Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Meta-Analysis Studies on Artificial Intelligence in Healthcare: A Bibliometric Analysis (2010-2025)

**Furkan Alp**

furkan.alp@medeniyet.edu.tr

Istanbul Medeniyet University

**Merve İnan Budak**

Istanbul Medeniyet University

**Research Article**

**Keywords:** Artificial Intelligence, Healthcare, Meta-Analysis, Bibliometric Analysis, VOSviewer, Bibliometrix

**Posted Date:** November 9th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-11640404/v1

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

**Additional Declarations:** No competing interests reported.

# META-ANALYSIS STUDIES ON ARTIFICIAL INTELLIGENCE IN HEALTHCARE: A BIBLIOMETRIC ANALYSIS (2010-2025)

**Furkan Alp[1]*, Merve İnan Budak[2]**

[1] Department of Health Management, Faculty of Health Sciences, Istanbul Medeniyet University, Istanbul, Türkiye, 0000-0001-6007-1212.

[2] Department of Psychiatric Nursing, Faculty of Health Sciences, Istanbul Medeniyet University, Istanbul, Türkiye, 0000-0002-4334-6779.

*Corresponding Author: Furkan Alp

Email: furkan.alp@medeniyet.edu.tr

Contributing Authors' Emails:

merve.inan@medeniyet.edu.tr

## ABSTRACT

**Aim:** Meta-analyses of artificial intelligence (AI) in healthcare have proliferated across many clinical domains, yet no study has mapped this literature as a body of scholarship in its own right. This study aims to conduct a bibliometric analysis of meta-analysis studies on AI in healthcare published between 2010 and 2025.

**Methods:** Following the PRISMA 2020 guidelines, 646 research and review articles indexed in the Web of Science Core Collection and Scopus databases (2010–2025) were analyzed for citation patterns, annual trends, and leading countries, institutions, authors, publishers, and journals. Co-authorship and keyword co-occurrence networks were constructed using VOSviewer (v.1.6.19.0) and Bibliometrix.

**Results:** Publication output remained minimal before 2019 but rose sharply thereafter, reaching 153 publications in 2024 and 280 in 2025, together accounting for about two-thirds of the dataset. The most cited article was Christodoulou et al. (2019, 1,663 citations), followed by Aggarwal et al. (2021, 848). China led in publication output (221 articles, 34.21%), while the United Kingdom led in citation impact (6,269 citations, 30.25%) despite ranking third in output. Tehran University of Medical Sciences was the most productive institution, Habibi, M.A. the most productive author, and the Journal of Medical Internet Research the leading journal. Author-level productivity was concentrated within a small set of interconnected teams, institutional collaboration was sparse, and country-level collaboration was broad and international. Keyword analysis identified artificial intelligence, machine learning, meta-

analysis, and deep learning as core terms, organized around three further axes: imaging-based diagnostic accuracy, outcome and risk prediction, and disease-specific evidence synthesis.

**Conclusion:** This study provides the first structural bibliometric mapping of the meta-analytic literature on AI in healthcare, revealing rapid recent growth alongside concentrated authorship and institutional collaboration. These findings suggest that publication growth has outpaced methodological consolidation and offer a reference point for future work in the field.

**Keywords:** Artificial Intelligence; Healthcare; Meta-Analysis; Bibliometric Analysis; VOSviewer; Bibliometrix.

## 1 INTRODUCTION

Artificial intelligence (AI) refers to computational systems that perform tasks once thought to require human cognition, such as recognizing patterns, making predictions, and supporting decisions (Yu et al., 2018). The idea of bringing such systems into medicine is not new. Rule-based expert systems for clinical diagnosis and decision support were already being proposed in the 1970s, but the computing power and data available at the time kept their use limited (Jiang et al., 2017). That changed over the past decade. Machine learning and deep learning methods matured just as health records and diagnostic imaging were becoming digitized on a large scale, giving these methods the volume of data they needed (Jiang et al., 2017). AI moved, in a relatively short period, from an experimental concept to a set of tools now evaluated across most areas of clinical practice, including diagnosis, prognosis, treatment planning, and health system operations (Al Kuwaiti et al., 2023). That shift explains much of the field's current scholarly attention: AI is now treated as a technology capable of changing how healthcare is delivered and managed, not as a marginal add-on (Bajwa et al., 2021). Reviews of the field are nonetheless consistent in noting real obstacles to adoption, among them data quality and algorithmic transparency (Aung et al., 2021).

This literature spans a wide range of clinical units rather than concentrating on a single specialty. In oncology, AI systems have been evaluated for ovarian cancer identification (Xu et al., 2022), lymphoma detection (Bai et al., 2024), and lung cancer screening (Thong et al., 2023); a related line of dental research has focused on caries detection (Luke & Rezallah, 2025). Cardiology researchers have examined AI models for detecting cardiovascular disease from wearable-device data (S. Lee et al., 2022) and for predicting cardiovascular outcomes more broadly (Baashar et al., 2022; Riipa et al., 2026). Time-critical settings have drawn comparable interest: AI-supported diagnostics have been tested in emergency and critical care (Sreedharan

et al., 2024), intensive care mortality prediction (Barboi et al., 2022), and emergency department disposition decisions (K.-M. Kuo & Chang, 2025). Gastroenterology offers another example, where computer-aided systems support colorectal polyp detection during colonoscopy (Nazarian et al., 2021). In surgery, models have been trained to analyze intraoperative video (Rafati Fard et al., 2026), and in psychiatry the focus has shifted toward diagnostic and therapeutic applications alike (Rony et al., 2025). Socially assistive robots have been evaluated for their effect on cognitive function in older adults (H. Lee et al., 2022), and a more recent line of work has examined AI-enhanced interventions for cancer patients in supportive care (Y. Li et al., 2026). This range extends beyond individual clinical units: AI has also been measured against health-related examinations and certification tests (Waldock et al., 2024), linked to broader indicators of healthcare quality (Alzghoul, 2024), and compared directly against healthcare professionals on diagnostic tasks (Kumar et al., 2026).

Alongside this growth in application-specific research, the AI-in-healthcare literature has itself come under closer methodological scrutiny. Broad systematic reviews have catalogued the range of AI tools reported across medicine and healthcare, along with their limitations and practical challenges (Younis et al., 2024). Automated methods have been developed to help screen this rapidly growing body of work (Feng et al., 2022), and reporting guidelines specific to AI-based systematic reviews and meta-analyses have been proposed to address recurring gaps in transparency and reproducibility (Kolbinger et al., 2024). Taken together, these developments suggest that the growth of this literature has been accompanied by increasing attention to how it should be reported and evaluated.

Despite this proliferation of unit-specific meta-analyses, comparatively little attention has been paid to the meta-analytic literature on AI in healthcare as a body of scholarship in its own right. Individual meta-analyses provide pooled estimates of AI performance within a single clinical unit, but they do not, by design, describe how this literature has grown over time, which institutions and countries have driven its production, or how its thematic focus has shifted as the underlying technology has matured. A bibliometric approach addresses this gap by treating the meta-analyses themselves as the unit of analysis. This makes it possible to map publication trends, collaboration networks, and keyword co-occurrence patterns across the field as a whole. The present study therefore applies bibliometric analysis to meta-analysis studies on artificial intelligence in healthcare published between 2010 and 2025, with the aim of characterizing the structure and evolution of this literature rather than its clinical conclusions.

## 2 METHODOLOGY

This study was designed to examine the academic literature on meta-analyses of artificial intelligence in healthcare through a bibliometric analysis approach. Reporting followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Page et al., 2021), which provide a standardized framework encompassing the identification, screening, eligibility assessment, and inclusion of studies. The study sample comprised research and review articles published in English between 2010 and 2025, retrieved from the Web of Science (WoS) and Scopus databases. A total of 4322 records were identified across the two databases (WoS = 1083, Scopus = 3239). Following removal of 857 duplicate records, 3465 records remained and were screened based on title and abstract. At this stage, 2794 records were excluded, comprising editorial articles, conference abstracts, letters, and similar publication types, as well as records not meeting the artificial intelligence, healthcare, or meta-analysis inclusion criteria, leaving 671 full-text articles assessed for eligibility. During eligibility assessment, a further 25 full-text articles were excluded as not relevant to the scope of the study. Consequently, 646 articles were included in the bibliometric analysis (Figure 1).

