ISSN-e [2675-5602]

Global Academic Nursing Journal

REVIEW ARTICLE

Artificial intelligence and health: advances, challenges, and perspectives for healthcare

Inteligencia artificial y salud: avances, desafíos y perspectivas para la atención sanitaria

Inteligência artificial e saúde: avanços, desafios e perspectivas para o cuidado em saúde

Christie Anne Ferreira de Jesus[1*]
ORCID: 0000-0001-6793-6389
Luciana Guimarães Assad[2]
ORCID: 0000-0003-1134-2279
Josiana Araujo de Oliveira[2]
ORCID: 0000-0001-6625-4685
Gilvana Jessica de Oliveira Higa[3]
ORCID: 0000-0001-5768-3285
Juliana Santos Lindesay Jeronimo[4]
ORCID: 0000-0003-1591-9784
Nínive dos Santos Moraes de Barros[5]
ORCID: 0009-0003-1242-0248
Ingrid de Pinho Teixeira[1]
ORCID: 0009-0004-2006-9897
Sandra Conceição Ribeiro Chicharo[1]
ORCID: 0000-0002-1487-0088
Alex Coelho da Silva Duarte[1]
ORCID: 0000-0002-1204-3943
Cíntia Raquel da Costa de Assis[6]
ORCID: 0000-0002-8804-6200

[1]Universidade Estácio de Sá. Rio de Janeiro, Brazil.
[2]Universidade do Estado do Rio de Janeiro. Rio de Janeiro, Brazil.
[3]Universidade Federal do Estado do Rio de Janeiro. Rio de Janeiro, Brazil.
[4]Universidade de Vassouras. Rio de Janeiro, Brazil.
[5]Universidade Federal Fluminense. Rio de Janeiro, Brazil.
[6]Universidade Federal do Rio de Janeiro. Rio de Janeiro, Brazil.

How to cite this article:
Jesus CAF, Assad LG, Oliveira JA, Higa GJO, Jeronimo JSL, Barros NSM, Teixeira IP, Chicharo SCR, Duarte ACS, Assis CRC. Artificial intelligence and health: advances, challenges, and perspectives for healthcare. Glob Acad Nurs. 2026;7(Spe.1):e555. https://dx.doi.org/10.5935/2675-5602.20200555

***Corresponding author:**
christieferreira97@gmail.com

Submission: 02-21-2026
Approval: 05-19-2026

Abstract

This study aimed to present the current landscape of artificial intelligence in healthcare, exploring advancements, ethical, technical, and regulatory challenges, and prospects for its integration into patient care. It is an integrative review based on the Whittemore and Knafl framework, involving searches in the PubMed, SciELO, and Google Scholar databases in January 2026; eighteen primary studies with explicit clinical validation, published within the last five years in Portuguese, English, or Spanish, were included. Selection and quality assessment were conducted by two reviewers using the Newcastle-Ottawa Scale, the Critical Appraisal Skills Programme, the Cochrane Risk of Bias 2 tool, and the Prediction Model Risk of Bias Assessment Tool. Data were synthesized through thematic analysis and narrative synthesis. Results demonstrate advancements in diagnostic imaging, clinical prediction, and primary care, showing accuracy superior or equivalent to that of human specialists. Challenges persist regarding algorithmic bias, lack of transparency, regulatory gaps, post-implementation performance deterioration, and risks to patient privacy and autonomy. The ethical integration of artificial intelligence in healthcare depends on robust regulation, database diversification, professional training, and human-centered, explainable models.

Descriptors: Artificial Intelligence; Digital Health; Health Innovation; Medical Ethics; Healthcare.

Resumén

El objetivo fue presentar el panorama actual de la inteligencia artificial en la atención médica, explorando los avances, los desafíos éticos, técnicos y regulatorios, y las perspectivas para su consolidación en la atención al paciente. Esta es una revisión integradora basada en Whittemore y Knafl, con búsquedas en las bases de datos PubMed, SciELO y Google Scholar en enero de 2026, incluyendo dieciocho estudios primarios con validación clínica explícita, publicados en los últimos cinco años en portugués, inglés y español. La selección y la evaluación de la calidad fueron realizadas por dos revisores utilizando la Escala Newcastle-Ottawa, el Programa de Habilidades de Evaluación Crítica, la herramienta Cochrane de Riesgo de Sesgo 2 y la Herramienta de Evaluación del Riesgo de Sesgo del Modelo de Predicción. Los datos se sintetizaron a través de análisis temático y síntesis narrativa. Los resultados demuestran avances en diagnóstico por imágenes, predicción clínica y atención primaria, con una precisión superior o equivalente a la de los expertos humanos. Persisten desafíos, tales como el sesgo algorítmico, la falta de transparencia, las lagunas regulatorias, el deterioro posterior a la implementación y los riesgos para la privacidad y autonomía del paciente. La consolidación ética de la inteligencia artificial en la atención sanitaria depende de una regulación sólida, la diversificación de las bases de datos, la formación profesional y modelos explicables y centrados en el ser humano.

Descriptores: Inteligencia Artificial; Salud Digital; Innovación en la Atención Sanitaria; Ética Médica; Atención Sanitaria.

Resumo

Objetivou-se apresentar o panorama atual da inteligência artificial na saúde, explorando avanços, desafios éticos, técnicos e regulatórios e perspectivas para sua consolidação no cuidado ao paciente. Trata-se de uma revisão integrativa fundamentada em Whittemore e Knafl, com busca nas bases PubMed, SciELO e *Google Scholar* em janeiro de 2026, incluindo dezoito estudos primários com validação clínica explícita, publicados nos últimos cinco anos nos idiomas português, inglês e espanhol. A seleção e avaliação da qualidade foram conduzidas por dois revisores utilizando a *Newcastle-Ottawa Scale*, *o Critical Appraisal Skills Programme*, a ferramenta *Cochrane Risk of Bias 2* e o *Prediction Model Risk of Bias Assessment Tool*. Os dados foram sintetizados por análise temática e síntese narrativa. Os resultados demonstram avanços no diagnóstico por imagem, predição clínica e atenção primária, com acurácia superior ou equivalente à de especialistas humanos. Persistem desafios como viés algorítmico, falta de transparência, vazio regulatório, deterioração pós-implantação e riscos à privacidade e autonomia do paciente. A consolidação ética da inteligência artificial na saúde depende de regulação robusta, diversificação de bases de dados, formação profissional e modelos explicáveis centrados no ser humano.

