Journal of Global Research in Multidisciplinary Studies (JGRMS)
Volume 2 | Issue 10 | October 2026 | Pages 30–37
DOI: https://doi.org/xxxxxxxxxxxxx
e-ISSN: 3049-2963
Review Article OPEN ACCESS Available online: https://saanvipublications.com/journals/index.php/jgrms

Artificial Intelligence for Healthcare Cybersecurity: A Comprehensive Survey of Threat Detection and Defense Mechanisms

Dr. Neha Upadhyay
Assistant Professor
MANIT Bhopal
nupadhyay@manit.ac.in

Corresponding Author: * Neha Upadhyay (nupadhyay@manit.ac.in)

Received: 15, October, 2026 | **Revised:** 20 October,2026 | **Accepted:** 28, October, 2026

***Abstract*—The increasing digitalization of healthcare systems has improved patient care, data management, and clinical decision-making, but it has also expanded the cybersecurity threat landscape. Healthcare organizations face growing risks from ransomware, phishing, malware, insider threats, supply-chain attacks, and vulnerabilities in Internet of Medical Things (IoMT) devices. These threats can compromise sensitive patient information, disrupt essential services, and affect the confidentiality, integrity, and availability of healthcare data. Artificial intelligence (AI) offers promising capabilities for strengthening healthcare cybersecurity through intelligent threat detection, behavioral analysis, anomaly identification, and automated response. This survey examines major cybersecurity threats and vulnerabilities in healthcare environments and reviews AI-based techniques, including machine learning, deep learning, and natural language processing, for identifying suspicious activities and potential attacks. It also discusses AI-integrated healthcare security architectures, cloud security, zero-trust principles, and privacy-preserving techniques such as homomorphic encryption, differential privacy, secure multiparty computation, and federated learning. Furthermore, the survey highlights challenges related to legacy systems, data privacy, access control, and secure deployment. This review provides an overview of AI-enabled detection and defense mechanisms for developing more secure, reliable, and privacy-preserving healthcare infrastructures.**

Keywords*—*Healthcare Cybersecurity, Artificial Intelligence, Machine Learning, Deep Learning, Threat Detection, Ransomware, Zero Trust, Internet of Medical Things (IoMT).

I. INTRODUCTION

The rapid adoption of digital technologies has transformed healthcare services by improving the accessibility, management, and exchange of medical information. Electronic Health Records (EHRs), connected medical devices, and web-based healthcare platforms enable healthcare professionals to access patient information and coordinate clinical activities more efficiently [1]. However, the increasing dependence on interconnected digital systems has also introduced significant cybersecurity challenges. Healthcare organizations manage sensitive patient information and critical medical infrastructure, making them attractive targets for cyberattacks. Unauthorized access, data theft, or the disruption of healthcare systems can compromise patient privacy, affect clinical operations, and undermine trust in healthcare providers [2]. Healthcare cybersecurity threats include ransomware, phishing, malware, network intrusions, and attacks targeting connected medical devices. These threats can exploit vulnerabilities in healthcare networks, outdated software, insufficient security controls, and human errors such as sharing passwords or responding to fraudulent communications. The consequences extend beyond financial and operational losses, as cyberattacks may interrupt medical services, compromise the integrity of patient records, and potentially affect patient safety. The growing complexity of healthcare environments, combined with the increasing number of connected devices and digital platforms, highlights the need for effective and reliable cybersecurity mechanisms [3].

Artificial Intelligence (AI) has emerged as a promising approach to strengthening healthcare cybersecurity by supporting the identification, analysis, and management of potential threats. Machine learning, deep learning, and anomaly detection techniques can analyze network traffic, system activities, and security-related data to identify suspicious patterns and deviations from normal behavior. These capabilities can support intrusion detection, malware identification, and the monitoring of connected medical devices and healthcare information systems [4]. In addition to threat detection, AI-based cybersecurity solutions can assist security teams through automated alerts, threat prioritization, and incident response, thereby improving the efficiency and responsiveness of healthcare security operations [5]. However, the effectiveness of these approaches depends on the availability of reliable data, appropriate model selection, and integration with existing healthcare infrastructure.

Despite its potential, the application of AI in healthcare cybersecurity presents several technical, operational, and ethical challenges. Limited availability of representative

Copyright © 2026 The Author(s).
This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, adaptation, and reproduction in any medium, provided the original author(s) and the source are properly credited.

DOI: https://doi.org/xxxxxxxxxxxx

security datasets, data privacy requirements, model interpretability, and computational costs can affect the development and deployment of AI-based solutions. Furthermore, adversarial attacks, model manipulation, and false-positive detections may reduce system reliability and create additional security concerns. Healthcare organizations must therefore consider the accuracy, robustness, scalability, and practical suitability of AI-driven cybersecurity mechanisms when protecting sensitive information and critical medical systems [6]. It examines the application of machine learning, deep learning, and anomaly detection approaches to identifying cybersecurity threats across healthcare information systems and connected medical devices [7][8].

A. *Structure of the Paper*

The remainder of this paper is organized as follows: Section II discusses cybersecurity threats, attack types, and vulnerabilities in healthcare systems. Section III reviews AI-based threat detection techniques, including machine learning and anomaly detection. Section IV presents AI-driven defense mechanisms, secure healthcare architectures, and privacy-preserving techniques. Section V summarizes related studies and research gaps, while Section VI concludes the paper and outlines future research directions.

II. Cybersecurity Threat Landscape in Healthcare Systems

As there is an increase in the digitalization process, healthcare device networking, and usage of patient information, healthcare organizations face various forms of cyberattacks today. Some of the major cybersecurity threats that have been witnessed are ransomware attacks, phishing attacks, malware attacks, insider attacks, and supply chain attacks.

A. *Cybersecurity Threats and Trends*

Ransomware, phishing, insider risks, IoMT device vulnerabilities, and supply-chain exposures dominate the threat environment in the United States healthcare sector, while new capabilities (AI, deepfakes, quantum threats) are reshaping the horizon [9].

