Original Article

DOI: 10.4274/uob.galenos.2025.2025.5.2
Bull Urooncol 2025;24(3):57-69

Hotspots and Research Trends in Machine Learning for Prostate Cancer: A Bibliometric Analysis and Visualization (1997-2025)

Tunahan Ateş[1], Nezih Tamkaç[2], İbrahim Halil Şükür[3], Fesih Ok[3], İsmail Önder Yılmaz[4], Mutlu Değer[4], Volkan İzol[4]

[1]*Defne State Hospital, Clinic of Urology, Hatay, Türkiye*
[2]*Hatay Mustafa Kemal University Faculty of Medicine, Department of Urology, Hatay, Türkiye*
[3]*University of Health Sciences Türkiye, Adana Training and Research Hospital, Department of Urology, Adana, Türkiye*
[4]*Çukurova University Faculty of Medicine, Department of Urology, Adana, Türkiye*

Abstract

Objective: This bibliometric analysis examines the evolution of prostate cancer (PCa) research and evaluates the impact of machine learning and artificial intelligence (AI) on its diagnosis, classification, and treatment.
Materials and Methods: Articles published between 1997 and 2025 were analysed using the Web of Science Core Collection database. VOSviewer and Bibliometrix software was utilized for bibliometric analysis. Terms such as "PCa", "machine learning (ML)", "deep learning" and "AI" were included in the search strategy. The number of publications, the most cited studies, author collaborations and country collaborations, thematic trends, and citation networks were visualised.
Results: A total of 3,277 articles were analysed. The in augural article was published in 1997. Over the past five years, there has been a significant increase in the number of articles published. The United States and China are the countries with the highest number of publications, and the most influential authors and institutions are concentrated in these countries. A marked upward trend has been observed in ML applications for PCa diagnosis, risk stratification, and treatment planning.
Conclusion: The use of AI and ML in PCa research has grown significantly over the last 20 years. However, most of the existing models have been tested with retrospective data, and more multicenter and prospective studies are needed for clinical applications. Comprehensive clinical validation is essential before AI-based systems can be reliably implemented.
Keywords: Prostate cancer, machine learning, artificial intelligence, bibliometric analysis, scientific trends

Introduction

Prostate cancer (PCa) ranks as the second most prevalent cancer among men globally and constitutes a substantial proportion of cancer-related mortality (1). This disease is particularly common in older men and may progress aggressively, with a high risk of metastasis if not detected early (1). Currently, the standard diagnostic methods for PCa include the prostate-specific antigen (PSA) test, multiparametric magnetic resonance imaging (mpMRI), and biopsy (2). Nevertheless, conventional diagnostic methods are not consistently definitive, and instances of false negatives or false positive results may occur (3). In this context, machine learning (ML) techniques offer innovative and promising approaches for the diagnosis and treatment of PCa, encompassing areas such as medical imaging analysis and biomarker discovery (3).

ML is a subset of artificial intelligence (AI) that enhances clinical decision-making support through the analysis of large-scale datasets. In recent years, various ML methodologies, including

Cite this article as: Ateş T, Tamkaç N, Şükür İH, et al. Hotspots and research trends in machine learning for prostate cancer: a bibliometric analysis and visualization (1997-2025). Bull Urooncol. 2025;24(3):57-69.

Address for Correspondence: Tunahan Ateş, MD, Defne State Hospital, Clinic of Urology, Hatay, Türkiye
E-mail: drtunahanates0101@gmail.com **ORCID:** orcid.org/0000-0001-9087-290X
Received: 15.05.2025 **Accepted:** 03.06.2025 **Publication Date:** 24.09.2025

supervised learning, unsupervised learning, and reinforcement learning, have been increasingly applied in the diagnosis and management of PCa treatment processes (4). Deep learning (DL) and convolutional neural networks (CNN), which are a class of neural networks specialized in extracting spatial features from image data, have demonstrated considerable success, particularly in the analysis of mpMRI (5). The utilization of these techniques in cancer diagnostics (6-8), which predict the aggressiveness of the disease (9,10) and facilitate risk classification (11,12) for the development of treatment plans, is becoming increasingly prevalent.

Bibliometric analysis serves as a quantitative method for evaluating publications in a specific research area, enabling the identification of prominent authors, institutions, countries, and emerging research trends (13). This analysis facilitates the identification of prominent topics within the literature, the journals that publish the most articles, and the studies that are cited most frequently (14). In recent years, software applications such as VOSviewer and Bibliometrix have gained significant traction for visualizing scientific networks and conducting bibliometric analyses. These tools enable researchers to examine the structures of collaborative networks within the scientific literature, trace the evolution of research themes, and forecast future trends (15).

The study will examine literature published between 1997 and 2025 by utilizing the Web of Science (WoS) Core Collection (CC) database and employing bibliometric analysis methods for data visualization. The objective of this research is to construct a comprehensive scientific map that delineates the relationship between PCa and ML. The findings of this study are anticipated to provide valuable insights for future research endeavors and to contribute to the advancement of AI-supported diagnostic and treatment systems in clinical applications. Furthermore, this research may foster increased international collaboration by analyzing scientific collaboration networks in the domain of ML-based PCa research.

Materials and Methods

Search Strategy

In this study, WoS CC was utilized as the primary data source. The WoS comprises numerous articles across various disciplines and is widely acknowledged by researchers as a high-quality database (16). The database is acknowledged as the most appropriate for conducting bibliometric analysis (15). We used the following search strategy: topic search (TS) = ("prostate cancer" or "prostate neoplasm" or "Gleason score" or "prostate carcinoma" or "PSA" or "prostate-specific antigen" or "multiparametric MRI" or "mpMRI") and TS = ("machine learning" or "supervised learning" or "unsupervised learning" or "reinforcement learning" or "reinforced learning" or "deep learning" or "transfer learning"). The timeframe for this search encompassed articles published up to the current year, with a submission deadline for queries set for 21 January 2025. The literature selected for this study was restricted to articles, review articles, and to those published in the English language. This search yielded a total of 3,277 articles. Full records and cited references were exported as plain text files for subsequent visualization and analysis. The search process is illustrated in Table 1.

Eligibility Criteria

This study employed specific criteria for the inclusion and exclusion of literature. The inclusion criteria encompassed original research articles and review articles published in relevant English language journals. Conversely, the following materials were excluded from the analysis: conference proceedings, meeting abstracts, early access publications, book chapters, editorial content, corrections, letters, retracted publications, books, and meeting reports. Additionally, duplicate articles were eliminated from consideration. The literature search was conducted independently by two reviewers to ensure comprehensive identification of all pertinent studies. In instances of discrepancies, the matter was referred to a third researcher for resolution. The process of the literature search is illustrated in Figure 1.

Data Analysis and Visualisation

Bibliometric analysis emerged in the twentieth century and was formally recognized as an independent discipline in 1969 (17). This study applies quantitative methods to analyze the existing literature in this field. This study involved the extraction of authors, keywords, journals, countries, references during the analytical process. Additionally, bibliometric analysis frequently employs the co-citation technique, which occurs when two articles are cited concurrently by one or more other articles. Co-citation analysis has been demonstrated to enhance data interpretation, thereby rendering the results more comprehensive.