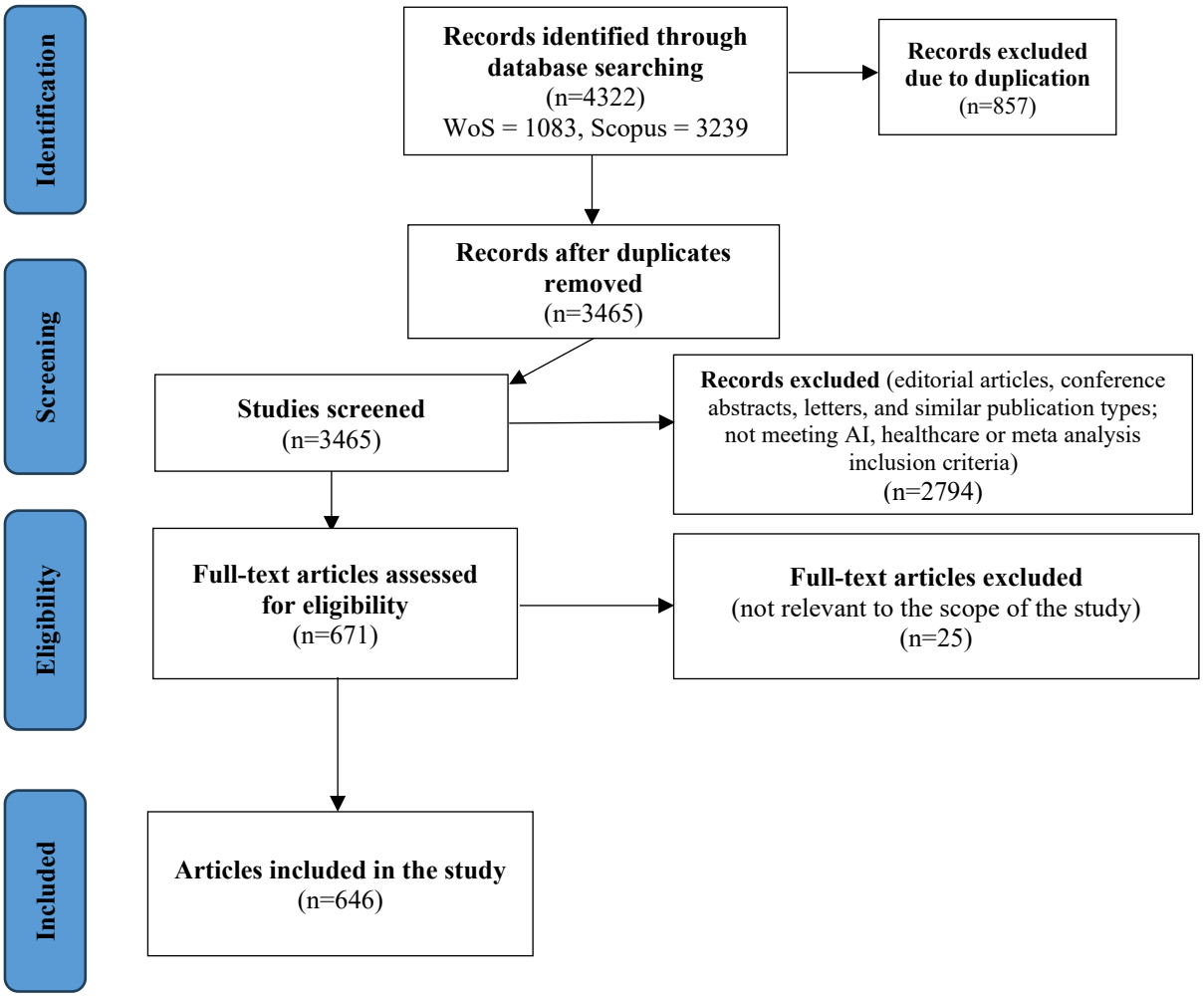


**FIGURE 1** PRISMA flow diagram

Data were retrieved from the Web of Science Core Collection and Scopus databases. The framework of the data collection process is presented in Table 1.

**TABLE 1** Search framework

| Parameters | Specification |
|---|---|
| Selection approach | Bibliometric analysis |
| Databases used | Web of Science and Scopus |
| Tools used for analysis | VOSviewer and Bibliometrix |
| Search query | ("meta-analysis" OR "meta-analyses" OR "meta analysis" OR "meta-analytic" OR "network meta-analysis" OR "pooled analysis" OR "meta-regression")<br>AND<br>("artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network*" OR "natural language processing" OR "large language model*" OR "LLM*" OR "generative AI" OR "computer vision" OR "predictive model*" OR "AI algorithm*" OR "AI-based" OR "chatbot*" OR "ChatGPT")<br>AND<br>("healthcare" OR "health care" OR "health service*" OR "clinical practice" OR "medicine" OR "medical" OR "patient care" OR "health system*" OR "hospital*" OR "clinical decision*") |
| Document type | Research and review articles |
| Time span | 2010-2025 |
| Language | English |
| Subject area | All fields |
| Publication stage | Final |
| Total number of publications for analysis | 646 |

The analysis was conducted on research and review articles published in English between 2010 and 2025 in the Web of Science and Scopus databases, using a search query in which meta-analysis-related, artificial intelligence-related, and healthcare-related terms co-occurred, yielding a final total of 646 publications.

The statistical analysis in this study was performed using VOSviewer (version 1.6.19.0) and the Bibliometrix software package. The analysis examined a range of bibliometric parameters, including the most highly cited studies, the yearly distribution of publications, the countries with the highest publication and citation counts, the most productive publishers and journals. Visual mapping techniques were also used to examine the structure of the literature, including the co-authorship network at the author, institutional, and country level, and the keyword co-occurrence network.

## 3 FINDINGS

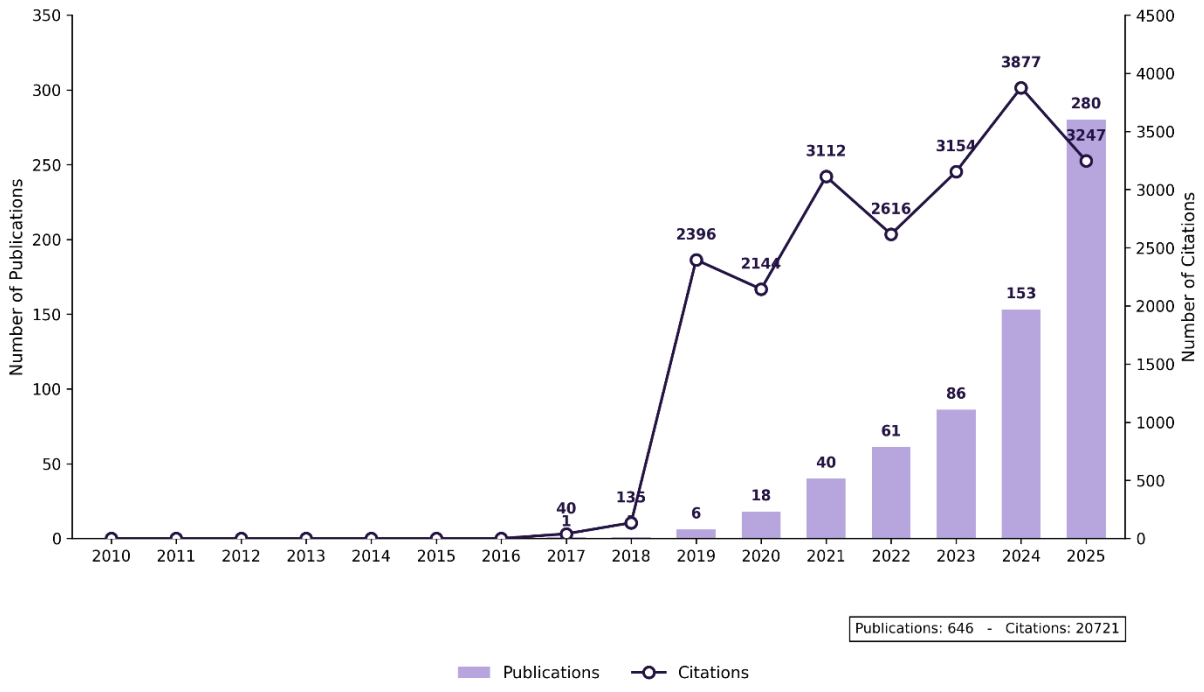



**FIGURE 2** Number of publications and citations by year

Figure 2 presents the annual distribution of publications and citations from 2010 to 2025. No publications meeting the inclusion criteria were retrieved for 2010-2016, and only one article appeared in each of 2017 and 2018. Output then rose from 6 publications in 2019 to 86 by 2023, before increasing sharply to 153 in 2024 and 280 in 2025, the two highest annual counts in the dataset. These two years alone account for approximately 67% of all 646 included studies. This

distribution is consistent with a rapidly emerging research area, although the concentration of studies in the last two years also means that citation counts for this period are still accumulating. Citation counts followed a less consistent trajectory across the same years: after an initial rise to 2396 in 2019, citations fell to 2144 in 2020, increased to 3112 in 2021, declined to 2616 in 2022, and then rose to 3154 in 2023 and 3877 in 2024, the highest annual total recorded. The lower figure of 3247 citations in 2025 most likely reflects the limited time available for recently published articles to accumulate citations, rather than an actual decrease in the field's scholarly impact.

