(Literature Review)

Optimizing Prostate Cancer Management: A Review on the Diagnostic Accuracy and Health Disparities using Machine Learning

Franklin Vincentius Malonda[1,*], Alvin Julian[2]

Received: 27-02-2026
Revised: 03-03-2026
Accepted: 03-03-2026
Published: 27-03-2026

Correspondence: franklin@ukwms.ac.id; Tel.: +6282122008224

[1] Widya Mandala Surabaya Catholic University, Surabaya, Indonesia; franklin@ukwms.ac.id
[2] Widya Mandala Surabaya Catholic University, Surabaya, Indonesia; alvinjulian@ukwms.ac.id

Highlights

This article shows the maturity of machine learning-based solutions for prostate cancer management, which has high diagnostic accuracy using multimodal data fusion. It also shows the disparities in the fairness of the solutions for diverse populations worldwide.

What are the main findings?

- Deep learning algorithms have high diagnostic precision with area under the receiver operating characteristic curve ranging from 0.84 to 0.91, which is comparable or superior to the precision of expert radiologists and pathologists.
- Multimodal data fusion significantly increases biopsy specificity and boosts the accuracy of the predictions for lymph node invasion and biochemical recurrence.
- The automated Gleason grading system has high reliability, with quadratic weighted kappa scores greater than 0.90 in the validation datasets.

What are the implications of the main findings?

- The integration of explainable AI tools into the urological workflow is seen as an important factor in building clinical trust through the provision of transparent and visually justifiable results to complex clinical decision-making processes.
- The need to address the racial performance disparities is crucial in the implementation of fairness-aware AI algorithms to ensure equitable care in the field of oncology and avoid any form of diagnostic inequalities among the patient population.
- The move from research to clinical practice requires the design of prospective clinical trials to assess the robustness of the AI models across the heterogeneous clinical protocols and equipment.

Abstract

Prostate cancer requires diagnostic tools that transcend conventional clinical paradigms. This systematic literature review aggregates primary research articles from 2021 to 2025 on the ap-plication of machine learning in prostate cancer management. We specifically selected studies on diagnostic accuracy and health disparities while excluding secondary sources. Overall, the results indicate a promising trend in multimodal fusion and explainable machine learning. All models had an area under the curve ranging from 0.84 to 0.91, which was either comparable to or even surpassing the performance of human experts such as radiologists and pathologists. Notably, the integrated models had a positive impact on biopsy specificity, while AI-based pathology had a kappa statistic of above 0.90 in Gleason grading. One of the issues, however, remains how to ensure the generalizability of the models to different racial and geographic populations. Overall, machine learning significantly enhances the accuracy of diagnosis and the efficiency of management. Therefore, it can be said that while machine learning significantly enhances the accuracy of diagnosis and management of prostate cancer, its effective application remains a matter of prospective validation of models and fairness in machine learning.

Keywords: Diagnostic accuracy, explainable AI, health disparities machine learning, multimodal fusion, prostate cancer, urology.

1. Introduction

Prostate cancer remains a significant health concern on a global scale, with projections indicating that in 2025, there will be over 313,000 newly diagnosed patients in the United States alone [1]. Despite advancements in screening modalities, traditional methods of diagnosis using prostate-specific antigen (PSA) and multiparametric magnetic resonance imaging (MRI) have been shown to have high inter-observer variability and low specificity [2]. This issue of diagnosis frequently leads to unnecessary procedures or underdiagnosis of clinically significant disease. Thus, there is a dire need for objective data-driven tools that can assist in the decision-making process for physicians.

The dawn of machine learning and deep learning techniques has provided a revolutionary breakthrough in the field of urological oncology [3]. Machine learning models have been shown to have a high degree of accuracy in identifying non-linear correlations in high-dimensional data. Recent advancements in machine learning have focused on a paradigm shift from single-modality analysis to multimodal analysis, in which machine learning models are able to analyze a combination of data types to provide a more holistic risk analysis [4]. Neural networks have been able to show promise in reducing unnecessary procedures when clinical variables are included in the model in relation to MRI analysis [5]. Additionally, multimodal models have been able to show prognostic accuracy in relation to standard risk grouping [6].

Recent studies from 2021 to 2025 have shown the effectiveness of these models in various clinical environments. For instance, multi-center studies have shown the capability of deep learning algorithms to perform at levels comparable to experienced radiologists in the identification of

high-grade lesions [7]. In addition, the implementation of computerized Gleason grading using digital pathology tools has shown high concordance with consensus results from experienced pathologists, which could help reduce the time required in the diagnostic process [8]. The implementation of specific deep learning architectures in abbreviated biparametric magnetic resonance imaging (bpMRI) allows real-time risk stratification and scanning [9]. All these studies illustrate the advancement of artificial intelligence tools and their readiness for clinical practice.

Significantly, there are controversies in the adoption of machine learning, particularly in its lack of transparency in complex networks. While high-performing algorithms provide better accuracy, their lack of transparency is of great concern, particularly in clinical applications [3]. Interactive explainable artificial intelligence (XAI) models have been developed to address these issues by utilizing established imaging features to explain their classification, aiming to improve physician confidence and trust in the model [10]. Notably, there is evidence that homogeneous model training may result in decreased model performance, leading to racial bias in diverse populations, complicating the quest for equal healthcare in all populations [11].

This literature review aims to bring together primary research from the years between 2021 and 2025 to provide a summary of the state of the art in terms of diagnostic accuracy and the effectiveness of different strategies in reducing health disparities. The conclusions drawn from the primary research suggest that, although machine learning has a major impact on the management of prostate cancer, the future path should be focused on the creation of transparent and fairness-conscious algorithms so that the benefits of artificial intelligence are equitably distributed across all patient populations.