Table 1. Criteria in the search process

Category	**Specific standard requirements**
Research database	Web of Science Core Collection
Citation indexes	SCIE, ESCI, SSCI, AHCI
Searching period	Database build to Jan 21, 2025
Language	English
Retrieval formula	**TS = ("prostate cancer" or "prostate neoplasm" or "gleason score" or "prostate carcinoma" or "PSA" or "prostate-specific antigen" or "multiparametric MRI" or "mpMRI") and TS = ("machine learning" or "supervised learning" or "unsupervised learning" or "reinforcement learning" or "reinforced learning" or "deep learning" or "transfer learning")**
Document types	"Articles", "Review articles"
Data extraction	The export should include comprehensive records along with cited references in plain text format, BibTeX, and tab-delimited file formats.
Final documentation	3,277

SCIE: Science Citation Index Expanded, ESCI: Emerging Sources Citation Index, SSCI: Social Sciences Citation Index, AHCI: Arts and Humanities Citation Index, TS: Topic search, mpMRI: Multiparametric magnetic resonance imaging

The following software applications were utilized for statistical and visualization analyses: VOSviewer version 1.6.20 (14) and R version 4.4.2, which includes the Bibliometrix R package and the Biblioshiny tool, accessible at https://www.bibliometrix.org/home/. These tools are widely employed in the medical research domain. The VOSviewer software, developed by Leiden University in the Netherlands, employs a chain-based data standardization method that offers diverse visualization perspectives on keywords, collaborating institutions, co-authors, and other relevant entities. These visualizations encompass mesh, overlap, and density views, characterized by notable features such as straightforward mapping and a visually informative structure (14). Different clusters within a network diagram are denoted by distinct colors and are indicative of collaborations, co-working relationships, and connectors. The size of each circle corresponds to the number of references, publications, and keywords associated with that cluster. The Bibliometrix R package, which is available in the R environment, was utilized alongside the Biblioshiny tool for conducting bibliometric analysis and visualization. This approach assists researchers in comprehending the prevailing research trends, focal areas of inquiry, and academic impact within a specific field (13). During the data processing phase, challenges such as name disambiguation (e.g., authors with similar names), keyword unification, and institutional name variations were encountered. Additionally, some inconsistencies in metadata (e.g., missing affiliations or citation counts) required manual checking. Despite these challenges, the use of VOSviewer and Bibliometrix allowed for effective network visualizations and thematic clustering.

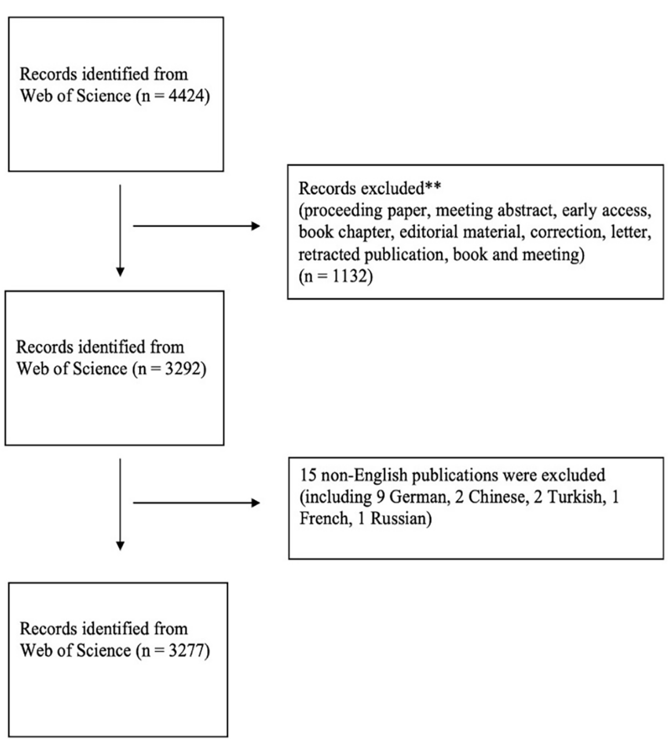


Figure 1. Flowchart of literature review

Statistical Analysis

Linear regression analysis of the number of publications by year was performed with IBM SPSS Statistics version 26.0.

Statements of Ethics

Not applicable, as this bibliometric analysis did not involve direct interaction with human participants or collection of personal data.

Results

General Informations

The initial search resulted in the identification of 4,424 articles. Following the application of selection criteria that included only original research articles, review articles, and articles published in the English language, a total of 3,277 articles were included in the final analysis. The flowchart illustrating the study's methodology is presented in Figure 1. Among the 3,277 articles, 2,911 are classified as original research articles, while 366 are categorized as review articles. Key findings from the analysis are depicted in Table 2 and Figure 2.

The inaugural article was published in 1997. Over the past five years, there has been a significant increase in the number of articles published. The increase in 2020 and beyond is reaching a significant level. The year 2024 recorded the highest output, with a total of 711 articles. Figure 2 shows that the number of publications on this topic has gradually increased over time ($R2=0.474$, $p<0.001$). Additionally, the annual average number of citations reached its zenith in 2019, during which each published article received an average of 10.74 citations. The trends in published articles and the average number of citations by year are illustrated in Figure 3.

Table 2. Main information

Main information about data	**Results**
Timespan	1997-2025
Sources (journals, books, etc)	927
Documents	3277
Annual growth rate %	14.38
Document average age	3.84
Average citations per doc	24.31
References	107675
Keywords plus (ID)	4909
Author's keywords (DE)	6401
Authors	18812
Authors of single-authored docs	38
Single-authored docs	40
Co-authors per doc	9.39
International co-authorships %	32,1
Article	2911
Review	366

Figure 2. Main information

Co-authorship Analysis

In the author-coauthorship analysis, Anant Madabhushi occupies the most central position (Supplementary Figure 1). He stands out as the author with the highest number of publications and citations. The authors with the highest number of publications are presented in Supplementary Table 1. Lotka's law posits that a small number of authors produce a large volume of articles, whereas a larger number of authors contribute only a few articles, with productivity following an inverse square law. The analysis conducted aligns closely with Lotka's law, achieving a near-perfect fit (Supplementary Figure 2).

In the analysis of co-authorship by country, the United States and China occupy central positions (Supplementary Figure 3). Notably, China's connections are more current. The United States leads in both the number of articles and citations, with a total of 1,216 articles and 41,635 citations, followed closely by China. The countries with the highest number of publications are detailed in Supplementary Table 2 and Supplementary Figure 4 illustrates the distribution of articles among the countries. Most countries in Africa have not published any articles.

In the analysis of co-authorship among the organizations, Case Western Reserve University and Emory University play a central role (Supplementary Figure 5). Case Western Reserve University distinguishes itself as the institution with the highest number of articles and citations, with 101 articles and 6,986 citations. The organizations with the most publications are presented in Supplementary Table 3. Notably, seven of the ten organizations with the highest publication counts are located in the United States, while only one of the top 10 organizations is based in China.

Source, Document and Keyword Analysis

The citation-source analysis shows that medical physics, and cancers journals are central (Supplementary Figure 6). Medical physics leads in both publications and citations, significantly outpacing other sources. Article and citation counts are detailed in Supplementary Table 4.

