



## OPEN ACCESS

## EDITED BY

Wenyi Jin,  
City University of Hong Kong,  
Hong Kong SAR, China

## REVIEWED BY

Yinhui Yao,  
Affiliated Hospital of Chengde Medical  
University, China  
Kui Wang,  
Shanghai Jiao Tong University, China

## \*CORRESPONDENCE

Hongcan Shi  
✉ shihongcan@163.com

†These authors have contributed equally to  
this work and share first authorship

RECEIVED 13 May 2024

ACCEPTED 19 June 2024

PUBLISHED 08 July 2024

## CITATION

Yuan L, Shen Z, Shan Y, Zhu J, Wang Q, Lu Y  
and Shi H (2024) Unveiling the landscape of  
pathomics in personalized immunotherapy  
for lung cancer: a bibliometric analysis.  
*Front. Oncol.* 14:1432212.  
doi: 10.3389/fonc.2024.1432212

## COPYRIGHT

© 2024 Yuan, Shen, Shan, Zhu, Wang, Lu and  
Shi. This is an open-access article distributed  
under the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other forums  
is permitted, provided the original author(s)  
and the copyright owner(s) are credited and  
that the original publication in this journal is  
cited, in accordance with accepted academic  
practice. No use, distribution or reproduction  
is permitted which does not comply with  
these terms.

# Unveiling the landscape of pathomics in personalized immunotherapy for lung cancer: a bibliometric analysis

Lei Yuan<sup>1,2,3†</sup>, Zhiming Shen<sup>1,2,3†</sup>, Yibo Shan<sup>1,2,3†</sup>, Jianwei Zhu<sup>1,2,3</sup>,  
Qi Wang<sup>1,2,3</sup>, Yi Lu<sup>1,2,3</sup> and Hongcan Shi<sup>1,2,3\*</sup>

<sup>1</sup>Department of Thoracic Surgery, Northern Jiangsu People's Hospital Affiliated to Yangzhou University, Yangzhou, China, <sup>2</sup>Institute of Translational Medicine, Medical College, Yangzhou University, Yangzhou, China, <sup>3</sup>Jiangsu Key Laboratory of Integrated Traditional Chinese and Western Medicine for Prevention and Treatment of Senile Diseases, Yangzhou University, Yangzhou, China

**Background:** Pathomics has emerged as a promising biomarker that could facilitate personalized immunotherapy in lung cancer. It is essential to elucidate the global research trends and emerging prospects in this domain.

**Methods:** The annual distribution, journals, authors, countries, institutions, and keywords of articles published between 2018 and 2023 were visualized and analyzed using CiteSpace and other bibliometric tools.

**Results:** A total of 109 relevant articles or reviews were included, demonstrating an overall upward trend; The terms "deep learning", "tumor microenvironment", "biomarkers", "image analysis", "immunotherapy", and "survival prediction", etc. are hot keywords in this field.

**Conclusion:** In future research endeavors, advanced methodologies involving artificial intelligence and pathomics will be deployed for the digital analysis of tumor tissues and the tumor microenvironment in lung cancer patients, leveraging histopathological tissue sections. Through the integration of comprehensive multi-omics data, this strategy aims to enhance the depth of assessment, characterization, and understanding of the tumor microenvironment, thereby elucidating a broader spectrum of tumor features. Consequently, the development of a multimodal fusion model will ensue, enabling precise evaluation of personalized immunotherapy efficacy and prognosis for lung cancer patients, potentially establishing a pivotal frontier in this domain of investigation.

## KEYWORDS

lung cancer, pathomics, artificial intelligence, deep learning, tumor microenvironment, immunotherapy

## 1 Introduction

Lung cancer remains one of the most prevalent malignancies and represents the foremost cause of cancer-related mortality worldwide (1, 2), the majority of lung cancers (80–90%) manifest as non-small cell lung cancer (NSCLC), often diagnosed at an advanced stage (65%), potentially with concurrent local or distant metastasis (3). Recent advances in immunotherapy, particularly the use of immune checkpoint inhibitors (ICIs), have shown promising outcomes in enhancing the prognosis of lung cancer patients (4). Nevertheless, not all patients experience the benefits of immunotherapy, highlighting the need for additional research into predictive biomarkers of immune response. These biomarkers, which may include substances, structures, or products of processes within the body, have the potential to facilitate personalized immunotherapy by enabling the monitoring of immune reactions.

Each lung cancer patient undergoes histopathological diagnosis, involving the preparation of biopsy tissues into pathological slides for examination. The traditional preservation method of using wax embedding techniques for pathological slides can now be digitized through computerization, archiving them as digital pathology images. This technological advancement serves as a foundation for applying big data analytics to digital pathology images. Consequently, the field of pathomics has emerged (5). Pathomics entails applying machine learning techniques to extract large-scale, objectively quantifiable, and readily analyzable datasets from digitally scanned pathological tissue images. Consistent with the pathological diagnostic requirements of diseases, morphological features, including size and shape of pathological images, along with multi-dimensional subtle features reflecting potential biological characteristics such as texture features and edge gradient features, are extracted. These features can be utilized for quantitative disease screening, diagnosis, prognosis prediction, and other applications (6).

In this study, CiteSpace (7) was utilized for the inaugural analysis of hotspots and trends in the application of pathomics in lung cancer. The objective is to provide valuable insights for scholars involved in research within this domain.

## 2 Materials and methods

### 2.1 Data collection

Web of Science Core Collection (WoSCC) database was chosen as the literature retrieval platform. The retrieval period spanned from 2018 to 2023, with the final search conducted on October 20, 2023. Subject terms were exclusively employed as the search method, and the search formula was: TS= (“Pathomics” OR “Pathomics” OR “Digital Pathology” OR “Whole-slide Imaging” OR “Whole Slide Imaging” OR “Computational Pathology”) AND TS=(“Lung Cancer” OR “Pulmonary Cancer” OR “Carcinoma of Lung” OR “Pulmonary Carcinoma” OR “Cancer of Lung” OR “Bronchogenic Carcinoma” OR “Bronchogenic” OR “Cancer of the Lung” OR “NSCLC” OR “SLC”), document type: Articles or Review Articles; a total of 109 documents were retrieved.

### 2.2 Statistical methods

Export the complete records and referenced bibliographies of the 109 documents retrieved from WoSCC in Text format, comprising 85 articles and 24 reviews. Conduct a comprehensive analysis of the literature using CiteSpace 6.2.R4 (64-bit) Basic, focusing on the country, institution, authorship, keywords, and cited references. The bibliometric online analysis platform, developed by the National Science Library of the Chinese Academy of Sciences, was employed to conduct a visual analysis of historical keywords and national collaborations.

## 3 Results

### 3.1 Annual publication volume in WoSCC

A total of 109 matching documents were retrieved, and the overall publication output exhibited a general upward trend, especially reaching a contribution rate of 26.61% in 2021 (Figure 1). The annual average publication output is approximately 21.8 articles. The results indicate a gradual increase in the attention to pathomics research in the context of lung cancer.

### 3.2 Distribution of source journals

The literature selected from the 109 studies on pathomics in the management of lung cancer has been indexed by 146 journals. For the top 10 journals in terms of publication output, detailed information on Journal Citation Reports (JCR) category, publication quantity, impact factor (IF), and their respective contribution percentages is provided in Table 1.

### 3.3 Visualization of collaborations between countries and institutions

Running the CiteSpace software for country analysis resulted in a knowledge graph with 35 nodes and 80 edges (Figure 2). Each

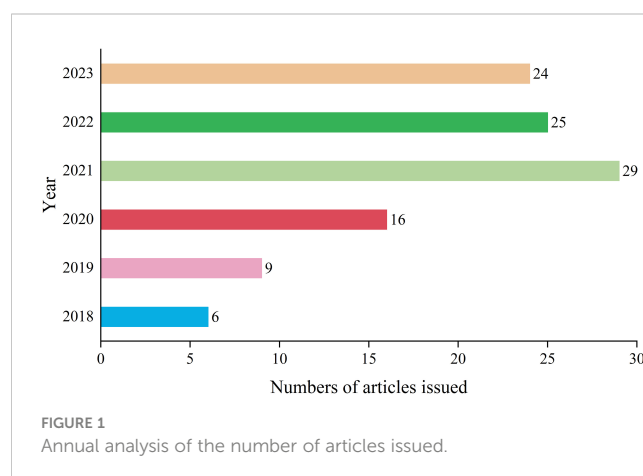


TABLE 1 Top 10 journals in terms of publication volume.

Journal Titles	JCR	Number	IF	Rate%
CANCERS	Q1	11	5.2	10.092
MODERN PATHOLOGY	Q1	6	7.5	5.505
EBIOMEDICINE	Q1	3	11.1	2.752
FRONTIERS IN ONCOLOGY	Q2	3	4.7	2.752
HISTOPATHOLOGY	Q1	3	6.4	2.752
IEEE ACCESS	Q2	3	3.9	2.752
MEDICAL IMAGE ANALYSIS	Q1	3	10.9	2.752
BIOINFORMATICS	Q1	2	5.8	1.835
COMPUTERS IN BIOLOGY AND MEDICINE	Q1	2	7.7	1.835
IEEE TRANSACTIONS ON MEDICAL IMAGING	Q1	2	10.6	1.835

circular node represents a country, with the size indicating the quantity of publications from that country. The connections between nodes represent collaborative relationships between countries, with the thickness of the connections reflecting the degree of collaboration. Different colors of nodes represent different time periods (8), the size of the purple circles reflects the centrality values indicating the influence of each country.

