


Trends and Current Topics in the Field of Artificial Intelligence in Hospitals: A Text Mining Analysis

Mahnaz Mohseni¹, Meisam Dastani^{2*} 

¹ Department of Knowledge and Information Science, Payame Noor University, Tehran, Iran

² Infectious Diseases Research Center, Gonabad University of Medical Sciences, Gonabad, Iran

*Corresponding Author: Meisam Dastani, Ph.D., Infectious Diseases Research Center, Gonabad University of Medical Sciences, Gonabad, Iran. Tel: +98-9306404852, Email: Meisam.dastani@gmail.com

Received September 11, 2025; Accepted December 23, 2025; Online Published March 30, 2026

Abstract

Background: Artificial Intelligence (AI), as a transformative technology, has found widespread applications in the health and hospital sectors.

Objectives: The present study aimed to analyze scientific articles related to AI in hospitals using text mining methods to identify dominant topics and emerging trends.

Methods: In the present study, text mining and topic modeling approaches were used to analyze research trends and identify dominant topics. The research steps included data collection from Scopus, text preprocessing, extraction of frequent words, topic modeling using Latent Dirichlet Allocation (LDA), and visualization. All steps were performed using the Python programming language and open-source libraries, such as NLTK, Gensim, Matplotlib, scikit-learn, and pyLDAvis.

Results: A total of 2238 records related to AI in hospitals were collected from Scopus since 2000. The terms "patient," "model," "machine learning," and "artificial intelligence" were identified as the most frequently used terms. The dominant topic clusters included "patient monitoring," "data-driven systems," "service innovation and emerging technologies," "clinical outcome prediction," "COVID-19 risk prediction," "mortality and hospitalization prediction," "health tourism," "management and implementation," and "hospital death prediction." Most articles were in the clusters "clinical outcome prediction modeling" (663 documents) and "mortality and hospitalization prediction" (335 documents). The publication trend has accelerated significantly since 2018, especially in the clusters "clinical outcome prediction" and "management and implementation."

Conclusion: Conclusion: Artificial intelligence in hospitals has grown rapidly over the last two decades. The shift from limited applications in modeling and prediction to interdisciplinary areas and innovative services indicates the gradual growth of this technology and its role in improving the quality of care, optimizing organizational processes, and developing new services.

Keywords: Artificial Intelligence, Hospital, Text Mining, Topic Modeling

1. Background

Artificial intelligence (AI) plays a crucial role in the development and improvement of healthcare systems, particularly in hospitals.¹ The scope of application of AI in the hospital environment is broad and ranges from early diagnosis of diseases and prediction of exacerbation of patients' clinical condition to optimization of hospital operations and improvement of patient-centered care.^{2,3} In clinical settings, AI aids in disease diagnosis, predicts treatment outcomes, and personalizes treatment plans.⁴ In hospital management, AI can significantly transform the provision of healthcare services by leveraging new technologies, including 5G, the Internet of Things (IoT), wearable devices, robots, and smart applications. These systems reduce the workload of healthcare staff and optimize resource management by automating processes, such as patient registration and discharge, sending timely reminders, and environmental cleaning. Moreover, by analyzing patient data and monitoring the status of

epidemics, such as COVID-19, AI plays a pivotal role in allocating resources, tracing the origin of the virus, and monitoring the quality of care.^{5,6}

The development of AI technologies and their diverse applications in the health and hospital sectors has led researchers to conduct extensive research in this area.¹ Research into the applications of AI in healthcare and within hospitals is considered an important area of innovation.⁷

Recent studies indicate a significant growth in the use of AI in hospitals to improve clinical and management processes. Despite the increasing growth of research and case implementations in hospitals, a holistic and comprehensive analytical view of current trends and prevailing issues in this field is essential. Most past studies have either focused on specific applications, such as image recognition, acute care, or management of high-risk patients. In this regard, Veldhuis et al. in a systematic review, examined the performance of AI-based models

for predicting mortality and clinical deterioration in non-critical care patients and reported that most of these models had high predictive accuracy.⁸ Similarly, Romero-Brufau et al. reported a 25% reduction in readmission rates by implementing an AI tool to predict readmission risk in a regional hospital, indicating the positive impact of this technology in the reduction of quality gaps.⁹ Moreover, Klumpp et al. examined diverse applications of AI in European hospitals, including data management, clinical decision support, and improving human-computer interaction, emphasizing the need for specialized solutions to overcome privacy and human acceptance challenges.²

