DIR

Diagn Interv Radiol 2026; DOI: 10.4274/dir.2026.263844

ABDOMINAL IMAGING

ORIGINAL ARTICLE

Global trends and hotspots of machine learning in the diagnosis of prostate cancer: a bibliometric analysis from 1997 to 2024

Yixiao Shi

Operating Room of Anesthesia Surgery Center, West China Hospital/West China School of Nursing, Sichuan University, Sichuan, China

PURPOSE

Prostate cancer (PCa) diagnosis has advanced with the integration of machine learning (ML). This study analyzes global trends in ML-based PCa diagnosis research using bibliometric methods.

METHODS

A systematic search was conducted in the Web of Science Core Collection database for articles published between 1997 and 2024. Bibliometric analysis was performed using VOSviewer (v1.6.20), CiteSpace (v6.3.R1), and R (v4.3.3).

RESULTS

The analysis included 1,045 articles, involving 6,704 authors from 4,762 institutions across 327 countries or regions. The number of publications increased over time, rising sharply from 2018. China published the highest number of articles (290), whereas the United States demonstrated the greatest research impact, leading in total citations (8,207) and international collaborations. The Berlin Institute of Health published the highest number of articles (120). Within this dataset, European Radiology had the highest H-index. Key authors included Stephan C, Jung K, and Cammann H. Keyword analysis identified "system," "MRI," and "guidelines" as prominent terms, with emerging trends focusing on "convolutional neural network," "data system," and "transfer learning."

CONCLUSION

ML in PCa diagnosis has advanced substantially, transitioning from fundamental biomarker investigations to sophisticated deep learning applications centered on medical imaging. Future directions emphasize the development of accurate, generalizable ML models integrated into clinical workflows with a continued focus on convolutional neural networks and transfer learning.

CLINICAL SIGNIFICANCE

This study delineates the global research evolution of maching learning for prostate cancer diagnosis, offering clear clinical guidance for the translation and routine application of artificial intelligence-assisted diagnostic models.

KEYWORDS

Bibliometrics, research trends, diagnosis, machine learning, prostate cancer

Prostate cancer (PCa) is one of the most common malignancies in men and remains a major contributor to cancer-related morbidity worldwide.[1] Recent estimates indicate that PCa is the second most frequently diagnosed cancer among men globally, with approximately 1.4 million new cases annually, and its incidence increases sharply after age 50.[2,3] Because early-stage PCa is often asymptomatic and diagnostic pathways are heterogeneous, improving diagnostic accuracy and efficiency remains clinically important.

Machine learning (ML), a key branch of artificial intelligence (AI), has increasingly been applied to medical diagnostics by identifying patterns in high-dimensional data that may be difficult to detect using conventional approaches.[4] To improve detection, risk stratification, and clinical decision support in PCa diagnosis, ML has been explored across multiple data modali-

Corresponding author: Yixiao Shi

E-mail: sdian314748761@163.com

Received 31 December 2025; revision requested 13 February 2026; accepted 30 March 2026.

Epub: 30.04.2026

Publication date:

DOI: 10.4274/dir.2026.263844

You may cite this article as: Shi Y. Global trends and hotspots of machine learning in the diagnosis of prostate cancer: a bibliometric analysis from 1997 to 2024. *Diagn Interv Radiol*. 30 April 2026 DOI: 10.4274/dir.2026.263844 [Epub Ahead of Print].

ties, including multiparametric magnetic resonance imaging (mpMRI), digital pathology, serum and urine biomarkers, and electronic health records.[5] However, translation into practice is constrained by issues such as data heterogeneity, limited external validation, interpretability, and regulatory and ethical requirements.[6] Therefore, a structured overview of how this research area has evolved and which themes are currently emerging may help guide future methodological and clinical work.

Bibliometric analysis quantitatively characterizes scientific literature by identifying influential contributors (authors, institutions, and journals), collaboration patterns, and evolving research topics using citation- and keyword-based mapping.[7,8] Several recent bibliometric studies have examined AI in PCa-related research. For example, Wei et al.[9] reported rapid growth in AI–PCa publications over the last decade, with the United States and China as major contributors and an increasing emphasis on deep learning for imaging tasks. However, existing bibliometric work often focuses on broader AI applications or more recent time windows and may not isolate diagnostic research as a primary theme.[10]

To extend and complement prior studies, we conducted a bibliometric analysis of ML-based PCa diagnosis research. Our objectives were to quantify publication and citation trends over a longer historical period; map global collaboration networks among countries, institutions, and authors; identify influential journals and references; and detect research hotspots and emerging topics through keyword co-occurrence and burst analyses. By explicitly focusing on diagnostic applications and extending the timeline, this study aims to clarify how current themes (e.g., MRI- and deep-learning–driven approaches) align with or diverge from prior bibliometric findings, and to provide context for future research directions and clinical translation.

Main points

- Global research on machine learning (ML) in prostate cancer diagnosis has grown rapidly, with publication output surging from 2018 and a diverse, international research community contributing across 327 regions.
- China leads in publication quantity, whereas the United States drives research impact, visibility, and collaboration, as reflected in citation counts and extensive international partnerships.
- Major institutions and authors shape the field, including the Berlin Institute of Health as a leading contributor and European Radiology as a top publisher by H-index; key researchers include Stephan C, Jung K, and Cammann H.
- Research themes center on imaging-driven ML, particularly magnetic resonance imaging and system-level approaches, with methodological emphasis on convolutional neural networks, data systems, and transfer learning.
- The field has evolved from biomarker-focused studies to advanced deep learning applications, with future work focusing on accurate, generalizable models integrated into clinical workflows.

Search strategies and data source

This study was designed as a descriptive, cross-sectional bibliometric analysis of published literature on ML in PCa diagnosis. The literature search was conducted in the Web of Science Core Collection (WoSCC), which was selected because it provides standardized citation metadata and cited-reference fields that are directly compatible with common bibliometric tools (e.g., VOSviewer, CiteSpace, and bibliometrix).

The search query applied was [WoSCC, topic search (TS)]: TS = ("prostat* cancer*" OR "prostat* carcinoma" OR "prostat* adenocarcinoma" OR "prostat* neoplasm") AND TS = ("machine learning" OR "artificial intelligence" OR "machine intelligence" OR "deep learn*" OR "neural network*" OR "natural language process*" OR "hybrid intelligent system*" OR "CNN" OR "LSTM" OR "RNN") AND TS = (diagnos*).[11] In WoSCC, TS searches the title, abstract, author keywords, and Keywords Plus. The wildcard diagnos* was used to capture multiple diagnostic-related word forms (e.g., diagnosis, diagnostic, diagnosed). As a verification step, we repeated the search using TS = (diagnosis OR diagnostic OR diagnosed) and confirmed that the retrieved record count was identical.

The search covered articles published between January 1, 1997, and October 8, 2024. To minimize bias from database updates, all records were retrieved and exported on a single day (October 8, 2024).

Records were included if they met all of the following criteria: 1. topic relevance: focused on PCa diagnosis and involved ML or AI methods (including deep learning or neural networks); 2. time window: published between 1997 and 2024 (up to October 8, 2024); 3. language: English; and 4. document type: Article (as classified by WoSCC).