1) *Ransomware and Data Extortion*

Ransomware has become the greatest disruptive cyber threat to the healthcare industry. In recent years, attacks have progressed from simple, opportunistic infections to highly targeted operations employing double- and triple-extortion techniques [10]. These attacks not only encrypt computers but also steal important data, threatening to reveal it unless additional payments are made.

2) *Phishing and Social Engineering*

Phishing remains the most common human-targeted attack vector in the healthcare industry. It is frequently used as the entry point for more serious attacks, such as credential theft, lateral movement, and ransomware deployment. A variety of factors contribute to vulnerability in hospital settings.

3) *Internet of Medical Things (Iomt) and Medical Device Vulnerabilities*

IoMT devices greatly increase the cyber-physical attack surface in healthcare contexts. Many of these devices were built for healthcare purposes rather than cybersecurity [11]. As a result, they frequently have lengthy operational lifecycles, limited patching capabilities, and complex vendor update procedures. Potential attack scenarios include firmware compromise, which changes device behavior.

4) *Supply Chain and Third-Party Risks*

Third-party breaches are a key cause of widespread record exposure. When suppliers such as billing processors, cloud providers, or imaging service partners are compromised, the consequences might spread to several healthcare companies [12]. This is because numerous suppliers rely on common platforms and related services.

5) *Emerging Threats*

Emerging technologies are altering cyber offense and defense. Attackers are increasingly using AI to generate more convincing phishing messages, automate vulnerability scanning, and accelerate exploit development [13]. Deepfakes also introduce new concerns, such as compromising identity verification and facilitating fraud against clinicians or administrative workers [14].

B. *Cyberattacks in Healthcare Systems*

Healthcare systems face cyberattacks that threaten patient data, medical devices, and essential services. Common threats include DoS, MITM, phishing, cryptographic attacks, SQL injection, malware, and privilege escalation [15]. Table I summarizes these attacks and their potential impacts, emphasizing the need for effective AI-based detection and defense mechanisms.

TABLE I. Types of Cyberattacks in Healthcare Systems

Type of Cyberattack	Description
Denial-of-Service (DoS)	Floods a network with excessive traffic, disrupting access to healthcare systems and critical network resources.
Man-in-the-Middle (MITM) [16]	Intercepts communication between two parties, potentially exposing sensitive healthcare information and compromising data integrity.
Phishing [17]	Uses fraudulent emails or websites to trick healthcare staff or patients into revealing sensitive information or downloading malicious software.
Cryptographic Attack [18]	Attempts to compromise cryptographic protections to access confidential information or recover encrypted data.
SQL Injection	Exploits database vulnerabilities to access, modify, or expose sensitive patient information, threatening data confidentiality and integrity.
Malicious Software [19](Malware)	Uses harmful programs, including viruses, worms, Trojans, spyware, and ransomware, to compromise healthcare systems and patient information.
Privilege Escalation [20]	Exploits system vulnerabilities to gain unauthorized permissions, potentially enabling attackers to access or modify patient records.

C. *Cybersecurity Threat Landscape in Healthcare*

Healthcare has become one of the most consistently targeted sectors for cyberattacks, driven by the high monetary and intelligence value of medical data and the critical requirement for uninterrupted service delivery [21].

1) *Insider Threats and Human Error*

Insider events in healthcare can include both purposeful behavior and unintended mistakes [22]. Malicious insiders may purposefully exfiltrate protected health information (PHI), but inadvertent exposures are frequently caused by human error, such as sending emails to the incorrect recipients, leaving file sharing insecure, or misplacing devices are shown in Fig 1.

J. Glob. Res. Multidiscip. Studies, 2 (10):30-37
DOI: https://doi.org/xxxxxxxxxxxx

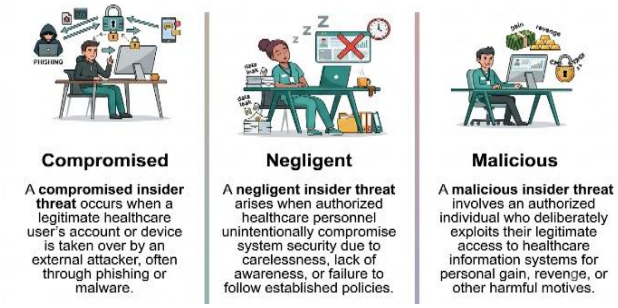


Fig. 1. Insider Threats in Healthcare Systems

- **Malicious Insiders:** authorized users who intentionally misuse their access to harm healthcare systems or organizations. In healthcare environments, such threats often involve the deliberate theft of electronic health records (EHRs), intellectual property, or research data for financial gain, espionage, or personal motives.
- **Negligent Insiders:** legitimate users who unintentionally cause security breaches due to carelessness or insufficient awareness. This category includes incidents such as lost devices, misdirected information, or improper handling of sensitive data.
- **Compromised Insiders:** Compromised insider threats occur when attackers hijack legitimate user credentials, often through phishing, malware, or credential theft. In healthcare environments, compromised accounts are particularly dangerous because attackers gain access to sensitive systems while appearing as trusted users.

2) Outsider Threats

Healthcare organizations increasingly face sophisticated external attacks, including phishing, ransomware, malware propagation, unauthorized network intrusion, and denial-of-service attacks. Such incidents are motivated by the high market value of medical information and the operational damage that can be inflicted by disabling clinical systems. The threats listed below represent common types of outsider attacks in healthcare systems.

D. Vulnerabilities in Healthcare Systems

Healthcare systems present vulnerabilities that can be broadly grouped into technical, operational, and organizational categories. Technical vulnerabilities are among the most visible [23]. Outdated operating systems, weak encryption protocols, and poorly configured firewalls create easy entry points for attackers the heavy reliance on legacy equipment, particularly in smaller hospitals, exacerbates these risks since devices are often incompatible with modern patches or security upgrades.

1) Outdated Medical Devices with Unpatched Software

Healthcare environments rely heavily on legacy medical devices such as infusion pumps, imaging systems, and patient monitors that often run on outdated operating systems and firmware [24]. Many clinical devices, such as imaging systems, infusion pumps, patient monitors, and laboratory equipment, rely on legacy operating systems that cannot be easily updated due to vendor restrictions, certification requirements, or the risk of disrupting clinical operations.

