Artificial Intelligence in Diagnostic and Therapeutic Interventions: A Systematic Review of Randomized Controlled Trials

Oscar Jimenez-Flores[1], Sandra Pajares-Centeno[2], Oscar Mejia-Sanchez[3], Rodrigo Flores-Palacios[4]
Escuela de Posgrado Newman, Tacna, Perú[1, 2, 3]
Universidad Nacional Jorge Basadre Grohmann, Tacna, Perú[4]

***Abstract*—Artificial intelligence (AI) is increasingly being integrated into diagnostic and therapeutic interventions, offering potential advantages in accuracy, efficiency, and clinical decision-making compared to conventional methods. This systematic review aimed to identify and characterise AI applications assessed in randomised controlled trials (RCTs), and to synthesise the reported clinical outcomes in comparison with standard approaches. A comprehensive search was conducted in PubMed, Scopus, and Web of Science for articles published between 2015 and June 2024. Eligible studies included randomised controlled trials involving patients with various medical conditions who received diagnostic or therapeutic interventions supported by AI technologies. Comparators included conventional diagnostic or treatment methods, placebo, or standard care. Two reviewers independently screened the studies, extracted data, and assessed risk of bias using the Cochrane RoB 1 tool. A total of 13 trials involving 10,566 participants met the inclusion criteria, spanning a range of medical specialties including gastroenterology, dermatology, radiology, oncology, neurology, and ophthalmology. While several trials reported improvements in diagnostic accuracy, treatment planning, or procedural efficiency, other studies showed inconsistent or limited benefits, highlighting the variability in outcomes depending on the clinical context and type of AI application. This review offers an updated synthesis of AI-based clinical interventions evaluated through randomised controlled trials and emphasises the need for further research to validate these tools, standardise their implementation, and assess their broader impact as health technology in modern healthcare systems.**

Keywords*—AI applications; diagnostic interventions, therapeutic interventions; randomised controlled trials; health technology

I. Introduction

In recent years, the field of artificial intelligence (AI) has experienced exponential growth, driving significant advancements across various disciplines, including medicine [1]. Medical applications of AI have begun to profoundly transform clinical practice and biomedical research, offering new possibilities to enhance diagnostic accuracy, personalize treatments, and optimize clinical data management [2–4]. These emerging technologies range from AI-assisted diagnostic systems to advanced predictive models and telemedicine applications, all aimed at improving the quality and efficiency of healthcare [5–7].

The implementation of AI in the medical field offers unprecedented opportunities for the personalization of treatments, dynamically adapting to the specific needs of each patient [8]. This ability not only improves clinical outcomes but also optimizes the utilization of medical resources, which is crucial in a globally strained healthcare environment [9]. AI-based technologies facilitate continuous monitoring and patient follow-up, providing innovative tools for early disease detection and proactive management of chronic conditions [10]. By analyzing large volumes of data in real time, AI systems can alert healthcare professionals to potential complications before they clinically manifest, enabling faster and more effective interventions [11].

Despite these significant advancements, the integration of AI into medical applications presents crucial challenges [12], particularly in terms of ethical considerations, privacy, and variability in outcomes [13]. The collection and analysis of extensive medical data sets raise serious concerns about patient confidentiality and data integrity [14]. Moreover, the effectiveness of AI-based interventions can vary significantly depending on the specific medical condition and clinical context, which raises questions about their reliability and comparability to conventional methods [15]. To ensure the safe and equitable use of AI, it is imperative to establish clear and rigorous policies that address these issues, including algorithmic biases and disparities in implementation across different healthcare settings [16].

Furthermore, the successful integration of AI in medicine requires that healthcare professionals be adequately prepared to utilize these advanced tools [17]. This necessitates the development of comprehensive educational and training programs that equip medical staff with the skills needed to leverage AI technologies fully [18]. Understanding the ethical principles and best practices for integrating AI into daily clinical practice is essential [19], as is the ability to interpret and apply AI-generated recommendations while recognizing their capabilities and limitations [20].

Given the rapid evolution and proliferation of AI applications in medicine, there is a clear need for a systematic review that explores how these technologies are being applied in diagnostic and therapeutic contexts [21]. While prior reviews have provided broad overviews of AI in healthcare, they have often lacked a targeted focus on AI-based interventions evaluated through randomised controlled trials (RCTs).

To address this gap, a systematic review was conducted to identify and characterise diagnostic and therapeutic interventions based on AI, as evaluated in RCTs.

This review offers a structured synthesis of current applications and reported outcomes compared to conventional methods, highlighting areas of clinical promise, heterogeneity across medical fields, and directions for future research and safe implementation [23].

II. Methodology

A. Design and Eligibility Criteria

This systematic review followed the PRISMA 2020 guidelines to ensure comprehensive and transparent reporting [24]. The PRISMA 2020 checklist is included as Supplementary Material 1. The objective of this review was to identify and synthesise randomised controlled trials (RCTs) that evaluated diagnostic or therapeutic interventions based on artificial intelligence (AI), compared to conventional approaches. The research question was structured using the PICO framework as follows:

- Population (P): Patients with specific medical conditions receiving AI-based diagnostic or therapeutic interventions.
- Intervention (I): AI-based diagnostic or therapeutic interventions.
- Comparison (C): Conventional diagnostic or therapeutic methods, placebo, or standard care.
- Outcome (O): Diagnostic accuracy, treatment efficacy, precision, and/or side effects.

1) Inclusion criteria were: (i) RCTs evaluating diagnostic or therapeutic interventions supported by AI; (ii) studies including patients of any age, sex, or clinical condition; (iii) trials comparing AI-based approaches with conventional diagnostic or treatment methods, placebo, or standard care; and (iv) studies reporting quantitative clinical outcomes.

