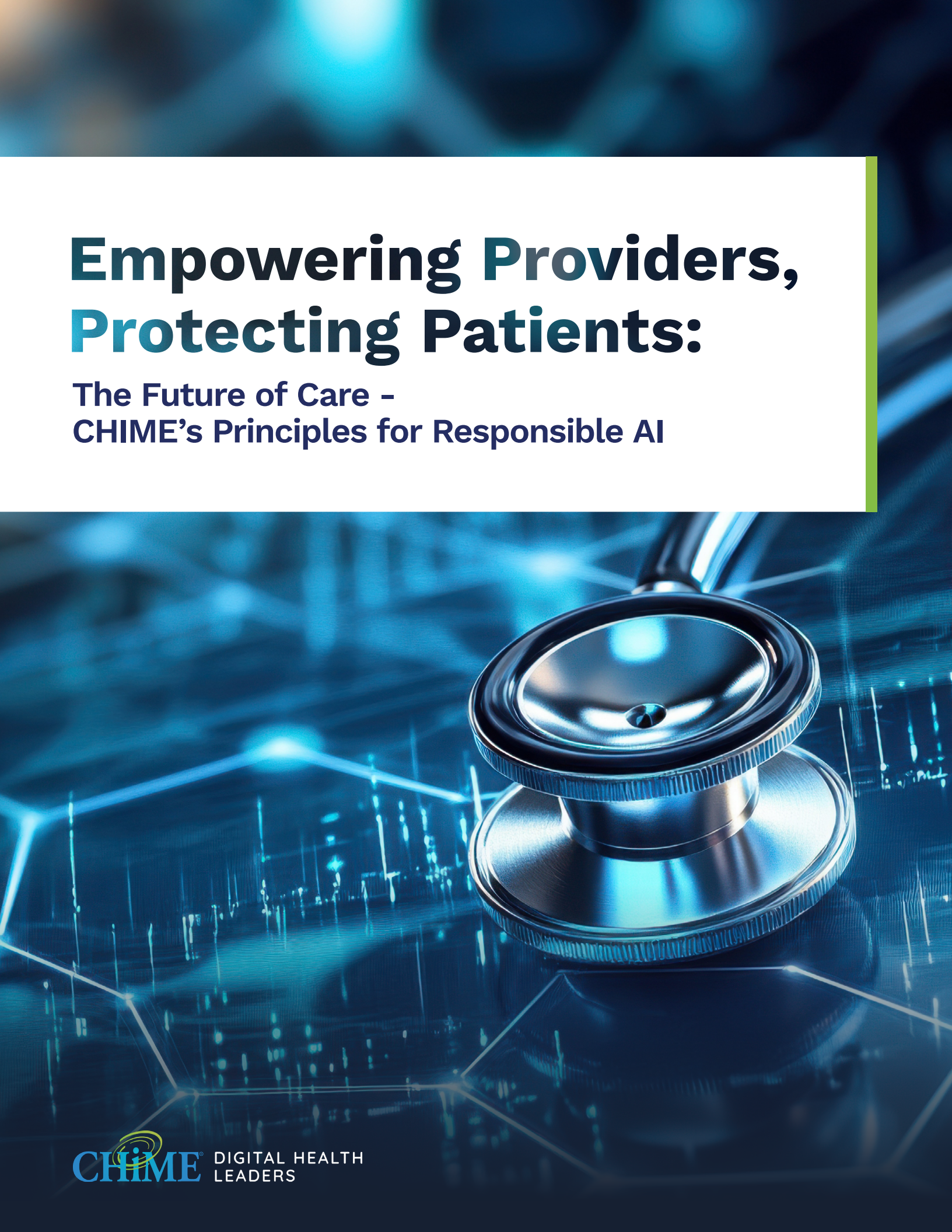
Empowering Providers, Protecting Patients:

The Future of Care – CHIME's Principles for Responsible AI



CHIME DIGITAL HEALTH LEADERS

FORWARD

Artificial Intelligence (AI) is no longer a distant concept. It is reshaping the very foundation of healthcare delivery.

From enhancing diagnostic accuracy to streamlining administrative workflows, AI offers unprecedented opportunities to improve patient outcomes, reduce clinician burden, and foster innovation across the care continuum. Yet, with these opportunities come profound responsibilities.

Healthcare providers stand at the intersection of technology and trust. As stewards of patient safety and privacy, they must navigate a rapidly evolving landscape where algorithms influence clinical decisions and data flows beyond traditional boundaries. This document serves as a guide to understanding and implementing AI responsibly anchored in principles of transparency, fairness, and patient-centered care.