**TABLE 2** Information on the top 10 most-cited articles

| Article Title | Authors | Year | Journal | Citation Count |
|---|---|---|---|---|
| A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models | Christodoulou, E., Ma, J., Collins, G. S., ... & Van Calster, B. | 2019 | Journal of Clinical Epidemiology | 1663 |
| Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis | Aggarwal, R., Sounderajah, V., Martin, G., ... & Darzi, A. | 2021 | npj Digital Medicine | 848 |
| Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy | Fleuren, L. M., Klausch, T. L. T., Zwager, C. L., ... & Elbers, P. W. G. | 2020 | Intensive Care Medicine | 627 |
| Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being | Li, H., Zhang, R., Lee, Y.-C., Kraut, R. E., & Mohr, D. C. | 2023 | npj Digital Medicine | 481 |
| Effectiveness and safety of using chatbots to improve mental health: Systematic review and meta-analysis | Abd-Alrazaq, A. A., Rababeh, A., Alajlani, M., Bewick, B. M., & Househ, M. | 2020 | Journal of Medical Internet Research | 381 |
| Artificial Intelligence in Fracture Detection: A Systematic Review and Meta-Analysis | Kuo, R. Y. L., Harrison, C., Curran, T.-A., ... & Furniss, D. | 2022 | Radiology | 304 |

| | | | | |
|---|---|---|---|---|
| Radiomics in nuclear medicine: robustness, reproducibility, standardization, and how to avoid data analysis traps and replication crisis | Zwanenburg, A. | 2019 | European Journal of Nuclear Medicine and Molecular Imaging | 266 |
| Artificial intelligence in digital pathology: a systematic review and meta-analysis of diagnostic test accuracy | McGenity, C., Clarke, E. L., Jennings, C., ... & Treanor, D. | 2024 | npj Digital Medicine | 228 |
| Performance of ChatGPT Across Different Versions in Medical Licensing Examinations Worldwide: Systematic Review and Meta-Analysis | Liu, M., Okuhara, T., Chang, X., ... & Kiuchi, T. | 2024 | Journal of Medical Internet Research | 186 |
| Prediction of sepsis patients using machine learning approach: A meta-analysis | Islam, M. M., Nasrin, T., Walther, B. A., ... & Li, Y. C. | 2019 | Computer Methods and Programs in Biomedicine | 176 |

Table 2 shows that the most highly cited studies were concentrated in a small number of digital health journals: npj Digital Medicine accounts for three of the ten entries (Aggarwal et al., 2021; H. Li et al., 2023; McGenity et al., 2024), and the Journal of Medical Internet Research for two (Abd-Alrazaq et al., 2020; Liu et al., 2024). Christodoulou et al. (2019), a systematic review comparing machine learning with logistic regression for clinical prediction models, ranks first with 1663 citations, nearly double the second-ranked study. This high count is notable because the study is methodological rather than applied, in contrast to most of the other entries, which report diagnostic or prognostic performance. Aggarwal et al. (2021), a meta-analysis of deep learning diagnostic accuracy in medical imaging, ranks second with 848 citations.

The ten studies span a range of clinical and methodological topics, including sepsis prediction (Fleuren et al., 2020; Islam et al., 2019), fracture detection (R. Y. L. Kuo et al., 2022), digital pathology (McGenity et al., 2024), radiomics standardization (Zwanenburg, 2019), chatbot-based mental health interventions (Abd-Alrazaq et al., 2020; H. Li et al., 2023), and a recent benchmark of large language models against human performance (Liu et al., 2024). Among the top-cited studies in this dataset, attention is therefore not limited to a single clinical domain but

spans diagnostic imaging, acute-care risk prediction, digital pathology, and generative AI applications.

All ten entries are meta-analyses or systematic reviews, consistent with the inclusion criteria applied in this study. Citation counts should be interpreted with some caution, since they are shaped by time since publication and journal visibility as well as by scholarly impact. With that caveat, two of the more recently published studies, McGenity et al. (2024) and Liu et al. (2024), had already reached 228 and 186 citations within about two years of publication, suggesting that recent meta-analytic work in this area can accumulate citations quickly. Overall, the ten most-cited studies span publication years 2019 to 2024 and cover a widening set of clinical applications, from established diagnostic-imaging work to more recent studies evaluating large language models.

**TABLE 3** Most productive authors

| Rank | Author | Number of Publications | Percentage (%) |
|---|---|---|---|
| 1 | Habibi, M.A. | 15 | 2.32% |
| 2 | Mohammadzadeh, I. | 12 | 1.86% |
| 3 | Hajikarimloo, B. | 11 | 1.70% |
| 4 | Eini, P. | 8 | 1.24% |
| 5 | Wang, T.-W. | 8 | 1.24% |
| 6 | Wu, Y.-T. | 8 | 1.24% |
| 7 | Niroomand, B. | 7 | 1.08% |
| 8 | Xue, P. | 6 | 0.93% |
| 9 | Hong, J.-S. | 5 | 0.77% |
| 10 | Mohammadzadeh, S. | 5 | 0.77% |

Examination of Table 3 shows that Habibi, M.A. is the most productive author in the field with 15 publications, followed by Mohammadzadeh, I. with 12 publications. Five of the ten most productive authors, Habibi, M.A.; Mohammadzadeh, I.; Hajikarimloo, B.; Eini, P.; and Niroomand, B., co-author a recurring set of systematic reviews and meta-analyses on machine learning-based outcome prediction in neurosurgery, which points to a strong concentration of the field's most productive authorship within a single research team. A second pair, Wang, T.-W. and Wu, Y.-T., co-authors of a series of deep learning segmentation reviews, shows a similar pattern of team-based productivity. Overall, the ten most productive authors account for

approximately 13% of the total publication output, suggesting that productivity in this area is concentrated among a relatively small number of researchers and research teams.

**TABLE 4** Most productive institutions

| Rank | Institution | Number of Publications | Percentage (%) |
|---|---|---|---|
| 1 | Tehran University of Medical Sciences | 35 | 5.42% |
| 2 | Shahid Beheshti University of Medical Sciences | 23 | 3.56% |
| 3 | Taipei Medical University | 10 | 1.55% |
| 4 | Imperial College London | 9 | 1.39% |
| 5 | Mayo Clinic | 9 | 1.39% |
| 6 | Isfahan University of Medical Sciences | 9 | 1.39% |
| 7 | Sichuan University | 8 | 1.24% |
| 8 | Tabriz University of Medical Sciences | 8 | 1.24% |
| 9 | Stanford University | 8 | 1.24% |
| 10 | University of Virginia | 8 | 1.24% |

According to Table 4, Tehran University of Medical Sciences emerges as the most productive institution with 35 publications, followed by Shahid Beheshti University of Medical Sciences with 23 publications and Taipei Medical University with 10 publications. Four of the ten institutions, Tehran University of Medical Sciences, Shahid Beheshti University of Medical Sciences, Isfahan University of Medical Sciences, and Tabriz University of Medical Sciences, are Iranian medical universities, consistent with the concentration of productive authorship already observed in Table 3. The remaining institutions span Europe, North America, and East Asia, including Imperial College London, Mayo Clinic, Stanford University, the University of Virginia, Taipei Medical University, and Sichuan University, indicating that research on artificial intelligence meta-analyses in healthcare draws on academic and clinical centers across multiple regions rather than a single country.

**TABLE 5** Most productive countries

| Rank | Country | Number of Publications | Percentage (%) |
|---|---|---|---|
| 1 | China | 221 | 34.21% |
| 2 | United States | 146 | 22.60% |
| 3 | United Kingdom | 79 | 12.23% |
| 4 | Iran | 70 | 10.84% |
| 5 | Germany | 40 | 6.19% |
| 6 | Australia | 38 | 5.88% |
| 7 | Italy | 34 | 5.26% |
| 8 | Canada | 34 | 5.26% |
| 9 | India | 30 | 4.64% |
| 10 | Taiwan | 29 | 4.49% |

Note: Because a single study may involve authors from more than one country, the percentages sum to more than 100%.

Examination of Table 5 shows that China holds a clear lead in the field with 221 publications, corresponding to 34.21% of the total, followed by the United States with 146 publications (22.60%) and the United Kingdom with 79 publications (12.23%). Iran ranks fourth with 70 publications (10.84%), consistent with the concentration of authorship observed in Table 3 and Table 4. The top five countries together account for more than 85% of total publication output, indicating that research on this topic is concentrated among a small number of countries with established artificial intelligence and biomedical research infrastructure.

**TABLE 6** Most cited countries

| Rank | Country | Number of Citations | Percentage (%) |
|---|---|---|---|
| 1 | United Kingdom | 6269 | 30.25% |
| 2 | United States | 4330 | 20.90% |
| 3 | China | 4230 | 20.41% |
| 4 | Netherlands | 2783 | 13.43% |
| 5 | Belgium | 2310 | 11.15% |
| 6 | Singapore | 1843 | 8.89% |
| 7 | Germany | 1660 | 8.01% |
| 8 | Iran | 1223 | 5.90% |
| 9 | Australia | 1219 | 5.88% |
| 10 | Italy | 1132 | 5.46% |

Note: Because a single study may involve authors from more than one country, the percentages sum to more than 100%.

According to Table 6, the United Kingdom accounts for the largest share of total citations (6269 citations, 30.25%), even though it ranks only third in publication output, indicating that a comparatively small number of UK-affiliated studies have had a disproportionately large influence on the field. The United States and China follow with 20.90% and 20.41% of total citations respectively, broadly consistent with their ranking in publication output. A notable finding is that the Netherlands, Belgium, and Singapore do not appear among the top 10 countries in terms of publication output but rank fourth, fifth, and sixth in citations, suggesting that studies affiliated with these countries, despite their relatively small number, have exerted a disproportionately high influence on the field.

**TABLE 7** List of publishers with the highest number of publications

| Rank | Publisher | Number of Publications | Percentage (%) |
|---|---|---|---|
| 1 | Springer Nature | 167 | 25.85% |
| 2 | Elsevier | 160 | 24.77% |
| 3 | Frontiers Media | 54 | 8.36% |
| 4 | JMIR Publications | 49 | 7.59% |
| 5 | MDPI | 41 | 6.35% |
| 6 | Wiley | 31 | 4.80% |
| 7 | Wolters Kluwer | 29 | 4.49% |
| 8 | Oxford University Press | 18 | 2.79% |
| 9 | PLOS | 12 | 1.86% |
| 10 | SAGE Publications | 11 | 1.70% |

Examination of Table 7 shows that Springer Nature is the leading publisher with 167 publications (25.85% of the total, including its BMC and Nature-branded journals), followed closely by Elsevier with 160 publications (24.77%, combining its various corporate imprints). Together, these two publishers account for more than half of all included studies, indicating a marked concentration of the field's literature within two major commercial publishing groups. The presence of open-access-oriented publishers such as MDPI, Frontiers Media, and JMIR Publications among the top ranks shows that research on artificial intelligence meta-analyses in healthcare is also widely disseminated through the open-access publishing model.