2) Network Misconfigurations

Network misconfigurations represent a critical vulnerability in healthcare systems, arising from improper setup of firewalls, access controls, segmentation policies, and communication protocols [25]. Common issues include overly permissive access rules, exposed services, default credentials, and insufficient network segmentation between clinical, administrative, and IoT device networks.

3) Default System Credentials

The continued use of default or weak system credentials remains a critical security weakness in healthcare environments. Many medical devices, network equipment, and healthcare applications are deployed with vendor-provided default usernames and passwords that are rarely changed due to operational oversight or usability concerns [26]. Attackers commonly exploit these credentials through automated scanning and brute-force techniques to gain unauthorized access.

4) Insecure Firmware

The insecurity of firmware is one major threat within the healthcare ecosystem because of its low-level nature and role in controlling the functions of medical equipment and network devices [27]. In cases where the firmware is poorly written, outdated, unsigned, and without integrity checks, an attacker can easily take advantage of it to get into devices.

5) Legacy Systems Lacking Modern Security Controls

Legacy healthcare systems often operate on outdated platforms that lack contemporary security mechanisms such as strong encryption, fine-grained access control [28], continuous monitoring, and compatibility with modern authentication protocols. These systems are frequently retained due to high replacement costs, regulatory constraints, or dependence on specialized clinical workflows. However, their limited security capabilities make them attractive targets for cyber adversaries.

III. AI-Based Threat Detection Techniques for Healthcare Cybersecurity

The AI-based methods for threat detection enable healthcare companies to detect cyber threats by analyzing network traffic, user activities, and other events in the system. Such methods as machine learning, deep learning, anomaly detection, and explainable AI provide the ability to detect intrusions, phishing attacks, and unusual access to the system.

A. AI and ML in Cybersecurity

The cybersecurity analysis process is greatly enhanced by AI and ML. It can analyze server logs, network traffic, and other disparate data points from an organization's digital infrastructure to reveal hidden dangers [29]. Most importantly, it can process this data more meticulously and within a shorter time than human analysts could by applying complex formulas.

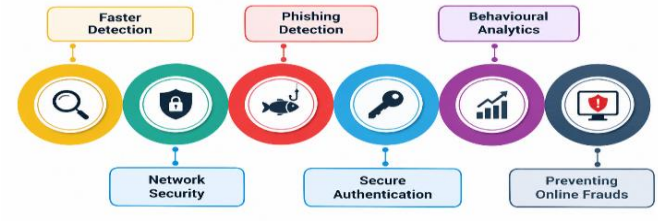


Fig. 2. AI in Cybersecurity

J. Glob. Res. Multidiscip. Studies, 2 (10):30-37
DOI: https://doi.org/xxxxxxxxxxxx

Fig 2 outlines the key areas where artificial intelligence (AI) is being used to enhance cybersecurity. Here's a breakdown of each element:

- **Faster Detection:** The real-time analysis of such large data sets, as the capability of most AI algorithms ensures the quick identification of potential risks like threats, weaknesses, diseases, among others. This speed is important in order to intercept attacks before they begin to inflict a lot of harm.
- **Phishing Detection:** Machine learning is capable of understanding the body of an email and the behavior of senders to flag phishing emails [30]. Phishing attempts may have a much lower success rate because of this.
- **Behavioral Analytics:** AI can monitor a user's behaviors and look for changes that may suggest that the system has been compromised. For instance, attempts to log in, or any pattern of data that needs more scrutiny could be flagged.
- **Network Security:** The traffic of networks could be monitored and analyzed by an AI to look for anomalies – more specifically, events such as attempts at unauthorized access or acts of malware. This helps in the battle against different forms of cyber threats in a network [31].
- **Secure Authentication:** Applying biometric data, for instance, fingerprints, face, or recognizing behavior patterns, AI might enhance the methods of user authentication and control the identity of a user. This could enhance security by often reaching it unprohibited users to achieve pedestal to systems.
- **Preventing Online Frauds:** AI can receive transaction information, identify certain trends, and inform about possible fraud, thus aiding in the detection of online fraud. Companies and individuals could gain from it since it lessens the chances of losing money.

B. AI-Based Threat Detection in Healthcare

AI provides unique capabilities for identifying and mitigating cyber threats that traditional methods cannot match. As cyber threats to healthcare systems have evolved in sophistication, the use of effective cyber analysis has become the bread and butter of robust cybersecurity practices [32]. Key AI technologies employed in healthcare data security include:

1) Predictive Maintenance and Anomaly Detection

The use of AI in healthcare facilities for anomaly detection and predictive maintenance [33]. Using ML, scientists have developed models that can foresee when machinery breaks down and identify any suspicious patterns in energy use.

2) Natural Language Processing (NLP)

NLP is another branch of AI that compels computers to understand and even read human speech. Applications in the healthcare sector include medical research, clinical documentation, and correspondence with the patient. Further, NLP may be adopted in developing artificial assistants such as virtual assistants and chatbots in the management of patients' health.

3) Machine Learning

It depends on inference and patterns from data. Using historical patient data, ML algorithms are used in forecasting their prognosis and disease evolution alongside the response to treatment so as to guide treatment and resource utilization.

IV. AI-Driven Defense Mechanisms and Secure Healthcare Architectures

The use of defense mechanisms with the assistance of AI provides the healthcare industry with increased security through the ability to detect threats automatically, secure access controls, encryption, and continuous monitoring. Some of these technologies include zero trust architecture, federated learning, and security information and event management systems [34].

A. AI-Integrated Healthcare Security Architecture

AI-enabled architecture for securing a healthcare system uses intelligence in monitoring, access control, encryption, and threat detection to enhance the security of the digital healthcare system infrastructure. AI-driven security architecture for the healthcare system links the networks of hospitals, clouds, medical devices, and patients' databases through secure communication channels.

1) Healthcare Cloud System Architecture

As infrastructure grows, security vulnerabilities increase with time. Cloud regulatory compliance and data privacy aren't accessible on the platform [35]. They need strict adherence to ethical security practices to provide a secure environment for data storage and code running, especially in healthcare systems. Fig 3 describes a general Architecture of a Secure Cloud Healthcare Architecture with different interconnection components.