Our goal is to empower providers and other stakeholders with actionable insights including:

- Protecting patients and their data through robust cybersecurity and privacy frameworks.
- Fostering patient-centered care by ensuring human oversight and ethical use of AI tools.
- Empowering providers with education, resources, and strategies to overcome barriers—especially for small and under-resourced organizations.

AI is not a replacement for human judgment; it is a tool to augment clinical expertise. By embracing innovation while safeguarding trust, healthcare leaders can ensure that AI fulfills its promise: improving health for all.

We invite you to explore our principles, challenge assumptions, and join us in shaping a future where technology and compassion work hand in hand.

Sincerely,

Russell P. Branzell, CHCIO, LCHIME
President and CEO
CHIIME

Table of Contents

CHIME DIGITAL HEALTH LEADERS

EXECUTIVE SUMMARY

AI is rapidly transforming healthcare, offering new opportunities to improve patient outcomes, streamline operations, and foster innovation. However, these advances bring complex challenges and responsibilities for healthcare organizations (HCOs).

The CHIME Principles for Responsible AI provide a comprehensive approach describing both challenges and potential solutions to help nurture responsible AI adoption. The document is organized into three interconnected sections:

I. Protecting Patients and Their Data

AI's integration into healthcare heightens the need for robust cybersecurity and privacy protections. Providers face increasing regulatory burdens, fragmented policies, and rising threats from AI-driven cyberattacks. This document calls for:

- Streamlined, incentive-based policies that prioritize investment in cybersecurity over punitive measures.
- Enhanced public-private partnerships to coordinate cyber threat response and shared best practices.
- Standardized security requirements for third-party technologies and medical devices.
- A comprehensive national privacy law to address gaps and inconsistencies across states and entities handling health data.
- Clear patient consent policies and transparent data use, including explainable and interpretable AI, clinician involvement, and robust anonymization protocols.

II. Fostering Patient-Centered Care

AI must be implemented with patient safety and human oversight at its core. These Principles emphasize:

- Adoption of voluntary frameworks and standards (e.g., NIST AI Risk Management Framework (AI RMF), National Academy of Medicine Code of Conduct) to guide ethical AI use.
- Risk-based oversight and ongoing monitoring of AI tools, with special attention to high-impact clinical applications.
- Safe harbor protections for providers using Food and Drug Administration (FDA)-authorized AI tools, and oversight to validate safety and effectiveness.
- Stress-testing AI models for fairness across diverse patient populations and maintaining transparency about training data and model limitations.
- Continuous human oversight to prevent over-reliance on AI and ensure clinical responsibility.

III. Empowering Providers

To realize AI's full potential, providers must be equipped with the knowledge, resources, and support to overcome adoption barriers. Our key recommendations include:

- Promoting transparency and trust through disclosure of model details, training data, and performance metrics.
- Standardizing model cards to facilitate informed purchasing and governance decisions.
- Investing in tailored education and workforce training to build AI literacy and critical thinking.
- Supporting small and under-resourced providers through targeted policies, funding, and collaborative purchasing options.
- Encouraging innovation while safeguarding patient safety, data integrity, and equitable access.

Conclusion

CHIME's Principles for Responsible AI call for a balanced approach to AI adoption: embracing innovation while safeguarding trust, privacy, and patient-centered care.

BY ALIGNING POLICY, PRACTICE, AND EDUCATION, HEALTHCARE LEADERS CAN ENSURE THAT AI FULFILLS ITS PROMISE TO IMPROVE HEALTH FOR ALL.

SECTION 1

Protecting Patients and Their Data

CYBERSECURITY

Top Challenges and Recommended Solutions

01 Regulatory Burdens

Regulatory uncertainty and overlapping laws and policies have created compliance burdens for providers, stifled innovation and unnecessarily increased provider liabilities. This has made it harder for providers to make needed cybersecurity investments and keep ahead of the rising cyber-attacks launched using AI.

SOLUTIONS

- Reducing regulatory burdens should be prioritized to foster innovation, cost savings and allow providers to make investments that will meaningfully improve their cybersecurity posture.
- Policies should prioritize incentives over penalties (i.e., “carrots” vs. “sticks”) to enable providers to invest their limited resources in cybersecurity tools that best meet their individual needs.
- The proposed Health Insurance Portability & Accountability Act (HIPAA) Security Rule should be rescinded and re-issued to remove duplicative and burdensome provisions.

  The Department of Health and Human Services’ (HHS’) Office for Civil Rights (OCR) should adopt flexible, incentive-based policies that encourage providers to implement recognized cybersecurity best practices, rather than penalizing them for experiencing (or being victims of) cyber incidents as supported under P.L. 116-321.
- Before it is finalized and released, the Cyber Incident Reporting for Critical Infrastructure Act (CIRCIA) proposed rule should be revised to reduce duplicative reporting burdens and better align with existing healthcare cybersecurity frameworks.
- The new Medicare Promoting Interoperability (PI) Program’s security risk management measure should be removed, as it is duplicative and burdensome.
- Voluntary adoption of the Cybersecurity Performance Goals (CPGs) which were co-developed in conjunction with HHS should continue to be prioritized.
- Updating the Stark and Anti-Kickback Statute (AKS) safe harbors for cybersecurity donation policies to address liability concerns would accelerate adoption of vital technologies, particularly by lesser-resourced organizations – thereby strengthening collective sector resilience.

