Revista Prevenção de Infecção e Saúde
The Official Journal of the Human Exposome and Infectious Diseases Network

ORIGINAL ARTICLE

DOI: 10.26694/repis.v11i1.7237

# Bibliometric Analysis on Artificial Intelligence applied to Infection Control

## Análise Bibliométrica sobre a Inteligência Artificial aplicada no Controle de Infecções

## Análisis bibliométrico sobre Inteligencia Artificial aplicada al Control de Infecciones

**Ellayne Beatriz Nunes de Sousa [1], Lorrana Maria Araújo de Almeida Nobre[1], Elisa Linhares Lima[2], Brenda Kelly da Silva Monte[1], Jonathan Wedson da Silva[1], Ana Raquel Batista de Carvalho[1]**

**How to cite this article:**
Sousa EBN, Nobre LMAA, Lima EL, Monte BKS, Silva JW, Carvalho ARB. Bibliometric Analysis on Artificial Intelligence applied to Infection Control. Rev Pre Infec e Saúde [Internet]. 2025; 11: 01. Disponível em: http://periodicos.ufpi.br/index.php/repis/article/view/7237. DOI: https://doi.org/10.26694/repis.v11i1.7237.

[1]Department of Medicine, Afya University Center UNINOVAFAPI, Teresina, Piauí, Brazil.

[2]Department of Business Management, Getúlio Vargas Foundation (FGV), São Paulo, SP, Brazil.



### ABSTRACT

**Introduction**: Infectious diseases remain a major global challenge, particularly in hospital settings. Artificial intelligence (AI) has emerged as a promising tool for the diagnosis, surveillance, and control of these infections. **Objective**: This study aimed to conduct a bibliometric analysis of the scientific literature on the use of AI in infection control. **Methods**: This is a bibliometric, descriptive study with a quantitative approach, based on data from the Web of Science™ Core Collection. The search was conducted in August 2024 and retrieved 1,189 articles. Analyses were performed using RStudio (Bibliometrix and Biblioshiny packages), VOSviewer, and CiteSpace. Variables analyzed included publication trends, most productive authors, institutions, countries, journals, and thematic structure. **Results**: A significant increase in scientific output was observed starting in 2016, with a peak in 2023. The articles were authored by 7,058 researchers and published in 624 journals, with *Scientific Reports* and *PLOS ONE* standing out. China and the United States led in the number of publications. The most productive institutions included the University of California System and Wuhan University. Thematic clustering revealed four main research areas: diagnosis and prognosis, epidemiological surveillance, critical care (such as sepsis), and clinical decision support. **Implications**: The bibliometric findings confirm AI as a consolidated and strategic tool in the fight against infections, particularly following the COVID-19 pandemic. The study highlights the need to increase the participation of low- and middle-income countries, to promote multicenter validations, and to further investigate the clinical applicability of AI-based solutions.

### DESCRITORES

Artificial Intelligence (AI), Infection Control, Bibliometrics, Infectious Diseases, Epidemiology.

**Corresponding Author:**
Ana Raquel Batista de Carvalho
Address: Rua Vitorino Orthiges Fernandes, 6123 - Uruguai, Teresina - PI, Brasil.
ZIP Code: 64073-505
Phone: (86) 2106-0700
Email: ana.raquel@uninovafapi.edu.br

Submitted: 11/07/2025
Accepted: 20/09/2025
Published: 08/10/2025

## INTRODUCTION

Infectious diseases remain one of the major global public health problems. Despite advances in prevention and treatment, they still represent a high burden of morbidity and mortality, being responsible for millions of deaths annually. In addition to the clinical impact, these diseases generate high economic and social costs, both in community and hospital settings, especially when associated with outbreaks and healthcare-associated infections[(1)].

Addressing these diseases requires constant surveillance, early diagnosis, and effective control strategies. However, significant challenges persist, such as underreporting of cases, limited resources, and the rapid spread of resistant pathogens. In this context, it is essential to integrate new technologies that enhance the response capacity to infectious threats[(2)].

In recent decades, technological advances applied to healthcare have transformed the way diseases are understood, prevented, and treated. The use of digital technologies and the application of big data in medicine enable large-scale analyses, pattern identification, and more targeted interventions. In this scenario, the emergence of artificial intelligence (AI) stands out as a tool that has been driving profound changes in different areas of healthcare[(3)].

AI has shown a significant impact on diagnosis, prognosis, and healthcare management. Machine learning models and neural networks, for example, have been applied in the interpretation of imaging tests, prediction of clinical outcomes, and support for medical decision-making. Such applications reinforce its potential as an innovative resource in addressing complex and multifactorial problems, such as infections [(4)].

Specifically in infection control, AI shows promising applications. Its use in epidemiological surveillance enables real-time outbreak monitoring and predictive modeling of infectious agent dissemination. In hospital settings, intelligent algorithms have been employed in the early detection of sepsis, identification of antimicrobial resistance, and prevention of surgical site infections[(2,4)].

The COVID-19 pandemic reinforced this role, acting as a catalyst for the expansion of AI in healthcare. Machine learning tools were applied on multiple fronts, from genomic analysis of the virus to the prediction of epidemiological trends, accelerating the recognition of the relevance of this technology in combating health emergencies[(1,3)].

Despite the advances, there is a lack of systematization regarding how scientific production in the field of AI applied to infection control has evolved. Few studies comprehensively explore publication trends, main thematic areas, collaborations between countries and institutions, as well as the most influential scientific actors. At this point, bibliometrics emerges as a methodological tool capable of mapping such aspects[(5)].

Understanding the evolution of research in this field is essential to guide efforts in innovation and public policy formulation. Bibliometric studies make it possible to identify areas of greatest impact, underexplored gaps, and opportunities for integration among researchers. In addition, they can support managers and healthcare professionals in defining strategic priorities for addressing infectious diseases[(5)]. In light of this, the present study aimed to conduct a bibliometric analysis of the scientific literature on the use of artificial intelligence in infection control.

## MÉTODOS

This is a bibliometric study, with a descriptive design, quantitative approach, and documentary nature. Bibliometrics consists of a quantitative research approach applied to the scientific literature, which makes it possible to measure and analyze the production and communication of scientific knowledge[(6)].

Through this methodology, it is possible to evaluate the productivity of authors and institutions, identify collaboration networks, map emerging areas, and measure the impact of publications in the investigated field. In addition, bibliometrics allows monitoring growth patterns and the dissemination of knowledge over time, addressing relevant questions about scientific progress in a given area[(7)].

When combined with statistical and visualization tools, the method also enables exploration of the thematic structure of studies, highlighting publication dynamics, emerging trends, and gaps in scientific production[(8)].

The present study was conducted in accordance with the five steps recommended by Donthu et al. (2021)[(9)] for bibliometric analyses, namely: (1) definition of the research question, (2) selection of the database, (3) formulation of the search strategy, (4) data extraction and analysis, and (5) interpretation and presentation of the results.

In addition, this work follows the guidelines of the Preferred Reporting Items for Bibliometric Analysis (PRIBA)[10] protocol, which establishes 25 items distributed across seven essential methodological domains, ensuring greater transparency, reproducibility, and standardization of bibliometric studies in the health field.

