Zhu *et al. Discover Oncology* (2025) 16:1536
https://doi.org/10.1007/s12672-025-03140-8

Discover Oncology

RESEARCH Open Access

Mapping the future: bibliometric analysis of omics research trends in non-small cell lung cancer

Yanqian Zhu[1,2†], Jiawei Chen[1,2†], Yufei Wang[1,2], Jinyang Hu[1,2], Chen Gao[1,2*] and Linyu Wu[1,2*]

†Yanqian Zhu and Jiawei Chen contributed equally to this work as co-first authors.

*Correspondence:
Chen Gao
doctor_gaochen@zcmu.edu.cn
Linyu Wu
wulinyu@zcmu.edu.cn
[1]Department of Radiology, The First Affiliated Hospital of Zhejiang Chinese Medical University (Zhejiang Provincial Hospital of Chinese Medicine), 54 Youdian Road, Shangcheng District, Hangzhou 310006, China
[2]The First School of Clinical Medicine, Zhejiang Chinese Medical University, Hangzhou 310053, China

Abstract

Purpose Omics technologies, such as genomics, transcriptomics, proteomics, and radiomics, play an increasingly important role in the diagnosis and treatment of non-small cell lung cancer (NSCLC). It is, therefore, essential to unveil the research landscape and future trends of relevant research. This study aims to explore the research fields based on omics technologies in NSCLC, elucidating the research status, hotspots, and trends from a bibliometric perspective.

Methods The Web of Science Core Collection was utilized to retrieve relevant publications in omics technologies and their applications in NSCLC. By using the bibliometric methods and tools ("bibliometrix" R package, VOSviewer, and CiteSpace), data and visualized analyses for annual publication outputs, countries, institutions, authors, journals, references, and keywords proceeded.

Results A total of 5,337 publications were involved in our analysis. These articles, written by 32,286 authors, originated in 5,863 institutions from 82 countries and were published in 797 journals. The *Journal of Thoracic Oncology* and *Clinical Cancer Research* were representative journals in omics-based research in NSCLC. "Survival," "adenocarcinoma," "mutation," "epidermal growth factor receptor," "resistance," and "chemotherapy" were the highest-frequency keywords. Liquid biopsy and deep learning were also trending topics in omics-related research, according to keyword clustering, trend topics, and citation burst analysis.

Conclusion Omics technologies, including genomics, transcriptomics, and proteomics, were widely used in the diagnosis, prognosis, and treatment of NSCLC. And innovative methods, including liquid biopsy and deep learning, demonstrate a profound impact on advancing the understanding and treatment strategies for NSCLC and warrant further investigation.

Keywords Bibliometric analysis, Non-small cell lung cancer, Omics, Visualized study

1 Introduction

Lung cancer continues to be the most common form of cancer and the primary cause of cancer-related fatalities globally, with almost 2.5 million new diagnoses and 1.8 million deaths in 2022 [1]. Non-small cell lung cancer (NSCLC), representing over 85% of

all lung cancer cases, is a histological classification that includes lung adenocarcinomas, squamous cell carcinomas, and other subtypes [2]. Enhancing survival rates and clinical outcomes for NSCLC patients has remained a central goal in research. Surgery is generally considered the preferred treatment option for this condition. In contrast, adjuvant therapy, such as radiotherapy and chemotherapy, can delay disease progression for inoperable patients with distant tumor metastasis or locally advanced tumors [3]. With an in-depth understanding of molecular mechanisms underlying tumor initiation and progression, molecular targeted therapy has garnered significant attention from researchers. Numerous therapeutic targets have been successfully identified and utilized, leading to substantial improvements in the prognosis of patients with NSCLC [4]. Immune checkpoint inhibitors (ICIs), targeting cytotoxic T lymphocyte associate protein-4 (CTLA-4) or programmed death receptor 1 (PD-1)/programmed cell death ligand 1 (PD-L1) pathways, comprise an indispensable part of immunotherapy and are essential components of first-line treatment for patients with advanced NSCLC without targetable mutations [5]. However, there are significant differences in the therapeutic effect of various drugs and treatment regimens in different patients [6]. Hence, it is necessary to devise the most suitable treatment strategy for NSCLC patients at the earliest possible stage to optimize therapeutic outcomes.

The application of omics technologies allows for a comprehensive and in-depth understanding of the carcinogenic mechanisms, thereby facilitating the discovery of therapeutic targets. Additionally, omics technologies enable the precise identification of biomarkers to guide diagnosis or treatment, contributing to personalized therapy and improving clinical outcomes. The concept of genomics was first proposed by American geneticist Roderick in 1986 [7]. When the Human Genome Project was completed in 2003, the information contained in DNA sequences was deciphered. On this basis, genomics has developed rapidly, enabling the precise identification of specific genomic characteristics in NSCLC patients, thereby marking a giant breakthrough in precision medicine for NSCLC. With the progress of high-throughput sequencing technology and bioinformatics, omics technologies can now elaborately unveil the molecular characteristics of NSCLC patients at multiple molecular levels, including DNA, RNA, proteins, and metabolites. Omics technologies such as transcriptomics, proteomics, metabolomics, and epigenomics have been successfully applied to identify biomarkers for NSCLC [8]. Since Lambin introduced the concept of radiomics in 2012 [9], this field has achieved significant advancements by leveraging high-throughput extraction of quantitative imaging features from medical images. Omics technologies, including radiomics and radiogenomics, have been widely applied in lung cancer research. The microbiome represents another critical direction of research within the omics disciplines in the context of genomics. Furthermore, multi-omics technology applied in NSCLC breaks the limitations of single-level omics approaches and enables a better comprehension of the mutual influence between different omics data on multiple scales [10], laying a foundation for identifying high-risk individuals, accurate diagnosis, prognosis prediction, and therapeutic effect prediction.

Bibliometrics is a statistical method that employs mathematical and statistical methods to conduct quantitative and qualitative analyses of publications in specific fields [11]. Extensively applied across diverse domains, bibliometric analysis is used to summarize the progress within the research field, explore the contributions of countries,

institutions, authors, and journals [12], and determine hotspots and research trends through an overview of scientific literature and visualization of bibliometric data [11]. VOSviewer, CiteSpace, and bibliometrix packages are commonly used bibliometric tools. The Omics approach has occupied a pivotal position in NSCLC research, with numerous studies summarizing the progress of academic research [13, 14]. Liang, et al. [15] carried out a bibliometric analysis on applying radiomics technology in lung cancer and revealed the research status and hotspots based on a single omics. While previous reviews and bibliometric studies have examined the application of omics technologies, they have not systematically mapped the overall progress and research trends in omics-based studies on NSCLC through bibliometric analysis. This study aims to fill that gap by conducting a comprehensive bibliometric analysis of omics research in NSCLC, assessing the current state of the field, identifying emerging hotspots, and highlighting key trends to guide future research.