02 Public-Private Partnerships

Public-private collaboration should be prioritized.

SOLUTIONS

- The Critical Infrastructure Partnership Advisory Council (CIPAC) authorities should be re-enabled to foster better coordination among the public and private sectors.
- The federal government should continue to support the Health Sector Coordinating Council's Cybersecurity Working Group (HSCC CWG) to foster support, cross-sector collaboration, and public-private partnerships like the 405(d) Program.
- A new public-private entity focused on addressing cyber threats arising from AI may be needed to protect national security.

03 Third-Party Security

Healthcare providers are increasingly hamstrung when purchasing and procuring applications, medical devices, AI solutions, and other digital technologies. Vendors often deliver products with insufficient baseline security assurances, forcing providers to shoulder the burden of conducting individual risk assessments that are costly, duplicative, and resource-intensive. This fragmented approach undermines efficiency, slows adoption of innovative tools, and can leave critical vulnerabilities unaddressed across the ecosystem.

SOLUTIONS

- A marketplace should be established for certified apps – including those leveraging AI – which have demonstrated their ability to meet a baseline set of security requirements. Establishing consistent, transparent, and enforceable security standards for third-party technologies is essential to reducing provider burden, strengthening supply chain resilience, and protecting patients from escalating cyber threats.
- Promote the interoperability of security processes between vendors to increase competition and reduce duplicative spending.
- Medical device manufacturers must fully comply with the Protecting and Transforming Cyber Health Care Act of 2022 (PATCH Act) signed into law as part of the 2023 Consolidated Appropriations Act and bear responsibility for "patching" their devices when cybersecurity vulnerabilities are found, rather than shifting the cost to providers, and ultimately patients.
- As federal policymakers consider ways to best support hospitals and healthcare systems, they should look to align with existing provider-supported cybersecurity policies at the state level.
- Policymakers should also consider how a single federal policy could offer efficiency by preempting the multiple laws and other requirements that providers must meet.
- Vendors should be required to disclose AI components, the type of AI and the data used, and their related security assessment, including addressing relevant threats and vulnerabilities.

04 Competition

Inefficient and fragmented federal policies are stifling competition and innovation in the healthcare sector. These policy-driven barriers limit market incentives, drive up costs, and ultimately hinder providers' ability to adopt technologies needed to efficiently and effectively mitigate and respond to evolving cyber threats.

SOLUTIONS

- Pro-business policies are needed to foster competition.
- Device manufacturers should adopt standards that enhance vendor speed to market, competition and innovation – and providers should request this from their vendors.
- The Cybersecurity Information Sharing Act of 2015 (CISA 2015) should be reauthorized to allow permitted threat sharing, which can improve our sector's posture. Reauthorizing this law could have a positive impact on competition and American cyber capabilities by improving awareness of cyber vulnerabilities.

Discussion

AI cannot be discussed in isolation from the growing impact of cybersecurity threats on healthcare. Providers and patients continue to bear the consequences of these attacks, which are increasing in frequency and sophistication with the use of AI.

The ability of providers to stay ahead of AI-driven cyber threats is becoming harder and harder. The magnitude of changes being ushered in by AI to the cyber threat landscape is changing extremely rapidly. McKinsey has coined AI the "greatest threat" to cybersecurity. They found that breakout times (the time it takes for an attacker to move laterally within a network) are now often under an hour, due to AI automation. According to another source, AI-generated phishing emails now account for over 80 percent of phishing attempts, with nearly 80 percent of recipients opening them due to their realism. The magnitude of changes being ushered in by AI to the cyber threat landscape is changing extremely rapidly.

Immediate, coordinated action from all stakeholders – both public and private – is essential to safeguard patient care, protect sensitive health data, and uphold national security. Effective cyber policies that empower and support, rather than penalize providers will form the foundation for a secure and trustworthy American AI ecosystem. This will ensure innovation and safety advance hand in hand.

Hospitals and healthcare systems – frequent victims of cyber-attacks – have been referred to as "target rich and cyber poor." Given how lucrative provider targets are for cybercriminals, healthcare unfortunately remains the top target for ransomware attacks.

ONE STUDY FOUND THAT NEARLY 70% OF PROVIDERS EXPERIENCED A RANSOMWARE INCIDENT IN 2024.