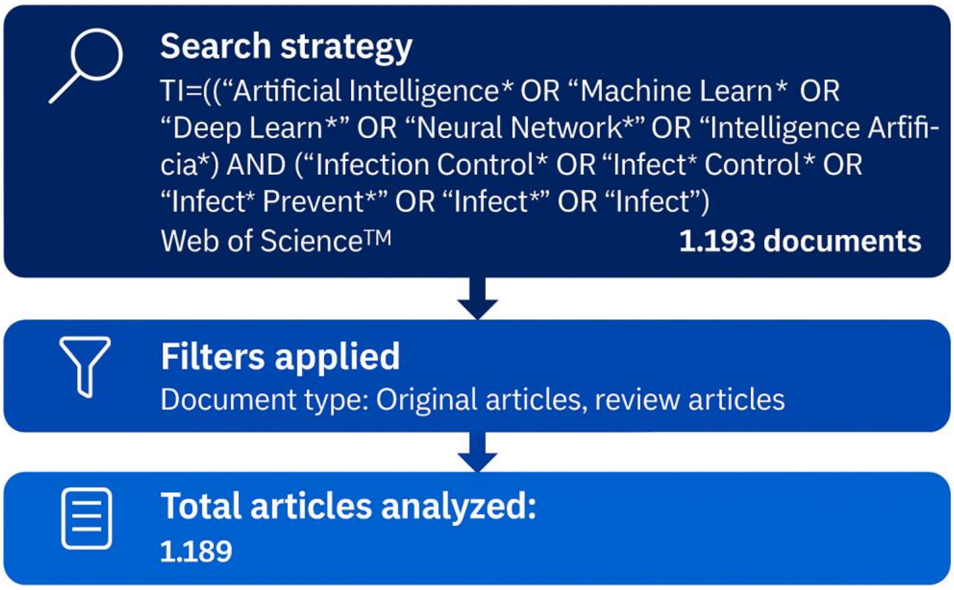
The research question that guided this study was: What are the bibliometric indicators of scientific production related to the use of artificial intelligence in infection control? The search was carried out in the Web of Science™ Core Collection (WoSCC) database on August 24, 2024, and covered all documents indexed up to that date. The WoSCC was selected for its wide coverage of high-impact scientific journals and the rigor in metadata indexing, which makes it particularly suitable for bibliometric studies[11,12].

To avoid bias resulting from automatic database updates and temporal fluctuations in citation counts, data extraction was performed on a single day, which ensures greater consistency and reliability of the results[13].

The construction of the search strategy was based on the identification of controlled and uncontrolled descriptors. For this purpose, the structured vocabulary of Medical Subject Headings (MeSH) was used, complemented by free terms and variants commonly employed in the scientific literature. The terms were combined using the Boolean operators AND and OR, and truncation (*) was applied to include morphological variations of the words. The search was performed in the title field (TI=), aiming for greater specificity and alignment with the study objectives. The choice of the title field was based on methodological evidence indicating that this filter reduces false positives and increases sample precision in bibliometric analyses[14].

Original and review articles were included without time restrictions, and opinion articles, reflections, editorials, and case studies were excluded, according to the predefined criteria. The initial search resulted in 1,193 documents retrieved from the WoSCC database. After screening and applying the eligibility criteria, 1,189 articles were selected to comprise the analysis sample. Figure 1 presents the search strategy and the selection process of the articles included in the study.

**Figure 1.** Search strategy used and selection of included articles.




The bibliographic data of the articles included in the final sample were exported from the Web of Science™ Core Collection (WoSCC) database in text format (.txt), using the Full Record and Cited References option, which allowed complete recovery of the information necessary for bibliometric analysis[15].

The files were imported into the RStudio Desktop environment (version 2023.06.2+561), integrated with R software (version 4.3.1), and processed using the Bibliometrix 4.3.3 package, a specific tool for advanced bibliometric analyses. The Biblioshiny application was also used, providing an interactive graphical interface for Bibliometrix, facilitating data exploration and visualization. Bibliometrix, a package developed in the R language, was used to perform mathematical and statistical calculations, such as publication frequencies, number and percentage of citations by journal, author, country, and institution. The Biblioshiny interface was employed to facilitate data exploration through interactive graphical tools(16).

In addition to these tools, VOSviewer (version 1.6.20) and CiteSpace (version 6.3.R1) were used, enabling dynamic and interactive visualization of knowledge networks. VOSviewer is a bibliometric analysis software developed by Nees Jan van Eck and Ludo Waltman for constructing and visualizing web-based bibliometric maps, from which relevant information can be extracted from a large set of publications(17).

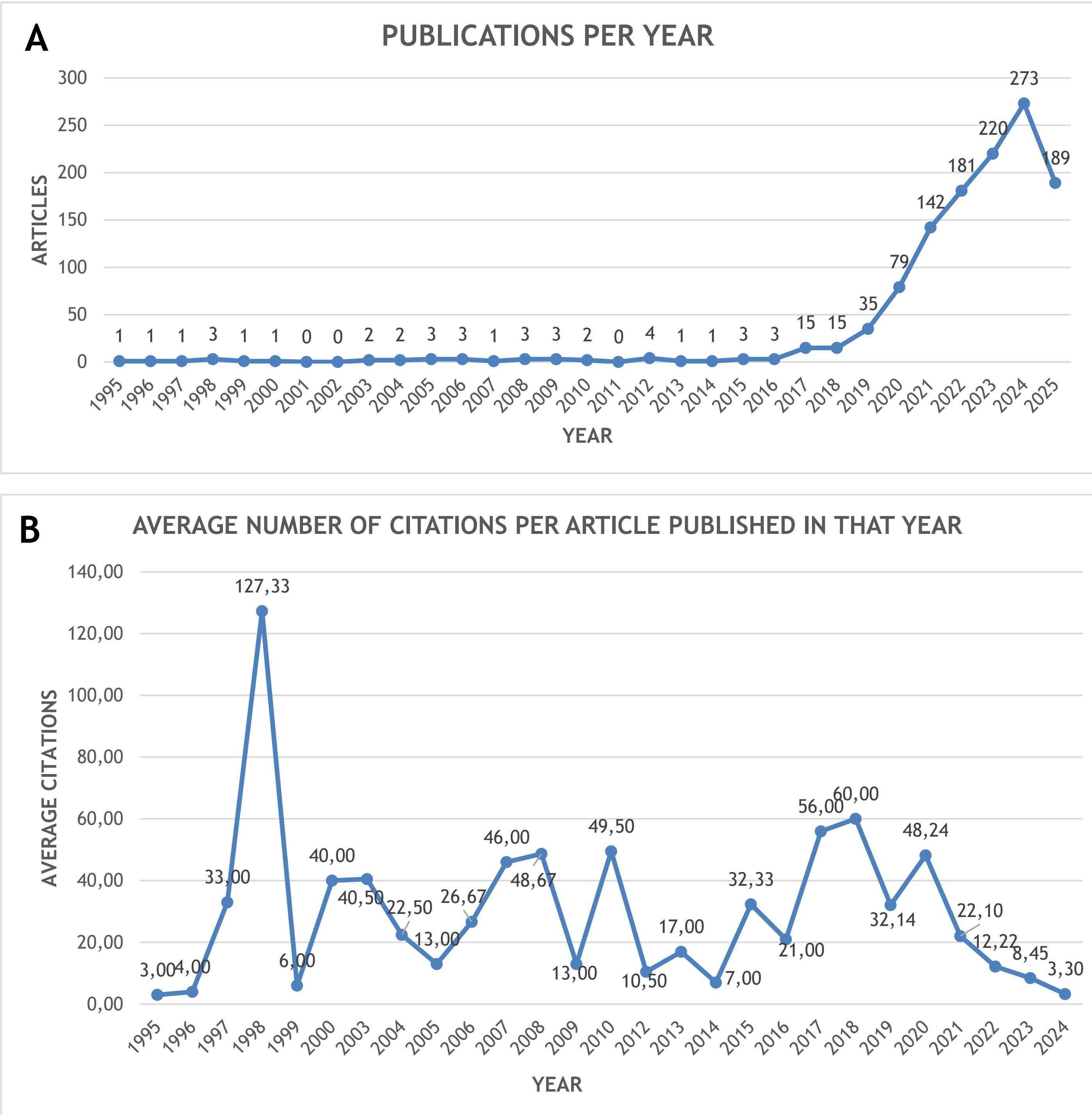
This program allows the creation of visual representations of scientific networks, including co-authorship, co-citation, keyword co-occurrence, and bibliographic coupling relationships. In the generated visualizations, different clusters are indicated by the color of the nodes, the number of publications by the size of the nodes, and the strength of the relationships by the thickness of the connecting lines.