In other areas, Varnosfaderani and Forouzanfar considered the role of AI in medical image analysis, patient monitoring through wearable devices, and hospital operations optimization, noting ethical challenges like data privacy and algorithmic biases.¹⁰ Wan emphasized the importance of interdisciplinary collaboration in developing theoretical and applied AI frameworks and introduced patient-centered approaches as a key trend.¹¹ Moreover, Rashid et al. highlighted the applications of AI in early disease diagnosis, robotic surgery, and personalized medicine in their review, emphasizing the need for investment in research and development to overcome security challenges and algorithmic transparency.³ Collectively, these studies demonstrate the diversity of applications of AI in hospitals and its potential to improve healthcare. Accordingly, given the broad and multidisciplinary nature of AI-related research in hospitals, extracting macro insights from published scientific articles in this field can provide a better understanding of advances, challenges, and research gaps. For instance, Wan demonstrates that partnerships among researchers, healthcare providers, and data scientists are crucial for the development of value-based AI-driven approaches to hospital care.¹¹

2. Objectives

The present study, using text mining and topic modeling methods, aims to draw a comprehensive picture of emerging trends, dominant topics, and upcoming opportunities in this field by analyzing the content of scientific articles related to AI in hospitals.

3. Methods

The present study aimed to analyze research trends and identify dominant topics in the field of AI in hospitals using text mining and topic modeling approaches. The research methodology consisted of five main steps: data collection, text preprocessing, extraction of frequent words, topic modeling using the Latent Dirichlet Allocation (LDA) method, and analysis of similarity between topics. All steps were carried out using the Python programming language and open-source libraries,

including NLTK, Gensim, Matplotlib, scikit-learn, and pyLDAvis.

3.1. Data Collection

The data for this study was retrieved from the Scopus database, one of the most comprehensive sources for indexing scientific articles, on August 1, 2025, based on the following search strategy:

(TITLE ("artificial intelligence" OR AI OR "computer reasoning" OR "computational intelligence" OR "computer vision*" OR "machine learning" OR "Deep learning" OR NLP OR "natural language processing") AND TITLE (hospital*)). The selected time frame was 2000 until the time of data collection. For better analysis of the texts, the Title and Abstract columns were merged and combined into a single text field.

3.2. Text Preprocessing

Text preprocessing is an essential step in natural language analysis.¹² In the present work, the preprocessing included the following steps:

- Converting all letters to lowercase;
- Replacing synonyms, such as "AI" with "artificial intelligence" and "ML" with "machine learning";
- Removing numbers and punctuation marks from the text;
- Tokenization using the NLTK library;
- Removing stopwords by combining the standard NLTK list with common words, useless for them in scientific articles, such as "method," "study," and "result";
- Lemmatization (reduction of words to their common roots) using the WordNet Lemmatizer;
- Detecting and creating bigrams (common two-word phrases) using the Phrases model in the Gensim library;
- Removing very short words (less than four letters) unless they are meaningful (e.g., "AI", "ML").

The preprocessed texts were stored in a separate column and used for subsequent analyses.

3.3. Word Frequency Analysis

The most frequent words were extracted using frequency counting. The top 50 keywords were then plotted graphically using the WordCloud library to provide an overview of the dominant concepts in the data.¹³

3.4. Topic Modeling Using LDA

The LDA algorithm was used to discover hidden semantic structures in a collection of texts.¹⁴ The steps of applying this method are as follows:

- Creating a dictionary and a bag-of-words (BoW) representation for each document;
- Determining the optimal number of topics by testing LDA models in a range of 2-20 topics and calculating the semantic coherence score using the c_v criterion;
- Training the final model with the optimal number of topics;

Figure 1 indicates that the words "patient," "model," "machine learning," and "artificial intelligence" were the most important and frequently used words in the scientific publications related to smart hospitals.

Then, in order to perform topic modeling, the desired number of clusters was obtained using the coherence score. Accordingly, Figure 2 demonstrates the coefficients of the number of topic clusters between 2 and 19 clusters.

The optimal number of topics was determined based on the highest semantic coherence score observed in Figure 2. Specifically, it selected nine topics, as this point represented the beginning of the plateau in the coherence curve, indicating an optimal balance between model complexity and interpretability.