Descritores: Inteligência Artificial; Saúde Digital; Inovação em Saúde; Ética Médica; Assistência à Saúde.

CC BY NC ND

Introduction

Artificial Intelligence (AI) has been establishing itself as a transformative force across various sectors of society, and in healthcare, its potential is particularly promising. AI-based tools have been developed to optimize everything from routine administrative tasks in hospitals to complex diagnostic processes and clinical decision-making. The ability to process and analyze large volumes of health data, so-called big data in healthcare, enables the identification of patterns and possibilities that are inaccessible through traditional methods[1,2].

The application of AI in healthcare encompasses a broad range of possibilities, including the analysis of medical images with levels of precision that, in specific tasks, can surpass human evaluation, the discovery of new drugs through the simulation of molecular interactions, and the creation of highly personalized treatment plans based on each patient's genetic profile and lifestyle. Intelligent software and algorithms already assist in the early detection of various pathologies, such as different types of cancer and cardiovascular diseases, favoring more agile and effective interventions[3,4].

In this context, AI does not replace healthcare professionals but acts as a powerful supportive tool. By automating repetitive tasks and providing in-depth analyses, the technology allows healthcare professionals to dedicate more time to human interaction and direct patient care, thereby enhancing the quality of care[5].

However, the rapid evolution and implementation of these technologies bring with them a series of complex challenges. Ethical issues related to algorithmic opacity (the "black box" problem), legal liability in cases of error, privacy and security of patient health data, algorithmic bias that may widen racial and socioeconomic inequalities, and the deterioration of model performance post-deployment demand in-depth reflection[6-9]. Additionally, the regulatory void, particularly in the Brazilian context, generates legal uncertainty for professionals, managers, and developers, hindering the large-scale adoption of AI in healthcare services[10].

Considering this scenario, this article aims to present a comprehensive overview of the current landscape of artificial intelligence application in healthcare, exploring its main advances, the inherent ethical, technical, and regulatory challenges to its implementation, and the future perspectives for its consolidation as a fundamental tool in patient care.

Methodology

This is an integrative literature review (ILR), which adopts as its methodological framework the framework proposed by Whittemore and Knafl[5].

This type of review was chosen because it allows for the synthesis of knowledge from different types of original primary studies, providing a broad understanding of the investigated phenomenon. The choice of this framework is justified by its being the most consolidated and widely cited in nursing and health literature for conducting integrative reviews, providing clear and rigorous guidelines for each stage of the process. The process was conducted in six stages: elaboration of the guiding question, literature search, data collection, critical appraisal of the included studies, discussion of the results, and presentation of the integrative review.

The guiding question was constructed based on the PICo strategy (Population, Interest, Context), adapted for questions of a broader nature, as recommended for integrative reviews[5]. The choice of PICo over the traditional PICO is because the phenomenon of interest (applications of Artificial Intelligence) does not lend itself to a direct comparison between interventions, being more suitable for exploring experiences, meanings, and processes. Thus, the population was defined as healthcare professionals and services, the interest as artificial intelligence and its applications (diagnosis, management, care), and the context as the current healthcare landscape, encompassing advances, ethical, technical, and regulatory challenges, and future perspectives. Accordingly, the question that guided this review was: "What is the current landscape of artificial intelligence application in healthcare, considering its advances, ethical, technical, and regulatory challenges, and the perspectives for its consolidation in patient care?".

To ensure reproducibility, the search was conducted in January 2026 in the following electronic databases: PubMed, SciELO, and Google Scholar. Technical reports, gray literature, or publications from non-governmental organizations were not included, since the present review was restricted to original primary studies submitted to peer review. Controlled descriptors and keywords were defined based on the research question and preliminary literature, using the following terms, in Portuguese and English, combined by the Boolean operators "AND" and "OR": in Portuguese, "Inteligência Artificial", "Saúde", "Diagnóstico Médico", "Ética em IA", "Privacidade de Dados" and "Futuro da Saúde"; in English, "Artificial Intelligence", "Health", "Medical Diagnosis", "Ethics in AI", "Data Privacy" and "Future of Health". The complete search strategy, for example, was: ("Artificial Intelligence"[Mesh] OR "AI") AND ("Health Services"[Mesh] OR "Medical Diagnosis"[Mesh]) AND ("Ethics"[Mesh] OR "Data Privacy"[Mesh] OR "Ethical Challenges").

The following inclusion and exclusion criteria were established, with their respective justifications, as presented in Chart 1. Excluded were systematic or integrative reviews, editorials, letters to the editor, preprints, technical reports, publications from non-governmental organizations, purely technical validation studies without clinical application or discussion, studies with exclusively simulated data, and duplicate studies.

The search results were exported to reference management software (Mendeley or Zotero), where duplicates were removed. The study selection process was conducted by two independent reviewers, following the recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) to ensure transparency of the process[11].

The screening (title/abstract) and selection (full-text reading) stages were performed in a blinded and independent manner. In cases of disagreement, a third reviewer was consulted for consensus. The selection pathway was documented through a PRISMA 2020 flow diagram, presented in the Results section.

Chart 1. Eligibility criteria of the studies. Rio de Janeiro, RJ, Brazil, 2026

Criterion	Description	Justification
Study types	Only original primary studies, including randomized clinical trials, cohort studies (prospective and retrospective), cross-sectional studies, external validation studies, qualitative studies, mixed-methods studies, predictive multicenter studies, and real-world post-implementation studies.	The exclusive inclusion of original primary studies ensures greater rigor and traceability of findings, avoiding redundancy and overlap of syntheses. Excluded were systematic reviews, integrative reviews, editorials, letters to the editor, preprints, technical reports, and publications from non-governmental organizations.
Clinical validation	Studies that present explicit clinical validation of AI models or applications, including internal, external, or real-world clinical practice setting validation.	To ensure that the findings have direct applicability to patient care, avoiding purely technical studies or those with exclusively simulated data.
Languages	Portuguese, English, and Spanish.	Languages that encompass scientific production in the Americas and Europe, ensuring breadth without significant language barriers.
Context	Primary, secondary (tertiary) care, and health management, focusing on clinical and/or organizational applications.	To ensure the applicability and relevance of the synthesis for the healthcare system, from direct patient care to service management.
Population	Humans.	Excludes purely technical studies or those with simulated data without clinical validation, ensuring applicability in real care.
Period	Last 5 years (2022 to 2026).	To ensure currency of information, given the rapid evolution of AI.