Leveraging the bibliometric online analysis platform, Figure 3 depicts the contributions of different countries in the field. Distinctly colored blocks represent the proportional contribution of each country. Table 2 presents the top 5 institutions in terms of publication output.

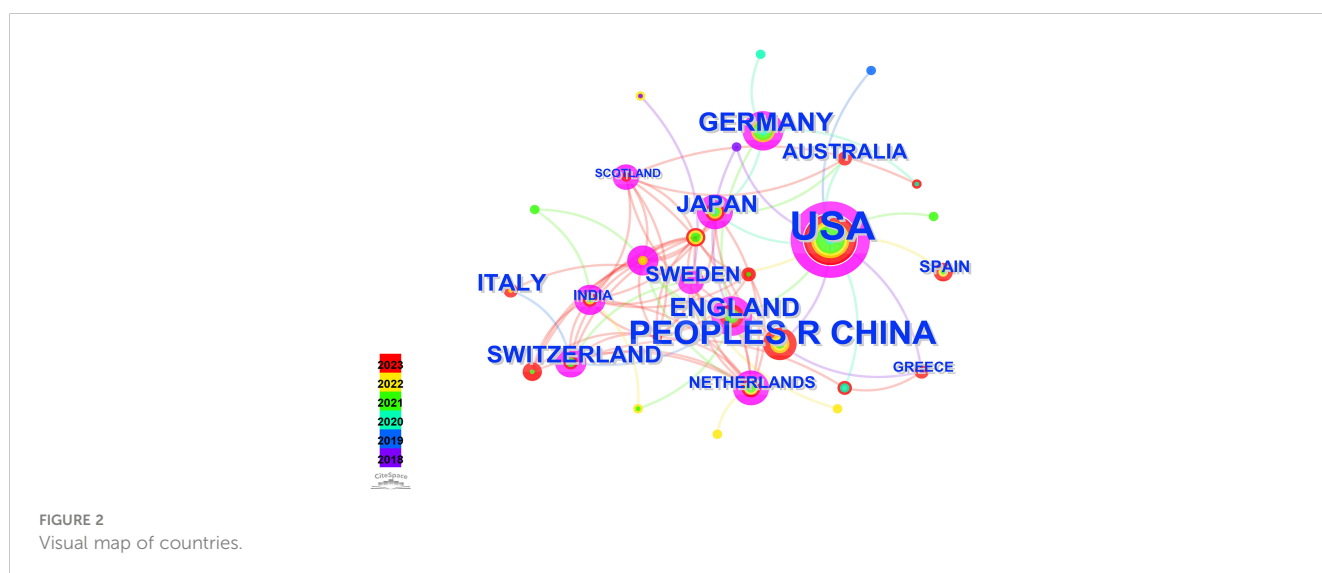
### 3.4 Visualization of author collaborations

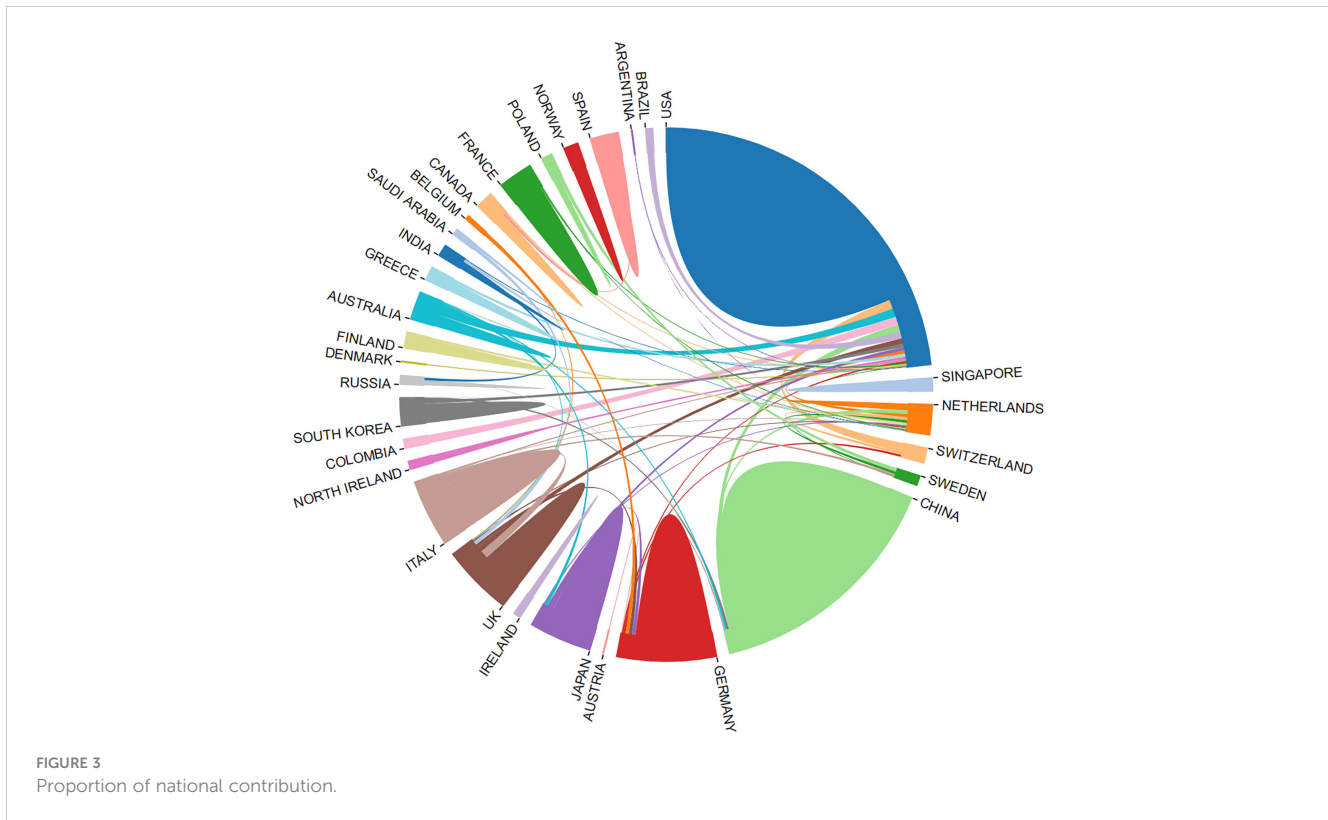
Running the CiteSpace software, author analysis resulted in a knowledge graph with 200 nodes and 383 edges (Figure 4). Each

circular node represents an author, and the connections between nodes represent collaborative relationships between authors. The thickness of the connections reflects the degree of collaboration. Different colors of nodes represent different time periods. Conducting a co-occurrence analysis on the author team collaboration network based on the literature retrieved from WoSCC, Table 3 is presented, listing the top 5 authors in terms of publication output along with their affiliated institutions in this research field.

### 3.5 Co-occurrence analysis of keywords

Keyword-related analysis, as manifested in the visualization of co-occurrence patterns, is crucial for delineating the research hotspots and frontiers within a given domain. Running the CiteSpace software with author keywords as node types, a co-occurrence network of keywords with 159 nodes and 334 edges





was generated (Figure 5). After removing redundant terms that overlap with the search strategy, an analysis of the co-occurrence frequency and centrality values of keywords in this field (Table 4) reveals that the prominent keywords include: deep learning, artificial intelligence (AI), computer-aided diagnosis, tumor microenvironment, feature extraction, image analysis, tumor mutation burden, survival prediction, markov random field, mixture model. Furthermore, Figure 6 illustrates the temporal frequency changes of different keywords over time. It highlights the research focal points in the past few years related to the application of AI-based pathomics in the diagnosis and treatment of lung cancer. These themes reflect the proactive role of pathomics in aiding diagnosis, classification, predicting treatment efficacy, risk assessment, exploring emerging biomarkers, and analyzing gene

expression levels in the context of lung cancer diagnosis and treatment.

### 3.6 Keyword cluster analysis

Keyword cluster analysis involves utilizing the log-likelihood rate (LLR) method to analyze the connection relationships among significant keyword nodes. This method reflects the hot topics within the research domain, with closely connected keywords in a cluster indicating higher research intensity. Larger node values within a cluster signify greater research interest. By examining these clusters, it is possible to predict the developmental patterns and emerging trends in the research field (9).

According to the keyword cluster analysis (Figure 7), researchers' studies are concentrated in the following 10 key areas: #0 parameter auto-tuning; #1 concordance study; #2 prognostic and predictive; #3 mixture model; #4 lung cancer slide cells; #5 non-small-cell lung cancer; #6 immunotherapy; #7 deep-learning microscopy; #8 telepathology; #9 radiology. By employing the clustering algorithm within CiteSpace software to organize title terms and visualize them (Figure 8), a clear sequential pattern emerges, encompassing: #0 spatial quantitative systems pharmacology platform spqsp-io; #1 adaptive radiotherapy; #2 patient survival; #3 pd-11 expression; #4 digital analysis; #5 Bayesian hidden Potts mixture model; #6 bayesian collaborative learning; #7 multi-stained feature matching; #8 oncology; #9 pathomics; Utilizing the clustering algorithm in CiteSpace to group subject categories and create a visual representation (Figure 9), a sequential progression of clusters is discernible,

TABLE 2 Top 5 institutions in terms of publication volume.