**TABLE 8** List of journals with the highest number of publications

| Rank | Journal | Number of Publications | Percentage (%) |
|---|---|---|---|
| 1 | Journal of Medical Internet Research | 34 | 5.26% |
| 2 | Frontiers in Oncology | 18 | 2.79% |
| 3 | npj Digital Medicine | 14 | 2.17% |
| 4 | BMC Medical Informatics and Decision Making | 11 | 1.70% |
| 5 | Journal of Clinical Medicine | 10 | 1.55% |
| 6 | Academic Radiology | 9 | 1.39% |
| 7 | European Journal of Radiology | 9 | 1.39% |
| 8 | International Journal of Medical Informatics | 9 | 1.39% |
| 9 | Computers in Biology and Medicine | 8 | 1.24% |
| 10 | Cancers | 7 | 1.08% |

Examination of Table 8 shows that the Journal of Medical Internet Research leads with 34 publications, followed by Frontiers in Oncology with 18 and npj Digital Medicine with 14. The prominence of the Journal of Medical Internet Research and npj Digital Medicine, both established outlets for digital health and AI-based research, shows that this literature is anchored in digital-health and general AI-in-medicine journals rather than in a single disease-specific specialty journal, since meta-analytic evidence on artificial intelligence in healthcare spans many clinical fields rather than one. Several of the remaining journals, including Academic Radiology, the European Journal of Radiology, and Computers in Biology and Medicine, focus on medical imaging and biomedical informatics, which is consistent with diagnostic imaging being one of the most frequently studied applications of AI in this literature. The presence of broad-scope journals such as the Journal of Clinical Medicine and Cancers among the top ranks indicates that these studies are also published in general clinical and oncology journals, not only in specialty AI or informatics outlets.

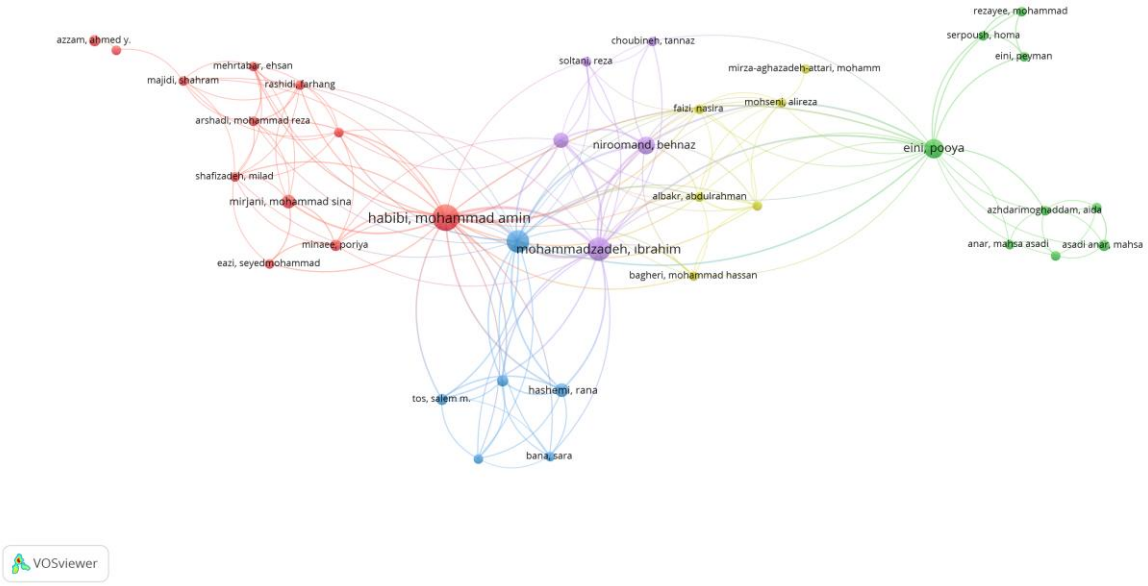



**FIGURE 3** Co-authorship network analysis

Figure 3 presents the co-authorship network for the 38 authors who met the co-authorship threshold, grouped into five clusters with a total link strength of 305. The largest and most densely connected cluster is centered on Habibi, M.A., who co-authors with a group of researchers including Azzam, A.Y.; Mehrtabar, E.; Majidi, S.; and Rashidi, F. The second-largest cluster is centered on Mohammadzadeh, I., linked to a smaller set of co-authors including Hashemi, R. and Najari, D. A visible connection links the Habibi and Mohammadzadeh clusters, suggesting some collaboration across the two groups, though this link appears weaker than the connections within each cluster. Two further clusters are centered on Niroomand, B. and Eini, P., each with their own set of co-authors, and a fifth, smaller cluster remains only loosely connected to the rest of the network.

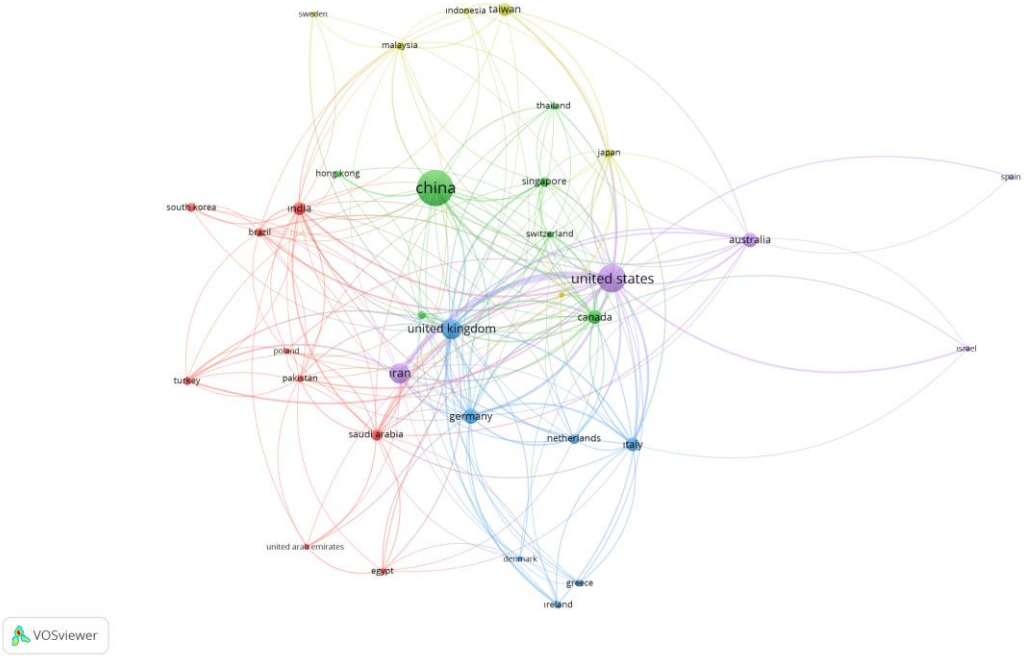



**FIGURE 4** Country-level co-authorship network analysis

Figure 4 presents the country-level co-publication network. According to Figure 4, 34 countries are grouped into 5 clusters, forming a densely interconnected network with 216 links (total link strength = 544). China occupies the largest and most central node, followed by the United States in second place, even though the United Kingdom recorded a higher total citation count than either, as shown in Table 6. The United Kingdom, together with the Netherlands, Italy, and Austria, forms a European-centered cluster. China clusters with Canada, Singapore, Japan, and several other East and Southeast Asian countries, forming the largest and most densely connected group in the network. The United States, Iran, and Australia form a separate cluster, while Turkey, Pakistan, Saudi Arabia, and several other countries form a further, more peripheral cluster. Nearly all countries in this network are connected to one another, indicating that a broad international collaboration structure exists at the country level even where individual research teams may work independently.

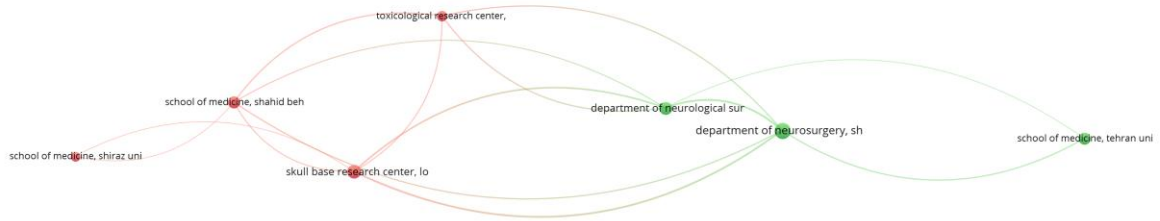



**FIGURE 5** Institutional co-authorship network analysis

Figure 5 presents the institutional co-authorship network. The network shows only 7 institutions divided into 2 clusters, with 14 links and a total link strength of 43, a considerably sparser structure than the author- and country-level networks. One cluster is centered on the Department of Neurosurgery at Shariati Hospital, Tehran University of Medical Sciences, connected to the Department of Neurological Surgery at the University of Virginia and the School of Medicine at Tehran University of Medical Sciences. The second cluster groups several units within Shahid Beheshti University of Medical Sciences, including its Toxicological Research Center and School of Medicine, together with the Skull Base Research Center at Loghman-Hakim Hospital and the School of Medicine at Shiraz University. Several other institutions with at least four publications, including Imperial College London and the Chinese Academy of Medical Sciences, showed no co-authorship link to any other institution in the selected set and are therefore absent from the visible network, which reinforces how limited cross-institutional collaboration is in this field.

