What Are the Cybersecurity Risks of AI in Healthcare?

Rasit Dinc

Rasit Dinc Digital Health & AI Research

Published: July 1, 2019 | Healthcare Data Privacy and Security

DOI: 10.5281/zenodo.17998820

Abstract

Artificial intelligence (AI) is rapidly transforming the healthcare landscape, offering unprecedented opportunities to improve diagnostics, personalize treat...

What Are the Cybersecurity Risks of AI in Healthcare?

Author: Rasit Dinc

Artificial intelligence (AI) is rapidly transforming the healthcare landscape, offering unprecedented opportunities to improve diagnostics, personalize treatments, and streamline administrative processes. However, the increasing integration of AI into healthcare systems also introduces a new frontier of cybersecurity risks that can have profound implications for patient safety and data privacy. For health professionals, understanding these risks is not just a matter of IT governance but a core component of modern clinical practice and risk management.

The Expanding Attack Surface: New Vulnerabilities in AI-Driven Healthcare

The adoption of AI in healthcare significantly expands the attack surface for malicious actors. AI systems, which often rely on vast datasets and complex algorithms, present unique vulnerabilities that differ from traditional IT systems. These can be broadly categorized into three main areas: risks related to data, risks associated with the operation of AI-powered devices, and systemic risks within the broader healthcare infrastructure.

Unauthorized Access to and Manipulation of Health Data

One of the most significant cybersecurity risks in AI-driven healthcare is the potential for unauthorized access to sensitive patient data. AI systems are data-hungry, and the large, centralized datasets they often require for training and operation can become high-value targets for cybercriminals. A breach of such a dataset could expose the personal and medical information of thousands or even millions of individuals, leading to identity theft, fraud, and a profound loss of patient trust. Beyond simple data theft, there is also the risk of data manipulation. Malicious actors could alter patient data to disrupt AI-powered diagnostic tools, leading to misdiagnoses and inappropriate treatments. For instance, a recent study highlighted the risk of unauthorized access to radiology reports through portable devices connected to a hospital network, demonstrating how data compromise can have direct and immediate impacts on clinical management and treatment decisions [1].

The Compromise of AI-Controlled Medical Devices

The proliferation of AI-powered medical devices, from insulin pumps to robotic surgical assistants, introduces another layer of risk. These devices are often connected to hospital networks and the internet, making them susceptible to hacking. A successful attack could have devastating consequences, allowing a malicious actor to alter the device's operation. For example, an attacker could change the dosage of an insulin pump or interfere with the signals of a pacemaker, with potentially lethal consequences for the patient. A chilling real-world example of the potential for such attacks occurred in 2020 at a German hospital. A ransomware attack crippled the hospital's IT systems, forcing them to divert a patient in critical condition to another facility. The delay in treatment resulted in the patient's death, a tragic illustration of how cyberattacks can have direct and fatal consequences in a healthcare setting [1].

System-Level Risks and the Challenge of Interoperability

Finally, the interconnected nature of modern healthcare creates systemic risks. Hospitals and clinics rely on a complex web of interconnected systems, including electronic health records (EHRs), picture archiving and communication systems (PACS), and various departmental information systems. The introduction of AI adds another layer of complexity to this ecosystem. A vulnerability in one system can create a cascade of failures across the entire network. The lack of secure interoperability between different systems is a major concern. If the interfaces between different systems are not designed with security in mind, a vulnerability in a seemingly minor component could be exploited to gain access to the entire network. Furthermore, human factors, such as a lack of cybersecurity training among staff, weak passwords, and the use of personal devices on hospital networks, can exacerbate these risks.

The Rise of Generative AI and Novel Threats

The emergence of generative AI models, such as large language models (LLMs), introduces a new set of challenges. These models can be used to create highly realistic "deepfakes" of medical images or to generate convincing but false medical advice, potentially leading to widespread misinformation and harm. There is also the risk of "prompt injection" attacks, where a malicious actor could manipulate the input to a generative AI model to bypass its safety features and generate harmful or biased outputs.

Conclusion: A Call for a New Paradigm in Healthcare Cybersecurity

The cybersecurity risks of AI in healthcare are real and multifaceted. They range from data breaches and device manipulation to systemic vulnerabilities and novel threats from generative AI. Addressing these risks requires a new paradigm in healthcare cybersecurity, one that goes beyond traditional IT security measures and embraces a more holistic, risk-based approach. This includes not only technical solutions, such as robust data encryption and access controls, but also a strong focus on staff training, secure software development practices, and the development of clear regulatory frameworks for the use of AI in healthcare. For health professionals, staying informed about these risks and advocating for a culture of cybersecurity within their organizations is essential to ensuring that the promise of AI in healthcare is realized safely and responsibly.

References

[1] Di Palma, G., Scendoni, R., Ferorelli, D., De Benedictis, A., Tambone, V., & De Micco, F. (2025). AI-Induced Cybersecurity Risks in Healthcare: A Narrative Review of Blockchain-Based Solutions Within a Clinical Risk Management Framework. *Risk Management and Healthcare Policy*, *18*, 3479–3497. https://doi.org/10.2147/RMHP.S544523