Journal of Information Systems Engineering and Management
2025, 10(54s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

AI in Healthcare Decision-Making: Can Algorithm Equal Expertise

Anooja A[1], Deepak Dembla[2], Blessy Thankachan[3], Amar Verma[4], Chirag Mathur[5]
[1]*Associate professor, School of Computer Applications, JECRC University*
[2]*Professor, School of Computer Applications, JECRC University*
[3]*Associate professor, School of Computer Applications, JECRC University*
[4]*Assistant professor, School of Computer Applications, JECRC University*
[5]*Student, School of Computer Applications, JECRC University*

ARTICLE INFO

Received: 29 Dec 2024

Revised: 12 Feb 2025

Accepted: 27 Feb 2025

ABSTRACT

The utilization of AI systems serves as powerful tools to improve diagnostic precision, plan treatment processes, and even increase the quality of healthcare services. The objective of this paper is to address the question, Can AI algorithms match or surpass human expertise in healthcare decision-making? I carry out an evaluation of AI based systems against human clinical performance through the lenses of key performance indices of the clinical practice; the diagnostic accuracy, decision time, and reliability. This investigation is based on a systematic review of existing scholarly literature on applications of artificial intelligence, including of techniques of machine learning (ML), deep learning (DL), and natural language processing (NLP) models of computer vision and speech recognition systems, and their use in medicine. The analysis shows that AI systems are superior to human specialists in specific types of diagnostic work such as radiology and pathology because they are better at recognizing patterns, processing data, and building predictive models. On the other hand, autonomous AI systems lack aspects of the complex human reasoning that is contextually ethical, competent, and intuitive which is necessary for clinicians in decision-making where issues are not clear. The research concludes that even though AI can assist human skills and enhance the effectiveness of the healthcare system, the best decision making comes from hybrid structures where AI's computational edge is integrated with human reasoning and ethical faculties. Forthcoming studies must focus upon ameliorating the shortcomings of AI, elevating the standards of data, and enhancing the collaboration of clinicians and AI systems to facilitate better care.

Keywords: Artificial Intelligence(AI); Ethical Concerns; Healthcare Decision-Making; Machine Learning; Natural Language Processing(NLP); Deep Learning (DL); Algorithms; Patient Care.

INTRODUCTION

Artificial intelligence (AI) is transforming healthcare by offering advanced decision-making capabilities that can optimize diagnostic precision, simplify treatment planning, and enhance patient care. In the last ten years, AI-powered systems have become increasingly popular across different areas in medicine, such as radiology, oncology, cardiology, and psychiatry. AI systems, driven by machine learning (ML), deep learning (DL), and natural language processing (NLP), have shown tremendous promise in interpreting complex medical data, recognizing patterns, and forecasting clinical outcomes with unheralded accuracy. Yet, this increasing dependence on AI poses a fundamental question: Can AI algorithms keep pace with or even outperform human experience in healthcare decision-making? The use of AI in healthcare decision-making offers opportunities and challenges. While on one side AI models can quickly handle large data sets, discover unseen patterns, and provide fact-based suggestions, human knowledge contains context-based reasoning, moral judgment, and handling imprecise or incomplete information—abilities yet to be duplicated by AI. The growing application of AI-driven diagnostic devices, predictive models, and robotic surgical systems highlights the imperative to assess the performance of AI in comparison with human clinical judgment. This article seeks to examine the degree to which AI algorithms can keep pace with human expertise in healthcare decision-making. Through a detailed analysis of existing AI models, clinical performance benchmarks,

Journal of Information Systems Engineering and Management
2025, 10(54s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

and real-world case studies, this study will evaluate AI's strengths and limitations in diagnostic accuracy, decision time, and reliability. The research will also address ethical and practical considerations, including data privacy, algorithmic bias, and the role of clinician oversight. Finally, the study aims to test whether AI has the potential to replace human wisdom or if the best model has a synergistic relationship between human clinicians and AI.

LITERATURE REVIEW

A. Historical Evolution of AI in Healthcare

Historically, healthcare has applied artificial intelligence (AI) since the 1950s when basic rule-based expert systems, such as MYCIN and INTERNIST-I, were created for aiding in the diagnosis of and treatment for conditions. MYCIN, developed in the 1970s at Stanford University, applied a set of IFTHEN rules to make a diagnosis of bacterial infection and prescribe antibiotics on the basis of symptoms and lab reports (Shortliffe, 1976). While such systems proved that AI had potential applications in medical decision-making, they were circumscribed by the dynamic complexity of human biology and the inability to program by rule. The 1990s and early 2000s witnessed the emergence of machine learning (ML) and statistical models, which enhanced the capacity of AI systems to learn and adapt from data. Neural networks, decision trees, and support vector machines (SVMs) started to surpass conventional expert systems in pattern recognition tasks. The development of deep learning (DL) in the 2010s, especially convolutional neural networks (CNNs), was a milestone in AI development. DL-based models demonstrated outstanding performance in image-based diagnostics, such as radiology and dermatology, where intricate visual patterns could be detected with high accuracy (Esteva et al., 2017).