2) Exclusion criteria were: (i) non-randomised studies (e.g., observational, cross-sectional, cohort, or case-control designs); (ii) RCTs without a comparator group using conventional methods; (iii) studies not reporting relevant clinical outcomes; (iv) duplicate publications; (v) editorials, conference abstracts, letters, reviews, or protocols.

The review protocol was prospectively registered in the International Prospective Register of Systematic Reviews (PROSPERO) under registration number CRD42024583396, on 26 August 2024.

B. Information Sources and Search Strategy

A systematic search was conducted across PubMed, Scopus, and Web of Science for studies published between January 2015 and June 2024. The search strategy combined controlled vocabulary and free-text terms related to artificial intelligence and clinical interventions. Core search terms included “artificial intelligence”, “machine learning”, “deep learning”, “neural networks”, “natural language processing” combined with clinical concepts such as “diagnosis”, “treatment”, “therapeutic”, “patient care”, and study design filters (“randomised controlled trial”, “clinical trial”).

Specific exclusion terms were also applied to remove non-relevant designs, such as reviews, meta-analyses, case reports, editorials, and observational studies. Only original articles classified as clinical trials were retained. No restrictions were applied regarding patient population, but searches were limited to articles published in English or Spanish.

The detailed search strategies for each database, including Boolean operators, limits, and filters, are provided in Supplementary Material 2. A total of 861 records were retrieved from PubMed, 1,359 from Scopus, and 306 from Web of Science.

C. Selection and Data Collection Process

All identified references were imported into Rayyan [25, 26]. where duplicates were automatically removed. Two reviewers (SPC, OMS) independently screened titles and abstracts, followed by full-text review of potentially eligible articles. Any disagreements were resolved through discussion with a third reviewer (RFP).

From the 2,526 records initially retrieved, 2,096 remained after duplicate removal. After screening 137 titles and abstracts, 91 full-text articles were assessed for eligibility. Of these, 73 were excluded for not meeting the inclusion criteria (detailed in Supplementary Material 3). Ultimately, 13 RCTs were included in the final analysis (Fig. 1).

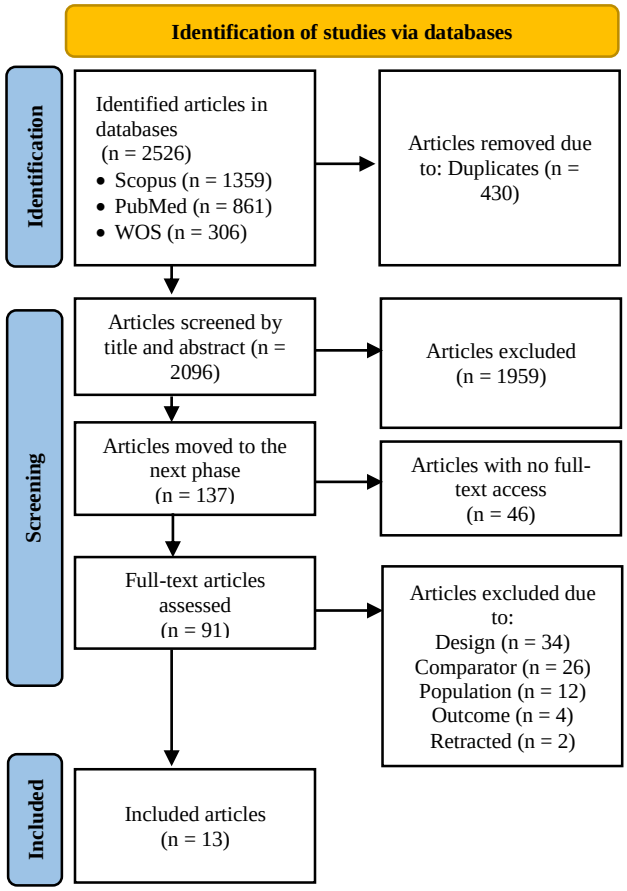

Fig. 1. Flow diagram of studies inclusion.

Data extraction was conducted independently by two reviewers using a pre-piloted form. Discrepancies were resolved by consensus.

D. Variables Collected

The primary outcomes of interest were diagnostic accuracy, treatment efficacy, and precision. Secondary outcomes included side effects and other clinically relevant metrics reported by the included studies. Each outcome was defined operationally according to the measures used in the original trials.

E. Risk of Bias Assessment

The risk of bias of the included RCTs was independently assessed by two reviewers (SPC, RFP) using the Cochrane RoB1 tool [27]. This tool evaluates six domains: random sequence generation, allocation concealment, blinding of participants, personnel and outcome assessors, incomplete outcome data, selective reporting, and other potential sources of bias. Disagreements were resolved with input from a third reviewer (OJF).

The overall risk of bias for each trial was categorised as low, unclear, or high, based on the assessment across domains.

III. Results

A. Study Selection

The breakdown of results by database followed the search strategy based on the following string: "(artificial intelligence" OR "AI" OR "machine learning" OR "deep learning" OR "neural network*" OR "natural language processing") AND ("medicine" OR "medical" OR "healthcare" OR "clinical" OR "health" OR "patient care" OR "diagnosis" OR "treatment" OR "biomedicine")". A total of 2,526 open-access records were identified, considering documents in English and Spanish from the years 2015 to 2024 for Scopus, WoS, and PubMed.

After removing duplicates, 2,096 records were screened by title and abstract. Following this initial screening, 137 articles were deemed potentially relevant and selected for full-text review. However, full text was not accessible for 46 of these articles. The remaining 91 articles underwent thorough full-text evaluation, leading to the exclusion of 73 studies due to various reasons, such as not meeting the inclusion criteria or lacking sufficient data for assessment. Ultimately, 13 randomised controlled trials (RCTs) with a combined total of 10,566 participants were included in the systematic review (Fig. 1).