Records were excluded if they met any of the following criteria: 1. not related to diagnosis (e.g., ML studies limited to prognosis, treatment planning, basic laboratory mechanisms, or non-diagnostic screening without diagnostic intent); 2. not specific to PCa (e.g., multi-cancer studies without PCa-specific diagnostic analysis); 3. non-article document types (e.g., reviews, editorials, letters, meeting abstracts) if not indexed as "Article" in WoSCC; and 4. duplicates identified during export or processing.

No additional filter was applied to remove "Early Access" items. Therefore, any records indexed by WoSCC as "Article"—including those labeled as Early Access or originating from conference proceedings but categorized as an Article—were eligible for inclusion. We acknowledge that indexing timing for Early Access items may influence publication counts in the most recent year.

The exported records included publication year, title, authors, affiliations, countries or regions, journal information, keywords, citation counts, and cited references, and were saved in plain-text format for subsequent bibliometric analysis. Because this study analyzed publicly available bibliographic metadata and did not involve human participants or identifiable private information, ethical approval and informed consent were not required.

Statistical analysis and visualization

Data preprocessing and cleaning

Before mapping, we performed data cleaning to reduce noise caused by inconsistent WoSCC metadata. During data processing, we encountered typical bibliometric challenges, including author name ambiguity (e.g., authors with similar names or different initials), keyword unification (synonyms and abbreviations), and institutional name variations (e.g., the same organization recorded under multiple spellings).

We therefore applied a combination of automatic normalization and manual checking. Specifically, author names were standardized by merging obvious variants based on consistent co-author patterns, affiliations, and research topics (e.g., abbreviated vs. full forms such as "Stephan C" and "Stephan Carsten" when they referred to the same individual). Keywords were normalized by merging common abbreviations and equivalent terms (e.g., "MRI" and "magnetic resonance imaging"), as well as plural and singular forms and closely related expressions with the same meaning. Institutional names were harmonized by consolidating clear spelling variants and alternative English names for the same institution.

In addition, we screened for metadata inconsistencies (e.g., missing affiliations or incomplete fields) and performed manual verification when necessary. Duplicate records were removed when identified.

Bibliographic records exported from WoSCC were analyzed using Microsoft Excel for descriptive statistics (e.g., annual publication counts, total citations, and averages). All bibliometric analyses were conducted using the same dataset exported on October 8, 2024. Visualization and network analyses were conducted using VOSviewer (Centre for Science and Technology Studies, Leiden University, Leiden, the Netherlands); CiteSpace (Drexel University, Philadelphia, PA, USA); R version 4.3.3 (R Foundation for Statistical Computing, Vienna, Austria) with the bibliometrix package.

When applicable, a VOSviewer thesaurus file was used to merge equivalent terms and improve the interpretability of keyword maps. VOSviewer was used to construct and visualize co-authorship and collaboration networks (countries, institutions, and authors), as well as keyword co-occurrence networks.[12] The association strength normalization method was used for network construction, and clusters were detected using VOSviewer's built-in clustering algorithm (based on modularity optimization). To improve readability and reduce sparsity, minimum-occurrence thresholds were applied; specifically, keywords were included in the co-occurrence analysis only if they appeared at least eight times. In VOSviewer maps, node size represents occurrence frequency (or publication output), link thickness represents total link strength, and colors indicate clusters or overlay attributes (e.g., average publication year).

CiteSpace was used for temporal analysis of research topics, including timeline and time-zone visualization and keyword burst detection.[13] To highlight major knowledge structures and emergent terms, the following settings were applied: time slicing: 1997–2024 (1 year per slice); node type: keywords; selection criteria: top N per slice; and pruning: Pathfinder (including pruning of merged networks). Burst strength and duration were reported to identify rapidly increasing keywords.

The bibliometrix package was used to compute bibliometric indices, including the H-index, G-index, and M-index, based on citations within the retrieved dataset.[14] The H-index is defined as the maximum value h such that an author or journal has at least h papers, each cited at least h times. The G-index is the largest value g such that the top *g* publications received at least g^2 total citations. The M-index is calculated as the H-index divided by the number of years since the first publication in the dataset, providing a time-normalized measure of impact.[15]

To describe the growth in annual publications, we fitted a simple regression model in Excel and R and reported the coefficient of determination (R^2). This model was used descriptively to summarize the overall growth pattern; no hypothesis testing (e.g., formal trend tests) was performed. Journal influence metrics [e.g., 2023 Journal Impact Factor and Journal Citation Reports quartile (Q1–Q4)] were obtained from the 2023 edition of Journal Citation Reports.

Results

Overview of publications

A total of 1,045 English-language articles on ML in PCa diagnosis published between 1997 and 2024 were analyzed (Figure 1). These studies involved 6,704 authors from 4,762 institutions across 327 countries or regions, were published in 352 journals, and cited 29,554 sources. Each publication had an average of 8.73 co-authors, and 30.81% involved international collaboration. The dataset included 2,052 keywords, and articles received an average of 24.61 citations (Figure 2A).

In the early years (1997–2012), research on ML in PCa diagnosis showed limited output. By 2013, annual publications exceeded 10, followed by a gradual increase from 2013 to 2018. From 2018, research output increased rapidly, peaking at 224 articles in 2023. A fitted curve (R^2: 0.8427) highlights the accelerating growth in publications over the period (Figure 2B).

Distribution and collaborative networks of countries

The 1,045 publications originated from 327 countries or regions. The top contributing countries or regions are summarized in Supplementary Table 1. China led in corresponding-author output with 290 articles (27.8%) and ranked second in total citations (4,762). The USA followed with 199 corresponding-author articles (19.0%), the highest number of multi-country publications (65), and the highest total citations (8,207). Germany ranked third with 66 corresponding-author articles (6.3%) and 1,838 citations (Figure 3A).

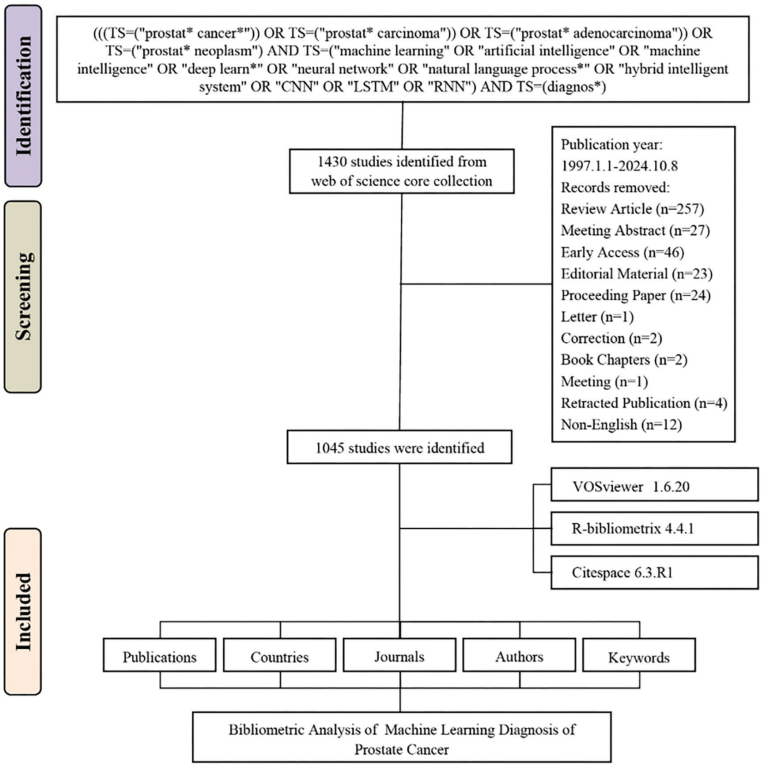


Figure 1. Flow diagram of the bibliographic retrieval process. TS, topic search; CNN, convolutional neural network; LSTM, long short-term memory; RNN, recurrent neural network.