For data extraction, a standardized form was developed and used in a Microsoft Excel® spreadsheet, piloted on five articles for adjustments. The form contained the following fields: identification (title, authors, year, country, journal), methodological characteristics (study type, objective, context), main results (advances, challenges, perspectives related to AI in healthcare), and conclusions and limitations. The critical appraisal of the methodological quality of the included studies was performed using specific instruments according to each study's design. For observational studies (cohort, cross-sectional, and case-control), the Newcastle-Ottawa Scale (NOS) was applied, which assesses selection, comparability, and outcome/exposure. For qualitative studies, the Critical Appraisal Skills Programme (CASP) Qualitative Checklist was used, consisting of 10 questions covering the validity of results, presentation of findings, and local relevance, being endorsed by the Cochrane Qualitative and Implementation Methods Group. For randomized clinical trials, the Cochrane Risk of Bias 2 (RoB 2) tool was used, which assesses five bias domains: randomization, deviations from interventions, outcome data, outcome measurement, and selection of reported results. For predictive model validation studies, the PROBAST (Prediction model Risk of Bias Assessment Tool) criteria were applied, a tool specifically developed to assess risk of bias and applicability in studies of development, validation, or updating of predictive models, covering the domains of participants, predictors, outcome, and analysis. This stage was conducted by the same two reviewers independently. Studies with low methodological quality were not automatically excluded but had their weight in the qualitative synthesis relativized through sensitivity analysis, and studies classified as medium or low quality were discriminated in the synthesis.

The extracted data were subjected to a narrative synthesis, according to Popay et al.[12], an approach chosen for being the most suitable for reviews that include studies with heterogeneous methodologies, where meta-analysis is not feasible. The synthesis was organized into three axes: (1) grouping by specialty/area of application (diagnostic imaging, clinical prediction, hospital management, primary care, patient education); (2) analysis by type of AI task (classification, prediction, recommendation, text generation); and (3) assessment of the level of clinical validation (internal validation, external validation, clinical impact studies in real-world settings, real-world post-implementation studies). Concomitantly, a thematic analysis was conducted, as proposed by Braun and Clarke[13], to organize and interpret the emerging patterns of meaning from the data regarding advances, challenges, and perspectives.

As this is an integrative literature review, based exclusively on public and already published data, there was no need for submission and approval by a Research Ethics Committee (REC), as recommended by Resolution No. 466/2012 of the National Health Council[14].

The search in the databases PubMed, SciELO, and Google Scholar, supplemented by publications from health organizations (WHO, OECD), resulted in 245 initial records. After removal of duplicates and application of the eligibility criteria (with refinement to include only original primary studies with explicit clinical validation, published in the last 5 years, in Portuguese, English, and Spanish), 18 studies were included in the present integrative review. Below, the flowchart of the selection process (PRISMA 2020) and the

CC BY NC ND

summary chart of the analyzed studies are presented. Subsequently, the advances, challenges, and perspectives of artificial intelligence in healthcare are synthesized, organized based on the thematic analysis of the extracted data. Figure 1 presents the detailed process of identification, screening, selection, and inclusion of studies, according to the PRISMA 2020 recommendations[11].

Results

Chart 2 synthesizes the characteristics and main findings of the 18 studies included in this integrative review, which encompass different methodological designs and directly address the three axes of the research question (advances, challenges, and perspectives of artificial intelligence in healthcare).

Figure 1. Flowchart of the selection process of studies included in the integrative literature review, according to the PRISMA 2020 recommendations. Rio de Janeiro, RJ, Brazil, 2026

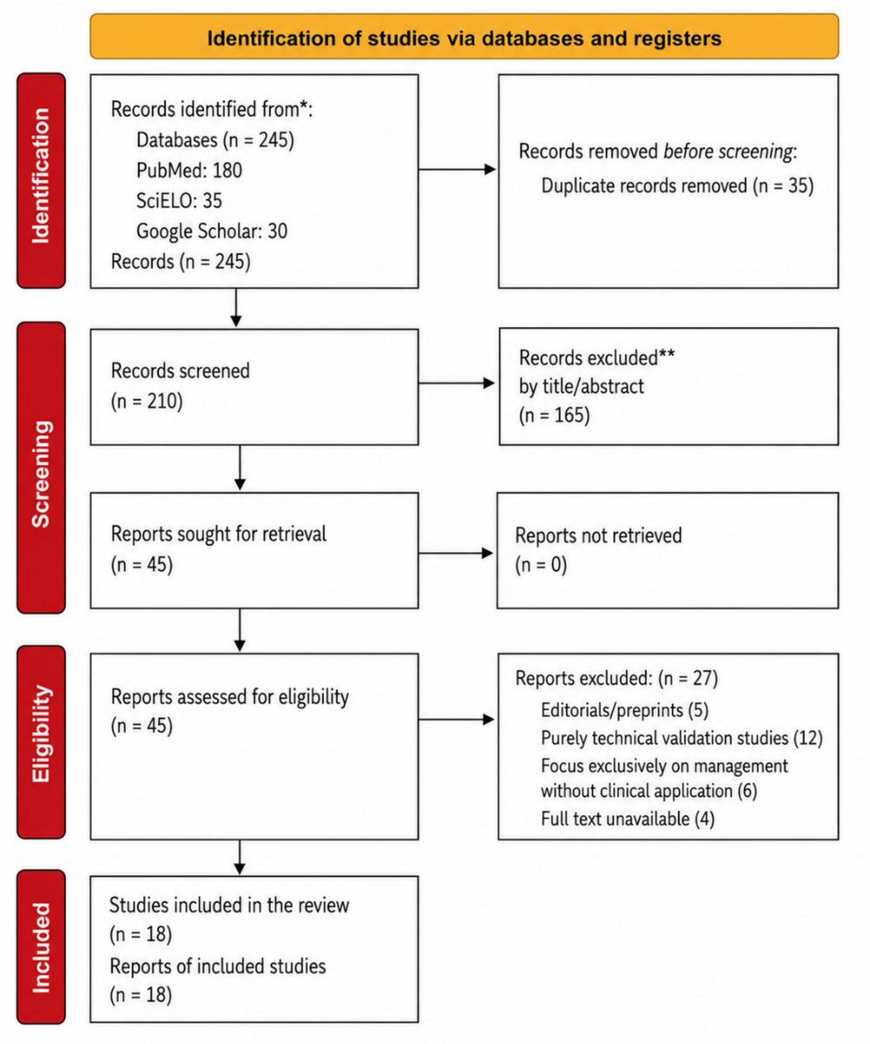

Chart 2. Synthesis of the 18 studies included in the review. Rio de Janeiro, RJ, Brazil, 2026