B. Risk of Bias

To enhance the rigor of the study, RoB 1 (Cochrane Tool) was employed for RCT. This comprehensive analysis covered various types of bias that could compromise the internal validity of the RCTs included in this review, including random sequence generation and allocation concealment to identify potential selection biases. Additionally, participant, personnel, and assessor blinding were examined to address performance and detection biases. Incomplete outcome data and selective reporting were also reviewed, as they might reflect attrition and reporting biases, respectively. Considering these factors, along with other potential biases, resulted in a thorough evaluation aimed at improving the study's quality and its conclusions, as detailed in Fig. 2. Table I shows characteristics of included randomized controlled trials.

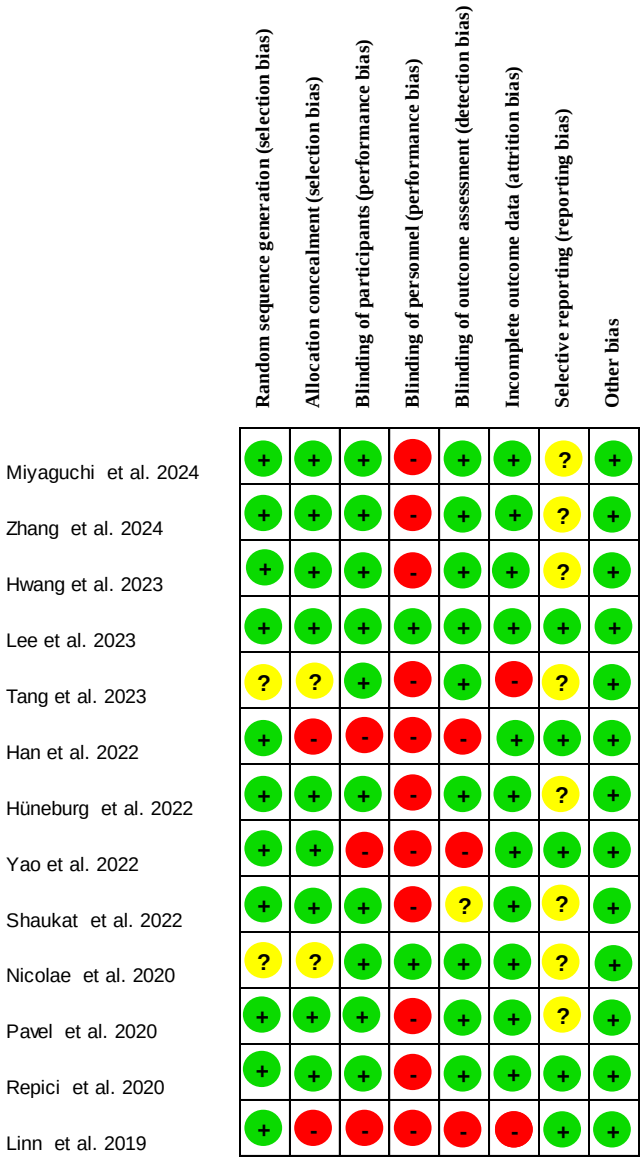

Fig. 2. Risk of bias summary across included studies. Note. Green represents low risk of bias; yellow represents unclear risk of bias; red represents high risk of bias.