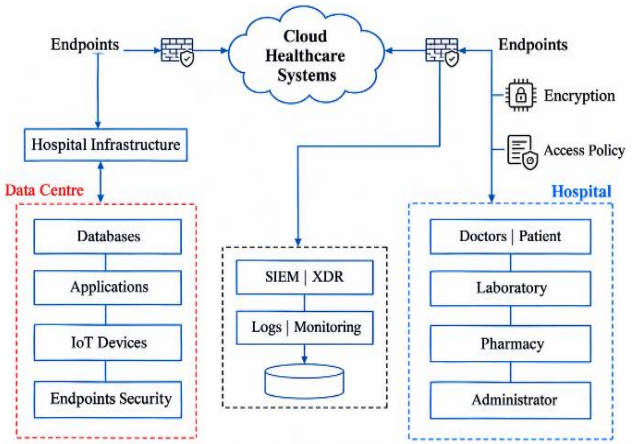


Fig. 3. Cloud Healthcare Secure Architecture

2) CIA Model for Healthcare Systems

Cloud infrastructure systems are increasingly popular among various innovative information technology companies as they adopt cloud policies [36]. These systems provide valuable and productive computing resources, delivering computation as a service with minimal effort and cost. Organizations are transitioning from traditional to web-based applications through a subscription-based approach, hosting services remotely. The CIA (Confidentiality, Integrity, Availability) model is critical in assessing and improving security in cloud-based enterprise systems, as shown in Fig 4. In healthcare environments where data confidentiality, integrity, and availability are paramount, understanding and implementing the CIA model is critical to protecting patient information and ensuring uninterrupted access to critical services.

J. Glob. Res. Multidiscip. Studies, 2 (10):30-37
DOI: https://doi.org/xxxxxxxxxxxxx

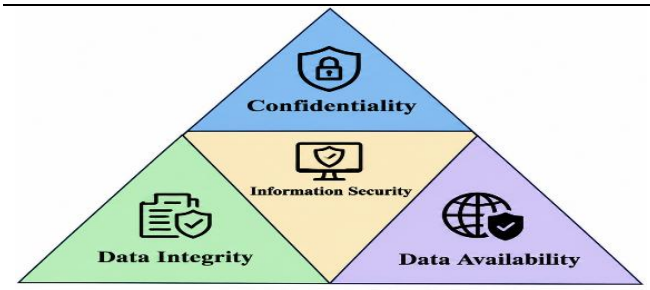


Fig. 4. Enterprise Cloud Infrastructure Systems Using the CIA Model

- **Data Confidentiality:** Data confidentiality has many implications, and data security is a mix of confidentiality, preventing unauthorized access to data, illegal changes or deletion of data, and preventing unlawful data retention [37].
- **Data Integrity:** An enterprise cloud system that uses the CIA model ensures data integrity, which is the most critical element of an information system [38]. Data integrity generally means protecting data from illegal deletion, alteration, or falsification.
- **Data availability:** Data availability refers to how users can retrieve and verify data [39]. Users can locate and access data effortlessly, which enhances their trust in the Cloud infrastructure system.

B. Privacy Preserving Techniques

Developing reliable healthcare AI systems using ML for different clinical tasks requires large amounts of carefully curated data [40]. A key barrier in the path of broader AI adoption is data access and its implementation at the commercial level. PPML ameliorates privacy techniques and infrastructure for improved security for data and ML, which is of utmost importance to overcome these barriers.

1) Cryptographic Techniques

The word cryptography comes from the Greek word cryptos, which means ‘‘hidden’’. Cryptography studies communication methods or strategies that assure secure data delivery from the sender to the recipient [41]. The content is safely shared between the sender and the recipient while maintaining its confidentiality and integrity. The primary mechanism employed in this strategy for privacy preservation is encryption.

- **Homomorphic encryption:** The technique in which the data owner encrypts its data and sends it for computation. The computational tasks are performed on the data without decrypting the data, and the output results of the encrypted data are sent to the data owner.
- **Secure Multiparty Computation (SMPC):** Cryptography that performs data computation by distributing the data between different parties [42]. Each party applies the algorithm to its secure data without knowing the rest of the data, resulting in privacy preservation.
- **Garbled circuits (GC):** A garbled circuit is a cryptographic approach that allows two distrusting parties to jointly evaluate a function over their private inputs without the requirement for a third party. Using the garbled circuit protocol, the function must be written as a Boolean circuit:
- **Secret sharing:** A method of disseminating information among a group of individuals, with each person receiving a piece of the information. Individual shares are useless on their own; thus, the secret can only be reconstructed when a huge number of them, maybe of diverse kinds, are united together [43].
- **Differential privacy (DP):** Privacy-preserving technique that provides data anonymity by introducing noise into the data [44]. All the communication analysis is done on the anonymous data without any client’s identity revelation.
- **Federated learning (FL):** Combined ML model based on data from different locations. In FL, there are two processes: model training and model inference. Information, but not data, can be transferred between parties throughout the model training process [45].

V. Literature of Review

Recent studies explore AI-driven healthcare cybersecurity using machine learning, deep learning, transfer learning, blockchain, and generative AI. These approaches address intrusion detection, anomaly identification, patient data protection, and automated responses. However, integration, scalability, privacy preservation, and real-world validation remain important challenges for effective healthcare security systems.

G. Akrap et al. (2026) explored the dual nature of AI in medicine: as both an enabler of progress and a potential vector of systemic risk, through the lenses of information and cybersecurity, intelligence, and resilience. It examines the technological and organizational dimensions of these challenges by jointly analyzing documented AI-enabled clinical infrastructures, data flows, and security controls alongside governance structures, institutional responsibilities, and human factors shaping system resilience [46].

