

A Bibliometric Analysis of Artificial Intelligence Applications in Infectious Diseases: A Decade of Research (2016–2025)

Çağlar Irmak

caglar_irmak08@hotmail.com

Izmir Tepecik Eğitim ve Araştırma Hastanesi

Ahmet Furkan Süner

Düzce Üniversitesi

İlkay Akbulut

Izmir Tepecik Eğitim ve Araştırma Hastanesi

Sabri Atalay

Izmir Tepecik Eğitim ve Araştırma Hastanesi

Research Article

Keywords: Artificial intelligence, infectious diseases, bibliometric analysis, machine learning, antimicrobial resistance, sepsis

Posted Date: July 2nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10131075/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Infection on August 13th, 2026. See the published version at <https://doi.org/10.1007/s15010-026-02934-4>.

Abstract

Purpose

Artificial intelligence (AI) is increasingly being applied in the field of infectious diseases, yet the global trajectory of this research domain has not been comprehensively characterized. This study aimed to map the evolution of scientific output on AI applications in infectious diseases over the decade 2016–2025, characterizing thematic clusters, collaboration networks, and emerging research trends to guide future clinical investigations.

Methods

A systematic bibliometric analysis was performed using the Web of Science Core Collection (WoSCC). The query combined AI-related terms (machine learning, deep learning, neural networks, large language models, and allied concepts) restricted to the WoSCC category of "Infectious Diseases" and an English-language filter, yielding 1,252 original articles published between January 2016 and December 2025. Bibliometric computations and visualizations were conducted using the Bibliometrix R package, VOSviewer, and Scimago Graphica. Keyword co-occurrence network analysis and temporal trend mapping were employed to identify thematic clusters and emerging research foci.

Results

The 1,252 articles were contributed by 9,160 authors from 124 countries across 2,850 institutions and published in 125 journals. Annual output grew more than 30-fold between 2016 and 2025, with 84.7% of all publications appearing in the final five years. The United States and China were the most productive nations and collectively dominated international collaboration networks. Harvard University was the leading institution. Keyword network analysis identified nine distinct thematic clusters. Recent trend analyses reveal a significant shift towards clinical applications, specifically highlighting antimicrobial resistance (AMR) surveillance and early sepsis prediction as dominant research hotspots.

Conclusion

Scientific production on AI applications in infectious diseases has expanded exponentially over the past decade, catalyzed principally by the COVID-19 pandemic. Sepsis management and AMR represent the most prominent areas for future investigation. These findings indicate that AI has a growing potential to become an integral component of infectious disease practice.

1. Introduction

Infectious diseases continue to pose a threat to global public health by emerging and re-emerging pathogens. Over the past 25 years, various epidemics and pandemics have affected many countries, resulting in substantial morbidity and mortality. Lower respiratory infections were responsible for approximately 2.50 million deaths globally in 2023, persisting as the world's leading infectious cause of death [1]. Compounding this burden is the escalating crisis of antimicrobial resistance (AMR), which was associated with an estimated 4.1 million deaths in 2021 alone, with projections warning of up to 10 million annual deaths by 2050 [2]. These clinical challenges highlight the urgent need for transformative approaches to the surveillance, diagnosis and clinical management of infectious diseases. [3].

Artificial intelligence (AI) is a field of computer science that aims to replicate the functions of human intelligence using algorithms. AI can process, analyze, and report on big data using various techniques such as machine learning (ML), deep learning (DL), and natural language processing (NLP) [4]. As in many other fields, AI has transformed healthcare in recent years. It has revealed the potential to improve patient care and quality of life by integrating it into healthcare systems [5, 6]. It offers significant advantages, including rapid and accurate diagnosis, support for clinical decision-making, reduced costs, and time savings while minimizing human errors [7]. Within infectious diseases specifically, AI is used to extend the spectrum from ML-based outbreak forecasting and AMR surveillance to DL-driven radiological screening and NLP-mediated early sepsis prediction [8]. In particular, the COVID-19 pandemic has accelerated the use of deep learning methods for radiological screening and genomic sequencing [9, 10].

Although literature on the use of AI in infectious diseases has increased, a comprehensive evaluation of the entire scientific landscape is critically lacking. No prior bibliometric analysis has mapped the full trajectory of AI research in infectious diseases across the decade. This study specifically targets the 2016–2025 timeframe. This period encompasses the maturation of fundamental machine learning algorithms, the impact of the COVID-19 pandemic, and the recent rise of large language models (LLMs) and advanced clinical decision support systems. Understanding the shift in research priorities during this period is crucial for determining the future direction of the field.

Bibliometric analysis is a method that uses mathematical and statistical techniques to analyze published literature, visualize collaboration networks and map out research trends [11]. It enables identification of knowledge gaps and provides evidence-based guidance for future research. Therefore, we conducted a comprehensive bibliometric analysis to map the evolution of the scientific landscape of AI applications in infectious diseases between 2016 and 2025, characterize dominant thematic clusters and international collaboration patterns, and identify emerging research areas to guide future investigations in this field.

2. Methods

2.1. Data source and search strategy

We conducted a comprehensive search for relevant publications related to the application of AI in infectious diseases within the Web of Science Core Collection (WoSCC) database on February 1, 2026. We utilized WoSCC as our research database, as it is widely recognized as the gold standard for bibliometric studies. It provides the structured metadata—such as comprehensive citation indices and institutional affiliations—essential for accurate bibliometric computations. Using a single, standardized source ensures data reliability and prevents duplicate records or incompatible formats. The search strategy used was as follows: TS= ("artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network" OR "supervised learning" OR "unsupervised learning" OR "reinforcement learning" OR "artificial neural network" OR "convolutional neural network" OR "recurrent neural network" OR "large language model"). The search was limited to the Web of Science Category of "Infectious Diseases". This category-based restriction has been implemented to maximize the specificity and reproducibility of the search. The time span was from January 1, 2016, to December 31, 2025, and only original articles written in English were included. Consequently, a total of 1,252 articles were analyzed. The search strategy and publication filtering process are shown in Table 1.

Table 1
The topic search query Web of Science Core Collection, 2016–2025

Set	Results	Search query
#1	1,423,178	TS=("artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network" OR "supervised learning" OR "unsupervised learning" OR "reinforcement learning" OR "artificial neural network" OR "convolutional neural network" OR "recurrent neural network" OR "large language model")
#2	269,828	Infectious Diseases (Web of Science Categories)
#3	1,700	#1 AND #2
#4	1,252	#3 AND Article (Document Types) AND English (Languages)

2.2. Data analysis and visualization

Data downloaded from WoSCC includes information on authors, journals, institutions, countries/regions, citation counts, keywords, and other relevant details essential for bibliometric analysis. The following software tools were employed for bibliometric analysis and visualization of this data: Microsoft Excel 2016, Bibliometrix R package in R-Studio Server 4.2.3 (RStudio, PBC, Boston), VOSviewer 1.6.20 (Leiden University, Netherlands) and Scimago Graphica 1.0.53 (SCImago Research Group S.L.).

Histogram and tables depicting the top 10 most productive countries and journals, as well as the top 10 most-cited papers were generated using Microsoft Excel and R-Studio. Additionally, the Bibliometrix R package was used to analyze trend topics and evaluate the chronological evolution of keywords. The impact factor and category quartiles were sourced from 2024 Journal Citation Reports. VOSviewer, a Java based software specializes in visualization relationships such as co-authorship, co-citation, bibliographic coupling, and keyword co-occurrence. In this study, VOSviewer was used to analyze collaboration networks and visualize among countries, institutions, and co-occurrence of keywords. To

ensure data consistency and eliminate redundancy, a custom thesaurus file was utilized in VOSviewer. This process merged synonymous keywords, spelling variants, and institutional sub-units into standardized terms before conducting the network analysis. Finally, Scimago Graphica software was used to generate geographical visualizations of the network of international cooperation between countries.