2. Materials and Methods

The systematic literature review was carried out in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standards to ensure a stringent and transparent review of contemporary literature. The methodology focuses on the identification of primary studies that explore the application of machine learning in prostate cancer management, including accuracy in diagnosis and health disparities.

2.1. Search Strategy and Data Sources

The primary objective of this review is to evaluate the diagnostic accuracy and health disparities in machine learning-based prostate cancer management. A systematic search was conducted for peer-reviewed articles published between January 2021 and February 2025. The search was performed across four databases: PubMed, Scopus, Google Scholar, and Mendeley. The search strings included combinations of keywords such as "Prostate Cancer", "Machine Learning", "Deep Learning", "Diagnostic Accuracy", and "Health Equity".

The study selection process followed the PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines to ensure transparency and reproducibility. Fig. 1 illustrates the PRISMA flow diagram, detailing the identification, screening, eligibility, and final inclusion of the studies.

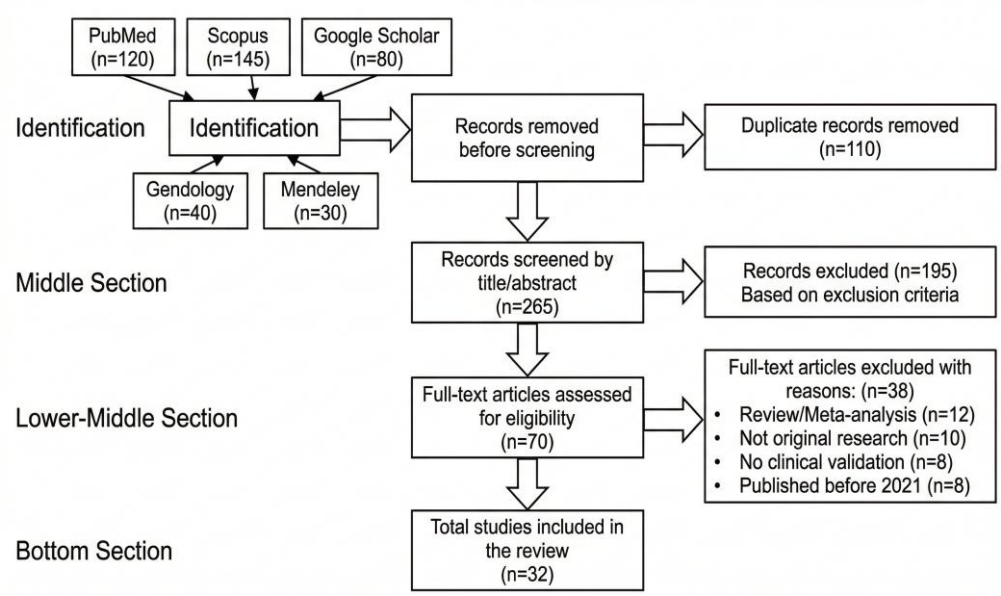

Fig. 1. PRISMA 2020 flow diagram illustrating the systematic study selection process.

The diagram details the identification of 375 records from four databases (PubMed, Scopus, Google Scholar, and Mendeley), the removal of duplicates, the screening of titles and abstracts, the exclusion of full-text articles with reasons, and the final inclusion of 32 original research studies for qualitative synthesis in the review.

2.2. Inclusion and Exclusion Criteria

Studies were included if they met the following criteria: (1) focus on machine learning or deep learning applications in prostate cancer; (2) use of medical imaging (MRI, PET-CT) or pathomics data; (3) reported performance metrics such as AUC, sensitivity, and specificity; and (4) written in English. Exclusion criteria involved: (1) review articles, editorials, or conference abstracts without full peer-reviewed data; (2) studies with a sample size of fewer than 20 patients; and (3) research published prior to 2021.

2.3. Data Extraction and Quality Assessment

Data were extracted from the included studies, focusing on the year of publication, dataset size, model architecture, and diagnostic performance. To address the methodological rigor requested for systematic reviews, each included study underwent a risk-of-bias assessment. Two reviewers independently evaluated the studies for potential bias in patient selection, index tests, reference standards, and flow and timing. Any disagreements were resolved through consensus. This critical appraisal ensures that the synthesized evidence reflects high-quality research and accounts for methodological heterogeneity across different AI models.

2.4. Synthesis of Results

The literature review results were categorized into various themes based on the clinical need and the corresponding literature. The themes were categorized into MRI-based detection approaches, digital pathology automation approaches, and the integration of multiple data modalities. The results were then analyzed to determine the trends in the overall performance of the models and the effectiveness of the approaches in reducing health disparities.

3. Results

The results provided in this section demonstrate a holistic summary of primary research literature on the application of machine learning in prostate cancer management, published from 2021 to 2025. The analysis provided here has been categorized to demonstrate a rational progression from the radiologic detection of cancer to definitive pathological confirmation and ultimate prognostic prediction. By providing these results according to specific clinical and technical domains, this section of the analysis highlights how machine learning addresses specific problems in urologic oncology, including inter-observer variability and the complexities of multi-modal data integration. The results provided also demonstrate performance on diverse populations to determine implications for health disparities.

3.1. Automated Lesion Detection and PI-RADS Refinement

The machine learning architectures that were developed were found to be highly effective in identifying and classifying suspicious lesions, especially in magnetic resonance imaging. The research that was carried out from 2021 to 2025 has shown that deep learning architectures were highly accurate when used on biparametric and multiparametric images, as can be seen from the synthesized performance data provided in Table I. Large-scale multicenter retrospective research has shown that deep learning architectures can even surpass the performance of expert radiologists in predicting clinically significant diseases [7], [12], [13].

TABLE I

Summary of key studies on automated lesion detection and PI-RADS refinement.