Among the remaining top contributors, the Netherlands had the highest citations per publication (53.8), indicating a comparatively higher citation impact per article. Among the 76 countries that contributed at least one article through international collaboration, the USA led with 357 partnerships, primarily with Germany, Canada, China, and Japan. Germany ranked second with 183 collaborations, mainly involving the United States, the United Kingdom, Austria, the Netherlands, and Switzerland. China ranked third with 163 collaborations, primarily with the United States and Australia (Figure 3B).

Distribution and collaborative networks of institutions

A total of 4,762 institutions contributed to publications on ML in PCa diagnosis. Among the top 10 institutions, the Berlin Institute of Health in Germany led with 120 articles, followed by Stanford University (79 articles) and the University of California system (76 articles), both in the United States (Figure 4A).

In total, 151 institutions participated in international collaborations, each with at least five articles, forming distinct research clusters. Soochow University in China had the highest number of collaborations (54), forming a central dark-blue cluster. Shanghai Jiao Tong University in China was next with 51 collaborations, forming a yellow cluster. The University of Toronto in Canada ranked third with 45 collaborations, represented by a light-blue cluster (Figure 4B).

Contributions and collaborative networks of journals and authors

The articles were published across 352 journals, with the leading outlets and indicators summarized in Supplementary Table 2. By publication volume (TP), Cancers ranked first (TP: 53), followed by Frontiers in Oncology (TP: 35) and Scientific Reports (TP: 30).

In terms of influence metrics, European Radiology had the highest H-index[15] within this dataset, followed by Cancers (H-index: 13) and Medical Image Analysis and Medical Physics (H-index: 12 each). Regarding Journal Citation Reports metrics, Medical Image Analysis and Cancer Research had the highest 2023 impact factors among high-impact journals, and most leading journals were ranked in Q1.

The journal co-occurrence network, which measures joint citations in articles, identified three journals with the highest total link strength: European Radiology (185), Cancers (179), and Lancet Oncology (151) (Figure 5A). A journal coupling network, based on shared references, showed that Cancers (14,447), Frontiers in Oncology (10,387), and European Radiology (8,302) had the strongest link strengths (Figure 5B).

A total of 6,704 authors contributed to the dataset. Stephan C exhibited the highest H-index,[12] G-index,[15] M-index (0.522), and the most total publications,[15] followed by Jung K and Cammann H (Supplementary Table 3). Among 138 authors involved in international collaborations (with at least four articles), Rusu Mirabela led with 75 collaborations, followed by Fan Richard E (72) and Shao Wei (68) (Figure 6).

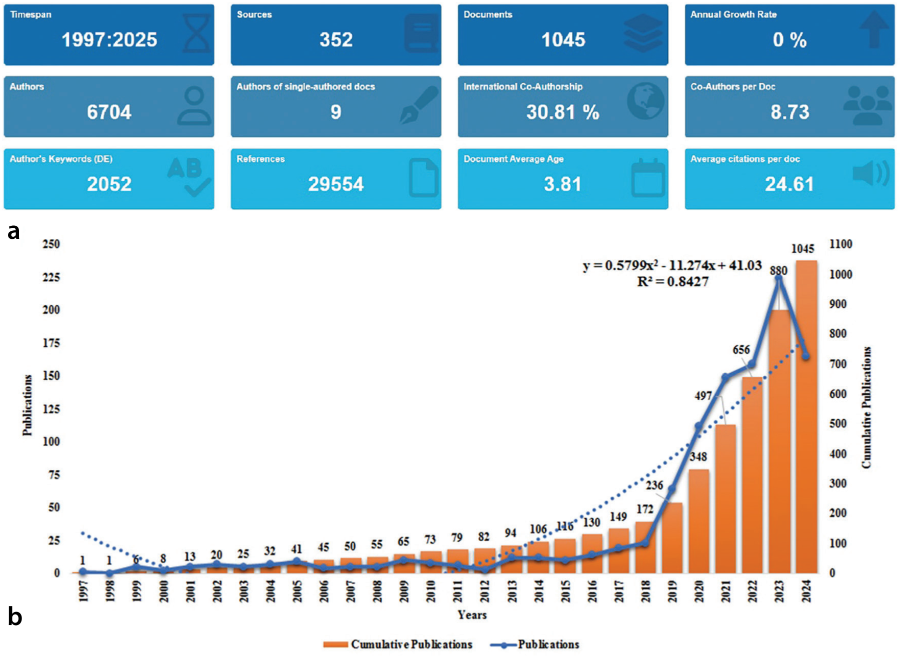
Keyword co-occurrence network and burst keywords

A total of 133 keywords with at least eight occurrences were identified for co-occurrence network analysis (Figure 7A). In the visualization, purple indicates earlier occurrences and yellow shows more recent ones. Around 2016, frequently occurring keywords included "antigen," "serum," and "identification" (purple). By 2018 (dark blue), terms such as "men," "radical prostatectomy," and "prediction" had emerged. Between 2018 and 2020 (dark green), "diagnosis" had the highest co-occurrence frequency (898), with "biopsy" and "classification" also prominent. Around 2020, newer terms such as "risk," "accuracy," and "features" appeared. Since 2022 (yellow), the most frequent keywords have been "system," "MRI," and "guidelines."

Keyword burst analysis identified terms with sharp increases in frequency, reflecting trends in the field (Figure 7B). The timeline uses gray to indicate periods without major activity, blue for ongoing research, and red for periods of intense activity. These bursts can be divided into two stages.

The first stage (1999–2005) showed prolonged bursts featuring keywords such as "antigen," "artificial neural network," "serum," "digital rectal examination," and "mass spectrometry." The second stage (2013–present) was characterized by shorter bursts and rapid keyword turnover. Between 2013 and 2018, terms such as "classification," "computer-aided diagnosis," and "ISUP Consensus Conference" were prominent. By 2019, new keywords, including "features," "convolutional neural network," "benign," and "data system," had gained prominence.

Since 2022, "transfer learning" has emerged as a key term and remains highly relevant. Many of these keywords persist in the blue phase, indicating ongoing research activity.



Timespan 1997:2025
Sources 352
Documents 1045
Annual Growth Rate 0 %
Authors 6704
Authors of single-authored docs 9
International Co-Authorship 30.81 %
Co-Authors per Doc 8.73
Author's Keywords (DE) 2052
References 29554
Document Average Age 3.81
Average citations per doc 24.61

Figure 2. Analysis of general information. (**a**) Summary of the quantitative analysis of publications. (**b**) Annual research output from 1997 to 2024.