## RESULTADOS

The bibliometric analysis encompassed a total of 1,189 documents published between 1995 and 2025, distributed across 624 different journals. These studies accumulated 44,520 references and presented an average of 13.6 citations per document, with an annual citation rate of approximately 2.89 citations per article. The evolution of scientific production is shown in Figure 2A, which demonstrates a steady increase in international interest in the topic, with modest output between 1995 and 2015, followed by a significant rise from 2016 onward, reaching its peak in 2023, with 273 publications.

Regarding impact, Figure 2B presents the average number of citations per article over the study period. It is observed that, in the early years, although the number of publications was limited, some studies achieved high visibility—for instance, in 1998, when a single article averaged 127 citations, representing the highest impact in the historical series. Conversely, in more recent years, such as 2018, the average impact also stood out, reaching 60 citations per article.

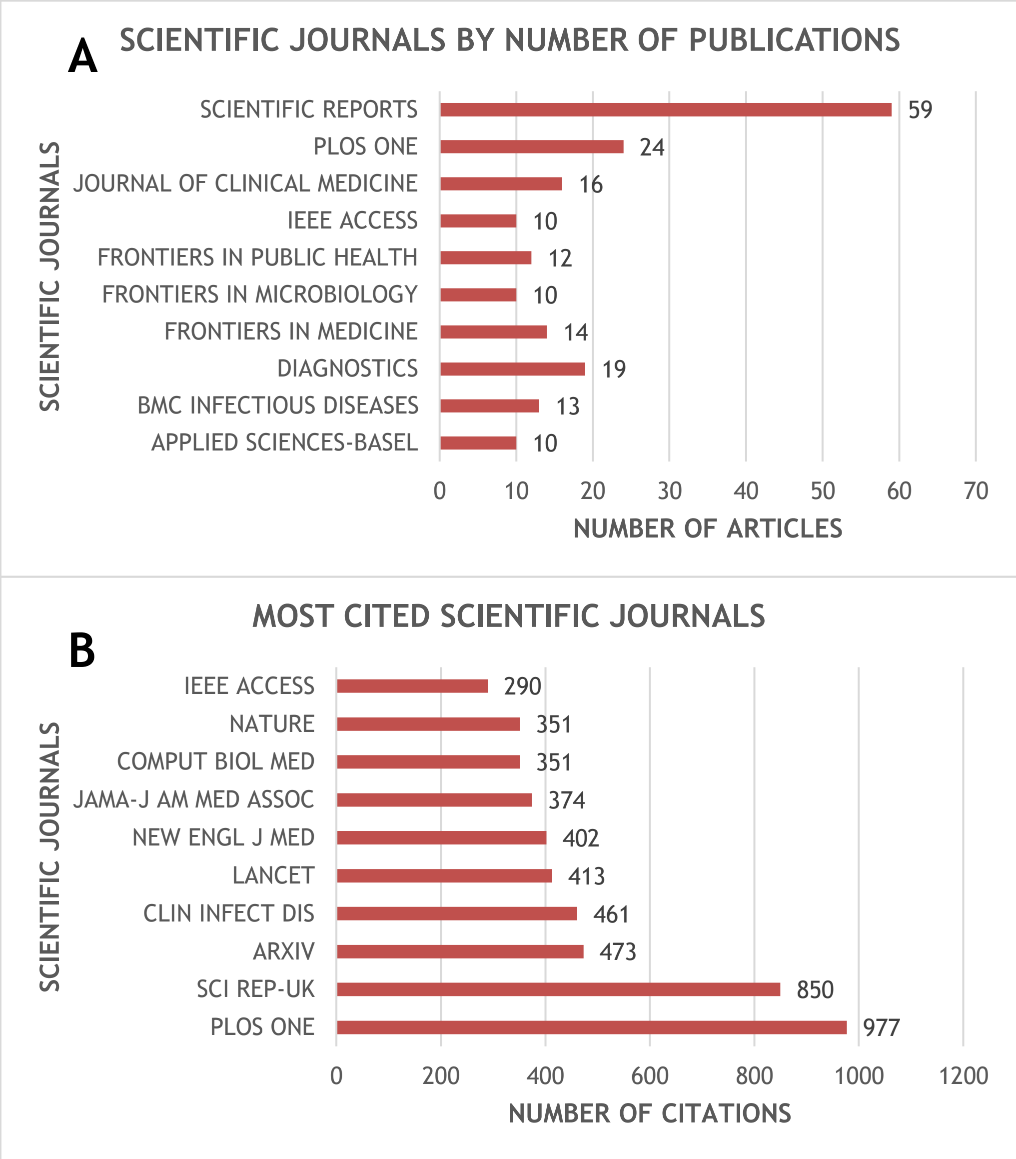
**Figure 2.** Annual distribution of articles according to year of publication. Teresina, Piauí, Brazil.




Among the most productive journals, the highlights include Scientific Reports (59 articles), PLOS ONE (24), Diagnostics (19), Journal of Clinical Medicine (16), Frontiers in Medicine (14), BMC Infectious Diseases (13), and Frontiers in Public Health (12), reflecting a strong presence of multidisciplinary and open-access journals.