Study	Modality	Focus	Key Performance Metrics
Cai et al. [12]	mpMRI	Fully automated detection	AUC 0.89 (matched radiologists)
Lee et al. [14]	bpMRI	Biopsy setting / Prospective	Specificity increased from 21% to 44%
Zhang et al. [15]	T2W MRIP	Ca vs. BPH differentiation	AUC 0.927 (ProZonaNet model)
Sun et al. [16]	DWI/T2W	PSA gray zone (4–10 ng/mL)	Patient-level sensitivity 96.0%
Jin et al. [17]	T2W Only	Non-invasive Gleason prediction	AUC 0.918 for PCa detection
Huang et al. [18]	bpMRI	Simulated contrast-enhanced MRI	10.5% cases upgraded to PI-RADS 4

Notably, the accuracy of these computational systems is often improved using multi-head attention mechanisms that integrate various modality-specific features using multiple MRI sequences [19]. Furthermore, the integration of gradient-weighted class activation maps (Grad-CAM) provides visual evidence of the localization of the tumor, thus meeting the transparency requirements applicable in clinical practice settings [12]. The precise segmentation and heatmap visualization of suspicious lesions using these deep learning models are shown in Fig. 2 [20].

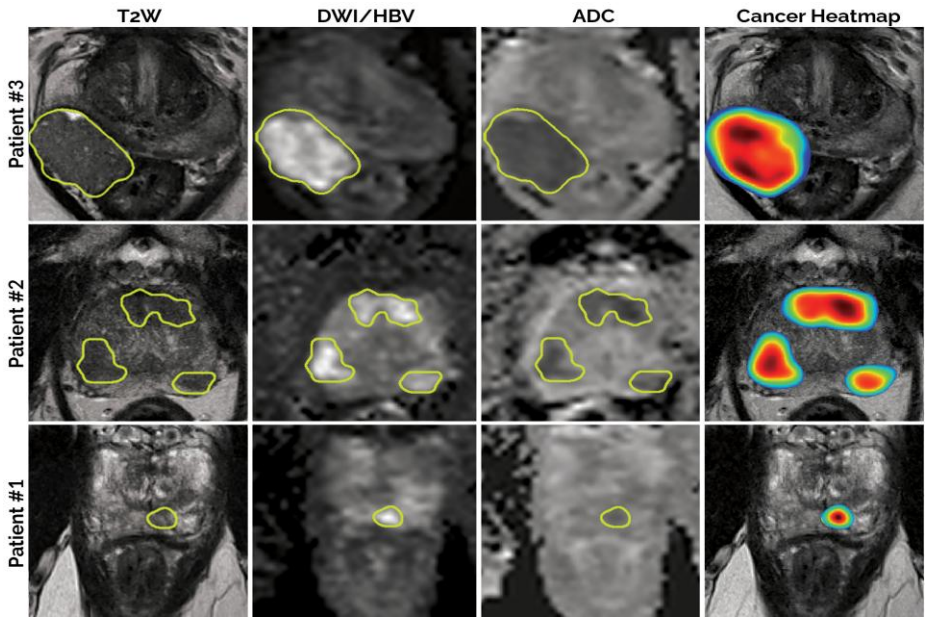

Fig. 2. Deep learning-based segmentation and Grad-CAM heatmap visualization of suspicious prostate lesions on multiparametric MRI. [20]

The application of cascaded deep learning models has been shown to be useful in the improvement of the interpretation of PI-RADS 3 lesions, which are often considered ambiguous in conventional clinical practice settings. The objective scoring system provided using these deep learning models is useful in the differentiation between indolent tissues and aggressive malignancies by clinicians. The accuracy of these deep learning models is increasingly being validated using whole-mount histopathology-referenced delineations, which are considered superior to biopsy-based samples in conventional clinical practice settings [21]. The integration of these deep learning-based scoring systems into clinical practice settings is shown to improve the overall specificity of the diagnosis from 21% to 44%, thus reducing the overall number of unnecessary biopsy procedures [14]. Notably, the integration of XAI techniques that use established imaging features to justify the classification results is shown to improve the confidence of non-expert readers while reducing the interpretation time by approximately one minute [10].

Apart from the conventional detection techniques, recent developments have also focused on overcoming specific clinical confounders and protocol-related issues. Fully automated detection tools using biparametric MRI have shown sensitivity levels of around 74% when compared with the conventional radiological interpretation process and have been established as a reliable tool to assist radiologists in the diagnosis process [2], [22]. Specific deep learning techniques using zonal volume-based biomarkers have shown an increased capability to differentiate between prostate cancer and BPH, with the AUC value rising from 0.758 to 0.927 in complex cases where the conditions are often interrelated [15]. In addition, it is now possible to generate contrast-enhanced MRI sequences from non-contrast images with high structural similarity using advanced algorithms, thus providing a safer and more cost-effective diagnostic option in cases where patients are at risk of contrast agent toxicity [18]. In order to assist in the faster development of the tools, specific techniques called semi-supervised learning have now been developed, which can achieve high diagnostic performance with fourteenfold fewer annotations using information from existing diagnostic reports [23].

The detection capabilities have been extended beyond MRI to other modalities, such as transrectal ultrasound (TRUS), which are more accessible. The use of transfer learning-based models, such as CNN models, specifically the EfficientNetV2 model, has been found to have exceptional detection potential in distinguishing prostate cancer from benign prostatic hyperplasia (BPH), with accuracy and sensitivity rates of over 0.99 [24]. Other studies, specifically those focused on the clinical gray zone of prostate-specific antigen (PSA) levels ranging from 4 to 10 ng/mL, have found that diffusion-weighted imaging (DWI) models can attain a sensitivity rate of 96.0% at the patient level, far outperforming the conventional PI-RADS methods for lesion localization [16]. The attempt to simplify the methods for prostate cancer detection includes single-modality approaches, in which the use of deep learning models for T2-weighted images has been found to have the potential for the detection of malignancy and the prediction of Gleason grade, with an area under the receiver operating characteristic curve (AUC) of 0.918, competing with the conventional multiparametric methods [17]. Other approaches, such as the CorrSigNIA, have been found to have the potential for the selective detection of indolent and aggressive components within lesions through radiology–pathology fusion [25].