TABLE I. CHARACTERISTICS OF INCLUDED RANDOMIZED CONTROLLED TRIALS

Author (year)	Medical Area	Population / Sample size	Intervention	Intervention Type	Comparator	AI Technology Used	Main Conclusion
Gastroenterology							
Miyaguchi (2024) [28]	Gastroenterology	Adults ≥20 undergoing colonoscopy / 800 (I:400, C:400)	Linked-color imaging colonoscopy with AI assistance	Diagnostic	Linked-color imaging colonoscopy without AI	CADEYE (Fujifilm EW10-EC02)	AI-assisted colonoscopy improved adenoma detection rate (ADR).
Zhang (2024) [29]	Gastroenterology	Adults 18–75 requiring colonoscopy / 1360 (I:680, C:680)	White-light colonoscopy with CADe	Diagnostic	White-light colonoscopy without CADe	EndoAngel (Wuhan ENDOANGEL)	Improved sensitivity and reduced polyp miss rate.
Tang (2023) [32]	Gastroenterology	Adults with pancreatic masses / 44 (I:21, C:23)	CH-EUS MASTER-guided EUS-FNA	Diagnostic	Conventional EUS-FNA	CH-EUS MASTER (AI-guided contrast-enhanced EUS)	Improved diagnostic accuracy and yield.
Hüneburg (2022) [34]	Gastroenterology	Adults ≥18 with Lynch syndrome / 96 (I:50, C:46)	AI-assisted colonoscopy	Diagnostic	High-definition white-light endoscopy (HD-WLE)	CAD-EYE (Fujifilm EW10-EC02)	Improved flat adenoma detection; overall ADR similar.
Yao (2022) [35]	Gastroenterology	Adults >18 undergoing colonoscopy / 1076 (I:805, C:271)	CADe, CAQ, and combined systems	Diagnostic	Colonoscopy without AI	CADe (YOLO V3), CAQ (AI quality control)	Combined AI improved ADR; CAQ was more effective than CADe alone.
Shaukat (2022) [36]	Gastroenterology	Adults ≥40 / 1369 (I:682, C:677)	Colonoscopy with CADe	Diagnostic	Standard colonoscopy	CADe (Computer-Aided Detection)	Increased ADR and adenomas per colonoscopy (APC).
Repici (2020) [39]	Gastroenterology	Adults 40–80 undergoing colonoscopy / 685 (I:341, C:340)	Colonoscopy with CADe	Diagnostic	Colonoscopy without CADe	GI-Genius (Medtronic)	Increased ADR without prolonging procedure time.
Radiology							
Hwang (2023) [30]	Radiology	Adults ≥19 with respiratory symptoms / 3576 (I:1761, C:1815)	Chest radiograph interpretation with AI-CAD	Diagnostic	Standard chest radiograph interpretation	Lunit INSIGHT CXR v2.0.2.0	No improvement in sensitivity or false positives.
Lee (2023) [31]	Radiology	Adults ≥20 undergoing chest X-ray / 329 (I:166, C:163)	AI-assisted interpretation of thoracic abnormalities	Diagnostic	Standard interpretation without AI	Lunit INSIGHT CXR v2.0.3.0	Improved non-radiologist diagnostic accuracy.
Dermatology							
Han (2022) [33]	Dermatology	Adults >19 with suspicious lesions / 576 (I:305, C:292)	Real-time AI-assisted diagnosis	Diagnostic	Diagnosis without AI	Model Dermatology (Multiclass AI)	Significantly improved accuracy, especially in non-experts.
Oncology							
Nicolae (2020) [37]	Oncology	Prostate cancer patients / 41 (I:21, C:20)	Machine learning-based treatment planning	Therapeutic	Manual planning	PIPA (Prostate Implant Planning Algorithm)	Comparable dosimetry, with time and efficiency gains.
Neurology							
Pavel (2020) [38]	Neurology	Neonates (36–44 weeks) / 264 (I:132, C:132)	EEG with ANSeR for seizure detection	Diagnostic	Conventional EEG monitoring	ANSeR (Algorithm for Neonatal Seizure Recognition)	Better seizure-hour identification; no difference at patient level.
Ophthalmology							
Linn (2019) [40]	Ophthalmology	Pediatric patients ≤14 years / 350 (I:350, C:350)	Diagnosis using CC-Cruiser	Diagnostic	Diagnosis by senior consultants	CC-Cruiser (Ophthalmic AI platform)	Lower diagnostic accuracy, but faster with higher satisfaction.

Note. I: Intervention group; C: Control group; AI: Artificial Intelligence; CADe: Computer-Aided Detection; CAQ: Computer-Aided Quality Improvement; CAD: Computer-Aided Diagnosis; ADR: Adenoma Detection Rate; ED: Emergency Department; CR: Chest Radiography; CXR: Chest X-ray; EUS-FNA: Endoscopic Ultrasound-Guided Fine-Needle Aspiration; CH-EUS MASTER: AI-assisted Contrast-Enhanced Harmonic Endoscopic Ultrasound system; HD-WLE: High-Definition White-Light Endoscopy; PIPA: Prostate Implant Planning Algorithm; ANSeR: Algorithm for Neonatal Seizure Recognition; EEG: Electroencephalogram.

IV. DISCUSSION

This systematic review synthesised the current landscape of diagnostic and therapeutic interventions based on artificial intelligence (AI) as evaluated in randomised controlled trials (RCTs). The included studies span multiple medical specialties, offering a structured overview of how AI technologies are being implemented in clinical settings and what outcomes have been reported in comparison with conventional approaches.

In gastroenterology, the integration of AI into endoscopic practice has consistently demonstrated clinically meaningful improvements. Multiple studies [28] and [39] confirmed that AI-assisted colonoscopy significantly increased adenoma detection rates (ADR), a well-established surrogate marker for colorectal cancer prevention. Complementarily, the use of Computer-Aided Detection (CADe) systems [29], [36] enhanced sensitivity for polyp detection without increasing unnecessary resections of non-neoplastic lesions, reinforcing their safety profile [22]. Importantly, AI demonstrated added value in specific high-risk populations: study [34] reported superior detection of flat adenomas in patients with Lynch syndrome, which are frequently overlooked by conventional methods. Beyond colonoscopy, AI has also optimized diagnostic yield in other domains, as evidenced by the CH-EUS MASTER system [32], which improved accuracy during EUS-FNA of pancreatic masses. Furthermore, the combined use of AI-driven detection and quality control systems [35] suggested a synergistic effect on real-time ADR, although this benefit was not replicated when each system was applied independently. Collectively, these findings not only highlight AI's capacity to enhance early lesion detection but also illustrate its potential to reshape preventive strategies in gastrointestinal oncology [41]. By enabling real-time visual analysis and identification of subtle mucosal changes often imperceptible to human observers, gastroenterology emerges as one of the most advanced fields for clinical AI implementation [42].

In radiology, the findings were more heterogeneous. Study [30] reported no significant improvement in sensitivity or false-positive rates when chest radiographs were interpreted with AI assistance for the diagnosis of acute thoracic diseases. Conversely, study [31] demonstrated that an AI-assisted diagnostic protocol based on deep learning significantly enhanced the performance of non-radiologist physicians in detecting pulmonary abnormalities. These discrepant results may be explained by variations in the maturity of the algorithms, the expertise of end-users, and the clinical contexts in which the tools were applied. Collectively, they highlight the pressing need for validation studies that account for both technological development and real-world implementation challenges.

In dermatology, study [33] showed that real-time AI assistance substantially improved diagnostic accuracy among non-expert physicians evaluating suspicious skin lesions. Interestingly, the greatest benefit was seen among trainees with limited dermatological experience, suggesting that AI functions not merely as a diagnostic adjunct but also as a pedagogical tool. This points to an underexplored opportunity for AI to strengthen medical education by supporting clinical reasoning and decision-making in early-career professionals.