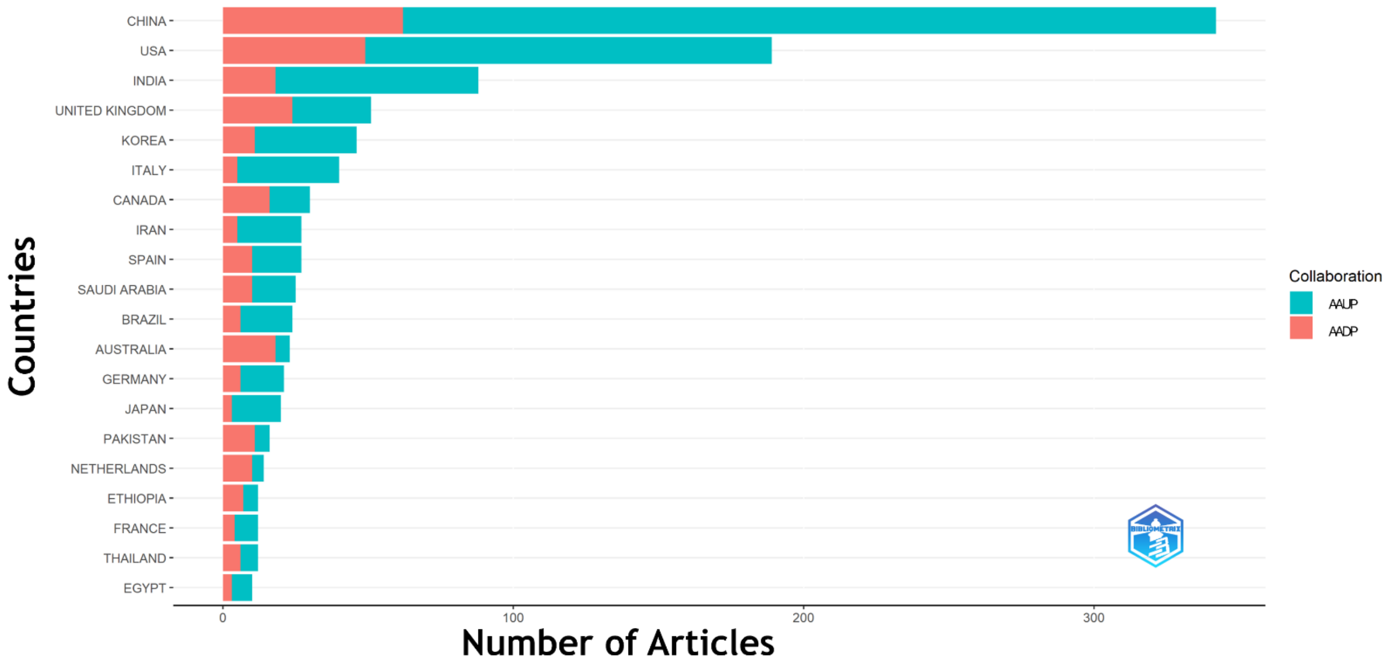
Regarding citation impact, there was a predominance of high-prestige international journals. PLOS ONE (977 citations) and Scientific Reports (850) accounted for the largest number of local citations, followed by arXiv (473), Clinical Infectious Diseases (461), The Lancet (413), New England Journal of Medicine (402), and JAMA (374).

**Figure 3.** Publications and citations of scientific journals. Teresina, Piauí, Brazil.




The authors of the analyzed articles were affiliated with 71 different countries. The analysis of the corresponding authors' countries of affiliation highlighted the leadership of China and the United States, followed by India and the United Kingdom. A predominance of single-country publications (SCP) was observed in the leading countries, indicating a strong capacity for independent scientific production. In contrast, countries such as the United Kingdom and South Korea showed a higher proportion of multiple-country publications (MCP) (Figure 4).

**Figure 4.** Collaboration among countries according to the corresponding author's affiliation. Teresina, Piauí, Brazil.

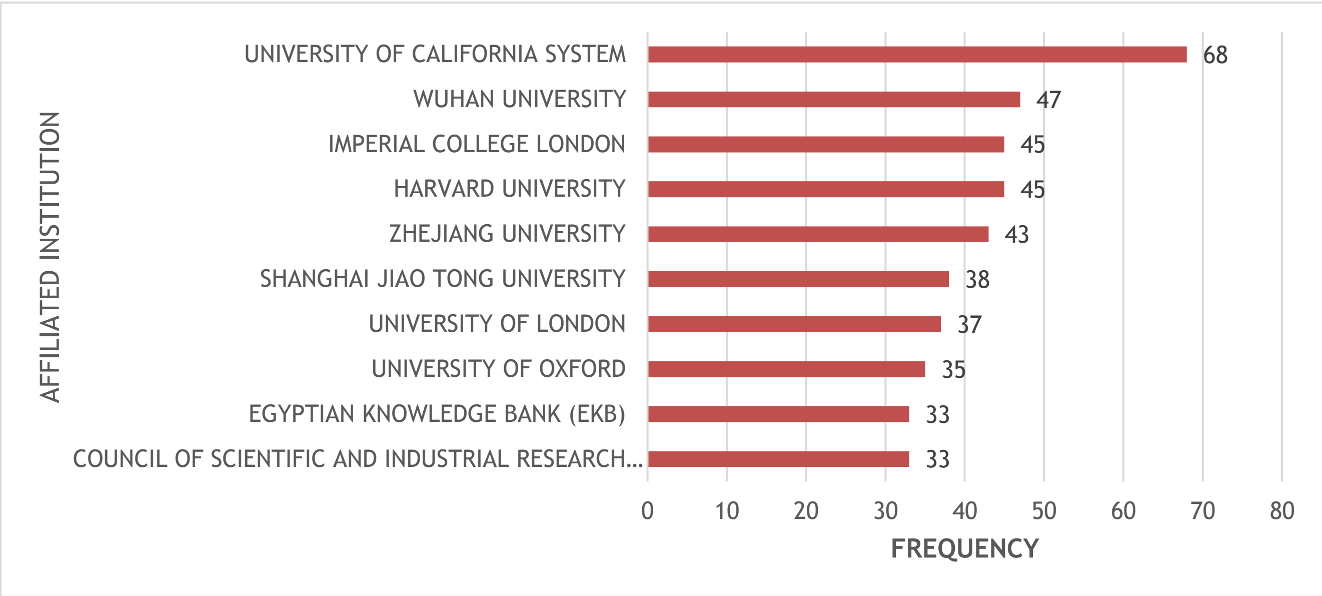



Legend: *SCP = Single-Country Publication; MCP = Multiple-Country Publication.*
**Source:** *Web of Science™*.

A total of 50 different institutions were identified in the authors' affiliations. The analysis revealed a strong concentration among certain universities of international prominence, while most institutions showed a low frequency of publications.

The University of California System led the ranking with 68 occurrences, followed by Wuhan University (47), Imperial College London (45), and Harvard University (45). Other relevant institutions include Zhejiang University (43), Shanghai Jiao Tong University (38), University of London (37), and University of Oxford (35), evidencing the prominent role of North American, European, and Asian academic centers in the scientific production on artificial intelligence applied to infection control (Figure 5).