3.2. Automated Histopathological Analysis and Gleason Grading

In digital pathology, machine learning has been used to minimize the inter-observer variation in tissue grading. Initial studies have shown that it is possible to automate Gleason grading using deep neural networks. Table II shows the principal performance metrics of the systems that have been developed. Quadratic weighted kappa scores above 0.90 indicate a high degree of agreement with the decisions of experienced urological pathologists [8].

TABLE II

Performance metrics of machine learning models in histopathological analysis and Gleason grading.

Study	Modality	Focus	Key Performance Metrics
Singhal et al. [8]	Core Needle Biopsy (WSI)	Automated Gleason Grading	$\kappa_{quad} > 0.90$
Erak et al. [11]	H&E Stained WSI	Genomic Subtype Prediction	Identification of ERG/ PTEN fusions
Yu et al. [26]	Biopsy Digital Images	Extraprostatic Extension (EPE)	AUC 0.886
Bhattacharya et al. [25]	Biopsy Cohorts	Mixed Lesions (CorrSigNIA)	AUC 0.86
Jin et al. [17]	T2-weighted MRI	Non-invasive Gleason Prediction	AUC 0.854

To reduce subjectivity in manual grading systems, recent studies have utilized explainable artificial intelligence, denoted by XAI, that follows the vocabulary established by pathologists and utilizes soft labels in the training process. These systems are able to address the uncertainties in the data, thus being able to segment images well in the presence of high observer variability in the data. The visual accuracy of these systems, in comparison to the interpretations made by experts, is depicted in Fig. 3 [27].

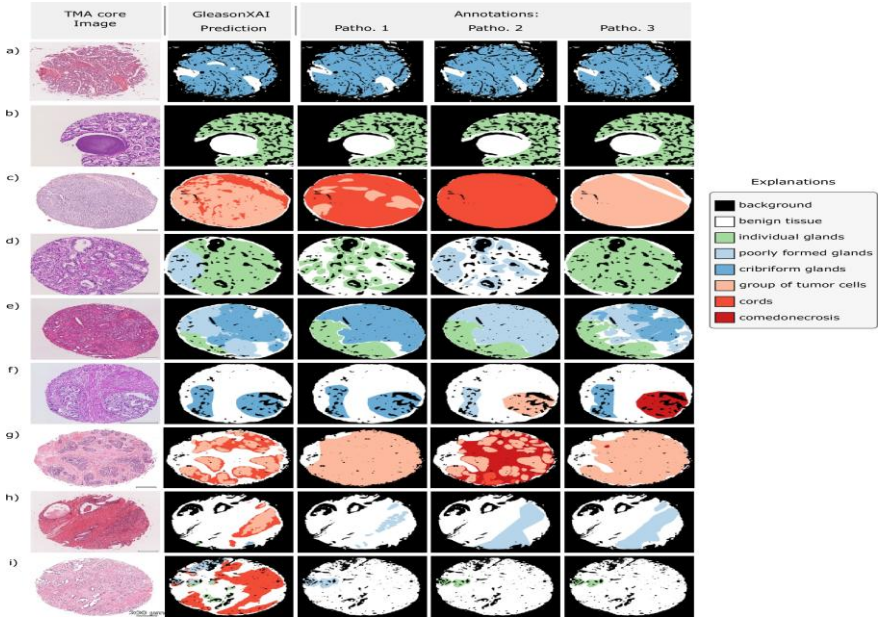

Fig. 3. Comparison of GleasonXAI segmentation results against expert annotations for Gleason patterns 3 (green), 4 (blue), and 5 (red). [27]

The strength of these algorithms lies in their ability to recognize minute morphological changes that might not be evident during the usual microscopic evaluation. In addition to the usual grading systems, recent advances have shown that deep learning-based H&E-stained WSI could be used for the detection of genomic changes such as ERG fusions and PTEN deletions, which previously lacked reproducible morphological correlations [11]. Further, the AI-based pathology evaluation has been able to make accurate predictions regarding the occurrence of extraprostatic extension (EPE) directly from preoperative biopsy slides, achieving an AUC of 0.886. This has provided a novel prognostic tool with a strong correlation with biochemical recurrence [26]. By expediting the evaluation of high-risk cases, artificial intelligence has proven to be a critical second opinion for the final diagnosis, making the process more efficient.

Recent frameworks have attempted to address the complexities associated with the presence of mixed lesions characterized by the presence of multiple grades of cancer. The CorrSigNIA framework, which utilizes a radiology-pathology fusion approach to detect the presence of aggressive Gleason Pattern 4 tissues along with indolent tissues in the tumor, has achieved an AUC of 0.86 on biopsy cohorts [25]. Such developments ensure precise targeted biopsy procedures and help clinicians avoid overtreatment of low-grade lesions. Further, non-invasive prediction of the Gleason grade using single-modality T2-weighted MRI imaging techniques has achieved an AUC of 0.854 in multi-center validation studies, thus emphasizing the need to reduce the need for invasive biopsy procedures using more precise risk assessment techniques prior to surgery [17].

3.3. Multimodal Data Fusion for Risk Stratification and Prognosis

Multimodal integration has been shown to possess superior performance in the prediction of systemic risks, for instance, in the prediction of lymph node invasion. Current models that combine PSMA PET/CT deep learning features with clinical parameters possess superior performance in the prediction of LNI, with an area under the receiver operating characteristic curve (AUC) of 0.89. This is superior to traditional clinical tools, which include the Briganti-2017 and MSKCC

nomograms [28]. The various integration approaches and the performance metrics reported in recent literature are summarized in Table III.