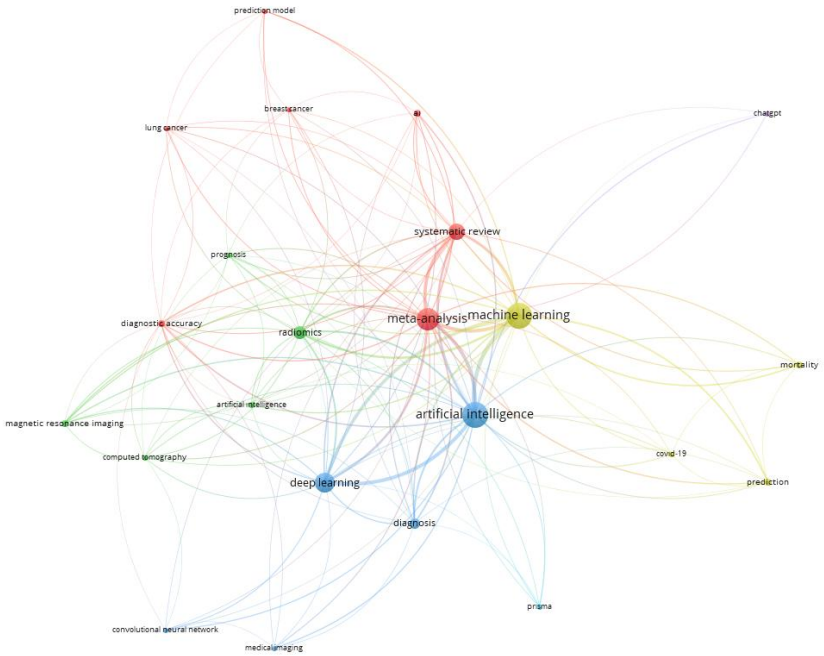



**FIGURE 6** Keyword network analysis

Figure 6 presents the keyword co-occurrence network, showing 23 keywords grouped into six clusters. Beyond the expected core terms of artificial intelligence, machine learning, meta-analysis, and deep learning, the network separates into three further axes. One cluster brings together radiomics, magnetic resonance imaging, computed tomography, diagnostic accuracy, and prognosis, reflecting an imaging-based diagnostic-accuracy axis. A second cluster groups mortality, covid-19, and prediction, an axis concerned with outcome and risk prediction. A third cluster combines systematic review, breast cancer, and lung cancer, pointing to disease-specific evidence synthesis. Two further terms, prisma and chatgpt, each form their own small, isolated cluster, with the latter marking the recent emergence of large language models as a distinct topic in this literature.

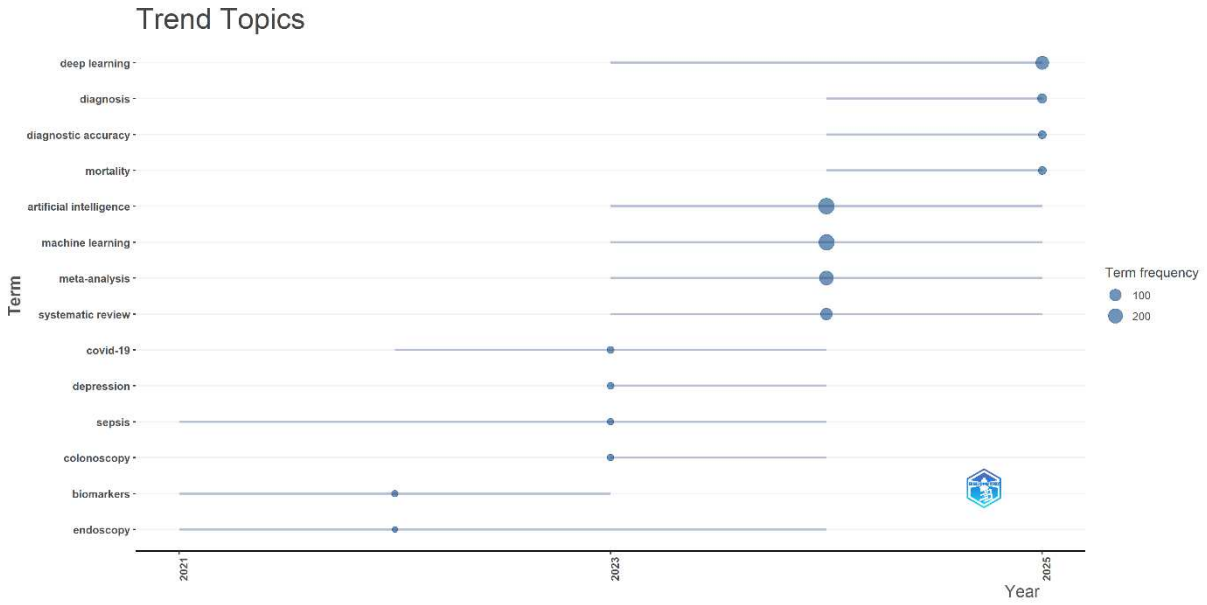



**FIGURE 7** Trend topics

Figure 7 presents the trend topics identified across the study period based on author keywords. The terms artificial intelligence, machine learning, meta-analysis, and systematic review form the largest and most persistent cluster, centered around 2023, which is expected given that these terms define the core scope of the dataset itself. Beyond this central cluster, the trend map reveals a shift in research focus over time. Early terms such as biomarkers and endoscopy, associated with 2021, point to initial applications of artificial intelligence in specific diagnostic procedures. By 2022, the topic range broadens to include covid-19, depression, sepsis, and colonoscopy, reflecting a diversification of clinical domains in which artificial intelligence-related meta-analyses were being conducted. The most recent terms, including deep learning, diagnosis, diagnostic accuracy, and mortality, emerge in 2024 and 2025 and suggest a shift toward more specific technical methods and clinical outcome measures, moving beyond broad application areas toward questions of diagnostic performance and patient outcomes.

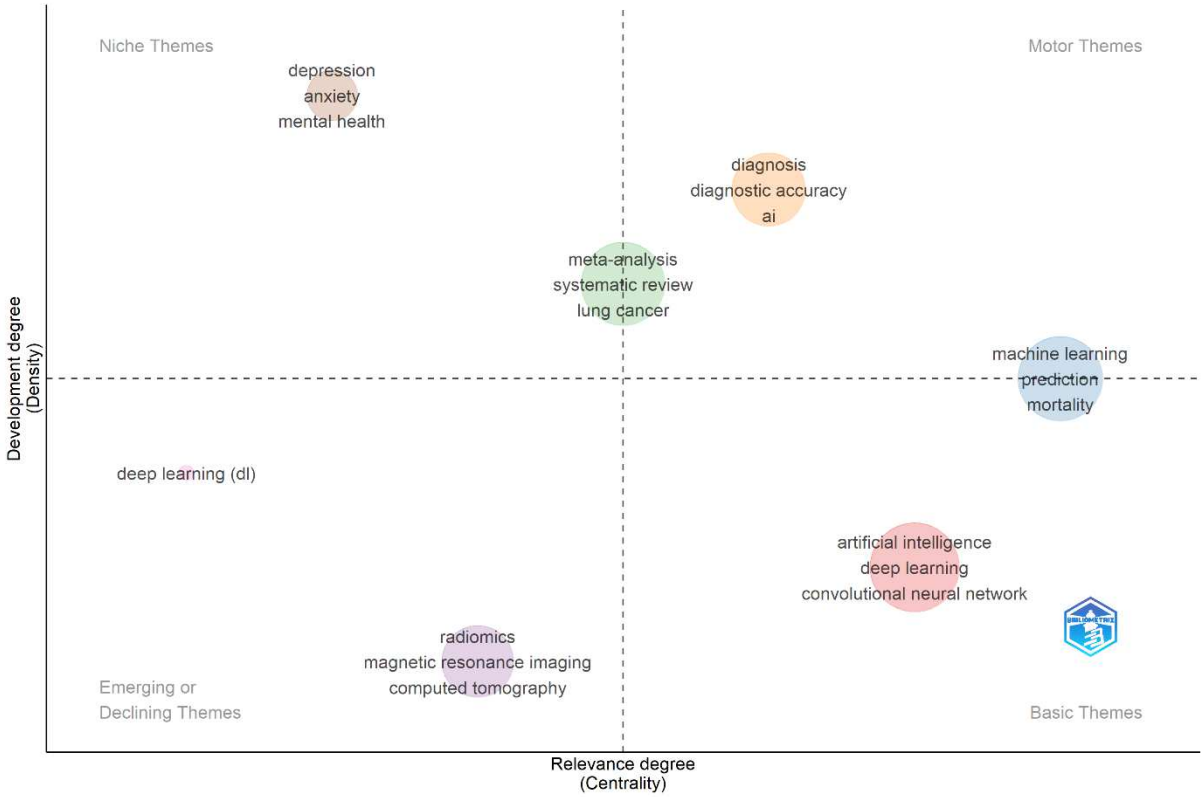



**FIGURE 8** Thematic map

Figure 8 presents the thematic map generated from the author keywords, positioning topic clusters according to their centrality and density. The cluster containing artificial intelligence, deep learning, and convolutional neural network, together with the cluster containing machine learning, prediction, and mortality, occupy the basic themes quadrant, reflecting their role as foundational and cross-cutting terms that underlie the majority of the dataset. The cluster of diagnosis, diagnostic accuracy, and ai falls within the motor themes quadrant, indicating a well-developed and central line of research on diagnostic performance. A distinct niche theme is formed by depression, anxiety, and mental health, positioned as a highly developed but relatively isolated cluster, pointing to a self-contained body of work on artificial intelligence applications in mental health. The cluster of radiomics, magnetic resonance imaging, and computed tomography appears in the emerging or declining themes quadrant, suggesting that imaging-based applications remain a peripheral and still-developing branch relative to the central themes of the field. The cluster of meta-analysis and systematic review, which also includes lung cancer, sits close to the intersection of the basic and motor themes quadrants, consistent with its role as a defining methodological anchor of the dataset.

