**Information Research Communications, 2027; 4(1):101-113.**
https://inforescom.org

Research Article

# Mapping the AI Revolution in Oncology: A Bibliometric Analysis of Trends, Collaborations, and Global Inequalities

**Manashree Chakraborty[1], Subhadip Mandal[2], Kanu Chakraborty[3,*], Swapan Kumar Patra[4], Simul Chakraborty[5]**

[1]Department of Medical Surgical Nursing, Apex College of Nursing, Varanasi, Uttar Pradesh, INDIA.
[2]Department of Library and Information Science, BHU, Varanasi, Uttar Pradesh, INDIA.
[3]Shreenivas Deshpande Library, IIT (BHU), Varanasi, Uttar Pradesh, INDIA.
[4]Department of Library & Information Science, Sidho-Kanho-Birsha University, Purulia, West Bengal, INDIA.
[5]Central Library, National Law University, Tripura, INDIA.

## ABSTRACT

**Background:** Artificial Intelligence (AI) is transforming oncology (cancer research) through advances in precision diagnostics, predictive modelling, and personalized therapies. Mapping the evolution of research in this rapidly expanding field is essential for identifying emerging trends, influential contributors, collaboration patterns, and research gaps. **Objectives:** This study aimed to analyze the global research landscape of AI applications in oncology using bibliometric methods, focusing on publication growth, core journals, author productivity, collaboration networks, citation impact, geographical distribution, and patterns of open-access article production. **Methodology:** A comprehensive dataset of publications related to AI and oncology was retrieved from the Scopus database. Quantitative analyses and visualization were performed using Biblioshiny, VOSviewer, and Python-based libraries. Bradford's Law was applied to examine the scattering of publications across journals, while Lotka's Law was used to assess author productivity patterns. The Gini coefficient and Lorenz curve were employed to evaluate inequality in the production of open-access publications across countries. Collaboration networks were further analyzed using network-based indicators, including average path length and degree centrality. **Results:** The field demonstrated rapid expansion, with a Compound Annual Growth Rate (CAGR) of 24.9%. Scientific Reports (Impact Factor=3.9, H-index=234, Q1) emerged as the leading journal. Wang was identified as the most productive author, contributing 0.05% of the total publications. The authorship collaboration network showed an average path length of 11.526, indicating relatively long-distance connectivity. The article by Esteva *et al.* (2017) was the most highly cited publication and achieved the highest altmetric attention score (2867). China was the leading contributing country, accounting for 4,399 publications (30.7%). **Conclusion:** Research on AI applications in oncology is expanding rapidly and demonstrates increasing global interest, interdisciplinary collaboration, and substantial scientific impact. The identified leading journals, authors, countries, and collaboration structures provide valuable insights into the development of the field. Despite its strong potential to transform cancer research and clinical practice, future progress will depend on addressing technical limitations as well as ethical, regulatory, and translational challenges.

**Keywords:** Altmetric Attention Score, Artificial Intelligence, Bibliometrics, Gini Coefficient, Lorenz Curve, Oncology.

**Correspondence:**
**Mr. Kanu Chakraborty**
Assistant Librarian, IIT (BHU),
Varanasi-221005, Uttar Pradesh, INDIA.
Email: kchakraborty.lib@iitbhu.ac.in

**Received:** 15-07-2026;
**Revised:** 03-09-2026;
**Accepted:** 22-10-2026.

## INTRODUCTION

The term 'Artificial Intelligence' was first coined by John McCarthy in 1950 (Cordeschi, 2007). AI is a structured computational system created by humans to replicate human actions, reasoning, and problem-solving. Artificial Intelligence (AI) is embedded across all sectors of modern society, influencing homes, hospitals, schools, workplaces, and in every application we use. Now it's not limited to the lab; it's become part of our daily lives. It reshapes the health sector by improving diagnosis, patient care, and treatment. According to the AIIR (Artificial Intelligence Index Report), the USA is investing $109.5 billion in AI, almost twelve times China's $9.3 billion.

Cancer is the most significant global health challenge that is exerting a tremendous physical, psychological, and economic burden on patients, their families, and the entire health system of the nation. According to the WHO, the world's second leading cause of death is cancer. Cancer is a broad group of diseases that can develop in any part of the body. Accordingly, the Global Cancer Observatory (GLOBOCAN) of the WHO, lung cancer

**DOI:** 10.5530/irc.4.1.9

**Copyright Information :**
Copyright Author (s) 2027 Distributed under Creative Commons CC-BY 4.0

**Publishing Partner :** Manuscript Technomedia. [www.mstechnomedia.com]

is the most common type of cancer for both sexes, followed by female breast cancer (World Health Organisation: WHO, 2024). As of 2022, China, the US, India, and Japan had the highest number of registered cancer cases worldwide (*World Cancer Research Fund*, 2025). China is the leading country in cancer research output, overtaking the USA for the first time in 2024, based on the 'Nature Index' (Nature Index, *2025).*

The remarkable strides in technologies such as artificial Intelligence, neural networks, machine learning, and modern medicine have improved cancer diagnosis, but in many cases, it remains an incurable disease. Artificial Intelligence has revolutionised cancer research, offering significant clinical opportunities. AI delivers significant outcomes by analysing large-scale data on human cancer reports. AI has shown remarkable potential across various types of cancer, such as lung, breast, cervical, brain, and colorectal cancers, by improving diagnostic accuracy and efficiency (Shirazi *et al*., 2024). Advanced AI automatically quantifies the radiographic pattern, detects, characterises, and monitors medical imaging data. AI-based Computer-Aided Detection (CADe) systems enhance radiological practices by reducing the likelihood of radiologist oversight. The rapid growth of AI and computer technology eliminates the need for a separate segmentation stage in the analysis of cancer lesions (Bi *et al*., 2019).

Open access is scientific information that is freely available and can be used without any barriers. Open access increases the visibility of research by providing broader, more impactful access. It reduces the gap between developed and developing nations, thereby advancing medical research that benefits both equally. The core principles of open access are 'more read, more cited.' It promotes cross-border and interdisciplinary collaboration, provides unrestricted access, encourages knowledge exchange, and supports innovation across diverse fields.

Information specialists and research analysts look for evidence that, on balance, a research field is progressing in a meaningful and impactful direction. Instead of examining a single study in isolation, bibliometricians collect and analyse data from all relevant publications, including articles, citations, keywords, authors, institutions, and collaboration networks. Through techniques such as bibliometric mapping and co-citation or co-word analysis, integrate findings from a large body of literature to create a broader, more comprehensive picture of a research domain. This approach provides a more precise understanding than any single paper can offer, revealing overall trends, influential contributors, knowledge gaps, and the evolving impact of a topic within the global scientific landscape.