2 Materials and methods

2.1 Data collection and cleaning

A comprehensive search of the literature has been conducted on the Web of Science (WoS) Core Collection up to October 9, 2024, using specific search terms like "carcinoma, non-small-cell lung," "omic*," "genome*," "transcriptome*," "proteome*," "metabolome*," "epigenome*," and "radiomic*." The detailed search strategy is listed in Table S1. Two authors (YZ and JC) performed the literature selection process collaboratively. First, we excluded meeting abstracts, editorial material, proceeding papers, corrections, letters, retracted publications, book chapters, data papers, and news items. Then, non-English items were also excluded. Secondly, a further selection was conducted to include literature limited to implicating omics technologies in NSCLC by examining the titles, abstracts, and original text when necessary. Data was exported in plain text file format on October 17, 2024, and it was comprised of publication year, countries, institutions, authors, journals, keywords, and cited references. The absence of duplicates was confirmed upon inspection and removal of duplicate entries. Then, we combined regions of certain countries, such as England, Scotland, Wales, and Northern Ireland, into the UK column and Hong Kong into the China column, displayed separately in the exported data. Keyword merging was conducted to address inconsistencies in singular and plural keyword usage, as well as the presence of synonyms and abbreviations.

2.2 Data statistics and visualized analysis

During the process of data analysis and visualization, we utilized the bibliometrix package for R (version 4.4.1), VOSviewer (version 1.6.20), and CiteSpace (version 6.4.R1) as our primary tools. Bibliometrix was employed to compute the annual publication outputs and number of citations. A trend topics map and a thematic evolution map were also generated in bibliometrix using the author's keywords. VOSviewer was used to analyze countries, institutions, authors, journals, references, and keywords, primarily involving their general information in the field. CiteSpace was used mainly for visualized analysis of authors, references, and keywords. We used VOSviewer and CiteSpace to analyze institutions' and authors' outputs and citation information to conduct visualization charts. Subsequently, we created a cooperative relationship network of these authors. Additionally, co-citation analysis and citation burst analysis on references were

also performed by CiteSpace. CiteSpace was also utilized to conduct co-occurrence, clustering, and citation burst analysis for keywords. A timeline chart of keyword clustering was also generated. In addition, information including Single Country Publications (SCP) and Multiple Countries Publications (MCP) of countries, along with the Impact Factor (IF) of journals in 2023, was directly obtained, analyzed, and exported from WoS. The flowchart for this study has been provided in Fig. S1.

3 Results

Seven thousand eight hundred thirty publications were retrieved from the WoS Core Collection until October 9, 2024. A total of 5,337 papers were included in our bibliometric analysis (4,897 articles and 440 review articles). While 1,347 items in other publication types or not written in English were excluded, so were the 1,146 papers not pertinent to omics technologies and their application in NSCLC (Fig. S1).

3.1 Annual publication outputs

A total of 5,337 papers related to omics technologies in NSCLC were included in this study from 2008 to 2024. Figure 1 illustrates the trends in the growth of publication outputs and the average number of citations per paper per year. The outputs demonstrated a slow increase before 2014, during the initial development stage of omics-related research in NSCLC. The period from 2015 to 2019 has witnessed a steady development in the fields of omics technologies in NSCLC. Then, the number of publications experienced rapid growth from 2020 to 2022, culminating in a peak of 774 papers in 2022, a 18-fold increase compared to 2008. Although the number of publications declined in the past two years, the annual publication outputs remained roughly equivalent to 2021. Notably, 2024 witnessed the publication of 578 papers up until October 9. Regarding citations, these studies collectively received 169,229 citations, with the average number of citations per paper per year peaking in 2015 at 8.05. This was primarily from the article "Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer," published in *Science*, with a citation of 6,167 times, accounting for 38.4% of the total citations in 2015.

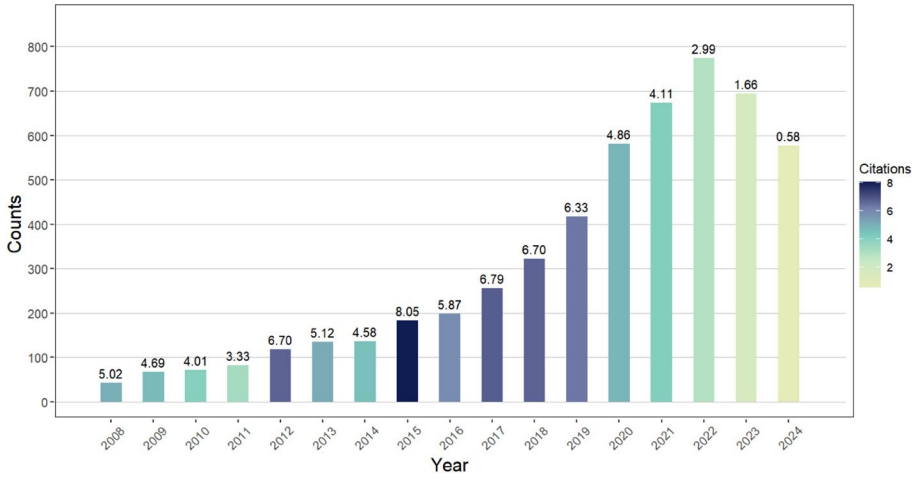

Fig. 1 Annual output and citation information of publications on omics technologies in NSCLC from 2008 to 2024. Citations represent the number of average citations per paper per year

3.2 Analysis of countries

Eighty-two countries contributed to the publication of these publications in total, with China (count = 2,531, 47.42%) and the USA (count = 1,639, 30.71%) being the two leading countries in terms of the number of publications, followed by Italy (count = 325, 6.09%), Japan (count = 258, 4.83%), and Germany (count = 243, 4.55%). Among the top 10 most productive countries, China produced the largest amount of literature, followed by the USA, possessing the highest count of citations on average, reaching 56.78. Detailed information about the most productive countries is provided in Fig. 2.

Based on the analysis of international collaboration, the UK (83.92%), Canada (73.91%), and France (70.22%) exhibited the highest proportions of MCP among the top 10 most productive countries (Fig. 2). Regarding the number of published papers, the USA had a relatively higher frequency participating in international collaborative studies. Using VOSviewer, we derived the collaboration data among 40 countries with a minimum publication output of 10 and created a collaboration network to illustrate the collaborations among these countries (Fig. 2). Each node represented a country, and the different colors represented the years. The line connecting the nodes represented the co-occurrence between countries, while the thickest one between China and the USA indicated the highest degree of cooperation.

3.3 Analysis of institutions

A total of 5,863 institutions have contributed to applying omics technologies in NSCLC. Figure 3 illustrates the number of publications and citations of each institution ranked in the top 10 by outputs. The UTMD Anderson Cancer Center (count = 178, citation = 13,422) stood out with the highest number of publications, followed by Shanghai Jiao Tong University (count = 174, citation = 4,330) and Nanjing Medical University (count = 168, citation = 4,884).