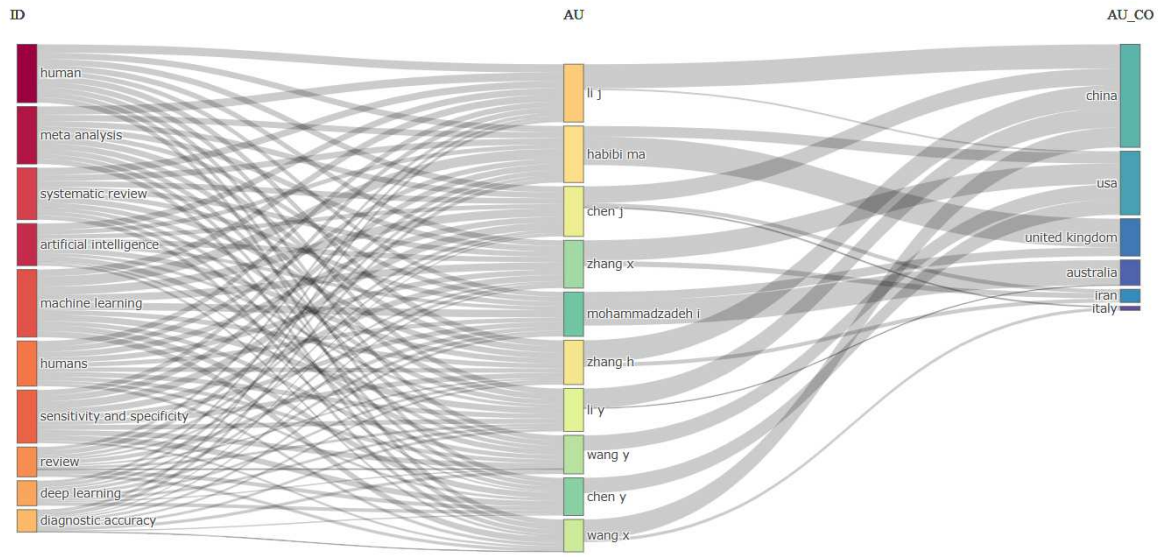



**FIGURE 9** Three-field plot

Figure 9 presents a three-field plot linking the most frequent author keywords, the most productive authors, and their countries of affiliation. The left side of the diagram is dominated by core methodological terms such as meta-analysis, systematic review, artificial intelligence, and machine learning, consistent with the defining scope of the dataset, while terms such as diagnostic accuracy and deep learning connect to a narrower subset of author nodes, reflecting more specific technical contributions. On the author level, Habibi M.A. and Mohammadzadeh I. appear among the central nodes, consistent with their standing as the most productive individual authors identified in Table 3. On the right side, China, the United States, and the United Kingdom emerge as the dominant countries linked to these authors and keywords, while Iran, Australia, and Italy form smaller but visible connections, reflecting the geographically concentrated distribution of contributing research groups already described in Table 5.

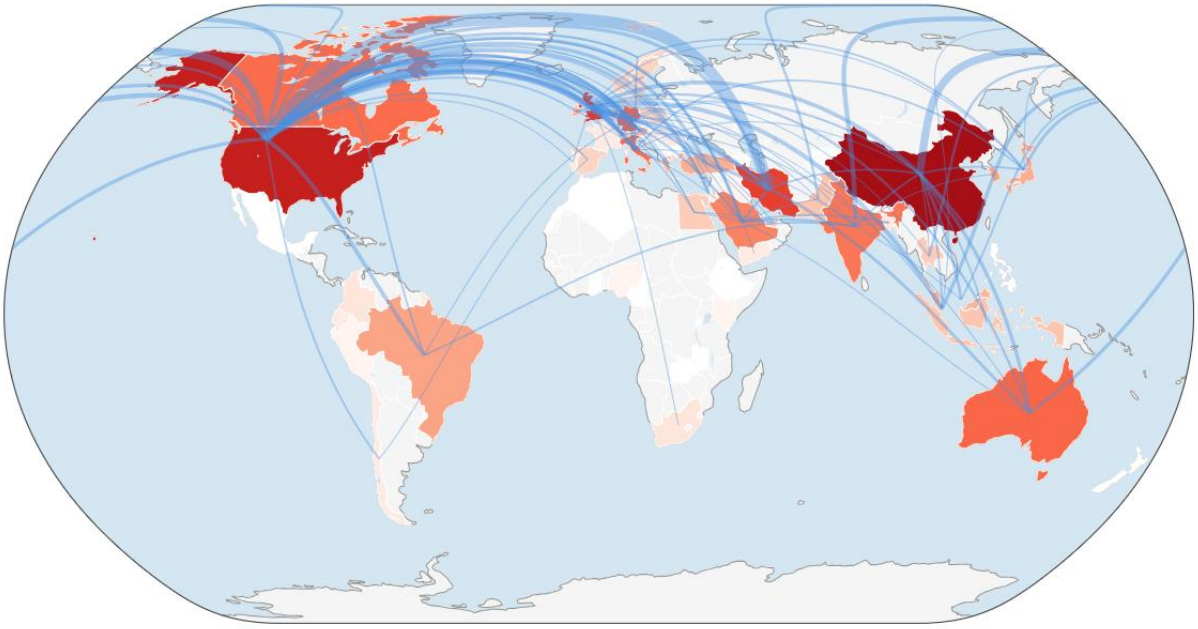


**FIGURE 10** Countries' collaboration world map

Figure 10 presents the countries' collaboration world map, illustrating the geographic distribution and intensity of international co-authorship across the dataset. The United States and China appear as the most intensely colored countries, consistent with their standing as the two leading contributors identified in Table 5, and both serve as central hubs connecting to a wide range of collaborating countries across North America, Europe, and Asia. The United Kingdom, Iran, Australia, and India also show a substantial number of collaborative ties, reflecting the geographically dispersed nature of the research groups contributing to this literature. The collaboration lines further reveal that intercontinental partnerships, particularly between North America, Europe, and East Asia, form the backbone of international cooperation in this field, while collaboration within Africa and South America remains comparatively limited.

## 4 DISCUSSION

The results point to a field that has grown very quickly. Nearly two-thirds of the 646 included studies were published in just the last two years, a pattern that mirrors the broader adoption curve described in the introduction, where machine learning and deep learning methods only became widely applicable to healthcare once large volumes of digitized clinical data were available. This surge in publication volume, however, does not appear to have been matched by an equivalent gain in methodological maturity. Ong et al. (2026), reviewing 31 randomized trials of AI-based cardiovascular decision support, found that the overall certainty of evidence was low to very low despite the scale of the underlying trial base. Daemi et al. (2025) reported a similarly cautionary picture across a broader set of diagnostic applications, where pooled accuracy was high but heterogeneity and publication bias were both substantial. The concern is not limited to newer or less-studied applications: Vasilev et al. (2024), reviewing 38 radiology-focused AI meta-analyses, found that methodological gaps remained common even within one of this dataset's most frequently studied clinical units. Read together, these three studies suggest that the field's rapid expansion in publication count has outpaced its methodological consolidation.

A related tension appears at the level of authorship and institutions. The most productive individual authors in this dataset are concentrated in a small number of research teams working on machine learning-based outcome prediction in neurosurgery, and the institutional co-authorship network is correspondingly sparse, with only seven institutions forming visible links. This is considerably narrower than the structure reported for the broader AI-in-clinical-medicine literature by Li et al. (2026), whose bibliometric analysis of 2,872 articles published between 2010 and 2025 found a well-developed collaboration network spanning 1,000 institutions, including universities, hospitals, and research institutes. The contrast suggests that meta-analytic synthesis work, the specific genre examined here, remains concentrated in a small set of teams even as the primary AI-in-medicine literature it draws on has grown far more collaborative. A related pattern appears at the country level. Li et al. (2026) found the United States leading in publication output, with Asia as a continent ahead in total citations, a mismatch between production and influence that parallels the present finding that the United Kingdom leads in citations despite ranking only third in publication output. Taken together, the two studies suggest that citation impact in this literature is not simply a function of how many papers a country produces.

The thematic evolution observed here, moving from broad early applications toward more specific diagnostic and outcome-related terms by 2024 and 2025, is echoed in the substantive literature, though not without exceptions worth noting. Karcıoğlu (2025) is a useful contrast case: working in health management rather than clinical medicine, it found that AI reduced medical errors and wait times across a large multi-institution sample, a perspective on the operational side of healthcare that is not clearly represented among this dataset's keyword clusters. Whether this reflects a genuine gap in the underlying literature or simply the diagnostic-accuracy orientation of the search terms used in the present study cannot be settled from bibliometric data alone. A second, smaller gap concerns workforce readiness. Amiri et al. (2024) found that only 44% of medical, dental, and nursing students demonstrated adequate AI knowledge despite largely positive attitudes toward it, and Kahraman et al. (2026) reported a moderate-to-high level of AI-related anxiety among healthcare professionals. Neither concern appears as a distinct cluster among the 23 keywords surfaced by the present analysis, which instead isolates mental health as its only self-contained niche theme, suggesting that the psychosocial dimensions of AI adoption remain comparatively underexamined relative to its technical performance. Furthermore, neglecting psychosocial and ethical dimensions is especially evident in the integration of AI within psychiatric units. While AI holds transformative potential for assessment, prediction, and optimization, the humanistic nature of psychiatric care presents distinct challenges for algorithmic deployment (Milasan & Scott-Purdy, 2025). Amplified by data bias and persistent ethical concerns (H. Li et al., 2024), these limitations demonstrate that technical success cannot substitute for therapeutic presence, clinical judgment, and empathy. Thus, interdisciplinary collaboration among clinicians, service users, and developers is essential to build person-centered tools that preserve the humanistic core of psychiatry alongside practical utility (H. Li et al., 2024; Milasan & Scott-Purdy, 2025).