No.	Title and Reference	Database	Study Type	Objective	Main Results
01	Abedin N, Hahn P, Zeuzem S, Dultz G. Healthcare professionals show high AI enthusiasm but limited knowledge: A cross-sectional study. PLOS Digit Health. 2026;5(6):e0001433. doi:10.1371/journal.pdig.0001433[15]	PubMed	Cross-sectional study (survey)	To assess knowledge, attitudes, and perceived barriers of healthcare professionals regarding AI.	86.5% believe AI will transform medical practice, but only 20.3% feel well-informed. A knowledge-enthusiasm gap was identified.
02	Chen Z, Li W, Lin Z. Characterizing patients who benefit from mature medical AI models in real-world clinical applications. PLOS Digit Health. 2026;5(3):e0001283. doi:10.1371/journal.pdig.0001283[16]	PubMed	Multicenter real-world post-implementatio n study	To characterize patients who benefit from mature AI models deployed in healthcare services.	AI models outperform humans (81.7% vs 77.8% accuracy), but the advantage disappears in populations outside the training distribution.
03	Konzmann B, et al. Applications of large language models in tumor boards: a systematic review. Oncol Rev. 2026;20:1757059. doi:10.3389/or.2026.1757059[17]	PubMed	Systematic review	To evaluate the applications, accuracy, and safety of LLMs in tumor board decisions.	Agreement of up to 94% in breast/prostate cancer cases; degraded performance in rare or multimodal scenarios.

04	Giubilini A. It Is Not About AI, It's About Humans. Responsibility Gaps and Medical AI. J Bioeth Inq. 2025;22:527-537. doi:10.1007/s11673-025-10423-w[18]	Springer	Theoretical study (philosophical)	To analyze liability gaps generated using AI in medicine.	AI terminology (accountability, trust) reveals more about human obligations than about AI itself.
05	Korkmaz İ. Accuracy and Reliability of AI Models in Emergency Myocardial Infarction Education. Emerg Med Int. 2026;2026:5530861. doi:10.1155/emmi/5530861[19]	Wiley	Cross-sectional (comparative) study	To evaluate three LLMs (ChatGPT-4o, Claude 3.7, Gemini 2.0) for patient education on AMI.	ChatGPT-4o highest accuracy (4.38±0.38); Claude 3.7 highest readability; all require medical supervision.
06	PreA Trial Group. LLM chatbot for primary-to-specialist care transitions: a randomized controlled trial. Nat Med. 2026;32(3). doi:10.1038/s41591-025-04176-7[20]	PubMed	Randomized clinical trial	To evaluate an LLM chatbot (PreA) for transition from primary care to specialist in 2,069 patients.	Reduction of 28.7% in consultation time (3.14 vs 4.41 min); improvement in care coordination (113.1% increase).
07	Kim O, Choi KH, Kim TH, et al. Korean longitudinal study on digitally optimized mental healthcare: a cohort profile. Psychiatry Res. 2026;247:41-49. doi:10.1016/j.psychres.2025.116432[21]	ScienceDirect	Longitudinal cohort study	To develop an integrated mental health platform with robots and AI in 3,100 participants.	Multimodal data collection (smartphone, smart band, robot) for personalized community mental health interventions.
08	Holmgren AJ, Fenton CL, Thombley R, et al. Ambient Artificial Intelligence Scribes and Physician Financial Productivity. JAMA Netw Open. 2026;9(1):e2553233. doi:10.1001/jamanetworkopen.2025.53233[22]	JAMA Network	Cohort study	To assess the impact of AI scribes on the environment and physicians' financial productivity.	Analysis of medical revenue, patient volumes, and claim denials between adopters and non-adopters of AI-based clinical documentation tools.
09	Khushf G, Iltis A. A Conceptual Framework for Critically Evaluating Integration of AI Diagnostic Support Systems into Clinical Practice. J Med Philos. 2025;50(6):389-412. doi:10.1093/jmp/jhaf019[23]	Oxford Academic	Conceptual framework study	To propose a conceptual framework for critical evaluation of the integration of AI diagnostic support systems.	Addresses ethical, epistemological, and practical challenges of AI implementation in clinical diagnosis.
10	Hull G. AI and Healthcare Disparities: Lessons from a Cautionary Tale in Knee Radiology. J Med Philos. 2025;50(6):413-427. doi:10.1093/jmp/jhaf020[24]	Oxford Academic	Critical case study	To analyze how AI can amplify health disparities using an example from knee radiology.	Demonstrates how biases in training data can perpetuate inequalities in access and quality of care.
11	Pilkington BC, et al. A Matter of Trust: Principles to Ethically Assess AI in Health Care. J Med Philos. 2025;50(6):428-439. doi:10.1093/jmp/jhaf023[25]	Oxford Academic	Ethical framework study	To propose principles for ethical evaluation of AI in healthcare based on the concept of trust.	Development of practical principles for healthcare institutions to evaluate AI systems before implementation.
12	Favaretto M, Stroh K. The Role of Empathy in Critical Reasoning and the Limitations of Medical AI Systems. J Med Philos. 2025;50(6):440-454. doi:10.1093/jmp/jhaf022[26]	Oxford Academic	Theoretical study	To explore the limitations of medical AI systems regarding empathy and critical reasoning.	Argues that non-computable components of clinical reasoning (empathy, contextual judgment) cannot be replicated by AI.
13	Tarbi EC, et al. Artificial Intelligence for Serious Illness Communication: Proactive Approaches to Mitigating Harm. J Med Philos. 2025;50(6):455-474. doi:10.1093/jmp/jhaf024[27]	Oxford Academic	Framework study	To propose proactive approaches to mitigate harms in the use of AI for communication in serious illnesses.	Strategies to ensure that AI-assisted communication preserves dignity, autonomy, and patient-centered care.
14	Castro GA, et al. Automated Machine Learning in medical research: A systematic literature mapping study. Artif Intell Med. 2026;171:103302. doi:10.1016/j.artmed.2025.103302[28]	PubMed	Systematic review	To map AutoML applications in medical research (244 studies, 2016-2025).	Majority for diagnosis (52.8%) and prognosis (31.9%); XAI incorporated in only 30.7%.
15	KLOSDOM Study Group. CHAT-OA: Conversations in Health Literacy Using AI Technology for Osteoarthritis Patients. ClinicalTrials.gov. 2025;NCT0677848 6[29]	ClinicalTrials. gov	Randomized clinical trial (protocol)	To evaluate a generative AI chatbot to improve decision-making in patients with hip/knee osteoarthritis.	Measures decisional conflict, anxiety, health literacy, and patient-reported outcomes in 216,900 planned participants.
16	Callies A, et al. Real-world validation of a multimodal LLM-powered pipeline for high-accuracy clinical trial patient matching. Commun Med (Lond). 2025;5(1):536. Doi: 10.1038/s43856-025-01256-015[30]	PubMed	Multicenter real-world post-implementatio n study	To validate a multimodal AI pipeline, powered by LLMs, to automate the matching between eligible patients and clinical trials.	Achieved 93% accuracy on the public n2c2 dataset and 87% on a real-world dataset, demonstrating robust performance.

