ¹⁹	external validation of the DeepDx ²⁰ model (detection and Gleason grading)	DNN, supervised learning	biopsy	external: 593 slides	detection: accuracy, 0.9831 GG grading [#] : κ 0.713, and κ_{quad} 0.922 GS grading [#] : κ 0.654, and κ_{quad} 0.904 GG1 or normal vs. GG2-5: Accuracy: 0.9376
Singhal et al. ²⁸	detection and automated GG grading	CNN, semi-supervised learning	biopsy	training: Local (580 slides), PANDA (Radboud) dataset external: PANDA (Karolinska) dataset	detection: AUC: 0.92 GG grading [#] : accuracy: 0.831; κ_{quad} 0.93 GG2 vs. GG3-5: AUC: 0.93
Pohjonen et al. ²⁹	improve generalization problem for the detection of prostate cancer	neural network trained with spectral decoupling	–	training: 90 patients from Helsinki external: PESO dataset	networks trained with spectral decoupling achieve up to 9.5% point higher accuracy on external datasets (The author did not report the exact value of accuracy.)
PANDA challenge ³⁰	detection and automated GG grading	various DL algorithms	biopsy	training: PANDA dataset external: two cohorts (714 and 330 slides)	detection: sensitivity 0.986 and 0.0.977; specificity 0.752 and 0.843 GG grading [#] : κ_{quad} 0.862 and 0.868
Silva-Rodríguez et al. ³¹	detection and automated Gleason grading	CNN, self-supervised learning	biopsy and RP	training: PANDA dataset external: SICAP, ARVANIT and GERTYCH datasets	detection: sensitivity 0.7568–0.9389; precision 0.9375–0.9720 GP grading [#] : accuracy 0.7226–0.8305; F1 score 0.7319–0.8202; κ_{quad} 0.7930–0.8303 GS grading [#] : κ_{quad} 0.8054–0.8299 GG grading [#] : κ_{quad} 0.8254–0.8854

(Continued on next page)

Table 2. Continued

Author	Purpose	Methods	Type	Dataset size	Performance on external dataset
Mun et al. ³²	detection and automated GG grading	weakly supervised DL	biopsy	training: HUMC 621 cases, 6,071 slides; KUGH 167 cases, 1,529 slides external: Gleason 2019 dataset	inter-institutional*: detection AUC 0.982 GG grading#: accuracy 0.674, κ 0.553, κ_{quad} 0.880 external: detection AUC 0.943 GG grading#: accuracy 0.545, κ 0.389, κ_{quad} 0.634
Li et al. ³³	detection of prostate cancer	multi-resolution MIL	biopsy	training: local (830 patients, 20,229 slides) external: SICAP-V1 dataset	AUC 0.994
Silva-Rodríguez et al. ³⁴	automated GP grading	CNN, supervised learning	biopsy	training: SICAPv2 external: ARVANIT and GERTYCH datasets	GP grading#: accuracy 0.5136–0.5861; F1 score 0.4753–0.5702; κ_{quad} 0.5116–0.6410
Nagpal et al. ³⁵	automated GG grading [†]	CNN, supervised learning	biopsy	training: 360 cases, 524 slides external: 322 slides	agreement rate: 0.801 (all biopsies) and 0.714 (tumor only)
Pantanowitz et al. ²¹	detection of prostate cancer and perineural invasion and automated Gleason grading	CNN, supervised learning	biopsy	training: 549 slides external [‡] : 100 consecutive cases, 1,627 slides	detection: AUC 0.991 GS 6 or ASAP vs. GS 7–10: AUC 0.941 GP3–4 or ASAP vs. GP5: AUC 0.971 perineural invasion: AUC 0.957
Tolkach et al. ²²	detection and automated GG grading	CNN, supervised learning	RP	training: TCGA-PRAD dataset external 1 ²² : 2 cohorts (592 and 279 patients) external 2 ²⁶ : 7473 cores, 423 patients from five centers	external 1: detection: AUC 0.9919 and 0.9918 GG grading [†] : κ 0.51–0.66 external 2: detection: sensitivity 0.971–1.000; specificity 0.875–0.976 grading: κ_{quad} 0.72–0.77
Ström et al. ³⁷	detection, measurement of tumor length and automated grading	DNN, supervised learning	biopsy	training: 1,069 patients, 6,953 slides external: Imagebase and Karolinska dataset (330 cores from 73 cases)	detection: AUC 0.986 GG grading: for Imagebase dataset, mean pairwise κ 0.62. for Karolinska dataset [#] , κ 0.70 and 0.76 (after calibrating). cancer length: correlation 0.87
Bulten et al. ³⁸	detection and automated GG grading	CNN, semi-supervised learning	biopsy	training: 1,243 cases, 5759 slides external: 245 tissue microarray cores	detection: AUC 0.985 GG grading#: κ_{quad} 0.715 GG1 or normal vs. GG2–5: AUC 0.875 GG1–2 or normal vs. GG3–5: AUC 0.875

(Continued on next page)

Table 2. Continued

Author	Purpose	Methods	Type	Dataset size	Performance on external dataset
Campanella et al. ¹⁵	detection of prostate cancer	MIL	biopsy	training: 836 patients, 12,132 slides external 1 ¹⁵ : 6,323 patients, 12,727 slides external 2 ¹³ : 1,876 cores from 118 consecutive patients external 3 ¹⁴ : 600 cores from 100 consecutive patients	external 1: AUC 0.932 external 2: sensitivity 0.977; PPV 0.979; specificity 0.993; NPV 0.992 external 3: sensitivity 0.99; NPV 1.0; specificity 0.93

Some studies included more than one task; we only collected the tasks that had independent external validation in the table. For specific methodological details, please refer to the original papers. GP, Gleason pattern; GS, Gleason score; GG, Gleason grade group; GG 0, benign; ASAP, atypical small acinar proliferation; NC, non-cancer; AUC, area under the receiver operating characteristic curve; PPV, positive predictive value; NPV, negative predictive value; μ IoU, mean intersection over union; κ_{quad} , quadratic-weighted kappa; MIL, multiple instance learning, belongs to weak supervised learning; MLP, multi-layer perceptron; GCN, graph convolution network; CNN, convolutional neural network; DNN, deep neural network; HUMC, dataset from Hanyang University Medical Center, Korea; KUGH, dataset from The Fourth Hospital of Hebei Medical University, China; HEBEI dataset contains 844 WSIs from 100 patients, which are not available to the public. ^aWhen reporting performance, the benign category is included. For example, for GG grading, benign and GG1-5 are included for a total of six categories. ^bWSIs from HUMC were used for the discovery, and WSIs from the KUGH were used for the validation. ^cGG4 and GG5 were combined into one class in this study. ^dThirty-two cases from this external center were used to calibrate the model prior to external validation.

the large-scale diagnostic labels available in pathological reports.^{15,31–33,49–52} Additionally, various new techniques have been proposed to improve AI model performance, such as knowledge distillation, deep quantum ordinal regression, and pyramid semantic parsing network.^{53–55}

AI APPLICATIONS BEYOND DIAGNOSIS AND GRADING

AI can effectively tackle various tasks beyond diagnosis and grading. For example, automatic measurement of cancer length and volume,^{20,23,37,56} quantification of GP percentage,^{43,57,58} recognition and quantification of perineural invasion,^{21,59} quantification of immunohistochemistry (IHC) staining,^{60,61} and detection and quantification of cribriform pattern.^{34,62,63} AI models have also been developed to assess tumor purity of PCa using frozen H&E-stained slides.⁶⁴ AI has significantly improved consistency and reproducibility of these tasks. AI can also streamline pathology laboratory workflows by identifying ambiguous cases that require additional IHC-stained examination and automatically requesting this before pathologist review, thus reducing turnaround times.⁶⁵ AI can also perform fully automatic quality control of WSI scans and identify scans that may require re-scanning or re-staining to improve quality.⁶⁶ In a cutting-edge field, Rana et al. attempted to add computationally generated H&E staining to unstained prostate biopsy images and found that the algorithm-generated images could accurately replicate prostate tumor characteristics and be used for pathological diagnosis, thus supporting early detection of abnormalities in non-stained tissue biopsies.⁶⁷

EMPOWERING PATHOLOGISTS: AI AS A COLLABORATIVE TOOL

Due to legal, ethical, and other considerations, it is unlikely that AI will replace human pathologists in the near future. However, AI can assist in standardization and quality control in pathological analysis^{68–70} (Figure 2). Multiple studies have shown that the performance of AI in conjunction with human pathologists is superior to either alone, and it can improve diagnostic accuracy and consistency, shorten diagnostic time, and reduce over-grading and over-treatment.^{12,57,71,72} Since a considerable proportion of prostate biopsies are benign,⁷³ AI as a first reader can help pathologists pre-screen and exclude benign biopsies, allowing them to focus on those with higher risk of malignancy. AI can also assist pathologists in quickly locating suspicious areas, thus improving diagnostic speed. Eloy et al. found that, with the assistance of AI, pathologists can reduce need for IHC and second opinions while maintaining accuracy.¹⁷ AI can also assess the difficulty of a case and prioritize difficult samples to be assigned to expert-level pathologists.⁷⁴

As a second reader, AI serves as an unbiased reviewing tool, marking samples that differ in results between pathologists and AI,¹³ and preventing pathologists from missing small tumor areas.^{15,16,28} In one study, AI assistance increased the sensitivity for detecting prostate cancer smaller than 0.6 mm from 46% to 83%.¹² Pantanowitz et al. developed a system that alerts pathologists to slides where AI detects potential cancer that was missed by the pathologist or indicates slides that were

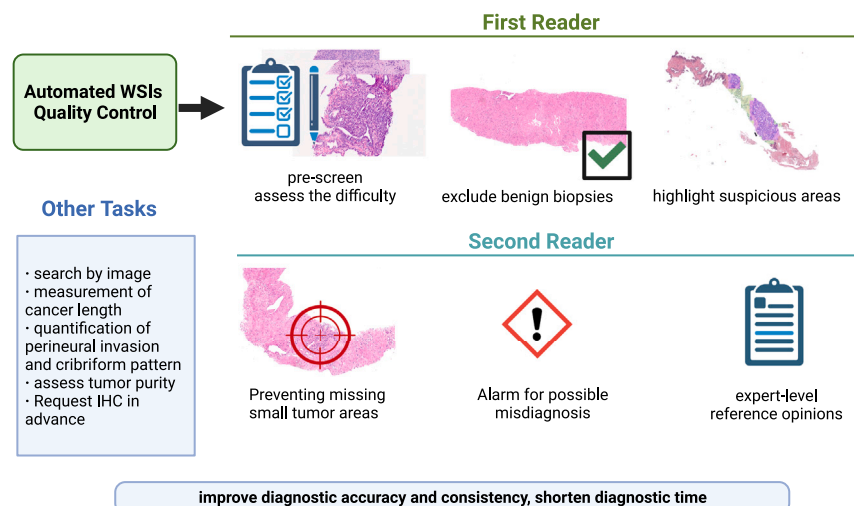


Figure 2. Interaction between pathology AI models and pathologists

on WSIs from PCa patients, achieving outstanding performance in predicting prostate cancer-specific mortality and 5-year distant metastasis events with AUCs of 0.77 and 0.83, respectively.