In oncology, study [37] showed that machine learning-based treatment planning for prostate cancer, specifically using the Prostate Implant Planning Algorithm (PIPA), improved time efficiency and streamlined workflow without compromising dosimetric quality compared to manual planning. These results highlight AI's potential to optimize operational processes while maintaining treatment precision.

In neurology, study [38] assessed the use of the ANSeR algorithm for neonatal seizure recognition. Although AI improved the identification of seizure hours, it did not enhance detection at the patient level when compared with conventional EEG. This suggests that, while AI can assist in data interpretation, its current clinical contribution to neonatal seizure diagnostics remains limited.

In ophthalmology, study [40] evaluated the CC-Cruiser system and found that, although it facilitated faster diagnostic processes and improved patient satisfaction, its diagnostic accuracy and therapeutic decision-making were inferior to those of senior consultants. These results underscore the potential of AI to enhance efficiency and patient experience, but they also reveal critical limitations in replicating expert-level judgement. This suggests that, at present, AI should be regarded as an adjunctive tool rather than a replacement in complex clinical decision-making, highlighting the need for further refinement and validation before broader clinical integration.

Overall, the included studies demonstrate a broad spectrum of AI applications across medical disciplines. However, the representation across specialties is uneven**,** with gastroenterology dominating the evidence base, while areas such as neurology [38] and ophthalmology [40] remain underrepresented. This limits the generalizability of findings to less-studied fields and highlights the need for more diverse and comprehensive RCTs in AI research. Additionally, the majority of included studies were conducted in technologically advanced, well-resourced hospital settings. Consequently, their applicability to low-resource or primary care environments remains uncertain.

The certainty of evidence also varied. In gastroenterology, robust findings from multiple high-powered RCTs (e.g., [28], [29], [39]) provide strong support for the use of AI. In contrast, the mixed results in radiology [30], [31] and limited data in oncology and neurology reduce confidence in the broader applicability of these technologies. The variability in AI system types, clinical indications, and user expertise further complicates the ability to draw uniform conclusions across studies.

Despite its strengths, this systematic review has several limitations. First, although the search strategy was extensive, it is possible that relevant studies were missed due to language restrictions or the stringent inclusion criteria. Second, heterogeneity across AI systems, patient populations, and outcome measures reduced the comparability of results and precluded meta-analysis. However, this same heterogeneity reflects the diversity and innovation within the field, illustrating the breadth of AI's applicability to various clinical challenges.

Notably, this review adheres to rigorous methodological standards, including AMSTAR-2 compliance, a predefined protocol, and a structured risk of bias assessment, which

collectively reinforce the validity and transparency of the findings. By restricting the review to randomised controlled trials, the study focused on high-quality evidence, ensuring that the conclusions are grounded in rigorous experimental data.

In addition to evaluating clinical performance, this review indirectly reflects the growing sophistication and practical deployment of machine learning algorithms in healthcare settings. Most of the AI-based interventions included in the analysed RCTs involve deep learning models, computer vision tools, or natural language processing systems integrated into real-time clinical workflows. These findings not only demonstrate the clinical utility of AI but also highlight the importance of interdisciplinary collaboration between clinicians, data scientists, and engineers. Future studies should document more precisely the algorithmic architecture, training datasets, performance metrics, and update protocols of the AI systems being evaluated, to ensure reproducibility and transparency across both medical and computational communities.

Collectively, these findings illustrate both the promise and the fragility of the current evidence base. Gastroenterology emerges as a model of successful integration of AI into routine practice, whereas other specialties remain in earlier stages of exploration, with evidence that is either inconsistent or limited in scope. This uneven development highlights the urgent need for methodological standardisation, greater representation of diverse healthcare contexts, and longitudinal validation studies to assess sustainability of outcomes over time. Addressing these gaps will not only advance the safe and equitable deployment of AI across medical disciplines but also ensure that its benefits extend beyond technologically advanced centres to the broader spectrum of global healthcare.

Future research should prioritise underexplored medical specialties, assess AI interventions across diverse healthcare settings—including low-resource and primary care environments—and examine the impact of user training and experience on clinical outcomes. Longitudinal and multicentre studies are encouraged to evaluate the sustainability and generalisability of AI-assisted interventions. Moreover, developing standardised methodological frameworks and reporting guidelines will facilitate comparability, reproducibility, and evidence-informed adoption of AI technologies across clinical disciplines.

V. Conclusion

This systematic review identified and synthesised randomised controlled trials evaluating AI-based diagnostic and therapeutic interventions compared to conventional methods. The findings highlight that AI technologies can improve diagnostic accuracy and clinical performance, particularly in fields such as gastroenterology and dermatology. However, the variability of outcomes observed in specialties such as radiology, neurology and ophthalmology suggest that the clinical utility of AI is highly context dependent and influenced by the nature of the task, technological design and user expertise.

These results underscore the importance of a critical and context-sensitive approach to the integration of AI in healthcare. While promising results support its implementation in specific domains, the observed heterogeneity indicates that AI is not a universal solution. A deeper understanding of the conditions under which AI adds the most value is essential, considering clinical context, provider training, technological maturity and patient characteristics.

To advance responsible and effective integration of AI into clinical practice, future research should expand evaluation across diverse healthcare settings and populations. The development of standardised protocols and implementation guidelines is needed to ensure consistent and evidence-informed use.

Additionally, studies should examine the impact of user training and experience on AI outcomes, and ensure that the deployment of these technologies promotes equity in access and benefit across all socioeconomic and geographic contexts. In summary, this systematic review provides an updated and evidence-based overview of AI applications in randomised controlled trials, offering a critical foundation to guide future research and the responsible integration of AI into clinical practice.