## LITERATURE REVIEW

Several bibliometric studies examined various aspects of cancer research (Patra *et al* 2003, Xiong *et al*. 2023) conducted a bibliometric study on liver cancer and found that China had the highest number of publications, and that computed tomography was the most commonly used diagnostic tool. Multi-type data fusion, multimodal treatment analysis, and Single-cell sequencing technology are recent trends in AI for liver cancer (Ji *et al*., 2024; Xiong *et al*., 2023; Yin *et al*. 2022) evaluate a bibliometric study on AI in Pancreatic Cancer (PC) research. The author highlights that, after 2018, research trends have been rapidly growing, and the USA is leading the field. AI uses preoperative and diagnostic data to predict surgical risk. AI is used in cancer research to predict patient outcomes, assess chemotherapy responses, identify early-stage disease, analyze images, and evaluate different types of therapy. Gencer (2024) analyzes the research trends in AI-assisted lung cancer, which have grown significantly over the last five years, with China and the US consistently leading this field. AI is an essential tool in oncological research, used for diagnosis, prognosis, medical imaging, risk assessment, and more. The large language model in machine learning and deep learning identifies suitable patterns for different types of therapy by analyzing radiological, pathological, and genomic data. Ekinci and Tekedere's (2025) bibliometric study on AI in breast cancer diagnosis highlights the sharp growth of publications after 2019. The thematic analysis of this topic focuses on Convolutional Neural Networks (CNN), histopathology, and deep learning architecture. The integration of machine learning, deep learning, neural networks, radiomics, advanced image detection systems, and computer-assisted diagnostic systems enhanced AI's ability to detect patterns in data, predict patient risk, and identify genomic and histopathological patterns.

Wang *et al*. (2023) describe that machine learning and deep learning are the most common keywords in the AI-based cancer research. This study found that AI-related cancer research has grown significantly since 2017. Research priorities varied by region; for example, American scholars mainly focus on medical imaging, while Japanese scholars mainly focus on endoscopy for gastric and oesophageal cancer. Future progress will require applying AI to cancer research. Chen and Ding's (2025) bibliometric analysis of cancer risk factors provides a comprehensive framework of current trends in cancer research. Future cancer research relies on big data and artificial Intelligence. AI-assisted systems predict individual cancer risk. Bibliometric analysis, as a statistical method, reveals significant patterns by highlighting the most applicable traditional methods, such as most influential authors, countries, institutions, most common keywords, themes, citation analysis, and global collaboration in a field. Xu *et al*. (2025) use bibliometric methods to analyse trends in AI in oncological nursing. Major keyword clusters include: prediction models, AI-assisted diagnosis and imaging, genomics, patient-centred care, etc. The study suggested that future research will be needed for AI decision-making systems, symptom identification, and related areas.

More than 200 types of cancer have been identified, but they are divided into 5 main groups. These are Carcinoma, Sarcoma, leukemia, Lymphoma a myeloma, and Brain and spinal cord cancer (Cancer Research UK, 2024). The most common types of cancer are breast, lung, colon, rectum, prostate, skin, stomach, and liver (WHO, 2025). Several researchers, including Zhang *et al.* (2023), Zhao *et al.* (2024), Xia *et al.* (2024), Gencer (2024), Ding *et al.* (2021), Liang *et al.* (2021), Zhou *et al.* (2022), have conducted bibliometric and scientometric analyses in the broad and narrow fields of AI in lung cancer. Each of these studies revealed that the USA is the leading country of publication, followed by China. These studies highlight that diagnostic modeling, AI-assisted image analysis, and AI-driven clinical tools are emerging trends in AI applications for lung cancer research. Similarly, in Breast cancer, which is the most common cancer in women, Karger and Kureljusic (2023), Khairi *et al.* (2021), Shen *et al.* (2022), Syed and Khan (2022), Wu *et al.* (2025), Zhang *et al.* (2022) conducted bibliometric studies on AI in breast cancer. After 2019, AI-related research on liver cancer gained significant momentum. China is the top publication country, and the USA has the highest H-index (Xiong *et al.*, 2023; Zhuo *et al.*, 2024). Computed tomography is the most commonly used diagnostic tool, and Convolutional neural networks are the most widely used computational technique in AI-assisted liver cancer research (Lee *et al.*, 2023; Rezaee-Zavareh *et al.*, 2024). Shen *et al.* (2022) highlight that the current research hotspots in AI for prostate cancer are deep learning and multiparametric MRI. The total keywords are divided into three broader areas: AI in surgery, AI in non-surgical settings, and AI-predicted diagnosis. AI can predict therapies, such as immunotherapy and chemotherapy, that patients are more likely to respond to (Elemento *et al.*, 2021). The integration of AI and big data enhances the precision of oncology, enabling more accurate diagnoses and personalised treatment. Zhao *et al.* (2024) highlight a broader need for international collaboration on AI in cervical cancer research, as most studies have been conducted in developed nations. The integration of imaging data with genomics can create a comprehensive framework for cancer assessments using liquid biopsy biomarkers (Bi *et al.*, 2019).

### Objectives

This bibliometric study aims to achieve the following objectives:

- To identify the scientific publication growth and types of published documents.
- To analyse the scattering pattern of literature in this field by applying Bradford's Law.
- To investigate the author productivity patterns in this field by applying Lotka's Law.
- To provide a comprehensive analysis using keyword analysis, word frequency, top relevant sources, actor-level centrality, top-cited articles, and publication distribution among countries.
- To measure the inequality of open access articles production among countries by using the Gini Coefficient and the Lorenz Curve.

## METHODOLOGY

### Data sources

This bibliometric study used the Elsevier Scopus database**,** a multidisciplinary, comprehensive, and standardised bibliographic database. It has a broader research coverage than the open-source medical database PubMed. It was the most frequently used bibliographic database in bibliometric studies because it indexed more than 28000 journals.

### Data Search

All data were retrieved from Scopus on November 11, 2025. The search query string is defined as follows:

(Title ("artificial Intelligence" OR "machine learning" OR "deep learning" OR "neural network") AND Title ("oncology" OR "cancer" OR "tumor" OR "malignancy" OR "neoplasm" OR "metastasis" OR "carcinogenesis" OR "radiomics")) Refined By "Open Access" and only "English" language.

We intentionally limited the search terms to 'Oncology,' 'tumor,' 'neoplasm,' 'malignancy, and 'Cancer' in this study, rather than using a broad range of keywords such as types of cancer. This decision was made to avoid bias in search results, maintain specificity, reduce redundancy, avoid overlap, and ensure consistency. This study also filtered the search to English-language written documents.

Developing countries invested heavily in their AI, ML, and Deep learning labs and in advanced computational research, which often imposes restrictions on the sharing of research outputs. Here, we applied the 'Open Access' filter to ensure only free and publicly accessible research. The 'open access' filter was selected to provide unrestricted access, increase reproducibility, promote equitable participation, improve transparency, and ensure accurate data retrieval for the study.

### Data Analysis

We extracted 14,355 open-access records from Scopus, a product of Elsevier, and analysed the data using Biblioshiny, VOSviewer, Excel, and various Python libraries. Biblioshiny is an open-source R package developed by Massimo Aria Corrado Cuccurullo. Biblioshiny is a web-based, user-friendly tool for analysing and visualising scientific publications. It is used to analyse the annual scientific growth, top-cited articles, and most relevant sources. VOSviewer is a software application for data visualisation, developed by Nees Jan Van Eck and Ludo Waltman at Leiden University in the Netherlands. It is used to analyse the

keyword cluster. We used the Matplotlib package of Python for data visualisation. Gini coefficient analysis for unequal scientific production among countries using NumPy, Python library, and Curve visualization using Matplotlib. The maps of scientific output across countries were created using iipmaps.com (Free version).

## Scientific publication and trends

This bibliometric study covered 14206 publications from 1988 to 2025, retrieved from the Scopus database. Figure 1 shows that AI in oncological research has increased every year. The year-wise distribution of publications reveals the upward trend in the literature. Based on the distribution of publications, the literature growth can be divided into 3 phases. The first phase began *research* in the area (from 1988 to 2007). This phase showed very slow growth in publications. Followed by the Moderate Growth Phase (2008-2016) and finally the Rapid Growth Phase (2017-2025). The Compound Annual Growth Rate (CAGR) from 1988 to 2025 was approximately 24.9%, indicating rapid expansion of research activity in the field (Rakshit and Bardhan, 2023). This pattern indicates exponential growth in the literature, a common feature of emerging interdisciplinary fields (e.g., AI, ontology, knowledge graphs).

