Artificial intelligence in precision medicine for lung cancer: A bibliometric analysis

Yuchai Wang[1] (iD), Weilong Zhang[1], Xiang Liu[1], Li Tian[1], Wenjiao Li[1], Peng He[1], Sheng Huang[1,2], Fuyuan He[3] and Xue Pan[3] (iD)

Abstract

Background: The increasing body of evidence has been stimulating the application of artificial intelligence (AI) in precision medicine research for lung cancer. This trend necessitates a comprehensive overview of the growing number of publications to facilitate researchers' understanding of this field.

Method: The bibliometric data for the current analysis was extracted from the Web of Science Core Collection database, CiteSpace, VOSviewer ,and an online website were applied to the analysis.

Results: After the data were filtered, this search yielded 4062 manuscripts. And 92.27% of the papers were published from 2014 onwards. The main contributing countries were China, the United States, India, Japan, and Korea. These publications were mainly published in the following scientific disciplines, including Radiology Nuclear Medicine, Medical Imaging, Oncology, and Computer Science Notably, Li Weimin and Aerts Hugo J. W. L. stand out as leading authorities in this domain. In the keyword co-occurrence and co-citation cluster analysis of the publication, the knowledge base was divided into four clusters that are more easily understood, including screening, diagnosis, treatment, and prognosis.

Conclusion: This bibliometric study reveals deep learning frameworks and AI-based radiomics are receiving attention. High-quality and standardized data have the potential to revolutionize lung cancer screening and diagnosis in the era of precision medicine. However, the importance of high-quality clinical datasets, the development of new and combined AI models, and their consistent assessment for advancing research on AI applications in lung cancer are highlighted before current research can be effectively applied in clinical practice.

Keywords

Lung cancer, artificial intelligence, bibliometric analysis, knowledge base, hotspots

Introduction

Lung cancer has been a serious threat to human health. Approximately 80–90% of lung cancers are caused by smoking, and are also associated with secondhand smoke, radon exposure, coal combustion, occupational exposure to carcinogens and cooking fumes, and air pollution.[1–3] The latest Global Cancer Statistics 2022 report has been shown that lung cancer was the most frequently diagnosed cancer in 2022, responsible for almost 2.5 million new cases, or 1 in 8 cancers worldwide (12.4% of all cancers globally).[4] Complications arising from lung cancer severely reduce the quality of life and life expectancy of patients. The 5-year survival rate for patients is only 10–20% after diagnosis, attributed to the fact that lung cancer

[1]Department of Pharmacy, Hunan University of Chinese Medicine, Changsha, Hunan Province, China
[2]Jiuzhitang Co., Ltd, Changsha, Hunan Province, China
[3]School of Pharmacy, Hunan University of Chinese Medicine, Changsha, Hunan Province, China

YW and WZ contributed equally to this work.

Corresponding authors:
Xue Pan, School of Pharmacy, Hunan University of Chinese Medicine, 300 Xueshi Road, Yuelu District, Hunan Province, China 410208.
Email:003920@hnucm.edu.cn

Fuyuan He, School of Pharmacy, Hunan University of Chinese Medicine, 300 Xueshi Road, Yuelu District, Hunan Province, China 410208.
Email: pharmsharking@163.com

is often not diagnosed until late in life and has a poor prognosis.[5] During the clinical workup of lung cancer, massive multidimensional datasets including text, images, vital sign data, genetic data, and other rich data types have been generated.[6,7] Thorough and iterative statistical data, analysis, and reading images or pathology slides to make clinical decisions lead to physician exhaustion. In addition, high false-positive and false-negative results,[8,9] cost-effectiveness,[10] and other issues in daily practice pose challenges for precision medicine.[11]

In recent years, the emerging artificial intelligence (AI) holds certain potential for solving these problems, the holistic definition of AI is quite broad, and it is regarded in the medical field as applications or technologies capable of learning and recognizing patterns and features from large amounts of representative data, which imitate the cognitive functions associated with human thought.[12,13] This information is then integrated into a domain-specific decision-making process (Figure 1). This application includes datasets for training, preprocessing methods, algorithms for generating predictive models and speeding up model construction, and pretrained models that inherit and utilize the experience of previous generations.[14] The core of AI is machine learning, which includes powerful algorithms such as deep learning, convolutional neural networks (CNNs), decision trees, etc.[15,16] These advanced AI algorithms build models that enhance medical image and data analysis, enabling them to efficiently analyze multidimensional datasets, enhance image analysis and interpretation, and provide decision support systems that allow researchers and clinicians to navigate the complex lung cancer data to provide valuable insights and recommendations for precision medicine.[17,18] Currently, the value of AI in clinical decision

making in lung cancer is being revealed in a growing number of clinical and experimental studies, including lung cancer screening,[19] assisting in lung cancer diagnosis,[20] prediction,[21] and assessing the treatment efficacy and prognosis.[22] This will attract both newcomers and seniors to consider research topics in this field, driving the field to an urgent need for a systematic description of the current state of research, development processes, and future research hotspots.

Bibliometrics is widely used in the fields of medicine, architecture, and psychology.[23,24] Different from traditional reviews with specific subjective characteristics, bibliometrics is an interdisciplinary discipline that analyzes knowledge carriers quantitatively via mathematical and statistical methods. It not only reveals important bibliometric indicators such as authoritative and productive countries, authors, journals, institutions, etc., to further identify research themes in the field, but also identifies highly cited key literature and keywords so as to explore research hotspots and frontier directions, and visually presents a panoramic view of the research field. This study aims to summarize global research trends and hotspots by identifying core contributing authors, institutions, countries, and regions in the field as well as visual measurements of keywords and cited literature to provide new perspectives on future directions for researchers and clinical decisions for clinicians.