References

[1] Silcox, C., Zimlichmann, E., Huber, K., Rowen, N., Saunders, R., McClellan, M., Kahn, C.N., Salzberg, C.A., Bates, D.W.: The potential for artificial intelligence to transform healthcare: perspectives from international health leaders. npj Digit. Med. 7, 88, 2024. https://doi.org/10.1038/s41746-024-01097-6.

[2] Da Silva, R.G.L.: The advancement of artificial intelligence in biomedical research and health innovation: challenges and opportunities in emerging economies. Global Health. 20, 44, 2024. https://doi.org/10.1186/s12992-024-01049-5.

[3] Gao, S., Xu, Z., Kang, W., Lv, X., Chu, N., Xu, S., Hou, D.: Artificial intelligence-driven computer aided diagnosis system provides similar diagnosis value compared with doctors' evaluation in lung cancer screening. BMC Med Imaging. 24, 141, 2024. https://doi.org/10.1186/s12880-024-01288-3.

[4] Knudsen, J.E., Ghaffar, U., Ma, R., Hung, A.J.: Clinical applications of artificial intelligence in robotic surgery. J Robotic Surg. 18, 102, 2024. https://doi.org/10.1007/s11701-024-01867-0.

[5] Bowman, M.A., Seehusen, D.A., Britz, J., Ledford, C.J.W.: Research to Improve Clinical Care in Family Medicine: Big Data, Telehealth, Artificial Intelligence, and More. J Am Board Fam Med. 37, 161–164, 2024. https://doi.org/10.3122/jabfm.2024.240050R0.

[6] Green, M.: Could artificial intelligence improve patient care and physician workload? Can Fam Physician. 70, 216–216, 2024. https://doi.org/10.46747/cfp.7003216.

[7] Zhang, L., Wong, C., Li, Y., Huang, T., Wang, J., Lin, C.: Artificial intelligence assisted diagnosis of early tc markers and its application. Discov Onc. 15, 172, 2024. https://doi.org/10.1007/s12672-024-01017-w.

[8] Chen, A., Wang, C., Zhang, X.: Reflection on the equitable attribution of responsibility for artificial intelligence-assisted diagnosis and treatment decisions. Intelligent Medicine. 3, 139–143, 2023. https://doi.org/10.1016/j.imed.2022.04.002.

[9] Cohen, R.Y., Kovacheva, V.P.: A Methodology for a Scalable, Collaborative, and Resource-Efficient Platform, MERLIN, to Facilitate Healthcare AI Research. IEEE J. Biomed. Health Inform. 27, 3014–3025, 2023. https://doi.org/10.1109/JBHI.2023.3259395.

[10] A, A., Dahan, F., Alroobaea, R., Alghamdi, Wael.Y., Mustafa Khaja Mohammed, Hajjej, F., Deema Mohammed Alsekait, Raahemifar, K.: A smart IoMT based architecture for E-healthcare patient monitoring system using artificial intelligence algorithms. Front. Physiol. 14, 1125952, 2023. https://doi.org/10.3389/fphys.2023.1125952.

[11] Akyüz, K., Cano Abadía, M., Goisauf, M., Mayrhofer, M.Th.: Unlocking the potential of big data and AI in medicine: insights from biobanking.

Front. Med. 11, 1336588, 2024. https://doi.org/10.3389/fmed.2024.1336588.

[12] Witkowski, K., Okhai, R., Neely, S.R.: Public perceptions of artificial intelligence in healthcare: ethical concerns and opportunities for patient-centered care. BMC Med Ethics. 25, 74, 2024. https://doi.org/10.1186/s12910-024-01066-4.

[13] Palmer, A., Schwan, D.: More Process, Less Principles: The Ethics of Deploying AI and Robotics in Medicine. Camb Q Healthc Ethics. 33, 121–134, 2024. https://doi.org/10.1017/S0963180123000087.

[14] Huang, S.-H., Peng, B.-R., Lin, C.-S., Tsai, H.-C., Pan, L.-F., Pan, L.-K.: Potential risk quantification from multiple biological factors via the inverse problem algorithm as an artificial intelligence tool in clinical diagnosis. THC. 31, 69–79, 2023. https://doi.org/10.3233/THC-236008.

[15] Aquino, Y.S.J., Carter, S.M., Houssami, N., Braunack-Mayer, A., Win, K.T., Degeling, C., Wang, L., Rogers, W.A.: Practical, epistemic and normative implications of algorithmic bias in healthcare artificial intelligence: a qualitative study of multidisciplinary expert perspectives. J Med Ethics. jme-2022-108850, 2023. https://doi.org/10.1136/jme-2022-108850.

[16] Yoo, J., Hur, S., Hwang, W., Cha, W.C.: Healthcare Professionals' Expectations of Medical Artificial Intelligence and Strategies for its Clinical Implementation: A Qualitative Study. Healthc Inform Res. 29, 64–74, 2023. https://doi.org/10.4258/hir.2023.29.1.64.

[17] Eskandarani, R., Almuhainy, A., Alzahrani, A.: Creating a master training rotation schedule for emergency medicine residents and challenges in using artificial intelligence. Int J Emerg Med. 17, 84, 2024. https://doi.org/10.1186/s12245-024-00657-7.

[18] Krajcer, Z.: Artificial Intelligence for Education, Proctoring, and Credentialing in Cardiovascular Medicine. Texas Heart Institute Journal. 49, e217572, 2022. https://doi.org/10.14503/THIJ-21-7572.