## Document types of publications

The publication types show that journal articles are the most prominent, with 11,313 publications (79.64%). Review papers are the second-largest category, with 1,426 documents (10.04%). Conference papers contribute 573 publications (4.03%). Other document types, such as editorials, letters, notes, and book chapters, collectively constitute a very small proportion. It can be concluded that journal articles are the primary outlet for this research area.

## Core journals

Tables 1 and 2 shows the journal name, the number of articles published in that journal, and the Cumulative Articles. According to Bradford's law, journals are categorised into one of three Bradford zones (Desai *et al.*, 2018). Bradford's Law suggests that a small core of journals (Zone 1) produces a significant portion of articles, followed by a larger set (Zone 2) producing roughly the same number of articles, and an even larger set (Zone 3) contributing the final third. 'Scientific Reports' is the top journal in this domain, with 730 articles published, followed by Frontiers in Oncology, with 698 articles and Cancers, with 510 articles.

## Bradford Curves

Figure 2 shows the Bradford Curve, where the X-axis represents the 'Cumulative Number of Journals' and the Y-axis represents the 'Cumulative Number of Articles'. This graph visually demonstrates Bradford's Law of Scattering. It typically shows a steep initial rise, indicating that a small number of core journals account for a one-third of articles. This segment represents the smallest number of journals (29) and accounts for approximately one-third of the total articles. These are the most productive journals in the field. Zone 2 represents a larger group of journals (174) that collectively account for the next one-third of the total articles. Zone 3 (Outer Zone) represents the largest group of journals (1785 journals), each publishing very few articles, yet together they account for the final one-third of the total articles. This zone typically contains a significantly higher number of journals than Zone 1 and Zone 2 combined. This plot makes it easier to visually identify the 'core' journals and how the productivity is distributed across the entire set of journals according to Bradford's Law.

## Top productive authors

The total number of unique authors in the dataset is 73,932. The top 20 authors by publication count and their percentage of total publications are given in Table 3. Wang is the most productive author, with 49 publications, accounting for 0.05% of total publications, followed by Kather, with 45 publications and Zhang, with 42 publications.

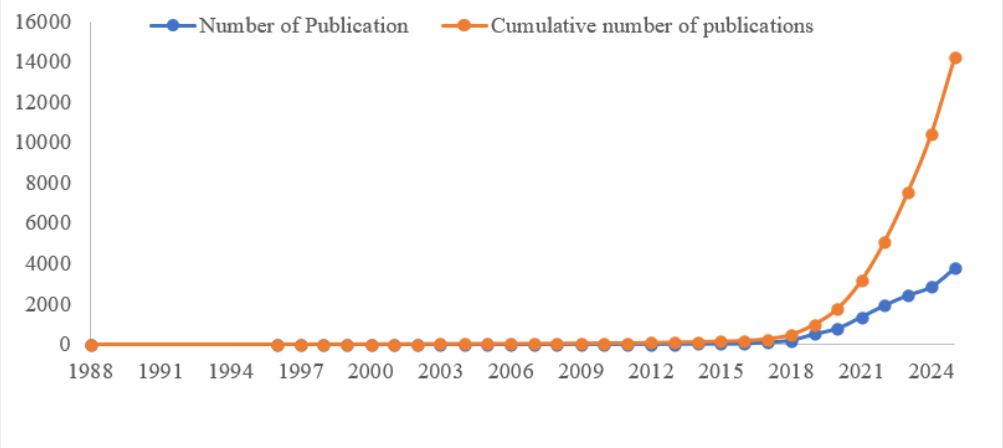



**Figure 1:** Growth of publications.

**Table 1: Document type.**

| Sl. No. | Document Type | No. of Document | Percentage |
|---|---|---|---|
| 1. | Article | 11313 | 79.64 |
| 2. | Book | 2 | 0.01 |
| 3. | Book chapter | 30 | 0.21 |
| 4. | Conference paper | 573 | 4.03 |
| 5. | Data paper | 9 | 0.06 |
| 6. | Editorial | 164 | 1.15 |
| 7. | Erratum | 344 | 2.42 |
| 8. | Letter | 138 | 0.97 |
| 9. | Note | 108 | 0.76 |
| 10. | Retracted | 65 | 0.46 |
| 11. | Review | 1426 | 10.04 |
| | Short survey | 32 | 0.23 |

## Lotka's law of scientific productivity

Publication frequencies (papers: authors), 1 publication: 58213 authors, 2 publications: 9730 authors, 3 publications: 3023 authors and so on. So, the author productivity pattern follows the typical inverse distribution described by *Lotka's Law*, where a large proportion of authors contribute only a single publication (Figure 3) (Kumar *et al.*, 1998). In the present dataset, 58,213 authors produced only one paper, while the number of authors decreases sharply with increasing publication frequency, reaching only 59 authors with ten publications. The calculated Lotka exponent (α=3.18 and calculated Constant (C): 84260.78 are higher than the classical value of 2, indicating a steeper decline in author productivity. It suggests that the field is characterised by a large number of occasional contributors and relatively few highly prolific researchers.

## Authorship collaboration network

Node 70,175, edges 83,167, average degree 2.37, average weighted degree 2.579, network diameter 38, connected component 3,418, graph density 0, average pathlength 11.526. The authors' collaboration networks have 70,267 nodes (Figure 4). This indicates a very large and complex network. In the network, the number of Edges is 83,259, indicating that many entities are connected, although the network is still relatively sparse, given the large number of nodes[1].

The network has an average degree of 2.37. It means that, on average, each node is connected to about two other nodes. This suggests that most entities have limited direct connections, a pattern typical of large bibliometric networks in which many participants collaborate only occasionally. An average weighted degree of 2.578 indicates that when connection strength is considered, nodes have slightly stronger interactions than the simple connection count suggests. A diameter of 38 indicates that the two most distant nodes in the network require 38 steps to reach each other through intermediate nodes. The presence of 3,418 connected components indicates that the network contains many isolated clusters, suggesting fragmented collaboration or thematic groups within the dataset. The graph density is very close to 0, which means that only a very small fraction of possible links actually exist, which is typical for large-scale networks. An average path length of 11.526 indicates that information or connections between nodes typically require about 11 intermediate links, suggesting relatively long-distance connectivity across the network.

Overall, the network exhibits the characteristics of a large, sparse, and highly fragmented structure. Although it contains a very large number of nodes, the low density and relatively small average degree indicate that most nodes have limited direct connections. The presence of numerous connected components suggests that the network is divided into many smaller clusters or communities. Additionally, the high network diameter and average path length indicate that interactions between distant nodes require several intermediate steps.

Such structural properties are commonly observed in large bibliometric networks, such as co-authorship, citation, or keyword networks, where collaboration and knowledge exchange primarily occur within smaller subgroups rather than across the entire network.

## Actor-level centrality analysis

The actor-level centrality measures highlight the main contributors within the networks using three metrics: degree, closeness, and betweenness (Table 4). In terms of degree centrality, HAGGENMÜLLER, SARAH, MARON, ROMAN CHRISTOPH R, and HEKLER, ACHIM are the most connected authors, indicating strong collaboration among them. HOSSEINZADEH, WAN, MOMEN-ROKNABADI, closeness centrality, suggesting that they are optimally positioned within the network to quickly

access or disseminate information to other actors. ZHANG (63551025.23), CHEN (56663494.27), and WANG (55086759.22) have the highest values, indicating their intermediate roles within different subgroups of the network.