Accordingly, the most important words of the obtained topic clusters are shown in Figure 3.

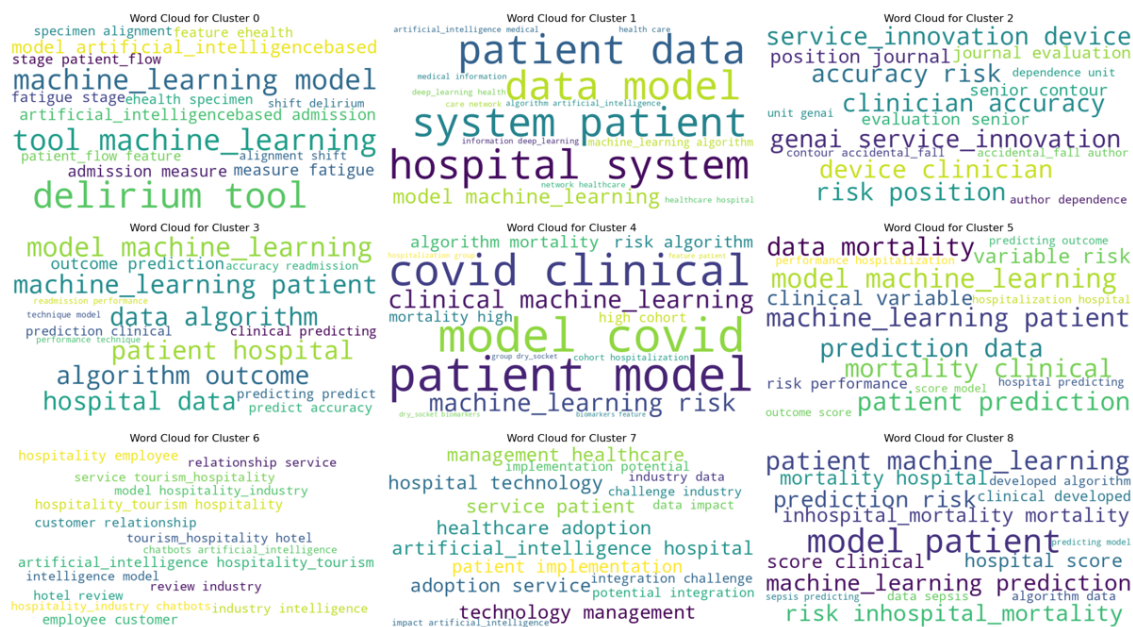


Figure 3. Word Cloud of Topic Clusters of Scientific Publications Related to AI and Hospitals.

According to Figure 3, the extracted topic clusters are named as patient monitoring, data-driven systems, service innovation and emerging technologies, clinical outcome prediction, COVID-19 risk prediction, mortality and hospitalization prediction, health tourism, management

and implementation, and hospital death prediction.

In addition, Table 1 presents the obtained topic clusters along with the most important keywords selected for each cluster and the number of documents in each cluster.

Table 1. Topic Clusters with Associated Keywords and Document Count

Topic	Important Keywords	Label	Num.
Topic 0	Delirium, tool, machine_learning, model, artificial_intelligencebased, admission, measure, fatigue, stage, patient_flow, feature, ehealth, specimen, alignment, shift.	Patient monitoring	25
Topic 1	Hospital, system, patient, data, model, machine_learning, algorithm, artificial_intelligence, medical, information, deep_learning, health, care, network, healthcare.	Data-driven systems	310
Topic 2	GenAI, service_innovation, device, clinician, accuracy, risk, position, journal, evaluation, senior, contour, accidental_fall, author, dependence, unit.	Service innovation and emerging technologies	18
Topic 3	Model, machine_learning, patient, hospital, data, algorithm, outcome, prediction, clinical, predicting, predict, accuracy, readmission, performance, technique.	Clinical outcome prediction	663
Topic 4	Patient, model, COVID, clinical, machine_learning, risk, algorithm, mortality, high, cohort, hospitalization, group, dry_socket, biomarkers, feature.	COVID-19 risk prediction	118
Topic 5	Model, machine_learning, patient, prediction, data, mortality, clinical, variable, risk, performance, hospitalization, hospital, predicting, outcome, score.	Mortality and hospitalization prediction	335
Topic 6	Artificial_intelligence, hospitality_tourism, hospitality, employee, customer, relationship, service, tourism_hospitality, hotel, review, industry, intelligence, model, hospitality_industry, chatbots.	Health tourism	177
Topic 7	Artificial_intelligence, hospital, technology, management, healthcare, adoption, service, patient, implementation, potential, integration, challenge, industry, data, impact.	Management and implementation	310
Topic 8	Model, patient, machine_learning, prediction, risk, in-hospital_mortality, mortality, hospital, score, clinical, developed, algorithm, data, sepsis, predicting.	Hospital death prediction	282

