ISSN: 0028-0712

Does AI benefit medical experts in complex real-world diagnoses and treatments?

Karim Hamda Farhat[1, *], Mostafa Ahmed Arafa[1, 2], Abdulaziz Mohammed Althunayan[1], Mohammed Khalid Bin Shuayl[3], Alanoud Abdullah Albekairi[4], Shahad Hamza Almonaie[5], Mashael Marzouq AlShebly[6], Danny Munther Rabah[1,7]

1- The Cancer Research Chair, Surgery Department, College of Medicine, King Saud University, Riyadh, Saudi Arabia.
2- Department of Epidemiology, High Institute of Public Health, Alexandria University, Alexandria, Egypt.
3- General Practitioner, General Directorate of Prisons Health, Ministry of Interior, Riyadh, Saudi Arabia.
4- Pediatric Resident, Prince Sultan Military Medical City, Riyadh, Saudi Arabia.
5- General Practitioner, Alfaisal University, Riyadh, Saudi Arabia.
6- Department of Obstetrics and Gynecology, College of Medicine, King Saud University and King Saud University Medical City, Riyadh, Saudi Arabia.
7- Department of Urology, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia.

*** Corresponding author**

Dr. Karim H. Farhat

Cancer Research Chair, Surgery department,

College of Medicine, King Saud University

Email: kfarhat@ksu.edu.sa

P.O Box: 76047, Riyadh 11922, KSA

Phone: 00966114679748

Abstract:

The integration of artificial intelligence (AI) into healthcare has sparked significant debate regarding its potential to outperform human physicians in the diagnosis and treatment of complex medical cases. While physicians rely on years of education, clinical experience, and intuition, AI platforms leverage vast datasets, machine learning algorithms, and computational power to analyze patterns and predict outcomes. This manuscript explores real-world scenarios in which AI and physicians have been pitted against each other or worked collaboratively, examining whether AI truly holds an advantage in managing complicated cases. By analyzing the strengths and limitations of both approaches, we aimed to provide insights into how AI can complement, rather than replace, physician expertise in modern medicine. This complementary relationship becomes particularly significant when considering the inherent complexity of medical practice, in which human expertise and technological capabilities must work together to address multifaceted clinical challenges. This study investigates the specifics of how AI can augment physicians' capabilities in handling complicated medical cases, setting the stage for detailed exploration.

Keywords: Artificial Intelligence, Medicine, Diagnosis, Complex Cases, Augmented Intelligence

Introduction:

The practice of medicine is inherently complex, requiring a nuanced understanding of biology, patient history, social determinants of health, and evolving scientific knowledge. In recent years, AI platforms such as IBM Watson Health and Google DeepMind have demonstrated remarkable capabilities in processing large volumes of data and identifying patterns that may even elude experienced clinicians [1-3]. However, it remains unclear whether this technological prowess translates into superior diagnostic accuracy and treatment efficacy in complicated cases.

Complicated cases often involve rare diseases, comorbidities, ambiguous symptoms, or conditions with atypical presentations. These scenarios challenge the limits of human cognition and highlight the need for tools to augment clinical decision-making [4]. This study investigates specific real-world examples to assess whether AI platforms possess unique advantages over physicians, and whether their utility lies primarily in collaboration.

Current Applications and Limitations of AI in Healthcare

AI has already made significant inroads into healthcare, particularly in areas that require the analysis of large datasets or the identification of subtle patterns. In radiology, AI algorithms are used to detect anomalies in medical images, such as early-stage cancers in mammograms and lung nodules in chest radiographs [3, 9]. In oncology, AI platforms assist in precision medicine by analyzing genomic data to recommend personalized treatment plans based on a patient’s unique tumor profile [7]. Furthermore, AI has been employed in genomics to identify genetic mutations linked to rare diseases, offering diagnostic insights where traditional methods may fail [5]. These applications demonstrate AI’s potential to enhance diagnostic accuracy and treatment efficacy in complex cases. Despite these promising applications, AI's effectiveness is tempered by several critical limitations.