2.3. Ethical declaration

Ethics committee approval was waived as this study analyzed previously published, publicly available data and did not involve human participants.

3. Results

3.1. Analysis of annual publications and citations

Between 1 January 2016, and 31 December 2025, a total of 9160 authors, 1252 articles, 125 journals, 2850 institutions, 124 countries /regions contributed to articles related to AI in infectious diseases in the Web of Science.

Prior to 2018, the annual publication rate in this field was minimal. However, following this, the number of papers increased, reaching 128 by 2021. This was over ten times the volume published in 2016. The number of articles published in 2025 peaked at 361, which is more than 30 times the number published in 2016 and more than three times the number published in 2021. The cumulative number of articles published over the previous five years accounted for 84.7% of all papers. The annual number of citations exhibited a comparable increase to the number of publications, exceeding 3000 in 2022. However, a relative decline was observed in the last 3 years (Fig. 1).

3.2. Countries analysis

Table 2 lists the top 10 countries/regions by publication ranked by the number of articles. The USA leads with 409 articles, followed by China, with 325 articles, and the UK, with 130. While the number of citations was highest in the USA (n = 6477), the rate of citations per publication was highest in India (19.1%).

Table 2
The top 10 most productive countries and their citations

Rank	Country	No. of articles	Total citations	Citation per publication	Total link strength*
1	USA	409	6477	15.8	460
2	China	325	2892	8.9	133
3	UK	130	1680	12.9	278
4	Germany	63	645	10.2	155
5	Australia	61	956	15.7	140
6	Switzerland	58	666	11.5	181
7	Brazil	56	415	7.4	108
8	India	54	1030	19.1	71
9	Spain	54	455	8.4	81
10	Canada	49	558	11.4	111
*The total link strength was analyzed from all countries using VOSviewer software.					

A total of 36 countries were included in the analysis, with the publication of more than 10 articles. The co-authorship collaboration network map among countries was constructed by using VOSviewer and Scimago Graphica (Fig. 2). Each node in the map represented a country/region, and the node size represented the number of publications. The lines between nodes represent cooperation between countries. The thickness of lines indicated the strength of cooperation (termed total link strength). The USA had the most publications and the highest total link strength (TLS).

3.3. Institutions analysis

Table 3 lists the top 10 institutions by publication ranked by the number of articles the Harvard University emerged as the leading institution, with the highest number of articles (n = 49), followed by the University of California San Francisco (n = 28), the University of Oxford (n = 24), the Nanjing Medical University (n = 22), the Sun Yat-sen University (n = 22), the Capital Medical University (n = 22) and the Johns Hopkins University (n = 22). Figure 3 shows the visualization map of the institutional collaboration network with the publication of more than 10 articles. (n = 46). While Harvard University had the most publications, most citations and the highest TLS, the rate of citations per publication was highest in Johns Hopkins University (33.3%).

Table 3
The top 10 most productive institutions and their citations

Rank	Institutions	No. of articles	Total citations	Citation per publication	Total link strength*
1	Harvard Univ.	49	1291	26.3	67
2	Univ. of California San Francisco	28	399	14.3	34
3	Univ. of Oxford	24	329	13.7	23
4	Johns Hopkins Univ.	22	733	33.3	26
5	Nanjing Medical Univ.	22	248	11.3	17
6	Sun Yat-sen Univ.	22	190	8.6	8
7	Capital Medical Univ.	22	183	8.3	6
8	Stanford Univ.	21	295	14.0	17
9	Imperial College London	20	344	17.2	12
10	Monash Univ.	19	281	14.8	21
*The total link strength was analyzed from all institutions using VOSviewer software.					

3.4. Journals analysis

The top 10 productive journals published a total of 562 (44.9%) articles. The publication count, citation frequency, impact factor (IF), and Journal Citation Reports (JCR) (2024) category of these journals are summarized in Table 4. BMC Infectious Diseases had the highest number of articles, while Clinical Infectious Diseases had the highest citations and highest IF with 7.3. Five journals ranked Q1 in JCR (2024), four journals ranked Q2, one journal ranked Q3.

Table 4

The top 10 productive journals publishing articles on artificial intelligence in infectious diseases.

Rank	Journals	No. of articles	% of (n/1252)	Total Citations	Citation per publication	IF ^a (2024)	JCR ^b Quartile
1	BMC Infectious Diseases	154	12.3	1172	7.61	3	Q2
2	PLOS Neglected Tropical Diseases	108	8.63	1709	15.8	3.4	Q1
3	Antibiotics-Basel	64	5.11	591	9.23	4.6	Q1
4	Clinical Infectious Diseases	47	3.75	1947	41.4	7.3	Q1
5	Journal Of Infectious Diseases	35	2.80	496	14.2	4.5	Q1
6	Open Forum Infectious Diseases	33	2.64	197	5.97	3.8	Q1
7	Malaria Journal	32	2.56	383	12.0	3	Q2
8	Infection And Drug Resistance	31	2.48	283	9.13	2.9	Q2
9	AIDS	30	2.40	187	6.23	3.1	Q2
10	Journal Of Acquired Immune Deficiency Syndromes	28	2.24	356	12.7	2.2	Q3

^aJournal Impact Factor, ^bJournal Citation Reports (2024)

3.5. Top cited references analysis

The analysis identified the 10 most-cited articles on artificial intelligence in infectious diseases (Table 5). The most-cited paper was *Classification of COVID-19 patients from chest CT images using multi-objective differential evolution-based convolutional neural networks*, published in 2020 in European Journal of Clinical Microbiology & Infectious Diseases, with 383 citations. The average number of citations for the top five articles was over 200. The 10 most-cited articles primarily focused on artificial intelligence models related to diagnosis and clinical decision-making in patients with COVID-19, tuberculosis, HIV, and dengue fever.

Table 5
The 10 most-cited articles

Authors	Article Title	Times cited, WoS Core	Citations average per year	Journal	Publication Year
Singh et all.	Classification of COVID-19 patients from chest CT images using multi-objective differential evolution-based convolutional neural networks	383	63.8	European Journal of Clinical Microbiology & Infectious Diseases	2020
Wiens et all.	Machine Learning for Healthcare: On the Verge of a Major Shift in Healthcare Epidemiology	337	42.1	Clinical Infectious Diseases	2018
Chia et all.	Dynamics of SARS-CoV-2 neutralising antibody responses and duration of immunity: a longitudinal study	215	43	Lancet Microbe	2021
Alzahrani et all.	Forecasting the spread of the COVID-19 pandemic in Saudi Arabia using ARIMA prediction model under current public health interventions	162	27	Journal Of Infection and Public Health	2020
Hwang et all.	Development and Validation of a Deep Learning-based Automatic Detection Algorithm for Active Pulmonary Tuberculosis on Chest Radiographs	161	23	Clinical Infectious Diseases	2019
Goodman et all.	A Clinical Decision Tree to Predict Whether a Bacteremic Patient Is Infected With an Extended-Spectrum β -Lactamase-Producing Organism	155	15.5	Clinical Infectious Diseases	2016
Guo et all.	Developing a dengue forecast model using machine learning: A case study in China	155	17.2	PLOS Neglected Tropical Diseases	2017

Authors	Article Title	Times cited, WoS Core	Citations average per year	Journal	Publication Year
Marcus et all.	Use of electronic health record data and machine learning to identify candidates for HIV pre-exposure prophylaxis: a modelling study	110	15.7	Lancet HIV	2019
Benedum et all.	Statistical modeling of the effect of rainfall flushing on dengue transmission in Singapore	106	13.3	PLOS Neglected Tropical Diseases	2018
Chintalapudi et all.	Sentimental Analysis of COVID-19 Tweets Using Deep Learning Models	102	20.4	Infectious Disease Reports	2021

3.6 Keyword co-occurrence and trend analysis

Keyword co-occurrence network analysis identified 95 keywords that met the minimum co-occurrence threshold of five or more. These keywords were organized into nine thematic clusters, with a total of 2,146 occurrences and a cumulative TLS of 3,792. Threshold value was selected to filter out isolated terms and focus on core research themes. Keywords and clusters resulting from the analysis were generated automatically by VOSviewer based on the strength of the associations between the nodes, with no manual intervention. Through this keyword analysis, the relationships between keywords and their clustering were identified. Each node in the network represents different keywords, with their size corresponding to the frequency of occurrence. The connecting lines between nodes represent co-occurrence relationships, while line thickness reflects the strength of these relationships. Furthermore, the node colors represent different clusters, each corresponding to distinct topics or subfields within the study.