Methods

Database sources and search strategy

We searched the PubMed database for Medical Subject Headings (MeSH) terms to help identify search terms. The
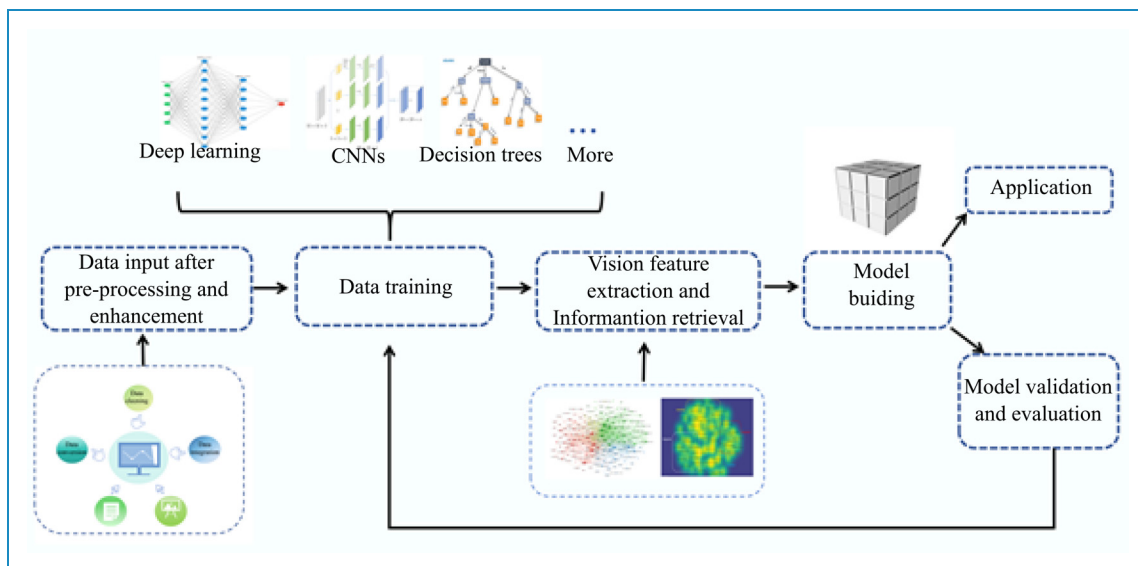

Figure 1. The general process of artificial intelligence model building.

Web of Science Core Collection was then selected as the original database for evaluation of the publications of more than 12,000 core journals. The retrieval time is set to 15 March 2024. Table 1 shows our search strategy. XL and WZ conducted a screening to include only original research articles and literature reviews and excluded publications that did not meet the inclusion criteria. The complete exported plain text records include the article title, author name, abstract, publication date, keywords, citations, etc. The retrieved files were deduplicated by CiteSpace to obtain a total of 4062 valid records, including 122,308 references. Figure 2 displays the flow of data extraction and analysis.

Statistical analysis

CiteSpace is a Java-based bibliometric software developed by Professor Chaomei Chen.[25] It enables quantitative analysis of domain-specific literature (collections) to explore valuable information and knowledge about the evolution of subject areas. CiteSpace parameter settings: set time partitioning parameters, the time slice is set to 1 year, TOP N is set to 50, the rest are kept as the system default, generate the author cooperation network; the threshold value (Top N% per slice) is selected as 30, generate the institutional cooperative network map and the burst keywords map.

VOSviewer, developed by Professors Van Eck and Waltman, is a document visualization software. It analyzes the frequency of co-occurrence of keywords and the co-citation frequency of cited literature to determine the relationship between topics. This clarification helps to understand the research content and structure of the field.[26,27] The analysis type is set to co-occurrence; the "complete count" option is selected; the minimum occurrence of the keyword is set to

Table 1. The topic search queries.

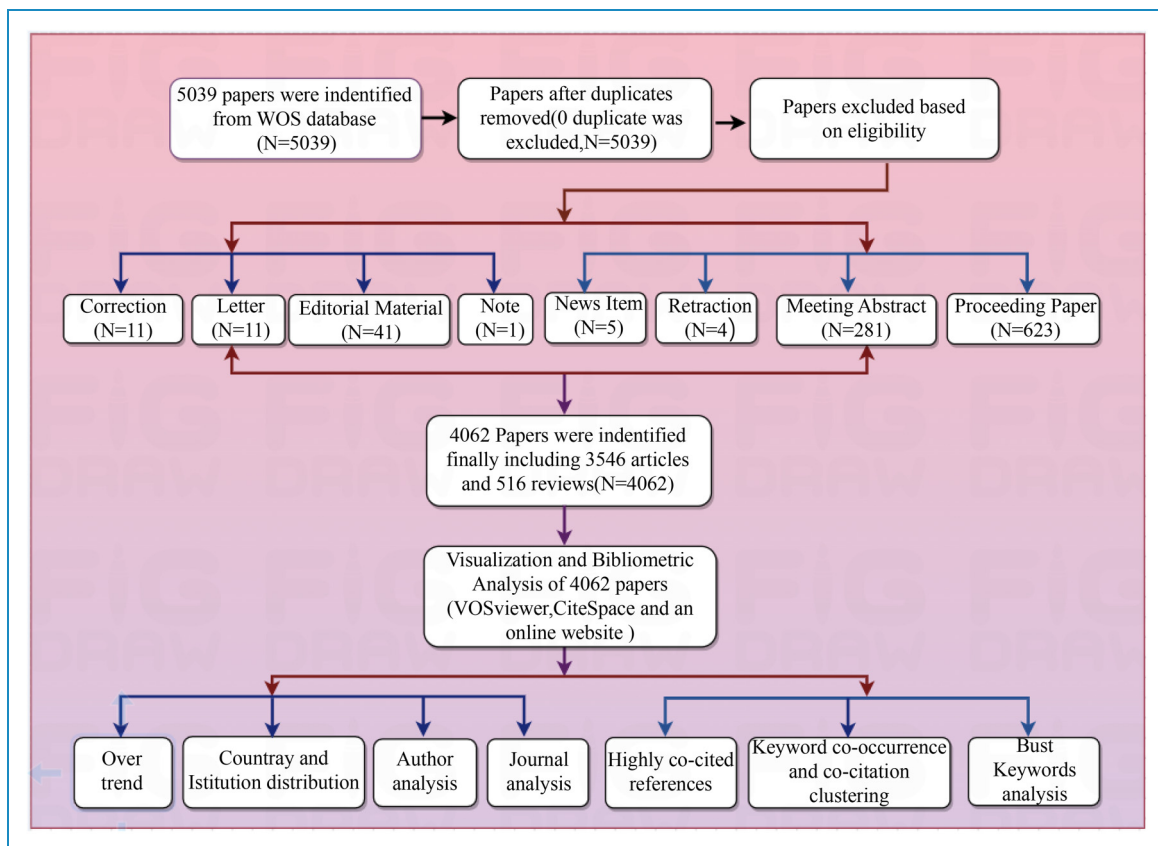| Set	Search query	Result
#1	TS = (AI OR "Artificial Intelligence" OR "Neural Network" OR "Transfer Learning OR Machine Learning" OR "Deep Learning" OR "Hierarchical Learning" OR "Machine Intelligence")	808,953
#2	TS = (Lung cancer OR Lung Neoplasm* OR Pulmonary Neoplasm* OR Pulmonary Cancer* OR "Cancer of Lung" OR Lung Carcinoma OR Pulmonary Carcinoma)	474,649
#3	#1 AND #2	5039

Figure 2. The flow of data extraction and analysis in the study (by Figdraw).

five based on the research requirements. Select network visualization to generate a keyword co-occurrence knowledge map; select density visualization to generate a co-occurrence knowledge map of the cited literature.[28,29]

An online bibliometrics website (http://bibliometric.com/) is used to visualize the national cooperation network. To further analyze the scientificity of the studies, the retrieved articles and journals were checked for the latest impact factor (IF) and the number of citations.

