Scientific Research Publishing

E-Health Telecommunication Systems and Networks, 2025, 14(4), 71-86
https://www.scirp.org/journal/etsn
ISSN Online: 2167-9525
ISSN Print: 2167-9517

A4 AI-Guard: Advancing AI Defense for Cybersecurity in Healthcare

Mostafa Rahmany

Cybersecurity Researcher, Dubai, UAE
Email: GlobalMostafa@gmail.com

How to cite this paper: Rahmany, M. (2025) A4 AI-Guard: Advancing AI Defense for Cybersecurity in Healthcare. *E-Health Telecommunication Systems and Networks*, **14**, 71-86.
https://doi.org/10.4236/etsn.2025.144007

Received: September 12, 2025
Accepted: October 21, 2025
Published: October 24, 2025

Abstract

Healthcare cyberattacks rise as attackers leverage system and human vulnerabilities. The existing security systems are technology-driven as they focus on system resilience, but they neglect human errors, which are the weakest link for cybersecurity in healthcare. A4 AI-Guard is an artificial intelligence-powered adaptive threat defense system which links AI systems to human security methods for the improvement of healthcare cybersecurity. The four adaptive pillars of the A4 AI-Guard system are, **Adaptive Behavioral Risk Profiling** for Adaptive Analysis, which monitors user activities to identify threats in advance, and **Active Threat Detection and Response**, which applies AI for cyber threat identification and prevention in real time, and **Augmented Role-Based Training**, which provides continuous EHR training to healthcare workers according to their role, and lastly, **Adaptive Cultural Reinforcement**, which uses AI reproducibility knowledge to help healthcare organizations build and maintain security culture. The value of A4 AI-Guard is validated by applying the system to two recent healthcare ransomware attacks, including Change Healthcare and Universal Health Services, to show how it would improve system availability and response time and deal with human security vulnerabilities. The evaluation metrics are threefold. First, the number of cyber incidents will be reduced while the detection and response times will be reduced. Second, the employee training completion rate will increase through more engagement and reasoning and exploration, and third, the organization will develop a stronger security culture through metrics and indicators. The results show that A4 AI-Guard is an adaptive defense system which provides both proactive protection and scalability as well as human understanding to help healthcare organizations protect themselves against complex, dynamic threats. The focus of the research is on A4 AI-Guard as an artificial intelligence cybersecurity framework which protects healthcare through ransomware defense, behavioral analytics and training, and EHR and IoMT systems. Operationalization of Adaptive Cultural Re-

inforcement requires quantifiable cultural metrics. For example, AI models can mine phishing-simulation results to measure incident-reporting responsiveness, analyze trends in security-incident narratives to detect sentiment shifts, and track completion/quality of peer-to-peer security discussions. These signals become dynamic inputs to reinforcement algorithms that recommend targeted awareness campaigns or micro-learning modules, creating measurable culture-improvement loops.

Keywords

A4 AI-Guard, AI-Driven Cybersecurity, Healthcare Cybersecurity, Ransomware Mitigation, Behavioral Analytics, Role-Based Training, Electronic Health Records (EHR), Internet of Medical Things (IoMT)

1. Introduction

The healthcare industry benefits from electronic health records, telemedicine, telemonitoring, and IoMT, but these innovations have made it easier for attackers to infiltrate healthcare systems [1]. The ransomware attacks on Change Healthcare in 2024 and Universal Health Services in 2020 have shown the destructive potential of cyberattacks, which led to patient data loss, service downtime, and financial loss [2] [3]. Healthcare providers are required to follow the HIPAA compliance requirements for data privacy and security, but they have been increasingly exposed to financial risks through fines, reputational loss, and even loss of patient care quality when their systems get compromised [4].

Healthcare organizations invest heavily in cybersecurity solutions, but human-related vulnerabilities have emerged as the primary drivers of security breaches. The breach log review revealed that phishing attacks, together with credential theft, insider threats, and even security culture incidents, were the most reported type of incidents [5]. The academic frameworks CARE and C4 are human-centered frameworks for employee security, but they both lack the adaptive AI-driven learning capabilities to fight against dynamic cyber threats [6] [7]. The cognitive gap in cybersecurity human vulnerabilities exists as these models are compliance-focused and awareness-driven, and they lack AI-driven adaptive capabilities. The current cyber defense technology provides effective cybersecurity, but human vulnerabilities still present exploitable threats which require anticipatory AI-driven human-centered solutions. A study published in Sensor's journal showed how AI has been used to protect IoMT infrastructure through anomaly detection and prediction models for the prevention of healthcare data breaches [8]. The healthcare industry has been developing new applications of federated learning and blockchain technology for privacy-preserving decentralized security approaches in networked healthcare systems [9]. The recent advancements in adaptive artificial intelligence support the development of A4 AI-Guard as a human-centered security model.

The solution to this challenge requires A4 AI-Guard as a new AI-driven threat defense system, which combines artificial intelligence with human-centered security approaches. The system A4 AI-Guard operates as a unified defense system, which integrates all four components instead of using the separate methods of predictive analytics, real-time threat detection, adaptive training, and cultural reinforcement. The system is built on the foundation of four adaptive pillars which form its infrastructure.