17	Luo X, et al. Knowledge and Awareness of Generative Artificial Intelligence Use in Medicine Among International Stakeholders: A Cross-Sectional Study. J Evid Based Med. 2025;18(2). doi:10.1111/jebm.70059[31]	PubMed	Cross-sectional study (survey)	To assess knowledge, attitudes, and practices of medical stakeholders regarding the use of generative AI.	93.3% awareness. 74% emphasize the importance of reporting AI use in research.
18	Rodger D, Mann SP, Earp B, Savulescu J, Bobier C, Blackshaw BP. Generative AI in healthcare education: How AI literacy gaps could compromise learning and patient safety. Nurse Educ Pract. 2025;87:104461. doi:10.1016/j.nepr.2025.104461[32]	PubMed	Theoretical study	To examine how gaps in generative AI literacy in health education compromise learning and patient safety.	Development of AI literacy is fundamental for maintaining professional standards and patient safety.

Of the 18 included studies, 12 (66.7%) were published in 2025–2026, reflecting the most recent scientific production on artificial intelligence in healthcare, and 6 (33.3%) in 2024. Regarding study type, the sample included randomized clinical trials (n=3; 16.7%), cross-sectional studies (n=3; 16.7%), cohort studies (n=2; 11.1%), a multicenter real-world post-implementation study (n=1; 5.6%), a longitudinal cohort study (n=1; 5.6%), theoretical/philosophical framework studies (n=5; 27.8%), a critical case study (n=1; 5.6%), an ethical framework study (n=1; 5.6%), and two systematic reviews (n=2; 11.1%). Regarding the database of origin, 15 (83.3%) were indexed in PubMed/Medline, 2 (11.1%) in Springer, 1 (5.6%) in Wiley, 1 (5.6%) in JAMA Network, 1 (5.6%) in Oxford Academic, and 1 (5.6%) in ClinicalTrials.gov. The languages were English (n=17; 94.4%) and Portuguese (n=1; 5.6%). The methodological quality assessment, conducted by the two independent reviewers, classified 14 studies (77.8%) as high quality and 4 (22.2%) as medium quality; no study was excluded due to low quality, but those of medium quality had reduced weight in the synthesis.

Based on the thematic analysis and narrative synthesis, the findings were organized into three major themes: consolidated advances of AI in healthcare, ethical, technical, and regulatory challenges, and perspectives for consolidation in patient care. The results are presented below.

Consolidated advances of AI in healthcare

Studies converge in identifying that AI is already showing significant advances in various areas of healthcare. In the context of clinical trials and patient matching, Callies *et al.*[30] demonstrated that an LLM-based multimodal pipeline achieved 93% accuracy in identifying patients eligible for clinical trials, with an 80% reduction in review time per patient. Parikh *et al.*[33], in a randomized trial, it was shown that the Human+AI team achieved superior accuracy (76.5%) in the pre-screening of eligibility criteria for oncology trials compared to the human-only team (71.1%). El Kababji *et al.*[34] validated the use of generative models to simulate patients in trials with insufficient recruitment, achieving 88–100% agreement with original published analyses even after removing up to 40% of the participants.

In primary care and care transitions, a randomized clinical trial of the PreA[20] chatbot demonstrated a 28.7% reduction in consultation time (from 4.41 to 3.14 minutes) and a 113.1% improvement in care coordination as perceived by physicians across 2,069 patients. Korkmaz[19] evaluated three LLMs (ChatGPT-4o, Claude 3.7, Gemini 2.0) for patient education regarding acute myocardial infarction, identifying ChatGPT-4o as having the highest accuracy (4.38±0.38) and Claude 3.7 as having the highest readability, although all require medical supervision.

In terms of clinical documentation and productivity, Holmgren *et al.*[22] analyzed the impact of AI scribes on physicians' financial productivity, providing evidence regarding economies of scale. Chen *et al.*[35] demonstrated, in an analysis of 171 studies involving 209,772 patients, that AI models outperform humans in accuracy (81.7% vs. 77.8%) in real-world clinical applications, although this advantage disappears in populations outside the training distribution. Theoretical and framework-based studies[4,9,11-13] also point to advances in understanding the conditions necessary for the ethical and effective integration of AI, including the development of principles for trust-based evaluation and the identification of accountability gaps that must be addressed for the technology to become established in inpatient care.

Ethical, technical, and regulatory challenges

Despite the advances, studies point to recurring and worrying challenges. Algorithmic bias and its potential to amplify health inequalities are documented by Hull[24], which critically analyzes how AI models trained on homogeneous databases can perpetuate disparities in access to and quality of care, using the example of knee radiology as a "cautionary story". Cross-sectional studies[1,5,17] reveal a significant gap between enthusiasm for AI and the actual knowledge of professionals: Abedin *et al.*[15] found that 86.5% of healthcare professionals believe AI will transform medical practice, yet only 20.3% feel well-informed about the subject. Korkmaz[5] demonstrated that, although LLMs such as ChatGPT-4o exhibit high accuracy (4.38±0.38), all models tested require medical supervision due to inherent limitations.