The terms “artificial intelligence” and “machine learning” formed the largest and most central nodes in the network; they were found to have strong connections with the terms “covid-19”, “hiv”, “prediction”, “tuberculosis”, “antimicrobial resistance”, “malaria”, “deep learning” and “sepsis”. Through cluster analysis, we identified nine core clusters (Fig. 4).

Cluster #1 (n = 16; occurrence = 640; TLS = 1,022), anchored by “machine learning” (occurrence = 506, TLS = 743), encompassed terms related to epidemiological surveillance and outbreak monitoring, including “epidemiology”, “big data”, “disease surveillance”, “outbreak detection” and “social media”.

Cluster #2 (n = 14; occurrence = 194; TLS = 398), anchored by “antimicrobial resistance” (occurrence = 69; TLS = 131), encompassed keywords related to “antimicrobial stewardship”, “MALDI-TOF” based

diagnostics, and “clinical decision support systems”, alongside specific multidrug-resistant pathogens including “*Klebsiella pneumoniae*”, “ESBL” and “carbapenem” resistant bacteria, as well as infection types such as “bacteremia”, “pneumonia”, and “urinary tract infection”.

Cluster #3 (n = 13; occurrence = 325; TLS = 631), anchored by “artificial intelligence” (occurrence = 159; TLS = 293), encompassed keywords spanning “large language model” in “public health” and “digital health” “surveillance”.

Cluster #4 (n = 11; occurrence = 277; TLS = 437) anchored by “covid-19” (occurrence = 141; TLS = 200), encompassed keywords related to pandemic “forecasting” and “time series analysis” of viral outbreaks, including “artificial neural network”, “influenza”, “dengue”.

Cluster #5 (n = 11; occurrence = 272; TLS = 484) co-anchored by “prediction” (occurrence = 103; TLS = 203) and “hiv” (occurrence = 100; TLS = 165), encompassed keywords related to predictive modeling and HIV/AIDS management, including “supervised learning”, “antiretroviral therapy” and “viral load”.

Cluster #6 (n = 9; occurrence = 140; TLS = 256) anchored by “malaria” (occurrence = 39; TLS = 57) and “sepsis” (occurrence = 29; TLS = 50), encompassed keywords related clinical severity scoring and mortality prediction, including “explainable artificial intelligence”, “SHAP”, “mortality”, “diagnosis”, and “nomogram”.

Cluster #7 (n = 8; occurrence = 87; TLS = 158), anchored by “random forest” (occurrence = 26; TLS = 51), encompassed keywords related to biomarker discovery in resource-limited settings, including “biomarker”, “schistosomiasis”, and “Africa”.

Cluster #8 (n = 8; occurrence = 65; TLS = 160), anchored by support vector machine (occurrence = 12; TLS = 32), encompassed keywords representing risk stratification and diagnostic “classification”, including “logistic regression”, “decision tree”, “feature selection” and “risk factor”.

Cluster #9 (n = 5; occurrence = 146; TLS = 246), co-anchored by tuberculosis (occurrence = 65; TLS = 103) and deep learning (occurrence = 58; TLS = 102), encompassed keywords related to image-based diagnostic modeling, including “chest radiography”, “radiomics”, and “screening”.

Trend topics analysis was used to visualize the temporal evolution of keywords over the study period, shown in Fig. 5. The terms “machine learning” and “artificial intelligence” have grown rapidly since 2020, reaching their peak frequency in 2025. “covid-19” suddenly entered the scientific literature in 2020 and have remained prevalent ever since. The presence of “hiv”, “tuberculosis” and “malaria” has remained steady over the years. Additionally, the terms ‘antimicrobial resistance’ and ‘sepsis’ have become more frequent in recent years.

4. Discussion

This study demonstrates that the use of artificial intelligence in the field of infectious diseases has grown exponentially over the past decade. Catalyzed significantly by the COVID-19 pandemic, the field has evolved from foundational predictive modelling to integrated clinical applications, with the USA and China leading the way. Current research hotspots are focusing increasingly on sepsis management and antimicrobial resistance surveillance.

4.1. General information

This bibliometric analysis systematically documents publication trends, international collaboration networks and areas of research focus in the field of artificial intelligence applications in infectious diseases between 2016 and 2025. Over this decade, 1252 original articles were published by 9160 researchers from 124 countries. The findings suggest that research in this field has advanced in parallel with the wider adoption of AI in medicine [12].

Scientific publications on infectious diseases and artificial intelligence have increased significantly since 2018. The number of annual publications peaked in 2025 at over 30 times the number in 2016, and the cumulative output of the last five years accounted for 84.7% of all articles. It appears that the surge was triggered by the COVID-19 pandemic [13, 14]. This aligns particularly with the development of radiological imaging [15], genomic sequencing [10] and predictive models [16] during the pandemic. Conversely, although the number of citations peaked in 2022, the relative attenuation observed over the following three years can be attributed to citation lag. Recently published articles have not yet accrued citations commensurate with their scientific impact, a limitation inherent to bibliometric analyses.

The geographical and institutional distribution of scientific output highlights the leadership of the USA and China in this field. This pattern is consistent with bibliometric analyses of AI applications in emergency medicine [17], critical care [18], and dermatology [19]. This dominance in high-income countries may be due to access to extensive digital health datasets and investment in technological infrastructure [20]. The USA is the leading player in terms of both productivity and impact, with the highest number of publications and citations. The fact that the USA ranks first in terms of total link strength highlights its central role in global cooperation networks. Notably, India has the highest citation rate per publication among all top ten countries. This may be due to a single highly cited article published by Singh et al. from India [21].

The top ten journals collectively accounted for 44.9% of publications. BMC Infectious Diseases published the most articles, while Clinical Infectious Diseases had the highest number of citations per publication and the highest impact factor (IF = 7.3). Clinical Infectious Diseases is also among the most cited journals in other bibliometric analyses investigating artificial intelligence applications in sepsis and tuberculosis [22, 23]. The presence of five Q1 journals among the top ten indicates that AI research in infectious diseases is being published in high-quality scientific journals. The inclusion of journals such as 'PLOS Neglected Tropical Diseases' and 'The Malaria Journal' on the list highlights the significant impact of AI research on diseases that affect low- and middle-income countries.

An analysis of the top 10 most-cited articles reveals that, prior to 2020, highly cited studies focused on specific issues such as image processing in tuberculosis [24], dengue fever surveillance [25, 26] and HIV prediction models [27]. The onset of the pandemic, the focus has suddenly shifted toward global respiratory infections. Singh et al.'s convolutional neural network–based COVID-19 classification study has become the most-cited paper [21]. The speed with which publications from the pandemic era reached record citation levels likely reflects the significant increase in knowledge dissemination during severe global health crises.