Result

Over trend

The first article was published back in 1992.[30] Nevertheless, the past 33 years can be divided into 2 periods (the search date for 2024 ends on March 15) (Figure 3(a)). From 1992 to 2013, when the number of publications was low and growing slowly, accounted for only 7.73% of the entire publications, with an average of 14 publications per year; from 2014 to 15 March 2024, when the number of publications accounted for 92.27% of the total, with an average of 344 publications per year, reaching 1003 publications in 2023. Overall, the number of publications on AI applications to lung cancer research has grown rapidly each year over the past decade, demonstrating the growing academic interest in the field (Figure 3(b) and (c)).

Country and institution distribution

Centrality is an important indicator for evaluating the importance of nodes in the network, and the higher the centrality, the larger the weight of the node in the network.[31] Eighty-nine countries/regions contributed to the publications, and the top four countries were China, the United States, India, and South Korea. The top four countries in terms of centrality (Table 2) are the United States (0.15), India (0.14), England (0.12), and Germany (0.09). China has produced the largest number of publications since 2008, accounting for 22.63%.
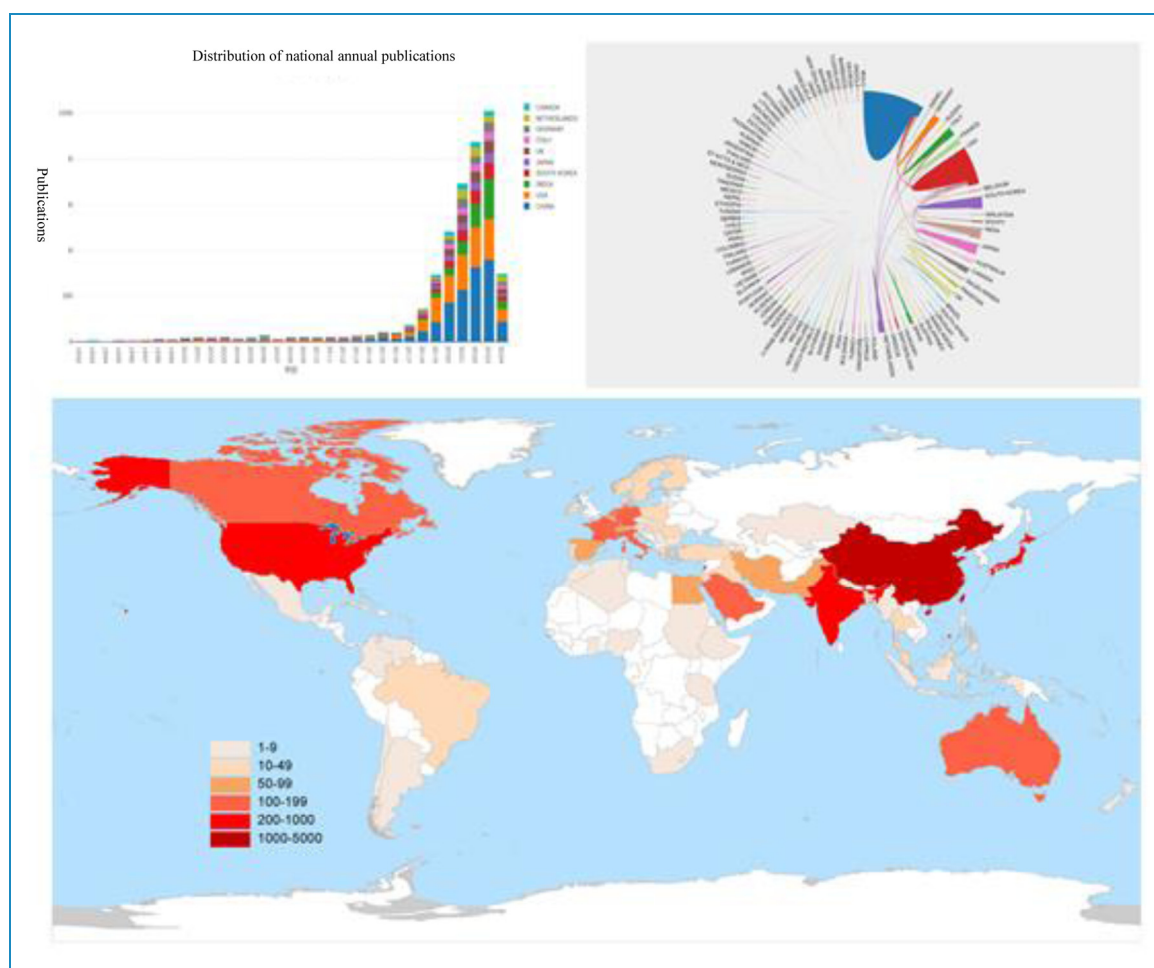

Figure 3. (a) Distribution of national annual publications. (b) Visual map of cross-country/regional collaborations. The thickness and quantity of boundaries between countries reflect the frequency of collaboration. (c) Geographical distribution: map of the geographical distribution based on the total number of publications in different countries/regions.

Table 2. Top 10 productive countries and institutions.

Rank	Country	Publications	Percentage	Centrality	Institution	Publications	Affiliation
1	China PR	1306	22.63	0.02	Shanghai Jiao Tong University	88	China PR
2	The United States	995	17.24	0.15	Chinese Academy of Sciences	76	China PR
3	India	459	7.95	0.14	Harvard University	74	The United States
4	South Korea	246	4.26	0.02	Fudan University	62	China PR
5	Japan	245	4.24	0.02	Stanford University	55	The United States
6	England	202	3.50	0.12	Seoul National University	53	South Korea
7	Italy	182	3.15	0.06	Maastricht University	48	The Netherlands
8	Germany	176	3.05	0.09	Memorial Sloan Kettering Cancer Counseling Center	46	The United States
9	The Netherlands	174	3.02	0.05	MassachusettS General Hospital	45	The United States
10	Canada	128	2.22	0.06	Northeastern University	43	The United States

Many countries around the world have participated in and enriched research in this field since the beginning of the twenty-first century, especially in East Asian countries.