Biochemical recurrence (BCR) is a clinical endpoint that patients with PCa may encounter in a relatively short time. AI has achieved promising results in directly predicting BCR from WSIs, outperforming traditional clinical indicators such as GS in several studies.^{83–90} For instance, Yamamoto et al. applied an unsupervised deep neural network (DNN) model to

diagnosed as Gleason score (GS) 6 by the pathologist but had a higher GS determined by AI.²¹ This system allows pathologists to view WSIs with AI-flagged regions for re-review, and it successfully identified a missed prostate cancer case during real-world deployment.²¹ Furthermore, AI systems with expert-level performance can provide general pathologists with expert-level reference opinions to assist in diagnosing challenging cases.^{19,72} Therefore, the second review provided by AI can improve diagnostic consistency, help achieve standardized grading, and bring expert-level grading to areas with limited medical resources.^{37,75}

AI can also enhance the utility of histopathology image datasets within a medical institution by providing a “search-by-image” function. Pathologists can match images from the dataset to the current patient and relevant pathological reports for decision support. This is especially helpful for rare or unusual cases.^{76–78}

AI for prediction of prognosis

One of the primary purposes of Gleason grading is to predict patient prognosis. Some studies have found that AI’s GSs are more effective in stratifying patient risk.^{22,42,43} In addition, Wulczyn et al. developed a model based on WSIs from RP specimens to quantify the percentages of GP4 and GP5. They then used a Cox proportional hazards regression model based on the obtained GP percentages to generate an AI risk score for predicting prostate cancer-specific mortality. The model achieved a C-index of 0.87, outperforming the predictive ability of the GG system (C-index of 0.75) and effectively stratified patient outcomes.⁷⁹ These results demonstrated that AI could surpass the accuracy of commonly used clinical models for risk prediction and stratification.

Furthermore, the goal of AI should not be limited to replicate pathologists’ assessments. Pathology slides contain a large amount of biological information not covered by the Gleason grading system. Instead of having AI learn the inherently biased Gleason grading system, it may be better to train AI directly on WSIs using long-term follow-up data to guide risk stratification.^{80,81} For example, Esteva et al.⁸² developed a model based

extract features from unannotated post-RP H&E slides to predict 1-year BCR, achieving an AUC of 0.845 in an external validation set, which was better than using only GS (AUC = 0.721).⁸⁷ Interestingly, combining AI with GS yielded the best performance (AUC = 0.884).⁸⁷

AS rather than RP is becoming the preferred management approach for low-risk PCa.⁹¹ Tsuneki et al. developed a DL model to determine whether patients should choose AS or RP based on the proportions of GP 4 and 5, achieving an AUC of 0.84 for identifying indolent tumors suitable for AS.⁹² However, accurately identifying AS patients at higher risk of disease progression remains challenging. Chandramouli et al. used WSIs from needle-core biopsy to predict progression in AS patients, achieving an AUC of 0.75, outperforming pro-PSA and Gleason.⁹³ AI can also predict other PCa prognosis indicators like lymph node metastasis (AUC = 0.68),⁹⁴ the development of castration-resistant prostate cancer (CRPC) after combined androgen blockade (CAB) therapy (sensitivity and specificity both 87.5%),⁹⁵ and adverse pathology outcomes (AUC = 0.72).⁹⁶ However, relying solely on morphological analysis may be insufficient for long-term outcome predictions, as randomness of mutations and the potential for escaped cells to acquire additional mutations can affect recurrence.^{68,86}

AI for prediction of molecular subtypes

In addition to pathological features, genetic sequencing has identified molecular markers associated with PCa progression. However, the cost and complexity have limited its clinical application. Genetic changes in tumor cells can lead to changes in their function, which can in turn affect the morphology of tumor cells.⁹⁷ The connection between the changes in tumor molecular composition and morphology can be found in almost every cancer and nearly every class of molecules,⁹⁸ providing a foundation for using AI to predict molecular characteristics from WSIs. Compared to genomic biomarkers, AI-based digital pathology systems are easier to obtain the input, more cost-effective, and more scalable for clinical application. For example, TMPRSS2-ERG gene fusion is a commonly observed genetic alteration in PCa, which is

associated with aggressive PCa⁹⁹ and correlates with certain morphological characteristics of PCa.¹⁰⁰ Dadhanian et al. established a CNN-based model to predict the presence of TMPRSS2-ERG gene fusion with WSIs of PCa patients, with an AUC of 0.82–0.84.¹⁰¹ Similarly, Erak et al. used a self-supervised DL model to predict ERG gene fusion and PTEN loss status with average AUCs of 0.83 and 0.76 over multiple independent external validation sets.¹⁰² The model can even predict the spatial patterns of subclonal PTEN deletions within tumor nodules. The SPOP mutation is a specific subtype of PCa. Schaumberg et al. used a ResNet model to determine the SPOP mutation status in PCa patients, achieving an AUC of 0.7589 in an external validation cohort.¹⁰³ Furthermore, a pan-cancer DNN model can predict TP53 and FOXA1 mutation status in PCa patients based on WSI, with AUCs of 0.685 and 0.762, respectively.⁹⁷

Furthermore, AI can predict gene expression from pathology images. Weitz et al. proposed a multioutput CNN model based on gene co-expression patterns to predict gene expression in PCa from WSIs. This model identified associations between morphology and gene expression for 5,419 genes.¹⁰⁴ To assess the clinical utility of the model, they calculated the CNN model's Prolaris Cell Cycle Progression (CCP) score, which is a commercial prognostic test based on gene expression values to assess the risk and progression of PCa. The CNN-derived CCP score achieved an AUC of 0.73 when discerning whether the RNA sequencing (RNA-seq)-based CCP score was above or below the median. In summary, the utilization of AI for predicting the molecular subtype of PCa can be particularly beneficial in low-resource settings, where molecular diagnostics are not readily available. However, the lack of large-scale PCa cohorts with both high-quality WSIs and corresponding genomic sequencing data poses a challenge for further validating the potential of AI for this task.

AI can also be applied to uncover connections between morphological changes and potential molecular alterations. Huang et al. used an AI model's morphological scoring system to identify ROIs associated with early recurrence of PCa.⁸⁹ Subsequently, multiplexed IHC was used to compare the expression of biomarkers within the high- and low-scoring ROIs identified by the AI model. They discovered that the high expression of TMEM173 within high-scoring ROIs may be a potential biomarker for driving early recurrence in PCa patients.⁸⁹ We believe that by further combining spatial omics techniques, we could uncover more molecular mechanisms behind the ROIs identified through automated AI feature extraction. Additionally, spatial omics annotates each region of pathology images with corresponding genetic information, which potentially refines AI's accuracy in predicting gene mutations and expressions. This advancement could also empower AI models to predict the spatial distribution of diverse molecular subtypes within tumors, thus providing deeper insights into intratumoral heterogeneity (ITH). AI can also help us understand ITH by visualizing the distribution of different GPs within the tumor and identifying patterns of tumor-infiltrating lymphocyte (TIL) distribution.¹⁰⁵ In conclusion, AI has the ability to go beyond simple morphological analysis and can extract more information from pathology slides than humans.

CHALLENGES OF APPLICATION OF AI IN CLINIC

Difficulty in generalization

Owing to the heterogeneity among datasets and the risk of overfitting, AI models may perform less effectively when applied to new datasets, limiting their widespread use.^{30,50} Heterogeneity can stem from various factors, such as staining variations, artifacts, and imaging differences between scanners.^{29,75,106,107} To overcome this, an ideal approach is to have a sufficiently large and diverse training set, such as continuously collecting all cases over a certain period of time from multiple institutions, in order to cover all possible variations in the real world to represent the entire target population. From the perspective of fully utilizing the existing data, data augmentation techniques such as rotation, flipping, and color enhancement can be applied to enhance the original training set.^{48,51} Alternatively, introducing histological artifacts into the training data¹⁰⁷ or fine-tuning model parameters using a subset of samples from the new dataset can help improve the generalization ability of the model.⁶¹ Additionally, tools such as color^{108,109} or style normalization¹¹⁰ can make new data more similar to the training set. For example, after using style normalization, the AUC of the AI model on an independent external validation set increased from an average of 0.875 to 0.975, which is close to the results of the training set cross-validation (AUC = 0.98),¹¹⁰ indicating that the generalization gap of the model can be compensated for by appropriate preprocessing. In summary, these methods can strengthen the generalization ability of AI models in the real world.

Lack of high-quality open-source data

Unlike many other medical imaging techniques, routine pathology workflows are rarely fully digital, requiring additional scanning to obtain WSIs. Moreover, publicly available datasets for PCa WSIs are scarce and stored sporadically, hindering AI progress in this area. AI models applied to PCa pathology analysis should be evaluated on independent test sets from different institutions; otherwise, it is difficult to assess whether AI can perform well in complex external clinical scenarios.¹¹¹ Despite an increasing number of relevant articles in recent years, the deployment of AI tools in clinical practice remains limited (Table 1). Many studies focused solely on algorithmic improvement using data from a single institution, lacking external validation or comparison with human pathologists. Additionally, variations in datasets and metrics used to evaluate model performance hinder direct comparisons among different studies. Therefore, we summarized the publicly available WSIs datasets of PCa (Table 3) and the source code of several AI models (Table S1), facilitating model validation with external datasets and enabling comparisons with established models.