Rank	Number	Institution	Country
1	8	CASE WESTERN RESERVE UNIVERSITY	USA
2	8	UNIVERSITY OF TEXAS SYSTEM	USA
3	6	LOUIS STOKES CLEVELAND VETERANS AFFAIRS MEDICAL CENTER	USA
4	6	UNIVERSITY OF TEXAS SOUTHWESTERN MEDICAL CENTER DALLAS	USA
5	6	US DEPARTMENT OF VETERANS AFFAIRS	USA

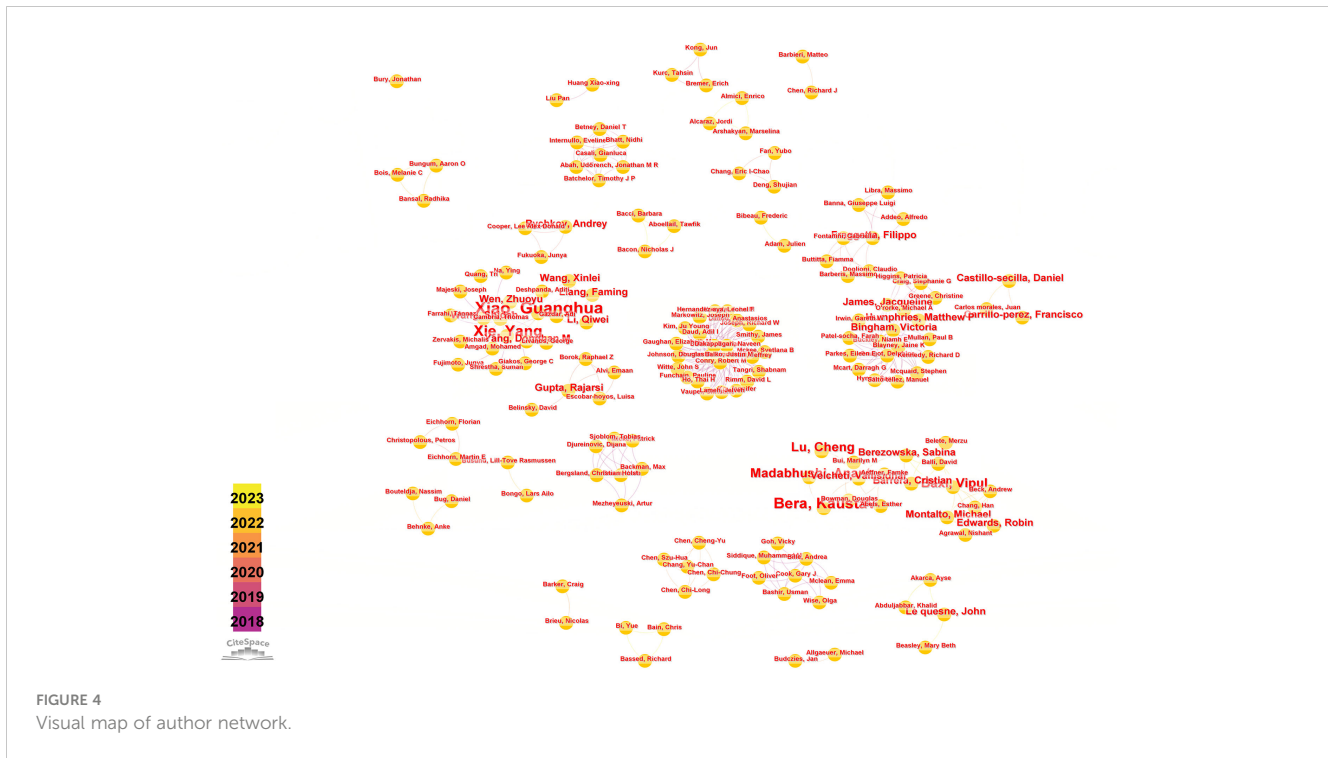


FIGURE 4  
Visual map of author network.

including: #0 Pathology; #1 Mathematics; #2 Medicine, Research & Experimental; #3 Computer Science, Theory & Methods; #4 Engineering, Multidisciplinary; #5 Statistics & Probability; #6 Imaging Science & Photographic Technology; #7 Biology; #8 Health Care Sciences Services; #9 Cell Biology. Employing the clustering algorithm for keywords and generating a graphical display (Figure 10), a sequential evolution of clusters is evident, incorporating: #0 digital pathology; #1 machine learning; #2 deep learning; #3 artificial intelligence; #4 lung cancer; #5 mixture model; #6 computational pathology; #7 scale invariant feature transform; #8 equity; #9 cancer immunopathology.

Each section is divided into 10 clustering modules, partial clustering blocks overlap with each other, suggesting close connections between these research areas. In addition to the research retrieval terms, other clusters demonstrate that pathomics in lung cancer research spans various fields, including medical experimental research, computer science, cell biology, statistics, and mathematics. Through advanced methods such as

AI and machine learning, pathomics involves in-depth digital analysis of tumor tissues and the tumor microenvironment based on patients' pathological tissue sections. It aims to construct hybrid models, identify a multitude of pathological features, conduct precise assessments, and predict tumor-related indicators, including programmed death-ligand 1 Tumor cell Proportion Score (PD-L1 TPS). The goal is to assist in personalized diagnosis and treatment for patients and contribute to clinical decision-making by leveraging the synergies between AI and clinical medicine.

### 3.7 Cited references

A total of 426 relevant articles were retrieved from WoSCC, accumulating a total of 10,174 citations. The average number of citations per article is 24. The top 10 most cited articles are listed in Table 5.

TABLE 3 Top 5 authors in terms of publication volume.

Rank	Author	Institution	Country	Number
1	Xiao, Guanghua	University of Texas Southwestern Medical Center	USA	5
2	Xie, Yang	University of Texas Southwestern Medical Center Clinic Science	USA	4
3	Bera, Kaustav	CASE WESTERN RESERVE University	USA	4
4	Baxi, Vipul	BRISTOL MYERS SQUIBB	USA	3
5	Lu, Cheng	University of ALBERTA	Canada	3

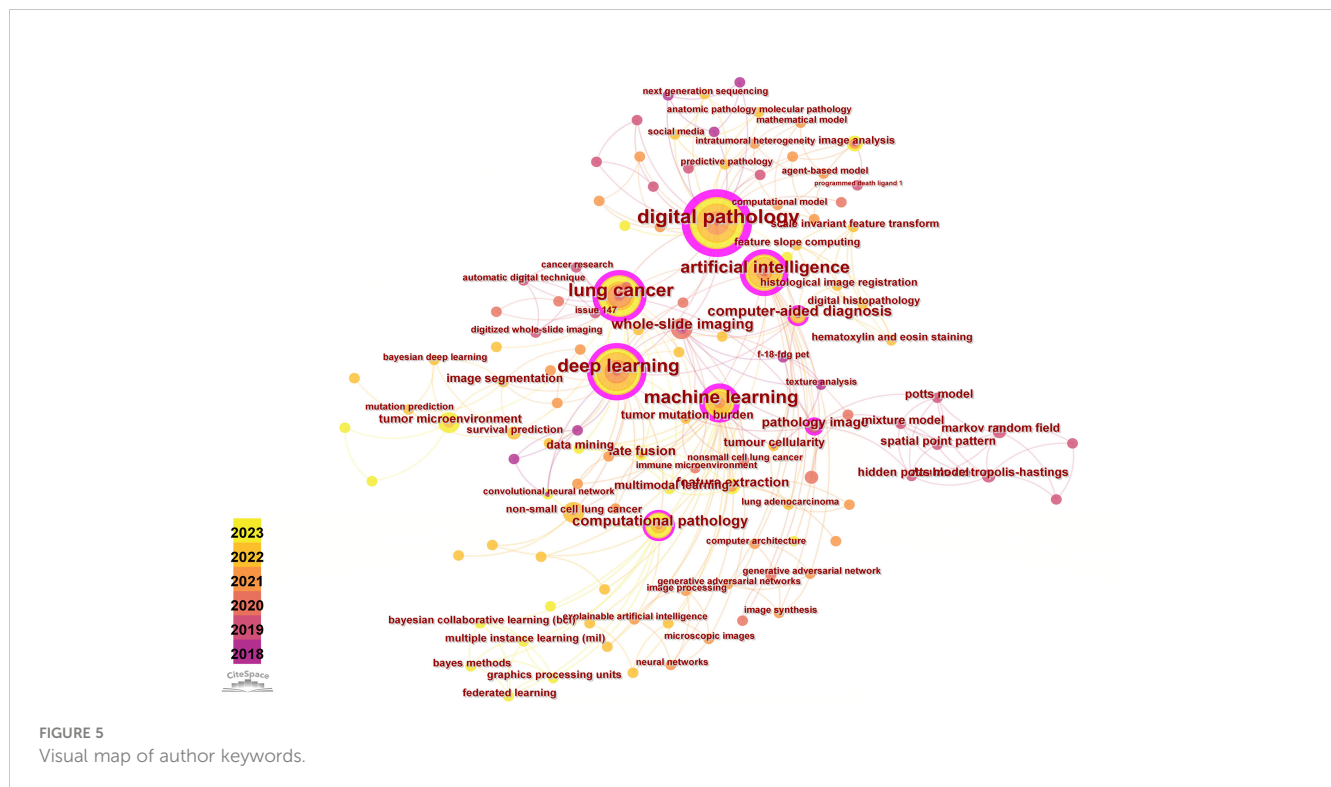


FIGURE 5  
Visual map of author keywords.