In total, 742 institutions are involved in the development of this research field. We list the 10 most productive institutions including specific information (Table 2). Shanghai Jiao Tong University is the most productive institution, with 88 publications, indicating its great contribution to this field.

Author analysis

The top 10 most prolific authors in AI applied to lung cancer research are listed in Table 3, as well as their H-index, total citations, and affiliation. The most productive author was Li, Weimin from West China Hospital, China, with 24 articles, an H-index of 44, and 380 citations. The most cited author was Aerts, Hugo J. W. L from the Netherlands, who works at Harvard Medical School and Maastricht University, and has published 21 articles with an average of 249.43 citations per article.

Journal analysis

The 4062 records cover 1046 journals. Table 4 listed the top 10 journals in which the research results of AI applications for lung cancer are mainly published. "Frontiers in Oncology" is ranked first for the number of publications. Moreover, "Medical

Physics" is ranked first in terms of citation frequency. The average Impact Factor (IF) of the top 10 journals was 4.57, and the average number of citations for these journals was 1451.2. In total, 1046 journals were involved in 135 categories of Radiology Nuclear Medicine, Medical Imaging, Oncology, Computer Science, Biomedical Engineering, etc.

Highly co-cited references analysis

The top 10 most highly cited papers in AI applied to lung cancer are listed in Table 5, and they have been co-cited more than 11,000 times. The most cited article was "Computational Radiomics System to Decode the Radiographic Phenotype," published by Van Griethuysen, JJM from the Netherlands Cancer Institute in Cancer Research in 2017, cited 3083 times in 9 years. This paper described the workflow and architecture of "PyRadiomics" and demonstrated its application in characterizing lung lesions. The second-ranked paper investigated how the performance of deep CNNs trained from scratch compared with that of pre-trained CNNs when fine-tuned in a layer-wise manner, specifically when applied to lung medical imaging tasks. The third-ranked paper trained a deep CNN (inception v3) to classify adenocarcinoma, carcinoma, and normal lung tissue. The fourth-ranked paper presented a deep learning algorithm that used a patient's current

Table 3. The top 10 prolific authors.

Rank	Author	Publications	H index	Total citations	Citations per item	Affiliation
1	Li, Weimin	24	44	380	15.83	China PR
2	Goo, Jin Mo	21	42	861	41.00	South Korea
3	Qian, Wei	21	45	669	31.86	China PR
4	Aerts, Hugo J. W. L	21	73	5238	249.43	The Netherlands
5	Park, Chang Min	19	50	805	43.37	South Korea
6	Wang, Chengdi	19	15	234	12.32	China PR
7	Kim, Hyungjin	16	86	393	24.56	South Korea
8	Qi, Shouliang	16	24	357	22.31	China PR
9	Tian, Jie	16	24	1287	80.44	China PR
10	Wang, Jing	16	8	283	17.69	China PR

and prior computed tomography volumes to predict the risk of lung cancer.

Analysis of keywords and co-citation clustering

Keywords directly reflect the central concept of certain literature. The more occurrences in the same literature, the hotter the research in the field. Closely linked keywords depict the core themes and contents of the field. In addition, co-citation clustering discovers the topics of the research field by visualizing the high co-citation relationships among a set of literature because the references constitute the knowledge base of the field.[42] We, therefore, grouped studies with high relevance to identify the central topics in the field of AI applied to lung cancer.[43] Keyword co-occurrence clustering and co-citation clustering graphs were constructed in VOSviewer (Figures 4 and 5). Table 6 lists the representative keywords featured in each module as well as the relevant literature to help understand better.

Cluster #1 (screening)

Cluster 1 focused on the application of AI in lung cancer screening, such as pulmonary nodule, molecular biomarkers, computer-aided detection, automatic detection, false-positive reduction. The earliest application of AI was in lung cancer screening. Various factors, primarily smoking, cause damage to lung tissue and trigger an inflammatory response, which can lead to the formation of nodules or other lesions, and chest X-rays, computed tomography (CT) scans or other imaging techniques are common preventive screening tools.[44]

Cluster #2 (diagnosis)

Cluster 2 focused on the application of AI in the diagnosis of lung cancer, such as diagnosis, computer-aided diagnosis, image segmentation, feature, and CT. AI has made significant strides in lung cancer diagnosis, and it has created a noninvasive way of detection. Thanks to the widespread use of whole-section imaging and imaging techniques applied to tissue sections for clinical applications, a wealth of high-resolution pathology images and medical images is available. These images can be used to train AI models in pathology tasks such as lung nodule segmentation, cancer cell identification, and cancer type classification.[45]

Cluster #3 (treatment)

Cluster 3 focused on the application of AI in the treatment of lung cancer, such as cell lung cancer, radiotherapy, therapy, image classification, and immunotherapy. AI affords the opportunity to model intelligent treatments through computer systems that hinge on staging, tumor location, histology, and genetic changes, thereby aiding in the interpretation of crucial information concerning a patient's disease. By providing pertinent evidence, AI assists doctors in formulating treatment plans and boosts clinical decision-making efficiency for patients.

Cluster #4 (prognosis)

Cluster 4 focused on the application of AI in the prognosis of lung cancer, such as survival, prognostic factor, risk, risk factor, benign, prognostic value, and survival prediction.

Table 4. The top 10 productive journals.

Rank	Journal	Publications	Citations	Impact Factor	Research somain
1	Frontiers in Oncology	123	1240	4.7	Oncology & Cancer Research
2	Medical Physics	120	3775	3.8	Radiology, Nuclear Medicine & Medical Imaging
3	Cancers	106	996	5.2	Oncology & Cancer Research
4	Scientific Reports	95	1637	4.6	Multidisciplinary Sciences
5	IEEE Access	87	1511	3.9	General Engineering & Computer Science & Materials Science
6	Diagnostics	73	600	3.6	Clinical Biochemistry
7	Physics in Medicine and Biology	65	1681	3.5	Medical Science & Biophysics
8	Computers in Biology and Medicine	64	2397	7.7	Computer Science Applications & Health Informatics
9	Multimedia Tools and Applications	51	316	3.6	Computer Science & Information System Science
10	Biomedical Signal Processing and Control	51	359	5.1	Engineering Technology & Biomedicine

Multiple factors are associated with lung cancer prognosis, however, improving prognostic outcomes based solely on these factors can be inefficient and subjective.[46] Predictive models for lung cancer combined with AI can effectively improve the survival of patients with lung cancer by predicting treatment outcomes and shaping personalized clinical care plans.[47]

To show the trends of the four clusters, a quantitative visualization of the annual publications for each cluster is given in Figure 6. The results of this visualization align with the findings from the cluster analysis. This may be attributed to AI facilitating the integration of multimodal data, including imaging, genomics, and clinical data, leading to a more comprehensive and accurate assessment of lung cancer risk and prognosis. Conversely, the treatment cluster has the fewest number of published articles because of its complex characteristics. Advances in screening and diagnosis are likely to have a large impact on the foundation of AI applications in lung cancer.