Table 1 indicates that the topics clinical outcome prediction modeling (cluster 3) with 663 documents and the topic mortality and hospitalization prediction (cluster 5) with 335 documents include the largest number of articles, reflecting the focus of research on developing prediction models for clinical outcomes, such as mortality, readmission, and treatment performance.

Moreover, the presence of topics such as COVID-19 (cluster 4), technology implementation and management in hospitals (cluster 7), and even health tourism (cluster 6) demonstrates that research in this area is diverse and has also expanded into interdisciplinary areas.

Figure 4 indicates the degree of similarity between the topic clusters.

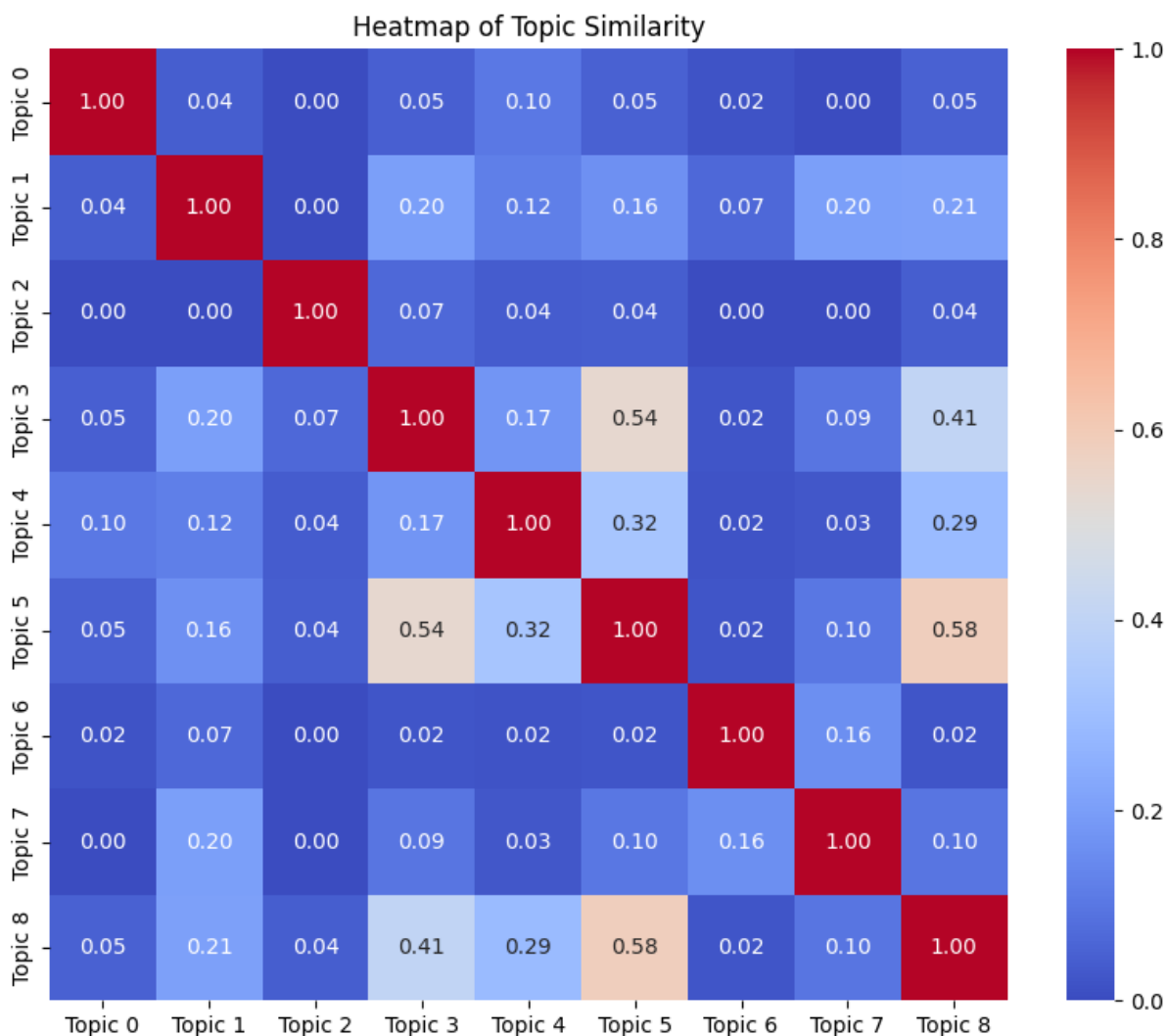


Figure 4. Similarity between Topic Clusters Obtained from Topic Modeling of Scientific Publications on AI and Hospitals.

Figure 4 indicates that the highest semantic similarity is observed between Cluster 5 (prediction of mortality and hospitalization) and Cluster 8 (prediction of death in hospital), with a value of 0.58. Another relatively strong relationship is observed between Cluster 3 (prediction of clinical outcome) and Cluster 5 (prediction of mortality and hospitalization), with a value of 0.54. There is also a moderate semantic relationship between Cluster 3 (prediction of clinical outcome) and Cluster 8 (prediction of death in hospital), as well as between Cluster 4 (prediction of COVID-19 risk) and Cluster 5 (prediction of mortality and hospitalization), with a value of 0.32. Other clusters, such as Cluster 0 (patient monitoring), Cluster 1 (data-driven systems), Cluster 2 (service

innovation and emerging technologies), and Cluster 6 (health tourism), showed minimal semantic relationships with other clusters.

Figure 5 indicates the publication trend of topics from 2000 onwards.

Figure 4 illustrates that the publication trend of topics from 2000 onwards has generally accelerated since 2018 and peaked from 2021 to 2024. In this regard, Cluster 3 (prediction of clinical outcomes) has experienced the most growth, with an upward trend since 2018, and reached its peak in 2024, comprising 156 documents. Cluster 7 (management and implementation) has experienced significant growth since 2020, reaching its highest levels in 2024 and 2025, with 119 and 104 documents, respectively.