## 4 Discussion

Pathomics is an innovative interdisciplinary field that combines digital pathology and AI. The rise of digital pathology has enabled the scanning of whole tissue slides, based on the fundamental principle of digitizing whole-slide images (WSI) using state-of-the-art whole-slide scanners. This technology can convert standard Hematoxylin-Eosin (H&E) staining glass slides into a digital format (WSI) (20). This allows for detailed spatial exploration of the entire tumor heterogeneity and its most invasive elements. It automatically extracts and classifies histological features, transforming this information into binary data. Finally, the

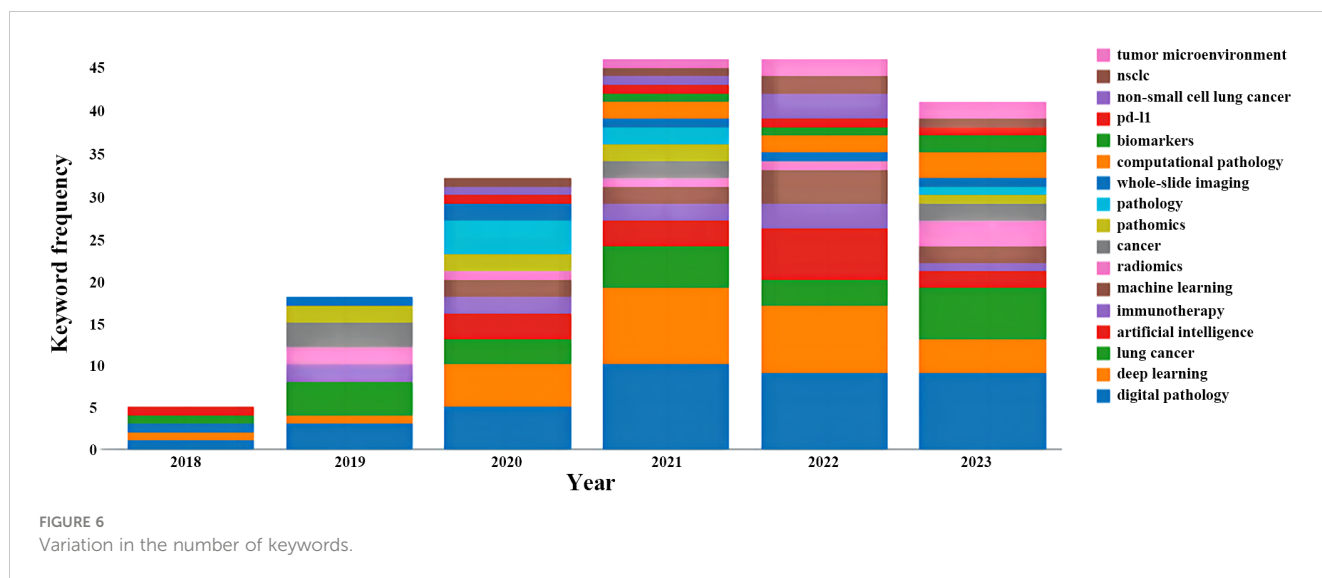
extracted features are processed through sophisticated computer algorithms to perform tasks such as cancer classification and outcome prediction (21). Computational analysis of digitized histological slides through pathomics can extract valuable information. Some research primarily focuses on predicting the prognosis of lung cancer (22), including improving clinical decisions for cancer immunotherapy and exploring biomarkers related to potential benefits from ICIs, such as microsatellite instability (MSI), PD-L1 TPS, and inflammatory genes, among others (23). Another significant research area involves the integration of pathomics with multiple omics disciplines to explore the classification of lung cancer and other related aspects. Alvarez-Jimenez C et al. demonstrated the potential existence of cross-scale correlations between pathomics and CT imaging, which could be used to identify relevant imaging and histopathological features (24).

The escalating demand for personalized cancer treatment necessitates more precise biomarker assessments and quantitative tissue pathology for accurate cancer diagnosis. Pathologists must be equipped with new methodologies and tools to enhance diagnostic sensitivity and specificity, ultimately contributing to more informed and improved treatment decisions (13). Recently, significant success has been achieved in the analysis of medical images using AI due to the rapid advancement of “deep learning” algorithms (16).

Recent breakthroughs in AI hold the promise of significantly changing the way we diagnose and stratify cancer in pathology. Deep learning technology represents a milestone in this transformation, with numerous deep learning architectures applied to pathology-focused research. Various modeling objectives have been pursued, and recent studies demonstrate the application of deep learning in pathology aiming to predict

TABLE 4 High frequency and centrality keywords.

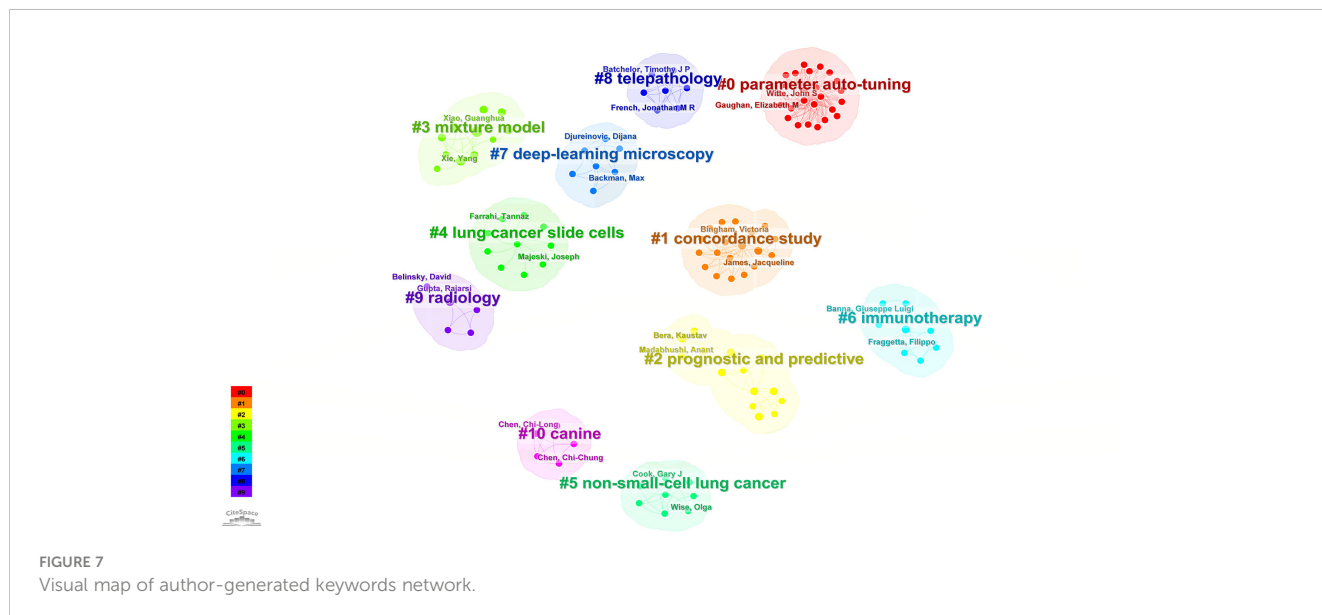
Rank	Keywords	Frequency	Centrality
1	deep learning	27	0.33
2	artificial intelligence	14	0.29
3	machine learning	10	0.22
4	computer-aided diagnosis	3	0.11
5	tumor microenvironment	4	0.07
6	feature extraction	3	0.04
7	image analysis	3	0.04
8	late fusion	2	0.02
9	tumor mutation burden	2	0.02
10	survival prediction	2	0.02

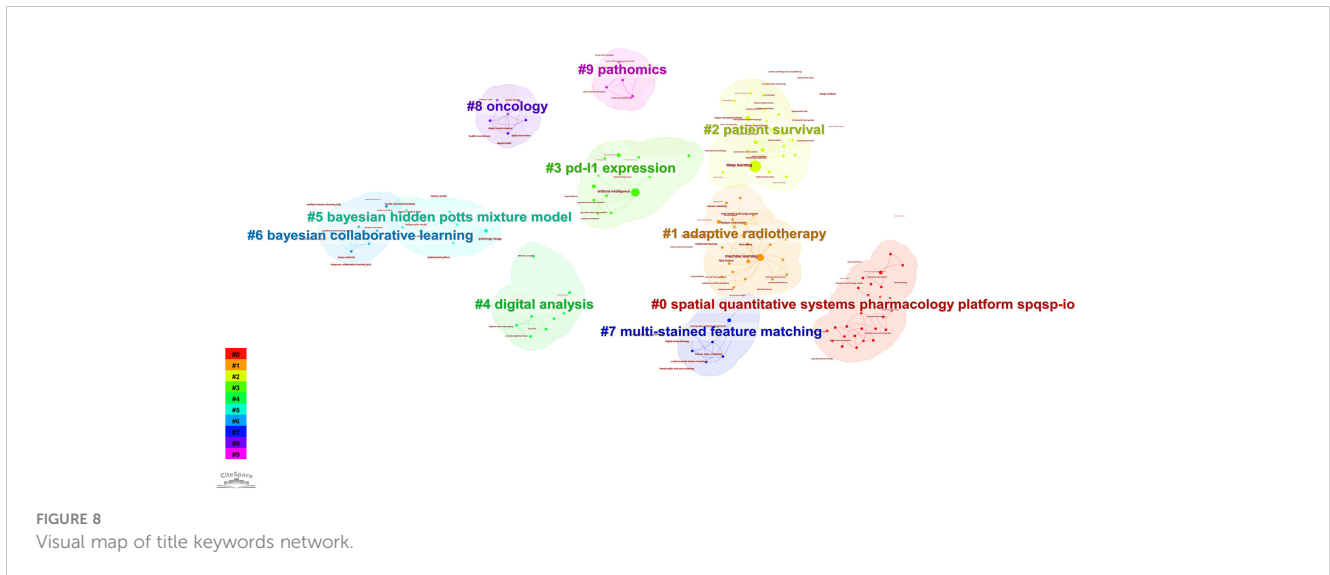


conventional diagnostic features used in pathology practice (such as distinguishing between diseases and normal tissues, defining tumor grades, and differentiating cancer types), leading to new insights into diseases (25, 26).