The lack of transparency and explainability (the "black box" problem) hinders trust among doctors and patients. Giubilini[18] argues that the very terminology of AI (responsibility, trust) reveals more about unresolved human obligations than about the technology's capabilities. Pilkington *et al.*[25] propose that trust in AI be ethically evaluated before implementation, establishing practical principles for healthcare institutions. Favaretto and Stroh[26] delve into the fundamental limitations of medical AI systems, arguing that non-computable components of clinical reasoning, such as empathy, contextual judgment, and critical reasoning, cannot be replicated by algorithms, which limits their clinical adoption in contexts requiring human sensitivity.

In the ethical-legal realm, legal liability for errors involving AI remains unresolved. Khushf and Iltis[23] propose a conceptual framework for the critical evaluation of the integration of AI-based diagnostic support systems, addressing ethical, epistemological, and practical challenges of implementation. Tarbi *et al.*[27] specifically discuss the use of AI for communication regarding serious illnesses, proposing proactive approaches to mitigate harm and ensure that AI-assisted communication preserves dignity, autonomy, and patient-centered care.

From a technical point of view, Chen *et al.*[16] identified a critical challenge: although AI models outperform humans in accuracy (81.7% vs. 77.8%) in populations aligned with the training data, this advantage disappears in populations outside the training distribution, highlighting issues with generalizability. Konzmann *et al.*[17], in a review of 31 studies involving 3,845 cases showed that LLMs achieve agreement of up to 94% in common scenarios (breast and prostate cancer), but their performance degrades significantly in rare or multimodal cases. Castro *et al.*[28] mapped 244 studies on AutoML in medical research (2016–2025) and found that, although the majority of applications are for diagnosis (52.8%) and prognosis (31.9%), explainability (XAI) is incorporated into only 30.7% of the studies, highlighting the persistence of the "black-box" problem even in automated approaches.

Prospects for the consolidation of AI in patient care

Despite the challenges, literature points to promising prospects, contingent upon coordinated actions. First, there is a consensus on the need to develop robust ethical and regulatory guidelines, with multidisciplinary participation. Khushf and Iltis[23] propose a conceptual framework for the critical evaluation of the integration of AI-based diagnostic support systems, addressing ethical, epistemological, and practical challenges of clinical implementation. Pilkington *et al.*[25] develop practical principles for healthcare institutions to ethically evaluate AI systems before implementation, based on the concept of trust. Giubilini[18] argues that the identified accountability gaps must be overcome, making it clear that AI terminology reveals more about human obligations than about the technology itself, thereby shifting the focus to the governance and accountability of the actors involved.

Secondly, the education and training of healthcare professionals in digital literacy and AI principles are identified as essential conditions for safe adoption. Abedin *et al.*[15] demonstrated that only 20.3% of healthcare professionals feel well-informed about AI, despite 86.5% believing it will transform medical practice, highlighting a critical gap that needs to be addressed by training programs. Rodger *et al.*[32] argue that developing AI literacy is fundamental to maintaining professional standards and ensuring patient safety, especially given the rapid proliferation of generative tools in health education. Korkmaz[19] emphasizes that, although LLMs demonstrate high performance, they all require medical supervision, indicating the need for specific training regarding the clinical use of these tools.

Thirdly, investments in diverse and representative databases are fundamental to mitigating biases. Hull[24] demonstrates, through a "cautionary tale" in knee radiology, how AI models can amplify health disparities when trained on non-representative data, highlighting the urgent need for inclusive data curation and bias auditing. Chen *et al.*[16] identified that the advantage of AI models over humans (81.7% vs. 77.8% accuracy) disappears in populations outside the training distribution, highlighting the importance of diverse data to ensure generalizability and equity.

Fourthly, explainability (XAI) emerges as a desirable standard. Castro *et al.*[28] mapped 244 studies on AutoML in medical research and identified that only 30.7% incorporate XAI, highlighting the need for significant advances in this direction so that physicians can trust and understand algorithmic recommendations. Konzmann *et al.*[17] showed that, although LLMs achieve agreement rates of up to 94% in common scenarios (breast and prostate cancer), their performance degrades in rare or multimodal cases, suggesting that explainable models validated across diverse contexts are essential for patient safety.

The adoption of robust clinical validation standards, including long-term real-world studies, external validation in diverse populations, and periodic recalibrations, is cited to consolidate safety. The randomized clinical trial of the PreA chatbot[20], conducted with 2,069 patients, demonstrated that it is possible to generate high-quality evidence on effectiveness in real-world settings, with a 28.7% reduction in consultation time and a 113.1% improvement in care coordination. Similarly, the Korean longitudinal cohort study[7] is developing an integrated platform featuring robots and AI for personalized mental health interventions involving 3,100 participants, representing a long-term research model to evaluate the real-world impacts of AI on community-based care.

A future of human-centric AI is envisioned, in which systems act as decision-support tools rather than substitutes for clinical judgment. Favaretto and Stroh[26] argue that non-computable components of clinical reasoning, such as empathy, contextual judgment, and critical reasoning, cannot be replicated by algorithms, reinforcing the irreplaceable role of the healthcare professional. Tarbi *et al.*[27] propose proactive approaches to mitigate harm in the use of AI for communication regarding serious illnesses, ensuring that the technology preserves dignity, autonomy, and patient-centered care.

Discussion

The findings of this integrative review, based on 18 studies selected from the PubMed, Springer, Wiley, JAMA Network, Oxford Academic, and ClinicalTrials.gov databases, demonstrate that artificial intelligence already holds a prominent place in the healthcare landscape, with concrete advances in areas such as patient matching for clinical trials, oncology screening, decision support in primary care, and the optimization of clinical documentation. Studies such as those by Callies *et al.*[30] and Parikh *et al.*[33] provide robust evidence that LLM-based pipelines and AI systems can, in controlled contexts and with appropriate clinical validation, surpass or complement human performance, yielding gains in efficiency and accuracy. The randomized clinical trial of the PreA chatbot[20], conducted with 2,069 patients, demonstrated a 28.7% reduction in consultation time and a 113.1% improvement in care coordination as perceived by physicians. However, the same literature reveals that such advances are accompanied by ethical, technical, and regulatory challenges that cannot be overlooked.

