

## The Role of Artificial Intelligence in Enhancing Patient Safety: Opportunities and Challenges


Farahzad Jabbari-Azad<sup>1\*</sup>

1. Allergy Research Center, Mashhad University of Medical Sciences, Mashhad, Iran

ARTICLE INFO	ABSTRACT
<p><b>Article type:</b> Editorial</p>	<p>Patient safety is a fundamental dimension of healthcare quality, and reducing medical errors remains a major objective of health systems worldwide (1,2). In recent years, advances in artificial intelligence (AI) have created new opportunities to improve patient safety across various clinical settings (3,4). Through the analysis of large volumes of healthcare data, identification of hidden patterns, and prediction of potential risks, AI has the potential to support clinical decision-making and reduce preventable adverse events (4,6). Nevertheless, the integration of AI into healthcare practice raises important concerns regarding ethics, data privacy, transparency, and overreliance on automated systems (9,10). This article reviews the potential contributions of AI to patient safety and discusses the challenges that must be addressed to ensure its safe and effective implementation.</p>
<p><b>Article History:</b> <b>Received:</b> 08 Jun 2026 <b>Accepted:</b> 13 Jun 2026</p>	
<p><b>Keywords:</b> Artificial Intelligence, Patient Safety, Medical Errors, Clinical Decision Support, Healthcare Quality</p>	
<p>► Please cite this paper as: Jabbari-Azad F. The Role of Artificial Intelligence in Enhancing Patient Safety: Opportunities and Challenges . <i>Journal of Patient Safety and Quality Improvement</i>. 2026; 14(3): 136-138. Doi: 10.22038/psj.2026.96307.1535</p>	

**\*Corresponding Author:**

Allergy Research Center, Mashhad University of Medical Sciences, Mashhad, Iran. Email: Jabbarif@mums.ac.ir

 Copyright©2026 Mashhad University of Medical Sciences. This work is licensed under a Creative Commons Attribution-Noncommercial 4.0 International License <https://creativecommons.org/licenses/by-nc/4.0/deed.en>

## **Introduction**

Patient safety has become one of the most important indicators of healthcare quality. Despite significant advances in medical science and technology, preventable adverse events continue to occur in healthcare organizations and remain a major source of patient harm (1,2). Efforts to improve safety have traditionally focused on system redesign, standardization of care processes, and strengthening safety culture (2). More recently, artificial intelligence has emerged as a promising tool that may further support these efforts (3,4).

Artificial intelligence refers to computer systems capable of performing tasks that typically require human intelligence, such as learning, pattern recognition, prediction, and decision support (4,6). The growing availability of electronic health records and large clinical datasets has accelerated the development of AI-based applications in healthcare, creating new possibilities for improving patient outcomes and reducing risks (4,6).

## **Applications of Artificial Intelligence in Patient Safety**

One of the most significant contributions of AI to patient safety is its ability to identify patients at risk of clinical deterioration. Machine learning algorithms can analyze physiological parameters, laboratory results, and clinical histories to predict complications such as sepsis, respiratory failure, or cardiac events before they become clinically apparent (6,7). Early detection enables healthcare professionals to intervene promptly and potentially prevent serious harm.

Another important application involves medication safety. Medication-related errors remain among the most common patient safety incidents worldwide (1,2). AI-powered clinical decision support systems can detect potential drug interactions, identify inappropriate dosages, and alert clinicians to patient-specific risks, thereby reducing the likelihood of adverse drug events (6,7).

AI also plays an increasing role in diagnostic accuracy. Diagnostic errors can lead to delayed treatment, inappropriate interventions, and poor clinical outcomes. Advanced AI algorithms have demonstrated considerable potential in interpreting medical images, recognizing subtle abnormalities, and assisting clinicians in making more accurate diagnoses (5,6). Rather than replacing healthcare professionals, these technologies can serve as valuable tools that complement clinical expertise (3).

Furthermore, AI can contribute to organizational learning by analyzing incident reports and patient safety data. Healthcare institutions generate large amounts of information related to adverse events, near misses, and safety concerns. AI systems can identify recurring patterns and underlying risk factors that may not be immediately evident through conventional analysis (7). Such insights can support targeted interventions and continuous quality improvement initiatives.

## **Challenges and Ethical Considerations**

Despite its considerable promise, the use of AI in healthcare presents several challenges. The effectiveness of AI systems depends heavily on the quality and representativeness of the data used for training and validation (6,9). Incomplete, inaccurate, or biased datasets may lead to unreliable predictions and potentially compromise patient safety.

Data privacy and cybersecurity also represent significant concerns. Healthcare organizations must ensure that patient information is protected and that AI systems comply with applicable ethical and legal standards (9,10). Maintaining public trust requires transparency regarding how patient data are collected, processed, and utilized.

Another challenge involves the interpretability of AI algorithms. Many advanced machine learning models operate as “black boxes,” making it difficult for clinicians to understand how specific

recommendations are generated (10). Lack of transparency may hinder acceptance and limit the practical integration of AI into clinical workflows.

Importantly, AI should not be viewed as a substitute for professional judgment. Clinical decision-making often requires contextual understanding, ethical reasoning, empathy, and communication with patients and families—capabilities that remain uniquely human (3,9). Therefore, AI should be considered a supportive technology that enhances, rather than replaces, the role of healthcare professionals.

### **Conclusion**

Artificial intelligence has the potential to transform patient safety by supporting early risk detection, improving diagnostic accuracy, reducing medication errors, and facilitating organizational learning (4–7). As healthcare systems continue to adopt digital technologies, AI-based solutions may become increasingly important in efforts to prevent harm and improve the quality of care.

However, successful implementation requires careful attention to ethical considerations, data quality, transparency, and human oversight (9,10). The future of patient safety will likely depend not on replacing clinicians with intelligent machines, but on developing effective partnerships between healthcare professionals and AI technologies (3). When appropriately designed and responsibly implemented, artificial intelligence can

serve as a powerful ally in the ongoing pursuit of safer healthcare.

### **References**

1. Kohn LT, Corrigan JM, Donaldson MS, editors. *To Err Is Human: Building a Safer Health System*. Washington (DC): National Academies Press; 2000.
2. World Health Organization. *Global Patient Safety Action Plan 2021–2030: Towards Eliminating Avoidable Harm in Health Care*. Geneva: World Health Organization; 2021.
3. Topol EJ. *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. New York: Basic Books; 2019.
4. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol*. 2017;2(4):230-43.
5. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115-8.
6. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med*. 2019;380(14):1347-58.
7. Sendak MP, D'Arcy J, Kashyap S, Gao M, Nichols M, Corey K, et al. A path for translation of machine learning products into healthcare delivery. *EMJ Innov*. 2020;4(1):44-52.
8. Bates DW, Singh H. Two decades since *To Err Is Human*: an assessment of progress and emerging priorities in patient safety. *Health Aff (Millwood)*. 2018;37(11):1736-43.
9. Char DS, Shah NH, Magnus D. Implementing machine learning in health care—addressing ethical challenges. *N Engl J Med*. 2018;378(11):981-3.
10. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med*. 2019;17(1):195