4.2. Research hotspots and future directions

The largest cluster in the network was anchored by machine learning and encompassed terms including epidemiology, big data, disease surveillance, and outbreak detection. Artificial intelligence (AI) is emerging as a powerful tool in public health, offering new opportunities to enhance infectious disease surveillance and early warning systems [28, 29]. The potential of AI to transform early outbreak detection and improve epidemiological modeling is attracting increasing research attention [30]. Machine learning-based models can process and analyze large datasets derived from big data repositories and social media platforms, enabling the identification of patterns and anomalies indicative of emerging public health threats [31, 32]. In a notable example, Goncalves et al. integrated electronic health records with machine learning algorithms across a dataset of 4.5 million patients and demonstrated that 33% of 123 outbreaks were detected one to 24 days earlier than conventional surveillance methods [33]. Realizing this potential will require coordinated efforts to foster collaboration between AI developers and public health experts.

Antimicrobial resistance formed one of the most clinically significant clusters in the keyword co-occurrence network, a finding corroborated by the trend topics analysis demonstrating an increase in publication frequency for this term in recent years. The strong potential of deep learning models in antimicrobial stewardship processes positions this domain as a priority research agenda for infectious disease practice [34]. Howard et al. published a roadmap for the application of artificial intelligence models in combating antimicrobial resistance [35], while Yang et al. developed clinical decision support systems to predict antibiotic resistance in complicated urinary tract infections, thereby guiding antibiotic selection [36]. Despite these advances, randomized prospective validation studies are needed before AI-assisted applications can be incorporated into international guidelines such as those of IDSA/SHEA and ESCMID. The integration of MALDI-TOF mass spectrometry with machine learning holds considerable promise for accelerating species identification and resistance profile prediction, ultimately improving empirical antibiotic selection. In this regard, Weis et al. demonstrated that a machine learning model can predict antimicrobial resistance directly from MALDI-TOF mass spectrometry profiles, representing a significant contribution to the field [37]. The application of deep learning models to novel antibiotic discovery represents a further promising direction, with landmark studies highlighting the potential of this approach [38, 39].

Early diagnosis and mortality prediction in acute infectious conditions such as sepsis and malaria constitute another clinically critical cluster in the network. Machine learning models for the early

detection of sepsis and mortality prediction in intensive care units have been shown to outperform traditional scoring systems such as SOFA and APACHE-II [40, 41]. In a prospective study, patients who received a Targeted Real-time Early Warning System (TREWS) alert within three hours had significantly reduced mortality, organ failure, and length of hospital stay compared with those who did not [42]. Nevertheless, challenges persist, including small sample sizes, high false-positive rates, insufficient external model validation, and limited transparency and interpretability. In this context, the co-occurrence of SHAP, nomogram, and explainable artificial intelligence within this cluster is particularly meaningful, reflecting a shift away from opaque black-box architectures toward models that clinicians can interrogate and trust. Gaouar et al. demonstrated that integrating deep learning models with explainable AI can support the early and interpretable detection of malaria in clinical settings [43]. Explainable AI-based research is therefore expected to gain considerable momentum in the coming years.

COVID-19 represents the central component of the cluster encompassing pandemic forecasting and time-series analysis of viral outbreaks. The pandemic served as a powerful catalyst for AI research in infectious diseases; Vaishya et al. demonstrated that AI applications could be used for cluster detection, outbreak prediction, and epidemic monitoring [44]. Subsequent studies applied time-series modeling to forecast COVID-19 cases with considerable accuracy [45, 46]. The methodological advances developed for COVID-19 have since been transferred to other viral diseases, including influenza and dengue fever, representing a tangible legacy of pandemic-era research. Artificial neural networks have emerged as particularly valuable tools for predicting the peak of seasonal and epidemic outbreaks [47, 48]. The ability to model transmission trajectories under acute crisis conditions is indispensable for healthcare capacity planning and resource allocation. Whether these methodologies can be systematically transferred to future novel outbreaks remains to be prospectively evaluated.

AI applications targeting pathogens with the highest global disease burden constitute the most established and mature clusters in the network. The sustained keyword frequency associated with HIV, tuberculosis, and malaria throughout the study period confirms that research activity in these domains predates the pandemic era and has continued uninterrupted. Within the HIV-related cluster, the co-occurrence of prediction, antiretroviral therapy, and viral load terms are particularly notable. Artificial intelligence has demonstrated considerable potential in addressing gaps in HIV care [49], improving diagnostic accuracy [50], monitoring clinical outcomes [51], supporting antiretroviral treatment decision-making [52], and enhancing treatment adherence strategies and therapeutic follow-up [53]. Nonetheless, challenges related to ethical considerations, data quality, and integration into existing health systems remain to be addressed.

The cluster encompassing image-based deep learning and radiomics analysis has consolidated around tuberculosis as its primary focus. Automated screening of chest radiographs using convolutional neural networks remains the most extensively validated application of computer vision in clinical infectious disease practice. Hwang et al., in one of the most cited articles identified in this analysis, demonstrated that deep learning-based pulmonary tuberculosis detection systems can achieve diagnostic performance comparable to that of expert radiologists [24]. In resource-limited settings where

radiologist availability is scarce and tuberculosis incidence remains high, these models offer a cost-effective and rapid solution for mass screening, thereby directly contributing to global tuberculosis elimination targets [54].

4.3. Limitations

Several methodological limitations should be considered when interpreting these findings. Firstly, to standardize the quality of the data analysis, it was limited to the WoSCC database. Other major databases such as PubMed, Scopus, and Embase were excluded. Second, articles published in languages other than English were excluded. Thus, we might have overlooked relevant studies published in other databases or in different languages. Finally, bibliometric and visualization tools such as VOSviewer and Scimago Graphica were used to quantify and visualize trends in the literature. While these tools provide valuable insights into publication distribution, authorship and citation impact, they may not accurately reflect the quality or clinical significance of studies. Relative to more established work, recent publications with high potential may be at a disadvantage in citation-based rankings.

This bibliometric analysis provides a comprehensive characterization of global research on artificial intelligence applications in infectious diseases between 2016 and 2025. Publication output has grown substantially over the past five years, and findings suggest this trend will continue. The USA and China collectively account for the majority of global scientific production in this field. Keyword co-occurrence analysis identified nine thematically distinct clusters, spanning clinical risk prediction, epidemiological surveillance, tuberculosis diagnostics, and antimicrobial resistance management. Analysis of trends indicates that sepsis and antimicrobial resistance are the most prominent issues for future investigation. Taken together, these findings suggest that artificial intelligence holds considerable potential to move beyond a research focus and become an integral part of clinical decision-making in infectious diseases.

Declarations

Acknowledgment

The authors thank all the authors of the original studies included in this analysis.

Funding

No funding for this research project.

Author information

Authors and Affiliations

Department of Infectious Diseases and Clinical Microbiology, University of Health Sciences, İzmir Tepecik Training and Research Hospital, Türkiye

Çağlar Irmak& İlkay Akbulut& Sabri Atalay

Ahmet Furkan Süner

Author Contributions

All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by ÇI and AFS. The first draft of the manuscript was written by ÇI. İA and SA took consultant roles, guided the design, and oversaw the drafting and revision of the manuscript. All authors approved and agreed on the final manuscript as submitted.

Ethics declarations

Ethics approval statement

Ethics committee approval was waived as this study analyzed previously published, publicly available data and did not involve human participants.

Consent to participate

Not applicable.