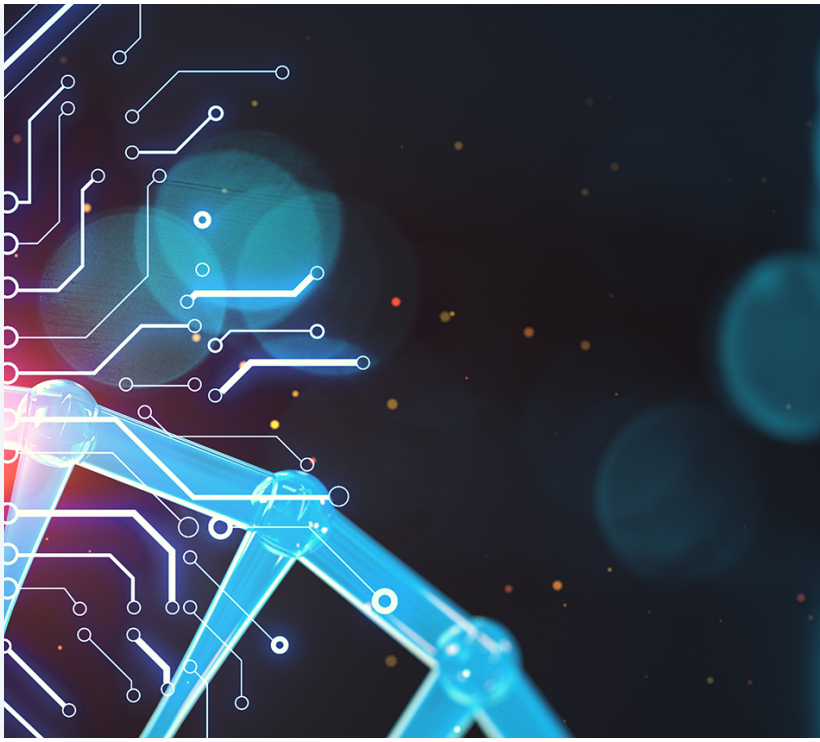
Despite healthcare organizations (HCOs) investing heavily in cybersecurity over the past several years and making meaningful progress, there are significant and persisting cybersecurity challenges and many remain outside of their immediate control. The challenges are many, and include a vastly more interconnected healthcare ecosystem, a burgeoning use of application programming interfaces (APIs), threats posed by third-party risk, a rapid rise in the availability and use of generative AI, and a growing level of sophistication of cyber criminals.

Together, these factors have amplified cybercriminals' ability to infiltrate provider networks and third-party applications they rely on, exploiting AI to carry out increasingly sophisticated attacks, including advanced social engineering and deepfake schemes. It has also allowed criminals to more quickly pinpoint and exploit weaknesses with the ability to not only do extensive harm, but to do so more quickly. This has created an environment that has made it exceptionally difficult for our members to stay ahead of cyber criminals and highlights the urgent need to work together with vendors, manufacturers, and other stakeholders including federal and state authorities, to ensure patient data remains secure. While some providers have begun to adopt AI to help improve their cyber posture, these solutions are costly, and organizations continue to grapple with complex privacy and liability challenges.

REGULATORY BURDENS

As the cyber landscape for healthcare providers has become infinitely more complex, so too has the legal and regulatory climate. Providers are navigating a policy landscape that is rife with uncertainty and marked by overlapping – and in many cases reactive – policies that have had the unintended effect of making it far more challenging for them to protect patient data.

This policy environment has created inefficiencies and hindered innovation. To date, policymakers have relied heavily on penalties, overlooking the need for proactive support to address the real cybersecurity challenges confronting providers.



For instance, HIPPAA breach notification policies are highly punitive with provider breaches often involving situations outside of their control – such as software or medical devices that are no longer patchable, or unknown vulnerabilities from the thousands of third-party products they have purchased and rely on. While the default response has been to layer additional mandates on providers, excessive regulatory prescriptions have not deterred ransomware attacks, and risk diverting resources from the very safeguards necessary to protect patient data and system integrity.

Two recently proposed rules amplify these concerns. The HIPAA Security Rule proposed in January 2025, and the Cyber Incident Reporting for Critical Infrastructure Act (CIRCIA) rule proposed in April 2024 each call for extensive new and burdensome policies. As detailed in our comment letters (here and here), if finalized, these policies would have a crushing impact on healthcare providers and squash their ability to adopt solutions best tailored to their individual needs without the intended outcome of a better protected ecosystem. Both rules take the approach that security can be solved with extensive documentation. This approach only serves to divert precious resources away from investments that could be used to truly improve their cyber posture or adopt much-needed patient care solutions. There are also varying and burdensome reporting mandates. For instance, while HIPAA requires breaches be reported "without unreasonable delay" and within 60 days to HHS, the CIRCIA statute requires a "substantial" cyber incident be reported within 72 hours, and any ransom paid must be reported within 24 hours to the Cybersecurity and Infrastructure Security Agency (CISA).