Deep learning encompasses various types of deep neural networks, and its application has achieved several breakthroughs in addressing current key challenges in pathology (27). Convolutional Neural Networks (CNN) are the most commonly used in the analysis of pathological images (28, 29). A standard CNN consists of an input layer, task-specific output layer, and multiple hidden layers. Each hidden layer is composed of numerous convolutional filters (parameters), which apply the same local transformation at different positions in their input images (30). Due to the shared parameters when applied locally in the image, effective parameterization of the CNN model is achieved. The typical implementation of CNN models offers a degree of translation invariance, allowing detected objects or patterns to

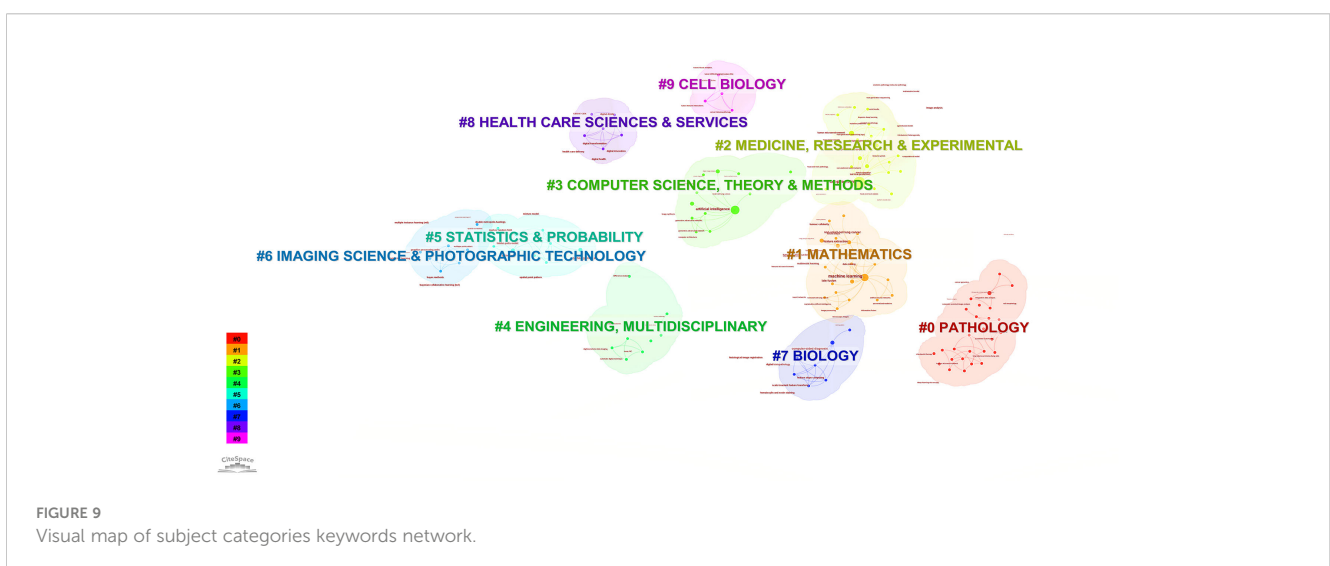
appear at any position within the image. Pooling layers are often included between convolutional layers to down-sample the intermediate outputs (feature maps) of the convolution function. Following the convolutional layers are fully connected layers, which flatten the output of the convolutional layers and generate the final representation for the input-output layers (30, 31). Each neuron in a CNN calculates its output by applying a weight vector and bias (parameters) to the input values from the previous layer. The optimization (training) of the CNN model involves iteratively adjusting these biases and weights to minimize the loss function. One advantage of CNNs over other image classification algorithms is their suitability for end-to-end learning (32). Another major advantage of CNNs is their flexibility and efficiency in learning patterns from image data. Currently, they represent state-of-the-art technology in the field of image analysis and classification, consistently outperforming earlier generations of image analysis methods (29, 32). Kao Y-S et al. conducted a study on the



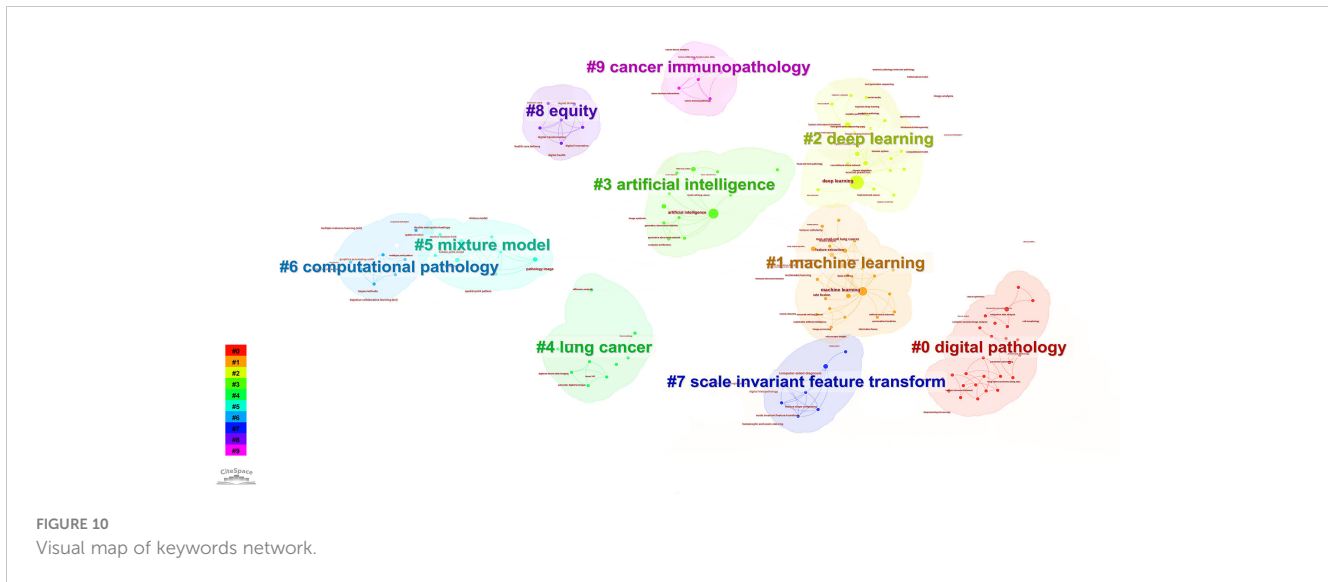


application of deep learning technology in histopathological tissue slices (deep pathomics) with the aim of predicting the response of stage III NSCLC to treatment (33). They assessed 35 digitalized tissue slices (biopsy or surgical specimens) from patients with stage IIIA or IIIB NSCLC. Based on the reduction in target volume observed in weekly CT scans during chemoradiotherapy, patients were categorized as responders (12/35, 34.7%) and non-responders (23/35, 65.7%). Employing a leave-two-out cross-validation method, they tested the digital tissue slices using 5 pre-trained CNNs-AlexNet, VGG, MobileNet, GoogLeNet, and ResNet, and evaluated the network performance. GoogLeNet was identified as the most effective CNN, accurately classifying 8/12 responders and 10/11 non-responders. Furthermore, deep pathomics exhibited a high level of specificity (True Negative Rate: 90.1) and considerable sensitivity (True Positive Rate: 0.75). Their data suggests that AI can surpass the capabilities of current diagnostic systems, providing additional insights beyond what is currently attainable in clinical practice.