Conflict of interest

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

References

1. GBD 2023 Lower Respiratory Infections and Antimicrobial Resistance Collaborators. Global burden of lower respiratory infections and aetiologies, 1990–2023: a systematic analysis for the Global Burden of Disease Study 2023. *Lancet Infect Dis.* 2026;26:343–61. [https://doi.org/10.1016/S1473-3099\(25\)00689-9](https://doi.org/10.1016/S1473-3099(25)00689-9).
2. GBD 2021 Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance 1990–2021: a systematic analysis with forecasts to 2050. *Lancet.* 2024;404:1199–226. [https://doi.org/10.1016/S0140-6736\(24\)01867-1](https://doi.org/10.1016/S0140-6736(24)01867-1).
3. Rabaan AA, Bakhrebah MA, Alotaibi J, et al. Unleashing the power of artificial intelligence for diagnosing and treating infectious diseases: A comprehensive review. *J Infect Public Health.* 2023;16:1837–47.
4. Pugliese R, Regondi S, Marini R. Machine learning-based approach: global trends, research directions, and regulatory standpoints. *Data Sci Manage.* 2021;4:19–29.
5. Alowais SA, Alghamdi SS, Alsuhebany N, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC Med Educ.* 2023;23:689.

6. Atalay S, Sönmez U. Digital twin in health care. In: Karaarslan E, Aydin Ö, Cali Ü, et al. editors. Digital twin driven intelligent systems and emerging metaverse. Singapore: Springer; 2023. pp. 209–31.
7. Al Kuwaiti A, Nazer K, Al-Reedy A, et al. A Review of the Role of Artificial Intelligence in Healthcare. *J Pers Med*. 2023;13:951.
8. Assudani PJ, Bhurgy AS, Kollem S, et al. Artificial intelligence and machine learning in infectious disease diagnostics: A comprehensive review of applications, challenges, and future directions. *Microchem J*. 2025;218:115802.
9. Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, et al. Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: Evaluation of the diagnostic accuracy. *Radiology*. 2020;296:E65–71.
10. Abubaker Bagabir S, Ibrahim NK, Abubaker Bagabir H, Hashem Ateeq R. Covid-19 and Artificial Intelligence: Genome sequencing, drug development and vaccine discovery. *J Infect Public Health*. 2022;15:289–96.
11. Aria M, Cuccurullo C. bibliometrix: An R-tool for comprehensive science mapping analysis. *J Informetrics*. 2017;11:959–75.
12. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019;25:44–56. <https://doi.org/10.1038/s41591-018-0300-7>.
13. Chang Z, Zhan Z, Zhao Z, et al. Application of artificial intelligence in COVID-19 medical area: a systematic review. *J Thorac Dis*. 2021;13:7034–53. <https://doi.org/10.21037/jtd-21-747>.
14. Abdullah R, Fakieh B. Health Care Employees' Perceptions of the Use of Artificial Intelligence Applications: Survey Study. *J Med Internet Res*. 2020;22:e17620. <https://doi.org/10.2196/17620>.
15. Zhang K, Liu X, Shen J, et al. Clinically Applicable AI System for Accurate Diagnosis, Quantitative Measurements, and Prognosis of COVID-19 Pneumonia Using Computed Tomography. *Cell*. 2020;181:1423–e143311. <https://doi.org/10.1016/j.cell.2020.04.045>.
16. Zheng N, Du S, Wang J, et al. Predicting COVID-19 in China Using Hybrid AI Model. *IEEE Trans Cybern*. 2020;50:2891–904. <https://doi.org/10.1109/TCYB.2020.2990162>.
17. Limon Ö, Bayram B, Çetin M, Limon G, Dirican N. A bibliometric analysis of clinical studies on artificial intelligence in emergency medicine. *Med (Baltim)*. 2025;104:e43282. <https://doi.org/10.1097/MD.00000000000043282>.
18. Luo Z, Lv J, Zou K. A bibliometric analysis of artificial intelligence research in critical illness: a quantitative approach and visualization study. *Front Med (Lausanne)*. 2025;12:1553970. <https://doi.org/10.3389/fmed.2025.1553970>.
19. Wang G, Meng X, Zhang F. Past, present, and future of global research on artificial intelligence applications in dermatology: A bibliometric analysis. *Med (Baltim)*. 2023;102:e35993. <https://doi.org/10.1097/MD.00000000000035993>.
20. Celi LA, Cellini J, Charpignon ML, et al. Sources of bias in artificial intelligence that perpetuate healthcare disparities-A global review. *PLOS Digit Health*. 2022;1:e0000022. <https://doi.org/10.1371/journal.pdig.0000022>.

21. Singh D, Kumar V, Vaishali, Kaur M. Classification of COVID-19 patients from chest CT images using multi-objective differential evolution-based convolutional neural networks. *Eur J Clin Microbiol Infect Dis*. 2020;39:1379–89. <https://doi.org/10.1007/s10096-020-03901-z>.
22. Tang M, Mu F, Cui C, et al. Research frontiers and trends in the application of artificial intelligence to sepsis: A bibliometric analysis. *Front Med (Lausanne)*. 2023;9:1043589. <https://doi.org/10.3389/fmed.2022.1043589>.
23. Cabanillas-Lazo M, Quispe-Vicuña C, Pascual-Guevara M, et al. Bibliometric analyses of applications of artificial intelligence on tuberculosis. *Int J Mycobacteriol*. 2022;11:389–93. https://doi.org/10.4103/ijmy.ijmy_134_22.
24. Hwang EJ, Park S, Jin KN, et al. Development and Validation of a Deep Learning-based Automatic Detection Algorithm for Active Pulmonary Tuberculosis on Chest Radiographs. *Clin Infect Dis*. 2019;69:739–47. <https://doi.org/10.1093/cid/ciy967>.
25. Guo P, Liu T, Zhang Q, et al. Developing a dengue forecast model using machine learning: A case study in China. *PLoS Negl Trop Dis*. 2017;11:e0005973. <https://doi.org/10.1371/journal.pntd.0005973>.
26. Benedum CM, Seidahmed OME, Eltahir EAB, Markuzon N. Statistical modeling of the effect of rainfall flushing on dengue transmission in Singapore. *PLoS Negl Trop Dis*. 2018;12:e0006935. <https://doi.org/10.1371/journal.pntd.0006935>.
27. Marcus JL, Hurley LB, Krakower DS, Alexeeff S, Silverberg MJ, Volk JE. Use of electronic health record data and machine learning to identify candidates for HIV pre-exposure prophylaxis: a modelling study. *Lancet HIV*. 2019;6:e688–95. [https://doi.org/10.1016/S2352-3018\(19\)30137-7](https://doi.org/10.1016/S2352-3018(19)30137-7).
28. Li C, Ye G, Jiang Y, Wang Z, Yu H, Yang M. Artificial Intelligence in battling infectious diseases: A transformative role. *J Med Virol*. 2024;96:e29355. <https://doi.org/10.1002/jmv.29355>.
29. MacIntyre CR, Chen X, Kunasekaran M, et al. Artificial intelligence in public health: the potential of epidemic early warning systems. *J Int Med Res*. 2023;51:3000605231159335. <https://doi.org/10.1177/03000605231159335>.
30. Ezech CJ, Anioke SC, Oyewole S, David MG. The role of predictive analytics in enhancing public health surveillance: Proactive and data-driven interventions. *World J Adv Res Reviews*. 2024;24:3059–77. <https://doi.org/10.30574/wjarr.2024.24.3.3909>.
31. Liscano Y, Anillo Arrieta LA, Montenegro JF, Prieto-Alvarado D, Ordoñez J. Early Warning of Infectious Disease Outbreaks Using Social Media and Digital Data: A Scoping Review. *Int J Environ Res Public Health*. 2025;22:1104. <https://doi.org/10.3390/ijerph22071104>.
32. Villanueva-Miranda I, Xiao G, Xie Y. Artificial intelligence in early warning systems for infectious disease surveillance: a systematic review. *Front Public Health*. 2025;13:1609615. <https://doi.org/10.3389/fpubh.2025.1609615>.
33. Goncalves A, Cadena J, Hu Y et al. (2025) AI-Enabled Diagnostic Prediction within Electronic Health Records to Enhance Biosurveillance and Early Outbreak Detection. *Res Sq rs.3.rs-6606632*. <https://doi.org/10.21203/rs.3.rs-6606632/v1>