R. Kant et al. (2026) provided an elaborate AI-based cybersecurity framework tailored to healthcare environment in particular. Based on real-world breach intelligence data provided by the U.S. Department of Health and Human Services (HHS), the Verizon Data Breach Investigations Report (DBIR), and the IBM XForce Threat Intelligence Index, study incidences of attacks and assess AI-enhanced intrusion detections, anomaly detections, and automated response systems. The suggested framework combines machine learning, deep learning, governance controls, and explainable AI to achieve the compliance of the regulations and the trust of the operations [47].

N. Neoaz, et al. (2025) The roles of AI systems in detecting threats in real-time, predictive analytics, and auto-response systems, as well as real-world examples of breaches and breakthroughs. Ethics are dealt with, including data privacy, bias, and transparency, as well as up-and-coming tendencies in resilient cyber-physical systems and cross-sector cooperation. The adoption of AI is presented not only as one of the defensive mechanisms, but rather, as a strategically valuable means to construct adaptive, intelligent, and ethically responsible security systems [48].

T. Ali et al. (2025) explored the integration of various artificial intelligence (AI) models and techniques to strengthen cybersecurity defenses in healthcare systems. By leveraging a

DOI: https://doi.org/xxxxxxxxxxxxx

multimodal approach that combines machine learning, deep learning, and anomaly detection algorithms, the method aims to significantly improve the speed and accuracy of detecting cyber threats in healthcare networks. The fusion of these diverse AI models enables the creation of a robust system capable of providing comprehensive protection against emerging threats, particularly those targeting sensitive healthcare data and infrastructure [49].

Z.Teo et al. (2024) explore the role of GenAI in healthcare, including clinical, operational, and research applications, and delve into the cybersecurity risks of this technology. Discuss risks such as data privacy risks, data poisoning attacks, the propagation of bias, and hallucinations. Recommend risk mitigation strategies to enhance cybersecurity in GenAI technologies and further explore the use of GenAI as a tool in itself to enhance cybersecurity across the various AI algorithms. GenAI is emerging as a pivotal catalyst across various industries, including the healthcare domain [50].

C.Chakraborty et al. (2023) proposed a cyber-attack detection system for the healthcare sector with centralized and federated transfer learning modes. Edge of Things (EoT) framework is developed in connection with the cloud and healthcare sectors to transmit the data efficiently, and the proposed Centralized with Multi-Source Transfer Learning (CMTL) algorithm is used for detection and classification of various threats such as information gathering, DoS/DDoS attacks, Malware attacks, Injection attacks, and man-in-the-middle attacks [51].

M. Alshehri, et al. (2023) The healthcare industry with blockchain-integrated cyber-security based on artificial intelligence (BICS-AI) in medical care to preserve medical-related things. Blockchain applications have the potential to consistently identify the most severe, potentially life-threatening mistakes in the medical field. The use of blockchain for decentralized data protection helps to protect patient health records from compromise. With the help of an access control provider (ACP), here came up with a lightweight solution that addresses this issue by allowing the delegation of security operations [52].

Despite advances in AI-based healthcare cybersecurity, several research gaps remain in developing reliable and scalable protection systems. Existing studies investigate machine learning, deep learning, transfer learning, blockchain, and generative AI, but their integration into a unified threat detection and defense framework remains limited. Many approaches lack comprehensive evaluation across hospital networks, cloud platforms, and medical IoT devices. Challenges related to adversarial attacks, data privacy, model interpretability, false alarms, and computational overhead also require further investigation. The comparative findings summarized in Table II highlight these limitations and emphasize the need for adaptive security solutions combining intelligent threat detection, explainable AI, privacy-preserving learning, and automated incident response. Future research should focus on real-world validation, cross-institutional threat intelligence sharing, and robust evaluation against evolving cyberattacks to improve healthcare cybersecurity.

TABLE II. SUMMARY OF KEY STUDIES ON ARTIFICIAL INTELLIGENCE FOR HEALTHCARE CYBERSECURITY

Author & Year	Objective / Approach	Technology / Method	Key Findings	Advantages	Limitations
G. Akrap et al. (2026)	Examined AI benefits and cybersecurity risks in healthcare.	AI-enabled clinical infrastructure, security controls, governance, and resilience.	Identified technological, organizational, and human factors affecting healthcare cybersecurity.	Covers security and governance challenges comprehensively.	Specific performance metrics and experimental validation are not reported in the provided summary.
R. Kant et al. (2026)	Developed an AI-based cybersecurity framework for healthcare.	ML, DL, intrusion detection, anomaly detection, and XAI.	Integrated threat detection, automated response, and governance controls using breach intelligence data.	Combines detection, response, explainability, and compliance.	Quantitative evaluation and implementation details are not specified in the provided summary.
N. Neoaz et al. (2025)	Investigated AI for real-time threat detection and predictive security.	AI-based detection, predictive analytics, and automated response.	Highlighted adaptive cybersecurity and concerns involving privacy, bias, and transparency.	Addresses emerging threats and ethical considerations.	Specific datasets, algorithms, and evaluation metrics are not provided in the summary.
T. Ali et al. (2025)	Explored combined AI techniques to improve healthcare threat detection.	ML, DL, and anomaly detection.	Described multimodal AI for improving the speed and accuracy of detecting healthcare cyber threats.	Combines complementary techniques for broader threat detection.	Numerical results and detailed experimental validation are not specified in the summary.
Z. Teo et al. (2024)	Examined healthcare GenAI applications and cybersecurity risks.	Generative AI, privacy protection, and risk mitigation strategies.	Identified data privacy, poisoning, bias, and hallucination risks while discussing AI-assisted security.	Covers both GenAI vulnerabilities and defensive opportunities.	Does not establish the performance of a dedicated healthcare threat detection system in the supplied summary.
C. Chakraborty et al. (2023)	Proposed a healthcare cyberattack detection system using transfer learning.	Centralized and federated transfer learning, EoT, cloud integration, and CMTL.	Addressed information gathering, DoS/DDoS, malware, injection, and man-in-the-middle attacks.	Supports multiple attack categories and distributed learning approaches.	Scalability and comparative performance metrics are not specified in the supplied summary.
M. Alshehri et al. (2023)	Proposed blockchain-integrated AI cybersecurity for healthcare data protection.	Blockchain, AI-based security, and access control provider (ACP).	Addressed decentralized protection of patient records and delegated security operations.	Supports access control and distributed data protection.	Quantitative security improvements and interoperability results are not specified in the supplied summary.