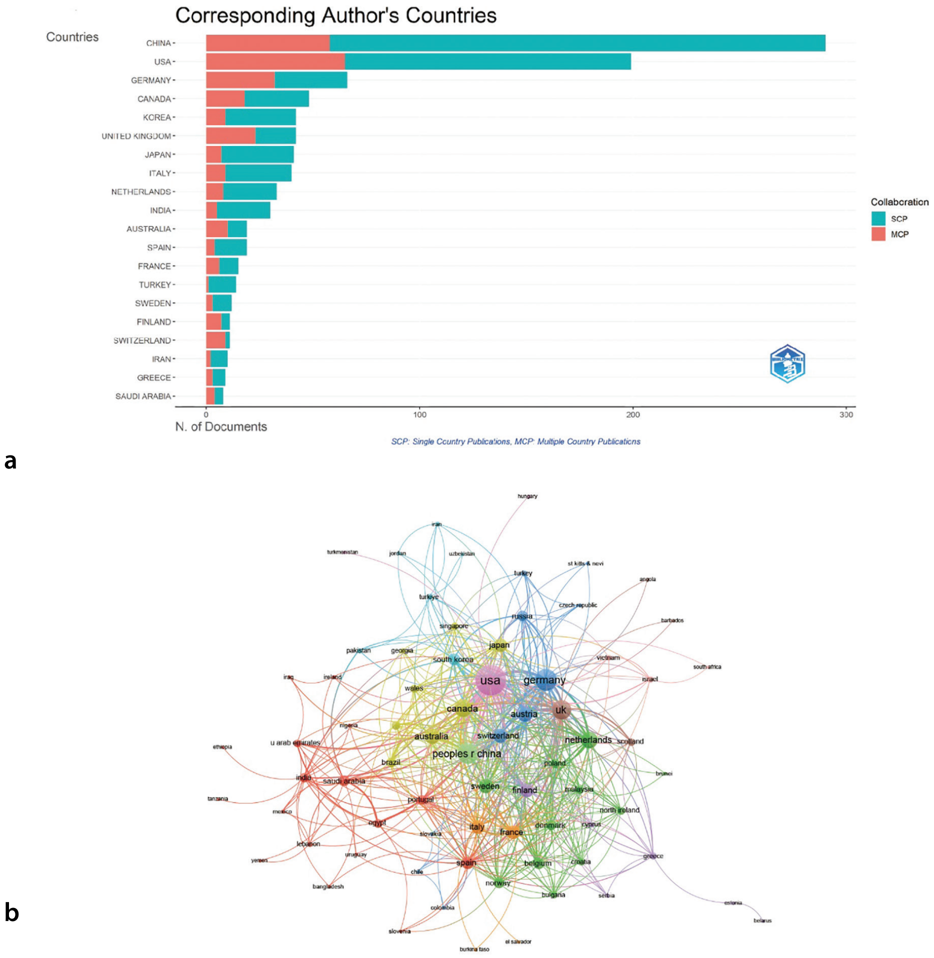


Figure 3. Analysis of countries. (**a**) Distribution of corresponding-author publications by country. (**b**) Visualization map depicting collaboration among countries.

Discussion

This bibliometric study summarizes ML research on PCa diagnosis from 1997 to late 2023. Publications increased rapidly from 2018 and peaked in 2023, suggesting growing interest and faster progress in ML methods and data use. China, the United States, and Germany were the main contributors. The United States and Germany show strong long-term foundations in biomedical and imaging research, and China's high output reflects continued investment in AI and health data science. The collaboration results also show that this field is increasingly driven by international partnerships and shared methods.

High-impact journals such as European Radiology, Cancers, and Medical Image Analysis show that research in this field primarily focuses on imaging, biomarkers, and AI-based diagnostic methods. The most cited study, "clinical-grade computational pathology using weakly supervised deep learning on whole slide images," highlights how deep learning can support digital pathology using slide-level labels without manual annotation.[16] Other influential works, such as the 2002 Cancer Research article on serum protein fingerprinting[17] and the 2019 review in Nature Reviews Clinical Oncology on AI for digital pathology,[18] further emphasize the central role of biomarker and imaging-based ML research in this domain.

In addition, our author-level productivity results indicate a highly skewed contribution pattern consistent with Lotka-type behavior, in which a relatively small group of authors accounts for a disproportionately large share of publications.[19] Within this concentrated expertise structure, key opinion leaders (e.g., Madabhushi and Yang) emerge as high-output, high-impact contributors, suggesting that mentorship and collaboration networks centered around leading groups may play an important role in shaping research directions and accelerating methodological diffusion across institutions.

Research hotspots and technological frontiers

Our keyword and burst analyses reveal a dynamic and evolving research landscape in ML-based PCa diagnosis, with thematic changes that track both clinical priorities and methodological breakthroughs. In the early years (1999–2005), research concentrated on traditional biomarker-oriented terms such as "antigen," "serum," and "identification",[20] reflecting the serum-based diagnostic strategies prevalent at the time. As the field matured, keywords such as "radical prostatectomy" and "prediction" emerged,[21] suggesting a shift toward clinically grounded modeling, including prognosis- and decision-relevant prediction tasks.

Between 2018 and 2020, the surge of "biopsy" and "classification" highlights an increased interest in ML methods that support core diagnostic decision points, particularly tissue confirmation and risk stratification.[22] During this period, frequent terms such as "accuracy" and "features" also indicate sus-

