Trustee Insights

AI AND CYBERSECURITY

5 keys to governing the convergence of AI, cybersecurity and clinical risk

Strategies and guidance for board decision-making, oversight and accountability across clinical and operational environments

BY LAYLA ALABDULKARIM, Ph.D.

Health care organizations operate in a uniquely complex cybersecurity and data protection environment due to the highly sensitive, mission-critical nature of the digital assets required to deliver patient-centered care. These assets — including patient health data, clinical systems and connected medical technologies — are essential to patient safety, regulatory compliance and community trust.

As artificial intelligence (AI) becomes increasingly embedded in core platforms such as electronic health records (EHRs), medical imaging systems and networked medical devices, it is transforming care delivery while simultaneously expanding the cybersecurity risk landscape. Cyberattacks targeting these AI-enabled environments can result in regulatory penalties, operational disruption, reputational damage and, critically, patient harm.

For health care boards, the convergence of AI, cybersecurity and clinical risk is no longer solely a technical issue — it is a core enterprise risk and governance priority requiring deliberate oversight, informed decision-making and clear accountability. This article outlines five key governance considerations and provides board-level questions throughout, followed by a structured oversight checklist for boards to guide action.

1. Strategic Alignment of AI and Health Care Organizational Objectives

AI is rapidly expanding across clinical, operational and administrative functions — introducing both measurable value and material risk. Boards must ensure that AI initiatives are aligned not only with organizational strategy, but also with cybersecurity resilience, patient safety priorities and regulatory compliance obligations.

In addition to established requirements, such as the Health Insurance Portability and Accountability Act (HIPAA), the Health Information Technology for Economic and Clinical Health Act, Office for Civil Rights guidance and U.S. Food and Drug Administration (FDA) frameworks, boards should be aware of emerging expectations under the Office of the National Coordinator's Health Data, Technology, and Interoperability (HTI-1) Final Rule. This rule introduces new requirements for transparency, evaluation and governance of AI-enabled clinical decision support tools, signaling increased accountability for the validity, effectiveness and safety of these technologies.

To operationalize alignment, organizations should leverage structured

frameworks such as:

- National Institute of Standards and Technology (NIST) AI Risk Management Framework
- Department of Health and Human Services (HHS) Cybersecurity Performance Goals (CPGs)
- Enterprise Risk Management integration models

AI should not be treated as a standalone innovation effort, but instead as an integrated component of enterprise cybersecurity strategy and governance.

2. Cybersecurity in an AI-Driven Environment

Artificial intelligence introduces risks to the confidentiality, integrity and availability of critical health care assets. These assets include EHRs, protected health information (PHI), medical devices, imaging systems such as Picture Archiving and Communication Systems (PACS), clinical decision-support tools, and operational platforms like scheduling and pharmacy systems, as well as Internet of Things (IoT)–enabled devices. Compromise of any of these systems can result not only in data loss, but also in operational disruption and potential harm to patients.

Effective stewardship governance requires coordinated attention to people, processes and technology, supported by a formal AI governance structure. This structure should involve stakeholders such as the CISO, CIO or IT leadership, legal and compliance, clinical leadership, finance, risk management and privacy officers. Many organizations establish cross-functional "management led", AI governance or stewardship councils that engage early — before AI acquisition or deployment — to identify and address cybersecurity, privacy, clinical, legal and financial risks.

From a **people** perspective, boards and executives must ensure accountability and coordination across stakeholders. This includes partnering with CIOs and CISOs to provide ongoing cybersecurity, privacy and AI training. Education should go beyond awareness to cover AI risks such as bias, overreliance, manipulation and improper data use. Oversight can be measured through metrics like training completion rates, AI-related tabletop exercises and documented governance discussions.

From a **process** perspective, AI development, procurement, deployment and monitoring should be standardized, auditable and aligned with frameworks such as the HIPAA Security Rule, NIST SP 800-66, FDA cybersecurity guidance, the NIST AI Risk Management Framework and HHS CPGs. Processes should include cybersecurity and privacy reviews in vendor selection, integrating AI into third-party risk management. Practical tools include pre-purchase risk checklists, vendor disclosures on data and model practices, and structured evaluation criteria for safety, efficacy and operational impact.