Difficulty in obtaining ground truth

Owing to inter-observer heterogeneity in the annotation of prostate WSI data among experts, it is difficult to obtain a perfect set of labels in practice. To ensure the label quality, a committee of experts is often relied upon.¹³⁷ Training and evaluating an AI model based solely on one pathologist's annotations may result in inaccurate ground truth or bias toward that pathologist's grading habits, leading to poor performance in

Table 3. Publicly available prostate pathology image datasets

Dataset	Data type	Sample type	Dataset size	Support data
PANDA Challenge ^{30,112}	WSI	biopsy	2,113 patients, 10,616 slides	annotations of stroma, benign, and GPs 3, 4, and 5 for data from Radboud; benign and cancerous tissue for data from Karolinska
TCGA-PRAD ¹¹³	WSI	RP	403 patients, 449 slides	clinical, various sequencing data, pathological report
SICAP-MIL ⁴⁹	WSI	biopsy	271 patients, 350 slides	annotations of GPs WSI-level GS including both primary and secondary GPs
Ibex ²¹	WSI	biopsy	210 patients, 2,501 slides	associated IHC slides
PESO ^{114,115}	WSI	RP	102 patients	annotations of benign and cancerous tissue
DiagSet ¹¹⁶	WSI and Patches	biopsy	5,151 slides	set A: patch-level annotations for 2.6 million patches extracted from 430 WSIs set B: slide-level annotations of cancer or non-cancer set C: slide-level annotations of cancer, non-cancer, or need IHC from nine pathologists
NADT-Prostate ¹¹⁷	WSI	biopsy and RP	39 patients, 1,404 slides	clinical, exome, genome, RNA-seq, slides stained with antibodies against p53, PTEN, AR, PSA, GR, Ki67, SYP, and PIN4-cocktail (p63 + CK5 + K18 + AMACR)
Gallo ¹¹⁸	WSI	biopsy	167 patients, 787 slides	annotations of tumor regions
PAIP ¹¹⁹	WSI	RP	600 slides	not reported
AGGC22 ¹²⁰	WSI	biopsy and RP	241 slides	pixel-level annotations of GPs 38 slides scanned by multiple scanners
PathPresenter ¹²¹	WSI and TMA	biopsy and RP	208 slides	slide-level diagnosis
Prostate Fused-MRI-Pathology ¹²²	WSI	RP	16 patients, 114 slides	annotations of cancer presence, MR
Tolkach ³⁶	WSI	biopsy	100 biopsies from two centers	GG grading from 10 or 11 pathologists
Oner ²⁷	WSI	biopsy and RP	99 slides	GS
PAIP 2021 Challenge ¹²³	WSI	RP	80 slides	annotations of perineural invasion
STHLM3 ^{37,124}	WSI	biopsy	60 slides are available	Gleason group, predictions of GPs generated by the AI
CMB-PCA ^{125*}	WSI	biopsy	eight patients, 10 slides	clinical, CT, MR, NM, genomic, phenotypic
Zhong ¹²⁶	TMA and WSI	RP	71 cores and two WSIs	clinical, PTEN DISH
Tolkach ²²	patches	RP	592 patients from Charite University Hospital and 279 patients from UKB	patch-level annotations of benign or cancerous tissue
Schömig-Markieffka ^{107,127,128}	patches	not reported	15, 18, and 51 patients from three centers, respectively	Each dataset contains 50,000 patches with tumor tissue, 50,000 patches with nonneoplastic glandular prostate tissue, and 20,000 patches with nonglandular tissue.
Prostate-MRI ¹²⁹	patches	RP	26 patients	MR
Arvaniti ^{42,130}	TMA	not reported	886 cores from five TMAs	clinical, annotations of GPs from two pathologists
Gleason 2019 Challenge ¹³¹	TMA	RP	231 patients, 331 cores	annotations of GPs from six pathologists
RINGS ^{132,133}	Images	biopsy	1,500 images extracted from WSIs of 150 patients	annotations of the contours of the glands

(Continued on next page)

Table 3. Continued

Dataset	Data type	Sample type	Dataset size	Support data
Gertych ¹³⁴	images	RP	210 images extracted from WSIs of 20 patients	annotations of stroma, benign epithelium, GP 3 or 4
Imagebase ¹³⁵	images	biopsy and RP	120 cases	GS reviewed by members of an international panel of 24 experts in each of the main fields of urological pathology
Kweldam ¹³⁶	images	not reported	60 cases (in supplemental information)	23 genitourinary pathologists' annotations of predominant GP per case (3, 4, or 5), and to indicate the predominant GP 4 growth pattern, if present

PRAD, prostate adenocarcinoma; RP, radical prostatectomy; WSI, whole-slide imaging; TMA, tissue microarray; NM, nuclear medicine; UKB, The University Hospital Bonn. *The dataset is ongoing. Note: different datasets have different terms of use. You should comply with the respective terms of use before conducting research.

external evaluations and limiting its ability to surpass the performance of that expert.^{30,45} Therefore, it is beneficial to train AI models on datasets with annotations from multiple experts, such as Imagebase¹³⁵ and PANDA,³⁰ enabling the model to learn more accurate ground truth.⁴⁵ Furthermore, evaluating the performance of the model by comparing it with annotations from multiple experts enables a more realistic assessment of its capabilities.⁴⁵ When faced with disagreements in expert annotations, the majority voting method is typically used to determine final labels,¹⁹ and the STAPLE algorithm can be used to merge annotations from multiple experts to construct the final ground-truth label.^{48,55} Some studies also employed automatic label-cleaning methods to refine annotations iteratively.³⁰ Additionally, Arvaniti et al. found that, due to the limited accuracy of manual annotations, pathologists sometimes include stromal tissue in the training area when annotating GP 3, causing AI to misidentify stromal regions as GP 3 during predictions.⁴² To improve annotation accuracy, it would be beneficial to pre-segment the glandular regions and then allow pathologists to specify the GPs.¹¹⁵

Poor interpretability of AI

One major drawback of AI pathology is its lack of interpretability. Many researchers attempted to enhance the interpretability of their models.¹³⁸ One common method is to use gradient-weighted class activation mapping (Grad-CAM) to visualize the ROI that the AI model focuses on.⁴² Another method involves pathologists reviewing AI model feature clustering results. In a study on predicting BCR, pathologists discovered that the model identified features of non-cancerous stroma as prognostic factors after reading representative features, which are not typically evaluated in prostate pathology analysis.⁸⁷ Pinckaers et al. used automatic concept explanations (ACEs) to explain which image features in their model were used to make decisions for predicting BCR.⁸⁶ Therefore, enhancing the interpretability of AI models may discover new histopathological features and increase our understanding of this disease.

In addition, introducing the common errors and advantages of AI to pathologists can help them better handle the predictions given by AI and alleviate potential over-reliance.^{16,47,57} Toledo-Cortés et al. quantified the uncertainty of AI predictions in their model, avoiding the limitations of previous models that only pro-

vided final prediction results. This can serve as a quality control tool for AI-based diagnoses, allowing pathologists to decide whether to trust the model's prediction.^{54,80}

Difficulty in tackling rare conditions, tumor heterogeneity, and ethnic variability

Although most PCa are acinar adenocarcinomas, there are many rare cases and confounding factors in clinical practice, such as inflammation, atrophy, atypical small acinar proliferation, atypical intraductal proliferation, and some diagnostically challenging histological subtypes (such as ductal and intraductal carcinoma of the prostate, prostatic adenocarcinoma with neuroendocrine differentiation), and treatment-related changes. During AI training, most studies excluded these rare conditions, or had limited cases, posing challenges when encountering such cases in real-world scenarios.^{13,16,30,33,39,68,72,139} Some researchers have proposed using generative adversarial network (GAN) technology to synthesize high-fidelity pathological images to compensate for the small sample size of certain types in the training dataset.^{64,140} Falahkheirkhah et al. employed this technology to synthesize PCa pathology images, and they mixed synthetic images with real images to train a model that classifies prostate tissue into epithelial and non-epithelial classes.¹⁴¹ The model outperformed the model trained using only real data on independent test data. In addition, pathologists were unable to distinguish between real and synthetic images in their study. Additionally, most studies rely on prostate biopsies to train and validate AI models (Table 2), but other tissue samples (e.g., transurethral resection or RP specimens) differ from biopsy tissue in many characteristics, raising questions about the accuracy of models trained on biopsy samples for other types of samples.

Tumor heterogeneity is another issue that needs attention. In clinical practice, more than one slide will be generated, whether it is a prostatectomy or biopsy specimen. However, most studies simply selected a representative slide per patient for training or validation. PCa is a highly heterogeneous tumor with multiple lesions presenting with different GPs. This especially poses challenges for predicting molecular subtyping and prognosis. Future research should incorporate data from all available slides of a patient. It would be of value to compare results generated from analyzing all slides versus single slides with the highest histologic grade.

Another key point is ensuring that the model is universally applicable to all patients, including specific subsets based on age, race, nationality, or other factors.¹⁴² For example, there are biological differences in PCa among races. Previous study found that the Asian population has a significantly higher frequency of FOXA1 mutations than European and American populations.¹⁴³ Currently, large international studies are limited to populations in Western countries.³⁰ Therefore, future investigation is warranted to explore the cross-ethnic applicability of AI pathology models to ensure that the models do not introduce discrimination or bias into clinical practice.^{69,80}

Scaling up the deployment of digital pathology

Clinical deployment of AI requires digital pathology labs; however, currently only a small portion of laboratories are fully digitized, even in high-resourced countries. To realize the benefits of computational pathology worldwide, we need to ramp up digital slide scanner availability and analysis capabilities cost-effectively. Pathology laboratories at regional medical centers could be prioritized for digitization, serving patients from their own institutions while also providing expert consultation services for other local institutions. Pathology laboratories could also choose to only scan a subset of challenging cases for AI solutions and teleconsultation before full digitization. This would allow patients with more intricate medical conditions to benefit from AI. As technology advances and economies of scale are achieved, the costs of deployment will decrease. Meanwhile, interim solutions such as computational-capable microscopes can fill the gap until whole-slide imaging catches up globally. For example, Chen et al. proposed an augmented-reality microscope (ARM) that overlays AI-based information in real time onto the current view of the sample, seamlessly integrating AI into routine workflows with relatively low-cost retrofitting.¹⁴⁴

Limitations of AI and the impact of AI on the behavior of pathologists

Despite AI's significant progress, potential flaws or biases may still lead to missed or incorrect diagnoses. Additionally, the misdiagnoses made by AI are not entirely without merits. In Tol-kach's study, pathologists recognized false-positive alerts from AI as useful warnings, as these highlighted areas required additional attention and immunostaining for further evaluation.³⁶ However, there is a risk that pathologists may overly rely on AI predictions without critically evaluating its predictions. Meyer's small-scale study showed that pathologists were willing to trust AI regardless of its accuracy.¹³⁸ This raises the question of whether AI would lead to overdependence and compromise pathologists' diagnostic skills. Therefore, similar to post-marketing surveillance for new drugs, continuous monitoring of pathology labs deploying AI tools is essential. This involves understanding how pathologists handle the results generated by AI and observing the impact of AI on clinical practice.

PROSPECTIVES

Recent advances in DL have brought AI to a level comparable to that of pathologists. AI has shown great potential in distinguishing between benign and malignant tumors, automated Gleason

grading, and the prediction of clinical prognosis and molecular subtypes. These tools can stratify patient risks and assist urologists in making clinical decisions. However, deploying AI systems in practice requires that AI systems be accurate and trustworthy and not contain biases or flaws that could lead to incorrect diagnoses or inappropriate treatment recommendations. Thus, efforts should focus on improving the generalizability of pathology AI and bridging the gap between regulatory testing and real-world clinical practice datasets. This entails prioritizing robust AI systems that are well-designed, rigorously tested, and continuously monitored for accuracy. Collaboration between AI experts, pathologists, and regulatory bodies is essential, along with ongoing training to effectively integrate pathology AI into practice.