TABLE III

Performance metrics of machine learning models in histopathological analysis and Gleason grading.

Study	Modality	Focus	Key Performance Metrics
Ma et al. [28]	PSMA PET/CT + Clinical	Lymph Node Invasion (LNI)	AUC 0.89
Bacchetti et al. [5]	bpMRI + Clinical Features	Biopsy Avoidance	43.4% reduction in biopsies
Yao et al. [29]	mpMRI + 18F-PSMA-PET/CT	Extraprostatic Extension (EPE)	AUC 0.82
Zhang et al. [30]	MRI Radiomics + Pathomics	Bone Metastasis Prediction	AUC 0.93
Esteva et al. [6]	Multi-modal (Phase III Trials)	Prognostic Prediction	9.2% to 14.6% over NCCN
Lee et al. [31]	mpMRI DL + Clinical	BCR-Free Survival	Improved long-term prediction

The integration of various data streams has been extended to include clinical data along with radiological data. Integrated nomograms that use the predictions of deep learning with clinical data, such as PSA density, patient age, and PI-RADS scores, have been shown to reduce unnecessary biopsies by a significant percentage [4], [5]. For example, the integration of clinical data with MRI-based data has the potential to avoid up to 43.4% of unnecessary biopsies while retaining high sensitivity to detect cancer [5]. Additionally, the use of various clinical trials of phase III to validate the performance of the multimodal architecture has resulted in relative improvements in the prognostic performance of the model ranging from 9.2% to 14.6% compared to the standard NCCN risk group stratification. This has the potential to provide highly personalized treatments based on predictions of long-term outcomes [6]. A conceptual framework of the integrated diagnostic pipeline has been presented in Fig. 4.

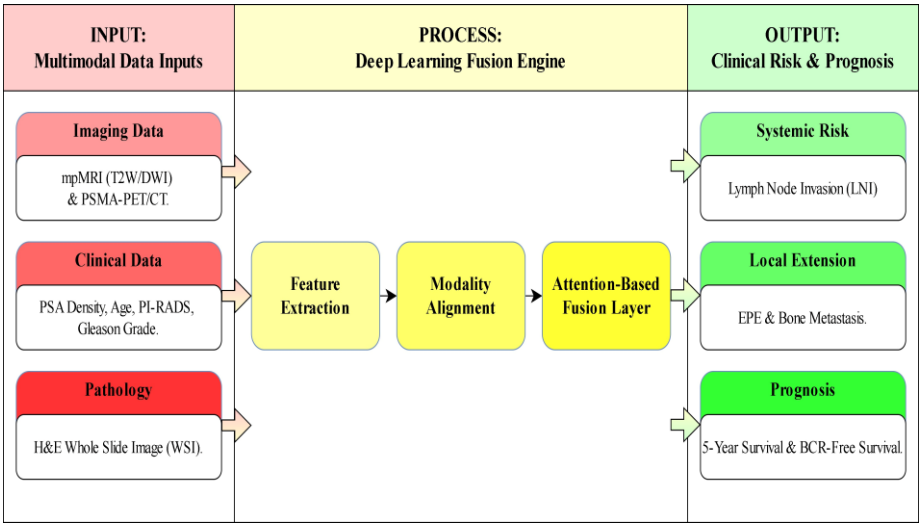

Fig. 4. Multimodal data fusion architecture for prostate cancer risk assessment and prognostic prediction.

In addition to conventional clinical information, the integration of advanced imaging and pathomics provides critical information for the progression and metastasis of cancer. The integration of mpMRI and 18F-PSMA-PET/CT significantly increases the ability to predict the probability of extraprostatic extension (EPE), which has been reported to achieve an area under the receiver operating characteristic curve (AUC) of 0.82, with a greater clinical net benefit compared to the assessment by a radiologist alone [29]. The integration of multimodal MRI radiomics and pathomics from tissue sections significantly increases the ability to predict the probability of bone metastasis, which has been reported to achieve an AUC of 0.93 [30]. This helps in identifying the potential risks at the systemic level that may be missed by conventional diagnostic approaches. Predictive models for the probability of biochemical recurrence (BCR) also benefit from this approach, in which clinical-deep learning models using features from mpMRI outperform conventional radiomics in the ability to forecast long-term postoperative survival [31].

3.4. Evaluation of Model Fairness and Mitigation of Health Disparities

Studies on the performance of such models have shown a significant performance gap in the diagnosis of diverse populations when models are trained on homogeneous data. Studies have shown that the performance of machine learning models decreases in non-white populations, which is a concern in the delivery of healthcare services to diverse populations [13]. This performance degradation of models is common in external validation in multicenter studies, where diverse populations are not represented in the training data [12], [13], [32].

By utilizing domain-agnostic feature learning, researchers have created models that maintain high diagnostic accuracy while promoting equal outcomes across a range of demographic populations [8]. The use of architectures that have undergone randomized phase III clinical trials is a significant improvement in the attempt to mitigate these disparities, as these trials provide a more representative sample of global diversity in patients compared to a single-retrospective cohort study population [6]. Additionally, the trend of multimodal fusion in incorporating race-specific clinical factors and family history has shown potential in mitigating the bias that is inherently present in traditional risk assessment tools [32]. Each study that addresses these disparities, along with the potential solutions, is outlined in Table IV.

TABLE IV

Summary of studies investigating model fairness and bias mitigation in prostate cancer.