Burst keywords analysis

A burst detection module in CiteSpace identifies significant changes in keywords within a period, determining if a topic is declining or rising. A high-breaking keyword indicates rapid growth in interest among researchers. Through burst analysis, research topics, and themes are revealed as they emerge, evolve, and decline, and research hotspots shift.

As shown in Figure 7, the top 25 keywords in AI applications for lung cancer from 1992 to 2024 have experienced a dynamic evolution. CNN, CT image, radiogenomics, COVID-19, generative adversarial network, artificial intelligence, deep learning, immunotherapy, deep, and immune checkpoint inhibitor are likely to be the research hotspots in the future.

Discussion

The great potential of AI in lung cancer research has led more researchers to consider the research topic in this field. A bibliometric study using information technology as a medium has presented current research results related to the field of AI application in lung cancer, with accurate and intuitive bibliometric indicators and knowledge maps to provide a more comprehensive and objective reference for the evolution process, scientific evaluation, and trend prediction of research topics.

The study reveals that the number of publications on the use of AI in lung cancer has been increasing during 1992 to 2024. The United States and China dominate in terms of the number of publications among the 10 countries. East Asian countries have an advantage in the volume of publications. It may be explained by the high demand for medical resources from their populations and the institutions' focus on academic collaboration and knowledge exchange. The United States occupies a leadership position in the global cooperation network. Developed countries collaborate more

Table 5. The top 10 co-cited publications.

Rank	Author	Year	Title	Origin	Citations	Average citations per year
1	Van Griethuysen et al.[32]	2017	Computational Radiomics System to Decode the Radiographic Phenotype	Cancer Research	3083	385.36
2	Tajbakhsh et al.[33]	2016	Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?	IEEE Transactions on Medical Imaging	1739	193.22
3	Coudray et al.[34]	2018	Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning	Nature Medicine	1352	193.14
4	Ardila et al.[35]	2019	End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography	Nature Medicine	903	150.50
5	Bi et al.[36]	2019	Artificial intelligence in cancer imaging: Clinical challenges and applications	CA-A Cancer Journal for Clinicians	798	133.00
6	Setio et al.[37]	2016	Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks	IEEE Transactions on Medical Imaging	777	86.33
7	MANGELS et al.[38]	1993	Carotenoid Content of Fruits and Vegetables—An Evaluation of Analytic Data	Journal of the American Dietetic Association	667	20.84
8	Rajpurkar et al.[39]	2018	Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists	PLoS Medicine	582	83.14
9	Alom et al.[40]	2019	Recurrent residual U-Net for medical image segmentation	Journal of Medical Imaging	567	94.5
10	Setio et al.[41]	2017	Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge	Medical Image Analysis	558	69.75

frequently and produce more publications. This phenomenon may be related to good economic support, a well-developed healthcare system, and excellent hardware and software, so that the demand for clinical research, big data collection, and AI model building could be satisfied.[48,49] This is a good trend for the wider application of AI in lung cancer.

Five of the top 10 institutions are located in the United States, suggesting that US research agencies are critical in the domain of AI in lung cancer applications and that they may be conducting deeper and more pioneering work. Interestingly, five of these are comprehensive universities, such as the Shanghai Jiao Tong University, the University of California system, and the Chinese Academy of Sciences. This indicates that multidisciplinary crossover and interinstitutional collaboration can increase the productivity and impact of research.

"Frontiers in Oncology" has the highest number of publications, "Medical Physics" has the potential to produce more high-quality papers in the future. The top 10 journals all have high impact factor, citation counts, and JCR divisions and are considered core journals. Remarkably, 116 JCR categories are covered, implying multidisciplinary collaboration for the flourishing of the field.