- The system monitors user activities through Adaptive Behavioral Risk Profiling to identify unauthorized logins and suspicious data access patterns [10].
- The system leverages Artificial Intelligence for the analysis of the captured behavioral data for real-time threat detection and response to minimize the time taken to respond to a cyberattack [11].
- The system offers personalized AI-driven training to healthcare personnel members through Augmented Role-Based Training, which educates them about their specific work-related security vulnerabilities [12].
- The system uses AI to measure organizational compliance and security culture, which allows the organization to build a sense of shared and adaptive cultural responsibility for long-term cybersecurity resilience [13].

Main Objectives:

- The paper will show how artificial intelligence has been used by healthcare organizations to improve their human-centered cybersecurity systems.
- The paper will assess the predictive, preventive, and mitigating capabilities of AI-based behavioral analytics and situational awareness systems when deployed in real healthcare settings.
- The paper will explore how A4 AI-Guard would have prevented two major ransomware attacks that have taken place in healthcare facilities [14].
- The paper will design a scalable and proactive system that has AI as its core, while behavior analysis and cultural evolution act as the key elements of a comprehensive cybersecurity framework.
- A4 AI-Guard protects healthcare organizations by linking system vulnerabilities and human system vulnerabilities through its adaptive defense mechanism against dynamic threats.

2. Literature Review

Healthcare cybersecurity requires a multidisciplinary approach, which brings together technological innovation with behavioral science. The current security frameworks CARE (Cognitive Awareness and Response Environment) and C4 (Cognitive Cybersecurity Framework) are primarily focused on the implementation of technical security controls and compliance requirements. The existing models have shown limited adaptability in the face of evolving human vulnerabilities [6] [7]. Healthcare organizations are at major risk, as human elements, phishing, social engineering, and insider threats remain their weakest links [5] [12].

2.1. Cognitive/Behavioral Aspects

Studies in traditional cognitive psychology have shown that employee decision-making is impaired by overconfidence biases and confirmation biases, which result in security issues by facilitating cyber threats [6] [12]. Healthcare organizations make incorrect risk judgments due to overconfidence, while their confirmation biases make them ignore security warnings that challenge their preexisting beliefs. The recognition of these aspects is essential for building cybersecurity models which guard against human vulnerabilities in stressful healthcare situations.

2.2. AI in Cybersecurity

The machine learning field has achieved impressive advances in computational algorithms by means of cognitively mutated algorithms, and this makes this work essential in the fight against both technical, human, and other vulnerabilities. Supervised learning algorithms, which consist of decision trees, support vector machines, and random forest algorithms, utilize labeled processed data to classify threats [1]. The unsupervised learning algorithms k-means clustering and autoencoders identify new attacks by isolating patterns that are dissimilar from normal user behavior [11]. The real-time optimization of the defensive moves by A4 AI-Guard is achieved through reinforcement learning (RL), which allows adaptivity for the protection against emerging threats and targeted attacks [4] [15].

The integration of neural networks with cross-validation and parameter tuning techniques improves system performance by minimizing false positives and at the same time enables transferability to different healthcare settings [2] [7]. The predictive model based on historical attacks combines with the adaptive models based on streaming data to form a hybrid model that defends against both known and unknown threats [1] [5]. Studies show that AI systems have been used to protect IoMT infrastructure through behavioral analysis, which identifies security threats before they mature into full-blown attacks. To provide context for these AI-driven tools, major industry evaluations have benchmarked the cybersecurity performance of healthcare organizations, identifying key areas of risk that require advanced solutions [16]. The proposed deep learning-based anomaly detection frameworks for IoMT healthcare systems have shown how predictive AI has enhanced active cyber defense capabilities [17] [18].

Research into federated learning and blockchain technology has continued to attract attention as these systems have shown promise for the protection of healthcare data privacy as well as enabling distributed network security. The study done by Abbas *et al.* [9] showed from their literature review that federated learning has the potential to protect patient data and at the same time share AI knowledge between healthcare facilities, and that blockchain technology can be used for secure data exchange in decentralized IoMT networks. The knowledge base reveals the need to construct A4 systems which support both adaptive and privacy-preserving operations.

2.3. Comparative Considerations

The majority of these efforts have been mostly technology-driven or compliance-oriented, and at the same time lacking human-centered defense features with AI-driven adaptivity [13] [19]. As shown in Table 1, A4 AI-Guard provides stronger real-time threat detection compared to the other frameworks CARE and C4.

Table 1. Comparative analysis of cybersecurity frameworks [19]-[22].

Feature	*A4 AI-Guard*	*CARE and C4 Frameworks*	*Traditional Cybersecurity Systems*
Focus	Human-centric + AI-driven	Human-centric	Technology-centric
Real-Time Threat Detection	AI-driven, automatic response	Limited, manual	Reactive, post-attack
Role-Specific Training	Adaptive, tailored to job roles	General, non-specific	Generic, one-size-fits-all
Cultural Reinforcement	AI monitors and improves culture	Not emphasized	Rarely incorporated

2.4. A4 AI-Guard

To address these limitations, A4 AI-Guard is proposed as a flexible, AI-enhanced defense system designed specifically for healthcare. It integrates machine learning, reinforcement learning, and cognitive insights to bridge the gap between technology and human vulnerabilities. As we describe in Table 2, the framework is structured around four adaptive pillars [10]:

Table 2. A4 AI-Guard key pillars.