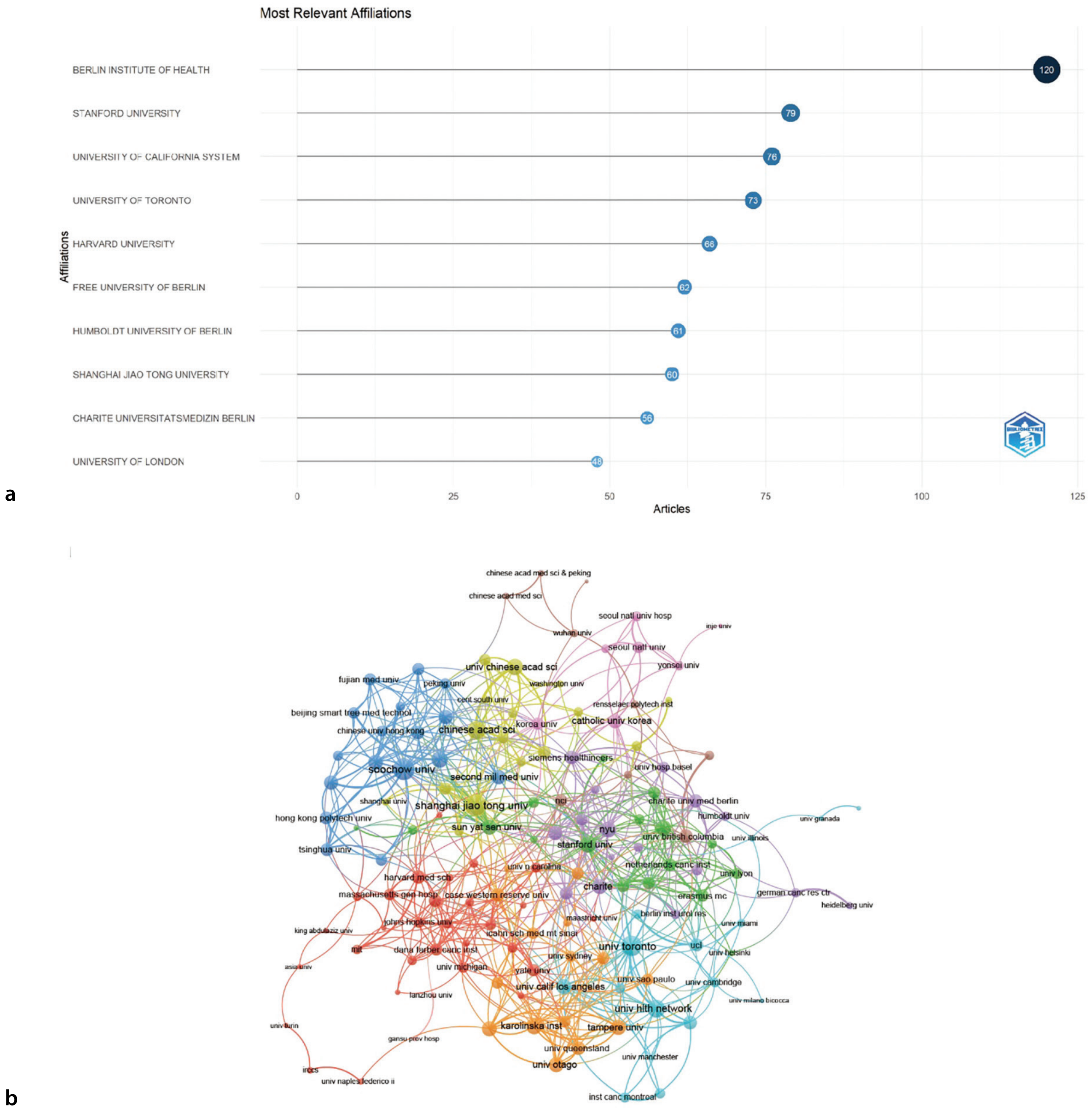


Figure 4. Analysis of institutions. (**a**) Top 10 institutions ranked by article count. (**b**) Visualization map depicting collaboration among institutions.

tained efforts to improve feature representation, model design, and performance evaluation practices.[23]

Why "MRI" has become dominant?

Since 2022, the increasing prominence of "MRI" aligns with the central role of mpMRI in contemporary PCa diagnostic pathways and with the suitability of MRI data for modern deep learning pipelines.[24] Multiparametric MRI provides rich spatial and functional information that supports tasks such as gland and lesion segmentation, lesion detection, and classification—areas in which convolutional neural network (CNN)-based methods have been particularly influential. The prominence of "MRI" in keyword analysis, therefore, likely reflects both its clinical importance in PCa workups and the scalability of imaging datasets for model development and benchmarking.

Contextualizing the "CNN" and "transfer learning" bursts with current technological developments

Burst analysis suggests a transition from earlier artificial neural networks and consensus-related terms (e.g., "ISUP Consensus Conference")[25] to a newer phase dominated by CNNs, data systems, and transfer learning.[26] This pattern is consistent with the broader deep-learning expansion following major CNN breakthroughs in computer vision, which subsequently catalyzed rapid uptake in medical imaging and computational pathology. In PCa diagnosis, CNN-based representation learning reduced reliance on manual feature engineering and enabled end-to-end learning on imaging and whole-slide data.[26]

The sustained "transfer learning" burst between 2022 and 2024[27] can be interpreted as

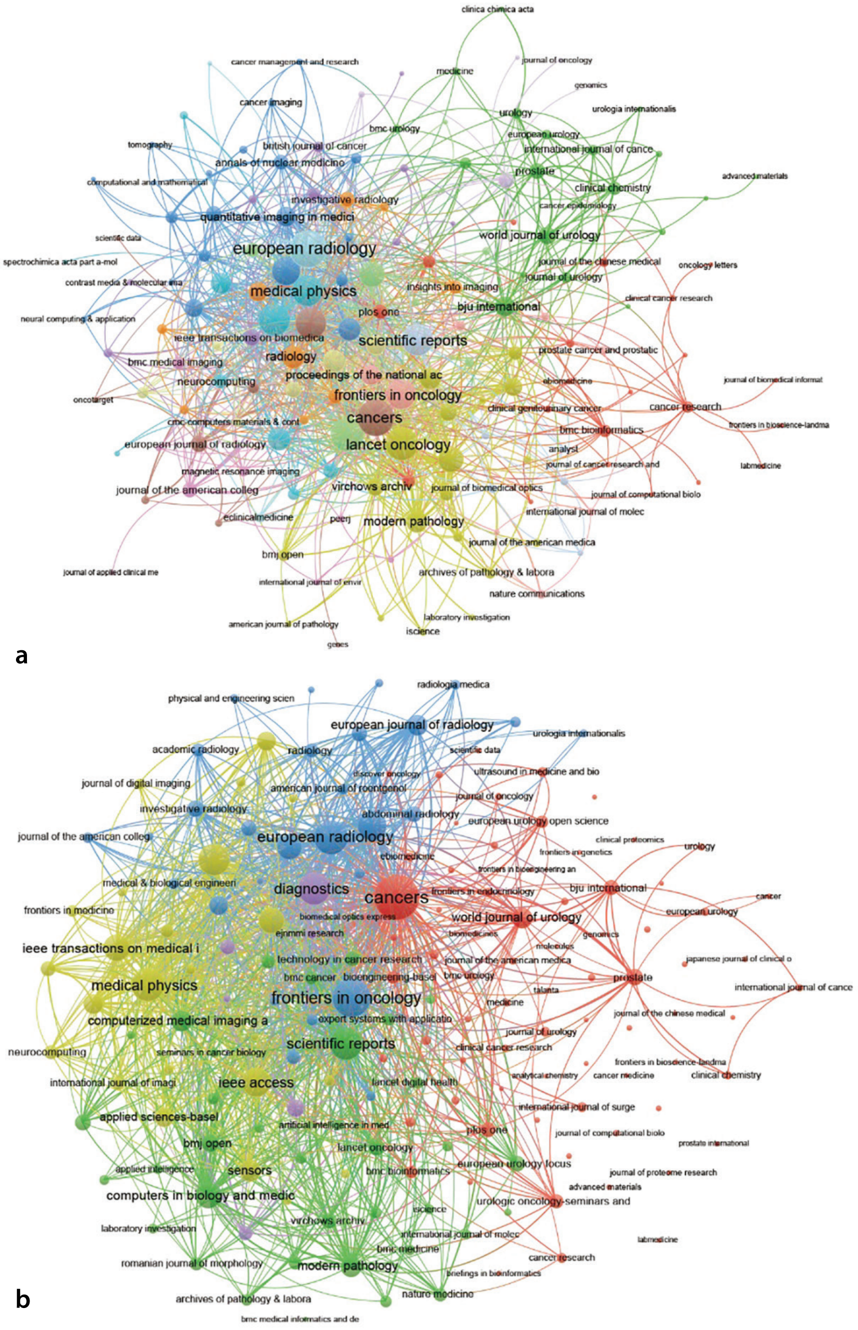


Figure 5. Analysis of journals. (**a**) Co-occurrence network of journals. (**b**) Coupling network of journals.