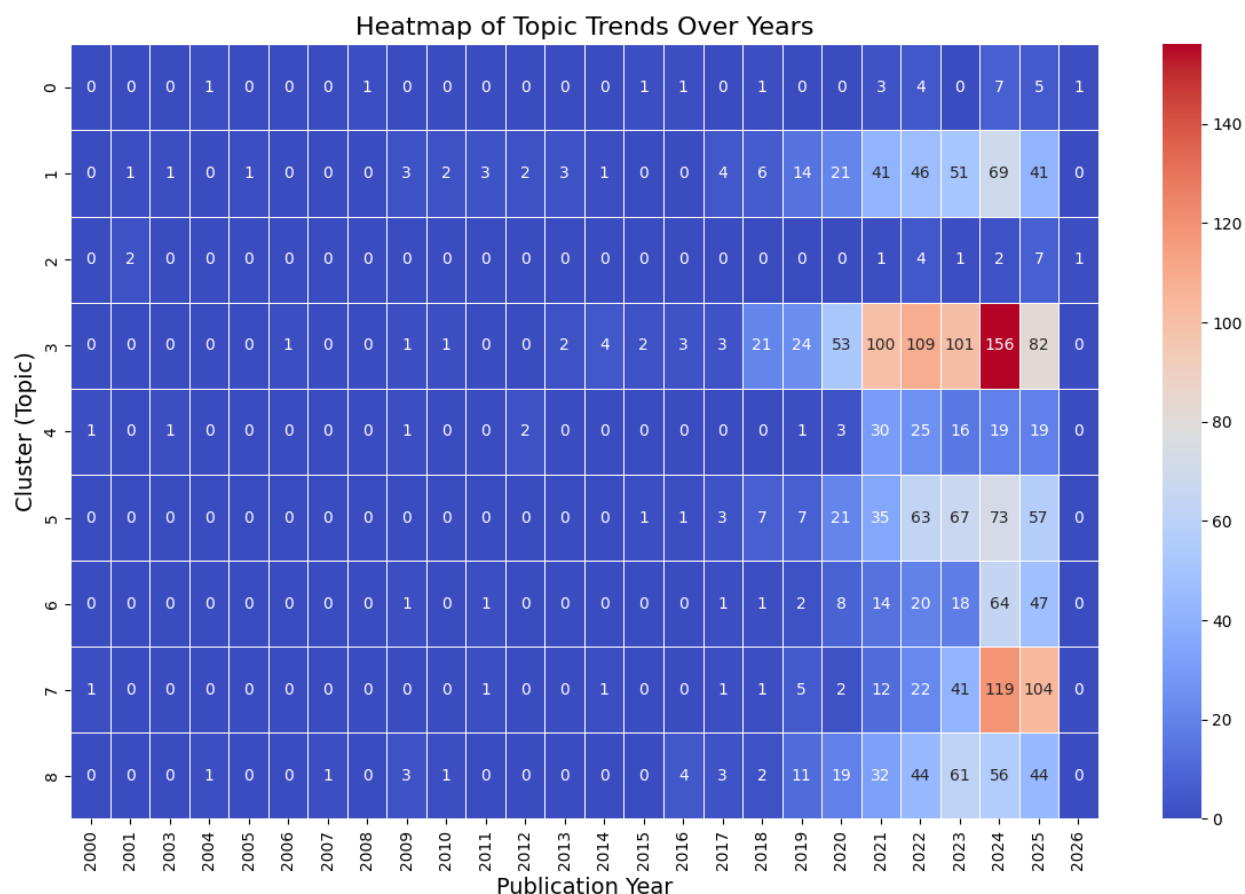


Figure 5. Scientific Publication Trend in the Field of AI and Hospitals.

Cluster 5 (prediction of mortality and hospitalization) has also shown an increasing trend since 2018 and reached its highest level (63 to 73 documents) between 2022 and 2024. Cluster 8 (prediction of in-hospital death) has demonstrated moderate growth since 2019 and peaked in 2023 with 61 documents. Cluster 1 (data-driven systems) has experienced significant growth since 2019, reaching 69 documents in 2024, highlighting the importance of data infrastructure in the digital transformation of hospitals. Cluster 6 (health tourism) has grown since 2018 and peaked at 64 documents in 2024. Cluster 4 (COVID-19 risk prediction) had the highest number of publications in 2021, with 30 documents. Cluster 0 (patient monitoring) and Cluster 2 (service innovation and emerging technologies) have much lower publication volumes than other clusters and do not show significant ups and downs.

5. Discussion

The results of the present study indicate that AI in hospitals has undergone a rapid evolutionary path over the past two decades, expanding from its initial limited areas of modeling and prediction to encompass a range of interdisciplinary topics. According to the results of the present study, the significant increase in scientific production since 2018, particularly in the areas of clinical