In author keyword co-occurrence analysis, the most central author keywords are ML, PCa, and DL (Supplementary Figure 7). The top three author keywords are ML, DL, and PCa. The most commonly used author keywords are presented in Supplementary Table 5. The presence of the keywords "magnetic resonance imaging" and "MRI" is among the top ten keywords,

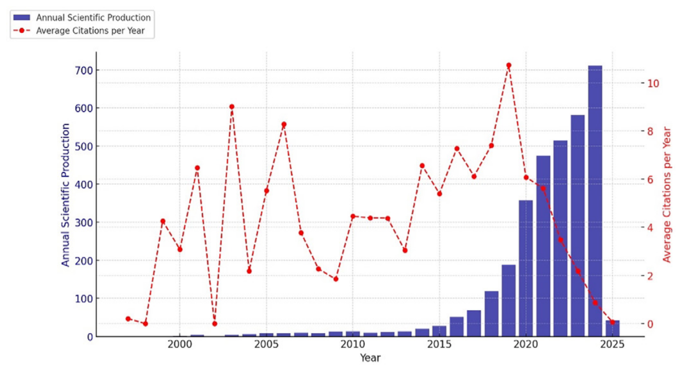


Figure 3. Annual scientific production and average citations per year

which indicates that ML is predominantly utilized in imaging techniques for PCa.

The most frequently cited articles are the studies titled "Clinical-Grade Computational Pathology Using Weakly Supervised Deep Learning on Whole Slide Images" and "Artificial Intelligence in Cancer Imaging: Clinical Challenges and Applications," both published in 2019. The prominence of these studies in the literature indicates that a significant portion of research in this field is dedicated to enhancing the accuracy and reliability of diagnoses. Furthermore, it underscores the central role of diagnostic AI applications in the management of PCa. The most cited articles are presented in Supplementary Table 6.

Thematic Map

The thematic map analysis using Bibliometrix identified three main themes in PCa research (Supplementary Figure 8). The first theme focuses on the diagnosis of PCa and ML methods, highlighting topics such as clinically relevant diagnosis, biochemical recurrence prediction, and genomic analyses. This theme includes ML-based approaches for processing clinical data, optimizing diagnostic processes, and discovering biomarkers. The second theme examines the integration of mpMRI with AI systems, covering AI algorithms for mpMRI data analysis, image segmentation, and enhancement of diagnostic accuracy. AI-based systems are increasingly used as clinical decision support tools in radiological evaluations. The third theme, the narrowest, centers on AI-supported applications for PCa treatment, particularly in radiotherapy. This involves AI models for radiotherapy planning, dose calculation optimization, and predicting treatment outcomes. The potential of AI-enabled systems to offer more precise and personalized treatment approaches is driving research in this area. Together, these themes illustrate the expanding role of AI and ML across the diagnostic, imaging, and therapeutic dimensions of PCa management.

Thematic Evaluation

The Sankey diagram created using Bibliometrix is based on abstracts and illustrates the evolution of research themes related to AI applications in PCa from 1997 through 2025, divided into three distinct time periods (Supplementary Figure 9). In the initial period (1997-2010), classical ML methods, including artificial neural networks and support vector machines, along with the use of clinical diagnostic parameters such as MRI and the Gleason score, were predominant in PCa studies. However, during the subsequent period (2011-2020), DL techniques, particularly CNN, gained prominence in fields such as MRI analysis. In the final period (2021-2025), PCa research has shifted towards multifaceted, high-tech, and multidisciplinary themes, such as treatment planning systems, radiomics analysis, and the cancer genome atlas, with an increased emphasis on integration into clinical applications.

Discussion

This bibliometric analysis underscores the growing importance of ML in the diagnosis and treatment of PCa, as reflected by a notable surge in related scientific publications over the past decade. The upward trajectory in research output not only signals increasing academic interest but also highlights the transformative potential of ML technologies within the field of urologic oncology.

ML has become increasingly integrated into several critical aspects of PCa management. It plays a pivotal role in early disease detection, enabling more accurate identification of clinically significant cancer cases. Moreover, ML contributes to risk stratification by distinguishing between indolent and aggressive forms of the disease, and supports personalized treatment planning through predictive modeling and data-driven decision support.

Over the past 25 years, ML has become a key driver of the information technology revolution, shaping various domains. As a subfield of AI, ML enables computers to learn from data without explicit programming, a concept introduced by Samuel (18). His application of ML in checkers pioneered the use of games as experimental platforms for evaluating ML algorithms (18).

Campanella's work, which is the most frequently cited in this analysis, represents a milestone in the field of computational pathology (4). The authors developed a clinical-grade decision support system using a weakly supervised DL approach, applying multiple instance learning to 44,732 whole slide images from 15,187 patients. A ResNet34-based CNN was used for tile-level feature extraction, and the extracted features were integrated via a recurrent neural network to produce slide-level predictions. The system achieved exceptional performance with area under the curves of 0.991 for PCa, 0.989 for basal cell carcinoma, and 0.965 for breast cancer metastases. Notably, the model maintained 100% sensitivity for PCa detection while reducing the number of slides requiring pathologist review by over 75%. This demonstrates the feasibility of deploying DL systems in clinical workflows without manual pixel-level annotations, thanks to the scale and diversity of the data used.

Over the past year, a significant number of bibliometric analyses have been conducted regarding the application of ML in various medical fields, including kidney diseases (19,20), Crohn's disease (21), cardiomyopathy (22), psychiatry (23), and gynecology (24). The findings of these papers, similar to those of our study, indicate that the United States, China, and various European countries are at the forefront of research in ML.

The United States is home to some of the world's most prestigious universities, well-funded research programs, and a strong academic infrastructure. Over the past two decades, China has significantly increased its investments in science and technology, which has led to the development of numerous international collaborations. Institutions such as the National Cancer Institute (NCI) in the United States (https://www.cancer.gov), the National Science Foundation (https://www.nsf.gov), and the National Natural Science Foundation (NSFC) in China (https://www.nsfc.gov.cn/english/site_1/index.html) have supported joint projects in cancer research. In China, both the NSFC and the state Council support AI-based biomedical initiatives. The launch of China's AI strategy in 2017 (https://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm) represents an effort to bridge the gap with United States leadership in this field. Scientific collaboration between China and the United States persists despite fluctuations in political relations; for instance, the NCI and the Chinese Cancer Institute have engaged in joint projects for many years. The prominence of these two nations in scientific research can be attributed to their ongoing investments in AI and their status as economically and technologically advanced countries.

ML research in the field of PCa is predominantly led by countries such as China, the United States: the United Kingdom, Germany, Spain, and Italy. The authors with the highest number of publications and citations are primarily affiliated with institutions in these countries. Notably, the United States is prominent due to its high publication volume, leading citation metrics, and centrality in international collaborations. However, recent years have witnessed a significant increase in contributions from countries like China and India, which have rapidly emerged in the research landscape concerning PCa. This trend can be attributed to the rise in global scientific collaborations and advancements in data sharing practices. Furthermore, the substantial number of recent studies underscores the contemporary relevance of this subject. The integration of AI into various domains has been extensively investigated within the medical field, and it is reasonable to assert that AI will play an increasingly integral role in the future of medicine.

The prominence of the keywords "magnetic resonance imaging" and "MRI" among the top ten highlights that medical imaging, particularly MRI, remains the central focus of ML applications in

PCa research. This suggests a strong reliance on non-invasive diagnostic approaches and reflects the maturity of image-based datasets available for training algorithms. For future research, this trend implies a need to refine ML models for imaging tasks—such as lesion segmentation, radiomic feature extraction, and image-based risk assessment—while also encouraging the integration of imaging with other data types like genomics and clinical notes, to enable more comprehensive and personalized diagnostic tools.