a response to persistent challenges in clinical AI, including limited annotated datasets, domain shifts across scanners and sites, and the need for robust generalization. Transfer learning enables the reuse of pretrained representations and adaptation to prostate imaging or pathology tasks with fewer labeled examples, supporting faster iteration and potentially improved external validity.[27] In the broader "foundation-model/GPT-era" context, the same underlying principle of leveraging large-scale pretraining to improve downstream performance has become increasingly central to modern AI development. Although the models and modalities differ across domains, this general trend helps explain why transfer learning remains a prominent and sustained hotspot in recent PCa diagnostic ML research.[27]

Although this analysis focuses on diagnosis, the observed keyword and thematic evolution can be interpreted as spanning three practical clusters that commonly co-develop in PCa ML research. First, diagnostic modeling and ML methods, including clinically relevant prediction problems tied to patient risk and outcomes (e.g., prediction-oriented keywords).[21] Second, imaging-driven ML (mpMRI-centric workflows), including detection, segmentation, and classification reflected by "MRI," "biopsy," and "classification".[22,24] Third, downstream clinical applications connected to treatment workflows, which tend to share enabling components such as segmentation, risk modeling, and data systems, and therefore appear adjacent to diagnosis-focused research streams in bibliometric structures.[26] This interpretation is consistent with the observed movement from algorithm keywords toward "system," suggesting increasing emphasis on end-to-end clinical utility rather than isolated model development.[24,26]

International collaboration and policy and infrastructure context

At the country level, China, the United States, and Germany dominate productivity and/or citation impact, and the collaboration network shows dense cross-national linkages. From an implementation perspective, these patterns likely reflect differences in long-term infrastructure (e.g., imaging and pathology digitization capacity, clinical research networks, and data engineering support) and the ability to sustain multidisciplinary teams combining urology, radiology, pathology, and AI. The emergence of highly connected collaboration hubs around established research systems is consistent with the need for large, well-curated datasets and rigorous validation to support clinically credible ML models.[28,29]

Importantly, collaboration results should not be interpreted solely as indicators of current research capacity. International co-authorship may also reflect historical funding structures, shared clinical trial networks, and established data ecosystems. In this context, the concentration of output in certain countries may reflect accumulated advantages in infrastructure and organization that facilitate multicenter studies and reproducibility—factors that are particularly important given ongoing challenges related to heterogeneity, interoperability, and clinical translation.[28,29]

Clinical implications

Bibliometric trends suggest that ML-based PCa diagnosis is moving from proof-of-concept algorithms toward systems designed to support real-world clinical decision-making. However, clinical impact depends on more than performance metrics. Models must be deployable within real workflows (radiology interpretation, biopsy decision-making, pathology review), interoperable with hospital information systems, and supported by

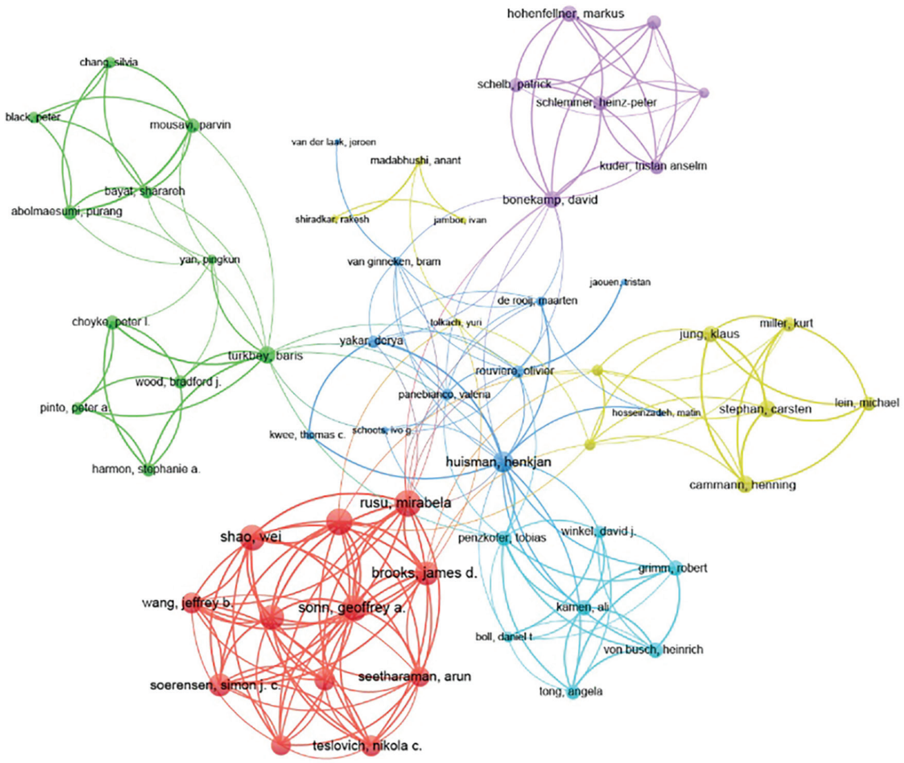


Figure 6. Visualization map depicting collaboration among authors.

user training and clear accountability structures.[28,30]

Although diagnostic accuracy and related metrics have improved over time,[29] further prospective and multicenter validation is needed to establish robustness across sites and populations. The increasing prominence of “system” in keyword trends reflects growing interest in integrated solutions but also highlights the complexity of implementation. Effective clinical integration requires attention to dataset shift, failure modes, and human–AI interaction, rather than focusing solely on retrospective performance.[28,29]

Regulatory, ethical, and clinical integration challenges

Despite rapid progress, regulatory and ethical considerations remain major barriers to clinical translation.[31] The adoption of ML in clinical workflows is constrained by requirements related to data privacy, patient safety, transparency, and governance. Ethical concerns, including algorithmic bias, explainability, reliability, and informed consent, must be proactively addressed to maintain trust and reduce disparities in performance across patient populations.[31]

The findings also suggest a gap between algorithm development and real-world implementation. The “black box” nature of deep learning and limited interpretability can hinder clinician acceptance and complicate error analysis and accountability.[28] As a result, explainable AI is receiving increasing attention. Interpretability methods can support clinical decision-making, facilitate auditing and bias assessment, and improve alignment with regulatory expectations for safety and transparency.

Technical performance and clinical validation

The literature indicates steady improvements in diagnostic accuracy, sensitivity, and specificity, particularly in imaging and pathology applications.[29] However, benchmarking and reporting remain inconsistent across studies, and large-scale prospective validation remains comparatively limited. This is a critical limitation given known challenges in data heterogeneity and domain shift.[28,29]