Furthermore, there are studies attempting to apply AI to histological images with the aim of discovering novel image-based prognostic and predictive biomarkers. Cao R et al. proposed a deep learning model based on histopathological images to predict microsatellite status, achieving area under curve (AUC) of 0.88 and 0.85, respectively. It is noteworthy that this model can identify five distinct pathological imaging features, which are associated with the mutation burden in the genome, DNA damage repair-related genotypes, and the anti-tumor immune activation pathway in the transcriptome. The predictive model provides the potential for multi-omics correlations through interpretability associated with pathology, genomics, and transcriptomics phenotypes (34). Wang X et al. developed a system capable of identifying high-risk recurrence in early-stage NSCLC patients with an accuracy ranging from 75% to 82% (22). In another study, Wang S et al. characterized a group of high-risk NSCLC patients and identified image-based tumor shape features as an independent prognostic factor (35). Rakae M et al. developed a machine learning-based scoring system







for tumor-infiltrating lymphocytes (TILs) to predict the response of NSCLC to immune checkpoint inhibitor therapy (36). Additionally, Coudray N, Ocampo PS et al. applied AI to digital pathology slides to predict the presence of mutations in lung adenocarcinoma (37). In summary, the development of these advanced deep learning algorithms enhances the capability of analyzing lung cancer pathology images, assisting pathologists in challenging diagnostic tasks such as tumor identification, metastasis detection, and analysis of the tumor microenvironment.

TME is primarily composed of tumor cells, lymphocytes, stromal cells, macrophages, blood vessels, and other components. The composition of the TME varies based on the relative

proportions of its different constituents, and its presence plays a crucial role in the growth and invasion of tumors.

Immune cells within the TME exhibit dual functions – on one hand, they identify and destroy tumor cells, while on the other hand, they also promote tumor growth and metastasis (38, 39). For instance, immune cells, including T cells, B cells, macrophages, and myeloid-derived suppressor cells, possess the ability to modulate the TME, thereby influencing tumor metastasis and pathological features (40, 41). Tumor Infiltrating Lymphocytes (TILs) in the TME involves a complex network of multiple cell types and cytokines and is a hallmark of immune recognition. Numerous studies have shown that activated CD8<sup>+</sup> T cells are the major players

TABLE 5 The top 10 cited articles.

Rank	Author	Year	Total Citations	Title
1	Pucci, Carlotta;	2019	343	Innovative approaches for cancer treatment: current perspectives and new challenges (10)
2	Lu, Ming Y;	2021	278	Data-efficient and weakly supervised computational pathology on whole-slide images (11)
3	Khosravi, Pegah;	2018	180	Deep Convolutional Neural Networks Enable Discrimination of Heterogeneous Digital Pathology Images (12)
4	Acs, B.;	2020	136	Artificial intelligence as the next step toward precision pathology (13)
5	Maibach, Fabienne;	2020	118	Tumor-Infiltrating Lymphocytes and Their Prognostic Value in Cutaneous Melanoma (14)
6	Mezheyeuski, Artur;	2018	114	Multispectral imaging for quantitative and compartment-specific immune infiltrates reveals distinct immune profiles that classify lung cancer patients (15)
7	Wang, Shidan;	2019	103	Artificial Intelligence in Lung Cancer Pathology Image Analysis (16)
8	Johnson, Douglas B.;	2018	91	Quantitative Spatial Profiling of PD-1/PD-L1 Interaction and HLA-DR/IDO-1 Predicts Improved Outcomes of Anti-PD-1 Therapies in Metastatic Melanoma (17)
9	Baxi, Vipul;	2022	80	Digital pathology and artificial intelligence in translational medicine and clinical practice (18)
10	Saw, Stephanie P. L.;	2021	56	Revisiting neoadjuvant therapy in non-small-cell lung cancer (19)

involved in anti-tumor immunity, and in a subset of tumors, cancer cells inhibit the activation of CD8<sup>+</sup> cytotoxic T cells through the expression of ligands such as PD-L1 that bind to inhibitory checkpoints, which has been suggested to be an important mechanism of immune escape for cancer cells (42). The expression of PD-L1 on TME immune cells, including myeloid cells (macrophages, dendritic cells) and T cells, appears to correlate more with the ICI response than expression on tumor cells. However, in NSCLC clinical practice, a limitation in histologically characterizing T lymphocyte infiltration is the scarcity of tumor tissue, which has hampered insight into the role of T lymphocytes in influencing the ICI response (43). Tumor-associated macrophages can promote angiogenesis and invasion by secreting cytokines, growth factors, and proteases (44). Cancer-associated fibroblasts (CAF) are pivotal in the formation of organs and the maintenance of tissue structure and function. They also play a significant role in tumor initiation, progression, metastasis, and the development of drug resistance through their potent immunosuppressive capabilities. Activated CAF possess the capability to secrete various substances, including extracellular matrix and vascular endothelial growth factor (VEGF), contributing to the complexity of the TME (45, 46). The markers associated with CAF are predominantly linked to T cell immunosuppression, inhibiting the functions of CD8<sup>+</sup> T cells and natural killer cells, particularly by secreting various chemokines and cytokines, notably interleukin-6 (IL-6), which leads to suboptimal clinical treatment outcomes. As research into the effects of CAF and the TME on immune cells and the efficacy of cancer immunotherapy advances, scientists can potentially develop novel compounds targeting these mechanisms, thereby offering innovative strategies for immunotherapy (47). It is noteworthy that research indicates a significant impact of the TME on the survival benefits of immunotherapy (48). The presence of immune cells in the TME, including the percentage of CD8<sup>+</sup> T cells, can serve as a predictive factor for the effectiveness of immunotherapy (49). The extracellular matrix can influence the mechanisms of tumorigenesis by affecting cell growth, metastasis, and immune evasion through the activation of signaling pathways. Additionally, tumor cells have the capability to release various growth factors, such as tumor growth factor, endothelial growth factor, and VEGF, contributing to the promotion of new blood vessel development (50). Angiogenesis is crucial for providing nutrients and oxygen to tumor cells, ultimately playing a critical role in tumor growth.

Therefore, TME plays a crucial role in tumor growth and metastasis. A comprehensive understanding of TME formation, investigating the interplay between immune cells and tumors, and exploring various genetic variations represent the future directions of TME research (51, 52). Additionally, selecting targeted therapeutic strategies based on TME subtypes can enhance the effectiveness of cancer treatment. To further emphasize this point, computer-assisted automatic detection of tumor cells in lymph nodes can significantly reduce the false-negative rate, thereby facilitating earlier detection and treatment of lung cancer, improving the accuracy of TNM staging, accelerating the examination process, and reducing the workload of pathologists. Moreover, tumor spread through air spaces (STAS) has been

identified as an important clinical factor associated with tumor recurrence and poor prognosis in patient survival. The identification and quantification of STAS require experienced pathologists to perform detailed examinations of entire tissue sections. Therefore, pathological image analysis tools that rapidly and accurately identifies STAS would be useful for pathologists (16). Quantitative characterization of TME and accurate prediction and classification of important TME components are essential for targeted tumor therapy and prognosis assessment (53), necessitating advanced data processing and analysis approaches.

Quantitative characterization of TME involves a crucial step of segmenting different types of tissue substructures and cells from pathological images. This segmentation forms the foundation for various image analysis tasks, including cellular composition, spatial organization, and morphology specific to substructures. Previous studies in oncology primarily focused on tumor cells, overlooking the pivotal role of TME in the initiation and progression of cancer. The TME of lung cancer is primarily composed of tumor cells, lymphocytes, stromal cells, macrophages, blood vessels, and other components. Studies in lung cancer have indicated that TILs are positive prognostic factors, while angiogenesis is negatively associated with survival outcomes. The role of stromal cells in prognosis is complex. Traditional image processing methods encompass feature definition, feature extraction or segmentation. These techniques have been employed to segment lymphocytes and analyze the spatial organization of TILs and stromal cells within the TME (54). Research associated with the quantitative characterization of TME has the potential to predict treatment outcomes and provides insights for the development of targeted therapeutic strategies. Innovative studies in immunotherapy, in particular, heavily rely on understanding the interactions among various components within the TME and the mechanisms of immune evasion.

Accurate characterization of specific structures and features of TME is crucial for evaluating tumor prognosis (55), enhancing clinical decisions, and advancing precision medicine. Radiomics can unveil the heterogeneity of tumor cells and TME, while genomics and pathomics explore the biological significance of imaging histological features. The integration of these three approaches contributes to a comprehensive understanding and decoding of TME characteristics in tumors, facilitating prognostic predictions (56). The interconnection between radiomics, pathomics, and genomics contributes to establishing and deepening our understanding of cancer biology and imaging features. Concurrently, powerful machine learning techniques can decipher the complex interactions between tumors and cancer treatments. The integration of machine learning technologies with digital imaging and novel methods for assessing TME at the molecular level significantly enhances our comprehension of TME and cancer prognosis assessment. Vanguri RS et al. employed machine learning to integrate multimodal features into a risk prediction model (57). By combining radiological, histopathological, and genomic features, they assessed the predictive capability of immunotherapy response in NSCLC. Their study revealed that the AUC value of the multimodal model was 0.80, surpassing any single variable. These findings establish a quantitative foundation for enhancing

the accuracy of predicting immunotherapy response in NSCLC patients through the integration of multimodal features and machine learning.