VI. Conclusion and Future Work

Healthcare cybersecurity has become increasingly important with the growing adoption of digital health records, cloud platforms, and connected medical devices. This review examines major cybersecurity threats, including ransomware, phishing, malware, insider attacks, supply-chain risks, and vulnerabilities in legacy systems. It discusses AI-based threat detection techniques, including machine learning, deep learning, behavioral analytics, and natural language processing, for identifying suspicious activities and potential attacks. Furthermore, AI-integrated security architectures, cloud security, the CIA model, and zero-trust principles provide important mechanisms for protecting healthcare infrastructure. Privacy-preserving techniques, such as homomorphic encryption, secure multiparty computation, differential privacy, and federated learning, support secure data processing while reducing privacy risks. However, challenges involving limited data availability, outdated devices, computational costs, and implementation complexity remain significant. Overall, integrating AI with robust security controls can strengthen threat detection, protect sensitive patient information, and improve healthcare system resilience. This review highlights the importance of developing reliable, scalable, and privacy-aware cybersecurity solutions for modern healthcare environments.

Future research should focus on developing robust and explainable AI models for real-time healthcare threat detection. Integrating federated learning, privacy-preserving cryptography, and zero-trust architectures may improve security across distributed healthcare networks. Further studies should evaluate these approaches using real-world datasets, diverse medical devices, and evolving attack scenarios while addressing scalability, computational efficiency, regulatory compliance, and adversarial robustness.

- **Funding:** This research received no external funding.
- **Institutional Review Board Statement:** Not applicable.
- **Informed Consent Statement:** Not applicable.
- **Data Availability Statement:** The data supporting the findings of this study are available from the corresponding author upon reasonable request.
- **Conflicts of Interest:** The author declares no conflict of interest.

How to cite this article:
N. Upadhyay, "Artificial Intelligence for Healthcare Cybersecurity: A Comprehensive Survey of Threat Detection and Defense Mechanisms," *Journal of Global Research in Multidisciplinary Studies*, vol. 2, no. 10, pp. 30–37, Oct. 2026, Doi XXXX

References

[1] G. Methari, Y. Hayat, and A. M. Khaja, "Artificial Intelligence for Cybersecurity in Healthcare Systems: A Simple Review of Applications, Challenges, and Future Directions," *Int. J. Innov. Res. Comput. Sci. Technol.*, vol. 13, no. 6, pp. 37–47, 2025, doi: 10.55524/ijircst.2025.13.6.5.

[2] A. R. A. Khan, M. I. Khan, A. Arif, N. Anjum, and H. Arif, "Intelligent Defense: Redefining OS Security with AI," *Int. J. Innov. Res. Comput. Sci. Technol.*, vol. 13, no. 1, pp. 85–90, 2025, doi: 10.55524/ijircst.2025.13.1.13.

[3] A. Raza, "Equality before Law and Equal Protection of Law: Contextualizing its Evolution in Pakistan," *SSRN Electron. J.*, pp. 1–17, February, 2025, doi: 10.2139/ssrn.5200799.

[4] A. Aziz, B. R., S. R. Kurakula, A. R. Mohammad, S. S. Basha, and S. Katta, "Federated and Explainable AI Models for Privacy-Preserving Predictive Analytics in Healthcare System," in *2026 International Conference on AI-Driven Smart Systems and Ubiquitous Computing (ICAUC)*, Pathum Thani, Thailand: IEEE, Jan. 2026, pp. 747–753. doi: 10.1109/ICAUC68182.2026.11441162.

[5] A. Raza, "Credit, Code, and Consequence: How AI Is Reshaping Risk Assessment and Financial Equity," *Euro Vantage J. Artif. Intell.*, vol. 2, no. 2, pp. 79–86, 2025, doi: 10.65923/tqt7fj71.

[6] A. L. Neves and J. Burgers, "Digital technologies in primary care: Implications for patient care and future research. 92nd EGPRN Meeting, Virtual Conference, 30 May-1 April," *Eur. J. Gen. Pract.*, vol. 28, no. 1, pp. 203–208, 2022, doi: 10.1080/13814788.2022.2052041.

[7] D. Giansanti, "Ten Years of TeleHealth and Digital Healthcare: Where Are We?," *Healthc.*, vol. 11, no. 6, 2023, doi: 10.3390/healthcare11060875.

[8] M. S. Ibrahim, H. Mohamed Yusoff, Y. I. Abu Bakar, M. M. Thwe Aung, M. I. Abas, and R. A. Ramli, "Digital health for quality healthcare: A systematic mapping of review studies," *Digit. Heal.*, vol. 8, p. 205520762210858, Jan. 2022, doi: 10.1177/20552076221085810.

[9] J. -, A. -, and N. -, "Cybersecurity Threat Landscape in the US Healthcare Sector: Trends, Risks, and National Implications," *Int. J. Technol. Appl. Sci.*, vol. 17, no. 1, pp. 1–12, Jan. 2026, doi: 10.71097/IJTAS.v17.i1.1157.

[10] J. X. Jiang, J. S. Ross, and G. Bai, "Ransomware Attacks and Data Breaches in US Health Care Systems," *JAMA Netw. Open*, vol. 8, no. 5, pp. 1–4, May 2025, doi: 10.1001/jamanetworkopen.2025.10180.

[11] J. Mulo *et al.*, "Navigating Challenges and Harnessing Opportunities: Deep Learning Applications in Internet of Medical Things," *Futur. Internet*, vol. 17, no. 3, pp. 1–48, 2025, doi: 10.3390/fi17030107.

[12] B. Avanzi, X. Tan, G. Taylor, and B. Wong, "On the Evolution of Data Breach Reporting Patterns and Frequency in the United States: A Cross-State Analysis," *North Am. Actuar. J.*, vol. 29, no. 4, pp. 833–864, Oct. 2025, doi: 10.1080/10920277.2025.2457491.