## Keyword Frequency

The keyword frequency analysis reveals the dominant research themes in this domain (Table 5). The 'cancer' is the top-ranked word with 31554 (3.66%), highlighting that oncology is the primary focus in this domain. Another popular keyword is learning, with a frequency of 28903, followed by Study with 16638.

The presence of the keyword "human" (1.50%) further indicates that a substantial portion of the studies are focused on human-related clinical or biomedical data. Overall, the distribution of keywords reveals a strong interdisciplinary trend combining healthcare, artificial Intelligence, and data-driven research methodologies.

## Keywords cluster

In Figure 5, we applied keyword and co-occurrence analyses using VOSviewer. This bibliometric approach analyses AI research trends in oncological studies by examining the frequency of article keywords. The map was structured into several clusters and dominant research topics, with their keyword occurrences

**Table 2: Core journals.**

| Sl. No. | Journal | Articles | Cumulative Articles | Cumulative Journals | Zone |
|---|---|---|---|---|---|
| 1. | Scientific Reports | 730 | 730 | 1 | 1 |
| 2. | Frontiers in Oncology | 698 | 1428 | 2 | 1 |
| 3. | Cancers | 510 | 1938 | 3 | 1 |
| 4. | Diagnostics | 310 | 2248 | 4 | 1 |
| 5. | IEEE Access | 257 | 2505 | 5 | 1 |
| 6. | PLOS ONE | 175 | 2680 | 6 | 1 |
| 7. | BMC Cancer | 149 | 2829 | 7 | 1 |
| 8. | Applied Sciences (Switzerland) | 134 | 2963 | 8 | 1 |
| 9. | Frontiers in Immunology | 122 | 3085 | 9 | 1 |
| 10. | Frontiers in Medicine | 110 | 3195 | 10 | 1 |
| 11. | BMC Medical Imaging | 97 | 3292 | 11 | 1 |
| 12. | Discover Oncology | 94 | 3386 | 12 | 1 |
| 13. | BioMed Research International | 87 | 3473 | 13 | 1 |
| 14. | Heliyon | 87 | 3560 | 14 | 1 |
| 15. | Journal of Clinical Medicine | 86 | 3646 | 15 | 1 |
| 16. | JCO Clinical Cancer Informatics | 82 | 3728 | 16 | 1 |
| 17. | Cancer Medicine | 81 | 3809 | 17 | 1 |
| 18. | International Journal of Molecular Sciences | 80 | 3889 | 18 | 1 |
| 19. | npj Precision Oncology | 80 | 3969 | 19 | 1 |
| 20. | European Radiology | 79 | 4048 | 20 | 1 |
| 21. | International Journal of Advanced Computer Science and Applications | 78 | 4126 | 21 | 1 |
| 22. | Journal of Physics: Conference Series | 75 | 4201 | 22 | 1 |
| 23. | Quantitative Imaging in Medicine and Surgery | 74 | 4275 | 23 | 1 |
| 24. | Journal of Healthcare Engineering | 74 | 4349 | 24 | 1 |
| 25. | Procedia Computer Science | 73 | 4422 | 25 | 1 |
| 26. | Sensors | 72 | 4494 | 26 | 1 |
| 27. | Journal of Cancer Research and Clinical Oncology | 72 | 4566 | 27 | 1 |
| 28. | Bioengineering | 69 | 4635 | 28 | 1 |
| 29. | Medical Physics | 65 | 4700 | 29 | 1 |

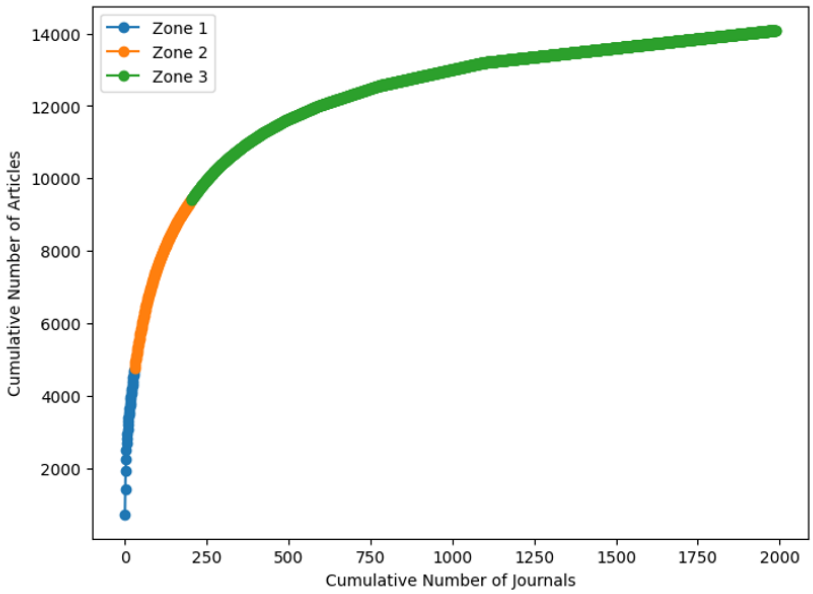



**Figure 2:** Bradford Curves.

and conceptual relatedness. The centre green colour of the map is dominated by controlled study, prediction, adult, male, female, and cohort study terms. This keyword theme highlights the population and clinically based oncological studies. The blue clusters on the map show the AI's adoption of oncological research, highlighting keywords such as deep learning, neural networks, image segmentation, decision trees, brain tumor, and diagnosis. The red color focused on computational biology and molecular oncology by highlighting the keywords such as prognosis, cancer prognosis, genetics, metabolism, neoplasms, colorectal cancer, etc. Overall, this map visualises a multidisciplinary approach to using modern AI technologies for cancer detection, diagnosis, and prediction of patient's recovery.

## Top-cited articles

Table 6 presents the top 10 most cited articles in AI oncology research. We compared traditional citations with modern altmetric scores. Traditional citation counts indicate the long-term academic impact, whereas Altmetric Attention Score (AAS) indicates the media attention a research output receives. High AAS indicates high media attention, whereas low AAS indicates less social sharing. The AAS was collected from Dimensions AI, the world's largest collection of free linked research data, a product of Digital Science. This comparative analysis shows that some highly cited articles received limited attention on social media; for example, HAVAEI's (2017) article is highly cited (2795) but has a low AAS (47). ESTEVA (2017) is the top-cited article with the highest altmetric attention score in this field, followed by HAVAEI (2017) and EHTESHAMI BEJNORDI (2017), which have 2679 citations and 742 AAS. COUDRAY's (2018) articles have a lower citation count (2246) than the top three articles, but they exhibit a notably high altmetric score of 1103.

## Publication distribution among Countries

Countries' publication distribution is a critical aspect of bibliometric study. Figure 6 shows all countries in open-access research publications on AI in oncological studies. The world map shows China leads with 4399 publications (30.7%), followed by the USA with 1710 publications (11.9%) and India with 1016 articles (7.1%). Overall, together these three countries account for half of the publications from all 115 countries.