Pillar	*Description*
Adaptive Behavioral Risk Profiling	Continuous AI monitoring of user behavior to detect anomalies and insider threats.
Active Threat Detection and Response	Real-time, AI-driven identification and containment of threats, leveraging adaptive learning to prevent the execution of malicious code.
Augmented Role-Based Training	Personalized training tailored to staff roles ensures situational preparedness against relevant risks.
Adaptive Cultural Reinforcement	AI analysis of compliance, incident reports, and feedback fosters a security-first organizational culture.

The four pillars function as an interdependent ecosystem. For instance, an anomaly detected by Adaptive Behavioral Risk Profiling can generate an automated alert to Active Threat Detection and Response, which in turn invokes Augmented Role-Based Training by issuing just-in-time learning prompts to relevant staff. Simultaneously, incident outcomes feed back to Adaptive Cultural Reinforcement to update organizational risk perception and resilience metrics. This closed feedback loop ensures that learning from one pillar continuously strengthens the others [10].

As illustrated in Figure 1, the A4 AI-Guard framework operates as a closed-loop ecosystem in which each adaptive pillar reinforces the others to enhance healthcare cybersecurity.

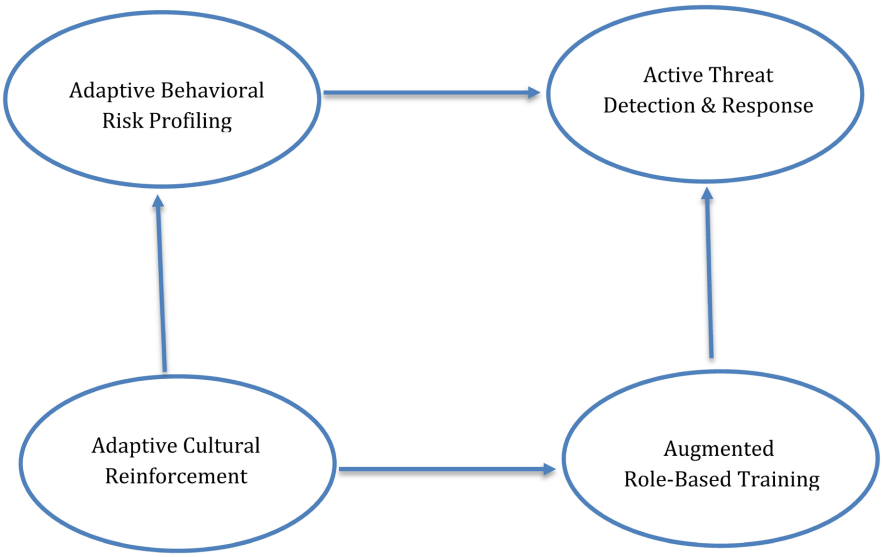

Figure 1. A4 AI Guard framework architecture.

The real-time interplay ensures that each pillar not only performs its specific role but also reinforces the others. For example, alerts from Adaptive Behavioral Risk Profiling can automatically trigger Active Threat Detection rules, prompt just-in-time learning through Augmented Role-Based Training, and feed cultural metrics back into Adaptive Cultural Reinforcement for continuous improvement.

2.5. Summary of Gaps

Upon completion of the literature search, three gaps were made evident based on the literature:

- Frameworks like CARE and C4 are compliance-driven and are not flexible or capable of triggering AI operations.
- AI research in the field of cybersecurity has primarily targeted technology-centric threats with little consideration for human and cultural factors.
- Absence of an established model describing a model that incorporates behavior analytics, real-time detection, personalized training, and culture reinforcement under one umbrella.

To address these gaps, this paper presents A4 AI-Guard as a scalable, proactive, and human-aware defense ecosystem for healthcare cyber.

3. The Method and Strategy

This study is best described as a conceptual-framework design validated through qualitative case-study analysis. Although we draw on a broad evidence base, the core contribution is the development of the A4 AI-Guard conceptual framework and its retrospective application to two major ransomware incidents [14]. This qualitative, case-anchored validation provides methodological rigor without implying quantitative mixed-method data collection.

The research method applied to the study has been rigorous, and the A4 AI-

Guard model has a well-grounded theoretical and practical basis.

The initial phase of the process involved an extensive review of the available literature on Artificial Intelligence (AI) and its use in healthcare cybersecurity. The review also recorded an increasing call to make human-centric paradigms integrated, since technology-based solutions could also contribute to human errors, such as responses to phishing or social engineering [12]. The paper, therefore, accepted this insight and proposed A4 AI-Guard, a cybersecurity framework that bridges both the predictive and adaptive capabilities of AI with behavioral analytics to help provide a more comprehensive defense mechanism for health organizations.

The framework was then based on four major pillars that are adaptive in nature. They are: AI-backed behavioral risk profiling, Active Threat Detection and Response, Augmented Role-Based Training, and Adaptive Cultural Reinforcement, all backed by AI insights. The approach I have proposed through the A4 AI-Guard framework is designed to address all of these pillars. This, in turn, is a direct response to the gaps that have been evidenced in the current models in place, as it is easy to see that a tool like behavioral risk profiling is an alarm system that continuously observes employee activity and detects behavioral changes, such as abnormal logging in, accessing atypical data, etc., for early warning signs of a potential breach [12]. Also, using AI in a real-time threat detection framework can minimize incident response time by automating the process and allowing for early identification and even containment of attacks.