The thematic mapping and keyword analyses conducted in this study reveal several underrepresented yet promising research directions in the field of ML-based PCa studies. Notably, while the majority of current research emphasizes diagnostic imaging—particularly MRI—there is considerable potential for expanding ML applications into areas such as treatment response prediction, active surveillance optimization, and long-term patient outcome modeling. Furthermore, the integration of radiomics with genomic data, often referred to as radiogenomics, represents an emerging field that remains insufficiently explored in PCa. This integration could facilitate more personalized risk stratification and treatment planning. Another crucial future direction involves enhancing the explainability and interpretability of ML models to ensure their acceptance in clinical settings. As AI systems become more complex, transparent mechanisms for decision-making and output justification will be essential for clinician trust and regulatory approval. Encouraging multidisciplinary collaborations, standardizing data formats, and establishing public repositories for high-quality annotated datasets will also be pivotal in driving innovation and clinical translation.

The analysis of the thematic map reveals that studies in urological oncology, particularly those focused on PCa detection, are increasingly influenced by AI and imaging techniques. It is crucial for urologists to identify clinically significant PCa. Numerous studies have been conducted on this topic (25,26). Additionally, there are publications that specifically examine the application of ML in this field (27,28). The clinical significance of this topic is substantial. Our thematic mapping indicates that studies pertaining to the diagnosis of PCa are encompassed by the primary themes. It is anticipated that CNN and DL will assume pivotal roles in the future diagnosis and treatment of PCa. This subject, which remains perpetually relevant, is expected to gain further prominence in the coming years.

The studies conducted in this domain are multidisciplinary, with the most commonly associated fields being radiology, nuclear medicine, and biochemistry. Bioinformatics methodologies, including ML, DL, radiomics, and gene expression analysis, are leading advancements in keyword analysis. Notably, there has been a substantial increase in both the volume of publications and citation density within this field over the past decade. Projections suggest that this trend will continue to escalate exponentially.

Here, we must also emphasize the necessity of thorough clinical validation before AI systems can be safely implemented in PCa care. Such validation processes generally include prospective studies, multi-center trials, and external validation using independent datasets that reflect clinical heterogeneity. These steps are crucial to ensure that models not only perform well on retrospective benchmarks but are also generalizable, reliable, and ethically sound for real-world use.

Although this study is bibliometric in nature and does not evaluate clinical or generative AI applications directly, the emergence of models such as ChatGPT underscores the accelerating pace and expanding scope of AI technologies across healthcare and biomedical research. While not the focus of this analysis, such developments contextualize the broader ecosystem in which ML-based PCa studies are evolving.

Recommendations for Future Research

One of the major concerns in ML-based PCa studies is the lack of comprehensive clinical validation, particularly through prospective and multicenter trials. Although many algorithms demonstrate strong performance with retrospective data, their clinical applicability remains limited without real-world validation. Additionally, our findings reveal that research collaborations are predominantly concentrated in a few countries, which restricts the global generalizability and adaptability of ML models. Beyond these structural limitations, several content-specific gaps have also been identified. For instance, ML applications in advanced stages of PCa—especially metastatic and treatment-resistant cases—are significantly underexplored. Moreover, while imaging data dominates current studies, there is a marked absence of research integrating diverse data modalities such as genomics, laboratory results, clinical narratives, and patient-reported outcomes. These data types hold the potential to enhance the predictive accuracy and clinical utility of AI-driven systems. Furthermore, limited attention has been given to developing real-time decision support tools for use during critical clinical procedures, including biopsy and radiotherapy planning. Addressing these gaps requires the development of interpretable, ethically responsible AI models that are rigorously validated and seamlessly integrated into routine clinical workflows.

Study Limitations

This study has several limitations. First, it relied solely on the WoS CC, thereby excluding studies indexed in other major databases such as PubMed and Scopus. Second, only English language articles were considered, potentially omitting high-quality research in other languages. It is noteworthy that there may be recent high-quality articles available in other languages. The citation of articles necessitates a systematic process, and the identification of quality studies will require time. Bibliometric analysis is based on publication counts and citation networks, which do not permit a direct evaluation of methodological quality. Future research should incorporate systematic reviews and meta-analyses to provide a more comprehensive evaluation of the field.

Conclusion

In conclusion, ML and AI applications present promising advancements in the management of PCa. However, further validation, use of large-scale datasets, and multidisciplinary collaborations are essential for the broader implementation of these technologies in clinical settings. Future research should prioritize the development of models that are validated with real-world data, promote extensive international collaborations, and expedite clinical validation processes. Notably, the integration of large language models and generative AI-based solutions into clinical practice represents a significant area for future investigation. Nevertheless, it is imperative to address concerns related to ethics, reliability, and the generalizability of these models.

Ethics

Ethics Committee Approval: Not applicable, as this bibliometric analysis did not involve direct interaction with human participants or collection of personal data.

Informed Consent: Not applicable, as this bibliometric analysis did not involve direct interaction with human participants or collection of personal data.

Acknowledgements

Publication: The results of the study were not published in full or in part in form of abstracts.

Contribution: There is not any contributors who may not be listed as authors.

Footnotes

Authorship Contributions

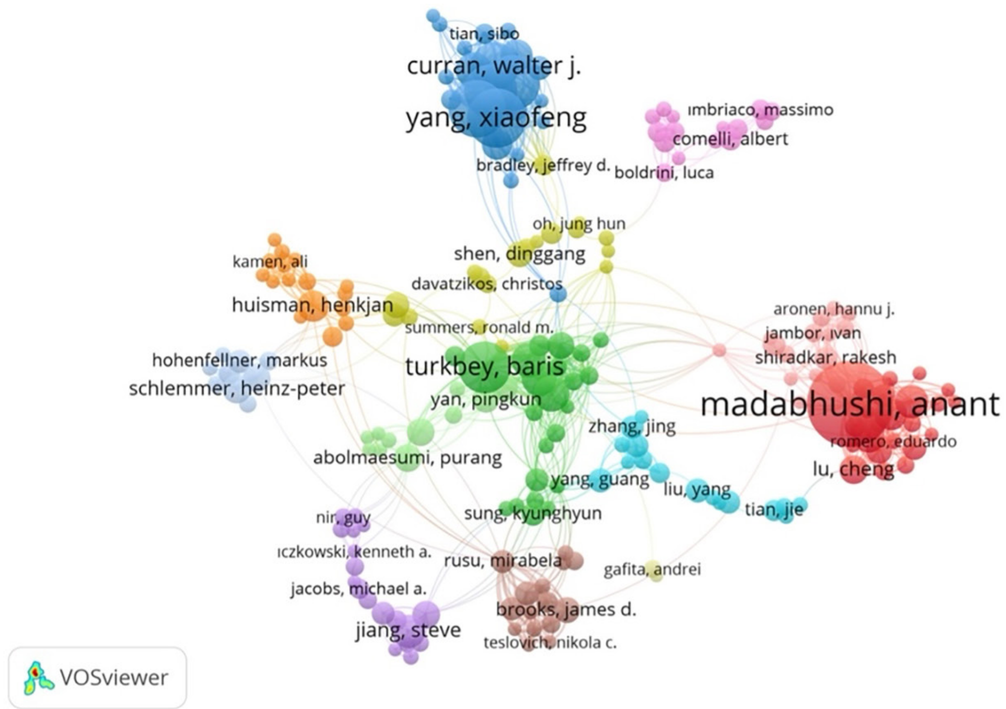
Surgical and Medical Practices: T.A., F.O., M.D., Concept: T.A., İ.H.Ş., Design: N.T., İ.Ö.Y., Data Collection or Processing: İ.H.Ş., V.İ., Analysis or Interpretation: V.İ., Literature Search: İ.H.Ş., İ.Ö.Y., Writing: T.A., F.O., İ.Ö.Y.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