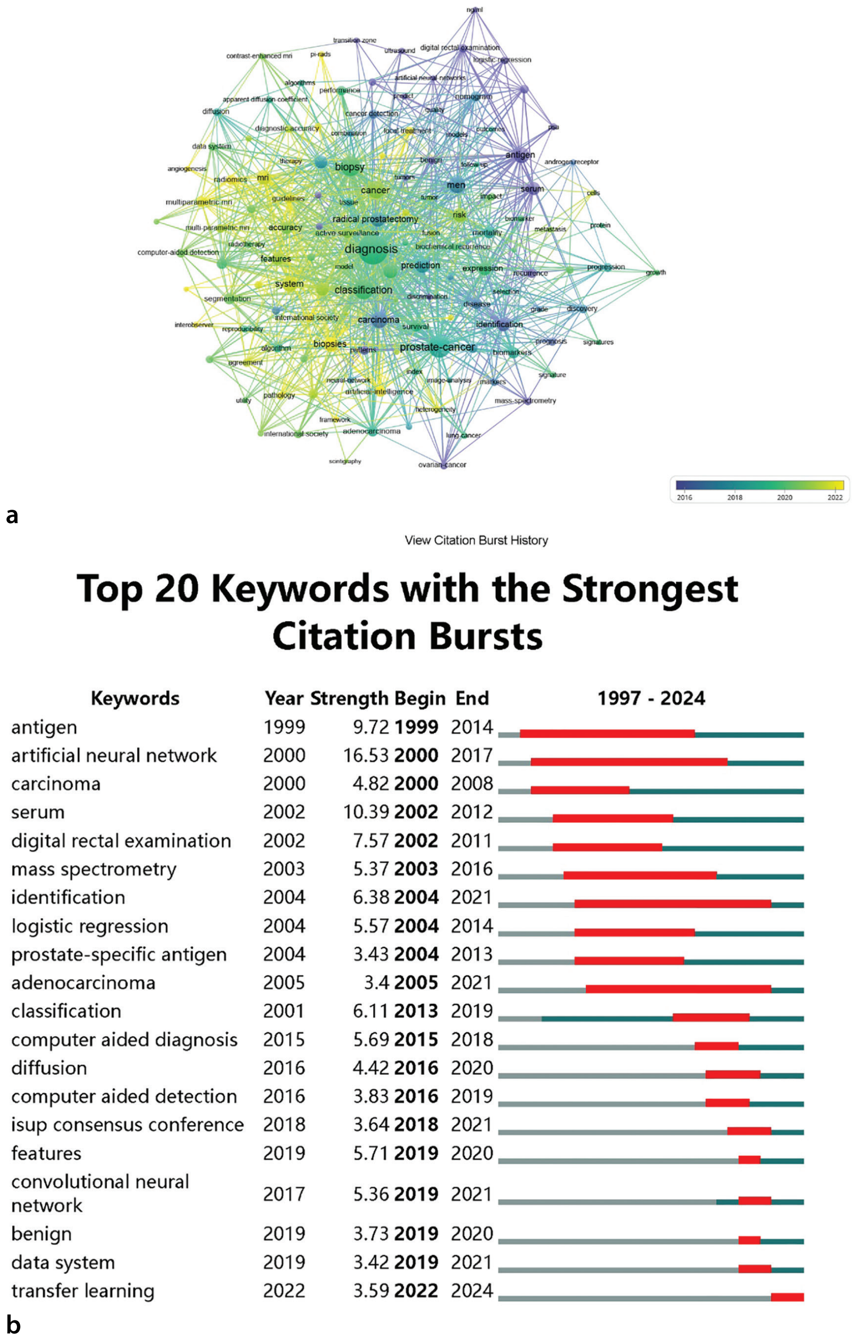
Multimodal ML that integrates imaging, pathology, genomics, and electronic health records offers potential for improved diagnostic precision but introduces additional challenges related to harmonization, missingness, and governance.[31] Approaches such as federated learning can support collaboration without centralizing sensitive patient data, and generative adversarial network–based augmentation may help address limited sample sizes. However, both approaches require careful validation to avoid amplifying bias or creating unrealistic signals.[32,33]

Future directions and recommendations

Based on the observed trends and ongoing barriers to clinical translation, future research should prioritize robust evaluation in clinically challenging settings, including higher-risk and heterogeneous patient subgroups, as limited generalizability across populations and institutions remains a key obstacle to real-world impact.[28,29]

The field should also progress from predominantly single-modality approaches toward multimodal integration where feasible, combining MRI with pathology and routinely collected clinical variables to improve diagnostic confidence and reduce failure modes associated with reliance on a single information source. Achieving this will require standardized data curation, harmonization, and quality-control pipelines.[31]

The increasing prominence of “system” alongside “MRI” in the keyword trends suggests that the next phase of impact will come from workflow-aware decision support tools designed for specific insertion points, such as MRI interpretation assistance and biopsy-related decision steps. These tools



Top 20 Keywords with the Strongest Citation Bursts

Keywords	Year	Strength	Begin	End	1997 - 2024
antigen	1999	9.72	**1999**	2014	
artificial neural network	2000	16.53	**2000**	2017	
carcinoma	2000	4.82	**2000**	2008	
serum	2002	10.39	**2002**	2012	
digital rectal examination	2002	7.57	**2002**	2011	
mass spectrometry	2003	5.37	**2003**	2016	
identification	2004	6.38	**2004**	2021	
logistic regression	2004	5.57	**2004**	2014	
prostate-specific antigen	2004	3.43	**2004**	2013	
adenocarcinoma	2005	3.4	**2005**	2021	
classification	2001	6.11	**2013**	2019	
computer aided diagnosis	2015	5.69	**2015**	2018	
diffusion	2016	4.42	**2016**	2020	
computer aided detection	2016	3.83	**2016**	2019	
isup consensus conference	2018	3.64	**2018**	2021	
features	2019	5.71	**2019**	2020	
convolutional neural network	2017	5.36	**2019**	2021	
benign	2019	3.73	**2019**	2020	
data system	2019	3.42	**2019**	2021	
transfer learning	2022	3.59	**2022**	2024	

Figure 7. Analysis of keywords. (**a**) Visual analysis of the keyword co-occurrence network. (**b**) Top 20 keywords with the strongest citation bursts.

should be evaluated for usability, reliability, and failure modes under realistic clinical conditions rather than only in retrospective experiments.[24,28,29]

In addition, interpretability and governance should be treated as core design requirements. Explainable AI methods and auditing processes should be incorporated early to support transparency, bias assessment, safety monitoring, and trust, thereby aligning development with clinical and regulatory expectations.[31,32]

Finally, more consistent validation and reporting standards are needed, including clear dataset descriptions, external validation, clinically meaningful performance metrics, and transparent reporting of limitations and negative findings. These practices will improve reproducibility and reduce overestimation of model performance.[29,32] Collectively, these directions support a transition from rapid research growth toward more reliable, generalizable, and clinically integrated ML systems for PCa diagnosis.[28,29,,31,32]

This study provides a detailed, data-driven overview of global research trends, technological developments, and emerging hotspots in ML-based PCa diagnosis. By incorporating co-occurrence and burst keyword analyses, it offers a nuanced understanding of the field's evolution and future directions.

However, several limitations should be noted. The use of a simplified regression model may not fully reflect the exponential growth in publications from 2018 onward. Restricting the analysis to English-language articles indexed in WoSCC may introduce language and database biases, potentially excluding relevant studies from other sources. In addition, the rapid pace of AI innovation may result in literature lag, whereby the latest advancements are underrepresented in bibliometric analyses.

Although alternative databases, such as Scopus and PubMed/MEDLINE, were considered, the analysis was limited to the WoSCC to ensure consistency in citation indexing and network analysis. This restriction should be considered a coverage-related limitation.

This bibliometric analysis highlights the evolution of ML applications in PCa diagnosis, showing a transition from early biomarker research to advanced deep learning techniques, particularly in medical imaging. Research output has increased substantially, with major contributions from the United States, China, and Germany. Recent trends emphasize CNNs and transfer learning.