From a **technology** perspective, AI tools should not introduce unmanaged vulnerabilities, opaque dependencies or excessive reliance on third parties. Due to the "black box" nature of many models, monitoring risks like data poisoning, model drift and manipulation often requires collaboration with vendors. Shared responsibilities for monitoring, incident response, auditability and transparency should be clearly defined.

Cybersecurity leaders must ensure their teams can address AI-specific risks alongside traditional threats by building internal expertise, using external assessments when needed and embedding AI into existing risk and compliance workflows. Clinical leadership must also be actively involved, as AI-related cybersecurity incidents can directly impact patient care. As a result, AI cybersecurity risk should be treated not only as a technical or compliance issue, but also as a patient safety concern requiring close collaboration among cybersecurity, IT, clinical leadership and governance bodies.

Board Questions to Consider:

- How does each AI initiative demonstrably improve patient outcomes or operational performance?
- Has cybersecurity and patient safety risk been assessed prior to deployment?
- What regulatory frameworks apply, including HTI-1 requirements for clinical decision support?
- What frameworks or controls are used to validate alignment and manage risk?

3. Human vs. AI in the Cybersecurity Chain

Human error has long been the weakest link in cybersecurity, and health care is no exception. Staff — often working under high pressure — are frequent targets of phishing, social engineering and credential

theft. This has driven greater reliance on AI-driven tools to automate detection, triage and response.

However, AI is not inherently more secure than humans. While it can outperform analysts in processing large volumes of data, operating at machine speed and reducing inconsistency or alert fatigue, it also introduces new risks. AI systems can be exploited, manipulated through adversarial inputs or misconfigured, sometimes increasing rather than reducing alert burdens. In health care, failures in AI-enabled cybersecurity decisions can extend beyond IT, directly impacting clinical operations and patient safety. As AHA National Advisor for Cybersecurity and Risk John Riggi noted, "Any cyberattack on the health care sector that disrupts or delays patient care … crosses the line from an economic crime to a threat-to-life crime."

The link between cybersecurity risk and patient safety further complicates the balance between human and AI roles. Cyberattacks on networked medical devices and operational technology can translate directly into physical harm. Boards should therefore rigorously evaluate the use of AI across the organization, including within cybersecurity functions—determining whether it genuinely mitigates human limitations or introduces new risks—and ensure strong oversight wherever AI influences clinical or operational decisions.

4. AI Risk Amplification: Clinical Ops, Cybersecurity and the Adversaries

The potential adverse impacts of AI in the health care sector are particularly pronounced, whether the risk arises from embedding AI in clinical operations, cybersecurity functions or even systems used by adversaries to launch cyberattacks. For example, AI-enabled medical systems, such as diagnostic imaging algorithms and sepsis prediction models, can introduce patient safety risks. A compromised algorithm, biased training data or software error could lead to misdiagnosis, incorrect treatment recommendations or delayed interventions, directly endangering patients.

In a similar manner, although AI can significantly enhance cybersecurity functions through faster threat detection, automated response and improved vulnerability management, it can also amplify risk rather than mitigate it when poorly managed or insufficiently tested. In health care settings, the consequences of such failures are often more severe due to the close coupling between digital systems and patient care. For example, automated incident response actions, such as isolating systems from the network, may protect data but inadvertently disrupt clinical workflows or delay care. Similarly, automated patching or system updates can unintentionally take down critical platforms such as EHRs or radiology systems if not rigorously validated in clinical contexts. Additionally,