Limitations of AI in Healthcare

- **Data Quality and Bias**: AI models are only as reliable as the data on which they are trained. If the training data are biased or incomplete, such as when underrepresenting certain populations, the AI's recommendations may be flawed or inequitable [15]. This can lead to disparities in care, particularly among marginalized groups.
- **Interpretability**: Many AI models, especially those based on deep learning, function as "black boxes," providing little insight into how they arrive at their conclusions [9]. This lack of transparency can make it difficult for clinicians to trust or act on AI recommendations, particularly in high-stake medical decision-making.
- **Contextual Understanding**: AI often struggles to incorporate the full context of a patient's situation, such as socioeconomic factors, cultural differences, or individual preferences, which are essential for holistic treatment decisions [14]. For example, AI may recommend treatment without considering whether the patient can afford it or whether it aligns with their values.
- **Ethical and Legal Concerns**: The use of AI in healthcare raises questions about accountability, particularly when AI-driven decisions lead to adverse outcomes in patients. Determining who is responsible, whether the developer, clinician, or institution, remains a complex issue [11]. Furthermore, the handling of sensitive health data by using AI systems raises privacy concerns.

These limitations underscore the need for further investigation into how AI can be effectively integrated into clinical practice, particularly in complex cases in which human judgment and contextual understanding are paramount. The methods employed in this study; case study analysis, literature review, comparative assessment, and expert interviews were chosen to address these gaps:

ISSN: 0028-0712

- **Case Study Analysis**: Provide concrete examples of AI's performance in real-world clinical scenarios, highlighting both its strengths and areas in which human expertise is indispensable.
- **Literature Review**: Synthesizing existing research to identify broader trends and challenges in AI applications, offering a foundation for understanding its current capabilities.
- **Comparative Assessment**: Quantifying AI's diagnostic and therapeutic performance compared to that of physicians in controlled settings, offering objective measures of its advantages and limitations.
- **Expert Interviews**: Capturing qualitative insights from clinicians and AI developers on the practical, ethical, and technical challenges of integrating AI into healthcare workflows.

By employing this multifaceted approach, this study aimed to explore not only whether AI can outperform physicians but also how AI and human expertise can work together to improve patient outcomes in complex medical cases.

Methods:

A mixed-method approach was employed to evaluate whether AI platforms have an advantage over physicians in diagnosing and treating complicated medical cases. This approach integrates qualitative and quantitative methods to provide a comprehensive assessment of AI's capabilities relative to human clinicians.

The methods of Case Study Analysis, Literature Review, Comparative Assessment, and Expert Interviews were selected to collectively address the research question by examining real-world applications, synthesizing existing evidence, benchmarking performance, and capturing expert insights. Each method is described below, highlighting its contribution to the study objective.

Case Study Analysis

To understand how AI platforms perform in real-world clinical scenarios, a detailed analysis of case studies was conducted. This method was chosen to provide concrete examples of AI's diagnostic and therapeutic contributions of AI in complex medical contexts, particularly for rare diseases, oncology, and emergency medicine. Relevant case studies were identified from peer-reviewed literature and clinical reports and were selected based on their focus on AI involvement in high-stake scenarios [5]. Each case was evaluated using the following framework:

- **Input Data**: The type and quality of data provided to the AI (e.g., genomic data, imaging results, and electronic health records).
- **AI Contribution**: Specific insights or recommendations generated by the AI platform.
- **Physician Role**: How clinicians interpret and act on AI outputs.
- **Outcome**: The clinical outcome for the patient and any lessons learned about the strengths or limitations of AI in that context [6].

This analysis grounded the study in practical applications, offering insights into AI's real-world effectiveness and its interaction with physician decision-making.

Literature Review

To contextualize the case study findings within the broader research landscape, a systematic literature review was performed. This method aimed to synthesize existing evidence on AI’s role in healthcare, identifying common themes, advantages, and challenges in comparison to physicians. A systematic search was performed across PubMed, Scopus, and Google Scholar using keywords such as "artificial intelligence in healthcare," "AI vs. physicians," "complicated medical cases," and "real-world AI applications." Studies published between 2017 and 2023 were included if they met the following criteria:

- Peer-reviewed articles or conference proceedings.
- AI’s role in diagnosis, treatment planning, and decision support was focused on.
- Clear comparison between AI and human clinician performance [7].

The findings were synthesized to highlight patterns in AI’s diagnostic accuracy, efficiency, and limitations of AI, providing a foundation for comparing real-world case studies with broader trends.

Comparative Assessment

To directly evaluate AI’s performance of AI against that of physicians in controlled settings, a comparative assessment was conducted using benchmark studies. This method was selected to quantify AI’s diagnostic and therapeutic capabilities of AI relative to those of human clinicians, offering an objective measure of its potential advantages. Key studies were reviewed, including the following:

- A 2020 study comparing AI systems to radiologists in breast cancer screening demonstrated AI’s ability to reduce false positives while maintaining high sensitivity [3].