3.4 Analysis of authors

Thirty-two thousand two hundred eighty-six authors published papers on omics-based research in NSCLC. Among them, the authors with the highest output were Christiani, David C (count = 46, citation = 997) and Wang, Jing (count = 37, citation = 2,341). Minna, John D (count = 32, citation = 3,495) and Wistuba, Ignacio I (count = 32, citation = 4,959) were tied for third place. The detailed information of the top 10 authors with the largest number of publications was visualized in Fig. 3. A visualized analysis of author collaboration was also conducted using CiteSpace. Six central authors possess a centrality exceeding 0.1: Wang, Jing (centrality = 0.19), Behrens, Carmen (centrality = 0.15), Awad, Mark M (centrality = 0.14), Minna, John D (centrality = 0.12), Gibbons, Don L (centrality = 0.11), and Soria, Jean Charles (centrality = 0.1).

3.5 Analysis of journals

Clinical Cancer Research (citation = 11,674) and *Journal of Thoracic Oncology* (citation = 7,129) were the most cited journals among the total of 797 journals. Despite only containing two papers, *Science* was cited 6,996 times and ranked third in the most cited journals, earning it the distinction of holding the highest average number of citations per paper. The most productive journal was *Frontiers in Oncology* (count = 217). Additionally, the *Journal of Thoracic Oncology* (count = 135) and *Clinical Cancer Research*

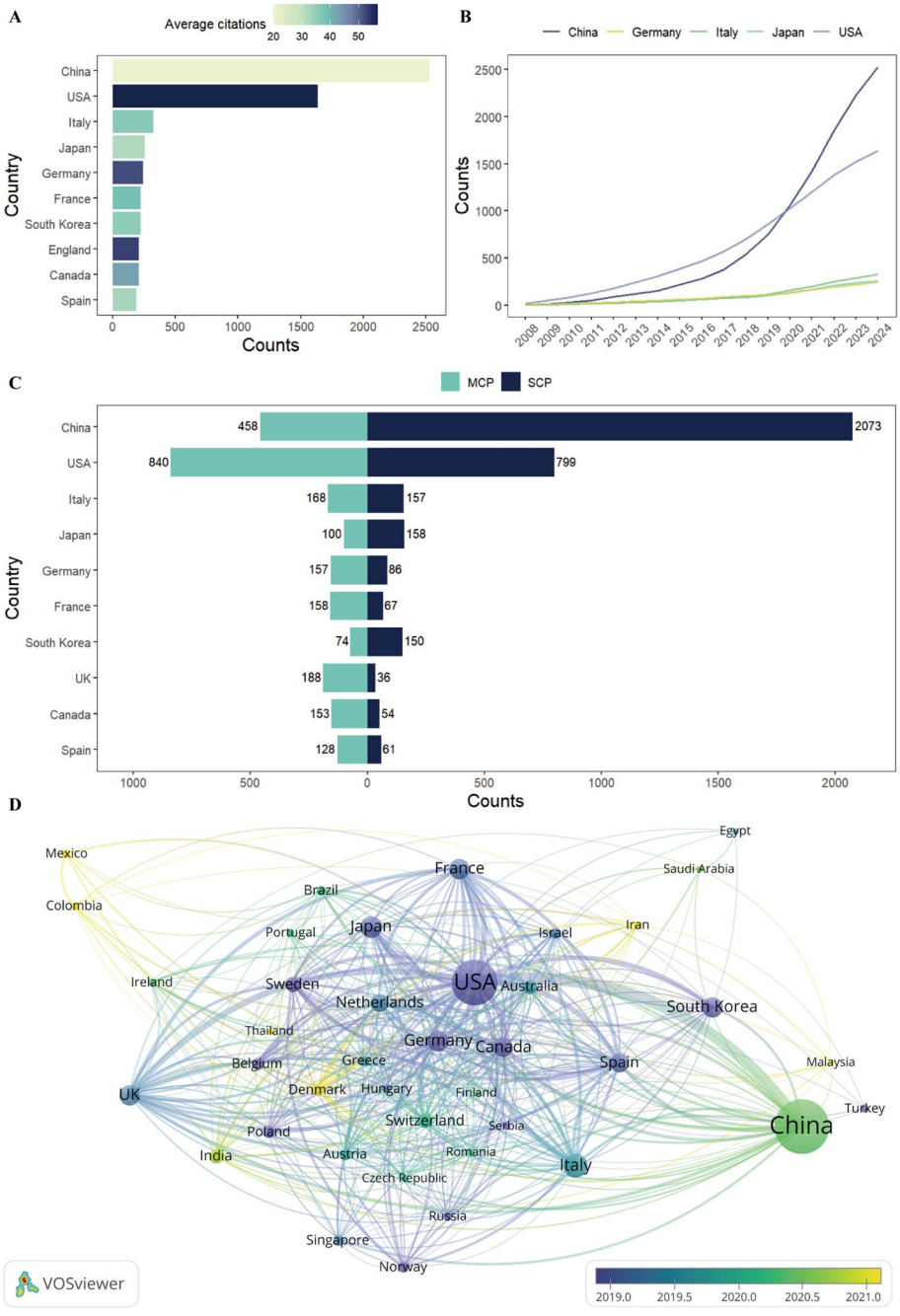

Fig. 2 Visualization maps on publications by different countries. **A** Publications and citations contribution of the top 10 most productive countries; **B** the annual variation trends in the cumulative number of publications among the top five countries; **C** collaboration status of the top 10 most productive countries; **D** collaboration network of the top 40 most productive countries by time. *SCP* single country publications; *MCP* multiple countries publications; *UK* United Kingdom; *USA* United States of America

(count = 131) were also highly productive journals, ranking in the top 5 in terms of the number of publications. The IF is a crucial metric for assessing the academic influence of a journal and its publications. Among the top 10 most frequently cited journals, *Science* held the highest IF of 44.8, followed by *Journal of Clinical Oncology* (IF = 42.1) and *Journal of Thoracic Oncology* (IF = 21.1). The information on the most cited journals is shown in Fig. 4.