This review provides an overview of the current developments and applications of AI in PCa management. One promising direction for future research is to predict molecular subtype based on pathology images or combining it with other omics techniques for precise diagnosis and treatment recommendations. In addition, generative AI tools such as ChatGPT also hold promise for future research. Drawing from its success in other fields, tools such as ChatGPT can be fine-tuned for specific tasks in PCa management, such as simplifying pathology reports for patient understanding.^{10,145} Integrating patient histories with WSIs for diagnosis and treatment recommendations is also feasible with updated multimodal analysis capabilities.^{146,147} However, careful evaluation is necessary to avoid hype and exaggeration.

DATA AND CODE AVAILABILITY

Not applicable.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.xcrm.2024.101506>.

ACKNOWLEDGMENTS

W.S.E.-D. is an American Cancer Society Research Professor and is supported by the Menco Family University Professorship at Brown University. Some WSI images in the figures were sourced from Bhattacharjee et al.¹⁴⁸ and Silva-Rodríguez et al.⁴⁹ Permission was obtained under the CC BY 4.0 license.

Funding

This study was supported by the National Natural Science Foundation of China (82272905), the Rising-Star Program of Science and Technology Commission of Shanghai Municipality (21QA1411500), the Natural Science Foundation of Science and Technology Commission of Shanghai (22ZR1478000), and Macao Polytechnic University grant (RP/FCA-15/2022). W.S.E.-D. is an American Cancer Society (ACS) Research Professor and is supported by the Menco Professorship in Medical Science at Brown University.

AUTHOR CONTRIBUTIONS

L.Z., J.P., and W.M., conceptualization, investigation, writing – original draft, methodology, literature review, writing – review & editing, and visualization; L.D., Y.Z., Y.W., G.P., E.H., B.A.C., M.H., W.S.E.-D., T.Y., T. Tan, T. Tong, N.T., Y.Z., Y.G., and Y.L.; conceptualization, writing – review & editing, and literature review; R.C., W.X., and L.C., conceptualization, literature review, project administration, supervision, resources, writing – review & editing,

and funding acquisition. All authors have read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

- Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., and Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A Cancer J. Clin.* 71, 209–249. <https://doi.org/10.3322/caac.21660>.
- Han, W., Johnson, C., Gaed, M., Gómez, J.A., Moussa, M., Chin, J.L., Pautler, S., Bauman, G.S., and Ward, A.D. (2020). Histologic tissue components provide major cues for machine learning-based prostate cancer detection and grading on prostatectomy specimens. *Sci. Rep.* 10, 9911. <https://doi.org/10.1038/s41598-020-66849-2>.
- Eichler, K., Hempel, S., Wilby, J., Myers, L., Bachmann, L.M., and Kleijnen, J. (2006). Diagnostic value of systematic biopsy methods in the investigation of prostate cancer: a systematic review. *J. Urol.* 175, 1605–1612. [https://doi.org/10.1016/S0022-5347\(05\)00957-2](https://doi.org/10.1016/S0022-5347(05)00957-2).
- Cheng, L., MacLennan, G.T., and Bostwick, D.G. (2020). *Urologic Surgical Pathology, Fourth edition* (Elsevier).
- Ozkan, T.A., Eruray, A.T., Cebeci, O.O., Memik, O., Ozcan, L., and Kuskonmaz, I. (2016). Interobserver variability in Gleason histological grading of prostate cancer. *Scand. J. Urol.* 50, 420–424. <https://doi.org/10.1080/21681805.2016.1206619>.
- Allsbrook, W.C., Mangold, K.A., Johnson, M.H., Lane, R.B., Lane, C.G., Amin, M.B., Bostwick, D.G., Humphrey, P.A., Jones, E.C., Reuter, V.E., et al. (2001). Interobserver reproducibility of Gleason grading of prostatic carcinoma: urologic pathologists. *Hum. Pathol.* 32, 74–80. <https://doi.org/10.1053/hupa.2001.21134>.
- Goodman, M., Ward, K.C., Osunkoya, A.O., Datta, M.W., Luthringer, D., Young, A.N., Marks, K., Cohen, V., Kennedy, J.C., Haber, M.J., and Amin, M.B. (2012). Frequency and determinants of disagreement and error in gleason scores: A population-based study of prostate cancer. *Prostate* 72, 1389–1398. <https://doi.org/10.1002/pros.22484>.
- Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L.H., and Aerts, H.J.W.L. (2018). Artificial intelligence in radiology. *Nat. Rev. Cancer* 18, 500–510. <https://doi.org/10.1038/s41568-018-0016-5>.
- Acs, B., Rantalainen, M., and Hartman, J. (2020). Artificial intelligence as the next step towards precision pathology. *J. Intern. Med.* 288, 62–81. <https://doi.org/10.1111/joim.13030>.
- Zhu, L., Mou, W., and Chen, R. (2023). 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.* 21, 269. <https://doi.org/10.1186/s12967-023-04123-5>.
- Bashashati, A., and Goldenberg, S.L. (2022). AI for prostate cancer diagnosis — hype or today's reality? *Nat. Rev. Urol.* 19, 261–262. <https://doi.org/10.1038/s41585-022-00583-4>.
- Raciti, P., Sue, J., Ceballos, R., Godrich, R., Kunz, J.D., Kapur, S., Reuter, V., Grady, L., Kanan, C., Klimstra, D.S., and Fuchs, T.J. (2020). Novel artificial intelligence system increases the detection of prostate cancer in whole slide images of core needle biopsies. *Mod. Pathol.* 33, 2058–2066. <https://doi.org/10.1038/s41379-020-0551-y>.
- Perincheri, S., Levi, A.W., Celli, R., Gershkovich, P., Rimm, D., Morrow, J.S., Rothrock, B., Raciti, P., Klimstra, D., and Sinard, J. (2021). An independent assessment of an artificial intelligence system for prostate cancer detection shows strong diagnostic accuracy. *Mod. Pathol.* 34, 1588–1595. <https://doi.org/10.1038/s41379-021-00794-x>.
- da Silva, L.M., Pereira, E.M., Salles, P.G., Godrich, R., Ceballos, R., Kunz, J.D., Casson, A., Viret, J., Chandarlapaty, S., Ferreira, C.G., et al. (2021). Independent real-world application of a clinical-grade automated prostate cancer detection system. *J. Pathol.* 254, 147–158. <https://doi.org/10.1002/path.5662>.
- Campanella, G., Hanna, M.G., Geneslaw, L., Mirafior, A., Werneck Krauss Silva, V., Busam, K.J., Brogi, E., Reuter, V.E., Klimstra, D.S., and Fuchs, T.J. (2019). Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat. Med.* 25, 1301–1309. <https://doi.org/10.1038/s41591-019-0508-1>.
- Raciti, P., Sue, J., Retamero, J.A., Ceballos, R., Godrich, R., Kunz, J.D., Casson, A., Thiagarajan, D., Ebrahimzadeh, Z., Viret, J., et al. (2023). Clinical Validation of Artificial Intelligence–Augmented Pathology Diagnosis Demonstrates Significant Gains in Diagnostic Accuracy in Prostate Cancer Detection. *Arch. Pathol. Lab Med.* 147, 1178–1185. <https://doi.org/10.5858/arpa.2022-0066-OA>.
- Eloy, C., Marques, A., Pinto, J., Pinheiro, J., Campelos, S., Curado, M., Vale, J., and Polónia, A. (2023). Artificial intelligence–assisted cancer diagnosis improves the efficiency of pathologists in prostatic biopsies. *Virchows Arch.* 482, 595–604. <https://doi.org/10.1007/s00428-023-03518-5>.
- Sandeman, K., Blom, S., Koponen, V., Manninen, A., Juhila, J., Rannikko, A., Ropponen, T., and Mirtti, T. (2022). AI Model for Prostate Biopsies Predicts Cancer Survival. *Diagnostics* 12, 1031. <https://doi.org/10.3390/diagnostics12051031>.
- Jung, M., Jin, M.-S., Kim, C., Lee, C., Nikas, I.P., Park, J.H., and Ryu, H.S. (2022). Artificial intelligence system shows performance at the level of uropathologists for the detection and grading of prostate cancer in core needle biopsy: an independent external validation study. *Mod. Pathol.* 35, 1449–1457. <https://doi.org/10.1038/s41379-022-01077-9>.
- Ryu, H.S., Jin, M.-S., Park, J.H., Lee, S., Cho, J., Oh, S., Kwak, T.-Y., Woo, J.I., Mun, Y., Kim, S.W., et al. (2019). Automated Gleason Scoring and Tumor Quantification in Prostate Core Needle Biopsy Images Using Deep Neural Networks and Its Comparison with Pathologist-Based Assessment. *Cancers* 11, E1860. <https://doi.org/10.3390/cancers11121860>.
- Pantanowitz, L., Quiroga-Garza, G.M., Bien, L., Heled, R., Laifenfeld, D., Linhart, C., Sandbank, J., Albrecht Shach, A., Shalev, V., Vecsler, M., et al. (2020). 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* 2, e407–e416. [https://doi.org/10.1016/S2589-7500\(20\)30159-X](https://doi.org/10.1016/S2589-7500(20)30159-X).
- Tolkach, Y., Dohmgörger, T., Toma, M., and Kristiansen, G. (2020). High-accuracy prostate cancer pathology using deep learning. *Nat. Mach. Intell.* 2, 411–418. <https://doi.org/10.1038/s42256-020-0200-7>.
- Vazzano, J., Johansson, D., Hu, K., Eurén, K., Elfving, S., Parwani, A., and Zhou, M. (2023). Evaluation of A Computer-Aided Detection Software for Prostate Cancer Prediction: Excellent Diagnostic Accuracy Independent of Preanalytical Factors. *Lab. Invest.* 103, 100257. <https://doi.org/10.1016/j.labinv.2023.100257>.
- Qu, L., Liu, S., Liu, X., Wang, M., and Song, Z. (2022). Towards label-efficient automatic diagnosis and analysis: a comprehensive survey of advanced deep learning-based weakly-supervised, semi-supervised and self-supervised techniques in histopathological image analysis. *Phys. Med. Biol.* 67, 20TR01. <https://doi.org/10.1088/1361-6560/ac910a>.
- Yang, Z., Wang, X., Xiang, J., Zhang, J., Yang, S., Wang, X., Yang, W., Li, Z., Han, X., and Liu, Y. (2023). The devil is in the details: a small-lesion sensitive weakly supervised learning framework for prostate cancer detection and grading. *Virchows Arch.* 482, 525–538. <https://doi.org/10.1007/s00428-023-03502-z>.
- Xiang, J., Wang, X., Wang, X., Zhang, J., Yang, S., Yang, W., Han, X., and Liu, Y. (2023). Automatic diagnosis and grading of Prostate Cancer with