Simultaneously, the quantitative characterization of TME in lung cancer poses certain challenges, including the following aspects: (1) Complexity and heterogeneity of lung cancer TME composition: In addition to the mentioned cell types, other structures such as bronchi, cartilage, and pleura often appear in pathological sections of the lung. This complexity and heterogeneity make segmentation and traditional feature definition challenging. (2) Cellular spatial organization (e.g., spatial distribution and interactions of different cell types): While playing a crucial role in TME, it is more challenging to capture than simply providing the quantity or ratio of different cell types. Current research mainly focuses on the proportion of different cell types, overlooking the intricate cellular spatial organization, which may result in limited and contradictory outcomes regarding the roles of different cell types in the TME. (3) For H&E-stained glass slides, there can be significant color variations based on staining conditions and the time gap between slide preparation and scanning. Traditional image processing methods based on manual feature extraction struggle to overcome these obstacles. (4) Multi-omics studies face the high dimensionality and heterogeneity of data, and integrating quantitative measurements of multi-modal data for prognosis prediction is a highly challenging task. In summary, pathomics, as a nascent research methodology, is presently undergoing preliminary investigation. Future studies utilizing extensive multi-omics datasets have the potential to advance the formulation of sophisticated integration strategies. These strategies would facilitate a more exhaustive evaluation, characterization, and elucidation of TME (58). Consequently, this advancement will yield profound insights into the imaging characteristics and the pathophysiological and biological underpinnings of tumor pathology.

In recent years, amidst the high incidence and mortality rates of lung cancer, the selection and implementation of treatment plans for advanced-stage lung cancer patients, as well as the creation of more precise platforms for predicting treatment responses, continue to face challenges. Pathomics not only synergizes with traditional pathological semantic information and clinical data to discover disease patterns but also interacts and integrates with various omics information, leveraging the unique advantages of each omics discipline. The development of these interdisciplinary approaches not only aids in identifying subtle lesions that may escape the naked eye and uncovering disease patterns beyond subjective judgment but also facilitates relatively objective and accurate assistance in disease screening, diagnosis, differential diagnosis, and prognosis assessment. Furthermore, it contributes to saving human and material resources, optimizing the utilization of limited medical resources to the maximum extent, and, on a broader scale, promoting the development of the personalized immune intervention.

## 5 Conclusion

In conclusion, this study systematically analyzed the literature on pathomics in the management of lung cancer indexed within the

WoSCC. It offers an initial overview of recent research trends and forecasts potential hotspots and frontiers for future inquiry, aiming to provide valuable insights and references for scholars and researchers involved in personalized immunotherapy efficacy and prognosis for lung cancer.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

## Author contributions

LY: Validation, Visualization, Writing – original draft, Writing – review & editing. ZS: Validation, Writing – original draft, Writing – review & editing. YS: Data curation, Validation, Writing – original draft, Writing – review & editing. JZ: Data curation, Software, Validation, Visualization, Writing – review & editing. QW: Data curation, Software, Validation, Visualization, Writing – review & editing. YL: Data curation, Visualization, Writing – review & editing. HS: Supervision, Validation, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX23\_3623) and National Natural Science Foundation of China (No. 82070020).

## Acknowledgments

All authors acknowledge the contributions from the CiteSpace, the bibliometric online analysis platform and the WoSCC.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2021. *CA: A Cancer J Clin.* (2021) 71:7–33. doi: 10.3322/caac.21654
- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* (2021) 71:209–49. doi: 10.3322/caac.21660
- Brozos-Vazquez EM, Diaz-Pena R, Garcia-Gonzalez J, Leon-Mateos L, Mondelomacia P, Pena-Chilet M, et al. Immunotherapy in nonsmall-cell lung cancer: current status and future prospects for liquid biopsy. *Cancer Immunol Immunother.* (2021) 70:1177–88. doi: 10.1007/s00262-020-02752-z
- Zhang Y, Yang Z, Chen R, Zhu Y, Liu L, Dong J, et al. Histopathology images-based deep learning prediction of prognosis and therapeutic response in small cell lung cancer. *NPJ Digital Med.* (2024) 7(1):15. doi: 10.1038/s41746-024-01003-0
- Liu YH, Cheng X, Han XA, Cheng X, Jiang S, Lin YR, et al. Global research landscape and trends of lung cancer immunotherapy: A bibliometric analysis. *Front Immunol.* (2022) 13:1032747. doi: 10.3389/fimmu.2022.1032747
- Dia AK, Ebrahimipour L, Yolchuyeva S, Tonneau M, Lamaze FC, Orain M, et al. The cross-Scale association between pathomics and radiomics features in immunotherapy-Treated NSCLC patients: a preliminary study. *Cancers.* (2024) 16(2):348. doi: 10.3390/cancers16020348
- Chen C. A glimpse of the first eight months of the COVID-19 literature on microsoft academic graph: themes, citation contexts, and uncertainties. *Front Res Metrics Analytics.* (2020) 5:607286. doi: 10.3389/frma.2020.607286
- Miao L, Zhang J, Zhang Z, Wang S, Tang F, Teng M, et al. A bibliometric and knowledge-map analysis of CAR-T cells from 2009 to 2021. *Front Immunol.* (2022) 13:840956. doi: 10.3389/fimmu.2022.840956
- Wang Q, Yang K-L, Zhang Z, Wang Z, Li C, Li L, et al. Characterization of global research trends and prospects on single-cell sequencing technology: bibliometric analysis. *J Med Internet Res.* (2021) 23(8):e25789. doi: 10.2196/25789
- Pucci C, Martinelli C, Ciofani G. Innovative approaches for cancer treatment: current perspectives and new challenges. *Ecancermedicalscience.* (2019) 13:961. doi: 10.3332/ecancer.2019.961
- Lu MY, Williamson DFK, Chen TY, Chen RJ, Barbieri M, Mahmood F. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nat Biomed Eng.* (2021) 5(6):555–70. doi: 10.1038/s41551-020-00682-w
- Khosravi P, Kazemi E, Imielinski M, Elemento O, Hajirasouliha I. Deep convolutional neural networks enable discrimination of heterogeneous digital pathology images. *EBioMedicine.* (2018) 27:317–28. doi: 10.1016/j.ebiom.2017.12.026
- Acs B, Rantalainen M, Hartman J. Artificial intelligence as the next step towards precision pathology. *J Internal Med.* (2020) 288:62–81. doi: 10.1111/joim.13030
- Maibach F, Sadozai H, Seyed Jafari SM, Hunger RE, Schenk M. Tumor-infiltrating lymphocytes and their prognostic value in cutaneous melanoma. *Front Immunol.* (2020) 11:2105. doi: 10.3389/fimmu.2020.02105
- Mezheyeuski A, Bergsland CH, Backman M, Djureinovic D, Sjöblom T, Bruun J, et al. Multispectral imaging for quantitative and compartmentalizing immune infiltrates reveals distinct immune profiles that classify lung cancer patients. *J Pathol.* (2018) 244:421–31. doi: 10.1002/path.5026
- Wang S, Yang DM, Rong R, Zhan X, Fujimoto J, Liu H, et al. Artificial intelligence in lung cancer pathology image analysis. *Cancers.* (2019) 11(11):1673. doi: 10.3390/cancers11111673
- Johnson DB, Bordeaux J, Kim JY, Vaupel C, Rimm DL, Ho TH, et al. Quantitative spatial profiling of PD-1/PD-L1 interaction and HLA-DR/IDO-1 predicts improved outcomes of anti-PD-1 therapies in metastatic melanoma. *Clin Cancer Res.* (2018) 24:5250–60. doi: 10.1158/1078-0432.CCR-18-0309
- Baxi V, Edwards R, Montalto M, Saha S. Digital pathology and artificial intelligence in translational medicine and clinical practice. *Modern Pathol.* (2022) 35:23–32. doi: 10.1038/s41379-021-00919-2
- Saw SPL, Ong B-H, Chua KLM, Takano A, Tan DSW. Revisiting neoadjuvant therapy in non-small-cell lung cancer. *Lancet Oncol.* (2021) 22:e501–16. doi: 10.1016/S1470-2045(21)00383-1
- Hanna MG, Parwani A, Sirintrapun SJ. Whole slide imaging: technology and applications. *Adv Anatomic Pathol.* (2020) 27:251–9. doi: 10.1097/PAP.0000000000000273
- Classe M, Lerousseau M, Scoazec J-Y, Deutsch E. Perspectives in pathomics in head and neck cancer. *Curr Opin Oncol.* (2021) 33:175–83. doi: 10.1097/CCO.0000000000000731
- Wang X, Janowczyk A, Zhou Y, Thawani R, Fu P, Schalper K, et al. Prediction of recurrence in early stage non-small cell lung cancer using computer extracted nuclear features from digital H&E images. *Sci Rep.* (2017) 7(1):13543. doi: 10.1038/s41598-017-13773-7
- Prelaj A, Miskovic V, Zanitti M, Trovo F, Genova C, Viscardi G, et al. Artificial intelligence for predictive biomarker discovery in immuno-oncology: a systematic review. *Ann Oncol.* (2024) 35:29–65. doi: 10.1016/j.annonc.2023.10.125
- Alvarez-Jimenez C, Sandino AA, Prasanna P, Gupta A, Viswanath SE, Romero E. Identifying cross-scale associations between radiomic and pathomic signatures of non-small cell lung cancer subtypes: preliminary results. *Cancers.* (2020) 12(12):3663. doi: 10.3390/cancers12123663
- Hosseini MS, Bejnordi BE, Trinh VQ, Chan L, Hasan D, Li X, et al. Computational pathology: A survey review and the way forward. *J Pathol Inform.* (2024) 15:100357. doi: 10.1016/j.jpi.2023.100357
- Jin W, Yang Q, Chi H, Wei K, Zhang P, Zhao G, et al. Ensemble deep learning enhanced with self-attention for predicting immunotherapeutic responses to cancers. *Front Immunol.* (2022) 13:1025330. doi: 10.3389/fimmu.2022.1025330
- Burrows L, Sculthorpe D, Zhang H, Rehman O, Mukherjee A, Chen K. Mathematical modelling and deep learning algorithms to automate assessment of single and digitally multiplexed immunohistochemical stains in tumoural stroma. *J Pathol Inform.* (2024) 15:100351. doi: 10.1016/j.jpi.2023.100351
- LeCun Y, Kavukcuoglu K, Farabet C. Convolutional networks and applications in vision. *Proc 2010 IEEE Int Symposium Circuits Syst.* (2010), 253–6. doi: 10.1109/ISCAS.2010.5537907
- Chang HY, Jung CK, Woo JI, Lee S, Cho J, Kim SW, et al. Artificial intelligence in pathology. *J Pathol Trans Med.* (2019) 53:1–12. doi: 10.4132/jptm.2018.12.16
- Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *J Pathol Inf.* (2016) 7:29. doi: 10.4103/2153-3539.186902
- Robertson S, Azizpour H, Smith K, Hartman J. Digital image analysis in breast pathology—from image processing techniques to artificial intelligence. *Trans Res.* (2018) 194:19–35. doi: 10.1016/j.trsl.2017.10.010
- Chaunzwa TL, Hosny A, Xu Y, Shafer A, Diao N, Lanuti M, et al. Deep learning classification of lung cancer histology using CT images. *Sci Rep.* (2021) 11(1):5471. doi: 10.1038/s41598-021-84630-x
- Nibid L, Greco C, Cordelli E, Sabarese G, Fiore M, Liu CZ, et al. Deep pathomics: a new image-based tool for predicting response to treatment in stage III non-small cell lung cancer. *PLoS One.* (2023) 18(11):e0294259. doi: 10.1371/journal.pone.0294259
- Cao R, Yang F, Ma S-C, Liu L, Zhao Y, Li Y, et al. Development and interpretation of a pathomics-based model for the prediction of microsatellite instability in Colorectal Cancer. *Theranostics.* (2020) 10:11080–91. doi: 10.7150/thno.49864
- Wang S, Chen A, Yang L, Cai L, Xie Y, Fujimoto J, et al. Comprehensive analysis of lung cancer pathology images to discover tumor shape and boundary features that predict survival outcome. *Sci Rep.* (2018) 8(1):10393. doi: 10.1038/s41598-018-27707-4
- Rakae M, Adib E, Ricciuti B, Sholl LM, Shi W, Alessi JV, et al. Association of machine learning-based assessment of tumor-infiltrating lymphocytes on standard histologic images with outcomes of immunotherapy in patients with NSCLC. *JAMA Oncol.* (2023) 9(1):51–60. doi: 10.1001/jamaoncol.2022.4933
- Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med.* (2018) 24:1559–67. doi: 10.1038/s41591-018-0177-5
- Deepak KGK, Vempati R, Nagaraju GP, Dasari VR, S N, Rao DN, et al. Tumor microenvironment: Challenges and opportunities in targeting metastasis of triple negative breast cancer. *Pharmacol Res.* (2020) 153:104683. doi: 10.1016/j.phrs.2020.104683
- Li L, Yu R, Cai T, Chen Z, Lan M, Zou T, et al. Effects of immune cells and cytokines on inflammation and immunosuppression in the tumor microenvironment. *Int Immunopharmacol.* (2020) 88:106939. doi: 10.1016/j.intimp.2020.106939
- Xia Z, Chen S, He M, Li B, Deng Y, Yi L, et al. Editorial: Targeting metabolism to activate T cells and enhance the efficacy of checkpoint blockade immunotherapy in solid tumors. *Front Immunol.* (2023) 14:1247178. doi: 10.3389/fimmu.2023.1247178
- Xiong J, Chi H, Yang G, Zhao S, Zhang J, Tran LJ, et al. Revolutionizing anti-tumor therapy: unleashing the potential of B cell-derived exosomes. *Front Immunol.* (2023) 14:1188760. doi: 10.3389/fimmu.2023.1188760
- Fan J, Zhang T. Multi-index comprehensive evaluation of the efficacy and response mechanism of immunotherapy in non-small cell lung cancer. *Clin Transl Oncol.* (2024). doi: 10.1007/s12094-024-03519-z
- Lei X, Lei Y, Li JK, Du WX, Li RG, Yang J, et al. Immune cells within the tumor microenvironment: Biological functions and roles in cancer immunotherapy. *Cancer Lett.* (2020) 470:126–33. doi: 10.1016/j.canlet.2019.11.009
- Vitale I, Manic G, Coussens LM, Kroemer G, Galluzzi L. Macrophages and metabolism in the tumor microenvironment. *Cell Metab.* (2019) 30:36–50. doi: 10.1016/j.cmet.2019.06.001
- Mao X, Xu J, Wang W, Liang C, Hua J, Liu J, et al. Crosstalk between cancer-associated fibroblasts and immune cells in the tumor microenvironment: new findings and future perspectives. *Mol Cancer.* (2021) 20(1):131. doi: 10.1186/s12943-021-01428-1
- Ren Q, Zhang P, Lin H, Feng Y, Chi H, Zhang X, et al. A novel signature predicts prognosis and immunotherapy in lung adenocarcinoma based on cancer-associated fibroblasts. *Front Immunol.* (2023) 14:1201573. doi: 10.3389/fimmu.2023.1201573