Study	**Population/ Cohort Diversity**	**Identified Fairness Challenge**	**Mitigation / Strategy**
Esteva et al. [6]	Randomized Phase III Trials	Global patient diversity	Multi-modal deep learning on trial data
Singhal et al. [8]	Multi-institutional sets	Site-specific data bias	Domain-agnostic feature learning
Zhao et al. [13]	7-hospital multicentre study	External validation gaps	Diverse retrospective validation
Esteban et al. [32]	Multicentre radiological data	Racial bias in risk tools	Clinical and radiological integration

Therefore, it is vital to address these health disparities at the system level to ensure the equitable distribution of the benefits of technological innovation in the field of oncology among all patients regardless of their racial or ethnic background. The advancement of inclusive AI is supported by the development of frameworks with high concordance rates among multi-institutional external datasets, which effectively eliminate the occurrence of bias resulting from the data characteristics of each institution [8], [25]. These findings provide a vital roadmap for the development of fairness-aware algorithms that prioritize health equity alongside diagnostic precision.

3.5. Comparative Synthesis of Clinical Performance

Overall, the results of the primary studies highlight the maturity of machine learning from being a research tool to a clinical decision support system. The synthesized results highlight the potential of the recent machine learning approaches with the integration of technical explainability and multimodal integration in optimizing the clinical management of patients with the highest level of accuracy and efficacy [10], [12], [14]. The recent machine learning approaches have the capability to provide robust, validated, and transparent results to help clinicians cope with the complexities associated with the diagnosis and treatment of prostate cancer. The machine learning approaches have the capability to perform at the level of experienced experts in the field of radiologic detection and pathologic grading of the disease and provide an unbiased second opinion to clinicians to eliminate inter-observer variability [8], [12], [17].

From the evidence, the most effective clinical applications of these systems are not limited to the use of single-modality systems. Integrated nomograms, as well as multimodal fusion systems, have proven to exhibit superior specificity, an attribute that is vital in minimizing the need for unnecessary procedures such as biopsies [4], [5], [14]. Moreover, the development of these systems to include prognostic capabilities, such as predicting extraprostatic extension, as well as biochemical recurrence, allows for a more comprehensive analysis of patients' risks along the continuum of care [6], [26], [31]. The transition towards the use of whole-mount histopathology as a gold standard, in addition to the implementation of prospective multicenter validations, has significantly increased the clinical reliability of these systems [13], [15], [21].

Lastly, the emphasis on fairness in models and the creation of domain-agnostic features underscore the importance of generalizability in clinical practice. Models that can maintain high performance on diverse populations and datasets are a critical roadmap to equitable healthcare delivery [8], [25], [32]. This evolution towards transparent, fair, and multimodal artificial intelligence marks a substantial milestone in the achievement of personalized and precision urology. The synthesized clinical maturity of these diverse applications is summarized in Table V.

TABLE V

Comparative synthesis of machine learning maturity across clinical domains.

Clinical Domain	**Key Technical Advantage**	**Maturity Level**	**Primary Evidence (Reference)**
Radiological Detection	PI-RADS 3 refinement, zonal biomarkers, and whole-mount referencing	High (Prospective)	J. C. Cai et al. [12], Y. J. Lee et al. [14], Z. Zhang et al. [15], D. Li et al. [21]

Clinical Domain	Key Technical Advantage	Maturity Level	Primary Evidence (Reference)
Pathological Grading	Automated Gleason grading with κ > 0.90 and non-invasive MRI prediction	High (Validated)	N. Singhal et al. [8], L. Jin et al. [17]
Risk Stratification	Multimodal fusion (MRI + Clinical) and radiology-pathology correlation	Emerging	A. Hiremath et al. [4], E. Bacchetti et al. [5], I. Bhattacharya et al. [25]
Prognostic Tracking	BCR survival, EPE prediction, and Phase III trial validation	Emerging	A. Esteva et al. [6], P. Yu et al. [26], H. W. Lee et al. [31]
Health Equity	Fairness-aware algorithms and multicenter bias mitigation	In-Development	L. Zhao et al. [13], L. M. Esteban et al. [32]
Clinical Decision Support	Interactive XAI to improve non-expert reader confidence and speed	Emerging	C. A. Hamm et al. [10]

4. Discussion

The results of this review suggest that machine learning has clearly moved beyond the realm of experimental proof of concept to a point of clinical validation. The high degree of accuracy achieved in each of these primary studies indicates that machine learning can greatly reduce the subjective nature of traditional prostate cancer screening methods [12], [13], [14]. Machine learning tools promise to provide a standard interpretation of complex data, which can be used to reduce inter-observer variability in radiologists' and pathologists' readings of data [8], [28]. This is a particularly important issue when discussing the detection of clinically significant prostate cancer, as the ability of machine learning to differentiate between indolent and aggressive disease can prevent morbidity from overdiagnosis and overtreatment of disease [25].

One of the prominent themes that have been observed in contemporary original research studies is the performance of multimodal fusion systems in comparison to single-modality systems. Integrating magnetic resonance imaging, genomic markers, clinical history, and advanced molecular imaging modalities such as PSMA PET/CT can provide a more detailed understanding of tumor biology [6], [28]. Studies have suggested that these multimodal systems are more suitable to cope with the heterogeneity of prostate cancer, providing more effective prognostic markers for biochemical recurrence and patient survival [8], [31]. From a clinical perspective, the emergence of multimodal systems in prostate cancer imaging can be regarded as a step towards personalized urology, in which treatment plans are tailored to the individual biological and radiological profiles of patients [6].

Despite the performance gains, the clinical adoption of machine learning has been limited by the lack of transparency. The lack of transparency of deep learning models has been cited as a major point of contention by various medical practitioners [10]. However, original research published between 2021 and 2025 has shown that XAI can be used to improve transparency by providing visual maps, such as Grad-CAMs, that match clinical reasoning [10], [12]. By making the

decision-making process of the algorithm transparent to the practitioner, trust can be fostered towards the adoption of AI in urological practice.