[13] A. Y. Balogun, O. O. Olaniyi, A. O. Olisa, M. O. Gbadebo, and N. C. Chinye, "Enhancing Incident Response Strategies in U.S. Healthcare Cybersecurity," *J. Eng. Res. Reports*, vol. 27, no. 2, pp. 114–135, Jan. 2025, doi: 10.9734/jerr/2025/v27i21399.

[14] S. Pawar, G. Patil, K. Patel, P. Pawar, S. Khedkar, and B. More, "Falsified News Detection Using Deep Learning Approach," in *2021 Asian Conference on Innovation in Technology (ASIANCON)*, PUNE, India: IEEE, Aug. 2021, pp. 1–5. doi: 10.1109/ASIANCON51346.2021.9544585.

[15] N. D. Bhandarwar, "AI-Powered Cybersecurity Models for Fraud Detection in IoT-Enabled Smart Healthcare Devices," in *2026 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, Gwalior, India: IEEE, Mar. 2026, pp. 1–6. doi: 10.1109/IATMSI68868.2026.11466129.

[16] B. Pingle, A. Mairaj, and A. Y. Javaid, "Real-World Man-in-the-Middle (MITM) Attack Implementation Using Open Source Tools for Instructional Use," in *2018 IEEE International Conference on Electro/Information Technology (EIT)*, IEEE, May 2018, pp. 0192–0197. doi: 10.1109/EIT.2018.8500082.

[17] M. F. Ansari, B. Dash, P. Sharma, and N. Yathiraju, "The Impact and Limitations of Artificial Intelligence in Cybersecurity: A Literature Review," *Ijarcce*, vol. 11, no. 9, pp. 81–90, 2022, doi: 10.17148/ijarcce.2022.11912.

[18] S. G. Gohwong, "The State of the Art of Cryptography-Based Cyber-Attacks," 2020, *SSRN*. doi: 10.2139/ssrn.3546334.

[19] H. Sinanovic and S. Mrdovic, "Analysis of Mirai malicious software," in *2017 25th International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, IEEE, Sep. 2017, pp. 1–5. doi: 10.23919/SOFTCOM.2017.8115504.

[20] M. Mehmood, R. Amin, M. M. A. Muslam, J. Xie, and H. Aldabbas, "Privilege Escalation Attack Detection and Mitigation in Cloud Using Machine Learning," *IEEE Access*, vol. 11, pp. 46561–46576, 2023, doi: 10.1109/ACCESS.2023.3273895.

[21] R. Qureshi and I. Koo, "A Comprehensive Survey of

Cybersecurity Threats and Data Privacy Issues in Healthcare Systems," *Appl. Sci.*, vol. 16, no. 3, 2026, doi: 10.3390/app16031511.

[22] P. Tabari, M. De Rosa, G. Costagliola, and V. Fuccella, "Invisible Threats: Rethinking Privacy in Digital Healthcare," in *Computational Science and Its Applications – ICCSA 2025 Workshops: Istanbul, Turkey, June 30 – July 3, 2025, Proceedings, Part VIII*, Berlin, Heidelberg: Springer-Verlag, 2025, pp. 267–284. doi: 10.1007/978-3-031-97645-2_18.

[23] B. O. Owolabi, "Exploring Systemic Vulnerabilities in Healthcare Digital Ecosystems Through Risk Modeling, Threat Intelligence, and Adaptive Security Control Mechanisms," *Int. J. Comput. Appl. Technol. Res.*, vol. 11, no. 12, pp. 687–699, 2025, doi: 10.7753/ijcatr1112.1029.

[24] A. Sharma and Gagandeep, "Securing the future of healthcare: Ensuring privacy and security in wearable and implantable medical devices," in *Security, Privacy, and Trust in WBANs and E-Healthcare*, Boca Raton: CRC Press, 2024, pp. 105–125. doi: 10.1201/9781032635101-8.

[25] S. Islam, S. Papastergiou, E. M. Kalogeraki, and K. Kioskli, "Cyberattack Path Generation and Prioritisation for Securing Healthcare Systems," *Appl. Sci.*, vol. 12, no. 9, 2022, doi: 10.3390/app12094443.

[26] M. K. Hasan *et al.*, "A review on security threats, vulnerabilities, and counter measures of 5G enabled Internet-of-Medical-Things," *IET Commun.*, vol. 16, no. 5, pp. 421–432, 2022, doi: 10.1049/cmu2.12301.

[27] S. Ul Haq, Y. Singh, A. Sharma, R. Gupta, and D. Gupta, "A survey on IoT & embedded device firmware security: architecture, extraction techniques, and vulnerability analysis frameworks," *Discov. Internet Things*, vol. 3, no. 1, p. 17, 2023, doi: 10.1007/s43926-023-00045-2.

[28] M. Khan *et al.*, "Modernization framework to enhance the security of legacy information systems," *Intell. Autom. Soft Comput.*, vol. 32, no. 1, pp. 543–555, 2022, doi: 10.32604/iasc.2022.016120.

[29] T. Petchiappan, "Ai-Driven Solutions for Enhancing Cybersecurity in Healthcare Systems: a Comprehensive Review," *Int. J. Core Eng. Manag.*, vol. 12, no. 7, pp. 1–8, 2024.

[30] B. S. Shaikh and H. Homayouni, "A Computer Vision Approach to Visual Fraud Detection in Phishing Websites Using YOLOv8," Oct. 2026. doi: 10.48550/arXiv.2609.34015.

[31] Y. Muni, "Advanced Multi-Protocol Label Switching (MPLS)-Enabled Networks: A Survey of Emerging Approaches Via Traffic Engineering," *Eng. Technol. J.*, vol. 11, no. 01, pp. 8605–8613, Jan. 2026, doi: 10.47191/etj/v11i01.22.

[32] S. Arefin, "Strengthening Healthcare Data Security with Ai-Powered Threat Detection," *Int. J. Sci. Res. Manag.*, vol. 12, no. 10, pp. 1477–1483, 2024, doi: 10.18535/ijsrm/v12i10.ec02.

[33] V. Dinesh, R. Kalli, and E. Jonathan, "AI-Driven Energy Management Solutions for Healthcare: Optimizing Medical Device Software," 2023.