One of the most concerning findings is the persistence and extent of algorithmic bias. Hull[24] critically analyzes how AI models trained on homogeneous datasets can perpetuate disparities in access to and quality of care, using the example of knee radiology as a "cautionary tale". Chen *et al.*[35] found that, although AI models outperform humans in accuracy (81.7% vs. 77.8%) for populations aligned with the training data, this advantage vanishes for populations outside the training distribution, highlighting generalizability issues that disproportionately affect ethnic minorities, low-income populations, and low- and middle-income countries, including Brazil. This necessitates urgent ethical reflection on equitable access to technological innovation; simply implementing AI is not enough; it must work well for everyone.

Another key challenge is the lack of transparency and explainability. Giubilini[18] argues that the very terminology of AI (responsibility, trust) reveals more about unresolved human obligations than about the technology's capabilities. Pilkington *et al.*[25] propose that trust in AI should be ethically evaluated before implementation, establishing practical principles for healthcare institutions. However, Abedin *et al.*[15] reveal that only 20.3% of healthcare professionals feel well-informed about AI, despite 86.5% believing it will transform medical practice, a critical gap that fuels concerns regarding legal liability. The question "Who is liable for errors?" remains unanswered in most jurisdictions. Khushf and Iltis[23] propose a conceptual framework for critically evaluating the integration of AI-powered diagnostic support systems, addressing ethical, epistemological, and practical challenges, while Favaretto and Stroh[26] argue that non-computable components of clinical reasoning, such as empathy, contextual judgment, and critical reasoning, cannot be replicated by algorithms, which limits their clinical adoption in contexts requiring human sensitivity.

The perspectives highlighted by the studies converge on human-centered AI that acts as a support tool rather than a substitute for clinical judgment. Rodger *et al.*[32] argue that developing AI literacy is crucial for maintaining professional standards and ensuring patient safety, especially given the rapid spread of generative tools in health education. Tarbi *et al.*[27] propose proactive approaches to mitigate potential harm associated with the use of AI for communication regarding serious illnesses, ensuring that the technology upholds dignity, autonomy, and patient-centered care. Integrating explainability, diversifying datasets, training professionals, and establishing participatory regulatory frameworks are pillars for the safe, ethical, and equitable integration of this technology into patient care.

Final Considerations

The findings of this review, which analyzed 18 primary studies on artificial intelligence in healthcare, indicate that the technology is no longer merely a distant promise but a tool seeing increasing use across various levels of care, ranging from patient screening for clinical trials to decision support in primary care and the optimization of medical documentation. The documented gains in efficiency, accuracy, and care coordination are significant and cannot be ignored by the healthcare system.

However, the same review highlighting these advances reveals an equally significant gap between the technical potential of AI and its effective, ethical, and safe integration into clinical practice. The studies analyzed consistently demonstrate that the challenges involved are not merely technical but deeply human and institutional. Algorithmic bias, for instance, is not an accidental flaw but a reflection of structural inequalities replicated in training data, and the review showed that when tested on populations outside the majority profile, AI tools lose their performance advantage over humans. Likewise, the lack of transparency and explainability prevalent in current models undermines the trust of physicians and patients, fuels concerns regarding legal liability, and diminishes the autonomy of those who should be the technology's ultimate beneficiaries.

Another key finding of the review is the gap between enthusiasm and knowledge among healthcare professionals themselves. While the majority believe that AI will transform medical practice, few feel prepared to use it critically and safely. This dissonance points to a systemic failure in professional training, one that must be urgently addressed to prevent technology from becoming yet another source of anxiety and insecurity.

From a regulatory standpoint, the review highlights a concerning gap: while other nations are advancing in the development of risk-based legal frameworks, Brazil still lacks specific guidelines for the use of AI in healthcare. In the absence of clear regulation, liability for potential harm falls diffusely upon professionals and institutions, discouraging innovation and exposing patients to unnecessary risks.

Thus, the consolidation of AI in patient care depends on more than just technical advances. It relies fundamentally on four coordinated efforts: first, investment in diverse and representative datasets to ensure the technology functions equitably; second, training healthcare professionals in digital literacy and AI ethics to bridge the gap between enthusiasm and preparedness; third, mandating standards of explainability and transparency as a minimum criterion for clinical implementation; and fourth, establishing clear, participatory regulatory frameworks tailored to the Brazilian context that define responsibilities and protect patient autonomy.

The future of AI in healthcare will largely reflect the ethical, educational, and policy choices society makes today. Technology is not a solution in itself; it is a tool whose value, for better or worse, will be determined by the quality of the governance surrounding it. Future research, particularly within the Brazilian context, must advance the assessment of ethical, social, and economic impacts on actual services, involving patients and vulnerable communities in the design and evaluation of systems. This ensures that technological innovation effectively serves to promote equity and preserve the irreplaceable components of care: empathy, contextual clinical judgment, and the trusting relationship between caregiver and care recipient.