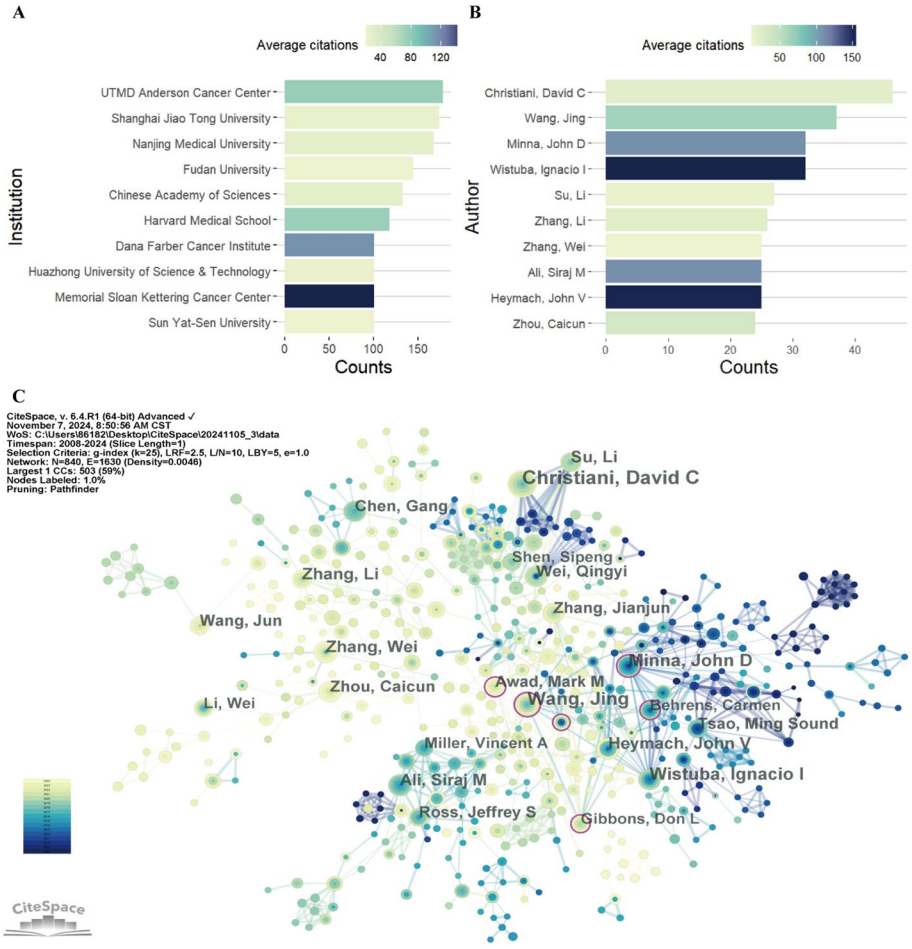

Fig. 3 Visualization maps on institutions and authors. **A** The top 10 most productive institutions; **B** the top 10 most productive authors; **C** the collaboration network of authors. The node size represents the authors' publication output. The nodes with purple rings represent the central authors, with the centrality reaching 0.1. The colors represent the years from 2008 to 2024

3.6 Highly cited publications and co-cited references

All 5,337 papers were cited 169,229 times. Table 1 lists the top 10 most cited papers, their publication years, total citation counts, countries, number of collaborative institutions, and source journals. Notably, the first five papers in this list were cited over 1,500 times. All 10 papers were authored in collaborations involving multiple institutions, with 7 in collaborations across multiple countries.

A total of 134,181 references were cited within these studies. Among the top 15 references with the strongest citation bursts among these references, the three most recent references that emerged were as follows: "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries" (strength = 94.72), published in *CA: A Cancer Journal for Clinicians* in 2021; "Radiomics: Images Are More than Pictures, They Are Data" (strength = 30.85), published in *Radiology* in 2016; and "Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer" (strength = 42.57), published in *New England Journal of Medicine* in 2016. The information on data citation burst analysis is available in Fig. S2.

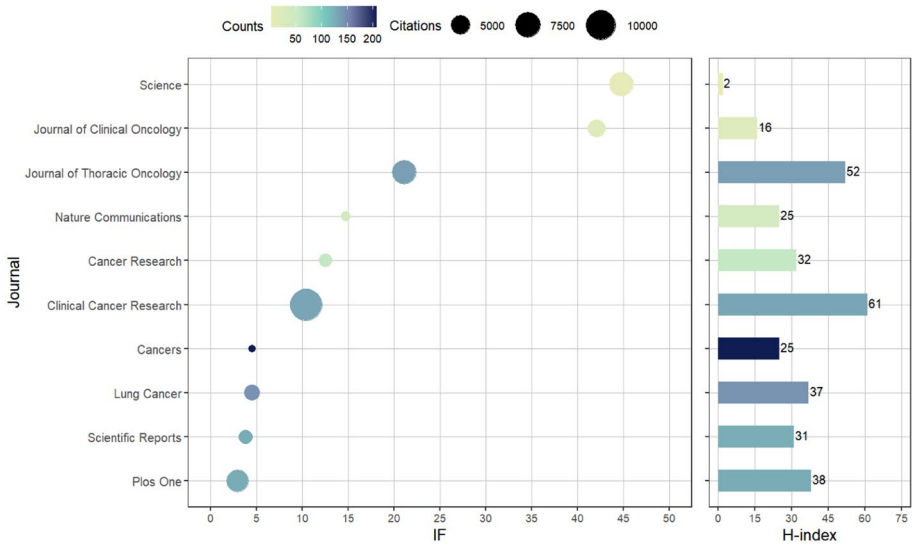

Fig. 4 The top 10 journals with the largest number of citations. The location of the nodes shows the IF (2023) of different journals, the node size represents the number of total citations, the length of the colored bar on the right illustrates the H-index of each journal in this field, and the color represents the number of publications. *IF* impact factor

Table 1 The top 10 most cited papers

Rank	Year	Total citations	Country	Number of institutions	Journal	References
1	2015	6167	USA, Netherlands	9	Science	[16]
2	2017	1600	USA	3	Molecular Cancer Therapeutics	[17]
3	2016	1569	USA, China	6	Genome Biology	[18]
4	2017	1564	UK, Denmark, USA, Hungary, Germany, Belgium	39	New England Journal of Medicine	[19]
5	2012	1529	USA, Netherlands	7	Magnetic Resonance Imaging	[20]
6	2012	1430	USA, South Korea, Germany	14	Cell	[21]
7	2013	1382	Hungary, Poland, Germany	6	Plos One	[22]
8	2018	1060	USA, Israel, China	19	Cancer Discovery	[23]
9	2018	1034	USA	7	Journal of Clinical Oncology	[24]
10	2012	909	USA	5	Cell	[25]

3.7 Analysis of keywords

CiteSpace extracted 692 keywords for analysis, creating a network visualization map of keywords (Fig. 5). The nodes of keywords such as "survival," "adenocarcinoma," "mutation," "epidermal growth factor receptor," "resistance," and "chemotherapy" possessed the larger size, which indicated that these keywords appeared more frequently in the research results, with their occurrence frequency peaking mostly between 2020 and 2022. There were 21 nodes with purple rings around them, representing the 21 central keywords, including "classification" (centrality = 0.18), "prognostic signature" (centrality = 0.18), and "smoking" (centrality = 0.17). A total of 20 clusters were formed through the keyword clustering analysis. The LLR algorithm was employed to generate labels for these clusters. The keyword citation burst analysis revealed that the three keywords