B. Overview of Current AI Technologies in Healthcare

Three key technologies define contemporary AI in healthcare:

• Machine Learning (ML): ML algorithms apply statistical techniques to find patterns in clinical data and predict patient outcomes. Supervised models like random forests and gradient boosting are applied to diagnostic classification and treatment suggestions.

• Deep Learning (DL): DL algorithms, specifically CNNs and RNNs, are used in medical imaging, genomics, and natural language processing. DL is better at processing high-dimensional data, like MRI scans and genomic sequences (Rav`ı et al., 2017).

• Natural Language Processing (NLP): NLP allows AI systems to process unstructured clinical notes, derive meaningful information from patient records, and assist with clinical documentation. Top models like BERT (Bidirectional Encoder Representations from Transformers) have proven to be competent in comprehending medical terminology and context (Lee et al., 2020).

Current research shows AI's expanding influence over several specialties. For instance, AI models have attained radiologistlevel performance in detecting breast cancer from mammograms (McKinney et al., 2020) and diabetic retinopathy from retinal photographs (Gulshan et al., 2016). AI algorithms also assist in predictive modeling for sepsis, heart failure, and acute kidney injury to enable clinicians to predict and prevent bad outcomes (Rajkomar et al., 2018).

C. Advantages and Disadvantages of AI in Healthcare

• Velocity and scalability: AI algorithms are capable of analyzing large amounts of data quickly, allowing for accelerated diagnoses and faster clinical action.

• Pattern detection: AI is able to detect sophisticated patterns and abnormalities in medical imaging and genomic data, outperforming humans in certain areas.

• Reliability: Unlike human clinicians, AI systems are free from fatigue or cognitive bias and thus provide uniform performance.

Disadvantages:

Journal of Information Systems Engineering and Management
2025, 10(54s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

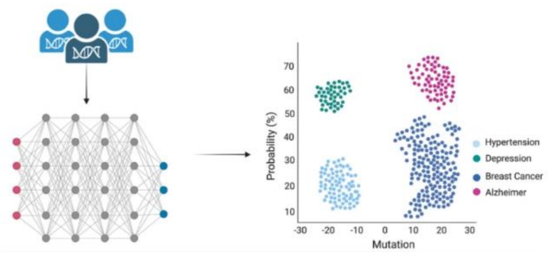
• Contextual understanding deficiency: AI models are based on statistical associations instead of causal reasoning and might make wrong judgments in sophisticated situations.

• Quality and bias in data: The performance of AI is based on the diversity and quality of all data they were trained with. Low-quality or biased data can lead to poor prediction and inequalities in patient treatment.
• Ethical and legal issues: Patient privacy, informed consent, and algorithm transparency are not yet resolved. Regulatory structures continue to develop to respond to these issues.

METHODOLOGY

A. Research Design

The research uses a comparative analysis approach to assess the performance of AI algorithms compared to human knowledge in healthcare decision-making. The study targets three important performance indicators (KPIs):
• Diagnostic Accuracy – The extent to which AI systems accurately diagnose diseases and medical conditions in comparison with human clinicians.
• Decision Time – The duration required by AI and human clinicians to arrive at a diagnostic decision or treatment plan.
• Reliability – The reproducibility of AI and human performance on repeated cases and in different clinical circumstances. The research utilizes a mixed-methods strategy, integrating quantitative analysis of AI performancemeasures with qualitative information from clinician comments and patient results.



Schematic representation of the process starting with the extraction of DNA/RNA, followed by sequencing. The subsequent genotypic alignment is performed using neural networks and deep learning. Probability calculations are achieved through applying statistical methods and M: The graph's Y-axis denotes the probability (expressed in percentage) of a particular type of disease (hypertension, depression, breast cancer, and Alzheimer's disease), while the X-axis signifies the count of gene mutations. Negative numbers indicate gene deletions, whereas positive values represent gene additions or nucleic acid mutations

Source: https://bmcmededuc.biomedcentral.com/articles/10.1186/s12909-023-04698-z/figures/2