An increasing number of providers have adopted a voluntary set of cybersecurity best practices developed in collaboration with HHS and the private sector under the 405(d) Program as required by the Cybersecurity Information Sharing Act of 2015 (CISA 2015). These best practices were co-developed over several years between industry and federal authorities under the Health Sector Coordinating Council's Cybersecurity Working Group (HSCC CWG) and published by HHS. P.L. 116-321, signed by President Trump on January 5, 2021, reinforced these standards by explicitly recognizing them in statute as a means for providers to earn credit for implementing cybersecurity best practices during a HIPAA audit or investigation – one of the few statutory incentives available to encourage proactive provider compliance. Based upon these best practices, HHS released ten essential and ten enhanced Cybersecurity Performance Goals (CPGs), offering a voluntary set of cybersecurity standards for use by healthcare providers and other sector stakeholders.

The future of these standards is uncertain as HHS appears to have shifted its focus away from these voluntary standards. The proposed HIPAA Security Rule seems to have removed the ability to leverage P.L. 116-321, and no resources have been published by HHS related to the 405(d) Program since October 2024. Further, in March the Department of Homeland Security (DHS) suspended indefinitely the Critical Infrastructure Partnership Advisory Council (CIPAC) framework, which for twenty years has exempted critical infrastructure public-private partnership meetings from public disclosure rules under the Federal Advisory Committee Act (FACA). The effect of this action has been to curtail the ability of designated CIPAC agencies – specifically HHS for the health sector – to actively coordinate planning in trusted, non-public settings with private industry sector coordinating councils on critical infrastructure policy and programs. This has constricted the frequency and fidelity of cyber threat and mitigation information sharing and incident response readiness

between the private sector and government.

The Centers for Medicare & Medicaid Services (CMS) also recently finalized a new measure in the Medicare Promoting Interoperability (PI) Program.

BEGINNING IN 2026, HOSPITALS AND HEALTH SYSTEMS WILL BE MANDATED TO ATTEST AFFIRMATIVELY THAT THEY HAVE CONDUCTED A COMPREHENSIVE SECURITY RISK MANAGEMENT PROCESS – A POLICY WHICH IS ALREADY REQUIRED UNDER HIPAA.

In addition to these rules, hospitals are also being asked to meet Joint Commission accreditation standards for cybersecurity.

States have taken action to fill some of the perceived gaps at the federal level by passing their own laws. In 2019, New York State passed (and subsequently amended) the Stop Hacks and Improve Electronic Data Security Act (SHIELD Act) established under S5575B requiring businesses in the state that experience a breach of private information – including medical and insurance information – to be reported within 30 days of the discovery of the breach to state authorities. The Texas Risk and Authorization Management Program (TX-RAMP), established in 2021 under S.B. No. 475, provides a standardized approach for security assessment, certification, and continuous monitoring of cloud computing services that process the data of Texas state agencies. Thus, third-parties looking to do business with the state must adhere to TX-RAMP standards, which includes AI vendors. Important lessons can be learned from what is occurring and working well at the state level as federal authorities consider new policies and streamline existing policies.

Nonetheless, given the growing number of federal and state cybersecurity policies that providers are expected to meet, a single set of policies that preempts the piecemeal and overlapping set of rules could bring greater efficiency for providers and free resources to make other needed cyber investments especially as the threat landscape grows more complex with the use of AI. If providers are to survive in a climate of growing cyber threats which are only increasing with the automation and sophistication from AI and one in which federal and state fiscal prudence and cost-cutting are prioritized, policies that support and reward providers for cybersecurity investments are needed, not ones that punish them. Castigating providers for cyber events outside their control will not improve their cyber posture.

Complicating matters further, until Congress reauthorizes CISA 2015, cyber threat sharing permissions enabled under the law will disappear, making it even harder for providers to guard against known cyber threats. With extremely limited resources and a regulatory climate that is highly punitive, providers will continue to experience headwinds. Reauthorizing CISA 2015 and reinstating CIPAC could help foster a more collaborative and secure environment for businesses and thus

enhance national security by bolstering the resilience and protection of the healthcare sector.

Another way that regulatory burdens could be lessened is by updating the CMS Stark and Office of the Inspector General (OIG) Anti-Kickback Statute (AKS) policies to clarify the provisions involving cybersecurity donation policies. Larger providers continue to express reluctance to donate cyber technology to other lesser resourced organizations because they fear they could incur liability if the smaller provider experiences a breach or cyber incident. These liability fears have deterred better resourced organizations from making these much-needed donations. Further, expanding the definition of what types of technology can be donated would also spur greater uptake.