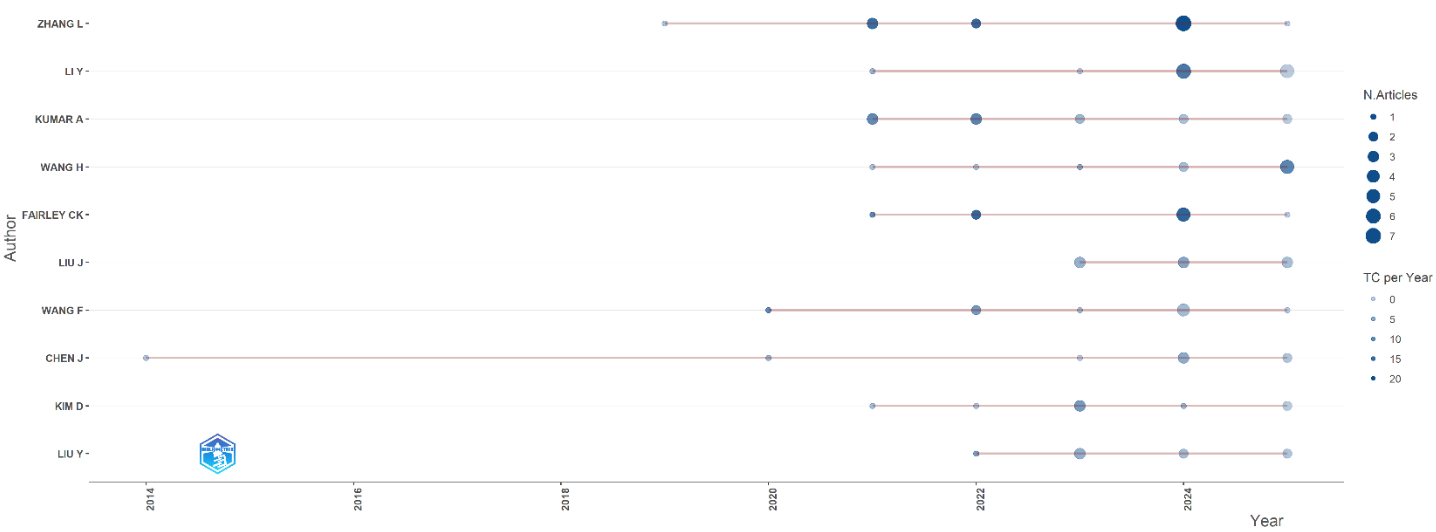
**Figure 5.** Leading institutions according to the frequency of authors' affiliations. Teresina, Piauí, Brazil.




A total of 7,058 different authors contributed to the analyzed articles, demonstrating the broad participation of researchers in the investigated topic. Figure 6 shows the evolution of author productivity over time. The size of the bubbles is proportional to the number of published articles (larger bubbles represent a higher number of publications), while the shades of blue are proportional to the number of citations (darker blue bubbles indicate greater citation impact).

In terms of productivity, Zhang L. (n = 7) stood out, followed by Li Y. and Kumar A. (n = 5 each), as well as other authors such as Wang H. and Fairley C.K., who also made relevant contributions in both publication volume and citation impact. It is noteworthy that Chen J. was the first author to publish on the topic within the analyzed period, establishing himself as a pioneer in the early studies in this field.

**Figure 6.** Leading authors according to the number of publications over time. Teresina, Piauí, Brazil.




A total of 1,189 articles were cited 12,403 times, resulting in an average of 10.43 citations per article. The 15 most cited articles accumulated between 339 and 120 citations, as presented in Table 1. These publications were disseminated across ten different journals, covering the period from 1998 to 2023.

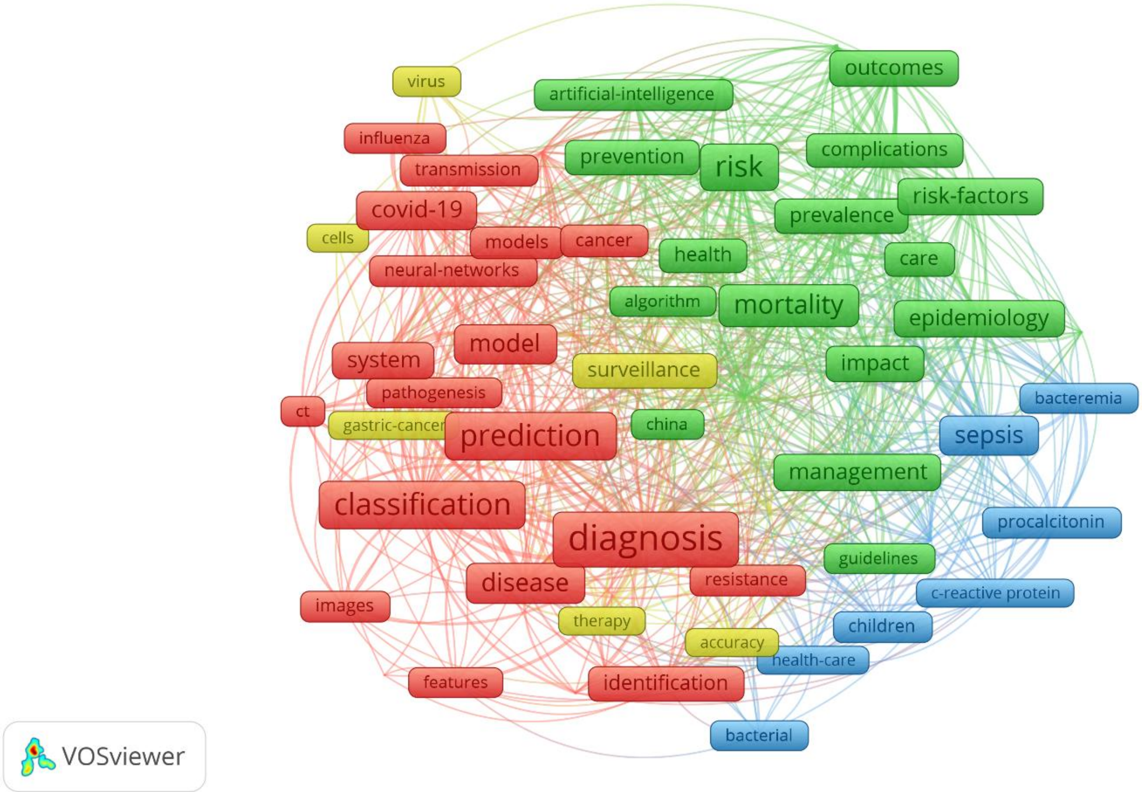
**Table 1.** Ranking of the most cited articles related to artificial intelligence applied to infection control. Teresina, Piauí, Brazil.