[34] B. P. Singh, "Convergence of AI and Zero Trust: Enabling Continuous Verification Across Hybrid Cloud Environments," *Int. J. Intell. Syst. Appl. Eng.*, vol. 14, no. 1, p. 339, Apr. 2026, doi: 10.17762/ijisae.v14i1s.8181.

[35] V. Bier and A. Gutfraind, "Risk analysis beyond vulnerability and resilience – characterizing the defensibility of critical systems," *Eur. J. Oper. Res.*, vol. 276, no. 2, pp. 626–636, Jul. 2019, doi: 10.1016/j.ejor.2019.01.011.

[36] A. Qureshi, W. Dashti, A. Jahangeer, and A. Zafar, "Security Challenges over Cloud Environment from Service Provider Prospective," *Cloud Comput. Data Sci.*, pp. 12–20, Apr. 2020, doi: 10.37256/ccds.112020318.

[37] A. K. M. B. Haque, B. Bhushan, and G. Dhiman, "Conceptualizing smart city applications: Requirements, architecture, security issues, and emerging trends," *Expert Syst.*, vol. 39, no. 5, p. e12753, Jun. 2022, doi: 10.1111/exsy.12753.

[38] A. H. Mohsin *et al.*, "Based Medical Systems for Patient's Authentication: Towards a New Verification Secure Framework Using CIA Standard," *J. Med. Syst.*, vol. 43, no. 7, p. 192, Jul. 2019, doi: 10.1007/s10916-019-1264-y.

[39] K. A. Torkura, M. I. H. Sukmana, F. Cheng, and C. Meinel, "CloudStrike: Chaos Engineering for Security and Resiliency in Cloud Infrastructure," *IEEE Access*, vol. 8, pp. 123044–123060, 2020, doi: 10.1109/ACCESS.2020.3007338.

[40] S. Al-Kuwari, "Privacy-Preserving AI in Healthcare," in *Multiple Perspectives on Artificial Intelligence in Healthcare: Opportunities and Challenges*, M. Househ, E. Borycki, and A. Kushniruk, Eds., Cham: Springer International Publishing, 2021, pp. 65–77. doi: 10.1007/978-3-030-67303-1_6.

[41] E. Sarkar, E. Chielle, G. Gürsoy, O. Mazonka, M. Gerstein, and M. Maniatakos, "Fast and Scalable Private Genotype Imputation Using Machine Learning and Partially Homomorphic Encryption," *IEEE Access*, vol. 9, pp. 93097–93110, 2021, doi: 10.1109/ACCESS.2021.3093005.

[42] T. Dugan and X. Zou, "A Survey of Secure Multiparty Computation Protocols for Privacy Preserving Genetic Tests," in *2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, 2016, pp. 173–182. doi: 10.1109/CHASE.2016.71.

[43] P. Sarosh, S. A. Parah, G. M. Bhat, A. A. Heidari, and K. Muhammad, "Secret Sharing-based Personal Health Records Management for the Internet of Health Things," *Sustain. Cities Soc.*, vol. 74, p. 103129, Nov. 2021, doi: 10.1016/j.scs.2021.103129.

[44] A. Vadavalli and R. Subhashini, "An Improved Differential Privacy-Preserving Truth Discovery approach In Healthcare," in *2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, 2019, pp. 1031–1037. doi: 10.1109/IEMCON.2019.8936141.

[45] T. S. Brisimi, R. Chen, T. Mela, A. Olshevsky, I. C. Paschalidis, and W. Shi, "Federated learning of predictive models from federated Electronic Health Records.," *Int. J. Med. Inform.*, vol. 112, pp. 59–67, Apr. 2018, doi: 10.1016/j.ijmedinf.2018.01.007.

[46] G. Akrap, I. Dumić-Čule, N. Parlov, and S. Jandroković, "Medicine in the Age of Artificial Intelligence: Cybersecurity, Hybrid Threats and Resilience," *Appl. Sci.*, vol. 16, no. 4, 2026, doi: 10.3390/app16041736.

[47] R. Kant, R. Rao Thallada, B. Pandey, and P. Srivastava, "AI-Based Cybersecurity in Healthcare: A Data-Driven, Governance-Aware Framework for Secure Clinical Systems," *2026 IEEE 5th Int. Conf. AI Cybersecurity, ICAIC 2026*, no. February, pp. 18–20, 2026, doi: 10.1109/ICAIC67076.2026.11395836.

[48] N. Neoaz, "Harnessing Artificial Intelligence for Cybersecurity in Healthcare and Food Processing: A Review of Emerging Trends and the Role of Generative Models like ChatGPT," *Glob. Trends Sci. Technol.*, vol. 1, no. 3, pp. 144–162, Jul. 2025, doi: 10.70445/gtst.1.3.2025.144-162.

[49] T. E. Ali, F. I. Ali, F. Eyvazov, and A. D. Zoltán, "Integrating AI Models for Enhanced Real-Time Cybersecurity in Healthcare: A Multimodal Approach to Threat Detection and Response," *Procedia Comput. Sci.*, vol. 259, pp. 108–119, 2025, doi: 10.1016/j.procs.2025.03.312.

[50] Z. L. Teo, C. W. N. Quek, J. L. Y. Wong, and D. S. W. Ting, "Cybersecurity in the generative artificial intelligence era," *Asia-Pacific J. Ophthalmol.*, vol. 13, no. 4, pp. 1–6, July–August, Jul. 2024, doi: 10.1016/j.apjo.2024.100091.

[51] C. Chakraborty, S. M. Nagarajan, G. G. Devarajan, T. V Ramana, and R. Mohanty, "Intelligent AI-based Healthcare Cyber Security System using Multi-Source Transfer Learning Method," *ACM Trans. Sens. Networks*, pp. 1–16, May 2023, doi: 10.1145/3597210.

[52] M. Alshehri, "Blockchain-assisted cyber security in medical things using artificial intelligence," *Electron. Res. Arch.*, vol. 31, no. 2, pp. 708–728, 2023, doi: 10.3934/era.2023035.