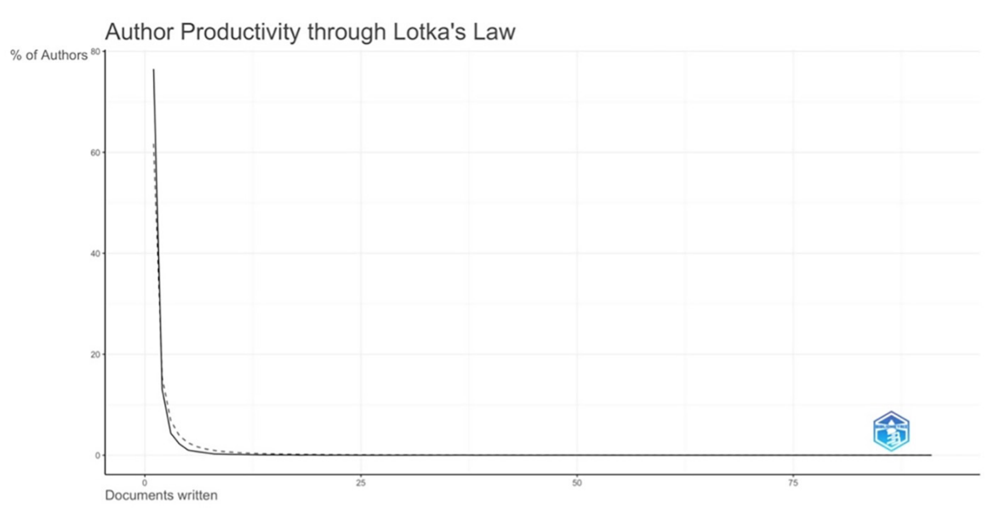
References

1. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. CA Cancer J Clin. 2023;73:17-48.
2. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017;42:60-88.
3. Bi WL, Hosny A, Schabath MB, et al. Artificial intelligence in cancer imaging: clinical challenges and applications. CA Cancer J Clin. 2019;69:127-157.
4. Campanella G, Hanna MG, Geneslaw L, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. Nat Med. 2019;25:1301-1309.
5. Liu Y. 3D Image Segmentation of MRI prostate based on a pytorch implementation of V-Net. J Phys Conf Ser. 2020;1549:1-6.
6. Abbasi AA, Hussain L, Awan IA, et al. Detecting prostate cancer using deep learning convolution neural network with transfer learning approach. Cogn Neurodyn. 2020;14:523-533.
7. Mylona E, Zaridis DI, Kalantzopoulos CN, et al. Optimizing radiomics for prostate cancer diagnosis: feature selection strategies, machine learning classifiers, and MRI sequences. Insights Imaging. 2024;15:265.
8. Nematollahi H, Moslehi M, Aminolroayaei F, et al. Diagnostic performance evaluation of multiparametric magnetic resonance imaging in the detection of prostate cancer with supervised machine learning methods. Diagnostics (Basel). 2023;13:806.
9. Bertelli E, Mercatelli L, Marzi C, et al. Machine and deep learning prediction of prostate cancer aggressiveness using multiparametric MRI. Front Oncol. 2021;11:802964.
10. Rodrigues A, Santinha J, Galvão B, et al. Prediction of prostate cancer disease aggressiveness using bi-parametric mri radiomics. Cancers (Basel). 2021;13:6065.
11. Cysouw MCF, Jansen BHE, van de Brug T, et al. Machine learning-based analysis of [18F]DCFPyL PET radiomics for risk stratification in primary prostate cancer. Eur J Nucl Med Mol Imaging. 2021;48:340-349.
12. T JMC, Arif M, Niessen WJ, et al. Automated classification of significant prostate cancer on MRI: a systematic review on the performance of machine learning applications. Cancers (Basel). 2020;12:1606.
13. Aria M, Cuccurullo C. Bibliometrix: an R-tool for comprehensive science mapping analysis. Journal of Informetrics. 2017;11:959-975.
14. Van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics. 2010;84:523-538.
15. Ding X, Yang Z. Knowledge mapping of platform research: a visual analysis using VOSviewer and CiteSpace. Electron Commer Res. 2022;22:787-809.
16. Thelwall M. Bibliometrics to webometrics. J Inf Sci. 2008;34:605-621
17. Pritchard A. Statistical bibliography or bibliometrics? J Doc. 1969;25:348-349.
18. Samuel AL. Some studies in machine learning using the game of checkers. IBM J Res Dev. 1959;3:206-226
19. Rao HH, Guo F, Tian J. Improving search strategies in bibliometric studies on machine learning in renal medicine. Int Urol Nephrol. 2025;57:1987-1988.
20. Li F, Hu C, Luo X. Research hotspots and frontiers of machine learning in renal medicine: a bibliometric and visual analysis from 2013 to 2024. Int Urol Nephrol. 2025;57:907-928.
21. Zhou Z, Yu C, Liu B, et al. Landscape of surgery in Crohn's disease across twenty years: insights from machine learning. Transl Gastroenterol Hepatol. 2024;9:64.
22. Chao T, Ge Y, Sun J, Wang C. Research landscape of genetics in dilated cardiomyopathy: insight from a bibliometric analysis. Front Cardiovasc Med. 2024;11:1362551.
23. Liu Z, Zhou Z, Ma J, et al. Major depressive disease research in BRICS: a bibliometric analysis of publications from 2003 to 2022. Asian J Psychiatr. 2024;92:103900.
24. Zhao Z, Hu B, Xu K, et al. A quantitative analysis of artificial intelligence research in cervical cancer: a bibliometric approach utilizing CiteSpace and VOSviewer. Front Oncol. 2024;14:1431142.
25. Rojo Domingo M, Conlin CC, Karunamuni R, et al. Utility of quantitative measurement of T2 using restriction spectrum imaging for detection of clinically significant prostate cancer. Sci Rep. 2024;14:31318.
26. Tamada T, Takeuchi M, Watanabe H, et al. Differentiating clinically significant prostate cancer from clinically insignificant prostate cancer using qualitative and semi-quantitative indices of dynamic contrast-enhanced MRI. Discov Oncol. 2024;15:770.
27. Zhao W, Hou M, Wang J, et al. Interpretable machine learning model for predicting clinically significant prostate cancer: integrating intratumoral and peritumoral radiomics with clinical and metabolic features. BMC Med Imaging. 2024;24:353.
28. Cai JC, Nakai H, Kuanar S, et al. Fully Automated deep learning model to detect clinically significant prostate cancer at MRI. Radiology. 2024;312:e232635.