[19] Bekbolatova, M., Mayer, J., Ong, C.W., Toma, M.: Transformative Potential of AI in Healthcare: Definitions, Applications, and Navigating the Ethical Landscape and Public Perspectives. Healthcare. 12, 125, 2024. https://doi.org/10.3390/healthcare12020125.

[20] Oniani, D., Hilsman, J., Peng, Y., Poropatich, R.K., Pamplin, J.C., Legault, G.L., Wang, Y.: Adopting and expanding ethical principles for generative artificial intelligence from military to healthcare. npj Digit. Med. 6, 225, 2023. https://doi.org/10.1038/s41746-023-00965-x.

[21] Ranjbar, A., Mork, E., Ravn, J., Brøgger, H., Myrseth, P., Østrem, H.P., Hallock, H.: Managing Risk and Quality of AI in Healthcare: Are Hospitals Ready for Implementation? RMHP. Volume 17, 877–882, 2024. https://doi.org/10.2147/RMHP.S452337.

[22] Macrae, C.: Managing risk and resilience in autonomous and intelligent systems: Exploring safety in the development, deployment, and use of artificial intelligence in healthcare. Risk Analysis. risa.14273 ,2024. https://doi.org/10.1111/risa.14273.

[23] Alami, H., Lehoux, P., Papoutsi, C., Shaw, S.E., Fleet, R., Fortin, J.-P.: Understanding the integration of artificial intelligence in healthcare organisations and systems through the NASSS framework: a qualitative study in a leading Canadian academic centre. BMC Health Serv Res. 24, 701, 2024. https://doi.org/10.1186/s12913-024-11112-x.

[24] 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., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D.: The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. n71 ,2021. https://doi.org/10.1136/bmj.n71.

[25] Valizadeh, A., Moassefi, M., Nakhostin-Ansari, A., Hosseini Asl, S.H., Saghab Torbati, M., Aghajani, R., Maleki Ghorbani, Z., Faghani, S.: Abstract screening using the automated tool Rayyan: results of effectiveness in three diagnostic test accuracy systematic reviews. BMC Med Res Methodol. 22, 160, 2022. https://doi.org/10.1186/s12874-022-01631-8.

[26] Yu, F., Liu, C., Sharmin, S.: Performance, Usability, and User Experience of Rayyan for Systematic Reviews. Proceedings of the Association for Information Science and Technology. 59, 843–844, 2022. https://doi.org/10.1002/pra2.745.

[27] Higgins, J.P., Savović, J., Page, M.J., Elbers, R.G., Sterne, J.A.: Assessing risk of bias in a randomized trial. In: Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., and Welch, V.A. (eds.) Cochrane Handbook for Systematic Reviews of Interventions. pp. 205–228. Wiley .2019. https://doi.org/10.1002/9781119536604.ch8.

[28] Miyaguchi, K., Tsuzuki, Y., Hirooka, N., Matsumoto, H., Ohgo, H., Nakamoto, H., Imaeda, H.: Linked-color imaging with or without artificial intelligence for adenoma detection: a randomized trial. Endoscopy. 56, 376–383, 2024. https://doi.org/10.1055/a-2239-8145.

[29] Zhang, H., Wu, Q., Sun, J., Wang, J., Zhou, L., Cai, W., Zou, D.: A computer-aided system improves the performance of endoscopists in detecting colorectal polyps: a multi-center, randomized controlled trial. Front. Med. 10, 1341259, 2024. https://doi.org/10.3389/fmed.2023.1341259.

[30] Hwang, E.J., Goo, J.M., Nam, J.G., Park, C.M., Hong, K.J., Kim, K.H.: Conventional Versus Artificial Intelligence-Assisted Interpretation of Chest Radiographs in Patients With Acute Respiratory Symptoms in Emergency Department: A Pragmatic Randomized Clinical Trial. Korean J Radiol. 24, 259, 2023. https://doi.org/10.3348/kjr.2022.0651.

[31] Lee, H.W., Jin, K.N., Oh, S., Kang, S.-Y., Lee, S.M., Jeong, I.B., Son, J.W., Han, J.H., Heo, E.Y., Lee, J.G., Kim, Y.J., Kim, E.Y., Cho, Y.J.: Artificial Intelligence Solution for Chest Radiographs in Respiratory Outpatient Clinics: Multicenter Prospective Randomized Clinical Trial. Annals ATS. 20, 660–667, 2023. https://doi.org/10.1513/AnnalsATS.202206-481OC.

[32] Tang, A., Tian, L., Gao, K., Liu, R., Hu, S., Liu, J., Xu, J., Fu, T., Zhang, Z., Wang, W., Zeng, L., Qu, W., Dai, Y., Hou, R., Tang, S., Wang, X.: Contrast - enhanced harmonic endoscopic ultrasound (CH - EUS) MASTER: A novel deep learning - based system in pancreatic mass diagnosis. Cancer Medicine. 12, 7962–7973, 2023. https://doi.org/10.1002/cam4.5578.

[33] Han, S.S., Kim, Y.J., Moon, I.J., Jung, J.M., Lee, M.Y., Lee, W.J., Won, C.H., Lee, M.W., Kim, S.H., Navarrete-Dechent, C., Chang, S.E.: Evaluation of Artificial Intelligence–Assisted Diagnosis of Skin Neoplasms: A Single-Center, Paralleled, Unmasked, Randomized Controlled Trial. Journal of Investigative Dermatology. 142, 2353-2362.e2. 2022. https://doi.org/10.1016/j.jid.2022.02.003.