ISSN: 0028-0712

- A 2021 trial pitting AI against oncologists in recommending personalized cancer therapies highlighted AI's proficiency in identifying overlooked biomarkers but also revealed discrepancies in practical applicability [8].

Expert Interviews

To capture nuanced perspectives on AI's integration into clinical practice, informal interviews were conducted with three practicing physicians and two AI developers specializing in healthcare technologies. This method was included to explore practical and ethical considerations, such as physician trust in AI and technical challenges in deployment, which are critical for understanding AI's real-world viability. The interviews covered:

- Physicians' perceptions of AI as a diagnostic aid.
- Developers' perspectives on integrating AI into clinical workflows.
- Concerns regarding data privacy, interpretability, and patient trust exist [10].

Having outlined our multifaceted approach in the Methods section, we now turn to practical insights garnered from our case studies that illustrate the nuanced interplay between AI and physician expertise in real-world contexts.

By combining these methods, this study systematically addressed the research question, leveraging real-world examples, existing research, performance benchmarks, and expert perspectives to evaluate AI's potential advantages over physicians in diagnosing and treating complex medical cases.

ISSN: 0028-0712

Results:

This study evaluated the performance of AI platforms compared to that of physicians in diagnosing and treating complex medical cases through case studies, literature synthesis, comparative assessments, and expert interviews. The findings from the case studies under specific clinical contexts are presented below, followed by a comparison of AI and physician performance, including AI's advantages, limitations, and complementary strengths of physicians.

Case Studies: AI vs. Physicians in Action

- **Rare Disease Diagnosis**

***Key Finding*: AI platforms can identify rare conditions by analyzing complex datasets; however, their insights require physician interpretation to be translated into effective treatments.

A notable example is a pediatric case study published in *Nature Medicine* (2019), in which an AI platform analyzed genomic data from a child with an undiagnosed metabolic disorder [5]. Despite multiple specialist consultations, the diagnosis remains unclear. AI cross-referenced the child's genetic profile with millions of research papers and identified a mutation linked to a rare enzyme deficiency. This enables physicians to tailor their treatment plans to stabilize the patient's condition. However, AI's role was limited to generating insights, with clinicians essential for interpreting and implementing actionable interventions, highlighting the symbiotic relationship between AI and human expertise [6].

- **Cancer Treatment Planning:**

***Key Findings*:** AI can enhance precision medicine by identifying overlooked biomarkers. However, its recommendations may not align with practical clinical constraints.

ISSN: 0028-0712

In oncology, precision medicine relies on the identification of mutations that drive tumor growth. A 2021 study compared treatment recommendations from oncologists with those from an AI platform trained on thousands of cancer cases [7]. In several instances, AI has suggested alternative therapies based on subtle biomarkers overlooked by physicians. For example, in the case of lung cancer, AI recommended immunotherapy instead of chemotherapy, leading to improved survival rates. However, AI recommendations sometimes fail to account for practical considerations, such as cost, drug availability, or patient preferences, factors that physicians routinely integrate into decision making [8].

- **Emergency Room Triage**

***Key Findings*:** AI can detect critical conditions in high-pressure environments; however, incomplete data can lead to errors, underscoring the need for human oversight.

AI systems have been deployed in emergency departments to assist with triage in high-stress situations. A pilot program in a major urban hospital used an AI-powered tool to prioritize patients based on their vital signs, laboratory results, and historical data [9]. In one scenario, the AI flagged a seemingly stable patient with early signs of sepsis, enabling early intervention that saved the patient's life. This demonstrates AI's ability to identify critical patterns amid chaos. However, there were instances in which AI misclassified low-risk patients as high-priority patients owing to incomplete data entry, emphasizing the importance of human oversight [10].

Comparison of AI and Physician Performance

Case studies illustrate both the potential and limitations of AI in complex medical scenarios. Below is a summary of the key findings from the case studies, followed by a structured comparison of AI's advantages and limitations of AI relative to those of physicians.