With augmented role-based adaptive training, relevant staff and employees can receive contextual education materials based on their roles and level of exposure in the organization, to help strengthen resilience at the organizational level. A more adaptive and dynamic model is preferred to traditional security awareness training, as the model can learn continuously based on feedback loops and reinforce behavior based on repeated anomalies. Culture reinforcement with AI insights and training, on the other hand, embeds a security cohesive culture that has an element of compliance as an actor's everyday decision-making process. This approach to culture improvement has been further articulated regarding how healthcare organizations can intrinsically make their overall cybersecurity stance more robust for the long run, through continuing feedback loops and adaptation.

This human factor has been ensured to be expressed as the largest vulnerability factor and is still seen in healthcare cybersecurity environments. This postulation has been supported by the many case studies included in this paper, which show that human errors are still a large contributor to the overall success of the major ransomware attacks. As such, the A4 AI Guard system has a much better probability of predicting, preventing, and mitigating the occurrence of these types of risk, more comprehensively than a technology-only solution [12].

In conclusion, improving the resilience factor within healthcare cybersecurity is dependent on the need to have adaptive, role-based training with cultural feedback, and real-time continuous monitoring systems to help meet the organiza-

tions' objectives. The three pillars have all been previously discussed to make A4 AI Guard a living and adaptive framework, not only in terms of being able to respond to present-day attack threats, but also being more capable of defending against the next new class of threats as they arise. This, therefore, means that organizations can go to sleep every night knowing that they are hardening their defenses against all ongoing attacks, but also doing so for the long term by creating and embedding an element of learning and culture change.

4. Real-World Use Cases

4.1. Case Study 1: Change Healthcare Ransomware Attack

In February 2024, Change Healthcare, one of the leading healthcare technology companies operating in the United States, experienced a large-scale ransomware attack [22]. Change Healthcare, whose operations for hospital pharmacies, help with obtaining medications, completing prescriptions, and supporting insurance processes was all in scramble mode when the company's systems were impacted by the ransomware attack. The culprits for the ransomware were known to be ALPHV/BlackCat, a notorious ransomware group that had stolen user credentials and gained access to their system [2].

With their backdoor, the bad actors were able to successfully move laterally within the network and began to steal large amounts of sensitive patient and financial data and also began propagating ransomware throughout the system for the entire country. As a result, more than 190 million people were said to be affected by the attack. The attack had led to patients potentially being unable to receive important medications and/or book appointments, and the hospitals experiencing significant financial disruption that came from loss of reimbursements [2]. The total losses as a result of the incident were said to be hundreds of millions of dollars. This is if one were to take into account reputation loss and more when accounting for legal implications.

In such a case, if A4 AI-Guard had been in place, the outcomes would have been significantly reduced. The Adaptive Behavioral Risk Profiling module would have flagged behavioral anomalies, like strange log-ins and access to suspicious data, much sooner than what happened. Active Threat Detection and Response would have also responded much more quickly to lateral movement in the network and would most likely have been contained in a matter of hours rather than days or weeks [3]. Role-based Adaptive Training would have also ensured that the IT admins and even general IT support staff would have been able not just to detect but also to report and respond more quickly to prevent the escalation of the incident as a result of credential taking [3]. The last pillar, that of Adaptive Cultural Reinforcement, would have also ensured that there was a much more conscious effort made towards being more security cognizant and less of a response to critical data being leaked, as well as cultural conditions in which being less disruptive in day-to-day work might have contributed to the overall increase in the attack surface of the company.

4.2. Case Study 2: Universal Health Services (UHS) Ransomware Attack

In September 2020, Universal Health Services (UHS), the largest healthcare provider in the United States with over 400 hospitals and practices for mental and behavioral health, was the subject of a large ransomware attack [23]. The attack came in the form of a successful phishing email in which one of the employees clicked a nefarious link, dropped the malware, and then it spread across their internal systems [2].

Within a few hours, all critical systems had been locked down, and the hospitals were made to operate as they had done many decades before—manually. Electronic Health Records (EHRs), emergency units, ambulatory service areas, and even surgical scheduling were all disabled, and even patient processing was difficult to achieve for a few days to weeks after the event. The disruptions were known to be so wide-reaching that this event has led to a recovery cost of over $67 million [2].

In such a situation as this, had A4 AI-Guard been a policy or regulation in place, at least three distinct pillars of defensive countermeasures that we have discussed in this paper would have at least lessened the impact of the attack. Adaptive Behavioral Risk Profiling would have flagged abnormal email activities, which would then have been able to trigger a set of additional response measures, and also account for anomaly detection in the systems that were originally connected to the phishing effort. Active Threat Detection and Response would then have automatically kicked in to quarantine any affected endpoints, which would have saved the organization from having the entire infection in the first place. [3]. Augmented Role-Based Training would also have made certain all frontline staff would have been trained to a degree that would have included administrative, technical, nursing, and many other staff. Such training would have equipped all this frontline staff with skills to be able to easily identify phishing indicators or, better still, report the case before significant time had even elapsed and infection had spread. [3]. The last pillar of Adaptive Cultural Reinforcement is one that would also have ensured that there was a more substantive sense of security awareness in the organization as a whole, which would have made it less likely an employee would not have been able to recognize a ransomware phishing email and cause an organization-wide system compromise.