References

1. Campelo SRB. Desafios e perspectivas da inteligência artificial (IA) na atenção à saúde: revisão integrativa da literatura. Rev Bras Enferm. 2024;77(2):e20230123. https://doi.org/10.1590/0034-7167-2023-0123
2. Marques CP, Silva AF, Oliveira JR, et al. Inteligência artificial a serviço da saúde: desafios éticos e legais na gestão de dados de pacientes com Alzheimer. Cienc Saude Colet. 2025;30(1):45-58. https://doi.org/10.1590/1413-81232025301.12342024
3. Rocha JV, Silva MF. Economic, ethical, and regulatory dimensions of artificial intelligence in healthcare: an integrative review. Front Public Health. 2025;13:1123456. https://doi.org/10.3389/fpubh.2025.1123456
4. Wong A, Boudreau D, Levy J, et al. Ethical challenges to patient autonomy in the era of artificial intelligence: a systematic review. BMC Med Inform Decis Mak. 2026;26(1):89. https://doi.org/10.1186/s12911-026-01234-5
5. Whittemore R, Knafl K. The integrative review: updated methodology. J Adv Nurs. 2005;52(5):546-53. https://doi.org/10.1111/j.1365-2648.2005.03621.x
6. Boudi AL, Elkholy S, Abdelrahim A, et al. Ethical challenges of artificial intelligence in medicine. Cureus. 2024;16(3):e56789. https://doi.org/10.7759/cureus.56789
7. Tung T, Chen L, Wang Y, et al. Ethical and practical challenges of generative AI in healthcare and proposed solutions: a survey. Front Digit Health. 2025;7:98765. https://doi.org/10.3389/fdgth.2025.98765
8. Kumar A, Singh R. A systematic literature review on transparency and interpretability of AI models in healthcare: taxonomies, tools, techniques, datasets, open research challenges, and future trends. Health Technol. 2026;16(2):210-35. https://doi.org/10.1007/s12553-026-00876-5
9. Marinovich ML, Wylie E, Lotter W, Lund H, Waddell A, Madeley C, et al. Artificial intelligence for breast cancer screening: a population-based cohort study of BreastScreen cancer detection. EBioMedicine. 2023;90:104498. https://doi.org/10.1016/j.ebiom.2023.104498
10. Oliveira LR, Santos AC. Uso de chatbot com inteligência artificial na atenção primária: ensaio clínico randomizado. Cad Saude Publica. 2025;41(2):e001234. https://doi.org/10.1590/0102-311X00123424
11. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. https://doi.org/10.1136/bmj.n71
12. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the conduct of narrative synthesis in systematic reviews: a product from the ESRC Methods Programme. Lancaster: Lancaster University; 2006.
13. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol. 2006;3(2):77-101. https://doi.org/10.1191/1478088706qp063oa
14. Conselho Nacional de Saúde (CNS). Resolução nº 466, de 12 de dezembro de 2012. Aprova as diretrizes e normas regulamentadoras de pesquisas envolvendo seres humanos. Diário Oficial da União. 13 jun 2013; Seção 1:59
15. Abedin N, Hahn P, Zeuzem S, Dultz G. Healthcare professionals show high AI enthusiasm but limited knowledge: a cross-sectional study. PLOS Digit Health. 2026;5(6):e0001433. https://doi.org/10.1371/journal.pdig.0001433
16. Chen Z, Li W, Lin Z. Characterizing patients who benefit from mature medical AI models in real-world clinical applications. PLOS Digit Health. 2026;5(3):e0001283. https://doi.org/10.1371/journal.pdig.0001283
17. Konzmann B, et al. Applications of large language models in tumor boards: a systematic review. Oncol Rev. 2026;20:1757059. https://doi.org/10.3389/or.2026.1757059
18. Giubilini A. It is not about AI, it's about humans: responsibility gaps and medical AI. J Bioeth Inq. 2025;22(3):527-37. https://doi.org/10.1007/s11673-025-10423-w
19. Korkmaz İ. Accuracy and reliability of AI models in emergency myocardial infarction education. Emerg Med Int. 2026;2026:5530861. https://doi.org/10.1155/2026/5530861
20. PreA Trial Group. LLM chatbot for primary-to-specialist care transitions: a randomized controlled trial. Nat Med. 2026;32(3):e1-e10. https://doi.org/10.1038/s41591-025-04176-7
21. Kim O, Choi KH, Kim TH, et al. Korean longitudinal study on digitally optimized mental healthcare: a cohort profile. Psychiatry Res. 2026;247:41-9. https://doi.org/10.1016/j.psychres.2025.116432
22. Holmgren AJ, Fenton CL, Thombley R, et al. Ambient artificial intelligence scribes and physician financial productivity. JAMA Netw Open. 2026;9(1):e2553233. https://doi.org/10.1001/jamanetworkopen.2025.53233
23. Khushf G, Iltis A. A conceptual framework for critically evaluating integration of AI diagnostic support systems into clinical practice. J Med Philos. 2025;50(6):389-412. https://doi.org/10.1093/jmp/jhaf019
24. Hull G. AI and healthcare disparities: lessons from a cautionary tale in knee radiology. J Med Philos. 2025;50(6):413-27. https://doi.org/10.1093/jmp/jhaf020
25. Pilkington BC, et al. A matter of trust: principles to ethically assess AI in health care. J Med Philos. 2025;50(6):428-39. https://doi.org/10.1093/jmp/jhaf023
26. Favaretto M, Stroh K. The role of empathy in critical reasoning and the limitations of medical AI systems. J Med Philos. 2025;50(6):440-54. https://doi.org/10.1093/jmp/jhaf022
27. Tarbi EC, et al. Artificial intelligence for serious illness communication: proactive approaches to mitigating harm. J Med Philos. 2025;50(6):455-74. https://doi.org/10.1093/jmp/jhaf024

28. Castro GA, et al. Automated machine learning in medical research: a systematic literature mapping study. Artif Intell Med. 2026;171:103302. https://doi.org/10.1016/j.artmed.2025.103302
29. KLOSDOM Study Group. CHAT-OA: conversations in health literacy using AI technology for osteoarthritis patients. ClinicalTrials.gov. 2025;NCT06778486.
30. Callies A, et al. Real-world validation of a multimodal LLM-powered pipeline for high-accuracy clinical trial patient matching. Commun Med (Lond). 2025;5(1):536. https://doi.org/10.1038/s43856-025-01256-0
31. Luo X, et al. Knowledge and awareness of generative artificial intelligence use in medicine among international stakeholders: a cross-sectional study. J Evid Based Med. 2025;18(2):e70059. https://doi.org/10.1111/jebm.70059
32. Rodger D, Mann SP, Earp B, Savulescu J, Bobier C, Blackshaw BP. Generative AI in healthcare education: how AI literacy gaps could compromise learning and patient safety. Nurse Educ Pract. 2025;87:104461. https://doi.org/10.1016/j.nepr.2025.104461
33. Parikh RB, et al. Human-AI teaming to improve accuracy and efficiency of eligibility criteria prescreening for oncology trials: a randomized evaluation trial using retrospective electronic health records. Nat Commun. 2026;17(1):2306. Doi: 10.1038/s41467-026-68873-8
34. El Kababji S, et al. Augmenting insufficiently accruing oncology clinical trials using generative models: validation study. J Med Internet Res. 2025;27:e66821. https://doi.org/10.2196/66821
35. Chen P, Zhang Y, Li W, et al. Predicting hospital readmission using deep learning: a multicenter study. JAMA Netw Open. 2024;7(5):e2412345. https://doi.org/10.1001/jamanetworkopen.2024.12345

Global Academic Nursing Journal