| AUTHOR (YEAR), SCIENTIFIC JOURNAL | TITLE | TOTAL CITATIONS | AVERAGE ANNUAL |
|---|---|---|---|
| Goodacre et al., (1998)[2], Microbiol-Sgm | Rapid identification of urinary tract infection bacteria using hyperspectral whole-organism fingerprinting and artificial neural networks | 339 | 12,11 |
| Peiffer-Smadja, (2020), Clin Microbiol Infec | Machine learning for clinical decision support in infectious diseases: a narrative review of current applications | 300 | 50,00 |
| Hu et al., (2020), IEEE Access | Weakly Supervised Deep Learning for COVID-19 Infection Detection and Classification From CT Images | 219 | 36,50 |
| Dechant et al., (2017), Phytopathology | Automated Identification of Northern Leaf Blight-Infected Maize Plants from Field Imagery Using Deep Learning | 210 | 23,33 |
| Shichijo (2017), Ebiomedicine | Application of Convolutional Neural Networks in the Diagnosis of Helicobacter pylori Infection Based on Endoscopic Images | 204 | 22,67 |
| Alakus et al., (2020), Chaos Soliton Fract | Comparison of deep learning approaches to predict COVID-19 infection | 195 | 32,50 |
| Brinati et al., (2020), J Med Syst | Detection of COVID-19 Infection from Routine Blood Exams with Machine Learning: A Feasibility Study | 190 | 31,67 |
| Chae et al., (2018) Int J Env Res Pub He | Predicting Infectious Disease Using Deep Learning and Big Data | 172 | 21,50 |
| Singh et al., (2019), IEEE Access | Multilayer Convolution Neural Network for the Classification of Mango Leaves Infected by Anthracnose Disease | 172 | 24,57 |
| Hassantabar et al., (2020) Chaos Soliton Fract | Diagnosis and detection of infected tissue of COVID-19 patients based on lung x-ray image using convolutional neural network approaches | 162 | 27,00 |
| Sedik et al., (2020), Viruses-Basel | Deploying Machine and Deep Learning Models for Efficient Data-Augmented Detection of COVID-19 Infections | 132 | 22,00 |
| Taylor et al., (2018) Plos One | Predicting urinary tract infections in the emergency department with machine learning | 126 | 15,75 |
| Itoh et al., (2018), Endosc Int Open | Deep learning analyzes Helicobacter pylori infection by upper gastrointestinal endoscopy images | 123 | 15,38 |
| Majid et al., (2020), Microsc Res Techniq | Classification of stomach infections: A paradigm of convolutional neural network along with classical features fusion and selection | 121 | 20,17 |
| Wong et al., (2023) Science | Leveraging artificial intelligence in the fight against infectious diseases | 120 | 40,00 |

The co-occurrence analysis of the 10 most frequent KeyWords Plus™ revealed the formation of four major thematic clusters. The red cluster is centered on *diagnosis*, *prediction*, *classification*, and *disease*, including terms such as *COVID-19*, *model*, *neural networks*, and *system*, which reflects the application of artificial intelligence in diagnostic and predictive models focused on emerging infectious diseases.

The green cluster is structured around *risk*, *mortality*, *epidemiology*, and *prevalence*, and is associated with *surveillance*, *complications*, *outcomes*, and *impact*, highlighting the role of AI in epidemiological surveillance, risk factor analysis, and clinical prognosis. The blue cluster comprises terms such as *sepsis*, *bacteremia*, *procalcitonin*, and *C-reactive protein*, emphasizing the use of algorithms for risk stratification, biomarker identification, and management of critically ill patients in intensive care units. Finally, the yellow cluster connects *identification*, *guidelines*, *accuracy*, and *therapy*, representing the efforts toward methodological standardization, validation, and clinical decision support based on artificial intelligence.

**Figure 7.** Conceptual structure of articles related to the application of artificial intelligence in infection control. Teresina, Piauí, Brazil.




## DISCUSSÃO

The temporal analysis revealed a significant growth in scientific production over the evaluated period. Between 1995 and 2015, production remained modest, with a limited number of annual publications. From 2016 onward, however, there was a progressive increase, reaching its peak in 2023, with 273 published articles.

This pattern aligns with the consolidation dynamics of emerging fields, in which the expansion of international interest results in a greater volume of publications and thematic diversification. However, the analysis of the average impact per article shows that this quantitative expansion was not uniformly accompanied by a proportional increase in scientific relevance.

In 1998, for instance, a single article reached an average of more than 120 citations, representing a highly influential outlier during a period of low productivity. This phenomenon is described in bibliometric studies as characteristic of pioneering works, whose originality or conceptual value gives them a central role in the subsequent development of the field[33].

Conversely, more recent years show an average impact more evenly distributed among a larger number of publications, indicating the consolidation of the field, but also the dilution of prestige among different contributions[34].

The substantial increase in scientific production within the scope of this research, particularly since the mid-2010s, may be related to technological advances and emerging health demands. The exponential rise in the availability of clinical and epidemiological data, from surveillance systems, electronic health records, and telemedicine platforms, created favorable conditions for the development and application of machine learning algorithms in infectious contexts, in line with the broader movement of AI in healthcare.

Furthermore, the COVID-19 pandemic played a catalytic role in this process, stimulating the use of

AI for transmission modeling, imaging-based diagnosis, contact tracing, and outbreak prediction, which accelerated the interdisciplinary convergence between AI and infection control[35].

Emerging technologies, such as deep learning applied to pattern recognition in laboratory and surveillance data, have further expanded the scope of investigations and fostered thematic diversification. Finally, the consolidation of AI as an indispensable tool in combating infectious diseases also stems from its role in complex contexts, where outbreak anticipation, optimization of control measures, and data-driven decision-making have proven fundamental to public health. Bibliometric literature reinforces that fields with strong practical applicability and social relevance tend to experience accelerated scientific growth precisely during systemic challenges, as observed during the COVID-19 pandemic[36,37].

The analysis of scientific journals revealed two complementary patterns. On one hand, production was concentrated in multidisciplinary and open-access journals, such as *Scientific Reports* and *PLOS ONE*, which stood out for their high publication volume, as well as in high-prestige journals, including *The Lancet*, *New England Journal of Medicine*, and *JAMA*, which accounted for a large number of citations. On the other hand, there was notable relevance of journals directly related to the investigated theme, such as *BMC Infectious Diseases*, *Frontiers in Public Health*, and *American Journal of Infection Control*, in addition to *Clinical Infectious Diseases*, which, although less productive, accumulated 461 local citations and established itself as a highly influential source.

Interdisciplinary journals focused on the application of artificial intelligence also stand out, such as *Computers in Biology and Medicine* and *Computers and Electronics in Agriculture*, which reinforce the interface between computational methods and infection control. This body of literature demonstrates both the broad dissemination of the topic in international outlets and its progressive specialization in journals dedicated to public health, infectious diseases, and computational sciences[38,39].

The diversification of journals, in turn, parallels the configuration of scientific collaboration networks among countries. The predominance of China and the United States in this field reflects not only their publication volume but also distinct research organization strategies. China is primarily focused on single-country authorship initiatives, driven by strong governmental investments in artificial intelligence applied to health, whereas the United States exhibits greater diversification of collaborations, including international partnerships, a characteristic of more decentralized and interconnected research systems[40,41].

This differentiation suggests that while China prioritizes the consolidation of self-sufficient scientific leadership, the United States expands its influence through global cooperation networks, which may favor greater impact and knowledge dissemination.

Comparative studies further indicate that European and Asian countries, such as the United Kingdom and South Korea, play a strategic role by integrating into international partnerships, establishing themselves as intermediary nodes in the global networks of scientific production [40,41].