4.3. Summary

In conclusion, both cases help to show the vulnerability of healthcare organizations that are often open to compromise from a human component as well as cultural incompetency and more, with already-volatile and ever-changing cyber threats. In both cases, it is clear that gaps in human behaviors, system monitoring, and organizational readiness in these healthcare corporations would not have been compromised or leveraged if the adaptive AI-enabled A4 AI GUARD defenses had been deployed. A4 AI-Guard is, therefore, a reinforcement framework that addresses the gap by offering a more comprehensive and, of course, value-added model of

enterprise defensible assurance approach that is clearly aligned to the four pillars that have been discussed earlier in the paper. It is also clear that the adoption of this new way of cybersecurity analysis, via its predictive analytics and real-time detection, tailored training, and adaptive cultural reinforcement, has the capacity to help organizations reduce, at the very least, the impact of such ransomware or data-theft attacks or, at best, ensure their systems are not compromised, and everything from capacity, delivery, financial, and reputational damage has been considered or, in this case, already realized.

Table 3 provides a side-by-side comparison of the two ransomware incidents, summarizing the exploited vulnerabilities, resulting impacts, and how A4 AI-Guard's four adaptive pillars could have mitigated each attack.

Table 3. Case study comparison.

No	*Incident*	*Vulnerability Exploited*	*Impact*	*How A4 AI-Guard Would Mitigate*
1	**Change Healthcare (2024)**	Stolen credentials → Undetected lateral movement	190M patients affected, nationwide service disruption, $100M+ in costs	Behavioral Profiling flags anomalies; Real-time Detection stops the spread; Role-based training prevents credential theft
2	**UHS (2020)**	Phishing email → Malware propagation	400 hospitals disrupted, EHR down, $67M+ recovery costs	Phishing detection; auto-quarantine infected endpoints; role-based training for staff; cultural reinforcement reduces risk

5. Cost-Benefit Analysis

The estimated costs of A4 AI-Guard implementation are based on industry benchmarks and real-world cases of large-scale enterprise cybersecurity ecosystem rollouts incorporating advanced AI-based technologies, primarily within the healthcare sector. Because no practical implementation of A4 AI-Guard has been undertaken as of yet, these estimates are based on the best-practice models and related technologies of extant AI-based cybersecurity technologies. System Integration Costs:

Deploying AI-driven cybersecurity applications typically represents a considerable investment in the hardware, software, and professional services required to customize them for the organization's needs and the AI-driven incident tracking workflows. In total, industry reports and case studies point to the costs of integration being within a range of $200,000 to $500,000. This includes the associated costs of deploying machine-learning infrastructure and data pipelines and, in the case of cloud-specific solutions, the costs of the cloud services. The costs will, of course, vary based on the size of the healthcare facility, the state of its existing IT systems, and the amount and complexity of AI Model(s) [24] [25].

- **Training Costs:**

Employee training is a very important component to ensuring that your AI-driven cybersecurity framework is suitably integrated into your organizational processes and strategic direction. Notably, the development of specified training programs and attendance at training sessions by employees depends on the maturity of the overall project. The overall costs for overall role-based training pro-

grams are estimated at a range of $50,000 to $100,000 based on prevailing industry practice for large healthcare organizations. This includes the necessary costs to develop the training modules, expenses for hiring subject-matter experts/consultants, as well as the payer's costs for the employee attending training sessions where they will be paid during the training time [26]. Spending on Augmented Role Based Training, as an example, will reflect direct staff training, which reflects a clear, concrete, and measurable return on investment as staff develop knowledge unique to their organizational role.

- **Maintenance Costs:**

AI systems require ongoing updates, patching, and optimizations to remain effective in defending against the dynamic landscape of cyber threats. The estimated continuing cost of maintaining these systems is $5,000 a month for cloud infrastructure maintenance, occasional upgrades, and system optimization [21]. The frequency of changes in the healthcare threat landscape necessitates that these ongoing costs are incurred to sustain the adaptability and resilience of A4 AI-Guard.

To highlight the economic feasibility of the proposed framework, **Figure 2** presents the cost-benefit comparison for A4 AI-Guard implementation, demonstrating how the anticipated long-term savings and risk-reduction benefits outweigh the initial deployment and maintenance costs.