34. Pinto-de-Sá R, Sousa-Pinto B, Costa-de-Oliveira S. Brave New World of Artificial Intelligence: Its Use in Antimicrobial Stewardship-A Systematic Review. *Antibiot (Basel)*. 2024;13:307. <https://doi.org/10.3390/antibiotics13040307>.
35. Howard A, Aston S, Gerada A, et al. Antimicrobial learning systems: an implementation blueprint for artificial intelligence to tackle antimicrobial resistance. *Lancet Digit Health*. 2024;6:e79–86. [https://doi.org/10.1016/S2589-7500\(23\)00221-2](https://doi.org/10.1016/S2589-7500(23)00221-2).
36. Yang J, Eyre DW, Lu L, Clifton DA. Interpretable machine learning-based decision support for prediction of antibiotic resistance for complicated urinary tract infections. *NPJ Antimicrob Resist*. 2023;1:14. <https://doi.org/10.1038/s44259-023-00015-2>.
37. Weis C, Cuénod A, Rieck B, et al. Direct antimicrobial resistance prediction from clinical MALDI-TOF mass spectra using machine learning. *Nat Med*. 2022;28:164–74. <https://doi.org/10.1038/s41591-021-01619-9>.
38. Stokes JM, Yang K, Swanson K, et al. A Deep Learning Approach to Antibiotic Discovery. *Cell*. 2020;181:475–83. <https://doi.org/10.1016/j.cell.2020.04.001>.
39. Wong F, Zheng EJ, Valeri JA, et al. Discovery of a structural class of antibiotics with explainable deep learning. *Nature*. 2024;626:177–85. <https://doi.org/10.1038/s41586-023-06887-8>.
40. Fleuren LM, Klausch TLT, Zwager CL, et al. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. *Intensive Care Med*. 2020;46:383–400. <https://doi.org/10.1007/s00134-019-05872-y>.
41. Goh KH, Wang L, Yeow AYK, et al. Artificial intelligence in sepsis early prediction and diagnosis using unstructured data in healthcare. *Nat Commun*. 2021;12:711. <https://doi.org/10.1038/s41467-021-20910-4>.
42. Adams R, Henry KE, Sridharan A, et al. Prospective, multi-site study of patient outcomes after implementation of the TREWS machine learning-based early warning system for sepsis. *Nat Med*. 2022;28:1455–60. <https://doi.org/10.1038/s41591-022-01894-0>.
43. Gaouar A, Cherif SH, Rahmoun A, Daho MEH. Explainable AI for early malaria detection using stacked-LSTM and attention mechanisms. *Inf Med Unlocked*. 2025;57:101667. <https://doi.org/10.1016/j.imu.2025.101667>.
44. Vaishya R, Javaid M, Khan IH, Haleem A. Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes Metab Syndr*. 2020;14:337–9. <https://doi.org/10.1016/j.dsx.2020.04.012>.
45. Somyanonthanakul R, Warin K, Amasiri W, et al. Forecasting COVID-19 cases using time series modeling and association rule mining. *BMC Med Res Methodol*. 2022;22:281. <https://doi.org/10.1186/s12874-022-01755-x>.
46. Chimmula VKR, Zhang L. Time series forecasting of COVID-19 transmission in Canada using LSTM networks. *Chaos Solitons Fractals*. 2020;135:109864. <https://doi.org/10.1016/j.chaos.2020.109864>.
47. Orang A, Berke O, Poljak Z, Greer AL, Rees EE, Ng V. Forecasting seasonal influenza activity in Canada—Comparing seasonal Auto-Regressive integrated moving average and artificial neural

- network approaches for public health preparedness. *Zoonoses Public Health*. 2024;71:304–13. <https://doi.org/10.1111/zph.13114>.
48. Akter T, Islam MT, Hossain MF, Ullah MS. A Comparative Study between Time Series and Machine Learning Technique to Predict Dengue Fever in Dhaka City. *Discrete Dynamics Nat Soc*. 2024;2024:2757381. <https://doi.org/10.1155/2024/2757381>.
49. Esber AL, Dear NF, King D, et al. Achieving the third 95 in sub-Saharan Africa: application of machine learning approaches to predict viral failure. *AIDS*. 2023;37:1861–70. <https://doi.org/10.1097/QAD.0000000000003646>.
50. Roche SD, Ekwunife OI, Mendonca R, et al. Measuring the performance of computer vision artificial intelligence to interpret images of HIV self-testing results. *Front Public Health*. 2024;12:1334881. <https://doi.org/10.3389/fpubh.2024.1334881>.
51. Domínguez-Rodríguez S, Serna-Pascual M, Oletto A, et al. Machine learning outperformed logistic regression classification even with limit sample size: A model to predict pediatric HIV mortality and clinical progression to AIDS. *PLoS ONE*. 2022;17:e0276116. <https://doi.org/10.1371/journal.pone.0276116>.
52. Akbulut I, Odemis I, Atalay S, Inkaya AC. Comparison of clinician preference and HIV-ASSIST recommendations in antiretroviral therapy decision-making: A single center experience from Türkiye. *Int J STD AIDS*. 2024;35:430–7. <https://doi.org/10.1177/09564624241229464>.
53. Ngcobo S, Mntla EM, Shock J, et al. Artificial intelligence for HIV care: a global systematic review of current studies and emerging trends. *J Int AIDS Soc*. 2025;28:e70045. <https://doi.org/10.1002/jia2.70045>.
54. Matteelli A, Lovatti S, Rossi B, Rossi L. Update on multidrug-resistant tuberculosis preventive therapy toward the global tuberculosis elimination. *Int J Infect Dis*. 2025;155:107849. <https://doi.org/10.1016/j.ijid.2025.107849>.