- weakly supervised learning on whole slide images. *Comput. Biol. Med.* 152, 106340. <https://doi.org/10.1016/j.compbiomed.2022.106340>.
27. Oner, M.U., Ng, M.Y., Giron, D.M., Chen Xi, C.E., Yuan Xiang, L.A., Singh, M., Yu, W., Sung, W.-K., Wong, C.F., and Lee, H.K. (2022). An AI-assisted tool for efficient prostate cancer diagnosis in low-grade and low-volume cases. *Patterns* 3, 100642. <https://doi.org/10.1016/j.patter.2022.100642>.
28. Singhal, N., Soni, S., Bonthu, S., Chattopadhyay, N., Samanta, P., Joshi, U., Jojera, A., Chharchhodawala, T., Agarwal, A., Desai, M., and Ganpule, A. (2022). A deep learning system for prostate cancer diagnosis and grading in whole slide images of core needle biopsies. *Sci. Rep.* 12, 3383. <https://doi.org/10.1038/s41598-022-07217-0>.
29. Pohjonen, J., Stürenberg, C., Rannikko, A., Mirtti, T., and Pitkänen, E. (2022). Spectral decoupling for training transferable neural networks in medical imaging. *iScience* 25, 103767. <https://doi.org/10.1016/j.isci.2022.103767>.
30. Bulten, W., Kartasalo, K., Chen, P.-H.C., Ström, P., Pinckaers, H., Nagpal, K., Cai, Y., Steiner, D.F., van Boven, H., Vink, R., et al. (2022). Artificial intelligence for diagnosis and Gleason grading of prostate cancer: the PANDA challenge. *Nat. Med.* 28, 154–163. <https://doi.org/10.1038/s41591-021-01620-2>.
31. Silva-Rodríguez, J., Colomer, A., Dolz, J., and Naranjo, V. (2021). Self-Learning for Weakly Supervised Gleason Grading of Local Patterns. *IEEE J. Biomed. Health Inform.* 25, 3094–3104. <https://doi.org/10.1109/JBHI.2021.3061457>.
32. Mun, Y., Paik, I., Shin, S.-J., Kwak, T.-Y., and Chang, H. (2021). Yet Another Automated Gleason Grading System (YAAGGS) by weakly supervised deep learning. *NPJ Digit. Med.* 4, 99. <https://doi.org/10.1038/s41746-021-00469-6>.
33. Li, J., Li, W., Sisk, A., Ye, H., Wallace, W.D., Speier, W., and Arnold, C.W. (2021). A multi-resolution model for histopathology image classification and localization with multiple instance learning. *Comput. Biol. Med.* 131, 104253. <https://doi.org/10.1016/j.compbiomed.2021.104253>.
34. Silva-Rodríguez, J., Colomer, A., Sales, M.A., Molina, R., and Naranjo, V. (2020). Going deeper through the Gleason scoring scale: An automatic end-to-end system for histology prostate grading and cribriform pattern detection. *Comput. Methods Progr. Biomed.* 195, 105637. <https://doi.org/10.1016/j.cmpb.2020.105637>.
35. Nagpal, K., Foote, D., Tan, F., Liu, Y., Chen, P.-H.C., Steiner, D.F., Manoj, N., Olson, N., Smith, J.L., Mohtashamian, A., et al. (2020). Development and Validation of a Deep Learning Algorithm for Gleason Grading of Prostate Cancer From Biopsy Specimens. *JAMA Oncol.* 6, 1372–1380. <https://doi.org/10.1001/jamaoncol.2020.2485>.
36. Tolkach, Y., Ovtcharov, V., Pryalukhin, A., Eich, M.-L., Gaisa, N.T., Braun, M., Radzhabov, A., Quas, A., Hammer, P., Dellmann, A., et al. (2023). An international multi-institutional validation study of the algorithm for prostate cancer detection and Gleason grading. *npj Precis. Oncol.* 7, 77–79. <https://doi.org/10.1038/s41698-023-00424-6>.
37. Ström, P., Kartasalo, K., Olsson, H., Solorzano, L., Delahunt, B., Berney, D.M., Bostwick, D.G., Evans, A.J., Grignon, D.J., Humphrey, P.A., et al. (2020). Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. *Lancet Oncol.* 21, 222–232. [https://doi.org/10.1016/S1470-2045\(19\)30738-7](https://doi.org/10.1016/S1470-2045(19)30738-7).
38. Bulten, W., Pinckaers, H., van Boven, H., Vink, R., de Bel, T., van Ginneken, B., van der Laak, J., Hulsbergen-van de Kaa, C., and Litjens, G. (2020). Automated deep-learning system for Gleason grading of prostate cancer using biopsies: a diagnostic study. *Lancet Oncol.* 21, 233–241. [https://doi.org/10.1016/S1470-2045\(19\)30739-9](https://doi.org/10.1016/S1470-2045(19)30739-9).
39. Dov, D., Assaad, S., Syedibrahim, A., Bell, J., Huang, J., Madden, J., Bentley, R., McCall, S., Henao, R., Carin, L., and Foo, W.C. (2022). A Hybrid Human-Machine Learning Approach for Screening Prostate Biopsies Can Improve Clinical Efficiency Without Compromising Diagnostic Accuracy. *Arch. Pathol. Lab Med.* 146, 727–734. <https://doi.org/10.5858/arpa.2020-0850-OA>.
40. Liu, B., Wang, Y., Weitz, P., Lindberg, J., Hartman, J., Wang, W., Egevad, L., Grönberg, H., Eklund, M., and Rantalainen, M. (2022). Using deep learning to detect patients at risk for prostate cancer despite benign biopsies. *iScience* 25, 104663. <https://doi.org/10.1016/j.isci.2022.104663>.
41. Nir, G., Hor, S., Karimi, D., Fazli, L., Skinnider, B.F., Tavassoli, P., Turbin, D., Villamil, C.F., Wang, G., Wilson, R.S., et al. (2018). Automatic grading of prostate cancer in digitized histopathology images: Learning from multiple experts. *Med. Image Anal.* 50, 167–180. <https://doi.org/10.1016/j.media.2018.09.005>.
42. Arvaniti, E., Fricker, K.S., Moret, M., Rupp, N., Hermanns, T., Fankhauser, C., Wey, N., Wild, P.J., Rüschhoff, J.H., and Claassen, M. (2018). Automated Gleason grading of prostate cancer tissue microarrays via deep learning. *Sci. Rep.* 8, 12054. <https://doi.org/10.1038/s41598-018-30535-1>.
43. Nagpal, K., Foote, D., Liu, Y., Chen, P.-H.C., Wulczyn, E., Tan, F., Olson, N., Smith, J.L., Mohtashamian, A., Wren, J.H., et al. (2019). Development and validation of a deep learning algorithm for improving Gleason scoring of prostate cancer. *NPJ Digit. Med.* 2, 48. <https://doi.org/10.1038/s41746-019-0112-2>.
44. Sahran, S., Albashish, D., Abdullah, A., Shukor, N.A., and Hayati Md Pauzi, S. (2018). Absolute cosine-based SVM-RFE feature selection method for prostate histopathological grading. *Artif. Intell. Med.* 87, 78–90. <https://doi.org/10.1016/j.artmed.2018.04.002>.
45. Nir, G., Karimi, D., Goldenberg, S.L., Fazli, L., Skinnider, B.F., Tavassoli, P., Turbin, D., Villamil, C.F., Wang, G., Thompson, D.J.S., et al. (2019). Comparison of Artificial Intelligence Techniques to Evaluate Performance of a Classifier for Automatic Grading of Prostate Cancer From Digitized Histopathologic Images. *JAMA Netw. Open* 2, e190442. <https://doi.org/10.1001/jamanetworkopen.2019.0442>.
46. Li, W., Li, J., Sarma, K.V., Ho, K.C., Shen, S., Knudsen, B.S., Gertych, A., and Arnold, C.W. (2019). Path R-CNN for Prostate Cancer Diagnosis and Gleason Grading of Histological Images. *IEEE Trans. Med. Imag.* 38, 945–954. <https://doi.org/10.1109/TMI.2018.2875868>.
47. Lucas, M., Jansen, I., Savci-Heijink, C.D., Meijer, S.L., de Boer, O.J., van Leeuwen, T.G., de Bruin, D.M., and Marquering, H.A. (2019). Deep learning for automatic Gleason pattern classification for grade group determination of prostate biopsies. *Virchows Arch.* 475, 77–83. <https://doi.org/10.1007/s00428-019-02577-x>.
48. Karimi, D., Nir, G., Fazli, L., Black, P.C., Goldenberg, L., and Salcudean, S.E. (2020). Deep Learning-Based Gleason Grading of Prostate Cancer From Histopathology Images-Role of Multiscale Decision Aggregation and Data Augmentation. *IEEE J. Biomed. Health Inform.* 24, 1413–1426. <https://doi.org/10.1109/JBHI.2019.2944643>.
49. Silva-Rodríguez, J., Schmidt, A., Sales, M.A., Molina, R., and Naranjo, V. (2022). Proportion constrained weakly supervised histopathology image classification. *Comput. Biol. Med.* 147, 105714. <https://doi.org/10.1016/j.compbiomed.2022.105714>.
50. Marini, N., Otálora, S., Müller, H., and Atzori, M. (2021). Semi-supervised training of deep convolutional neural networks with heterogeneous data and few local annotations: An experiment on prostate histopathology image classification. *Med. Image Anal.* 73, 102165. <https://doi.org/10.1016/j.media.2021.102165>.
51. Otálora, S., Marini, N., Müller, H., and Atzori, M. (2021). Combining weakly and strongly supervised learning improves strong supervision in Gleason pattern classification. *BMC Med. Imag.* 21, 77. <https://doi.org/10.1186/s12880-021-00609-0>.
52. Pinckaers, H., Bulten, W., van der Laak, J., and Litjens, G. (2021). Detection of Prostate Cancer in Whole-Slide Images Through End-to-End Training With Image-Level Labels. *IEEE Trans. Med. Imag.* 40, 1817–1826. <https://doi.org/10.1109/TMI.2021.3066295>.
53. Hassan, T., Shafay, M., Hassan, B., Akram, M.U., ElBaz, A., and Werghi, N. (2022). Knowledge distillation driven instance segmentation for grading prostate cancer. *Comput. Biol. Med.* 150, 106124. <https://doi.org/10.1016/j.compbiomed.2022.106124>.