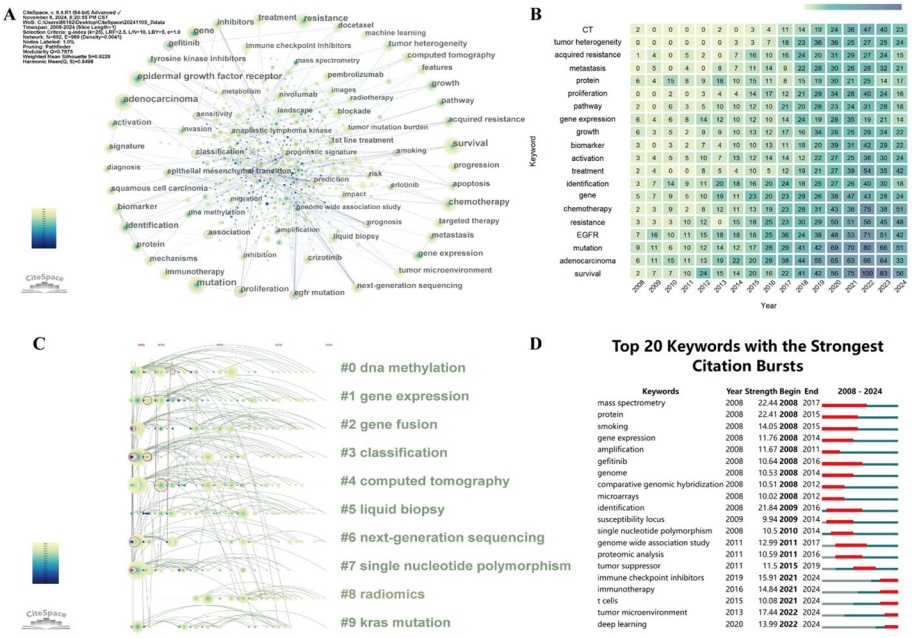

Fig. 5 Analysis of all keywords in the application of the omics technologies in NSCLC. **A** Network visualization of keywords; **B** annual fluctuations in publications for the top 20 high-frequency keywords; **C** the timeline chart of keyword clustering; **D** the top 20 keywords with the strongest citation bursts

with the highest burst strength were "mass spectrometry," "protein," and "identification." Additionally, the latest burst keywords identified in the analysis include "deep learning," "tumor microenvironment," "T cells," "immunotherapy," and "immune checkpoint inhibitors."

Additionally, we generated a trend topics map in the bibliometric package using the author's keywords. As shown in Fig. 6, the lines in the trend topics map indicating "gut microbes," "Traditional Chinese Medicine," "18F-FDG PET/CT," "deep learning," and "network pharmacology" commence in 2022 and conclude in 2024, suggesting that these topics were prominent since 2022 in this research domain. Similarly, "radiomics" and "CT" received the attention of researchers between 2020 and 2023. Furthermore, we divided the data into two distinct periods to create a thematic evolution map (Fig. 6), aiming to depict the evolution trajectory of each stage within the omics disciplines in NSCLC research. Between 2022 and 2024, research in NSCLC primarily focused on immunotherapy, *EGFR*, and *ALK*. This diagram also revealed an evolution from topics such as "biomarkers," "next-generation sequencing," and "adenocarcinoma" to an emphasis on "immunotherapy." Developing biomarkers for immunotherapy became a key area of exploration within omics technologies applied to NSCLC.

4 Discussion

This study conducted a bibliometric analysis to elucidate the present research landscape on omics technologies in NSCLC, including genomics, transcriptomics, proteomics, and radiomics, to illustrate the research hotspots and trends in these research fields. The results indicated that omics approaches were mainly used in diagnosing, prognosing, and treating NSCLC. At the same time, radiotherapy, chemotherapy, targeted therapy, and immunotherapy were the major applications of omics-based research from the therapeutic aspect of NSCLC. The results also emphasized that liquid biopsy and deep

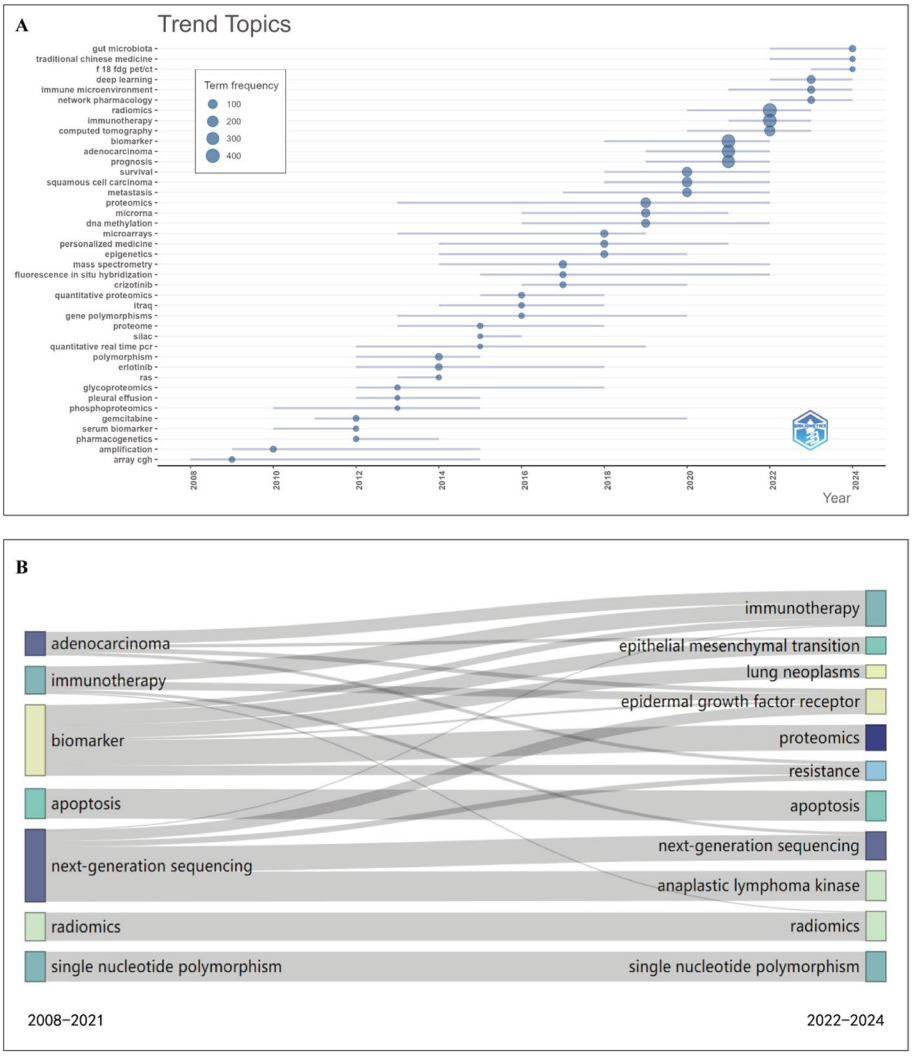

Fig. 6 Visualized analysis of author's keywords. **A** Trend topics map from 2008 to 2024; **B** thematic evolution by time. The nodes represent different topics, and the edges between nodes represent the flowing data. The time segmentation point was established in 2021

learning are pivotal tools for future development and play an essential role in omics-related research and clinical applications.

Based on the publication trend from 2008 to 2024, the application of omics technologies in NSCLC has consistently remained a focal point of research, with a steady increase in the number of publications as of October 9, 2024. In line with previous studies on radiomics [15], the top 10 most productive countries, mostly in East Asia, North America, and Western Europe, hold relatively high Human Development Index (HDI). Epidemiological studies have shown that the incidence and mortality rates of lung cancer are higher in East Asia, North America, and Europe, especially in countries with higher levels of economic development, as assessed by the HDI [1, 2]. This phenomenon explains that these countries allocate greater attention and investment to the research fields based on omics technologies in NSCLC, leading to a higher output of relevant publications. The *Journal of Thoracic Oncology* and *Clinical Cancer Research* rank in the top 10 regarding the number of papers published and total citations received, with an average citation per paper exceeding 50, suggesting them as the representative journals.