The top 10 highly cited references mostly appeared post-2016, with research topics including the proposal
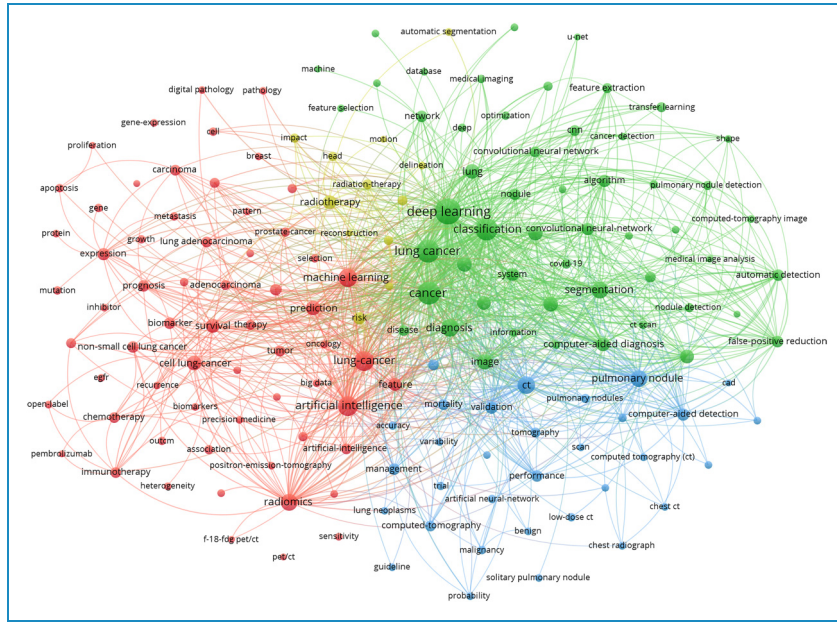
Figure 4. Co-occurrence keyword clustering. The size of the circles represents the total frequency of keyword occurrences, the lines indicate the strength of the association between keywords, and the same color means that their co-occurrence is under the same cluster.
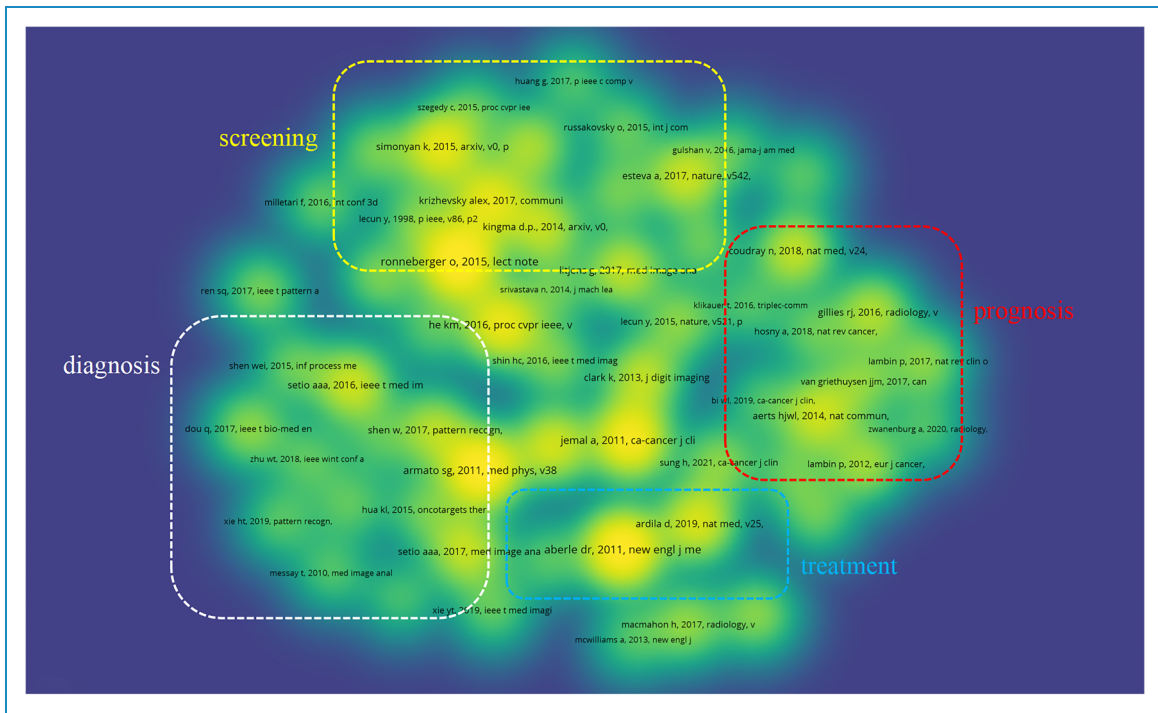

Figure 5. Co-citation clustering. Each heading is a reference, and references with relevance form color blocks that can be defined as a cluster.

and application of multiple novel deep learning frameworks, and the application of AI in lung cancer screening and diagnosis. The presentation of new AI models can often contribute to the flourishing of the field.

From the perspective of keyword co-occurrence and co-citation clusters, we can observe that the knowledge base of AI applications in lung cancer is divided into four clusters that are mostly understood: screening, diagnosis,

Table 6. Four theme clusters and their representative keywords and cited literature.

Clusters	Keywords	Cited references
Screening	Pulmonary nodule, molecular biomarkers, computer aided detection, automatic detection, false positive reduction	1. A deep learning system to screen novel coronavirus disease 2019 Pneumonia 2. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge 3. Biomarkers in lung cancer screening: Achievements, promises, and challenges
Diagnosis	Diagnosis, computer aided diagnosis, image segmentation, feature, computed tomography	1. Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks 2. Data-efficient and weakly supervised computational pathology on whole-slide images 3. Early-stage lung cancer diagnosis by deep learning-based spectroscopic analysis of circulating exosomes
Treatment	Cell lung cancer, radiotherapy, therapy, image classification, immunotherapy	1. Predicting response to cancer immunotherapy using noninvasive radiomic biomarkers 2. Prediction of respiratory tumour motion for real-time image-guided radiotherapy 3. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study
Prognosis	Survival, prognostic factor, risk, risk factor, benign, prognostic value, survival prediction	1. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma 2. Deep learning for lung cancer prognostication: A retrospective multi-cohort radiomics study 3. Deep learning predicts lung cancer treatment response from serial medical imaging

treatment, and prognosis. The statistical results of the annual publication volume for the four clusters reveal that more research is used more widely for lung cancer screening and diagnosis. Novel AI models are being actively applied in the field of research, screening, and diagnostic interpretation of image information. This is attributed to the increasing number of patients undergoing lung cancer screening and early-stage diagnosis, as well as the lower occurrence of complications compared to patients in advanced stages. These factors facilitate patient cooperation with researchers and their understanding of the utilization of these data. The process of standardized data collection ensures the quality and consistency of the data, enabling the models to analyze and interpret diverse image data more reliably. These advancements indicate substantial breakthroughs in the future and may revolutionize early clinical detection and characterization of lung cancer.