PUBLIC-PRIVATE PARTNERSHIPS

The HSCC CWG has been instrumental in advancing the cyber posture of the healthcare sector and CHIME strongly supports these collaborative efforts. The HSCC CWG acts as a vital public-private partnership that brings together over 400 healthcare and cybersecurity professionals to strengthen the sector's resilience. As the recognized body representing healthcare within the nation's critical infrastructure framework, the HSCC CWG plays a pivotal role in developing actionable resources, fostering cross-sector coordination, and advancing cybersecurity as a shared responsibility. CHIME and our members actively contribute to the HSCC CWG's leadership and task groups, ensuring that provider voices shape national strategies and that federal policies reflect the operational realities of healthcare delivery. The HSCC CWG has launched a bold new task group devoted to AI with a strong focus

on helping organizations address emerging risks associated with the use of AI, drafting recommendations to mitigate those risks, as well as crafting guidelines, standards, and best practices to ensure safe use of AI in the healthcare sector.

Given the growing sophistication around AI and its weaponization to launch increasingly sophisticated attacks and threats to national security, it may be necessary to establish a new multi-stakeholder collaborative body focused on tackling these issues head on. This new entity could include federal agencies and private stakeholders such as the Open Worldwide Application Security Project (OWASP), a nonprofit foundation and global community focused on improving the security of software applications, and who are analyzing the threat vectors in

detail and providing threat data, guidance, development standards, and top 10 lists of vulnerabilities for various AI systems.

THIRD-PARTY SOLUTIONS

Providers continue to grapple with substantial vulnerabilities inherent in third-party solutions, underscoring the need for remedial action to safeguard the confidentiality, integrity, and availability of Americans' health data. Providers need confidence that the technologies they purchase are secure; however, in the absence of a uniform set of standards that must be met, our members are forced to conduct their own risk assessments when procuring essential medical devices, apps, and AI solutions. A rigorous review of third-party devices and applications is needed to ensure compliance and security best practices. Providers are saddled with these time-consuming and costly processes as they conduct their own assessments on the same technology being purchased by thousands of other providers, which is inefficient, redundant, and a waste of precious healthcare resources. When delivering updates to software that deliver new AI capabilities or updates or changes AI components, vendors should be required to disclose those AI components, the type of AI and the data used, as well as disclose their related security assessment including addressing relevant threats and vulnerabilities.

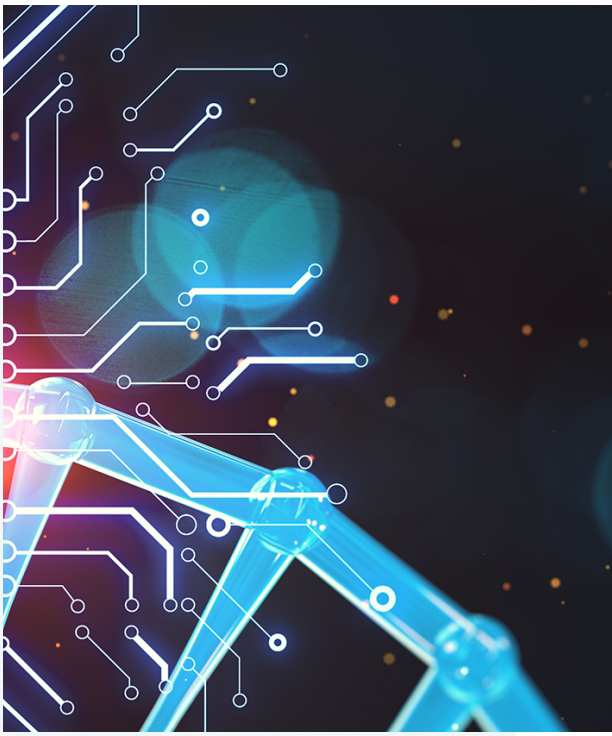
COMPETITION

By instituting a baseline set of security standards, healthcare providers will have greater confidence that what they are buying meets a baseline set of security standards. If manufacturers adopted standards, this could enhance their speed to market and improve competition and innovation. This could be done incrementally by focusing on one segment at a time such as managed detection and response (MDR). Furthermore, this would bring efficiency to the marketplace and assist under-resourced providers including small and rural entities

to purchase third-party products and services with confidence.

Today, providers are conducting these reviews using varying standards – creating inefficiencies for providers and vendors. While some providers have the buying power to require vendors to meet a baseline set of security controls (i.e., Indiana University of Health), this is far from commonplace and many vendors simply will not agree to this. In many cases providers must choose between two vendors performing similar functions – it is very hard to move from one solution to another which inhibits competition. Vendor solutions that prioritize security and can adhere to a standard set of security controls could help drive competition.