Figures

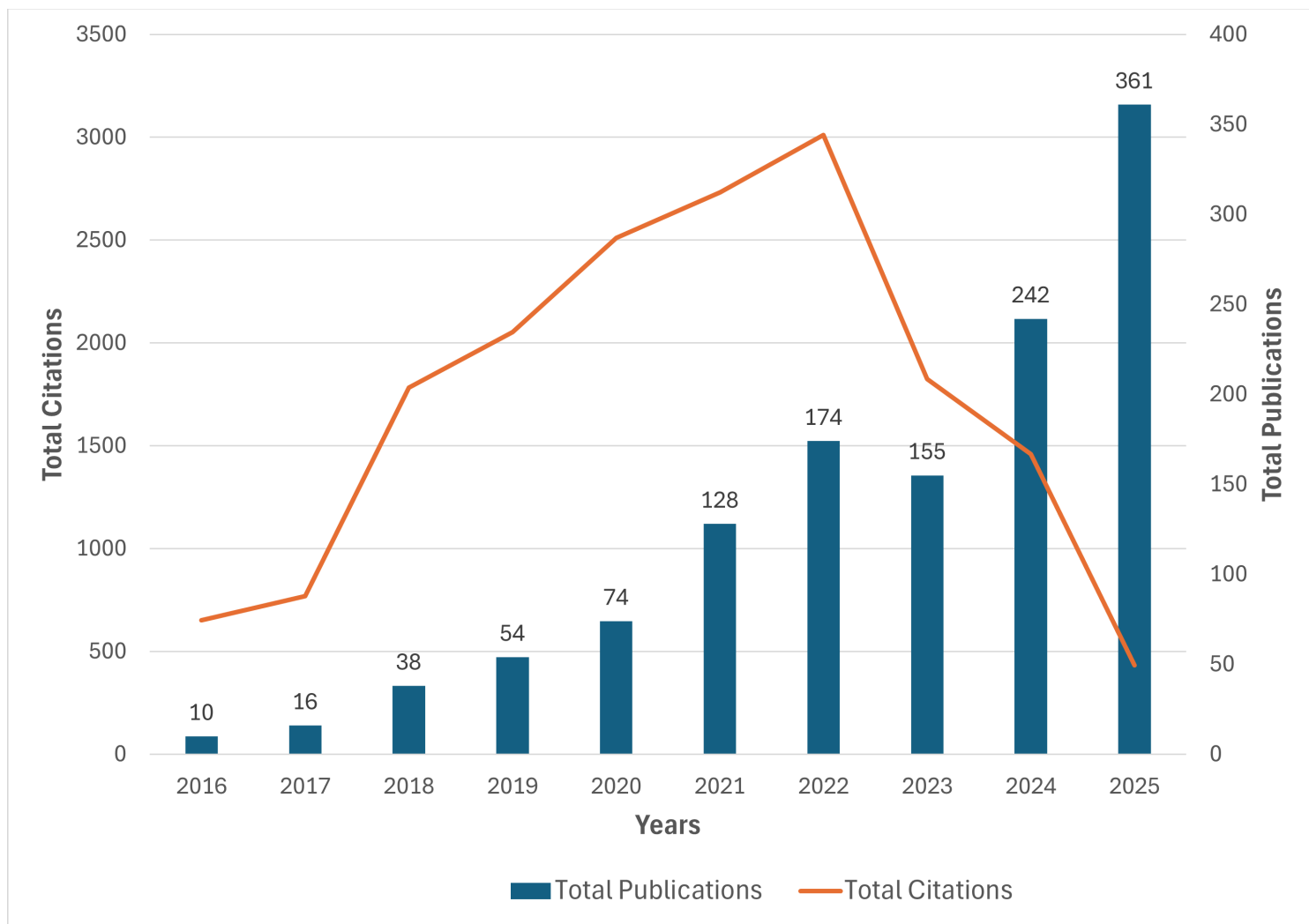


Figure 1

The annual number of publications and citations, 2016–2025

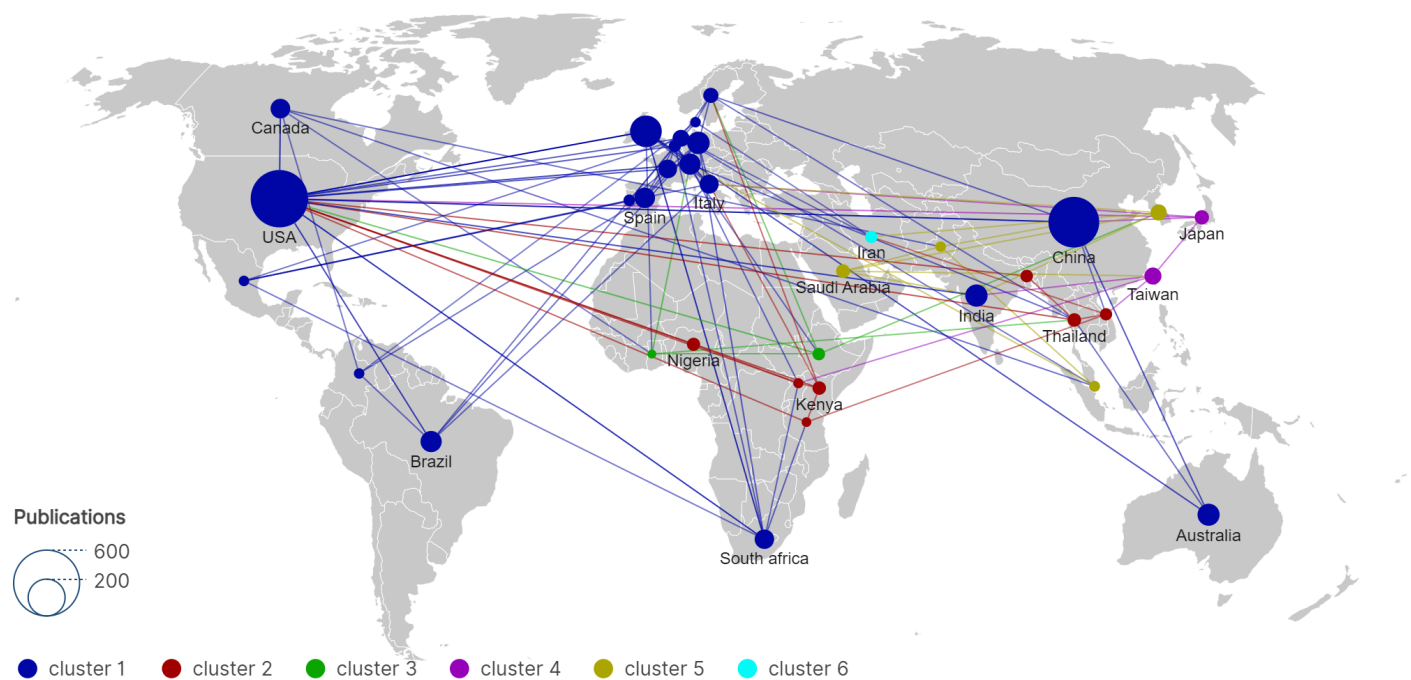


Figure 2

Geographical visualization of research networks of countries/regions.