## Countries Scientific Production

We used the Gini Coefficient to measure the distribution of scientific output across countries by comparing it to the area under the perfect equality line, also called the 45-degree line. The Gini coefficient measures inequality in a distribution. It was developed by Italian statistician Corrado Gini in 1913. It was measured on a scale from 0 to 1, with 0 indicating perfect equality and 1 indicating inequality. The Gini coefficient is mainly used to measure income and financial inequality in a society. However, it is frequently applied across various fields, such as measuring inequality in global research outcomes (Cho, 2020), assessing journal citation inequality (Chatterjee *et al.*, 2016), and examining inequality in science journals and academic institutions (Ghosh *et al.*, 2014).

The Lorenz curve is a graphical representation of the distribution of research output across countries. It was developed by Max O. Lorenz to represent income inequality. The Gini coefficient is defined as the ratio of the areas under the Lorenz curve. The yellow line in the Figure 7 indicates the perfect equality line, and the blue curve shows the perfect inequality. Gini Coefficient: 0.8467

The Gini coefficient is calculated using this standard formula:

$$G = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i - x_j|}{2n^2\bar{x}}$$

Here, $x_i$ represents the number of publications by country I, n means the total number of countries, and $\bar{x}$ represents the mean of publications.

The Gini coefficient for the distribution of countries' publications is g=0.8467. For the total of 14355 open-access publications from 120 countries on AI applications in oncological studies, the Gini coefficient is 0.87, indicating an extreme level of country inequality in publication output. There are specific countries with large volumes of documents, such as China (4399), the USA (1710), India (1016), South Korea (555), Italy (416), Germany

**Table 3:** **Top Authors.**

| Sl. No. | Name of the author | Publication Count | Percentage |
|---|---|---|---|
| 1. | Wang, Wei | 49 | 0.05% |
| 2. | Kather, Jakob Nikolas | 45 | 0.04% |
| 3. | Zhang, Wei | 42 | 0.04% |
| 4. | Tian, Jie | 38 | 0.04% |
| 5. | Zhang, Yang | 33 | 0.03% |
| 6. | Wang, Ying | 32 | 0.03% |
| 7. | Wang, Jing | 31 | 0.03% |
| 8. | Aerts, Hugo J.W.L. | 31 | 0.03% |
| 9. | Liu, Zaiyi | 31 | 0.03% |
| 10. | Li, Jing | 31 | 0.03% |
| 11. | Zhang, Yu | 30 | 0.03% |
| 12. | Turkbey, Ismail Baris | 28 | 0.03% |
| 13. | Brinker, Titus Josef | 27 | 0.03% |
| 14. | Zaidi, Habib N. | 26 | 0.02% |
| 15. | Wang, Yu | 26 | 0.02% |
| 16. | Zhang, Lei | 25 | 0.02% |
| 17. | Zhang, Jing | 24 | 0.02% |
| 18. | Fuller, Clifton David | 23 | 0.02% |
| 19. | Zhang, Tao | 23 | 0.02% |
| 20. | Pearson, Alexander T. | 22 | 0.02% |

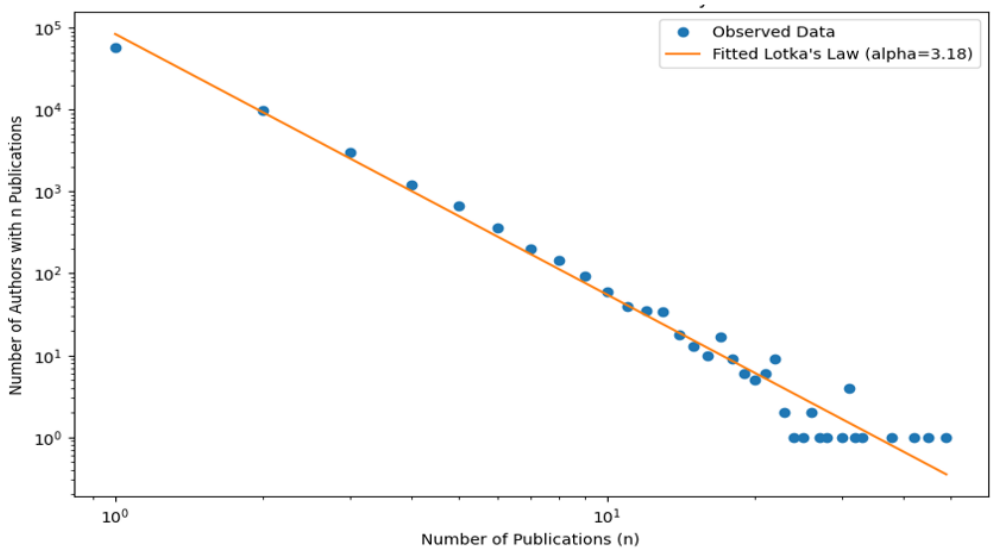



**Figure 3:** Lotka's law of Scientific productivity.

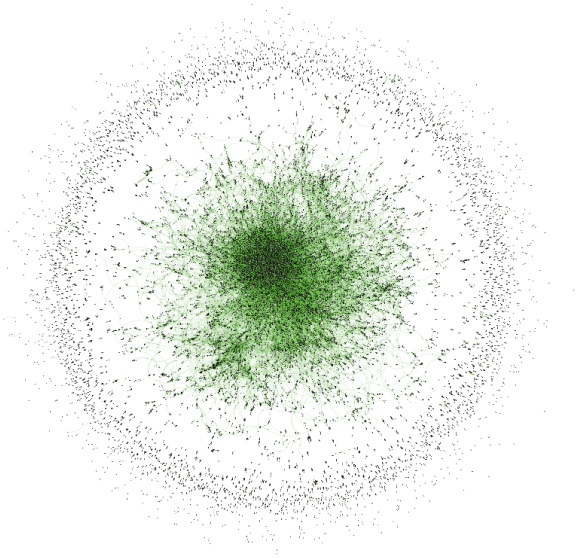


**Figure 4:** Authorship Collaboration Networks.

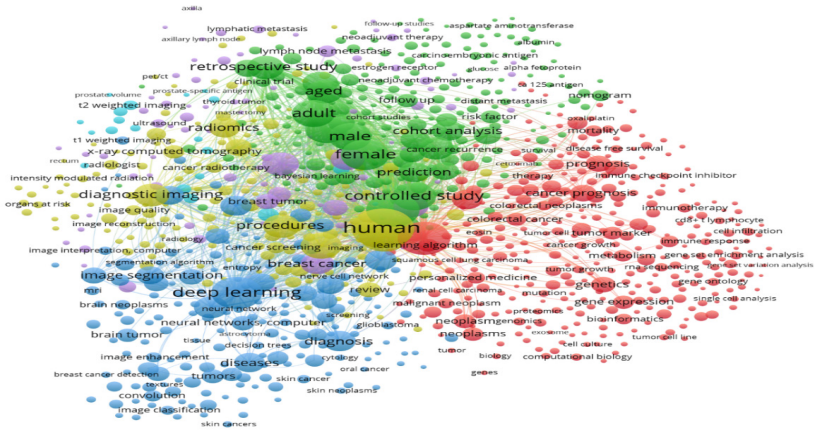



**Figure 5:** Keyword Clusters.

**Table 4: Actor-level centrality measures of the top 10 actors.**

| Label | Degree | Label | Closeness | Label | Betweenness |
|---|---|---|---|---|---|
| Haggenmüller, Sarah | 119 | Hosseinzadeh, Farideh | 1 | Zhang, Wei | 63551025.23 |
| Maron, Roman Christoph R. | 118 | Wan, Nathan | 1 | Chen, Hao | 56663494.27 |
| Hekler, Achim | 116 | Momen-Roknabadi, Amir | 1 | Wang, Jing | 55086759.22 |
| Way, Gregory P. | 104 | Palmer, Rebecca Natalie | 1 | Ehteshami Bejnordi, Babak | 44774452.02 |
| Anania, Gabriele | 101 | Heo, Joonnyung | 1 | Saltz, Joel Haskin | 36538953.77 |
| Saltz, Joel Haskin | 101 | Cardoso, Marcella Regina | 1 | Wang, Wei | 33389989.41 |
| Jiao, Wei | 100 | Croce, Giancarlo | 1 | Thagaard, Jeppe | 32830026.37 |
| Khadhouri, Sinan | 99 | Oura, Hirotaka | 1 | Wang, Yu | 30963098.33 |
| Wang, Joshua M. | 99 | Statsenko, Yauhen | 1 | Jiang, Yuming | 30252781.12 |
| Thagaard, Jeppe | 99 | Chauvie, Stéphane | 1 | Zhang, Yang | 30123321.82 |