47. Akai M, Noma K, Kato T, Nishimura S, Matsumoto H, Kawasaki K, et al. Fibroblast activation protein-targeted near-infrared photoimmunotherapy depletes immunosuppressive cancer-associated fibroblasts and remodels local tumor immunity. *Br J Cancer*. (2024) 130:1647–58. doi: 10.1038/s41416-024-02639-1
48. Pitt JM, Marabelle A, Eggermont A, Soria JC, Kroemer G, Zitvogel L. Targeting the tumor microenvironment: removing obstruction to anticancer immune responses and immunotherapy. *Ann Oncol*. (2016) 27:1482–92. doi: 10.1093/annonc/mdw168
49. Sun D, Liu J, Zhou H, Shi M, Sun J, Zhao S, et al. Classification of tumor immune microenvironment according to programmed death-Ligand 1 expression and immune infiltration predicts response to immunotherapy plus chemotherapy in advanced patients with NSCLC. *J Thorac Oncol*. (2023) 18:869–81. doi: 10.1016/j.jtho.2023.03.012
50. de Visser KE, Joyce JA. The evolving tumor microenvironment: From cancer initiation to metastatic outgrowth. *Cancer Cell*. (2023) 41:374–403. doi: 10.1016/j.ccell.2023.02.016
51. Fang X, Li D, Wan S, Hu J, Zhang P, Jie D, et al. Insights into the heterogeneity of the tumor microenvironment in lung adenocarcinoma and squamous carcinoma through single-cell transcriptomic analysis: implications for distinct immunotherapy outcomes. *J Gene Med*. (2024) 26:e3694. doi: 10.1002/jgm.3694
52. Chen X, Li R, Yin YH, Liu X, Zhou XJ, Qu YQ. Pan-cancer prognosis, immune infiltration, and drug resistance characterization of lung squamous cell carcinoma tumor microenvironment-related genes. *Biochem Biophys Rep*. (2024) 38:101722. doi: 10.1016/j.bbrep.2024.101722
53. Fukuda H, Arai K, Mizuno H, Nishito Y, Motoi N, Arai Y, et al. Molecular subtypes of lung adenocarcinoma present distinct immune tumor microenvironments. *Cancer Sci*. (2024) 115:1763–77. doi: 10.1111/cas.16154
54. Ichikawa T, Aokage K, Sugano M, Miyoshi T, Kojima M, Fujii S, et al. The ratio of cancer cells to stroma within the invasive area is a histologic prognostic parameter of lung adenocarcinoma. *Lung Cancer*. (2018) 118:30–5. doi: 10.1016/j.lungcan.2018.01.023
55. Huang M, Cha Z, Liu R, Lin M, Gafoor NA, Kong T, et al. Enhancing immunotherapy outcomes by targeted remodeling of the tumor microenvironment via combined cGAS-STING pathway strategies. *Front Immunol*. (2024) 15:1399926. doi: 10.3389/fimmu.2024.1399926
56. Kang W, Qiu X, Luo Y, Luo J, Liu Y, Xi J, et al. Application of radiomics-based multiomics combinations in the tumor microenvironment and cancer prognosis. *J Transl Med*. (2023) 21:598. doi: 10.1186/s12967-023-04437-4
57. Vanguri RS, Luo J, Aukerman AT, Egger JV, Fong CJ, Horvat N, et al. Multimodal integration of radiology, pathology and genomics for prediction of response to PD-(L)1 blockade in patients with non-small cell lung cancer. *Nat Cancer*. (2022) 3:1151–64. doi: 10.1038/s43018-022-00416-8
58. Lucena-Sánchez E, Hicke FJ, Clara-Trujillo S, García-Fernández A, Martínez-Máñez R. Nanoparticle-mediated tumor microenvironment remodeling favors the communication with the immune cells for tumor killing. *Advanced Ther*. (2024) 7(5). doi: 10.1002/adtp.202400004