The release of America's AI Action Plan in July is aimed at positioning America first. Pro-business policies like establishing a baseline set of security standards and removing regulatory burdens and punitive policies for providers following a cyber incident are needed to foster competition and support providers.



Considering that hospital volume according to experts has been found to drop between 17-26 percent during the initial week following a ransomware incident and patient mortality is found to increase by 34-38 percent for patients admitting during the attack, the costs and safety implications to hospitals are enormous.

Congress passed the Protecting and Transforming Cyber Health Care Act (PATCH Act) in 2022 which gives the FDA additional authority over cybersecurity of medical devices, a step widely supported by providers. However, providers report some manufacturers are skirting the law and cost-shifting cybersecurity updates for medical devices onto them. The Federal Risk and Authorization Management Program (FedRAMP®) provides a standardized, reusable approach to security assessment and authorization for cloud service offerings, but this law

1

PRIVACY

Top Challenges and Recommended Solutions

01 Multiple Laws

Varying sets of state laws present complexity, burden and added costs to the healthcare system.

SOLUTIONS

- The U.S. needs a comprehensive national data privacy law to better protect consumers' sensitive health information and inform consumers of how their data is being used without duplicating what is already required under HIPAA.
- A federal privacy law and a federal AI law that preempts state laws are needed to reduce burden and improve efficiency.

02 HIPAA Oversight

Some vendors try to avoid risk by refusing to agree to certain provider terms and provisions in business associate agreements (BAAs) shifting more risk to providers.

SOLUTIONS

Greater oversight is needed to ensure that business associates (BAs) with access to protected health information (PHI) are meeting their obligations under HIPPA, and to promote a more equitable allocation of risk exposure between covered entities (CEs).

- Establishing a standardized HIPAA BAA could create greater contractual consistency and prevent hospitals and healthcare systems from disproportionately bearing liability.
- Vendors should not be permitted to skirt risk through opaque and questionable BAA contract practices.
- HIPAA's BAA requirements should be enhanced to include a clear, enforceable definition of de-identified data that explicitly prohibits its use in any manner that could lead to re-identification.

03 Health Data Residing Outside of HIPAA

Many entities who handle Americans' health data are either not governed by HIPAA or are not in compliance. The current federal legal and regulatory environment disadvantages both patients and healthcare providers by creating inconsistent protections across different types of health data.

SOLUTIONS

- Any entity with access to healthcare data should be required to meet and be held accountable for privacy and security standards pursuant to the Federal Trade Commission's (FTC) authority under the Health Breach Notification Rule (HBNR) and other policies.
- Those who are not required to meet HIPAA but have access to health data should employ robust privacy and security protocols similar to what is required of HIPAA CEs.
- Congress should consider passing a law that requires that PHI be handled as such regardless of whether the other party is a traditional healthcare provider or not.

Discussion

The amount of health data being created using AI is staggering in size. Advanced models analyzing x-rays, magnetic resonance imaging (MRIs), clinical notes, and genomic data contribute massive amounts of unstructured data.

This data is in addition to the data being created by smart devices that are continuously collecting health metrics. A single major hospital can generate upwards of 50 petabytes of data annually which translates to approximately 170 years of HD video footage according to the World Economic Forum. The exponential growth of health data requires heightened responsibility from healthcare providers to safeguard its security, integrity, and responsible application.

MULTIPLE LAWS

The existing state patchwork of privacy laws has created an enormously complicated legal landscape for healthcare providers and other stakeholders. Varying sets of state laws present complexity, burden and added costs to the healthcare system. Thus, the need for a national data privacy law has become increasingly important. Five years ago, there were five states with privacy laws enacted; today 20 states have privacy laws. As providers move to adopt AI and the types of tools they are adopting are more complex, the privacy challenges are only mounting. A federal law that preempts the myriad of state laws could bring greater efficiency to providers and reduce administrative burden and complexity.

Some states have also passed AI legislation further complicating an already challenging legal and regulatory environment. California's governor just signed into law SB 53, the Transparency in Frontier Artificial Intelligence Act (TFAIA), instituting AI guard rails. This fragmented approach at the state level can lead to inconsistencies in how patient data is protected and how AI tools are deployed. There have been efforts to flatten state AI laws by freezing states' ability to pass or enforce local laws and centralizing AI governance at the federal level though. The President's AI Action Plan instructed the Office of Management and Budget (OMB) to, "work with Federal agencies that have AI-related discretionary funding programs to ensure, consistent with applicable law, that they consider a state's AI regulatory climate when making funding decisions and limit funding if the state's AI regulatory regimes may hinder the effectiveness of that funding or award."