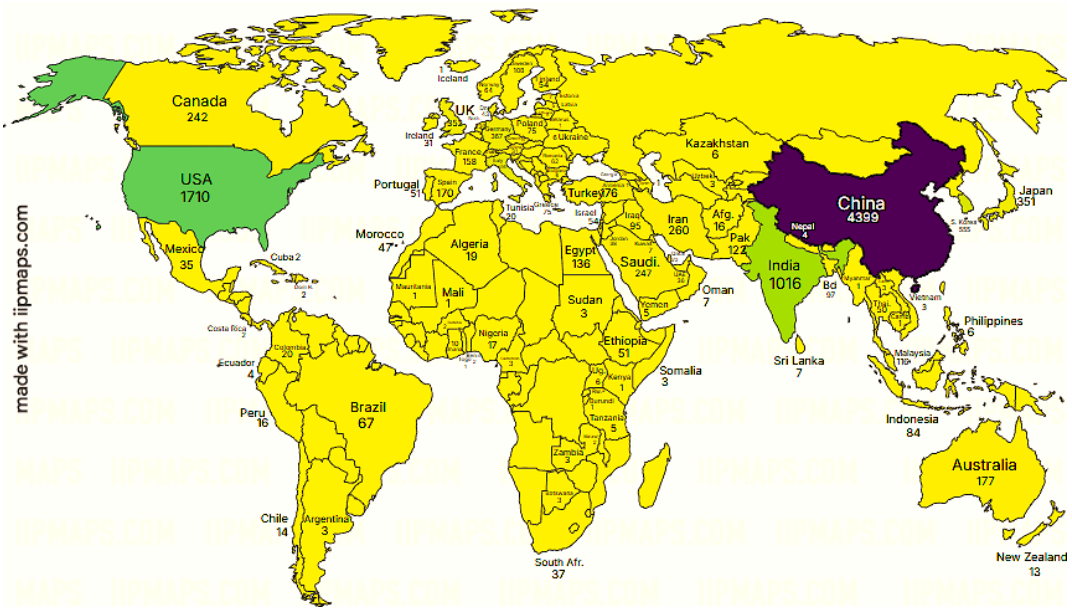



**Figure 6:** Publication Distributions across Countries.

**Table 5:** **Keyword frequency.**

| Rank | Keyword | Frequency | Percentage |
|---|---|---|---|
| 1 | Cancer | 31554 | 3.66% |
| 2 | Learning | 28903 | 3.36% |
| 3 | Study | 16638 | 1.93% |
| 4 | Machine | 15621 | 1.81% |
| 5 | Tumor | 14084 | 1.64% |
| 6 | Human | 12903 | 1.50% |
| 7 | Deep | 12597 | 1.46% |
| 8 | Imaging | 10886 | 1.26% |
| 9 | Neural | 10181 | 1.18% |
| 10 | Analysis | 9708 | 1.13% |
| 11 | Aged | 8851 | 1.03% |
| 12 | Artificial | 8782 | 1.02% |
| 13 | Clinical | 8508 | 0.99% |
| 14 | Breast | 8384 | 0.97% |
| 15 | Female | 8014 | 0.93% |
| 16 | Diagnostic | 7814 | 0.91% |
| 17 | Network | 7793 | 0.90% |
| 18 | Image | 7704 | 0.89% |
| 19 | Cell | 7508 | 0.87% |
| 20 | Intelligence | 7148 | 0.83% |

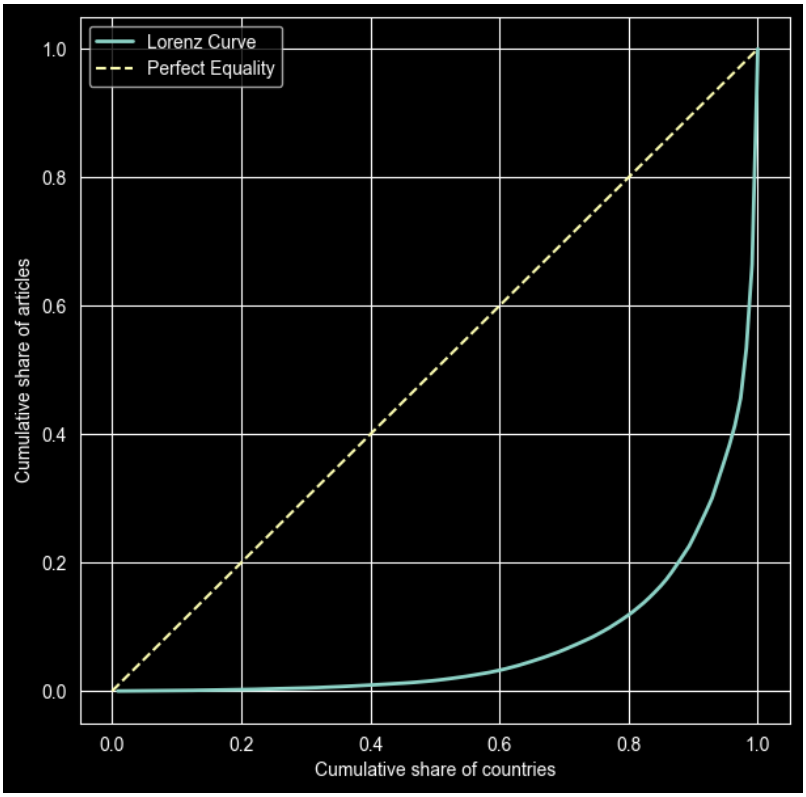



**Figure 7:** Lorenz Curve.

**Table 6: Highly cited articles.**

| Paper | DOI | Total Citations | TC per Year | Normalized TC | Altmetric Score |
|---|---|---|---|---|---|
| Esteva, (2017). Nature | 10.1038/nature21056 | 10874 | 1,087.40 | 39.66 | 2867 |
| Havaei, (2017). Med Image Anal | 10.1016/j.media.2016.05.004 | 2795 | 279.50 | 10.19 | 47 |
| Ehteshami Bejnordi, 2017, Jama | 10.1001/jama.2017.14585 | 2679 | 267.90 | 9.77 | 742 |
| Kourou, (2015). Comput Struct BiotechnoL J | 10.1016/j.csbj.2014.11.005 | 2612 | 217.67 | 20.53 | 115 |
| Khan, 2001, Nat Med | 10.1038/89044 | 2351 | 90.42 | 1.98 | 38 |
| Coudray, (2018). Nat Med | 10.1038/s41591-018-0177-5 | 2246 | 249.56 | 19.80 | 1103 |
| BI, 2019, Ca Cancer J Clin | 10.3322/caac.21552 | 1622 | 202.75 | 20.96 | 90 |
| Cruz, (2006). Cancer informatics | 10.1177/117693510600200030 | 1073 | 51.10 | 1.00 | 140 |
| Kather, (2019). Nat Med | 10.1038/s41591-019-0462-y | 1062 | 132.75 | 13.72 | 280 |
| Araújo, 2017, Plos One | 10.1371/journal.pone.0177544 | 863 | 86.30 | 3.15 | 27 |

(367), Japan (351), and the UK (352), while other countries contribute only marginally.