The concentration of scientific output in institutions such as the University of California System, Harvard University, Imperial College London, and Wuhan University reflects the pattern described in bibliometric analyses, whereby centers with robust infrastructure and greater funding capacity tend to accumulate scientific prominence [42].

This scenario highlights the strategic role of North American, European, and Asian universities in leading research on artificial intelligence applied to health, consolidating large-scale, multidisciplinary projects[43].

The presence of Chinese institutions among the most productive aligns with national policies aimed at leadership in artificial intelligence, reinforced by the context of the COVID-19 pandemic, which intensified the demand for innovative solutions in surveillance, diagnosis, and infection control[43].

In the same vein, the analysis of authors reinforces China's centrality in this research field, evidenced both by the volume of publications and by the influence of its researchers. Notable names include Zhang L. and Fairley C.K., whose productivity and citation impact have established them as foundational references.

Among the most cited authors, Chen J.[44] stands out as the first to publish on the topic in 2014, in a study that used artificial neural networks to predict risk factors for nosocomial infection in patients with lung cancer. In this study, which involved 609 individuals, the main predictors identified were length of hospitalization, advanced age, and clinical stage of the disease.

His contribution became one of the most influential works, demonstrating that foundational articles can exert disproportionate impact, as well as showing the feasibility and advantage of artificial intelligence over conventional statistical methods in infection prediction, establishing itself as a foundational and highly influential milestone for subsequent research in the field.

The concentration of citations in a restricted group of articles reflects the phenomenon described by Price (1963)[45], whereby science advances through highly influential contributions that shape the direction of subsequent studies. Among the 15 most cited articles, stand out publications that combine the development of deep learning and machine learning models with clinically relevant problems, such as COVID-19 diagnosis, bacterial infection identification, and risk prediction in hospital settings [19,24,25].

The pioneering study by Goodacre et al. (1998)[18], which employed artificial neural networks to identify bacteria in urinary infections using advanced spectroscopy, inaugurated a new methodological direction. Meanwhile, the study by Peiffer-Smadja et al. (2020)[19] consolidated the role of narrative reviews by examining the use of machine learning in clinical decision support for infectious diseases, highlighting areas such as diagnosis, sepsis prediction, and antimicrobial resistance. Complementarily, the publication by Wong et al. (2023)[32] addressed how AI has driven significant advances in the diagnosis, treatment, and understanding of infectious diseases, consolidating its global applicability.

The thematic cohesion of these studies, which range from spectroscopy-based neural network methods to emergency clinical applications and synthetic reviews, illustrates both the historical depth and contemporary evolution of the field. This reinforces the notion that scientific impact depends on both pioneering contributions and integrative works that consolidate and disseminate knowledge[46].

In this context, the first cluster revealed the predominance of applications focused on the use of artificial intelligence in diagnostic and prognostic processes, particularly in high-impact clinical scenarios, such as the COVID-19 pandemic. This focus demonstrates how machine learning algorithms have been directed toward supporting clinical decision-making and optimizing the early detection of infectious diseases.

The second axis comprised investigations related to epidemiological surveillance and risk factor analysis, highlighting the role of AI in outbreak monitoring, prevalence forecasting, and identification of adverse health outcomes in public health. The remaining clusters, in turn, point to more specific and emerging areas, such as risk stratification in critical conditions (sepsis, bacteremia) and the development of standardized models for clinical validation and application.

This configuration demonstrates that, although the field has advanced rapidly in diagnostic solutions, there remains a need to expand research to evaluate economic impact, cost-effectiveness, and applicability across different health systems. Future studies that prioritize multicenter validations and integration between data science, epidemiology, and clinical practice may strengthen algorithm robustness and promote their large-scale adoption.

## LIMITATIONS

Although the bibliometric analysis provided a comprehensive overview of the scientific production on artificial intelligence applied to infection control, some limitations must be acknowledged.

The choice of Web of Science (WoS) as the sole data source may have restricted the scope of the research, since other relevant databases, such as Scopus, PubMed, and Embase, were not included. While this decision ensures rigor and standardization, it may have resulted in the exclusion of significant articles indexed in other repositories. Another aspect to consider is the search strategy, which was applied only to article titles. This restriction ensures greater specificity to the topic but may exclude studies whose main focus is described only in the abstract or keywords.

Finally, the results also reflect the indexing criteria and algorithms specific to WoS, which may introduce biases related to journal selection or thematic classification. Nevertheless, consistent methodological procedures were established, including well-defined descriptors, clear inclusion and exclusion criteria, and the use of established tools for bibliometric analysis. These measures strengthen the reliability and robustness of the findings presented.

## CONCLUSION

The bibliometric analysis revealed a progressive growth in scientific production on artificial intelligence applied to infection control, with fluctuations over the years and a significant increase in 2024. The studies are concentrated in multidisciplinary, open-access journals and in countries with a strong research tradition, such as China and the United States, while the thematic clusters revealed three main axes: diagnosis and prognosis, epidemiological surveillance, and applications in critical conditions, such as sepsis.

For future research, it is essential to expand the participation of low- and middle-income countries, conduct multicenter validation of models, and investigate not only diagnostic accuracy, but also effectiveness, cost-effectiveness, and clinical applicability of artificial intelligence–based solutions. Another key point is the strengthening of algorithm transparency and interpretability, which is central to their adoption in clinical practice and public health policies.

Therefore, this study contributes by mapping the main trends and gaps in the literature, providing foundations for scientific advancement and the development of innovative strategies for infection prevention and control. These findings may guide future investigations and support policymakers and healthcare professionals in implementing safer and more effective technologies.

## REFERENCES

1. Zhao Y, Li J, Li X, Xu J. Application of artificial intelligence in the prevention and control of infectious diseases. China CDC Wkly. 2024;6(30):642-6. doi:10.46234/ccdcw2024.192.

2. Chen Y, Li Y, Narayan S, Subramanian N, Roland T. Application of artificial intelligence in infectious disease management: advances and prospects. Front Pharmacol. 2022;13:1079566. doi:10.3389/fphar.2022.1079566.

3. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol. 2017;2(4):230-43. doi:10.1136/svn-2017-000101.

4. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. Nat Med. 2019;25(1):30-6. doi:10.1038/s41591-018-0307-0.

5. Ferreira VLR, Silva Júnior JB, Garcia LP. Inteligência artificial aplicada ao enfrentamento de epidemias: revisão e perspectivas. Rev Saude Publica. 2020;54:80. doi:10.11606/s1518-8787.2020054002475.

6. Hutchins BI, Yuan X, Anderson JM, Santangelo GM. Relative Citation Ratio (RCR): a new metric that uses citation rates to measure influence at the article level. PLoS Biol. 2016;14(9):e1002541.

7. Vošner HB, Kokol P, Bobek S, Železnik D, Završnik J. A bibliometric retrospective of the Journal Computers in Human Behavior (1991–2015). Comput Human Behav. 2016;65:46-58.

8. Zupic I, Čater T. Bibliometric methods in management and organization. Organizational Research Methods. 2015;18(3):429-472.

9. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM. How to conduct a bibliometric analysis: an overview and guidelines. J Bus Res. 2021;133:285–96.