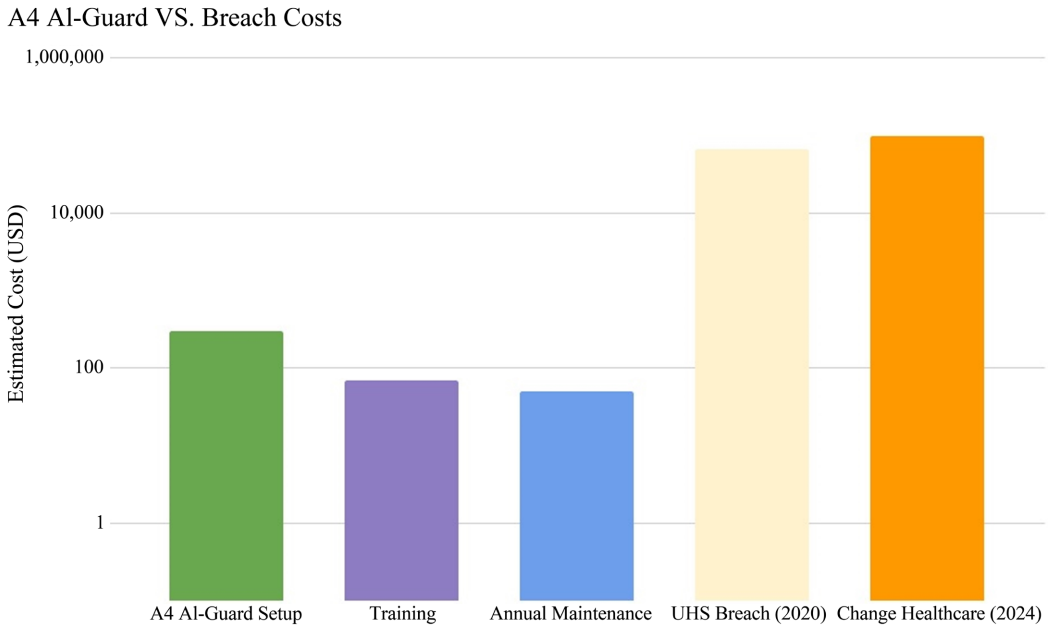

Figure 2. Presents the cost-benefit comparison for A4 AI-Guard implementation.

6. Discussion and Implications

A4 AI-Guard is a substantial step forward in healthcare cybersecurity. Action4 is a frame of reference that legitimizes advancing beyond the compliance-based approaches to a (preemptive) preparedness model for cyber-attacks utilizing AI, and as such, can have benefits in enhancing the overall system's defensive infrastruc-

ture. The A4 AI-Guard framework intentionally recognizes that while technology is important, cybercriminals target human behaviour and organizational culture most often [13].

Collectively, these four adaptive pillars act as a single, interdependent ecosystem, where detection events can automatically drive real-time responses, trigger targeted staff training, and feed cultural metrics back into the system for continuous improvement.

6.1. Benefits

The framework offers healthcare organizations a wide array of pragmatic benefits, including:

- Proactive Threat Identification: The A4 AI-Guard System leverages AI-enabled monitoring to identify abnormal activities at their onset. Rapid identification allows for swift action to be taken to mitigate the damage of ransomware, phishing, and insider attacks [19] [27].
- Role-Specific Training. The A4 AI-Guard System will utilize Augmented Role-Based Training to provide staff with continuous education that is specific to them and allows each (like all roles) of the organization's staff to be prepared at all levels (frontline - IT) [28].
- Culture changes. AI-generated data and insights, experiences recognizable using the system, will understand its role as fostering a "security" culture. Enhancing the shared responsibility of cyber security is dependent on it being acknowledged as a business priority and not only a technical ask [8].

6.2. Implications

The potential implications of A4 AI-Guard are significant and go beyond facilitating its implementation in healthcare.

- **Scalability:** The A4 AI-Guard framework is written to be agile. Scalability is achieved through a modular micro-service architecture and cloud-native deployment. Each pillar can be adopted independently. For example, a small clinic might start with Augmented Role-Based Training, while larger hospital systems can deploy all four pillars orchestrated through containerized services and elastic cloud infrastructure. This modularity allows cost-controlled, incremental adoption without sacrificing interoperability or protection depth.
- **Cost-Effectiveness:** A4 AI-Guard reduces the cost of cybersecurity incidents through deterrence or damage limitation, thus preventing or minimizing a costly breach. Given the economic impact of attacks such as Change Healthcare and UHS, the framework represents a sustainable long-term investment [22].
- **Strategic Resilience:** A4 AI-Guard not only protects against immediate threats, but it also improves long-term resilience by embedding continuous learning and cultural reinforcement. This positions healthcare organizations to be able to withstand future threats, including those which cannot yet be anticipated through conventional means.

7. Limitations and Future Work

Despite the above advantages, there are several limitations of A4 AI-Guard to acknowledge. The framework is heavily dependent on a behavioral and system data set that is of sufficiently high quality, which means that bias, incompleteness, or noise in the inputs can compromise detection. Additionally, the ongoing evolution of healthcare cyber threats presents a continual need for updates, retraining AI models, and optimizing systems. AI transparency and explainability have yet to be resolved—the less clinicians and IT trust any sort of black box prediction, the less they will rely on its recommendations. The cost of integration and organizational cultural adoption may also become a hindrance for smaller healthcare providers with inherently limited funding and resources [29].

Public safety, privacy, and data security are also paramount. Because healthcare data is especially sensitive, there is a significant need for compliance with HIPAA, GDPR, and related regulations. Advanced methods, including federated learning, offer collaborative training between organizations and institutions without having to centralize data, and blockchain offers a secure and auditable way to share medical records. While a recent review done by Abbas *et al.* [9] highlights the potential for combining these approaches to create privacy-preserving, decentralized security infrastructures.

Effective operation requires high-quality, multimodal data, such as EHR access logs, network flow records, endpoint telemetry, and user-activity metadata. While these inputs are vital for accurate detection and behavioral modeling, they raise significant privacy and regulatory concerns. Our design anticipates mitigations such as federated learning to keep raw patient data local, differential privacy in model training, and HIPAA/GDPR-compliant access controls to ensure lawful processing.