Furthermore, this review also focuses on the ethical need for addressing the issue of disparity within the medical informatics community. Empirical research has shown that the generalizability of the model may be restricted when the dataset used for training the model is narrow or homogeneous, thus resulting in diagnostic disparity [13]. However, the development of fairness-aware machine learning models is a significant step towards addressing the issue. Recent research has shown that the use of domain-agnostic feature learning and the deliberate inclusion of diverse cohorts in the model may result in fair diagnostic performance for diverse demographics [8], [25]. This focus on equity is a critical need in order for the advancements in technology to benefit all patients, irrespective of their race or ethnicity.

This will be particularly important in the validation of machine learning models, as they move from the research to the clinical practice stage. Although the results of the present study are promising, many models still need to be validated in different real-world settings, characterized by different machines and approaches [14]. Future studies will have to address the challenges of implementing such models in a way that helps, as opposed to hindering, the clinician. To conclude, the continued development of transparent, multimodal, and equitable artificial intelligence will be at the heart of future strategies in the management of prostate cancer.

5. Conclusions

In conclusion, the literature on original research articles published within the time frame of 2021 to 2025 indicates that machine learning has advanced significantly in the management of prostate cancer. The literature indicates that, particularly, deep learning models have achieved considerable accuracy in the detection of clinically significant prostate cancer, with the accuracy metrics being at par with those of experienced radiology and pathology experts. The use of data fusion techniques has been identified as the most effective technique in the improvement of the accuracy of machine learning models. The development of explainable artificial intelligence has been identified as a crucial technique in addressing the transparency of machine learning models.

However, there are considerable gaps that need to be addressed in the integration of machine learning in the management of prostate cancer. For instance, there have been concerns over the racial bias of machine learning models, as indicated in various studies published within the last five years. The literature indicates that fair machine learning models are crucial in the provision of equitable healthcare to all. Although there have been considerable validations of machine learning tools, there is a need for further studies to confirm the validity of machine learning tools. The potential of machine learning in the management of prostate cancer cannot be overstated. However, the potential of machine learning in the management of prostate cancer can be achieved through transparency and equitable patient care.

References

[1] R. L. Siegel, T. B. Kratzer, A. N. Giaquinto, H. Sung, and A. Jemal, “Cancer statistics, 2025,” *CA. Cancer J. Clin.*, vol. 75, no. 1, pp. 10–45, Jan. 2025, doi: 10.3322/caac.21871.

[2] N. Debs, A. Routier, A. Bône, and M.-M. Rohé, “Evaluation of a deep learning prostate cancer detection system on biparametric MRI

against radiological reading," *Eur. Radiol.*, vol. 35, no. 6, pp. 3134–3143, 2025, doi: 10.1007/s00330-024-11287-1.

[3] A. Esteva *et al.*, "Deep learning-enabled medical computer vision.," *NPJ Digit. Med.*, vol. 4, no. 1, p. 5, Jan. 2021, doi: 10.1038/s41746-020-00376-2.

[4] A. Hiremath *et al.*, "An integrated nomogram combining deep learning, Prostate Imaging-Reporting and Data System (PI-RADS) scoring, and clinical variables for identification of clinically significant prostate cancer on biparametric MRI: a retrospective multicentre study," *Lancet Digit. Heal.*, vol. 3, no. 7, pp. e445–e454, Jul. 2021, doi: 10.1016/S2589-7500(21)00082-0.

[5] E. Bacchetti *et al.*, "A Deep Learning Model Integrating Clinical and MRI Features Improves Risk Stratification and Reduces Unnecessary Biopsies in Men with Suspected Prostate Cancer," *Cancers (Basel).*, vol. 17, pp. 1–17, 2025, doi: 10.3390/cancers17132257.

[6] A. Esteva *et al.*, "Prostate cancer therapy personalization via multi-modal deep learning on randomized phase III clinical trials," *npj Digit. Med.*, vol. 5, no. 1, p. 71, 2022, doi: 10.1038/s41746-022-00613-w.

[7] J. M. Castillo T. *et al.*, "Classification of Clinically Significant Prostate Cancer on Multi-Parametric MRI: A Validation Study Comparing Deep Learning and Radiomics," 2022. doi: 10.3390/cancers14010012.

[8] N. Singhal *et al.*, "A deep learning system for prostate cancer diagnosis and grading in whole slide images of core needle biopsies," *Sci. Rep.*, vol. 12, no. 1, p. 3383, 2022, doi: 10.1038/s41598-022-07217-0.

[9] P. M. Johnson *et al.*, "Prostate Cancer Risk Stratification and Scan Tailoring Using Deep Learning on Abbreviated Prostate MRI," *J. Magn. Reson. Imaging*, vol. 62, no. 3, pp. 858–866, Sep. 2025, doi: 10.1002/jmri.29798.

[10] C. A. Hamm *et al.*, "Interactive Explainable Deep Learning Model Informs Prostate Cancer Diagnosis at MRI," *Radiology*, vol. 307, no. 4, p. e222276, Apr. 2023, doi: 10.1148/radiol.222276.

[11] E. Erak *et al.*, "Predicting Prostate Cancer Molecular Subtype With Deep Learning on Histopathologic Images.," *Mod. Pathol.*, vol. 36, no. 10, p. 100247, Oct. 2023, doi: 10.1016/j.modpat.2023.100247.

[12] J. C. Cai *et al.*, "Fully Automated Deep Learning Model to Detect Clinically Significant Prostate Cancer at MRI," *Radiology*, vol. 312, no. 2, p. e232635, Aug. 2024, doi: 10.1148/radiol.232635.

[13] L. Zhao *et al.*, "Predicting clinically significant prostate cancer with a deep learning approach: a multicentre retrospective study," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 50, no. 3, pp. 727–741, 2023, doi: 10.1007/s00259-022-06036-9.