- **Summary of Case Study Findings**:
  - In rare disease diagnosis, AI excels at processing vast datasets to identify rare genetic mutations but requires physician expertise to interpret and act on these insights.
  - In cancer treatment planning, AI has identified subtle biomarkers that improve treatment recommendations but fail to consider practical constraints, such as cost and patient preferences.
  - In emergency room triage, AI detected critical conditions early but made errors when the data were incomplete, highlighting the need for human oversight.
- **Advantages of AI Platforms**:
  - **Data Processing Power**: AI excels in analyzing massive datasets far beyond human capacity, enabling the rapid identification of correlations and anomalies [1].
  - **Unbiased Analysis**: Unlike humans, AI lacks cognitive biases (e.g., confirmation bias) that can influence clinical judgment [11]. However, it is important to note that AI can inherit algorithmic biases from training data. For instance, if trained on non-representative datasets, AI may perpetuate care disparities. To mitigate this, diverse and representative training data, regular bias audits, and transparency in AI development are essential. Although AI can reduce certain human biases, it is not inherently unbiased and requires ongoing vigilance to ensure fairness and accuracy.
  - **Continuous Learning**: AI systems improve over time as they are exposed to more data, whereas physician knowledge may plateau without ongoing education [12].
  - **Scalability**: AI can serve remote or underserved populations by providing expert-level analyses, regardless of geographic location [13].

ISSN: 0028-0712

- **Limitations of AI Platforms**:
    - **Lack of Contextual Understanding**: AI struggles to account for socioeconomic factors, cultural nuances, and patient values that shape treatment choices [14].
    - **Dependence on Quality Data**: AI outputs are only as reliable as the data fed into them; biased or incomplete datasets can lead to erroneous conclusions [15].
    - **Ethical Concerns**: Questions regarding accountability arise when AI-driven decisions result in adverse outcomes [11].
    - **Interpretability Challenges**: Many AI models operate as "black boxes," making it difficult for clinicians to understand or trust their reasoning [9].
- **Physician Strengths Complementing AI**
    - **Empathy and Communication**: Physicians build rapport with patients, fostering trust and adherence to treatment plans [8].
    - **Holistic Perspective**: Clinicians consider the entire person, including psychological, emotional, and environmental factors [6].
    - **Adaptability**: Physicians can quickly pivot when new information emerges, whereas AI systems may struggle unless explicitly retrained [13].
    - **Ethical Judgment**: Humans navigate moral dilemmas with nuances, balancing competing priorities in ways that AI cannot replicate [11].

Real-World Implications

The future of healthcare is likely to lie in hybrid models, where AI augments rather than replaces physician expertise. For example:

- **Decision Support Systems**: AI tools can act as virtual assistants, offering evidence-based suggestions while clinicians make the final decisions [2].

- **Remote Monitoring**: AI-powered wearable devices can alert physicians to deteriorating conditions before patients notice their symptoms [10].
- **Training Enhancement**: Medical students can use AI simulations to diagnose rare diseases or manage emergencies [7].

Discussion:

This study aimed to investigate whether AI technologies benefit medical professionals in the diagnosis and treatment of complex medical conditions, particularly in real-world applications. Through an analysis involving case studies, literature reviews, comparative assessments, and expert interviews, the findings indicate that AI platforms can offer significant advantages, particularly in processing large datasets and identifying patterns that may not be immediately apparent to human clinicians. For example, in the diagnosis of rare diseases, AI can identify genetic mutations that elude multiple specialists [5], and in oncology, it can suggest personalized treatment options based on the presence of biomarkers [7]. These examples underscore AI's potential to improve diagnostic accuracy and therapeutic efficacy in complex cases. Moreover, they provide concrete illustrations of how AI can be integrated into clinical workflows to support medical professionals and reveal practical challenges that must be addressed.

However, this study also highlights the critical limitations of AI that must be addressed. AI's inability to incorporate contextual factors, such as patient preferences, socioeconomic constraints, and cultural nuances, was evident in an oncological case study in which its recommendations sometimes ignored practical considerations [8]. Furthermore, AI's performance of AI is highly dependent on the quality and representativeness of the data on which it is trained, raising concerns about bias and generalizability [15]. The "black box" nature of many AI algorithms further complicates their integration into clinical practice, as clinicians may struggle to understand or trust the reasoning behind AI recommendations [9]. Ethical concerns, particularly regarding accountability for AI-driven decisions, remain uncertain [11].

These findings suggest that although AI can serve as a powerful tool in healthcare, it is not a substitute for human expertise. Physicians possess essential qualities that AI lacks, including empathy, general understanding, flexibility, and ethical judgment [6, 8, 13]. Therefore, the most effective approach is likely to be a hybrid model in which AI augments physician decision making rather than replacing it. In such models, AI can provide data-driven insights, and physicians can integrate these insights with their clinical judgement and patient-specific knowledge to make informed decisions [2].

The practical implications of this study are significant for medical professionals, healthcare administrators, and policymakers. For clinicians, this highlights the importance of implementing AI as a complementary tool to enhance their capabilities, particularly in data-intensive areas. These care systems should invest in AI technologies, while ensuring robust data governance and ethical oversight. Furthermore, this study highlights the potential of AI to improve access to high-quality care in remote or underserved areas, where specialist expertise may be uncommon [13].