Future work should focus on the piloting of A4 AI-Guard in actual healthcare environments to assess its functionality in real time, and to benchmark against currently used cybersecurity frameworks. To do this effectively, XAI (explainable AI) methods will need to be incorporated to promote trust and buy-in from healthcare stakeholders. It will also be important to validate the scalability of A4 AI-Guard in different healthcare ecosystems—from small practices, to smaller networks, to larger national hospital networks—to assess its longer-term viability.

This study outlined a robust, flexible, human-centered cybersecurity solution for healthcare in the context of an evolving cyber threat environment, but there are factors that must be addressed to ensure the appropriate technology solution can be deployed that encompasses the emerging technology as discussed and designed in earlier sections, to promote visible adaptation and mitigation of AI, cybersecurity, and human behaviors on a human-centered adaptive model with significant reinforcement to improve healthcare cyber resilience.

8. Conclusions

This rise in cyber threats provides an opportunity to seek solutions that focus on both human vulnerabilities and technological vulnerabilities to offer real risk mit-

igation for the healthcare system as a whole. This paper proposes A4 AI-Guard, which enables a comprehensive understanding and perspective of cybersecurity with a human focus, using artificial intelligence and human-centered approaches to address human vulnerabilities. A4 AI-Guard is not an average approach to cybersecurity; instead of focusing solely on AI (artificial intelligence) perspectives, A4 AI-Guard incorporates four adaptive pillars to deliver a holistic and evolving open approach to targeting ransomware, phishing, insider threats, and other methods aimed at healthcare or in the healthcare industry through adaptive behavioral risk profiling, active threat detection and incident response, augmented role-based development, and adaptive cultural reinforcement.

Using the Change Healthcare [22] and Universal Health Services [23] ransomware situations, this paper also highlighted what A4 AI-Guard could have done to limit the scope and impact of the incidents brought to the owner's attention by providing reduced detection latency, limiting lateral movement, and developing ready, willing, and able staff. A4 AI-Guard's reliance on systemic and time-driven continuous AI adaptability for recognition/training and its capacity to right-size the cultural reinforcement to be feasible and memorable in a dynamic, trivially impacting, and complex healthcare environment lends itself as a viable method to drive systemic adaptability.

Future validation of A4 AI-Guard through real implementation is a next step, and further empirical testing should be carried out through piloting the A4 AI-Guard in varying healthcare environments, benchmarking with existing cybersecurity frameworks, and further exploring the addition of blockchain and federated learning, data sharing, and further privacy of underlying AI models. As well, addressing the challenges to mitigate AI concerns (bias, etc.) to ensure privacy in workloads and ensuring security possesses scalability will be a limitation to be aware of when looking to validate the practical uptake when looking to deploy the A4 AI-Guard as a solution.

In conclusion, A4 AI-Guard is an adaptable, scalable, hierarchical, proactive, and human-centric defense ecosystem. Through the merged approaches of AI and further humanization of behavioral observation and cultural transformational changes, A4 AI-Guard seeks to enable healthcare organizations to develop the resilience to survive the evolving nature of cyberthreats that are becoming increasingly more complex.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] He, Y., Aliyu, A., Evans, M. and Luo, C. (2021) Health Care Cybersecurity Challenges and Solutions under the Climate of COVID-19: Scoping Review. *Journal of Medical Internet Research*, **23**, e21747. https://doi.org/10.2196/21747

[2] Shryock, T. (2024) Ransomware Surge Highlights Critical Cybersecurity Gaps in Health Care.

https://www.medicaleconomics.com/view/ransomware-surge-highlights-critical-cybersecurity-gaps-in-health-care

[3] Alder, S. (2025) Change Healthcare Increases Ransomware Victim Count to 192.7 Million Individuals. https://www.hipaajournal.com/change-healthcare-responding-to-cyberattack/

[4] Zhan, Y., Ahmad, S.F., Irshad, M., Al-Razgan, M., Awwad, E.M., Ali, Y.A., *et al.* (2024) Investigating the Role of Cybersecurity's Perceived Threats in the Adoption of Health Information Systems. *Heliyon*, **10**, e22947. https://doi.org/10.1016/j.heliyon.2023.e22947

[5] Burke, W., Stranieri, A., Oseni, T. and Gondal, I. (2024) The Need for Cybersecurity Self-Evaluation in Healthcare. *BMC Medical Informatics and Decision Making*, **24**, Article No. 133. https://doi.org/10.1186/s12911-024-02551-x

[6] Rahmany, M. and Selvi, A. (2025) C4 Framework for Healthcare Cybersecurity Defense: A Human-Centric, Socio-Technical Approach. *E-Health Telecommunication Systems and Networks*, **14**, 31-38. https://doi.org/10.4236/etsn.2025.143004

[7] Akhtar, Z.B. and Tajbiul Rawol, A. (2024) Enhancing Cybersecurity through Ai-Powered Security Mechanisms. *IT Journal Research and Development*, **9**, 50-67. https://doi.org/10.25299/itjrd.2024.16852

[8] Yigit, Y., Ferrag, M.A., Ghanem, M.C., Sarker, I.H., Maglaras, L.A., Chrysoulas, C., *et al.* (2025) Generative AI and LLMs for Critical Infrastructure Protection: Evaluation Benchmarks, Agentic AI, Challenges, and Opportunities. *Sensors*, **25**, Article 1666. https://doi.org/10.3390/s25061666