54. Toledo-Cortés, S., Useche, D.H., Müller, H., and González, F.A. (2022). Grading diabetic retinopathy and prostate cancer diagnostic images with deep quantum ordinal regression. *Comput. Biol. Med.* **145**, 105472. <https://doi.org/10.1016/j.compbiomed.2022.105472>.
55. Qiu, Y., Hu, Y., Kong, P., Xie, H., Zhang, X., Cao, J., Wang, T., and Lei, B. (2022). Automatic Prostate Gleason Grading Using Pyramid Semantic Parsing Network in Digital Histopathology. *Front. Oncol.* **12**, 772403. <https://doi.org/10.3389/fonc.2022.772403>.
56. Huang, W., Randhawa, R., Jain, P., Iczkowski, K.A., Hu, R., Hubbard, S., Eickhoff, J., Basu, H., and Roy, R. (2021). Development and Validation of an Artificial Intelligence-Powered Platform for Prostate Cancer Grading and Quantification. *JAMA Netw. Open* **4**, e2132554. <https://doi.org/10.1001/jamanetworkopen.2021.32554>.
57. Steiner, D.F., Nagpal, K., Sayres, R., Foote, D.J., Wedin, B.D., Pearce, A., Cai, C.J., Winter, S.R., Symonds, M., Yatziv, L., et al. (2020). Evaluation of the Use of Combined Artificial Intelligence and Pathologist Assessment to Review and Grade Prostate Biopsies. *JAMA Netw. Open* **3**, e2023267. <https://doi.org/10.1001/jamanetworkopen.2020.23267>.
58. Chen, Y., Loveless, I.M., Nakai, T., Newaz, R., Abdullah, F.F., Rogers, C.G., Hassan, O., Chitale, D., Arora, K., Williamson, S.R., et al. (2023). Convolutional Neural Network Quantification of Gleason Pattern 4 and Association With Biochemical Recurrence in Intermediate-Grade Prostate Tumors. *Mod. Pathol.* **36**, 100157. <https://doi.org/10.1016/j.modpat.2023.100157>.
59. Kartasalo, K., Ström, P., Ruusuvaari, P., Samaratunga, H., Delahunt, B., Tsuzuki, T., Eklund, M., and Egevad, L. (2022). Detection of perineural invasion in prostate needle biopsies with deep neural networks. *Virchows Arch.* **481**, 73–82. <https://doi.org/10.1007/s00428-022-03326-3>.
60. Blessin, N.C., Yang, C., Mandelkow, T., Raedler, J.B., Li, W., Bady, E., Simon, R., Vettorazzi, E., Lennartz, M., Bernreuther, C., et al. (2023). Automated Ki-67 labeling index assessment in prostate cancer using artificial intelligence and multiplex fluorescence immunohistochemistry. *J. Pathol.* **260**, 5–16. <https://doi.org/10.1002/path.6057>.
61. Harmon, S.A., Patel, P.G., Sanford, T.H., Caven, I., Iseman, R., Vidotto, T., Picanço, C., Squire, J.A., Masoudi, S., Mehralivand, S., et al. (2021). 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.* **34**, 478–489. <https://doi.org/10.1038/s41379-020-00674-w>.
62. Leo, P., Chandramouli, S., Farré, X., Elliott, R., Janowczyk, A., Bera, K., Fu, P., Janaki, N., El-Fahmawi, A., Shahait, M., et al. (2021). Computationally Derived Cribriform Area Index from Prostate Cancer Hematoxylin and Eosin Images Is Associated with Biochemical Recurrence Following Radical Prostatectomy and Is Most Prognostic in Gleason Grade Group 2. *Eur. Urol. Focus* **7**, 722–732. <https://doi.org/10.1016/j.euf.2021.04.016>.
63. Ambrosini, P., Hollemans, E., Kweldam, C.F., Leenders, G.J.L.H.v., Stallinga, S., and Vos, F. (2020). Automated detection of cribriform growth patterns in prostate histology images. *Sci. Rep.* **10**, 14904. <https://doi.org/10.1038/s41598-020-71942-7>.
64. Brendel, M., Getseva, V., Assaad, M.A., Sigouros, M., Sigaras, A., Kane, T., Khosravi, P., Mosquera, J.M., Elemento, O., and Hajirasouliha, I. (2022). Weakly-supervised tumor purity prediction from frozen H&E stained slides. *EBioMedicine* **80**, 104067. <https://doi.org/10.1016/j.ebiom.2022.104067>.
65. Chatrian, A., Colling, R.T., Browning, L., Alham, N.K., Sirinukunwattana, K., Malacrino, S., Haghighat, M., Aberdeen, A., Monks, A., Moxley-Wyles, B., et al. (2021). Artificial intelligence for advance requesting of immunohistochemistry in diagnostically uncertain prostate biopsies. *Mod. Pathol.* **34**, 1780–1794. <https://doi.org/10.1038/s41379-021-00826-6>.
66. Haghighat, M., Browning, L., Sirinukunwattana, K., Malacrino, S., Khalid Alham, N., Colling, R., Cui, Y., Rakha, E., Hamdy, F.C., Verrill, C., and Rittscher, J. (2022). Automated quality assessment of large digitised histology cohorts by artificial intelligence. *Sci. Rep.* **12**, 5002. <https://doi.org/10.1038/s41598-022-08351-5>.
67. Rana, A., Lowe, A., Lithgow, M., Horback, K., Janovitz, T., Da Silva, A., Tsai, H., Shanmugam, V., Bayat, A., and Shah, P. (2020). Use of Deep Learning to Develop and Analyze Computational Hematoxylin and Eosin Staining of Prostate Core Biopsy Images for Tumor Diagnosis. *JAMA Netw. Open* **3**, e205111. <https://doi.org/10.1001/jamanetworkopen.2020.5111>.
68. Egevad, L., Delahunt, B., Samaratunga, H., Tsuzuki, T., Yamamoto, Y., Yaxley, J., Ruusuvaari, P., Kartasalo, K., and Eklund, M. (2021). The emerging role of artificial intelligence in the reporting of prostate pathology. *Pathology* **53**, 565–567. <https://doi.org/10.1016/j.pathol.2021.04.002>.
69. Oszwald, A., Wasinger, G., Pradere, B., Shariat, S.F., and Compérat, E.M. (2021). Artificial intelligence in prostate histopathology: where are we in 2021? *Curr. Opin. Urol.* **31**, 430–435. <https://doi.org/10.1097/MOU.0000000000000883>.
70. Egevad, L., Ström, P., Kartasalo, K., Olsson, H., Samaratunga, H., Delahunt, B., and Eklund, M. (2020). The utility of artificial intelligence in the assessment of prostate pathology. *Histopathology* **76**, 790–792. <https://doi.org/10.1111/his.14060>.
71. Bulten, W., Balkenhol, M., Belinga, J.-J.A., Brilhante, A., Çakır, A., Egevad, L., Eklund, M., Farré, X., Geronatsiou, K., Molinié, V., et al. (2021). Artificial intelligence assistance significantly improves Gleason grading of prostate biopsies by pathologists. *Mod. Pathol.* **34**, 660–671. <https://doi.org/10.1038/s41379-020-0640-y>.
72. Egevad, L., Swanberg, D., Delahunt, B., Ström, P., Kartasalo, K., Olsson, H., Berney, D.M., Bostwick, D.G., Evans, A.J., Humphrey, P.A., et al. (2020). Identification of areas of grading difficulties in prostate cancer and comparison with artificial intelligence assisted grading. *Virchows Arch.* **477**, 777–786. <https://doi.org/10.1007/s00428-020-02858-w>.
73. Abraham, N.E., Mendhiratta, N., and Taneja, S.S. (2015). Patterns of repeat prostate biopsy in contemporary clinical practice. *J. Urol.* **193**, 1178–1184. <https://doi.org/10.1016/j.juro.2014.10.084>.
74. Parwani, A.V. (2019). Commentary: Automated Diagnosis and Gleason Grading of Prostate Cancer - Are Artificial Intelligence Systems Ready for Prime Time? *J. Pathol. Inf.* **10**, 41. https://doi.org/10.4103/jpi.jpi_56_19.
75. Thomas, T. (2020). Automated systems comparable to expert pathologists for prostate cancer Gleason grading. *Nat. Rev. Urol.* **17**, 131. <https://doi.org/10.1038/s41585-020-0294-z>.
76. Kalra, S., Tizhoosh, H.R., Shah, S., Choi, C., Damaskinos, S., Safarpour, A., Shafiei, S., Babaie, M., Diamandis, P., Campbell, C.J.V., and Pantanowitz, L. (2020). Pan-cancer diagnostic consensus through searching archival histopathology images using artificial intelligence. *NPJ Digit. Med.* **3**, 31. <https://doi.org/10.1038/s41746-020-0238-2>.
77. Hegde, N., Hipp, J.D., Liu, Y., Emmert-Buck, M., Reif, E., Smilkov, D., Terry, M., Cai, C.J., Amin, M.B., Mermel, C.H., et al. (2019). Similar image search for histopathology: SMILY. *NPJ Digit. Med.* **2**, 56. <https://doi.org/10.1038/s41746-019-0131-z>.
78. Sparks, R., and Madabhushi, A. (2016). Out-of-Sample Extrapolation utilizing Semi-Supervised Manifold Learning (OSE-SSL): Content Based Image Retrieval for Histopathology Images. *Sci. Rep.* **6**, 27306. <https://doi.org/10.1038/srep27306>.
79. Wulczyn, E., Nagpal, K., Symonds, M., Moran, M., Plass, M., Reihs, R., Nader, F., Tan, F., Cai, Y., Brown, T., et al. (2021). Predicting prostate cancer specific-mortality with artificial intelligence-based Gleason grading. *Commun. Med.* **1**, 10. <https://doi.org/10.1038/s43856-021-00005-3>.
80. Kartasalo, K., Bulten, W., Delahunt, B., Chen, P.-H.C., Pinckaers, H., Olsson, H., Ji, X., Mulliqi, N., Samaratunga, H., Tsuzuki, T., et al. (2021). Artificial Intelligence for Diagnosis and Gleason Grading of Prostate Cancer in Biopsies—Current Status and Next Steps. *Eur. Urol. Focus* **7**, 687–691. <https://doi.org/10.1016/j.euf.2021.07.002>.