## CONCLUSION

AI has changed every field of research, especially in medicine. The medical discipline that is highly dependent on image analysis for disease identification is where AI is rapidly evolving. Machine learning, deep learning, neural networks, and Large Language Models (LLMs) help cancer research by analysing large amounts of medical data, enabling drug discovery, predicting treatment outcomes, and identifying new approaches to cancer diagnosis. The scientific mapping of AI in oncological studies systematically analyzes research publication trends, publication by country, the top journals in this field, and the top-cited articles. Also applying modern bibliometric approaches, such as topic modelling with BERTopic, keyword analysis, topic heatmaps, and the Gini Coefficient to measure the inequality in the distribution of scientific output across countries.

Research on AI in Oncological studies began after 2000, but it entered a period of rapid development in 2016. The segmented linear growth analysis for the period 2017 to 2025 yielded $R^{2=}0.9658$. The $R^2$ value shows an excellent model fit, indicating that the model captures the data from 2017 to 2025 very well. The highest publication, 3079, came in 2025. China is the most productive country, followed by the USA and India. Scientific Reports (IF=3.9, H-index=234, Q1) is the most preferred periodical among researchers, followed by Frontiers in Oncology (IF=3.3, H-index=137, Q2) and Cancers (IF=4.4, H-index=158, Q2). However, Frontiers in Immunology has the highest impact factor (5.9, Q1) of any journal on this list. The "ESTEVA (2017). NATURE" is the top-cited article with 10874 citations and 2867 Altmetric Attention Score (AAS). The keyword cluster focuses on AI-based clinical oncological studies, highlighting keywords such as human, controlled study, prediction, cancer recurrence, and deep learning.

Bradford's Curve typically shows a steep initial rise, indicating that a small number of core journals account for a large share of articles. Lotka's Law suggests that there are many occasional contributors and relatively few highly prolific researchers. The authorship collaboration network exhibits the characteristics of a large, sparse, and highly fragmented structure.

The Gini coefficient (G=0.8467) indicates perfect inequality in the distribution of scientific production across countries. The Lorenz curve in the Figure shows significant inequality in the distribution of scientific output across countries. The curve lies far below the line of perfect equality (45-degree line), indicating that research production is highly concentrated among a limited number of countries.

AI has strong potential to support cancer diagnosis by performing accurate data analysis of large medical datasets, image analysis, and providing diagnosis suggestions based on prior data. Machine learning, a part of AI, helps recognise images, pathological data, and genomic data. This AI contribution helps in faster clinical decision-making. Overall, AI demonstrates strong potential to transform cancer research and clinical practice, provided that technical, ethical, and regulatory challenges are carefully addressed.

## ABBREVIATIONS

**AI:** Artificial Intelligence; **ML:** Machine Learning; **DL:** Deep Learning; **CNN:** Convolutional Neural Network; **LLM:** Large Language Model; **CAGR:** Compound Annual Growth Rate; **AAS:** Altmetric Attention Score; **CADe:** Computer-Aided Detection; **GLOBOCAN:** Global Cancer Observatory; **WHO:** World Health Organization; **PC:** Pancreatic Cancer.

## CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

## REFERENCES


Bi, W. L., Hosny, A., Schabath, M. B., Giger, M. L., Birkbak, N. J., *et al.* (2019). Artificial Intelligence in cancer imaging: clinical challenges and applications. *CA: a cancer journal for clinicians*, *69*(2), 127-157.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993-1022. https://doi.org/10.5555/944919.944937

Cancer Research UK. (2024, November 21). *Understanding statistics in cancer research*. https://www.cancerresearchuk.org/about-cancer/what-is-cancer/understanding-statistics-in-cancer-research

Chatterjee, A., Ghosh, A., & Chakrabarti, B. K. (2016). Universality of citation distributions for academic institutions and journals. *PLoS ONE*, *11*(1), e0146762. https://doi.org/10.1371/journal.pone.0146762

Chen, S., & Ding, Y. (2025). A bibliometric analysis of the risk factors of cancer. *Genes, Chromosomes, and Cancer*, *64*(1), e70019. https://doi.org/10.1002/gcc.70019

China overtakes the United States in cancer research output. (2025). *Nature*, *640*(8060), S65. https://doi.org/10.1038/d41586-025-01154-4

Cho, J. (2020). Analysis of global inequality in research outcome using the Gini coefficient. *Performance Measurement and Metrics*, *22*(1), 25-37. https://doi.org/10.1108/pmm-05-2020-0027

Cordeschi, R. (2007). AI TURNS FIFTY: REVISITING ITS ORIGINS. *Applied Artificial Intelligence*, *21*(4-5), 259-279. https://doi.org/10.1080/08839510701252304

Desai, N., Veras, L., & Gosain, A. (2018). Using Bradford's law of scattering to identify the core journals of pediatric surgery. *Journal of Surgical Research*, *229*, 90-95. https://doi.org/10.1016/j.jss.2018.03.062

Ding, H., Wu, C., Liao, N., Zhan, Q., Sun, W., Huang, Y., *et al.* (2021). Radiomics in Oncology: A 10-Year Bibliometric Analysis. *Frontiers in Oncology*, *11*, 689802. https://doi.org/10.3389/fonc.2021.689802

Ekinci, B., & Tekedere, H. (2025). Bibliometric analysis of research on artificial Intelligence applications in breast cancer diagnosis. *Technology and Health Care*, 9287329251362602. https://doi.org/10.1177/09287329251362602

Elemento, O., Leslie, C., Lundin, J., & Tourassi, G. (2021). Artificial Intelligence in cancer research, diagnosis and therapy. *Nature Reviews Cancer*, *21*(12), 747-752.

Gencer, A. (2024). Bibliometric analysis and research trends of artificial Intelligence in lung cancer. *Heliyon*, *10*(2), e24665. https://doi.org/10.1016/j.heliyon.2024.e24665

Ghosh, A., Chattopadhyay, N., & Chakrabarti, B. K. (2014). Inequality in societies, academic institutions and science journals: Gini and k-indices. *Physica a Statistical Mechanics and Its Applications*, *410*, 30-34. https://doi.org/10.1016/j.physa.2014.05.026

*Global cancer data by country | World Cancer Research Fund*. (2025, July 28). World Cancer Research Fund. https://www.wcrf.org/preventing-cancer/cancer-statistics/global-cancer-data-by-country/

Ji, Y., An, Q., Wen, X., Xu, Z., Xia, Z., *et al.* (2024). Liver cancer from the perspective of single-cell sequencing: a review combined with bibliometric analysis. *Journal of Cancer Research and Clinical Oncology*, *150*(6), 316. https://doi.org/10.1007/s00432-024-05855-7

Karger, E., & Kureljusic, M. (2023). Artificial Intelligence for Cancer Detection-A bibliometric analysis and Avenues for Future Research. *Current Oncology*, *30*(2), 1626-1647. https://doi.org/10.3390/curroncol30020125