outcome prediction and management, as well as implementation, indicates the impact of technical advances in deep learning, big data processing, and the expansion of hospital digital infrastructure. Additionally, although the findings reveal the presence of nine distinct thematic clusters, the primary focus of the research has been on predicting clinical outcomes, mortality, and readmission. This issue indicates that the development of machine learning (ML)-based predictive models remains the core of innovations in hospitals. Veldhuis et al. also noted that the most common clinical applications of AI were in the early identification of patient deterioration and the prediction of treatment outcomes. Their study also showed the high accuracy of AI models in clinical prediction and mortality,⁸ which is consistent with the results of the present study. In addition, Kashani and Dastahi showed the growth trend of research in the field of AI in medical sciences since 2000. Moreover, the terms "machine learning" and "deep learning" have been proposed as key concepts in this field,¹ which is also consistent with the results of the present work.

One of the highlights of the current study is the inclusion of emerging topics, such as COVID-19 risk prediction and health tourism, on the topic map. Although the topic of COVID-19 peaked in 2021, it has had a smaller share of scholarly output since the pandemic

subsided, a pattern also observed in bibliometric studies.¹⁶ In contrast, health tourism, as an interdisciplinary field that links AI technologies with cross-border healthcare services,¹⁷ has seen steady growth, reflecting a gradual shift in research towards commercialization models and innovative services.