81. Shmatko, A., Ghaffari Laleh, N., Gerstung, M., and Kather, J.N. (2022). Artificial intelligence in histopathology: enhancing cancer research and clinical oncology. *Nat. Can. (Ott.)* 3, 1026–1038. <https://doi.org/10.1038/s43018-022-00436-4>.
82. Esteve, A., Feng, J., van der Wal, D., Huang, S.-C., Simko, J.P., DeVries, S., Chen, E., Schaeffer, E.M., Morgan, T.M., Sun, Y., et al. (2022). Prostate cancer therapy personalization via multi-modal deep learning on randomized phase III clinical trials. *NPJ Digit. Med.* 5, 71. <https://doi.org/10.1038/s41746-022-00613-w>.
83. Ren, J., Karagoz, K., Gatz, M.L., Singer, E.A., Sadimin, E., Foran, D.J., and Qi, X. (2018). Recurrence analysis on prostate cancer patients with Gleason score 7 using integrated histopathology whole-slide images and genomic data through deep neural networks. *J. Med. Imaging* 5, 047501. <https://doi.org/10.1117/1.JMI.5.4.047501>.
84. Donovan, M.J., Fernandez, G., Scott, R., Khan, F.M., Zeineh, J., Koll, G., Gladoun, N., Charytonowicz, E., Tewari, A., and Cordon-Cardo, C. (2018). Development and validation of a novel automated Gleason grade and molecular profile that define a highly predictive prostate cancer progression algorithm-based test. *Prostate Cancer Prostatic Dis.* 21, 594–603. <https://doi.org/10.1038/s41391-018-0067-4>.
85. Bhargava, H.K., Leo, P., Elliott, R., Janowczyk, A., Whitney, J., Gupta, S., Fu, P., Yamoah, K., Khani, F., Robinson, B.D., et al. (2020). Computationally Derived Image Signature of Stromal Morphology Is Prognostic of Prostate Cancer Recurrence Following Prostatectomy in African American Patients. *Clin. Cancer Res.* 26, 1915–1923. <https://doi.org/10.1158/1078-0432.CCR-19-2659>.
86. Pinckaers, H., van Ipenburg, J., Melamed, J., De Marzo, A., Platz, E.A., van Ginneken, B., van der Laak, J., and Litjens, G. (2022). Predicting biochemical recurrence of prostate cancer with artificial intelligence. *Commun. Méd.* 2, 64. <https://doi.org/10.1038/s43856-022-00126-3>.
87. Yamamoto, Y., Tsuzuki, T., Akatsuka, J., Ueki, M., Morikawa, H., Numata, Y., Takahara, T., Tsuyuki, T., Tsutsumi, K., Nakazawa, R., et al. (2019). Automated acquisition of explainable knowledge from unannotated histopathology images. *Nat. Commun.* 10, 5642. <https://doi.org/10.1038/s41467-019-13647-8>.
88. Lee, G., Veltri, R.W., Zhu, G., Ali, S., Epstein, J.I., and Madabhushi, A. (2017). Nuclear Shape and Architecture in Benign Fields Predict Biochemical Recurrence in Prostate Cancer Patients Following Radical Prostatectomy: Preliminary Findings. *Eur. Urol. Focus* 3, 457–466. <https://doi.org/10.1016/j.euf.2016.05.009>.
89. Huang, W., Randhawa, R., Jain, P., Hubbard, S., Eickhoff, J., Kummar, S., Wilding, G., Basu, H., and Roy, R. (2022). A Novel Artificial Intelligence-Powered Method for Prediction of Early Recurrence of Prostate Cancer After Prostatectomy and Cancer Drivers. *JCO Clin. Cancer Inform.* 6, e2100131. <https://doi.org/10.1200/CCI.21.00131>.
90. Harder, N., Athelogou, M., Hessel, H., Brieu, N., Yigitsoy, M., Zimmermann, J., Baatz, M., Buchner, A., Stief, C.G., Kirchner, T., et al. (2018). Tissue Phenomics for prognostic biomarker discovery in low- and intermediate-risk prostate cancer. *Sci. Rep.* 8, 4470. <https://doi.org/10.1038/s41598-018-22564-7>.
91. Pastor-Navarro, B., Rubio-Briones, J., Borque-Fernando, Á., Esteban, L.M., Dominguez-Escrig, J.L., and López-Guerrero, J.A. (2021). Active Surveillance in Prostate Cancer: Role of Available Biomarkers in Daily Practice. *Int. J. Mol. Sci.* 22, 6266. <https://doi.org/10.3390/ijms22126266>.
92. Tsuneki, M., Abe, M., Ichihara, S., and Kanavati, F. (2023). Inference of core needle biopsy whole slide images requiring definitive therapy for prostate cancer. *BMC Cancer* 23, 11. <https://doi.org/10.1186/s12885-022-10488-5>.
93. Chandramouli, S., Leo, P., Lee, G., Elliott, R., Davis, C., Zhu, G., Fu, P., Epstein, J.I., Veltri, R., and Madabhushi, A. (2020). Computer Extracted Features from Initial H&E Tissue Biopsies Predict Disease Progression for Prostate Cancer Patients on Active Surveillance. *Cancers* 12, E2708. <https://doi.org/10.3390/cancers12092708>.
94. Wessels, F., Schmitt, M., Kriehoff-Henning, E., Jutzi, T., Worst, T.S., Waldbillig, F., Neuberger, M., Maron, R.C., Steeg, M., Gaiser, T., et al. (2021). Deep learning approach to predict lymph node metastasis directly from primary tumour histology in prostate cancer. *BJU Int.* 128, 352–360. <https://doi.org/10.1111/bju.15386>.
95. Nakata, W., Mori, H., Tsujimura, G., Tsujimoto, Y., Gotoh, T., and Tsujihata, M. (2022). Pilot study of an artificial intelligence-based deep learning algorithm to predict time to castration-resistant prostate cancer for metastatic hormone-naïve prostate cancer. *Jpn. J. Clin. Oncol.* 52, 1062–1066. <https://doi.org/10.1093/jco/hyac089>.
96. Paulson, N., Zeevi, T., Papademetris, M., Leapman, M.S., Onofrey, J.A., Sprenkle, P.C., Humphrey, P.A., Staib, L.H., and Levi, A.W. (2022). Prediction of Adverse Pathology at Radical Prostatectomy in Grade Group 2 and 3 Prostate Biopsies Using Machine Learning. *JCO Clin. Cancer Inform.* 6, e2200016. <https://doi.org/10.1200/CCI.22.00016>.
97. Kather, J.N., Heij, L.R., Grabsch, H.I., Loeffler, C., Echle, A., Muti, H.S., Krause, J., Niehues, J.M., Sommer, K.A.J., Bankhead, P., et al. (2020). Pan-cancer image-based detection of clinically actionable genetic alterations. *Nat. Can. (Ott.)* 1, 789–799. <https://doi.org/10.1038/s43018-020-0087-6>.
98. Fu, Y., Jung, A.W., Torne, R.V., Gonzalez, S., Vöhringer, H., Shmatko, A., Yates, L.R., Jimenez-Linan, M., Moore, L., and Gerstung, M. (2020). Pan-cancer computational histopathology reveals mutations, tumor composition and prognosis. *Nat. Can. (Ott.)* 1, 800–810. <https://doi.org/10.1038/s43018-020-0085-8>.
99. Zoma, M., Curti, L., Shinde, D., Albino, D., Mitra, A., Sgrignani, J., Mapelli, S.N., Sandrini, G., Civenni, G., Merulla, J., et al. (2021). EZH2-induced lysine K362 methylation enhances TMPRSS2-ERG oncogenic activity in prostate cancer. *Nat. Commun.* 12, 4147. <https://doi.org/10.1038/s41467-021-24380-6>.
100. Mosquera, J.-M., Perner, S., Demichelis, F., Kim, R., Hofer, M.D., Mertz, K.D., Paris, P.L., Simko, J., Collins, C., Bismar, T.A., et al. (2007). Morphological features of TMPRSS2-ERG gene fusion prostate cancer. *J. Pathol.* 212, 91–101. <https://doi.org/10.1002/path.2154>.
101. Dadhania, V., Gonzalez, D., Yousif, M., Cheng, J., Morgan, T.M., Spratt, D.E., Reichert, Z.R., Mannan, R., Wang, X., Chinnaiyan, A., et al. (2022). Leveraging artificial intelligence to predict ERG gene fusion status in prostate cancer. *BMC Cancer* 22, 494. <https://doi.org/10.1186/s12885-022-09559-4>.
102. Erak, E., Oliveira, L.D., Mendes, A.A., Dairo, O., Ertunc, O., Kulac, I., Baena-Del Valle, J.A., Jones, T., Hicks, J.L., Glavaris, S., et al. (2023). Predicting Prostate Cancer Molecular Subtype With Deep Learning on Histopathologic Images. *Mod. Pathol.* 36, 100247. <https://doi.org/10.1016/j.modpat.2023.100247>.
103. Schaumberg, A.J., Rubin, M.A., and Fuchs, T.J. (2018). H&E-stained Whole Slide Image Deep Learning Predicts SPOP Mutation State in Prostate Cancer. Preprint at bioRxiv. <https://doi.org/10.1101/064279>.
104. Weitz, P., Wang, Y., Kartasalo, K., Egevad, L., Lindberg, J., Grönberg, H., Eklund, M., and Rantalainen, M. (2022). Transcriptome-wide prediction of prostate cancer gene expression from histopathology images using co-expression based convolutional neural networks. *Bioinformatics* 38, 3462–3469. <https://doi.org/10.1093/bioinformatics/btac343>.
105. Zormpas-Petridis, K., Noguera, R., Ivankovic, D.K., Roxanis, I., Jamin, Y., and Yuan, Y. (2020). SuperHistopath: A Deep Learning Pipeline for Mapping Tumor Heterogeneity on Low-Resolution Whole-Slide Digital Histopathology Images. *Front. Oncol.* 10, 586292. <https://doi.org/10.3389/fonc.2020.586292>.
106. Wong, A.N.N., He, Z., Leung, K.L., To, C.C.K., Wong, C.Y., Wong, S.C.C., Yoo, J.S., Chan, C.K.R., Chan, A.Z., Lacambra, M.D., and Yeung, M.H.Y. (2022). Current Developments of Artificial Intelligence in Digital Pathology and Its Future Clinical Applications in Gastrointestinal Cancers. *Cancers* 14, 3780. <https://doi.org/10.3390/cancers14153780>.
107. Schömig-Markieffka, B., Prylukhin, A., Hulla, W., Bychkov, A., Fukuoka, J., Madabhushi, A., Achter, V., Nieroda, L., Büttner, R., Quaas, A., and