This study had some limitations. The reliance on a selected number of case studies may limit the generalizability of the findings, as these cases may not represent the full spectrum of complex medical conditions. While the informal nature of expert interviews provides valuable qualitative insights, it lacks the rigor of a systematic survey, potentially introducing bias. Furthermore, the rapidly evolving nature of AI technology means that the findings may need to be revisited as new advancements are made.

Future research should focus on addressing the current limitations of AI in healthcare. This includes developing AI systems that can better integrate contextual information, such as advanced natural language processing or multimodal data analysis. Efforts to improve the transparency and interpretation of AI algorithms are crucial for building clinician trust and facilitating their adoption in clinical practice. Furthermore, establishing clear ethical guidelines and accountability frameworks is essential for navigating the complex moral landscape of AI-driven healthcare [11]. Research on bias mitigation strategies, such as ensuring diverse and representative training datasets, is critical for preventing AI from perpetuating or exacerbating healthcare disparities [15].

Conclusions:

AI platforms undoubtedly have transformative potential in the field of medicine, particularly for handling complex cases that are characterized by complexity and uncertainty. Their unparalleled ability to process data and uncover hidden patterns offers several advantages. However, these systems are not infallible and lack contextual awareness, empathy, and ethical judgment intrinsic to human practitioners.

Rather than framing the debate as "AI versus physicians," the focus should shift toward fostering collaboration between the two. By combining the analytical precision of AI with the wisdom and compassion of physicians, healthcare systems can achieve better patient outcomes. As technology continues to evolve, the ultimate goal must remain clear: to leverage innovation to enhance rather than overshadow the art and science of healing.

Declarations:

Acknowledgements: The authors are grateful to the Deanship of Scientific Research, King Saud University, Riyadh, Kingdom of Saudi Arabia, for funding through the Vice Deanship of Scientific Research Chairs, Cancer Research Chair.

Competing interests: The authors declare no competing interests.

ISSN: 0028-0712

References:

1. Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44–56. doi: 10.1038/s41591-018-0300-7
2. Esteva, A., et al. (2019). A guide to deep learning in healthcare. *Nature Medicine*, 25(1), 24-29. doi: 10.1038/s41591-018-0316-z
3. McKinney, S. M., et al. (2020). International evaluation of an AI system for breast cancer screening. *Nature*, 577(7788), 89-94. doi: 10.1038/s41586-019-1799-6
4. Jiang, F., et al. (2017). Artificial intelligence in healthcare: Past, present, and future. *Stroke and Vascular Neurology*, 2(4), 230–243. doi: 10.1136/svn-2017-000101
5. Quazi, S. (2022). Artificial intelligence and machine learning in precision and genomic medicine. *Medical Oncology*, 39(8), 120. doi: 10.1007/s12032-022-01711-1
6. Ghassemi, M., et al. (2020). A review of challenges and opportunities in machine learning for health. *AMIA Joint Summits on Translational Science Proceedings*, 2020, 191-200.
7. Li, J., et al. (2021). Artificial intelligence-assisted decision making for prognosis and drug efficacy prediction in lung cancer patients: a narrative review. *Journal of Thoracic Disease*, 13(12), 7021-7033. doi: 10.21037/jtd-21-864
8. Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the future: big data, machine learning, and clinical medicine. *The New England Journal of Medicine*, 375(13), 1216–1219. doi: 10.1056/NEJMp1606181
9. Rajpurkar, P., et al. (2018). Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Medicine*, 15(11), e1002686. doi: 10.1371/journal.pmed.1002686

10. Reddy, S., Fox, J., & Purohit, M. P. (2019). Artificial intelligence-enabled healthcare delivery. *Journal of the Royal Society of Medicine*, 112(1), 22-28. doi: 10.1177/0141076818815510

11. Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing machine learning in health care - addressing ethical challenges. *The New England Journal of Medicine*, 378(11), 981-983. doi: 10.1056/NEJMp1714229

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

13. Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in healthcare. *JAMA*, 319(13), 1317-1318. doi: 10.1001/jama.2017.18391

14. Amisha, et al. (2019). Overview of artificial intelligence in medicine. *Journal of Family Medicine and Primary Care*, 8(7), 2328-2331. doi: 10.4103/jfmpc.jfmpc_440_19

15. Ting, D. S. W., et al. (2017). Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA*, 318(22), 2211-2223. doi: 10.1001/jama.2017.18152