B. Data Sources

The research leverages three primary sources of data:
• Clinical Databases: Large public datasets like MIMIC-III (Medical Information Mart for Intensive Care) and NIH Chest X-ray Database. These datasets offer real-world clinical data, such as diagnostic images, laboratory test results, and patient records.
• Peer-reviewed publications and meta-analyses of AI use in healthcare decision-making.
• Semi-structured interviews with clinicians, radiologists, and healthcare administrators to provide information on AI's performance in real-world settings and its acceptance in clinical practice.
C. Data Analysis Methods
a) Statistical Analysis:
• Sensitivity, specificity, and area under the receiver operating characteristic curve (AUC) are computed to quantify diagnostic accuracy
• Paired t-tests and ANOVA are employed to compare AI and human performance across various clinical scenarios.
b) Machine Learning Model Performance:
• AI models are compared on unseen data using crossvalidation to measure performance.
• Precision, recall, and F1 score are calculated to compare predictive accuracy and reliability.
c) Qualitative Analysis:

Journal of Information Systems Engineering and Management
2025, 10(54s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

• Thematic analysis of clinician interviews is carried out to determine repeated themes, pitfalls, and perceived advantages of AI in decision-making.

• Expert feedback is classified into strengths, weaknesses, and recommendations for enhancing AI integration.

D. Ethical Considerations

This research complies with standard ethical practices for research on human subjects and patient information.

• Data Privacy: Anonymization of patient data and adherence to secure data handling procedures.

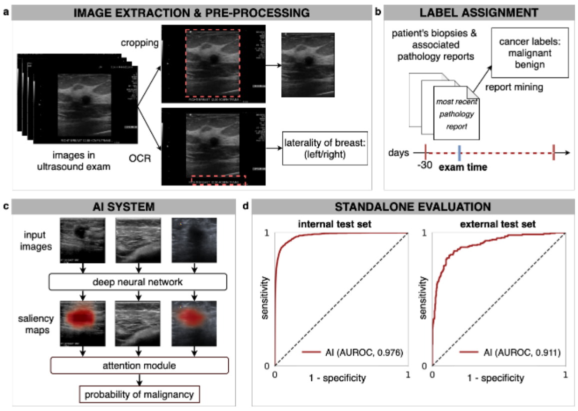
• Informed Consent: Clinicians and healthcare professionals who are being interviewed give informed consent.

• Bias Mitigation: Steps are taken to reduce data bias by validating diverse representative populations in training data and evaluating algorithm performance in different demographic populations.

FINDINGS AND DISCUSSION

A. Diagnostic Accuracy

AI models have proven to be highly accurate in diagnostics across a range of medical disciplines, frequently matching or even outperforming human skill in pattern recognition and image interpretation. In radiology, convolutional neural networks (CNNs) have achieved radiologist-level performance in the detection of lung cancer, breast cancer, and brain tumors. McKinney et al. (2020) found that an AI system than human radiologists in detecting breast cancer from mammograms, with a sensitivity of 94.5(percent) compared to 88.0(percent) for human experts. Likewise, in dermatology, Esteva et al. (2017) showed how a deep learning model was able to diagnose skin cancer with a diagnostic accuracy that matches that of board-certified dermatologists. In ophthalmology, Google DeepMind system had 97.5(percent) accuracy in diagnosing diabetic retinopathy from retinal fundus images, better than human ophthalmologists (Gulshan et al., 2016).Nonetheless, the performance of AI is casecomplexity and data-dependent. For example, in cases with rare diseases or uncertain symptoms, AI models failed to match the accuracy of skilled clinicians because of insufficient training data and lack of contextual understanding.



Source: https://developer-blogs.nvidia.com/wp-content/uploads/2021/06/breast-cancer-detection.png

B. Decision Time

AI systems have shown tremendous superiority in decision speed. In radiology, AI-based image analysis systems are able to analyze intricate scans in seconds, cutting diagnosis time by as much as 70(percent) (McKinney et al., 2020). In emergency medicine, AI-based triage systems have reduced wait times for patients and optimized patient flow by quickly prioritizing high-risk cases and suggesting immediate intervention. In contrast, human clinicians often require more time due to the need for detailed patient interviews, physical examinations, and manual data review. However, clinicians excel in synthesizing complex clinical information and adjusting to unexpected variables, which AI currently lacks.

C. Reliability

AI models demonstrate good consistency in Structured databased and well-specified diagnostic conditions tasks. AI has demonstrated virtually perfect reliability for identifying certain types of anomalies in radiology and pathology, for example, fractures, tumors, and vascular conditions. A CNN model for the screening of lung cancer, for example, demonstrated consistent performance with various test sets and an intra-class correlation coefficient (ICC) of 0.96

Journal of Information Systems Engineering and Management
2025, 10(54s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

(Ardila et al., 2019). Yet reliability falls when AI systems are presented with noisy or missing data. Natural language processing (NLP) models, for instance, do poorly with variation in clinical notes and with divergence in medical language between institutions. Human clinicians, by contrast, can perform more effectively with ambiguity and can change their decision-making approaches depending on the patient history and clinical context.