Although diagnostic capabilities have improved, challenges remain in clinical translation, including data heterogeneity, ethical and regulatory considerations, and the need for rigorous clinical validation. Advancing multimodal integration, explainable AI, and standardized evaluation frameworks will be essential to fully realize the potential of ML for early detection and effective management of PCa.

Footnotes

Conflict of interest disclosure

The author declared no conflicts of interest.

Supplementary Tables 1-3: https://d2v96fxpocvxx.cloudfront.net/bda9171a-fae8-4995-8276-2138323f1e16/content-images/1f64efde-a00e-44e8-9cc2-a14a32825a34.pdf

References

1. Rebello RJ, Oing C, Knudsen KE, et al. Prostate cancer. *Nat Rev Dis Primers*. 2021;7(1):8. [Crossref]
2. Sekhoacha M, Riet K, Motloung P, Gumenku L, Adegoke A, Mashele S. Prostate cancer review: genetics, diagnosis, treatment options, and alternative approaches. *Molecules*. 2022;27(17):5730. [Crossref]
3. Schaeffer EM, Srinivas S, Adra N, et al. Prostate cancer, version 4.2023, NCCN Clinical Practice Guidelines in oncology. *J Natl Compr Canc Netw*. 2023;21(10):1067-1096. [Crossref]
4. Suarez-Ibarrola R, Hein S, Reis G, Gratzke C, Miernik A. Current and future applications of machine and deep learning in urology: a review of the literature on urolithiasis, renal cell carcinoma, and bladder and prostate cancer. *World J Urol*. 2020;38(10):2329-2347. [Crossref]
5. Tong Y, Tan Z, Wang P, Gao X. A machine learning method for predicting biomarkers associated with prostate cancer. *Front Biosci (Landmark Ed)*. 2023;28(12):333. [Crossref]
6. Castaldo R, Cavaliere C, Soricelli A, Salvatore M, Pecchia L, Franzese M. Radiomic and genomic machine learning method performance for prostate cancer diagnosis: systematic literature review. *J Med Internet Res*. 2021;23(4):e22394. [Crossref]
7. Choudhri AF, Siddiqui A, Khan NR, Cohen HL. Understanding bibliometric parameters and analysis. *Radiographics*. 2015;35(3):736-746. [Crossref]
8. Wang Y, Zhang J, Zhang Y, Yao J. Bibliometric analysis of global research profile on ketogenic diet therapies in neurological diseases: beneficial diet therapies deserve more attention. *Front Endocrinol (Lausanne)*. 2023;13:1066785. [Crossref]
9. Wei Y, Mei Z, Xie C, Yuan F, Xu D. Mapping the application landscape of artificial intelligence in prostate cancer: a global bibliometric analysis. *Int J Surg*. 2026;112(2):2988-3002. [Crossref]
10. Musa IH, Afolabi LO, Zamit I, et al. Artificial intelligence and machine learning in cancer research: a systematic and thematic analysis of the top 100 cited articles indexed in Scopus database. *Cancer Control*. 2022;29:10732748221095946. [Crossref]
11. Zeng N, Sun JX, Liu CQ, et al. Knowledge mapping of application of image-guided surgery in prostate cancer: a bibliometric analysis (2013-2023). *Int J Surg*. 2024;110(5):2992-3007. [Crossref]
12. van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*. 2010;84(2):523-538. [Crossref]
13. Zhang C, Zhang M, Cao X, et al. Navigating the landscape of MANF research: a scientometric journey with CiteSpace analysis. *Cell Mol Neurobiol*. 2023;43(8):3897-3913. [Crossref]
14. Arruda H, Silva ER, Lessa M, Proença D Jr, Bartholo R. VOSviewer and Bibliometrix. *J Med Libr Assoc*. 2022;110(3):392-395. [Crossref]
15. Garg K, Ranjan M, Krishna V, Singh M, Rezai A. A scientometric analysis of the 100 most cited articles on magnetic resonance guided focused ultrasound. *Front Hum Neurosci*. 2022;16:981571. [Crossref]
16. Campanella G, Hanna MG, Geneslaw L, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med*. 2019;25(8):1301-1309. [Crossref]
17. Adam BL, Qu Y, Davis JW, et al. Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. *Cancer Res*. 2002;62(13):3609-3614. [Crossref]
18. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol*. 2019;16(11):703-715. [Crossref]
19. Qiu J, Zhao R, Yang S, Dong K. Informetrics: theory, methods and applications. Singapore: Springer; 2017. [Crossref]
20. Wei C, Tian L, Jia B, et al. Association between serum triglycerides and prostate specific antigen (PSA) among U.S. males: National Health and Nutrition Examination Survey (NHANES), 2003-2010. *Nutrients*. 2022;14(7):1325. [Crossref]
21. Jović S, Miljković M, Ivanović M, Šaranović M, Arsić M. Prostate cancer probability prediction by machine learning technique. *Cancer Invest*. 2017;35(10):647-651. [Crossref]
22. Omer A, Lamb AD. Optimizing prostate biopsy techniques. *Curr Opin Urol*. 2019;29(6):578-586. [Crossref]
23. Rammal A, Assaf R, Goupil A, Kacim M, Vrabie V. Machine learning techniques on homological persistence features for prostate cancer diagnosis. *BMC Bioinformatics*. 2022;23(1):476. [Crossref]
24. van Leenders GJLH, van der Kwast TH, Grignon DJ, et al. The 2019 International Society of Urological Pathology (ISUP) Consensus Conference on grading of prostatic carcinoma. *Am J Surg Pathol*. 2020;44(8):e87-e99. [Crossref]
25. Hamm CA, Baumgärtner GL, Biessmann F, et al. Interactive explainable deep learning model informs prostate cancer diagnosis at MRI. *Radiology*. 2023;307(4):e222276. [Crossref]
26. Islam R, Imran A, Rabbi MF. Prostate cancer detection from MRI using efficient feature extraction with transfer learning. prostate cancer. 2024;2024:1588891. [Crossref]
27. Poddar M, Marwaha JS, Yuan W, Romero-Brufau S, Brat GA. An operational guide to translational clinical machine learning in academic medical centers. *NPJ Digit Med*. 2024;7(1):129. [Crossref]
28. Khalifa M, Albadawy M. AI in diagnostic imaging: revolutionising accuracy and efficiency. *Comput Methods Programs Biomed Update*. 2024;5:100146. [Crossref]
29. Ye Y, Liu Z, Zhu J, et al. Development trends and knowledge framework in the application of magnetic resonance imaging in prostate cancer: a bibliometric analysis from 1984 to 2022. *Quant Imaging Med Surg*. 2023;13(10):6761-6777. [Crossref]
30. Tan X, Su AT, Hajiabadi H, Tran M, Nguyen Q. Applying machine learning for integration of multi-modal genomics data and imaging data to quantify heterogeneity in tumour tissues. *Methods Mol Biol*. 2021;2190:209-228. [Crossref]
31. Gangwal A, Ansari A, Ahmad I, Azad AK, Wan Sulaiman WMA. Current strategies to address data scarcity in artificial intelligence-based drug discovery: a comprehensive review. *Comput Biol Med*. 2024;179:108734. [Crossref]
32. Mennella C, Maniscalco U, De Pietro G, Esposito M. Ethical and regulatory challenges of AI technologies in healthcare: a narrative review. *Heliyon*. 2024;10(4):e26297. [Crossref]