The analysis of burst keywords has shown clear evolutionary progress in the application of AI in lung cancer research. The early stage was exploratory due to the limitations of computer hardware and software technology. The United States is at the forefront of research on AI applications in lung cancer, and researchers' attention is focused on the development and selection of algorithmic models, with the artificial neural network being the most popular algorithm; exploring the feasibility of automated AI detection in cancer, mainly by screening lung nodules in chest CT scans and computer-aided study of P53 tumor suppressor gene. The emergent words include automated detection, P53, solitary pulmonary nodule, carcinoma, cancer, and computer-aided detection.
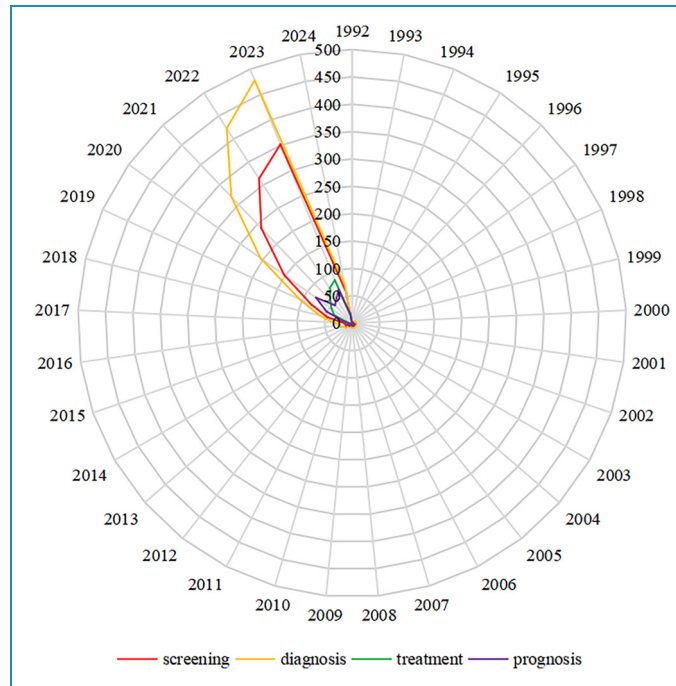
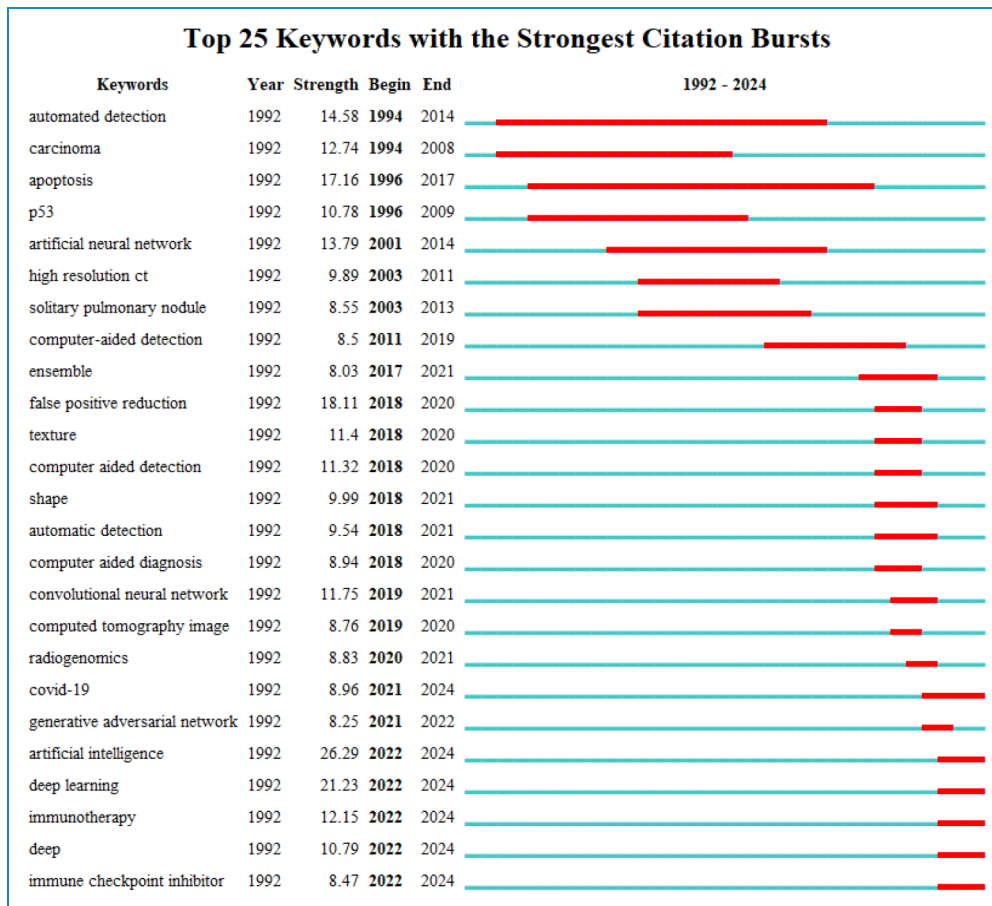**Figure 6.** Timeline of publications in four clusters.

Figure 7. Twenty-four keywords with strong bursts. Time interval is represented by the blue line, and burst keywords by the red line.

Limitation

This study has some limitations. First, this study may only include literature from specific databases with specific keywords. This study only covers literature written in English, which may lead to the oversight of research findings published in other languages. However, we believe that these differences may not have altered the overall trend of this study.

Conclusion

This bibliometric analysis reveals a global expansion of research on the application of AI in lung cancer. The substantial increase in publications after 2014 reflects the growing importance of this research field. This study identifies the top institutions, researchers, and journal worldwide involved in the application of AI in lung cancer research. Shanghai Jiao Tong University is the most productive institution of articles, Hugo J. W. L is the most influential author and "Frontiers in Oncology" is the most active journal. Key research areas include screening, diagnosis, treatment, and prognosis. Research hotspots identified include lung nodules, hepatocellular carcinoma, computer-aided diagnosis, image analysis, and the consistency of AI algorithms. In summary, this study provides insights into current trends, key contributors, and research hotspots for AI applications in lung cancer. These findings contribute to the understanding of the field and provide valuable guidance for future AI research in precision medicine for lung cancer and other cancers.

ORCID iDs: Yuchai Wang [ID] https://orcid.org/0009-0003-6851-8113
Xue Pan [ID] https://orcid.org/0009-0004-6324-6398

References

1. Huang J, Deng Y, Tin MS, et al. Distribution, risk factors, and temporal trends for lung cancer incidence and mortality: a global analysis. *Chest* 2022; 161: 1101–1111.
2. Chen S and Wu S. Identifying lung cancer risk factors in the elderly using deep neural networks: quantitative analysis of web-based survey data. *J Med Internet Res* 2020; 22: e17695.
3. Pallis AG and Syrigos KN. Lung cancer in never smokers: disease characteristics and risk factors. *Crit Rev Oncol Hematol* 2013; 88: 494–503.
4. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2024; 74: 229–263.
5. Xu K, Zhang C, Du T, et al. Progress of exosomes in the diagnosis and treatment of lung cancer. *Biomed Pharmacother* 2021; 134: 111111.
6. El Aboudi N and Benhlima L. Big data management for healthcare systems: architecture, requirements, and implementation. *Adv Bioinformatics* 2018; 2018: 4059018.
7. Toumazis I, Bastani M, Han SS, et al. Risk-based lung cancer screening: a systematic review. *Lung Cancer* 2020; 147: 154–186.
8. Bartlett EC, Silva M, Callister ME, et al. False-negative results in lung cancer screening-evidence and controversies. *J Thorac Oncol* 2021; 16: 912–921.
9. Sato A, Hamada S, Urashima Y, et al. The effect of false-positive results on subsequent participation in chest X-ray screening for lung cancer. *J Epidemiol* 2016; 26: 646–653.
10. Liu Q, Tan C, Yi L, et al. Cost-effectiveness analysis of pembrolizumab plus chemotherapy as first-line therapy for extensive-stage small-cell lung cancer. *PLoS ONE* 2021; 16: e0258605.
11. Holman A, Kross E, Crothers K, et al. Patient perspectives on longitudinal adherence to lung cancer screening. *Chest* 2022; 162: 230–241.
12. Kann BH, Hosny A and Aerts H. Artificial intelligence for clinical oncology. *Cancer Cell* 2021; 39: 916–927.
13. Hamet P and Tremblay J. Artificial intelligence in medicine. *Metab: Clin Exp* 2017; 69s: S36–s40.
14. Garcia-Vidal C, Sanjuan G, Puerta-Alcalde P, et al. Artificial intelligence to support clinical decision-making processes. *EBioMed* 2019; 46: 27–29.
15. Hosny A, Parmar C, Quackenbush J, et al. Artificial intelligence in radiology. *Nat Rev Cancer* 2018; 18: 500–510.
16. Dong S, Wang P and Abbas K. A survey on deep learning and its applications. *Comput Sci Rev* 2021; 40: 22.
17. Esteva A, Robicquet A, Ramsundar B, et al. A guide to deep learning in healthcare. *Nat Med* 2019; 25: 24–29.