In addition, based on the results of the present study, there is a substantial similarity between the clusters of clinical outcome prediction, mortality and hospitalization prediction, and in-hospital death prediction. This topic convergence reflects methodological commonalities and the use of similar data (e.g., electronic health records and real-time clinical data). Previous studies have also demonstrated that AI and ML models have performed significantly in predicting clinical outcomes, such as in-hospital mortality, length of stay, and disease recurrence.¹⁸⁻²⁰

From a time perspective, the growth in publications since 2018, particularly in the clinical outcomes prediction cluster, reflects the impact of technological developments and the increasing availability of big data in hospitals. This trend is consistent with the findings of Bates et al. (2020), who suggest that one of the reasons for the rapid growth of AI applications in healthcare is related to algorithmic advances and data infrastructures.²¹ However, the significant growth of the management and implementation cluster since 2020 is likely due to a greater focus on the operational and organizational challenges associated with integrating AI technologies in hospitals. The topic of "health tourism," as an interdisciplinary area, reflects the expansion of AI applications beyond clinical care. However, the relatively low volume of publications in this cluster compared to the clinical clusters indicates the need for more research in this area. In this regard, evidence suggests that despite the relatively small share of this sector in the tourism industry, it has a high capacity to generate income and improve health services. The use of AI can contribute to the sustainable development of this industry by improving access, analyzing patient data, and optimizing the management of medical and residential centers.²²

6. Conclusion

The results of the present study indicate that AI in hospitals has evolved from limited, model-based applications toward diverse interdisciplinary domains. The findings derived from topic modeling and word frequency analyses clearly demonstrate that the dominant research focus has been on clinical outcome prediction, mortality and hospitalization prediction, and management and implementation. This confirms that predictive modeling remains the core area of innovation in hospital-based AI research.

Furthermore, the emergence of topics such as COVID-19 risk prediction, health tourism, and data-

driven systems reflects the growing integration of AI into both clinical and managerial contexts, extending beyond traditional medical applications. The temporal trend analysis also revealed a substantial growth in publications since 2018, particularly in clusters related to outcome prediction and hospital management, which aligns with global advances in deep learning, big data analytics, and hospital digital transformation.

These findings collectively suggest that AI has become not only a technological innovation but also a strategic enabler in transforming healthcare delivery, improving operational efficiency, and fostering innovation in hospital environments. The evidence obtained from this study supports the notion that the future of AI in hospitals will increasingly focus on implementation challenges, ethical considerations, and interdisciplinary collaboration to achieve sustainable and human-centered innovation in healthcare.

Research Highlights

What Is Already Known?

- AI is widely used in hospitals for tasks such as diagnosis, patient monitoring, and operational optimization.
- Bibliometric analyses of AI in healthcare exist, but few focus specifically on thematic evolution within hospital contexts.

What Does This Study Add?

- Text mining of 2,238 scientific articles reveals a sharp increase in hospital-based AI research since 2018.
- Nine core topic clusters were identified, with "Clinical Outcome Prediction" and "Mortality & Hospitalization Prediction" as the most prevalent themes.

Author Contributions

Authors contributed equally to this work.

Conflict of Interest Disclosures

All authors declared that they have no conflict of interest.

Declaration of Generative AI and AI-assisted Technologies

The authors used ChatGPT (GPT-4.0) to improve English language clarity and grammar in some sections.

References

1. Kashani M, Dastani M. Evolution of artificial intelligence in medical sciences: a comprehensive scientometrics analysis. *Glob Knowl Mem Commu*. 2025. doi:10.1108/GKMC-09-2024-0586
2. Klumpp M, Hintze M, Immonen M, Rydenas-Rigla F, Pilati F, Aparicio-Martínez F, et al. Artificial Intelligence for Hospital Health Care: Application Cases and Answers to Challenges in European Hospitals. *Healthcare*. 2021;9(8):961. doi:10.3390/healthcare9080961
3. Rashid Z, Ahmed H, Nadeem N, Zafar SB, Yousaf MZ. The paradigm of digital health: AI applications and transformative trends. *Neural Comput Appl*. 2025; 37(17):11039-70. doi:10.1007/s00521-025-11081-0
4. DuBois KN. Deep medicine: how artificial