Board-level AI and Cybersecurity Oversight Action Checklist

- **Validate strategic alignment:** Confirm AI initiatives align with organizational strategy, cybersecurity priorities, patient safety objectives and applicable regulatory obligations, including emerging requirements for AI transparency, validity, effectiveness and safety in clinical decision support.
- **Strengthen stewardship or oversight governance and accountability:** Ensure clear ownership, escalation pathways and reporting structures for AI-related cyber and patient safety risks. Consider whether a dedicated AI or AI/cybersecurity Board function or subcommittee is warranted, and ensure clinical leadership is meaningfully involved in oversight.
- **Expand cybersecurity risk oversight:** Ensure visibility into AI-specific risks such as data integrity issues, model drift or failure, inappropriate or unintended use of AI systems, and dependencies on third-party vendors, in addition to traditional cybersecurity threats.
- **Assess AI impact on care:** Verify that AI-enabled cybersecurity and clinical systems incorporate appropriate safeguards, human oversight and rigorous testing that account for real-world clinical workflows and patient safety implications.
- **Confirm preparedness and financial exposure:** Ensure incident response, resilience planning and enterprise risk management frameworks explicitly address AI-related failures. Evaluate potential financial exposure, including whether insurance coverage (e.g., cyber liability, malpractice and technology errors and omissions (E&O)) is sufficient and appropriately tailored to AI-related risks.

AI-powered phishing and social engineering tools enable attackers to target hospital staff at scale with increasing sophistication.

To address these risks, organizations should consider establishing a dedicated AI and Cybersecurity Subcommittee at the board or executive level to provide focused oversight. AI governance teams play a critical role in minimizing risk and supporting the responsible use of AI tools by defining clear accountability structures, implementing validation and monitoring processes, and ensuring alignment with patient safety priorities. Boards should focus on evaluating AI use cases based not only on efficiency gains but also on clinical impact, resilience and risk exposure.

A key governance challenge is distinguishing between responsible and reckless use of AI. Responsible use is characterized by rigorous testing, transparency, human oversight and alignment with patient safety and operational continuity. In contrast, reckless use prioritizes speed or innovation without sufficient validation, risk assessment or safeguards. Boards should ensure that AI use across cybersecurity and clinical systems includes strong oversight, clear escalation paths and explicit assessment of patient safety impacts prior to deployment.

Furthermore, boards and CISOs should remain mindful of AI risks that extend beyond their direct control, particularly those stemming from AI-enabled cyberattacks. This awareness should be reflected in the organization's cybersecurity strategy and corresponding security controls, ensuring preparedness not only for internal AI risks but also for increasingly sophisticated external threats.

5. Financial Risk and Brand Damage from AI-Driven Cyber Threats

AI-related cybersecurity incidents can expose health care organizations to a broad range of financial and reputational risks, including data breaches, intellectual property theft, regulatory noncompliance, operational disruption, erosion of patient trust and direct patient harm. While many of these risks predate AI adoption, the integration of AI across health care functions increases their scale, speed and complexity, making incidents more difficult to predict and contain.

Boards should ensure that these risks are explicitly modeled within the organization's enterprise risk management framework. This includes understanding the potential financial impact of AI-related incidents, evaluating preparedness for regulatory investigations and confirming that incident response and business continuity plans address failures of AI and automated systems. In addition, boards should assess whether existing liability insurance coverage adequately reflects AI-related risks, including potential gaps in cyber liability, malpractice, and technology errors and omissions coverage, and consider adjustments to ensure sufficient protection against emerging exposures.

Conclusion

The integration of AI into health care is reshaping care delivery, cybersecurity, data privacy and patient safety risk. As AI becomes embedded in clinical and operational workflows, its failure or compromise can lead not only to data loss, but also to regulatory exposure, operational disruption and patient harm. For boards, this convergence highlights the need for deliberate stewardship, informed oversight and disciplined decision-making.

To support this, this article contains questions and and a practical checklist to guide board discussion and and decisions. These focus areas are designed to ensure AI adoption strengthens organizational resilience and trust while aligning with evolving regulatory expectations, including clinical decision support and algorithm transparency requirements under HTI-1 Final Ruls and anticipated updates such as HTI-5.

Layla Alabdulkarim, Ph.D., CISSP, CISM, CISA, (layla.abdul@gmail.com), is a cybersecurity researcher and consultant.

Please note that the views of the authors do not always reflect the views of their organizations or of the AHA.