18. Bera K, Braman N, Gupta A, et al. Predicting cancer outcomes with radiomics and artificial intelligence in radiology. *Nat Rev Clin Oncol* 2022; 19: 132–146.

19. Gillies RJ and Schabath MB. Radiomics improves cancer screening and early detection. *Cancer Epidemiol Biomark Prev: A Publicat Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol* 2020; 29: 2556–2567.

20. Tunali I, Gillies RJ and Schabath MB. Application of radiomics and artificial intelligence for lung cancer precision medicine. *Cold Spring Harbor Perspect Med* 2021; 11: a039537.

21. Li Y, Chen D, Wu X, et al. A narrative review of artificial intelligence-assisted histopathologic diagnosis and decision-making for non-small cell lung cancer: achievements and limitations. *J Thorac Dis* 2021; 13: 7006–7020.

22. Acs B, Rantalainen M and Hartman J. Artificial intelligence as the next step towards precision pathology. *J Intern Med* 2020; 288: 62–81.

23. Tran BX, Latkin CA, Vu GT, et al. The current research landscape of the application of artificial intelligence in managing cerebrovascular and heart diseases: a bibliometric and content analysis. *Int J Environ Res Public Health* 2019; 16: 2699.

24. Xiong J, Qi W, Liu J, et al. Research progress of ferroptosis: a bibliometrics and visual analysis study. *J Healthc Eng* 2021; 2021: 2178281.

25. Liu X, Zhao S, Tan L, et al. Frontier and hot topics in electrochemiluminescence sensing technology based on CiteSpace bibliometric analysis. *Biosens Bioelectron* 2022; 201: 113932.

26. Huang L, Shi X, Zhang N, et al. Bibliometric analysis of trends and issues in traditional medicine for stroke research: 2004-2018. *BMC Complement Med Ther* 2020; 20: 39.

27. Yu Y, Li Y, Zhang Z, et al. A bibliometric analysis using VOSviewer of publications on COVID-19. *Ann Transl Med* 2020; 8: 816.

28. Alam BF, Nayab T, Ali S, et al. Current scientific research trends on salivary biomarkers: a bibliometric analysis. *Diagnostics* 2022; 12: 1171.

29. Li T, Zeng Y, Fan X, et al. A bibliometric analysis of research articles on midwifery based on the web of science. *J Multidiscip Healthc* 2023; 16: 677–692.

30. Bianciardi L, Rinaldi M, Depaula U, et al. A decision support in lung-cancer diagnosis and therapy—lung-cancer expert system luca. *J Exp Clin Cancer Res* 1992; 11: 153–160.

31. Lim M, Carollo A, Neoh MJY, et al. Mapping miRNA research in schizophrenia: a scientometric review. *Int J Mol Sci* 2022; 24: 436.

32. Van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* 2017; 77: e104–e107.

33. Tajbakhsh N, Shin JY, Gurudu SR, et al. Convolutional neural networks for medical image analysis: full training or fine tuning? *IEEE Trans Med Imaging* 2016; 35: 1299–1312.

34. Coudray N, Ocampo PS, Sakellaropoulos T, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med* 2018; 24: 1559–1567.

35. Ardila D, Kiraly AP, Bharadwaj S, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med* 2019; 25: 954–961.

36. 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.

37. Setio AA, Ciompi F, Litjens G, et al. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. *IEEE Trans Med Imaging* 2016; 35: 1160–1169.

38. Mangels AR, Holden JM, Beecher GR, et al. Carotenoid content of fruits and vegetables: an evaluation of analytic data. *J Am Diet Assoc* 1993; 93: 284–296.

39. Rajpurkar P, Irvin J, Ball RL, et al. Deep learning for chest radiograph diagnosis: a retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Med* 2018; 15: e1002686.

40. Alom MZ, Yakopcic C, Hasan M, et al. Recurrent residual U-Net for medical image segmentation. *J Med Imaging* 2019; 6: 014006.

41. Setio AAA, Traverso A, de Bel T, et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge. *Med Image Anal* 2017; 42: 1–13.

42. Yuan M, Lin H, Wu H, et al. Community engagement in public health: a bibliometric mapping of global research. *Arch Public Health* 2021; 79: 6.

43. Chen D, Wu Y, Li H, et al. Treatment on patients with spastic cerebral palsy in the past 30 years: a systematic review and bibliometric analysis. *Medicine* 2022; 101: e30535.

44. Ewals LJS, van der Wulp K, van den Borne B, et al. The effects of artificial intelligence assistance on the radiologists' assessment of lung nodules on CT scans: a systematic review. *J Clin Med* 2023; 12: 3536.

45. Pei Q, Luo Y, Chen Y, et al. Artificial intelligence in clinical applications for lung cancer: diagnosis, treatment and prognosis. *Clin Chem Lab Med* 2022; 60: 1974–1983.

46. Xu L, Zhou H, Wang G, et al. The prognostic influence of histological subtypes of micropapillary tumors on patients with lung adenocarcinoma ≤ 2 cm. *Front Oncol* 2022; 12: 954317.

47. Chen ZH, Lin L, Wu CF, et al. Artificial intelligence for assisting cancer diagnosis and treatment in the era of precision medicine. *Cancer Commun* 2021; 41: 1100–1115.

48. Shankar A, Saini D, Dubey A, et al. Feasibility of lung cancer screening in developing countries: challenges, opportunities and way forward. *Transl Lung Cancer Res* 2019; 8: S106–S121.

49. Jiwnani S, Penumadu P, Ashok A, et al. Lung cancer management in low and middle-income countries. *Thorac Surg Clin* 2022; 32: 383–395.