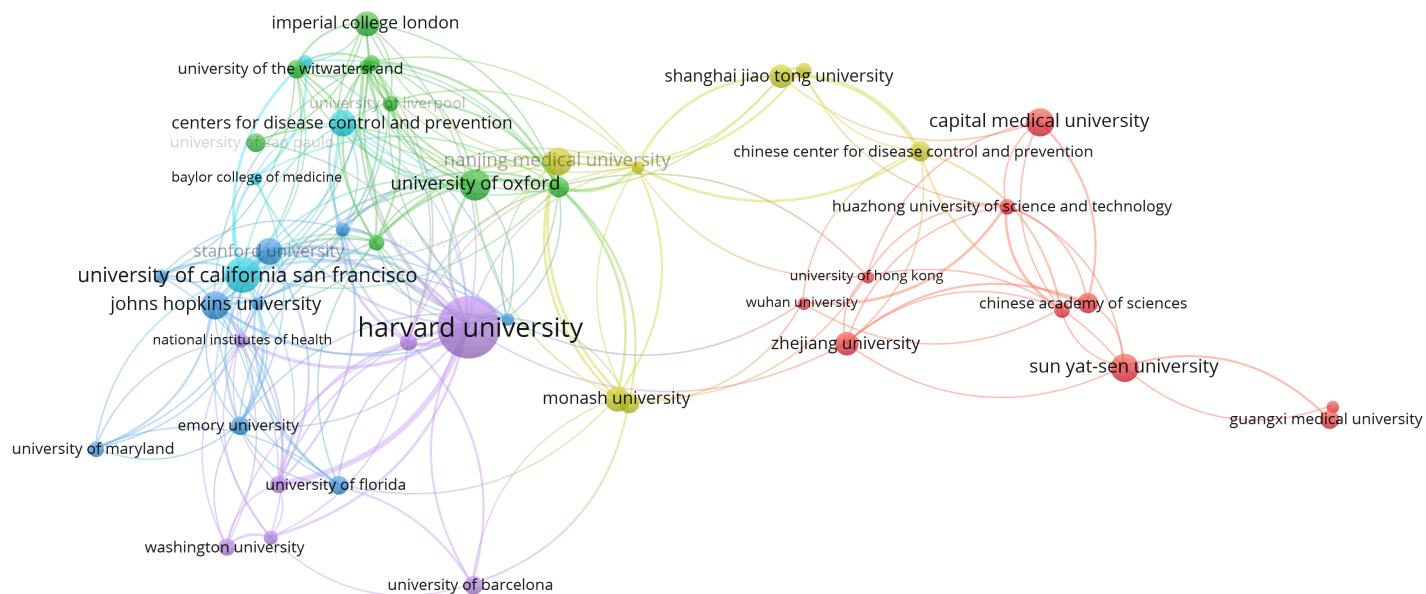
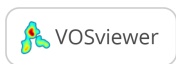
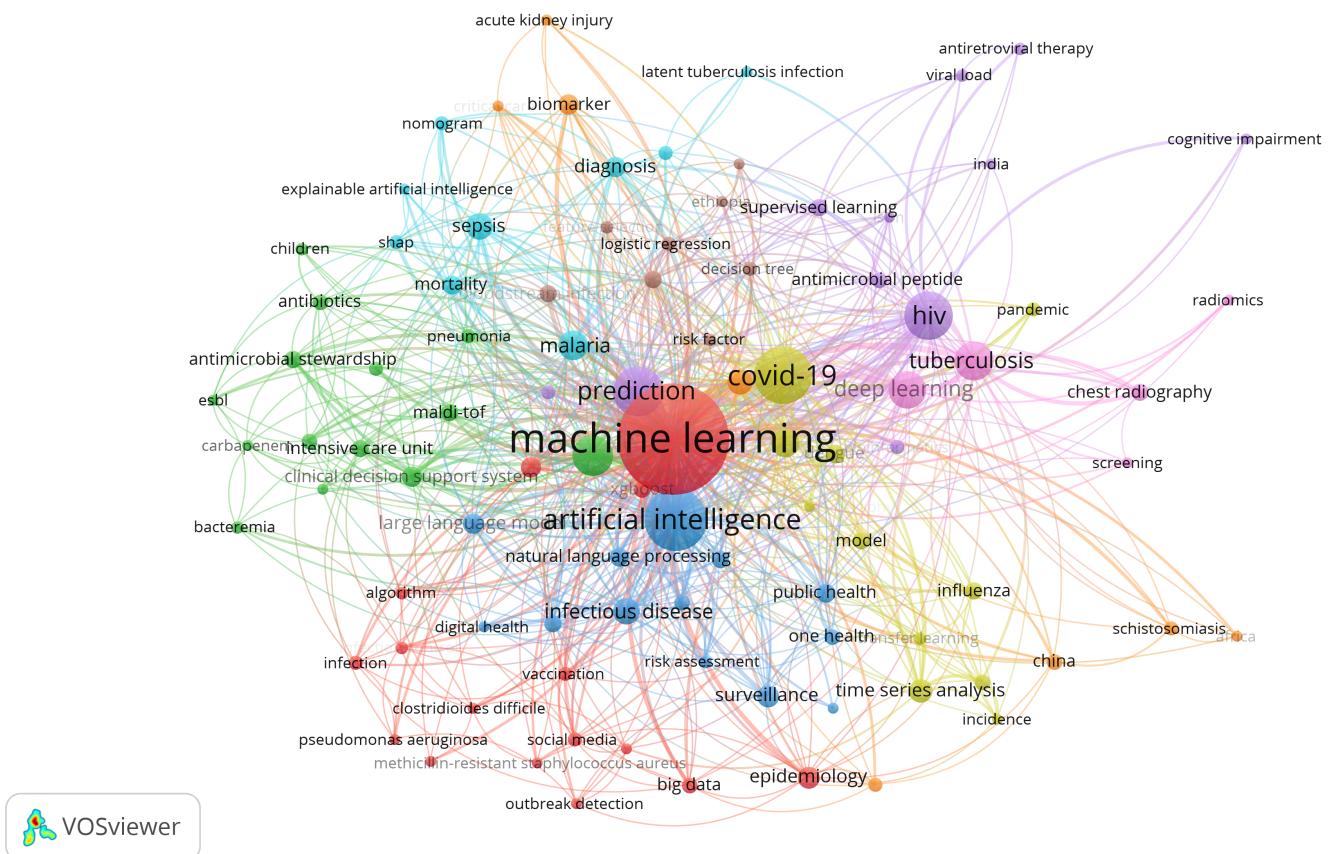


Figure 3

Visualization of research networks of institutions.



Keyword co-occurrence network.

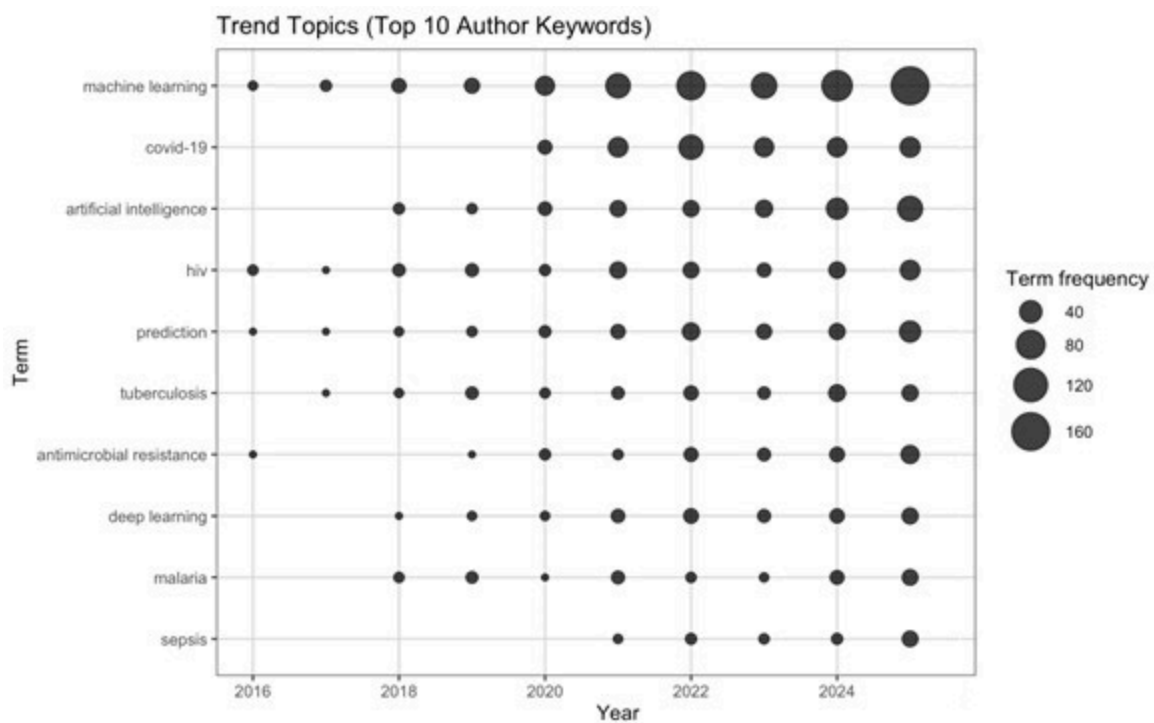


Figure 5

Trend topics of the top 10 author keywords, 2016–2025