The most cited papers are mainly published in top journals, and multiple countries and institutions co-author most of these papers. Collaboration across multiple countries and institutions, involving a wider range of study participants from various nations, regions, and ethnic backgrounds, can significantly enhance the reliability and generalization of research findings. Furthermore, such collaboration between multiple countries and institutions typically provides access to more research resources and expertise, thereby improving academic research.

Genomics, transcriptomics, proteomics, and metabolomics have greatly advanced our knowledge of NSCLC. These omics technologies serve as crucial tools for elucidating the mechanisms of initiation and progression in NSCLC, identifying therapeutic targets or drug resistance, and enhancing the accuracy of diagnosis, prognosis prediction, and therapeutic prediction. Keyword co-occurrence and clustering analysis indicate that genomic alterations, including gene mutation and gene fusion, are the hotspots in omics-based research in NSCLC. Although significant progress has been made in detecting genomic alterations in NSCLC, with several key driver genes identified, including *EGFR*, *ALK*, *KRAS*, *TP53*, and *ROS1* [25, 26], additional and full-scale investigation of gene alterations still make sense, exploring the biological characterization and novel therapeutic targets in lung cancer [21]. And it is crucial for understanding the carcinogenic mechanisms and further guiding personalized diagnosis and treatment. Gene fusions can result from structural rearrangements or the cis-splicing and trans-splicing of pre-mRNAs [27]. These alterations have also been realized as important drivers of cancer, while fusion transcripts have been confirmed as promising therapeutic targets [27]. Since some fusions may be omitted by DNA-based detection, using comprehensive panel tests and analyzing both DNA and RNA samples will facilitate the discovery of rare novel fusions [28].

The application of epigenomics in NSCLC primarily concentrates on DNA methylation, histone modifications, and the regulation of non-coding RNAs. Keyword clustering analysis shows that DNA methylation is a significant focus of attention for researchers. Methylation markers are valuable in identifying high-risk individuals, facilitating early diagnosis, predicting treatment response, and optimizing prognosis prediction [29]. Notably, the diagnostic value of plasma methylation markers in occult lymph node (LN) metastasis has been confirmed in NSCLC, contributing to the preoperative judgment of LN status and the enhancement of treatment strategies [30]. Specific methylation profiles can also represent targets for therapeutic intervention, with drugs discovered and developed, such as decitabine and azacytidine, playing a significant role in the precision treatment of NSCLC [31]. The timeline chart of keyword clustering indicates that the radiomics application emerged relatively later than other omics technologies. Radiomics can provide supplementary information on tumor heterogeneity at the whole tumor level in a non-invasive and repeatable way [9], making it a promising tool for improving the efficacy of diagnosis, prognosis prediction, and therapeutic prediction in NSCLC. Previous research has shown that radiomics features are significantly associated with outcome (overall survival, progression-free survival, and recurrence) and tumor biological characteristics (genomic or pathological features and signaling pathways) in NSCLC [32].

Citation and keyword analysis indicated that relevant applications mainly focus on radiotherapy, chemotherapy, targeted therapy, and immunotherapy, roughly in accord

with the previous study [15]. Targeted therapy primarily acts on specific molecular targets or signal pathway, inhibiting tumor cells' growth or proliferation and inducing apoptosis, for therapeutic purposes. The application of omics technologies enables the discovery of novel therapeutic targets for targeted therapy. *EGFR* mutations and *ALK* rearrangements are well-characterized oncogenic drivers. Tyrosine kinase inhibitors for *EGFR* and *ALK* alterations have improved progression-free survival compared with chemotherapy [33, 34]. In addition, oncogenes such as *BRAF*, *MET*, *ALK*, *ROS1*, *RET*, and *NTRK* have also been successfully identified; targeted therapies for these oncogenes are available and have become the first-line therapy for advanced NSCLC with actionable oncogenic alterations [5]. However, the occurrence of primary or acquired resistance is a major limitation of targeted therapy, and comprehensive molecular profiling and sufficient characterization of tumor heterogeneity make great sense in identifying and monitoring drug resistance.

Regarding immunotherapy, relevant studies based on omics technologies mainly concentrated on ICIs. Determining which patients will benefit from treatment is critical, given the potential for severe side effects and high costs [5]. PD-L1 protein expression, detected through immunohistochemistry, is the most widely used predictive biomarker for ICIs in clinical trials. However, its reliability is limited due to its inducibility and temporal and spatial heterogeneity [35]. Therefore, the extensive exploration of biomarkers across a wide range of omics technologies is still ongoing [16, 17, 24, 36, 37]. In addition, omics technologies have been applied to facilitate a comprehensive characterization of molecular features and cellular compositions in the tumor microenvironment (TME) [18]. Indeed, single-cell analysis has been integrated with omics technologies, such as genomics, epigenomics, transcriptomics, and proteomics [38, 39], offering a higher resolution way to investigate the TME. These applications grant insights into the mechanisms of anti-tumor immune response and tumor immune escape, thereby driving the improvement of immunotherapy strategies and treatment response rate [23]. Omics technologies also possess significant application value for radiotherapy and chemotherapy of NSCLC. Omics biomarkers hold great promise in predicting the efficacy and toxicity of radiotherapy or chemotherapy [40–42]. Research has demonstrated that radiomics features, single nucleotide polymorphisms, epigenomic features, and specific proteins or peptides can be biomarkers for predicting radiotherapy toxicity [40]. Furthermore, omics biomarkers are associated with chemotherapy resistance [41]. An intensive exploration of these biomarkers is significant for better patient selection, enhancing treatment efficacy, and minimizing toxicities [5].

The development and application of omics technologies in diagnosing, prognosing, and treating NSCLC depend on the progression of related technologies. The analysis of keyword clusters and the timeline chart of keyword clustering demonstrates that liquid biopsy has been a hot topic in omics-based research applied to NSCLC. Liquid biopsy mainly investigates circulating tumor cells, circulating tumor DNA (ctDNA), and extracellular vesicles from liquid samples, including blood, pleural effusion, cerebrospinal fluid, saliva, and urine [43]. This approach enables the non-invasive detection of molecular characteristics in patients with NSCLC and facilitates the dynamic monitoring of specific diagnostic, prognostic, or therapeutic markers. Notably, ctDNA has considerable sensitivity in detecting actionable gene fusions and genomic rearrangements in various cancer types [44] and has exhibited potential as a powerful tool for identifying

postoperative molecular residual disease [45]. Citation burst analysis of keywords and trend topics map indicate that deep learning continues to gain attention in recent years, consistent with previous research results [15]. Deep learning, a subset of artificial intelligence based on neural network structure, stands out in discovering intricate structures of high-dimensional data [46] and is thus applicable to the processing and analysis of medical imaging data. Significantly, deep learning can integrate imaging scans at multiple time points, improving clinical outcomes' prediction ability [47]. Moreover, deep learning enables the analysis of large-scale omics data, enhancing multi-omics research by effectively capturing nonlinearities and complex relationships within the data [48]. It has been developed to address common challenges, such as missing modalities in samples and the integration of imaging-based omics data with molecular omics data, demonstrating its advantages in multi-omics research [49]. Further advancements and optimization of deep learning technology are essential for driving progress in omics-based NSCLC research.