Khairi, S., Bakar, M., Alias, M., Bakar, S., Liong, C., *et al.* (2021). Deep Learning on Histopathology Images for Breast Cancer Classification: A Bibliometric analysis. *Healthcare*, *10*(1), 10. https://doi.org/10.3390/healthcare10010010

Kumar, S., Sharma, P., & Garg, K. (1998). Lotka's law and institutional productivity. *Information Processing & Management*, *34*(6), 775-783. https://doi.org/10.1016/s0306-4573(98)00027-2

Lee, K., Hwang, J. W., Sohn, H. J., Suh, S., & Kim, S. (2023). A systematic review of progress on hepatocellular carcinoma research over the past 30 years: a machine-learning-based bibliometric analysis. *Frontiers in Oncology*, *13*, 1227991. https://doi.org/10.3389/fonc.2023.1227991

(Nature Index, 2025)- China overtakes the United States in cancer research output. (2025). Nature, 640(8060), S65. https://doi.org/10.1038/d41586-025-01154-4

Patra, S. K., & Bhattacharya, P. (2003). Bibliometric Study of Cancer Research in India. DESIDOC Journal of Library & Information Technology, 25(2). https://doi.org/10.14429/djlit.25.2.3650

Rakshit, D., & Bardhan, D. (2023b). *CAGR: Compound annual growth rate* [Dataset]. https://doi.org/10.32614/cran.package.cagr

Rezaee-Zavareh, M. S., Kim, N., Yeo, Y. H., Kim, H., Lee, J. M., *et al.* (2024). Artificial Intelligence in liver cancer research: a scientometrics analysis of trends and topics. *Frontiers in Oncology*, *14*, 1355454. https://doi.org/10.3389/fonc.2024.1355454

Shen, Z., Hu, J., Wu, H., Chen, Z., Wu, W., Lin, J., Xu, Z., Kong, J., & Lin, T. (2022). Global research trends and foci of artificial Intelligence-based tumor pathology: a scientometric study. *Journal of Translational Medicine*, *20*(1), 409. https://doi.org/10.1186/s12967-022-03615-0

Shen, Z., Wu, H., Chen, Z., Hu, J., Pan, J., Kong, J., & Lin, T. (2022). The Global Research of Artificial Intelligence on Prostate Cancer: A 22-Year Bibliometric Analysis. *Frontiers in Oncology*, *12*, 843735. https://doi.org/10.3389/fonc.2022.843735

Syed, A. H., & Khan, T. (2022). Evolution of research trends in artificial Intelligence for breast cancer diagnosis and prognosis over the past two decades: A bibliometric analysis. *Frontiers in Oncology*, *12*, 854927. https://doi.org/10.3389/fonc.2022.854927

Wang, S., Chen, G., Zhong, X., Lin, T., Shen, Y., Fan, X., & Cao, L. (2023). Global development of artificial Intelligence in the cancer field: a bibliometric analysis ranging from 1983 to 2022. *Frontiers in Oncology*, *13*, 1215729. https://doi.org/10.3389/fonc.2023.1215729

World Health Organization: WHO. (2024, February 1). Global cancer burden growing, amidst mounting need for services. NO. https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing--amidst-mounting-need-for-services#:~:text=Data%20covers%20185%20countries%20and,of%20the%20total%20new%20cases.

World Health Organization: WHO. (2025, February 3). *Cancer*. https://www.who.int/news-room/fact-sheets/detail/cancer

Wu, X., Xia, Y., Lou, X., Huang, K., Wu, L., & Gao, C. (2025). Decoding breast cancer imaging trends: the role of AI and radiomics through bibliometric insights. *Breast Cancer Research*, *27*(1), 29. https://doi.org/10.1186/s13058-025-01983-1

Xia, Y., Zhong, S., Lou, X., Hua, M., Wu, L., & Gao, C. (2024). Research trends of artificial Intelligence and radiomics in lung cancer: a bibliometric analysis. *Quantitative Imaging in Medicine and Surgery*, *14*(12), 8771-8784. https://doi.org/10.21037/qims-24-1316

Xiong, M., Xu, Y., Zhao, Y., He, S., Zhu, Q., *et al.* (2023). Quantitative analysis of artificial Intelligence on liver cancer: A bibliometric analysis. *Frontiers in Oncology*, *13*, 990306. https://doi.org/10.3389/fonc.2023.990306

Xu, M., Chen, Y., Wu, T., Chen, Y., Zhuang, W., Huang, Y., & Chen, C. (2025). Global research trends in the application of artificial Intelligence in oncology care: a bibliometric study. *Frontiers in Oncology*, *14*, 1456144. https://doi.org/10.3389/fonc.2024.1456144

Yin, H., Zhang, F., Yang, X., Meng, X., Miao, Y., Hussain, M. S. N., Yang, L., & Li, Z. (2022). Research trends of artificial Intelligence in pancreatic cancer: a bibliometric analysis. *Frontiers in Oncology*, *12*, 973999. https://doi.org/10.3389/fonc.2022.973999

Zhang, J., Zhu, H., Wang, J., Chen, Y., Li, Y., Chen, X., Chen, M., Cai, Z., & Liu, W. (2023). Corrigendum: Machine learning in non-small cell lung cancer radiotherapy: a bibliometric analysis. *Frontiers in Oncology*, *13*, 1197217. https://doi.org/10.3389/fonc.2023.1197217

Zhang, Y., Yu, C., Zhao, F., Xu, H., Zhu, C., & Li, Y. (2022). Landscape of Artificial Intelligence in Breast Cancer (2000-2021): A bibliometric analysis. *Frontiers in Bioscience-Landmark*, *27*(8), 224. https://doi.org/10.31083/j.fbl2708224

Zhao, Z., Hu, B., Xu, K., Jiang, Y., Xu, X., & Liu, Y. (2024). A quantitative analysis of artificial Intelligence research in cervical cancer: a bibliometric approach utilizing CiteSpace and VOSviewer. *Frontiers in Oncology*, *14*, 1431142. https://doi.org/10.3389/fonc.2024.1431142

Zhao, Z., Xu, K., Jiang, Y., Xu, X., & Liu, Y. (2024). The role of artificial Intelligence in immune checkpoint inhibitor research: A bibliometric analysis. *Human Vaccines & Immunotherapeutics*, *20*(1), 2429893. https://doi.org/10.1080/21645515.2024.2429893

Zhou, W., Deng, Z., Liu, Y., Shen, H., Deng, H., & Xiao, H. (2022). Global Research Trends of Artificial Intelligence on Histopathological Images: A 20-Year Bibliometric Analysis. *International Journal of Environmental Research and Public Health*, *19*(18), 11597. https://doi.org/10.3390/ijerph191811597

Zhuo, E., Yang, W., Wang, Y., Tang, Y., Wang, W., Zhou, L., Chen, Y., Li, P., Chen, B., Gao, W., & Liu, W. (2024). Global trends in machine learning applied to clinical research in liver cancer: Bibliometric and visualization analysis (2001-2024). *Medicine*, *103*(49), e40790. https://doi.org/10.1097/md.0000000000040790

**Cite this article:** Chakraborty M, Mandal S, Chakraborty K, Patra SK, Chakraborty S. Mapping the AI Revolution in Oncology: A Bibliometric Analysis of Trends, Collaborations, and Global Inequalities. Info Res Com. 2027;4(1):101-13.