Finally, Wenderott et al. (2024) offers a useful caution for interpreting this literature as a whole. Although most individual studies in their review reported that AI improved workflow efficiency in medical imaging, pooling these studies at the meta-analytic level erased the effect entirely. Zheng et al. (2021) reached a related conclusion for tumor-detection accuracy, where pooled performance approached that of clinicians but the authors still called for standardization and external validation before the results could be considered clinically reliable. Both studies illustrate why the growth documented in this bibliometric analysis should be read as growth in research activity rather than as settled evidence of clinical benefit.

## 5 CONCLUSION

This study mapped the meta-analytic literature on artificial intelligence in healthcare through bibliometric methods, not through an evaluation of AI's clinical performance itself. The field is young. It is also expanding fast: most of the 646 included studies were published in just the last two years. Even so, authorship and institutional collaboration remain concentrated in a small number of teams, while country-level networks are broad and international. The topics covered have shifted over time too, moving from a handful of early diagnostic applications toward a wider and more technically specific set of outcomes. A few areas, health management and workforce readiness among them, remain only lightly represented relative to diagnostic-accuracy research.

That so many meta-analyses have accumulated on this topic within a single decade is itself a sign of how central artificial intelligence has become to healthcare research, whatever the strength of the underlying evidence. These patterns carry a practical message. A rapidly growing number of meta-analyses does not, on its own, mean that the underlying evidence on AI in healthcare is settling into consensus. Several of the specific meta-analyses discussed here found pooled results still limited by heterogeneity, small samples, or a lack of external validation, even in well-studied units such as imaging and diagnostics. Bibliometric growth and evidentiary maturity are not the same thing, and the two should not be conflated when this literature is used to guide clinical or policy decisions.

Future work could extend this analysis in two directions. One is temporal: repeating this mapping in a few years, once the 2024-2025 publication surge has had time to accumulate citations, would show whether the field's current momentum holds. The other is structural: a closer comparison between meta-analyses in this space and the broader primary literature on AI in clinical medicine could clarify whether the concentrated authorship observed here is a temporary feature of a young sub-field or a more lasting one.

### Study Limitations

Restricting the search to Web of Science and Scopus may have excluded meta-analyses indexed only in PubMed or regional databases, potentially underrepresenting research from lower- and middle-income countries. Citation-based indicators favor older publications, likely underestimating the influence of the 2024-2025 articles that make up two-thirds of this dataset. The keyword-based search may not fully capture terminology that emerged late in the study period, such as foundation models, which were not included as separate search terms. Finally,

as a bibliometric rather than a systematic review, this study does not assess the methodological quality of the individual meta-analyses it maps.

**Funding Statement**

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**Conflict of Interest Statement**

The authors declare that there is no conflict of interest.

**Author Contributions**

These authors contributed equally to this work.

**Data Availability Statement**

All data analyzed in this study are publicly available and can be accessed through the Web of Science and Scopus databases.

**Ethics Statement**

As this is a bibliometric analysis using publicly available data, ethical review and approval were not applicable.

**REFERENCES**

Abd-Alrazaq, A. A., Rababeh, A., Alajlani, M., Bewick, B. M., & Househ, M. (2020). Effectiveness and safety of using chatbots to improve mental health: Systematic review and meta-analysis. *Journal of Medical Internet Research*, *22*(7), e16021. https://doi.org/10.2196/16021

Aggarwal, R., Sounderajah, V., Martin, G., Ting, D. S. W., Karthikesalingam, A., King, D., Ashrafian, H., & Darzi, A. (2021). Diagnostic accuracy of deep learning in medical imaging: A systematic review and meta-analysis. *Npj Digital Medicine*, *4*(1), 65. https://doi.org/10.1038/s41746-021-00438-z

Al Kuwaiti, A., Nazer, K., Al-Reedy, A., Al-Shehri, S., Al-Muhanna, A., Subbarayalu, A. V., & Al-Muhanna, F. A. (2023). A review of the role of artificial intelligence in healthcare. *Journal of Personalized Medicine*, *13*(6), 951. https://doi.org/10.3390/jpm13060951

Alzghoul, B. (2024). Impact of artificial intelligence on healthcare quality: A systematic review and meta-analysis. *The Open Public Health Journal*, *17*(1). https://doi.org/10.2174/0118749445181059240201054546

Amiri, H., Peiravi, S., Rezazadeh Shojaee, S. S., Rouhparvarzamin, M., Nateghi, M. N., Etemadi, M. H., & Asadi Anar, M. (2024). Medical, dental, and nursing students' attitudes and knowledge towards artificial intelligence: A systematic review and meta-analysis. *BMC Medical Education*, *24*(1), 412. https://doi.org/10.1186/s12909-024-05406-1

Aung, Y. Y., Wong, D. C. S., & Ting, D. S. W. (2021). The promise of artificial intelligence: A review of the opportunities and challenges of artificial intelligence in healthcare. *British Medical Bulletin*, *139*(1), 4–15. https://doi.org/10.1093/bmb/ldab016

Baashar, Y., Alkawsi, G., Alhussian, H., Capretz, L. F., Alwadain, A., Alkahtani, A. A., & Almomani, M. (2022). Effectiveness of artificial intelligence models for cardiovascular disease prediction: Network meta-analysis. *Computational Intelligence and Neuroscience*, *2022*(1), 5849995. https://doi.org/10.1155/2022/5849995

Bai, A., Si, M., Xue, P., Qu, Y., & Jiang, Y. (2024). Artificial intelligence performance in detecting lymphoma from medical imaging: A systematic review and meta-analysis. *BMC Medical Informatics and Decision Making*, *24*(1), 13. https://doi.org/10.1186/s12911-023-02397-9

Bajwa, J., Munir, U., Nori, A., & Williams, B. (2021). Artificial intelligence in healthcare: Transforming the practice of medicine. *Future Healthcare Journal*, *8*(2), e188–e194. https://doi.org/10.7861/fhj.2021-0095

Barboi, C., Tzavelis, A., & Muhammad, L. N. (2022). Comparison of severity of illness scores and artificial intelligence models that are predictive of intensive care unit mortality: Meta-analysis and review of the literature. *JMIR Medical Informatics*, *10*(5), e35293. https://doi.org/10.2196/35293

Christodoulou, E., Ma, J., Collins, G. S., Steyerberg, E. W., Verbakel, J. Y., & Van Calster, B. (2019). A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *Journal of Clinical Epidemiology*, *110*, 12–22. https://doi.org/10.1016/j.jclinepi.2019.02.004

Daemi, A., Kalami, S., Tahiraga, R. G., Ghanbarpour, O., Barghani, M. R. R., Hooshiar, M. H., & Yönden, Z. (2025). Revolutionizing personalized medicine using artificial intelligence: A meta-analysis of predictive diagnostics and their impacts on drug development. *Clinical and Experimental Medicine*, *25*(1), 255. https://doi.org/10.1007/s10238-025-01723-x

Feng, Y., Liang, S., Zhang, Y., Chen, S., Wang, Q., Huang, T., & Pan, H. (2022). Automated medical literature screening using artificial intelligence: A systematic review and meta-analysis. *Journal of the American Medical Informatics Association*, *29*(8), 1425–1432. https://doi.org/10.1093/jamia/ocac066

Fleuren, L. M., Klausch, T. L. T., Zwager, C. L., Schoonmade, L. J., Guo, T., Roggeveen, L. F., Swart, E. L., Girbes, A. R. J., Thoral, P., Ercole, A., Hoogendoorn, M., & Elbers, P. W. G. (2020). Machine learning for the prediction of sepsis: A systematic review and

meta-analysis of diagnostic test accuracy. *Intensive Care Medicine*, *46*(3), 383–400. https://doi.org/10.1007/s00134-019-05872-y

Islam, M. M., Nasrin, T., Walther, B. A., Wu, C.-C., Yang, H.-C., & Li, Y.-C. (2019). Prediction of sepsis patients using machine learning approach: A meta-analysis. *Computer Methods and Programs in Biomedicine*, *170*, 1–9. https://doi.org/10.1016/j.cmpb.2018.12.027

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*, *2*(4), 230–243. https://doi.org/10.1136/svn-2017-000101

Kahraman, H., Yüceler Kaçmaz, H., & Akutay, S. (2026). Artificial intelligence anxiety levels among healthcare professionals: A systematic review and meta-analysis. *Journal of Nursing Management*, *2026*(1), 5410088. https://doi.org/10.1155/jonm/5410088

Karcıoğlu, U. B. (2025). The role of artificial intelligence in healthcare management on patient safety and operational efficiency: A meta-analytic evaluation. *Journal of Healthcare Management and Leadership*, (1), 8–21. https://doi.org/10.35345/johmal.1670414