- Tolkach, Y. (2021). Quality control stress test for deep learning-based diagnostic model in digital pathology. *Mod. Pathol.* 34, 2098–2108. <https://doi.org/10.1038/s41379-021-00859-x>.
108. Macenko, M., Niethammer, M., Marron, J.S., Borland, D., Woosley, J.T., Guan, X., Schmitt, C., and Thomas, N.E. (2009). A method for normalizing histology slides for quantitative analysis. In 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 1107–1110. <https://doi.org/10.1109/ISBI.2009.5193250>.
109. Anghel, A., Stanisavljevic, M., Andani, S., Papandreou, N., Rüschoff, J.H., Wild, P., Gabrani, M., and Pozidis, H. (2019). A High-Performance System for Robust Stain Normalization of Whole-Slide Images in Histopathology. *Front. Med.* 6, 193. <https://doi.org/10.3389/fmed.2019.00193>.
110. Swiderska-Chadaj, Z., de Bel, T., Blanchet, L., Baidoshvili, A., Vossen, D., van der Laak, J., and Litjens, G. (2020). Impact of rescanning and normalization on convolutional neural network performance in multi-center, whole-slide classification of prostate cancer. *Sci. Rep.* 10, 14398. <https://doi.org/10.1038/s41598-020-71420-0>.
111. van Leenders, G.J.L.H., van der Kwast, T.H., Grignon, D.J., Evans, A.J., Kristiansen, G., Kweldam, C.F., Litjens, G., McKenney, J.K., Melamed, J., Mottet, N., et al. (2020). The 2019 International Society of Urological Pathology (ISUP) Consensus Conference on Grading of Prostatic Carcinoma. *Am. J. Surg. Pathol.* 44, e87–e99. <https://doi.org/10.1097/PAS.0000000000001497>.
112. Prostate cANcer graDe Assessment (PANDA) Challenge <https://kaggle.com/competitions/prostate-cancer-grade-assessment>.
113. Zuley, M.L., Jarosz, R., Drake, B.F., Rancilio, D., Klim, A., Rieger-Christ, K., and Lemmerman, J. (2016). The Cancer Genome Atlas Prostate Adenocarcinoma Collection (TCGA-PRAD). Version 4 (The Cancer Imaging Archive). <https://doi.org/10.7937/K9/TCIA.2016.YXOGLM4Y>.
114. Bulten, W., Bándi, P., Hoven, J., Loo, R.v.d., Lotz, J., Weiss, N., Laak, J.v.d., Ginneken, B.v., Hulsbergen-van de Kaa, C., and Litjens, G. (2021). PESO: Prostate Epithelium Segmentation on H&E-stained Prostatectomy Whole Slide Images. <https://doi.org/10.5281/zenodo.5137717>.
115. Bulten, W., Bándi, P., Hoven, J., Loo, R.v.d., Lotz, J., Weiss, N., Laak, J.v.d., Ginneken, B.v., Hulsbergen-van de Kaa, C., and Litjens, G. (2019). Epithelium segmentation using deep learning in H&E-stained prostate specimens with immunohistochemistry as reference standard. *Sci. Rep.* 9, 864. <https://doi.org/10.1038/s41598-018-37257-4>.
116. Koziarski, M., Cyganek, B., Olborski, B., Antosz, Z., Żydak, M., Kwolek, B., Wąsowicz, P., Bułko, A., Swadźba, J., and Sitkowski, P. (2021). DiagSet: a dataset for prostate cancer histopathological image classification. Preprint at arXiv. <https://doi.org/10.48550/arXiv.2105.04014>.
117. Wilkinson, S., Ye, H., Karzai, F., Harmon, S.A., Terrigino, N.T., VanderWelle, D.J., Bright, J.R., Atway, R., Trostel, S.Y., Carrabba, N.V., et al. (2020). Nascent prostate cancer heterogeneity drives evolution and resistance to intense hormonal therapy. Version 1 (The Cancer Imaging Archive). <https://doi.org/10.7937/K9/TCIA.JHQD-FR46>.
118. Gallo, M., Krajčanský, V., Nenutil, R., Holub, P., and Brázdil, T. (2023). Shedding light on the black box of a neural network used to detect prostate cancer in whole slide images by occlusion-based explainability. *N. Biotechnol.* 78, 52–67. <https://doi.org/10.1016/j.nbt.2023.09.008>.
119. Kang, Y., Kim, Y.J., Park, S., Ro, G., Hong, C., Jang, H., Cho, S., Hong, W.J., Kang, D.U., Chun, J., et al. (2021). Development and operation of a digital platform for sharing pathology image data. *BMC Med. Inf. Decis. Making* 21, 114. <https://doi.org/10.1186/s12911-021-01466-1>.
120. AGGC22 - Grand Challenge Gd.-Challengeorg. <https://aggc22.grand-challenge.org/AGGC22/>.
121. pathpresenter - slide library <https://pathpresenter.net/slide-library>.
122. Madabhushi, A., and Feldman, M. (2016). Fused Radiology-Pathology Prostate Dataset. Version 1 (The Cancer Imaging Archive). <https://doi.org/10.7937/K9/TCIA.2016.TLPMR1AM>.
123. PAIP2021 - Grand Challenge Gd.-Challengeorg. <https://paip2021.grand-challenge.org/>.
124. TissUMaps - PROSTATE CANCER PROJECT STHLM3 <https://tissuumsaps.research.it.uu.se/sthlm3/>.
125. Cancer Moonshot Biobank (2022). Cancer Moonshot Biobank - Prostate Cancer Collection (CMB-PCA). Version 2 (The Cancer Imaging Archive). <https://doi.org/10.7937/25T7-6Y12>.
126. Zhong, Q., Guo, T., Rechsteiner, M., Rüschoff, J.H., Rupp, N., Fankhauser, C., Saba, K., Mortezaei, A., Poyet, C., Hermanns, T., et al. (2017). A curated collection of tissue microarray images and clinical outcome data of prostate cancer patients. *Sci. Data* 4, 170014. <https://doi.org/10.1038/sdata.2017.14>.
127. Tolkach, Y. (2021). Datasets Digital Pathology and Artifacts, Part 1. <https://doi.org/10.5281/zenodo.4789576>.
128. Tolkach, Y. (2021). Datasets Digital Pathology and Artifacts, Part 2. <https://doi.org/10.5281/zenodo.4904569>.
129. Choyke, P., Turkbey, B., Pinto, P., Merino, M., and Wood, B. (2016). Data from PROSTATE-MRI. Version 1 (The Cancer Imaging Archive). <https://doi.org/10.7937/K9/TCIA.2016.6046GUDV>.
130. Arvaniti, E., Fricker, K., Moret, M., Rupp, N., Hermanns, T., Fankhauser, C., Wey, N., Wild, P., Rüschoff, J.H., and Claassen, M. (2018). Replication Data for: Automated Gleason grading of prostate cancer tissue microarrays via deep learning, (Harvard Dataverse). 10.7910/dvn/ocycmp. <https://doi.org/10.7910/dvn/ocycmp>.
131. Gleason2019 - Grand Challenge Gd.-Challengeorg. <https://gleason2019.grand-challenge.org/>.
132. Salvi, M., Bosco, M., Molinaro, L., Gambella, A., Papotti, M.G., Acharya, U.R., and Molinari, F. (2021). RINGS Algorithm Dataset, 1. <https://doi.org/10.17632/h8bdwrtnr5.1>.
133. Salvi, M., Bosco, M., Molinaro, L., Gambella, A., Papotti, M., Acharya, U.R., and Molinari, F. (2021). A hybrid deep learning approach for gland segmentation in prostate histopathological images. *Artif. Intell. Med.* 115, 102076. <https://doi.org/10.1016/j.artmed.2021.102076>.
134. Gertych, A., Ing, N., Ma, Z., Fuchs, T.J., Salman, S., Mohanty, S., Bhele, S., Velásquez-Vacca, A., Amin, M.B., and Knudsen, B.S. (2015). Machine learning approaches to analyze histological images of tissues from radical prostatectomies. *Comput. Med. Imag. Graph.* 46 Pt 2, 197–208. <https://doi.org/10.1016/j.compmedimag.2015.08.002>.
135. Prostate cancer grading | Grading of Prostatic Adenocarcinoma: ISUP Grade 1 – 5. <https://isupweb.org/pib/>.
136. Kweldam, C.F., Nieboer, D., Algaba, F., Amin, M.B., Berney, D.M., Billis, A., Bostwick, D.G., Bubendorf, L., Cheng, L., Compérat, E., et al. (2016). Gleason grade 4 prostate adenocarcinoma patterns: an interobserver agreement study among genitourinary pathologists. *Histopathology* 69, 441–449. <https://doi.org/10.1111/his.12976>.
137. Li, W., Li, J., Wang, Z., Polson, J., Sisk, A.E., Sajed, D.P., Speier, W., and Arnold, C.W. (2022). PathAL: An Active Learning Framework for Histopathology Image Analysis. *IEEE Trans. Med. Imag.* 41, 1176–1187. <https://doi.org/10.1109/TMI.2021.3135002>.
138. Meyer, J., Khademi, A., Tétu, B., Han, W., Nippak, P., and Remisch, D. (2022). Impact of artificial intelligence on pathologists' decisions: an experiment. *J. Am. Med. Assoc.* 29, 1688–1695. <https://doi.org/10.1093/jama/ocac103>.
139. Hammouda, K., Khalifa, F., El-Melegy, M., Ghazal, M., Darwish, H.E., Abou El-Ghar, M., and El-Baz, A. (2021). A Deep Learning Pipeline for Grade Groups Classification Using Digitized Prostate Biopsy Specimens. *Sensors* 21, 6708. <https://doi.org/10.3390/s21206708>.
140. Goldenberg, S.L., Nir, G., and Salcudean, S.E. (2019). A new era: artificial intelligence and machine learning in prostate cancer. *Nat. Rev. Urol.* 16, 391–403. <https://doi.org/10.1038/s41585-019-0193-3>.
141. Falahkheirkhah, K., Tiwari, S., Yeh, K., Gupta, S., Herrera-Hernandez, L., McCarthy, M.R., Jimenez, R.E., Cheville, J.C., and Bhargava, R. (2023).

- Deepfake Histologic Images for Enhancing Digital Pathology. *Lab. Invest.* 103, 100006. <https://doi.org/10.1016/j.labinv.2022.100006>.
142. Chauhan, C., and Gullapalli, R.R. (2021). Ethics of AI in Pathology: Current Paradigms and Emerging Issues. *Am. J. Pathol.* 191, 1673–1683. <https://doi.org/10.1016/j.ajpath.2021.06.011>.
143. Li, J., Xu, C., Lee, H.J., Ren, S., Zi, X., Zhang, Z., Wang, H., Yu, Y., Yang, C., Gao, X., et al. (2020). A genomic and epigenomic atlas of prostate cancer in Asian populations. *Nature* 580, 93–99. <https://doi.org/10.1038/s41586-020-2135-x>.
144. Chen, P.-H.C., Gadepalli, K., MacDonald, R., Liu, Y., Kadowaki, S., Nagpal, K., Kohlberger, T., Dean, J., Corrado, G.S., Hipp, J.D., et al. (2019). An augmented reality microscope with real-time artificial intelligence integration for cancer diagnosis. *Nat. Med.* 25, 1453–1457. <https://doi.org/10.1038/s41591-019-0539-7>.
145. Amin, K.S., Davis, M.A., Doshi, R., Haims, A.H., Khosla, P., and Forman, H.P. (2023). Accuracy of ChatGPT, Google Bard, and Microsoft Bing for Simplifying Radiology Reports. *Radiology* 309, e232561. <https://doi.org/10.1148/radiol.232561>.
146. Yang, Z., Li, L., Lin, K., Wang, J., Lin, C.-C., Liu, Z., and Wang, L. (2023). The Dawn of LMMs: Preliminary Explorations with GPT-4V(ision). Preprint.
147. Zhu, L., Mou, W., Lai, Y., Chen, J., Lin, S., Xu, L., Lin, J., Guo, Z., Yang, T., Lin, A., et al. (2024). Step into the era of large multimodal models: A pilot study on ChatGPT-4V(ision)'s ability to interpret radiological images. *Int. J. Surg.* <https://doi.org/10.1097/JS9.0000000000001359>.
148. Bhattacharjee, S., Ikromjanov, K., Carole, K.S., Madusanka, N., Cho, N.-H., Hwang, Y.-B., Sumon, R.I., Kim, H.-C., and Choi, H.-K. (2021). Cluster Analysis of Cell Nuclei in H&E-Stained Histological Sections of Prostate Cancer and Classification Based on Traditional and Modern Artificial Intelligence Techniques. *Diagnostics* 12, 15. <https://doi.org/10.3390/diagnostics12010015>.