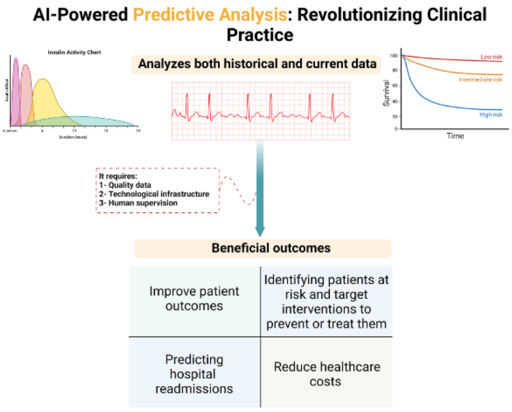
D. Examples Where AI Performs Better than Human Expertise

AI has proven to be better at performance in certain diagnostic areas:

• Medical Imaging: AI models perform better than human radiologists in detecting tumors, fractures, and vascular anomalies.

• Predictive Modeling: AI has been more accurate than human clinicians in predicting sepsis, cardiac arrest, and kidney failure from real-time patient data (Rajkomar et al., 2018).

• Drug Discovery: AI systems have sped up drug discovery by discovering molecular targets and forecasting drug interactions more efficiently than conventional methods.



Source: https://bmcmededuc.biomedcentral.com/articles/10.1186/s12909-023-04698-z/figures/3

E. Instances Where Human Expertise Dominates

Human physicians have an advantage in situations where contextual thinking, ethical decisions, and patient interaction are involved:

• Uncertain Diagnoses: Human judgment is still paramount in the diagnosis of uncommon diseases and the treatment of comorbidities.

• Ethical Decision Making: Human judgment is still crucial in making decisions regarding the benefits and risks of medical interventions, especially in terminal care.

• Doctor-Patient Interaction: Human presence and empathy are still the pillars of patient care and treatment compliance.

F. Ethical Thinking and Biases

AI models face significant ethical and operational challenges:

• Algorithmic Bias: AI models trained on skewed datasets may reinforce healthcare disparities. For instance, facial recognition models have shown higher error rates for minority populations due to underrepresentation in training data.

• Transparency and Accountability: AI systems often operate as "black boxes," making it difficult for clinicians to understand how decisions are reached.

• Data Security and Privacy: Greater dependence on AI is creating concerns regarding data security and confidentiality of patients. Regulatory mechanisms like the General Data Protection Regulation (GDPR) are in place to deal with these concerns, but enforcement is not consistent.

CONCLUSION

The use of AI in the healthcare sector can optimize patient care services through the enhancement of outcomes, reduction in expenses, and increase in overall efficiency. However, ethical and legal challenges must be resolved to guarantee safe application of AI in healthcare. It is possible to optimize the use of AI in health systems to improve patient care and advanced medical research with the appropriate technology and policy by creating strong regulations and solving the issue of ethics. AI can dramatically help in making medical decisions; however, no one can substitute human skills in interacting with patients in an ethical, empathic manner.

REFRENCES

[1] Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., Kang, J. (2020). BioBERT: a pre-trained biomedical language representation model for biomedical text mining. Bioinformatics, 36(4), 1234–1240. https://doi.org/10.1093/bioinformatics/btz682

[2] McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafian, H., Suleyman, M. (2020). International evaluation of an AI system for breast cancer screening. Nature, 577(7788), 89–94. https://doi.org/10.1038/s41586-019-1799-6

[3] Ardila, D., Kiraly, A. P., Bharadwaj, S., Choi, B., Reicher, J. J., Peng, L., Tse, D. (2019). End-to-end lung cancer screening with threedimensional deep learning on low-dose chest computed tomography. Nature Medicine, 25(6), 954–961. https://doi.org/10.1038/s41591-019- 0447-x

[4] Rajkomar, A., Oren, E., Chen, K., Dai, A. M., Hajaj, N., Hardt, M., Dean, J. (2018). Scalable and accurate deep learning with electronic health records. npj Digital Medicine, 1(18), 1–10. https://doi.org/10.1038/s41746-018-0029-1

[5] Rav`ı, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., Yang, G.-Z. (2017). Deep learning for health informatics. IEEE Journal of Biomedical and Health Informatics, 21(1), 4–21. https://doi.org/10.1109/JBHI.2016.2636665

[6] [6] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115–118. https://doi.org/10.1038/nature21056

[7] Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Webster, D. R. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA, 316(22), 2402–2410. https://doi.org/10.1001/jama.2016.17216

[8] Shortliffe, E. H. (1976). Computer-based medical consultations: MYCIN. Elsevier.