10. Koo M, Lin SC. An analysis of reporting practices in the top 100 cited health and medicine-related bibliometric studies from 2019 to 2021 based on a proposed guidelines. Heliyon. 2023 May 27;9(6):e16780.

11. Moura MEB, et al. Global trends from original research on COVID-19 and coinfection. Revista

Prevenção de Infecção e Saúde. 2022;8(1).

12. Julien A, et al. Analysis of scientific publications related to Hepatitis B prevention in the last decade: 2013-2023. Revista Prevenção de Infecção e Saúde. 2024;10(1).

13. Lin HY, Deng X, Song D. Research trends of global artificial intelligence application in obstetrics and gynecology from 1999 to 2025: a bibliometric analysis based on Web of Science. J Robot Surg. 2025;19(1):606.

14. Portela DMMC, et al. Treatment of Parkinson's disease by deep brain stimulation: a bibliometric analysis. Sao Paulo Med J. 2024;142(5):e2023187.

15. Aria M, Cuccurullo C. bibliometrix: An R-tool for comprehensive science mapping analysis. J Informetrics. 2017;11(4):959-975.

16. Büyükkıdık S. A bibliometric analysis: A tutorial for the bibliometrix package in R using IRT literature. J Meas Eval Educ Psychol. 2022;13(3):164-193.

17. van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics. 2010;84:523–538.

18. Goodacre S, Rooney K, Slater K, et al. Rapid identification of urinary tract infection bacteria using hyperspectral whole-organism fingerprinting and artificial neural networks. Microbiol-Sgm. 1998;144:339–349.

19. Peiffer-Smadja N. Machine learning for clinical decision support in infectious diseases: a narrative review of current applications. Clin Microbiol Infect. 2020;26:300–310.

20. Hu S, Gao Y, Wang L, et al. Weakly Supervised Deep Learning for COVID-19 Infection Detection and Classification From CT Images. IEEE Access. 2020;8:219–225.

21. Dechant R, Poland S, Wisser RJ, et al. Automated Identification of Northern Leaf Blight-Infected Maize Plants from Field Imagery Using Deep Learning. Phytopathology. 2017;107:210–220.

22. Shichijo S. Application of Convolutional Neural Networks in the Diagnosis of Helicobacter pylori Infection Based on Endoscopic Images. Ebiomedicine. 2017;18:204–213.

23. Alakus TB, et al. Comparison of deep learning approaches to predict COVID-19 infection. Chaos Soliton Fract. 2020;139:195–205.

24. Brinati D, et al. Detection of COVID-19 Infection from Routine Blood Exams with Machine Learning: A Feasibility Study. J Med Syst. 2020;44:190.

25. Chae S, et al. Predicting Infectious Disease Using Deep Learning and Big Data. Int J Environ Res Public Health. 2018;15:172.

26. Singh D, et al. Multilayer Convolution Neural Network for the Classification of Mango Leaves Infected by Anthracnose Disease. IEEE Access. 2019;7:172–184.

27. Hassantabar S, et al. Diagnosis and detection of infected tissue of COVID-19 patients based on lung x-ray image using convolutional neural network approaches. Chaos Soliton Fract. 2020;138:162–172.

28. Sedik A, et al. Deploying Machine and Deep Learning Models for Efficient Data-Augmented Detection of COVID-19 Infections. Viruses. 2020;12:132.

29. Taylor RA, et al. Predicting urinary tract infections in the emergency department with machine learning. PLOS ONE. 2018;13:e126.

30. Itoh T, et al. Deep learning analyzes Helicobacter pylori infection by upper gastrointestinal endoscopy images. Endosc Int Open. 2018;6:123–130.

31. Majid M, et al. Classification of stomach infections: A paradigm of convolutional neural network along with classical features fusion and selection. Microsc Res Tech. 2020;83:121–131.

32. Wong A, et al. Leveraging artificial intelligence in the fight against infectious diseases. Science. 2023;379:120–130.

33. Chen C, Dubin R, Kim MC. Emerging trends and new developments in regenerative medicine: a scientometric update (2000–2014). Expert Opin Biol Ther. 2014;14(9):1295-317. doi: 10.1517/14712598.2014.920813.

34. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM. How to conduct a bibliometric analysis: An overview and guidelines. J Bus Res. 2021;133:285-96. doi: 10.1016/j.jbusres.2021.04.070.

35. Vaishya R, Javaid M, Khan IH, Haleem A. Artificial Intelligence (AI) applications for COVID-19 pandemic. Diabetes Metab Syndr. 2020;14(4):337-339. doi: 10.1016/j.dsx.2020.04.012.

36. L Arab R, Adel R, et al. Inteligência artificial na prevenção de infecções hospitalares: uma revisão integrativa. Front Public Health. 2025;13:1547-450.

37. Pham Q-V, Nguyen DC, Huynh-The T, Hwang WJ, Pathirana PN. Artificial intelligence (AI) and big data for coronavirus (COVID-19) pandemic: a survey on the state-of-the-arts. IEEE Access. 2020;8:130820-130839.

38. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56.

39. Piwowar H, Priem J, Larivière V, Alperin JP, Matthias L, Norlander B, et al. The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. PeerJ. 2018;6:e4375.

40. Kitajima K, et al. The altering landscape of US–China science collaboration: From convergence to divergence. Palgrave Communications / Nature (Humanities & Social Sciences Comms) 2025;12.

41. Zhu Y, Kim D, Yan E, Kim MC, Qi G. Analyzing China's research collaboration with the United States in high-impact and high-technology research. Quantitative Science Studies. 2021;2(1):363-375.

42. Xie Y, Zhai Y, Lu G. Evolution of artificial intelligence in healthcare: a 30-year bibliometric study. Front Med (Lausanne). 2025;11:1505692.

43. Kitajima K, Okamura K. Evolving landscape of US-China science collaboration: Convergence and divergence. *arXiv [Preprint].* 2023; arXiv:2309.05033.

44. Chen J, Zhang L, Hu J, Liu L, Chen W. Use of an artificial neural network to predict risk factors of nosocomial infection in lung cancer patients. *Asian Pac J Cancer Prev.* 2014;15(13):5349-5353.

45. Price DJ de S. Little Science, Big Science. New York: Columbia University Press; 1963.

46. Cao S, Wei Y, Yue Y, Wang D, Xiong A, Yang J, Zeng H. Uncovering the scientific landscape: A

bibliometric and visualized analysis of artificial intelligence in Traditional Chinese Medicine. Heliyon. 2024;10(18):e37439.

**ARTICLE ORIGIN**

Original article.

**AUTHOR CONTRIBUTIONS**

Sousa EBN and Nobre LMAA made substantial contributions to the conception of the research, methodological and organizational design of the study, as well as to the critical review of the intellectual content. Lima EL, Monte BKS, Silva JW made substantial contributions to the data analysis and to the review of the intellectual content. Carvalho ARB made substantial contributions to the writing of the manuscript and to the discussion of the data.

**RESEARCH ETHICS COMMITTEE APPROVAL**

Not applicable.

**CONFLICT OF INTERESTS**

The authors declare no conflict of interest.