[9] Abbas, S.R., Abbas, Z., Zahir, A. and Lee, S.W. (2024) Federated Learning in Smart Healthcare: A Comprehensive Review on Privacy, Security, and Predictive Analytics with IoT Integration. *Healthcare*, **12**, Article 2587. https://doi.org/10.3390/healthcare12242587

[10] Balla, E. (2025) Pillars of Intelligence: Key Fields at the Forefront of AI. https://aijourn.com/pillars-of-intelligence-key-fields-at-the-forefront-of-ai/

[11] Gupta, S., Kapoor, M. and Debnath, S.K. (2025) AI-Based Cybersecurity Solutions for Healthcare. In: *Artificial Intelligence-Enabled Security for Healthcare Systems*, Springer, 65-86. https://doi.org/10.1007/978-3-031-82810-2_4

[12] Sudheer, S. (2024) Ransomware Attacks and Their Evolving Strategies: A Systematic Review of Recent Incidents. *Journal of Technology and Systems*, **6**, 32-59. https://doi.org/10.47941/jts.2399

[13] Javaid, M., Haleem, A., Singh, R.P. and Suman, R. (2023) Towards Insighting Cybersecurity for Healthcare Domains: A Comprehensive Review of Recent Practices and Trends. *Cyber Security and Applications*, **1**, Article 100016. https://doi.org/10.1016/j.csa.2023.100016

[14] Universal Health Services Inc (2020) Form 8-K: Information Technology Security Incident. U.S. Securities and Exchange Commission. https://www.sec.gov/Archives/edgar/data/352915/000156459020044863/uhs-8k_20200927.htm

[15] Binhammad, M., Alqaydi, S., Othman, A. and Abuljadayel, L.H. (2024) The Role of AI in Cyber Security: Safeguarding Digital Identity. *Journal of Information Security*, **15**, 245-278. https://doi.org/10.4236/jis.2024.152015

[16] Censinet, K.L.A.S., American Hospital Association, Health-ISAC and HPHSCC (2025) 2024 Healthcare Cybersecurity Benchmarking Study. https://health-isac.org/partnered-report-healthcare-cybersecurity-benchmarking-study-2024/

[17] International Research Journal of Multidisciplinary Studies (2025) Deep Learning for Anomaly Detection in IoT Healthcare Systems. https://www.irjms.com/journal/deep-learning-for-anomaly-detection-in-iot-healthcare-systems

[18] Alalwany, E., Alsharif, B., Alotaibi, Y., Alfahaid, A., Mahgoub, I. and Ilyas, M. (2025) Stacking Ensemble Deep Learning for Real-Time Intrusion Detection in IoMT Environments. *Sensors*, **25**, Article 624. https://doi.org/10.3390/s25030624

[19] Ewoh, P. and Vartiainen, T. (2024) Vulnerability to Cyberattacks and Sociotechnical Solutions for Health Care Systems: Systematic Review. *Journal of Medical Internet Research*, **26**, e46904. https://doi.org/10.2196/46904

[20] Neprash, H.T., McGlave, C.C., Cross, D.A., Virnig, B.A., Puskarich, M.A., Huling, J.D., *et al.* (2022) Trends in Ransomware Attacks on US Hospitals, Clinics, and Other Health Care Delivery Organizations, 2016-2021. *JAMA Health Forum*, **3**, e224873. https://doi.org/10.1001/jamahealthforum.2022.4873

[21] Mishra, S. (2023) Exploring the Impact of AI-Based Cyber Security Financial Sector Management. *Applied Sciences*, **13**, Article 5875. https://doi.org/10.3390/app13105875

[22] Pults, K. (2025) The Change Healthcare Breach: What Changed, What Didn't, and What Must. https://www.securitymagazine.com/articles/101394-the-change-healthcare-breach-what-changed-what-didnt-and-what-must

[23] Newman, L. (2020) A Ransomware Attack Has Struck a Major US Hospital Chain. https://www.wired.com/story/universal-health-services-ransomware-attack

[24] IBM (2025) Cost of a Data Breach: The Healthcare Industry. https://www.ibm.com/think/insights/cost-of-a-data-breach-healthcare-industry

[25] National Cyber Security Centre (NCSC) (2025) Annual Review 2024. https://www.ncsc.gov.uk/files/NCSC_Annual_Review_2024.pdf

[26] Ponemon Institute (2024) The 2024 Study on Cyber Insecurity in Healthcare: The Cost and Impact on Patient Safety and Care. https://ponemonsullivanreport.com/2024/10/the-2024-study-on-cyber-insecurity-in-healthcare-the-cost-and-impact-on-patient-safety-and-care/

[27] Health-ISAC (2024) Partnered Report: Healthcare Cybersecurity Benchmarking Study 2024. https://health-isac.org/partnered-report-healthcare-cybersecurity-benchmarking-study-2024/

[28] Mennella, C., Maniscalco, U., De Pietro, G. and Esposito, M. (2024) Ethical and Regulatory Challenges of AI Technologies in Healthcare: A Narrative Review. *Heliyon*, **10**, e26297. https://doi.org/10.1016/j.heliyon.2024.e26297

[29] Asan, O., Bayrak, A.E. and Choudhury, A. (2020) Artificial Intelligence and Human Trust in Healthcare: Focus on Clinicians. *Journal of Medical Internet Research*, **22**, e15154. https://doi.org/10.2196/15154