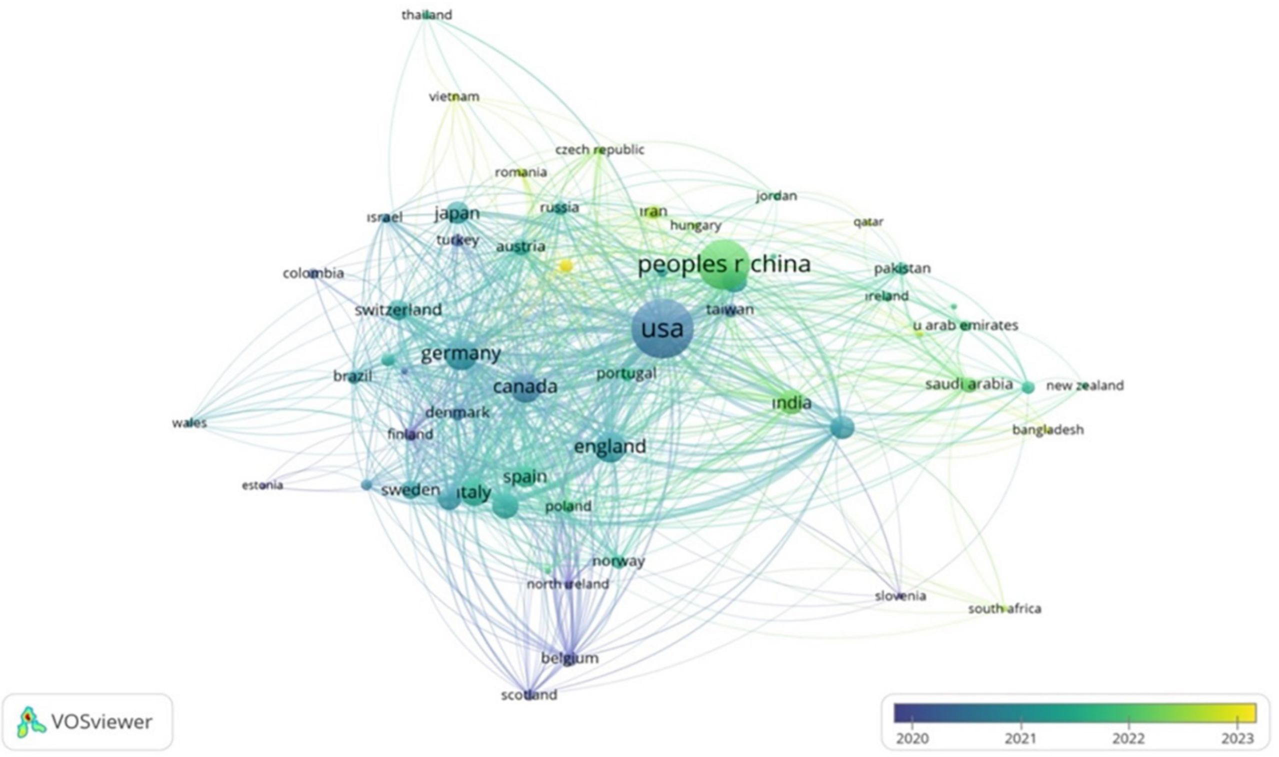
Supplementary Figure 1. Author co-authorship visualization

Articles with more than 25 authors were excluded and the complete count method was selected. The minimum number of articles and citations for an author was set at 5 and 100, respectively. These criteria were met by 259 authors. Weighting was based on the number of publications and the circle size on the map represents the number of articles



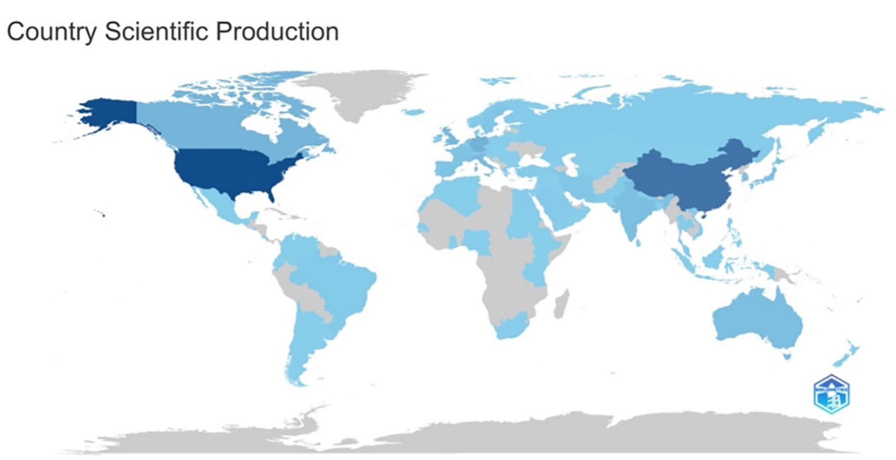
Supplementary Figure 2. Lotka's law graphics

Dashed lines indicate Lotka's law, while solid lines indicate the author productivity of the analysis



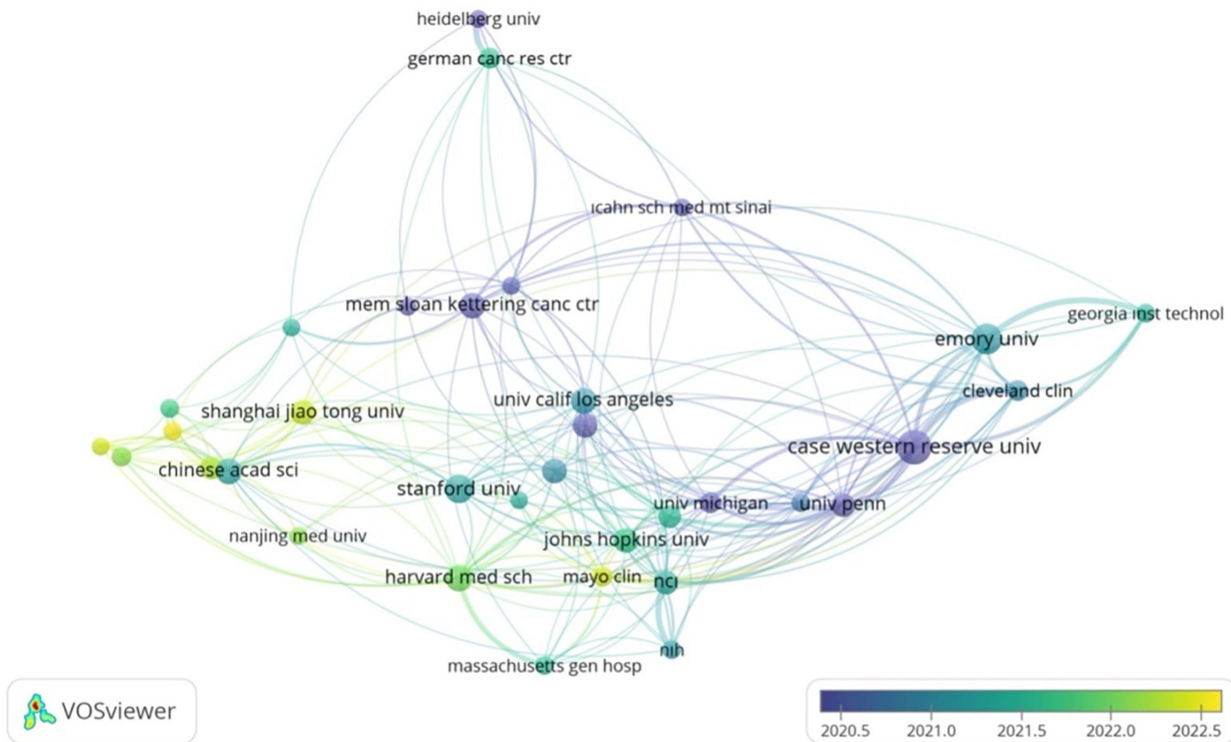
Supplementary Figure 3. Country co-authorship visualization

Articles with authors from more than 25 different countries were excluded and the full count method was selected. The minimum number of articles and citations for a country was set as 5. These criteria were met by 59 countries. The weighting was based on the number of publications and the circle size on the map represents the number of articles