- intelligence can make healthcare human again. *Perspect Sci Christ Faith*. 2019;71(3):199-201.
5. Mi D, Li Y, Zhang K, Huang C, Shan W, Zhang J. Exploring intelligent hospital management mode based on artificial intelligence. *Front Public Health*. 2023;11:1182329. doi:10.3389/fpubh.2023.1182329
 6. Gourisaria MK, Agrawal R, Singh V, Rautaray SS, Pandey M. AI and IoT enabled smart hospital management systems. In: *Data Science in Societal Applications: Concepts and Implications*. Cham: Springer; 2022. pp. 77-106. doi:10.1007/978-981-19-5154-1_6
 7. Halawa F, Madathil SC, Gittler A, Khasawneh MT. Advancing evidence-based healthcare facility design: a systematic literature review. *Health Care Manag Sci*. 2020;23(3):453-80. doi:10.1007/s10729-020-09506-4
 8. Veldhuis LI, Woittiez NJ, Nanayakkara PW, Ludikhuizen J. Artificial intelligence for the prediction of in-hospital clinical deterioration: a systematic review. *Crit Care Explor*. 2022;4(9):e0744. doi:10.1097/CCE.0000000000000744
 9. Romero-Brufau S, Wyatt KD, Boyum P, Mickelson M, Moore M, Cognetta-Rieke C. Implementation of artificial intelligence-based clinical decision support to reduce hospital readmissions at a regional hospital. *Appl Clin Inform*. 2020;11(4):570-7. doi:10.1055/s-0040-1715827
 10. Maleki Varnosfaderani S, Forouzanfar M. The role of AI in hospitals and clinics: transforming healthcare in the 21st century. *Bioengineering*. 2024;11(4):337. doi:10.3390/bioengineering11040337
 11. Wan TT. Convergence of artificial intelligence research in healthcare: trends and approaches. *J Integr Des Process Sci*. 2023;26(3-4):291-305. doi:10.3233/JID-200002
 12. Hickman L, Thapa S, Tay L, Cao M, Srinivasan P. Text Preprocessing for Text Mining in Organizational Research: Review and Recommendations. *Organ Res Methods*. 2022;25(1):114-46. doi:10.1177/1094428120971683
 13. Heimerl F, Lohmann S, Lange S, Ertl T. Word cloud explorer: Text analytics based on word clouds. In: 2014 47th Hawaii International Conference on System Sciences. IEEE; 2014. pp. 1833-42. doi:10.1109/HICSS.2014.231
 14. Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. *J Mach Learn Res*. 2003;3:993-1022.
 15. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. *J Mach Learn Res*. 2011;12:2825-30.
 16. Danesh F, Ghavidel S. Coronavirus: Scientometrics of 50 Years of Global Scientific Productions. *Iran J Med Microbiol*. 2020;14(1):1-16. doi:10.30699/ijmm.14.1.1
 17. Monroy-Rodriguez S, Caro-Carretero R. Mapping the landscape of health tourism: a bibliometric review and its alignment with the SDGs. *Int J Spa Wellness*. 2025;8(1):76-105. doi:10.1080/24721735.2024.2444014
 18. Hilton CB, Milinovich A, Felix C, Vakharia N, Crone T, Donovan C, et al. Personalized predictions of patient outcomes during and after hospitalization using artificial intelligence. *NPJ Digit Med*. 2020;3(1):51. doi:10.1038/s41746-020-0249-z
 19. Pettit RW, Fullem R, Cheng C, Amos CI. Artificial intelligence, machine learning, and deep learning for clinical outcome prediction. *Emerg Top Life Sci*. 2021;5(6):729-45. doi:10.1042/ETLS20210246
 20. Huang C-C, Kuo W-Y, Shen Y-T, Chen C-J, Lin H-J, Hsu C-C, et al. Artificial intelligence prediction of In-Hospital mortality in patients with dementia: A multi-center study. *Int J Med Inform*. 2024;191:105590. doi:10.1016/j.ijmedinf.2024.105590
 21. Bates DW, Auerbach A, Schulam P, Wright A, Saria S. Reporting and implementing interventions involving machine learning and artificial intelligence. *Ann Intern Med*. 2020;172(11 Suppl):S137-44. doi:10.7326/M19-0872
 22. Lukovic V. Data, Artificial Intelligence, and Health Tourism. In: *Innovation and Creativity in Tourism, Business and Social Sciences*. Cham: Springer Nature Switzerland; 2025. doi:10.1007/978-3-031-78471-2_36