Kolbinger, F. R., Veldhuizen, G. P., Zhu, J., Truhn, D., & Kather, J. N. (2024). Reporting guidelines in medical artificial intelligence: A systematic review and meta-analysis. *Communications Medicine*, *4*(1), 71. https://doi.org/10.1038/s43856-024-00492-0

Kumar, P., Alnaimi, N. A., Soman, S., Suansing, L., Ryan Arriola, D., & Jamea, L. A. (2026). Meta-analysis on comparison of diagnostic accuracy between artificial intelligence and healthcare professionals. *Sci*, *8*(4), 73. https://doi.org/10.3390/sci8040073

Kuo, K.-M., & Chang, C.-S. (2025). A meta-analysis of the diagnostic test accuracy of artificial intelligence predicting emergency department dispositions. *BMC Medical Informatics and Decision Making*, *25*(1), 187. https://doi.org/10.1186/s12911-025-03010-x

Kuo, R. Y. L., Harrison, C., Curran, T.-A., Jones, B., Freethy, A., Cussons, D., Stewart, M., Collins, G. S., & Furniss, D. (2022). Artificial intelligence in fracture detection: A systematic review and meta-analysis. *Radiology*, *304*(1), 50–62. https://doi.org/10.1148/radiol.211785

Lee, H., Chung, M. A., Kim, H., & Nam, E. W. (2022). The effect of cognitive function health care using artificial intelligence robots for older adults: Systematic review and meta-analysis. *JMIR Aging*, *5*(2), e38896. https://doi.org/10.2196/38896

Lee, S., Chu, Y., Ryu, J., Park, Y. J., Yang, S., & Koh, S. B. (2022). Artificial intelligence for detection of cardiovascular-related diseases from wearable devices: A systematic review and meta-analysis. *Yonsei Medical Journal*, *63*(Suppl), S93. https://doi.org/10.3349/ymj.2022.63.S93

Li, H., Zhang, R., Lee, Y.-C., Kraut, R. E., & Mohr, D. C. (2023). Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. *Npj Digital Medicine*, *6*(1), 236. https://doi.org/10.1038/s41746-023-00979-5

Li, H., Zhu, G., Zhong, Y., Zhang, Z., Li, S., & Liu, J. (2024). Applications of artificial intelligence in psychiatric nursing: A scope review. *Innovation in Applied Nursing Informatics, Studies in Health Technology and Informatics*, *315*, 74–80. https://doi.org/10.3233/SHTI240109

Li, M., Chen, S., Liu, S., Yang, J., Qin, Y., Chen, Y., & Tai, X. (2026). A bibliometric analysis of the global research landscape on artificial intelligence applications in clinical medicine (2010–2025). *Digital Health*, *12*. https://doi.org/10.1177/20552076261443381

Li, Y., Li, Y., Zhang, W., & Li, J. (2026). The effectiveness of artificial intelligence-enhanced interventions for cancer patients: A meta-analysis of randomized controlled trials.

*Worldviews on Evidence-Based Nursing*, *23*(1), e70117. https://doi.org/10.1111/wvn.70117

Liu, M., Okuhara, T., Chang, X., Shirabe, R., Nishiie, Y., Okada, H., & Kiuchi, T. (2024). Performance of ChatGPT across different versions in medical licensing examinations worldwide: Systematic review and meta-analysis. *Journal of Medical Internet Research*, *26*, e60807. https://doi.org/10.2196/60807

Luke, A. M., & Rezallah, N. N. F. (2025). Accuracy of artificial intelligence in caries detection: A systematic review and meta-analysis. *Head & Face Medicine*, *21*(1), 24. https://doi.org/10.1186/s13005-025-00496-8

McGenity, C., Clarke, E. L., Jennings, C., Matthews, G., Cartlidge, C., Freduah-Agyemang, H., Stocken, D. D., & Treanor, D. (2024). Artificial intelligence in digital pathology: A systematic review and meta-analysis of diagnostic test accuracy. *Npj Digital Medicine*, *7*(1), 114. https://doi.org/10.1038/s41746-024-01106-8

Milasan, L. H., & Scott-Purdy, D. (2025). The Future of Artificial Intelligence in Mental Health Nursing Practice: An Integrative Review. *International Journal of Mental Health Nursing*, *34*(1), e70003. https://doi.org/10.1111/inm.70003

Nazarian, S., Glover, B., Ashrafian, H., Darzi, A., & Teare, J. (2021). Diagnostic accuracy of artificial intelligence and computer-aided diagnosis for the detection and characterization of colorectal polyps: Systematic review and meta-analysis. *Journal of Medical Internet Research*, *23*(7), e27370. https://doi.org/10.2196/27370

Ong, A. Q. C., Ang, C. S., Bojic, I., Johnson, C. L., Aggour, H., Ng, F. S., & Lai, N. M. (2026). Artificial intelligence in cardiovascular care: A systematic review and meta-analysis of randomised controlled trials. *EClinicalMedicine*, *98*. https://doi.org/10.1016/j.eclinm.2026.104123

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, n71. https://doi.org/10.1136/bmj.n71

Rafati Fard, A., Williams, S. C., Smith, K. J., Dhaliwal, J. K., Ferreira, T., Das, A., & Marcus, H. J. (2026). Comparing artificial intelligence and healthcare professional performance in surgical and interventional video analysis: A systematic review and meta-analysis. *Npj Digital Medicine*, *9*(1), 323. https://doi.org/10.1038/s41746-026-02401-2

Riipa, Md. B., Ahmed, F., Rony, M. K. K., Hossain, A., Islam, A., Utsho, Md. R., & Tasnim, A. F. (2026). The role of artificial intelligence in predicting cardiovascular outcomes: A systematic review and meta-analysis. *Biostatistics & Epidemiology*, *10*(1), e2670804. https://doi.org/10.1080/24709360.2026.2670804

Rony, M. K. K., Das, D. C., Khatun, M. T., Ferdousi, S., Akter, M. R., Khatun, M. A., & Akter, F. (2025). Artificial intelligence in psychiatry: A systematic review and meta-analysis of diagnostic and therapeutic efficacy. *Digital Health*, *11*, 20552076251330528. https://doi.org/10.1177/20552076251330528

Sreedharan, J. K., Saleh, F., Alqahtani, A., Albalawi, I. A., Gopalakrishnan, G. K., Alahmed, H. A., & Karthika, M. (2024). Applications of artificial intelligence in emergency and critical care diagnostics: A systematic review and meta-analysis. *Frontiers in Artificial Intelligence*, *7*, 1422551. https://doi.org/10.3389/frai.2024.1422551

Thong, L. T., Chou, H. S., Chew, H. S. J., & Lau, Y. (2023). Diagnostic test accuracy of artificial intelligence-based imaging for lung cancer screening: A systematic review and meta-analysis. *Lung Cancer*, *176*, 4–13. https://doi.org/10.1016/j.lungcan.2022.12.002

Vasilev, Y. A., Vladzimirskyy, A. V., Omelyanskaya, O. V., Reshetnikov, R. V., Blokhin, I. A., Kodenko, M. M., & Nanova, O. G. (2024). Review of meta-analyses on the use of artificial intelligence in radiology. *Medical Visualization*, *28*(3), 22–41. https://doi.org/10.24835/1607-0763-1425

Waldock, W. J., Zhang, J., Guni, A., Nabeel, A., Darzi, A., & Ashrafian, H. (2024). The accuracy and capability of artificial intelligence solutions in health care examinations and certificates: Systematic review and meta-analysis. *Journal of Medical Internet Research*, *26*, e56532. https://doi.org/10.2196/56532

Wenderott, K., Krups, J., Zaruchas, F., & Weigl, M. (2024). Effects of artificial intelligence implementation on efficiency in medical imaging: A systematic literature review and meta-analysis. *Npj Digital Medicine*, *7*(1), 265. https://doi.org/10.1038/s41746-024-01248-9

Xu, H.-L., Gong, T.-T., Liu, F.-H., Chen, H.-Y., Xiao, Q., Hou, Y., & Wu, Q.-J. (2022). Artificial intelligence performance in image-based ovarian cancer identification: A systematic review and meta-analysis. *EClinicalMedicine*, *53*. https://doi.org/10.1016/j.eclinm.2022.101662

Younis, H. A., Eisa, T. A. E., Nasser, M., Sahib, T. M., Noor, A. A., Alyasiri, O. M., & Younis, H. A. (2024). A systematic review and meta-analysis of artificial intelligence tools in medicine and healthcare: Applications, considerations, limitations, motivation and challenges. *Diagnostics*, *14*(1), 109. https://doi.org/10.3390/diagnostics14010109

Yu, K.-H., Beam, A. L., & Kohane, I. S. (2018). Artificial intelligence in healthcare. *Nature Biomedical Engineering*, *2*(10), 719–731. https://doi.org/10.1038/s41551-018-0305-z

Zheng, Q., Yang, L., Zeng, B., Li, J., Guo, K., Liang, Y., & Liao, G. (2021). Artificial intelligence performance in detecting tumor metastasis from medical radiology imaging: A systematic review and meta-analysis. *EClinicalMedicine*, *31*. https://doi.org/10.1016/j.eclinm.2020.100669

Zwanenburg, A. (2019). Radiomics in nuclear medicine: Robustness, reproducibility, standardization, and how to avoid data analysis traps and replication crisis. *European Journal of Nuclear Medicine and Molecular Imaging*, *46*(13), 2638–2655. https://doi.org/10.1007/s00259-019-04391-8