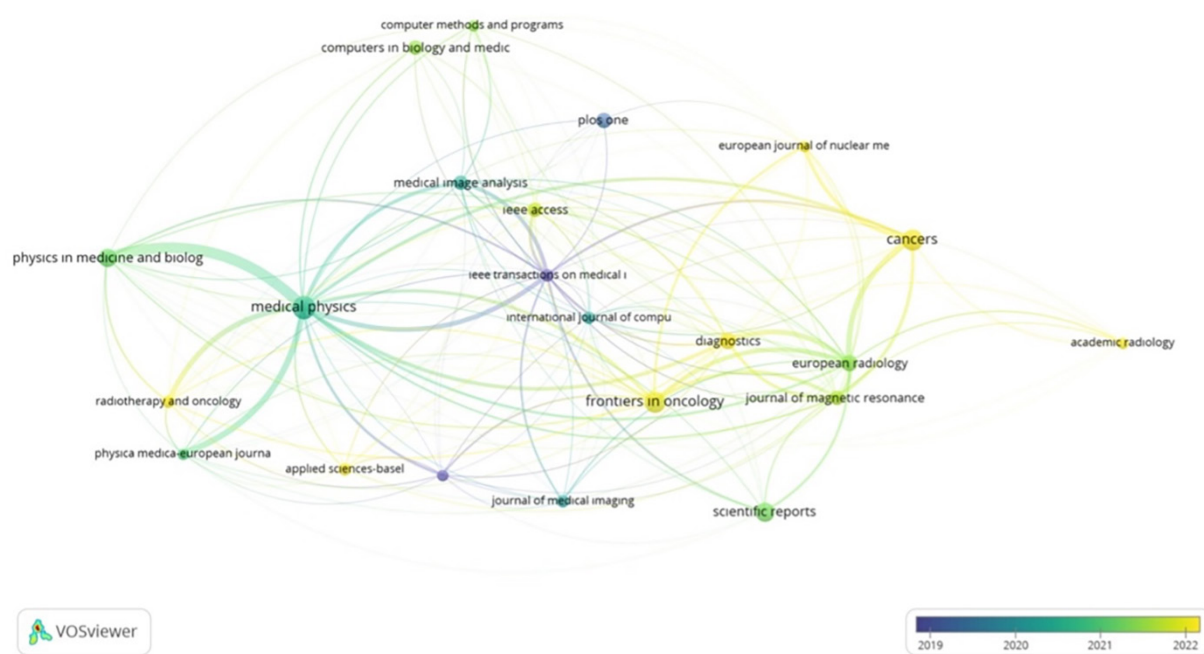
Supplementary Figure 4. Article status of countries

The countries depicted in grey do not have an associated article, while the darker shades represent those countries with the highest number of articles



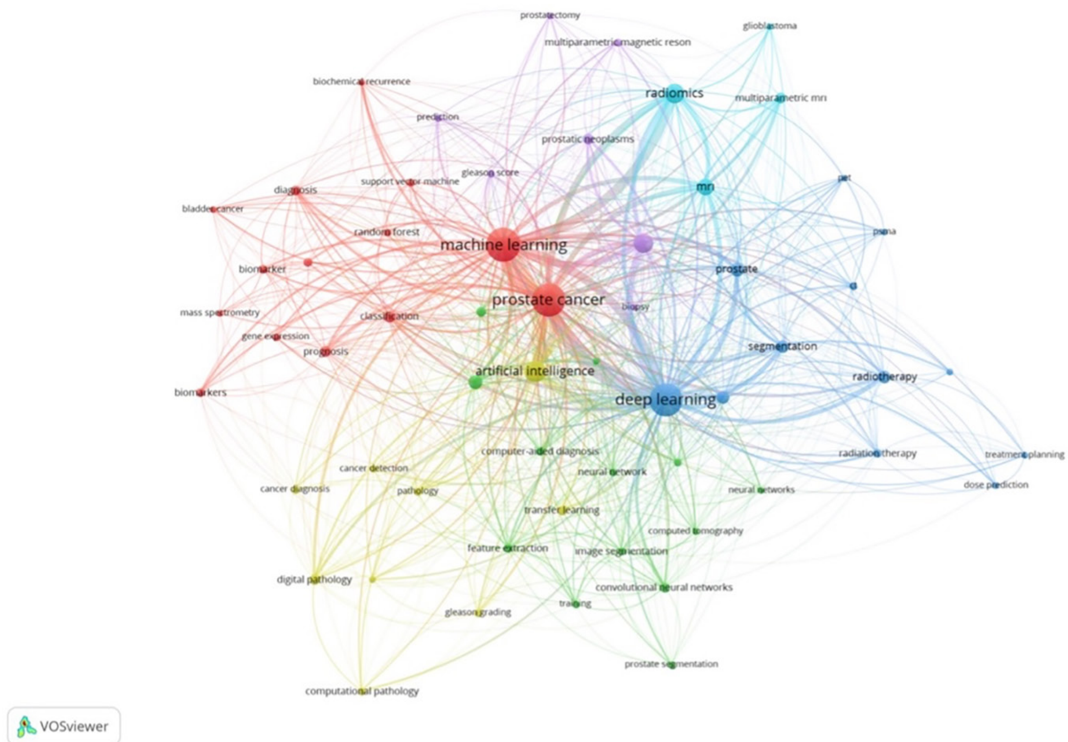
Supplementary Figure 5. Organization co-authorship visualization

Articles with authors from more than 25 different institutions were excluded and a complete count method was selected. The minimum number of articles and citations for an institution was set at 30 and 50, respectively. These criteria were met by 34 institutions. Weighting was based on the number of publications and the circle size on the map represents the number of articles



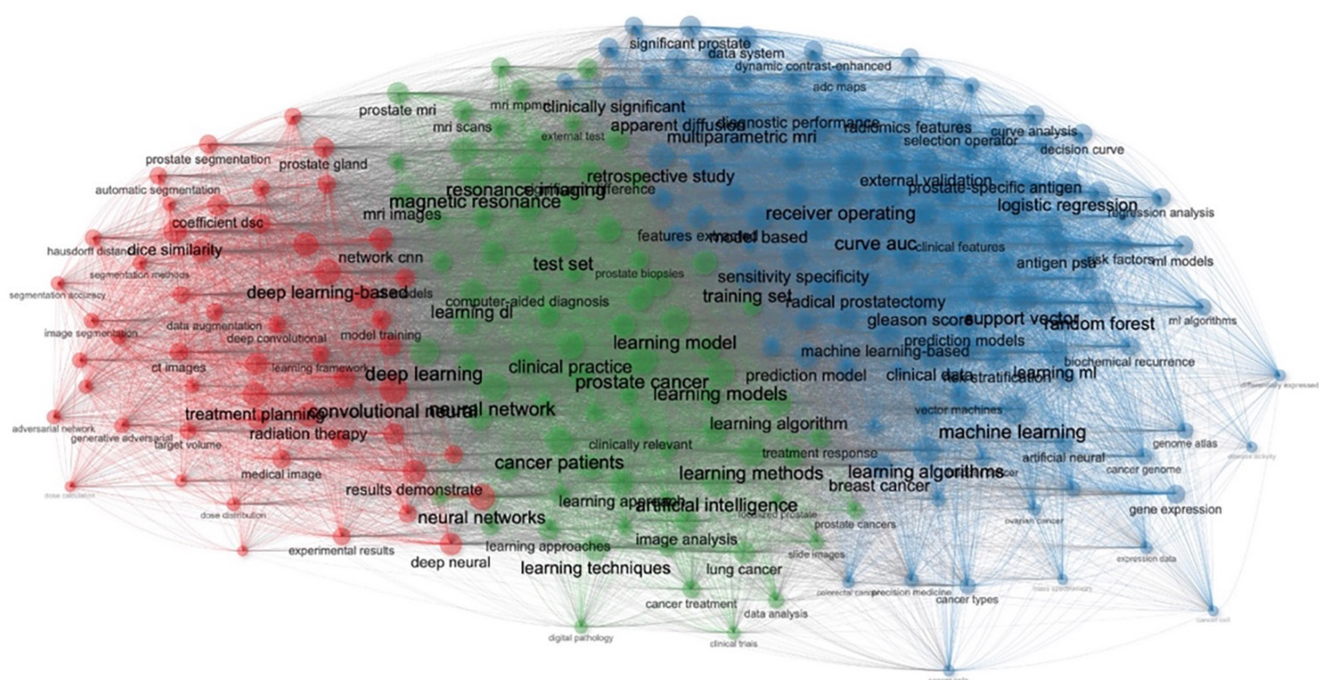
Supplementary Figure 6. Citation-source analysis visualization