[14] Y. J. Lee *et al.*, "MRI-based Deep Learning Algorithm for Assisting Clinically Significant Prostate Cancer Detection: A Bicenter Prospective Study," *Radiology*, vol. 314, no. 3, p. e232788, Mar. 2025, doi: 10.1148/radiol.232788.

[15] Z. Zhang *et al.*, "A deep learning derived prostate zonal volume-based biomarker from T2-weighted MRI to distinguish between prostate cancer and benign prostatic hyperplasia," *Med. Phys.*, vol. 52, no. 8, p. e18053, Aug. 2025, doi: 10.1002/mp.18053.

[16] Z. Sun *et al.*, "Deep-Learning Models for Detection and Localization of Visible Clinically Significant Prostate Cancer on Multi-Parametric MRI," *J. Magn. Reson. Imaging*, vol. 58, no. 4, pp. 1067–1081, Oct. 2023, doi: 10.1002/jmri.28608.

[17] L. Jin, Z. Yu, F. Gao, and M. Li, "T2-weighted imaging-based deep-learning method for noninvasive prostate cancer detection and Gleason grade prediction: a multicenter study," *Insights Imaging*, vol. 15, no. 1, p. 111, 2024, doi: 10.1186/s13244-024-01682-z.

[18] H. Huang *et al.*, "Deep Learning to Simulate Contrast-Enhanced MRI for Evaluating Suspected Prostate Cancer," *Radiology*, vol. 314, no. 1, p. e240238, Jan. 2025, doi: 10.1148/radiol.240238.

[19] Y. Xu, R. Wang, Z. Fang, and J. Tang, "Feasibility study of AI-assisted multi-parameter MRI diagnosis of prostate cancer," *Sci. Rep.*, vol. 15, no. 1, p. 10530, 2025, doi: 10.1038/s41598-024-84516-8.

[20] A. Saha, M. Hosseinzadeh, and H. Huisman, "End-to-end prostate cancer detection in bpMRI via 3D CNNs: Effects of attention mechanisms, clinical priori and decoupled false positive reduction," *Med. Image Anal.*, vol. 73, p. 102155, 2021, doi: 10.1016/j.media.2021.102155.

[21] D. Li *et al.*, "Deep Learning in Prostate Cancer Diagnosis Using Multiparametric Magnetic Resonance Imaging With Whole-Mount Histopathology Referenced Delineations," *Front. Med.*, vol. Volume 8-, 2022, doi: 10.3389/fmed.2021.810995.

[22] S. Mehralivand *et al.*, "Deep learning-based artificial intelligence for prostate cancer detection at biparametric MRI.," *Abdom. Radiol. (New York)*, vol. 47, no. 4, pp. 1425–1434, Apr. 2022, doi: 10.1007/s00261-022-03419-2.

[23] J. S. Bosma, A. Saha, M. Hosseinzadeh, I. Slootweg, M. de Rooij, and H. Huisman, "Semisupervised Learning with Report-guided Pseudo

Labels for Deep Learning–based Prostate Cancer Detection Using Biparametric MRI," *Radiol. Artif. Intell.*, vol. 5, no. 5, p. e230031, Jul. 2023, doi: 10.1148/ryai.230031.

[24] T.-L. Huang *et al.*, "Transfer learning with CNNs for efficient prostate cancer and BPH detection in transrectal ultrasound images," *Sci. Rep.*, vol. 13, no. 1, p. 21849, 2023, doi: 10.1038/s41598-023-49159-1.

[25] I. Bhattacharya *et al.*, "Selective identification and localization of indolent and aggressive prostate cancers via CorrSigNIA: an MRI-pathology correlation and deep learning framework," *Med. Image Anal.*, vol. 75, p. 102288, 2022, doi: 10.1016/j.media.2021.102288.

[26] P. Yu *et al.*, "Deep learning model for predicting extraprostatic extension of prostate cancer based on H&E-stained biopsy digital images," *Ann. Med.*, vol. 57, no. 1, p. 2547094, Dec. 2025, doi: 10.1080/07853890.2025.2547094.

[27] G. Mittmann *et al.*, "Pathologist-like explainable AI for interpretable Gleason grading in prostate cancer," *Nat. Commun.*, vol. 16, no. 1, p. 8959, 2025, doi: 10.1038/s41467-025-64712-4.

[28] Q. Ma *et al.*, "PSMA PET/CT based multimodal deep learning model for accurate prediction of pelvic lymph-node metastases in prostate cancer patients identified as candidates for extended pelvic lymph node dissection by preoperative nomograms," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 52, no. 6, pp. 2063–2075, 2025, doi: 10.1007/s00259-024-07065-2.

[29] F. Yao *et al.*, "Multimodal imaging deep learning model for predicting extraprostatic extension in prostate cancer using MpMRI and 18 F-PSMA-PET/CT," *Cancer Imaging*, vol. 25, no. 1, p. 103, 2025, doi: 10.1186/s40644-025-00927-4.

[30] Y.-F. Zhang *et al.*, "Deep learning algorithm-based multimodal MRI radiomics and pathomics data improve prediction of bone metastases in primary prostate cancer," *J. Cancer Res. Clin. Oncol.*, vol. 150, no. 2, p. 78, 2024, doi: 10.1007/s00432-023-05574-5.

[31] H. W. Lee, E. Kim, I. Na, C. K. Kim, S. I. Seo, and H. Park, "Novel Multiparametric Magnetic Resonance Imaging-Based Deep Learning and Clinical Parameter Integration for the Prediction of Long-Term Biochemical Recurrence-Free Survival in Prostate Cancer after Radical Prostatectomy," 2023. doi: 10.3390/cancers15133416.

[32] L. M. Esteban *et al.*, "Integrating radiological and clinical data for clinically significant prostate cancer detection with machine learning techniques," *Sci. Rep.*, vol. 15, no. 1, p. 4261, 2025, doi: 10.1038/s41598-025-88297-6.