[34] Hüneburg, R., Bucksch, K., Schmeißer, F., Heling, D., Marwitz, T., Aretz, S., Kaczmarek, D.J., Kristiansen, G., Hommerding, O., Strassburg, C.P., Engel, C., Nattermann, J.: Real - time use of artificial intelligence (CADEYE) in colorectal cancer surveillance of patients with Lynch syndrome—A randomized controlled pilot trial (CADLY). UEG Journal. 11, 60–68, 2023. https://doi.org/10.1002/ueg2.12354.

[35] Yao, L., Zhang, L., Liu, J., Zhou, W., He, C., Zhang, J., Wu, L., Wang, H., Xu, Y., Gong, D., Xu, M., Li, X., Bai, Y., Gong, R., Sharma, P., Yu, H.: Effect of an artificial intelligence-based quality improvement system on efficacy of a computer-aided detection system in colonoscopy: a four-group parallel study. Endoscopy. 54, 757–768, 2022. https://doi.org/10.1055/a-1706-6174.

[36] Shaukat, A., Lichtenstein, D.R., Somers, S.C., Chung, D.C., Perdue, D.G., Gopal, M., Colucci, D.R., Phillips, S.A., Marka, N.A., Church, T.R., Brugge, W.R., Thompson, R., Chehade, R., Loew, B., Downing, J., Vermillion, J., Borges, L., Rajbhandari, R., Schafer, T., Coban, S., Richter, J., Carolan, P., Colizzo, F., Jeong, T., DelSignore, M., Asher, S., McCabe, R., Van Handel, D., Cinnor, B., Mitlyng, B., Sherman, C., Feldshon, S.D., Lounsbury, A., Thompson, A., Duggirala, A., Davies, I., Huang, C., Bliss, C., Mohanty, A., Sina, O., Mendez, J., Iwan, A., Stromberg, J., Ng, J., Erisson, L., Golland, P., Wang, D., Wlodkowski, E., Carlin, J., Javia, P., Chavali, N., Wang, A., Little, J., Hunsberger, C.: Computer-Aided Detection Improves Adenomas per Colonoscopy for Screening and Surveillance Colonoscopy: A Randomized Trial. Gastroenterology. 163, 732–741, 2022. https://doi.org/10.1053/j.gastro.2022.05.028.

[37] Nicolae, A., Semple, M., Lu, L., Smith, M., Chung, H., Loblaw, A., Morton, G., Mendez, L.C., Tseng, C.-L., Davidson, M., Ravi, A.: Conventional vs machine learning–based treatment planning in prostate brachytherapy: Results of a Phase I randomized controlled trial. Brachytherapy. 19, 470–476, 2020. https://doi.org/10.1016/j.brachy.2020.03.004.

[38] Pavel, A.M., Rennie, J.M., De Vries, L.S., Blennow, M., Foran, A., Shah, D.K., Pressler, R.M., Kapellou, O., Dempsey, E.M., Mathieson, S.R., Pavlidis, E., Van Huffelen, A.C., Livingstone, V., Toet, M.C., Weeke, L.C., Finder, M., Mitra, S., Murray, D.M., Marnane, W.P., Boylan, G.B.: A machine-learning algorithm for neonatal seizure recognition: a multicentre, randomised, controlled trial. The Lancet Child & Adolescent Health. 4, 740–749, 2020. https://doi.org/10.1016/S2352-4642(20)30239-X.

[39] Repici, A., Badalamenti, M., Maselli, R., Correale, L., Radaelli, F., Rondonotti, E., Ferrara, E., Spadaccini, M., Alkandari, A., Fugazza, A., Anderloni, A., Galtieri, P.A., Pellegatta, G., Carrara, S., Di Leo, M., Craviotto, V., Lamonaca, L., Lorenzetti, R., Andrealli, A., Antonelli, G., Wallace, M., Sharma, P., Rosch, T., Hassan, C.: Efficacy of Real-Time Computer-Aided Detection of Colorectal Neoplasia in a Randomized Trial. Gastroenterology. 159, 512-520.e7. 2020. https://doi.org/10.1053/j.gastro.2020.04.062.

[40] Lin, H., Li, R., Liu, Z., Chen, J., Yang, Y., Chen, H., Lin, Z., Lai, W., Long, E., Wu, X., Lin, D., Zhu, Y., Chen, C., Wu, D., Yu, T., Cao, Q., Li, X., Li, J., Li, W., Wang, J., Yang, M., Hu, H., Zhang, L., Yu, Y., Chen, X., Hu, J., Zhu, K., Jiang, S., Huang, Y., Tan, G., Huang, J., Lin, X., Zhang, X., Luo, L., Liu, Y., Liu, X., Cheng, B., Zheng, D., Wu, M., Chen, W., Liu, Y.: Diagnostic Efficacy and Therapeutic Decision-making Capacity of an Artificial Intelligence Platform for Childhood Cataracts in Eye Clinics: A Multicentre Randomized Controlled Trial. EClinicalMedicine. 9, 52–59, 2019. https://doi.org/10.1016/j.eclinm.2019.03.001.

[41] Uchikov, P., Khalid, U., Vankov, N., Kraeva, M., Kraev, K., Hristov, B., Sandeva, M., Dragusheva, S., Chakarov, D., Petrov, P., Dobreva-Yatseva, B., Novakov, I.: The Role of Artificial Intelligence in the Diagnosis and Treatment of Ulcerative Colitis. Diagnostics. 14, 1004, 2024. https://doi.org/10.3390/diagnostics14101004.

[42] Stafie, C.S., Sufaru, I.-G., Ghiciuc, C.M., Stafie, I.-I., Sufaru, E.-C., Solomon, S.M., Hancianu, M.: Exploring the Intersection of Artificial Intelligence and Clinical Healthcare: A Multidisciplinary Review. Diagnostics. 13, 1995, 2023. https://doi.org/10.3390/diagnostics13121995.