There are some limitations in our study. Firstly, since the retrieve was confined to the WoS Core Collection, and only the papers written in English were included, the literature's coverage may not be comprehensive enough. Further studies may necessitate integrating data extracted from additional databases, such as PubMed and Scopus. Secondly, the citation data did not exclude self-citations, which may cause some biases. Additionally, our analysis was limited to the keywords and citations, without incorporating a full-text analysis, potentially resulting in omitting some important information only reflected in the text.

5 Conclusions

Omics technologies have significantly shaped the landscape of NSCLC research, offering valuable insights into diagnosis, prognosis, and treatment. This bibliometric analysis highlights the continuous evolution of omics-based studies, with contributions from genomics, transcriptomics, proteomics, and radiomics, for example. Emerging trends, such as liquid biopsy and deep learning, underscore the shift toward more precise and personalized approaches in NSCLC management. However, radiotherapy, chemotherapy, targeted therapy, and immunotherapy are the main directions of omics-based research related to the treatment of NSCLC. While advancements in omics technologies continue to enhance our understanding of tumor biology, further research is needed to address existing challenges and optimize their clinical applications. Moving forward, the more disparate disciplines' involvement and technological innovations will be critical in driving the next phase of omics research in NSCLC.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1007/s12672-025-03140-8.

Supplementary Material 1

Acknowledgements
None.

Author contributions
YZ and JC: investigation, formal analysis, data curation, visualization, validation, writing - original draft, writing - review and editing. YW: investigation, writing - review and editing. JH: investigation, writing - review and editing. CG and LW: conceptualization, methodology, investigation, data curation, writing - review and editing, funding acquisition, supervision, project administration. All authors read and approved the final manuscript.

Funding
This study has received funding from the Medical and Health Science and Technology Project of Zhejiang Province (Grant no. 2024KY129, 2025KY093, 2024KY132, 2022KY230), the Zhejiang Province Traditional Chinese Medicine Science and Technology Plan Project (Grant no. 2025ZL302, 2025ZS012), and the Research Project of Zhejiang Chinese Medical University (Grant no. 2022FSYYZY08).

Data availability
All data generated or analyzed during this study are included in this published article and its supplementary information files.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing Interests
The authors declare no competing interests.

Received: 6 May 2025 / Accepted: 2 July 2025
Published online: 12 August 2025

References

1. 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;74:229–63. https://doi.org/10.3322/caac.21834.
2. Leiter A, Veluswamy RR, Wisnivesky JP. The global burden of lung cancer: current status and future trends. Nat Rev Clin Oncol. 2023;20:624–39. https://doi.org/10.1038/s41571-023-00798-3.
3. Le Chevalier T, Arriagada R, Quoix E, et al. Radiotherapy alone versus combined chemotherapy and radiotherapy in nonresectable non-small-cell lung cancer: first analysis of a randomized trial in 353 patients. J Natl Cancer Inst. 1991;83:417–23. https://doi.org/10.1093/jnci/83.6.417.
4. Herrera-Juárez M, Serrano-Gómez C, Bote-de-Cabo H, et al. Targeted therapy for lung cancer: beyond EGFR and ALK. Cancer. 2023;129:1803–20. https://doi.org/10.1002/cncr.34757.
5. Meyer ML, Fitzgerald BG, Paz-Ares L, et al. New promises and challenges in the treatment of advanced non-small-cell lung cancer. Lancet. 2024;404:803–22. https://doi.org/10.1016/s0140-6736(24)01029-8.
6. Schwaederle M, Zhao M, Lee JJ, et al. Impact of precision medicine in diverse cancers: a meta-analysis of phase II clinical trials. J Clin Oncol. 2015;33:3817–25. https://doi.org/10.1200/jco.2015.61.5997.
7. Kuska B, Beer. Bethesda, and biology: how genomics came into being. J Natl Cancer Inst. 1998;90:93. https://doi.org/10.1093/jnci/90.2.93.
8. Vargas AJ, Harris CC. Biomarker development in the precision medicine era: lung cancer as a case study. Nat Rev Cancer. 2016;16:525–37. https://doi.org/10.1038/nrc.2016.56.
9. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer. 2012;48:441–6. https://doi.org/10.1016/j.ejca.2011.11.036.
10. Menyhárt O, Győrffy B. Multi-omics approaches in cancer research with applications in tumor subtyping, prognosis, and diagnosis. Comput Struct Biotechnol J. 2021;19:949–60. https://doi.org/10.1016/j.csbj.2021.01.009.
11. Cheng K, Li Z, Sun Z, et al. The rapid growth of bibliometric studies: a call for international guidelines. Int J Surg. 2024;110:2446–8. https://doi.org/10.1097/js9.0000000000001049.
12. Agarwal A, Durairajanayagam D, Tatagari S, et al. Bibliometrics: tracking research impact by selecting the appropriate metrics. Asian J Androl. 2016;18:296–309. https://doi.org/10.4103/1008-682x.171582.
13. Ge Y, Ye T, Fu S, et al. Research progress of extracellular vesicles as biomarkers in immunotherapy for non-small cell lung cancer. Front Immunol. 2023;14:1114041. https://doi.org/10.3389/fimmu.2023.1114041.
14. Mei T, Wang T, Zhou Q. Multi-omics and artificial intelligence predict clinical outcomes of immunotherapy in non-small cell lung cancer patients. Clin Exp Med. 2024;24:60. https://doi.org/10.1007/s10238-024-01324-0.
15. Liang H, Chen Z, Wei F, et al. Bibliometrics research on radiomics of lung cancer. Transl Cancer Res. 2021;10:3757–71. https://doi.org/10.21037/tcr-21-1277.
16. Rizvi NA, Hellmann MD, Snyder A, et al. Mutational landscape determines sensitivity to PD-1 Blockade in non-small cell lung cancer. Science. 2015;348:124–8. https://doi.org/10.1126/science.aaa1348.
17. Goodman AM, Kato S, Bazhenova L, et al. Tumor mutational burden as an independent predictor of response to immunotherapy in diverse cancers. Mol Cancer Ther. 2017;16:2598–608. https://doi.org/10.1158/1535-7163.Mct-17-0386.
18. Li B, Severson E, Pignon JC, et al. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. Genome Biol. 2016;17:174. https://doi.org/10.1186/s13059-016-1028-7.
19. Jamal-Hanjani M, Wilson GA, McGranahan N, et al. Tracking the evolution of non-small-cell lung cancer. N Engl J Med. 2017;376:2109–21. https://doi.org/10.1056/NEJMoa1616288.
20. Kumar V, Gu Y, Basu S, et al. Radiomics: the process and the challenges. Magn Reson Imaging. 2012;30:1234–48. https://doi.org/10.1016/j.mri.2012.06.010.
21. Imielinski M, Berger AH, Hammerman PS, et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell. 2012;150:1107–20. https://doi.org/10.1016/j.cell.2012.08.029.