The minimum number of articles and citations of a source was determined as 20 and 50, respectively. These criteria were met by 22 sources. Weighting was based on the number of publications and the circle size on the map represents the number of articles



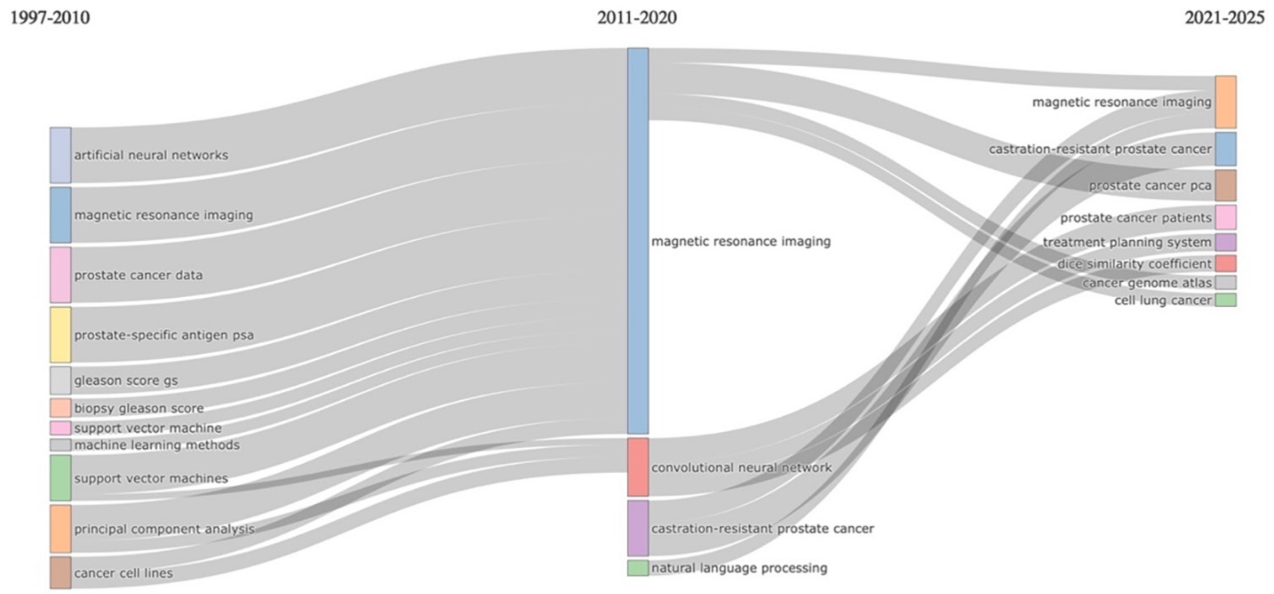
Supplementary Figure 7. Author Figure 6. keyword co-occurence visualization

The visualisation was based on the total number of occurrences and included author keywords that appeared at least 20 times. This criterion was met by 59 author keywords. Weighting was based on occurrence



Supplementary Figure 8. Thematic map

Field: Abstract, N-gram: Bigram, Word stemming: No



Supplementary Figure 9. Thematic evolution

Field: Abstract, N-gram: Trigram

Supplementary Table 1. Top ten authors with the highest number of publications

Author	Documents	Citations
Madabhushi, Anant	91	6,161
Yang, Xiaofeng	50	2,860
Liu, Tian	47	2,834
Lei, Yang	45	2,831
Wang, Tonghe	41	2,698
Turkbey, Baris	40	1,192
Curran, Walter J.	36	2,763
Patel, Pretesh	23	1,174
Choyke, Peter L.	23	499
Wood, Bradford J.	22	649

Supplementary Table 2. Top ten countries with the highest number of publications

Country	Documents	Citations
United States	1,216	41,635
China	794	14,092
Germany	262	8,339
England	245	7,529
Canada	235	8,405
Italy	192	4,256
Netherlands	163	7,347
India	147	2,138
South Korea	136	2,203
Australia	133	3,950

Supplementary Table 3. Top ten instutions with the highest number of publications

Organization	Documents	Citations	Country
Case Western Reserve University	101	6,986	USA
Emory University	79	3,904	USA
Stanford University	70	2,016	USA
Harvard Medical School	62	3,456	USA
Chinese Academy of Sciences	58	1,850	China
Memorial Sloan Kettering Cancer Center	58	3,390	USA
University of California Los Angeles	57	1,683	USA
University Toronto	55	2,356	Canada
University British Columbia	55	3,096	Canada
Johns Hopkins University	54	1,341	USA

USA: United States of America

Supplementary Table 4. Top ten sources with the highest number of publications

Source	WoS Index	WoS Quartil	Documents	Citations
Medical Physics	SCIE	Q1	140	4,383
Cancers	SCIE	Q1	113	1,338
Frontiers in Oncology	SCIE	Q2	110	1,225
Scientific Reports	SCIE	Q1	88	2,948
Physics in Medicine and Biology	SCIE	Q1	70	2,253
Diagnostics	SCIE	Q1	57	964
European Radiology	SCIE	Q1	52	2,055
PLos One	SCIE	Q1	44	1,389
Journal of Magnetic Resonance Imaging	SCIE	Q1	42	1,211
IEEE Access	SCIE	Q2	41	518

SCIE: Science Citation Index Expanded, WoS: Web of Science, IEEE: Institute of Electrical and Electronics Engineers

Supplementary Table 5. Top ten occurence author keywords

Keyword	Occurrences
Machine learning	900
Deep learning	846
Prostate cancer	817
Artificial intelligence	298
Radiomics	242
Magnetic resonance imaging	228
MRI	147
Prostate	104
Cancer	95
Convolutional neural network	84

MRI: Magnetic resonance imaging

Supplementary Table 6. Top ten documents with the highest number of citations

Document	DOI number	Citations
Campanella (2019)	DOI:10.1038/s41591-019-0508-1	1295
Bi (2019)	DOI:10.3322/caac.21552	990
Bera (2019)	DOI:10.1038/s41571-019-0252-y	758
Lu (2021)	DOI:10.1038/s41551-020-00682-w	736
Cruz (2006)	N/A	718
Litjens (2016)	DOI:10.1038/srep26286	680
Mobadersany (2018)	DOI:10.1073/pnas.1717139115	616
Xu (2016)	DOI:10.1109/TMI.2015.2458702	597
Su (2001)	N/A	536
Choy (2018)	DOI:10.1148/radiol.2018171820	493

N/A: Not applicable