22. Győrffy B, Surowiak P, Budczies J, et al. Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. PLoS One. 2013;8:e82241. https://doi.org/10.1371/journal.pone.0082241.
23. Skoulidis F, Goldberg ME, Greenawalt DM, et al. STK11/LKB1 mutations and PD-1 inhibitor resistance in KRAS-mutant lung adenocarcinoma. Cancer Discov. 2018;8:822–35. https://doi.org/10.1158/2159-8290.Cd-18-0099.
24. Rizvi H, Sanchez-Vega F, La K, et al. Molecular determinants of response to anti-programmed cell death (PD)-1 and anti-programmed death-ligand 1 (PD-L1) Blockade in patients with non-small-cell lung cancer profiled with targeted next-generation sequencing. J Clin Oncol. 2018;36:633–41. https://doi.org/10.1200/jco.2017.75.3384.
25. Govindan R, Ding L, Griffith M, et al. Genomic landscape of non-small cell lung cancer in smokers and never-smokers. Cell. 2012;150:1121–34. https://doi.org/10.1016/j.cell.2012.08.024.
26. Shi J, Hua X, Zhu B, et al. Somatic genomics and clinical features of lung adenocarcinoma: a retrospective study. PLoS Med. 2016;13:e1002162. https://doi.org/10.1371/journal.pmed.1002162.
27. Latysheva NS, Babu MM. Discovering and Understanding oncogenic gene fusions through data intensive computational approaches. Nucleic Acids Res. 2016;44:4487–503. https://doi.org/10.1093/nar/gkw282.
28. Suda K, Mitsudomi T. Emerging oncogenic fusions other than ALK, ROS1, RET, and NTRK in NSCLC and the role of fusions as resistance mechanisms to targeted therapy. Transl Lung Cancer Res. 2020;9:2618–28. https://doi.org/10.21037/tlcr-20-186.
29. Li P, Liu S, Du L, et al. Liquid biopsies based on DNA methylation as biomarkers for the detection and prognosis of lung cancer. Clin Epigenetics. 2022;14:118. https://doi.org/10.1186/s13148-022-01337-0.
30. Chen Z, Xiong S, Li J, et al. DNA methylation markers that correlate with occult lymph node metastases of non-small cell lung cancer and a preliminary prediction model. Transl Lung Cancer Res. 2020;9:280–7. https://doi.org/10.21037/tlcr.2020.03.13.
31. Ansari J, Shackelford RE, El-Osta H. Epigenetics in non-small cell lung cancer: from basics to therapeutics. Transl Lung Cancer Res. 2016;5:155–71. https://doi.org/10.21037/tlcr.2016.02.02.
32. Fornacon-Wood I, Faivre-Finn C, O'Connor JPB, et al. Radiomics as a personalized medicine tool in lung cancer: separating the hope from the hype. Lung Cancer. 2020;146:197–208. https://doi.org/10.1016/j.lungcan.2020.05.028.
33. Mok TS, Wu YL, Thongprasert S, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. N Engl J Med. 2009;361:947–57. https://doi.org/10.1056/NEJMoa0810699.
34. Shaw AT, Kim DW, Nakagawa K, et al. Crizotinib versus chemotherapy in advanced ALK-positive lung cancer. N Engl J Med. 2013;368:2385–94. https://doi.org/10.1056/NEJMoa1214886.
35. Lantuejoul S, Sound-Tsao M, Cooper WA, et al. PD-L1 testing for lung cancer in 2019: perspective from the IASLC pathology committee. J Thorac Oncol. 2020;15:499–519. https://doi.org/10.1016/j.jtho.2019.12.107.
36. Mino-Kenudson M, Schalper K, Cooper W, et al. Predictive biomarkers for immunotherapy in lung cancer: perspective from the international association for the study of lung cancer pathology committee. J Thorac Oncol. 2022;17:1335–54. https://doi.org/10.1016/j.jtho.2022.09.109.
37. Trebeschi S, Drago SG, Birkbak NJ, et al. Predicting response to cancer immunotherapy using noninvasive radiomic biomarkers. Ann Oncol. 2019;30:998–1004. https://doi.org/10.1093/annonc/mdz108.
38. Lin J, Xue X, Wang Y, et al. scNanoCOOL-seq: a long-read single-cell sequencing method for multi-omics profiling within individual cells. Cell Res. 2023;33:879–82. https://doi.org/10.1038/s41422-023-00873-5.
39. Liu Y, Chen X, Zhang Y, et al. Advancing single-cell proteomics and metabolomics with microfluidic technologies. Analyst. 2019;144:846–58. https://doi.org/10.1039/c8an01503a.
40. Yang WC, Hsu FM, Yang PC. Precision radiotherapy for non-small cell lung cancer. J Biomed Sci. 2020;27:82. https://doi.org/10.1186/s12929-020-00676-5.
41. Bonanno L. Predictive models for customizing chemotherapy in advanced non-small cell lung cancer (NSCLC). Transl Lung Cancer Res. 2013;2:160–. https://doi.org/10.3978/j.issn.2218-6751.2013.03.07. 71.
42. Hertz DL, Lustberg MB, Sonis S. Evolution of predictive risk factor analysis for chemotherapy-related toxicity. Support Care Cancer. 2023;31:601. https://doi.org/10.1007/s00520-023-08074-x.
43. Yang Y, Liu H, Chen Y, et al. Liquid biopsy on the horizon in immunotherapy of non-small cell lung cancer: current status, challenges, and perspectives. Cell Death Dis. 2023;14:230. https://doi.org/10.1038/s41419-023-05757-5.
44. Kasi PM, Lee JK, Pasquina LW, et al. Circulating tumor DNA enables sensitive detection of actionable gene fusions and rearrangements across cancer types. Clin Cancer Res. 2024;30:836–48. https://doi.org/10.1158/1078-0432.Ccr-23-2693.
45. Chaudhuri AA, Chabon JJ, Lovejoy AF, et al. Early detection of molecular residual disease in localized lung cancer by Circulating tumor DNA profiling. Cancer Discov. 2017;7:1394–403. https://doi.org/10.1158/2159-8290.Cd-17-0716.
46. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521:436–44. https://doi.org/10.1038/nature14539.
47. Xu Y, Hosny A, Zeleznik R, et al. Deep learning predicts lung cancer treatment response from serial medical imaging. Clin Cancer Res. 2019;25:3266–75. https://doi.org/10.1158/1078-0432.Ccr-18-2495.
48. Leng D, Zheng L, Wen Y, et al. A benchmark study of deep learning-based multi-omics data fusion methods for cancer. Genome Biol. 2022;23:171. https://doi.org/10.1186/s13059-022-02739-2.
49. Ballard JL, Wang Z, Li W, et al. Deep learning-based approaches for multi-omics data integration and analysis. BioData Min. 2024;17:38. https://doi.org/10.1186/s13040-024-00391-z.

Publisher's note