In the absence of a federal law on AI, and interest in AI at an all-time high, states continue passing their own laws which have addressed mental health AI tools (i.e., chatbots), AI in insurance and claims, clinical decision support systems, patient data privacy and algorithmic transparency. Similar to the need for a federal privacy law that preempts state laws, a single federal AI law that preempts state laws is needed.

HIPAA OVERSIGHT

HIPAA privacy protections extend only to HIPAA covered entities (CEs) which include providers, payers and healthcare clearinghouses, as well as any business associates (BAs) of these entities. Companies that have access to PHI and do business with a HIPAA CE are required by law to safeguard the data. However, some vendors try to avoid risk by refusing to agree to certain provider terms and provisions in BAAs, shifting more risk to providers. As providers ink agreements with AI vendors, they have reported vendors attempting to shift liability by refusing to accept key terms in the BAA or overriding restrictions in the BAA – exposing them to additional risk. Our members also report vendors using contracts that append the BAA as an attachment and use language in the main contract that overrides all other language as it pertains to data use. Some AI vendors have refused to sign BAAs with providers, while others have elected to train their models using de-identified data instead.

Although HIPAA requires a standard BAA, the regulatory language sets only a minimum threshold—stating that a BA may use or disclose PHI solely as permitted by the agreement with the CE. The challenge arises when vendors embed broad data use rights into the main service agreements, effectively circumventing the intent of the BAA. Many vendors assert that de-identified data falls outside HIPAA's scope, using this claim to justify expansive data use, often in ways that make providers uncomfortable. However, de-identification is highly subjective and context-dependent, and vendor language—such as assurances that "de-identified data will not be used in any way that could identify a natural human"—raises serious concerns about enforceability and the potential for re-identification.

TO STRENGTHEN PROTECTIONS, HIPAA'S BAA REQUIREMENTS SHOULD BE ENHANCED TO INCLUDE A CLEAR, ENFORCEABLE DEFINITION OF DE-IDENTIFIED DATA THAT EXPLICITLY PROHIBITS ITS USE IN ANY MANNER THAT COULD LEAD TO RE-IDENTIFICATION.

In a perfect world, the original intent of HIPAA would be reinforced: data should be used solely for the purposes outlined in the agreement and only in ways that directly support the services for which the provider has contracted.

Unfortunately, many providers cannot afford to walk away from certain vendors, placing them in an impossible bind—underscoring the urgent need for increased oversight by the Office for Civil Rights (OCR) to ensure that the spirit of HIPAA is upheld.

Examples of the challenges faced by our members are reflected in the below testimonials:

- *Organizations that are large and influential refuse to alter their BAA.*
- *Our legal teams spend 6+ months redlining and negotiating any data sharing provisions and sometimes we still end up not being able to use the technology.*
- *Another frustrating BAA experience was when we had a vendor who wanted to charge us for using our BAA. We walked away from the vendor.*
- *The vendor is literally brokering with you to get a product for free or deeply discounted if you will not do a BAA and allow them unbridled re-use of the data. The vendor is hoping some Executive will override legal and compliance to get the service they want cheap.*
- *That they can use our data for any purpose, mostly de-identified but sometimes not, and in fact some go so far as to say with them all rights and royalties to our data.*
- *Over liability clauses that supersedes anything else stated in the contract limiting their out of pocket liability to only what you paid for one year of service. As you can imagine that is often very low compared to the cost of a breach*
- *I'm looking at restricted use of health data and a BAA and the AI vendor said no to our terms. The vendor also wants us to sign an NDA.*
- *One vendor proposed capping their liability for a PHI breach at $50,000, citing the limited value of the contract. This completely ignored the downstream risk to the health system and our patients. We rejected the language, but the negotiation dragged on for weeks, consuming time and legal resources. Ultimately, I had to walk away from the partnership.*

As more data is being created across the healthcare ecosystem, a growing share of health data is now held by entities not required to comply with HIPAA.

HEALTH DATA RESIDING OUTSIDE OF HIPAA

This is due to the proliferation of consumer-facing technologies, applications, products, and services that access, produce, and manage health information but are not bound by HIPAA regulations. These entities include mobile health apps, fitness trackers, and other digital health wearables and tools – including genetic ones – that collect and store vast amounts of health data. The Federal Trade Commission (FTC) exercises oversight over non-HIPAA CEs that hold health data in the event of a breach under their Health Breach Notification Rule (HBNR), however the requirements are considerably narrower in scope than those imposed under HIPAA. The HBNR defines health data outside of HIPAA as "personal health record (PHR) identifiable health information" which includes data collected by apps, devices, and online services that are not CEs under HIPAA.

With federal policies governing health data differently depending on the entity holding or accessing it, there is a substantial risk that consumer data may be stored, processed, or transmitted without adequate protection. This includes the potential for sensitive data not subject to HIPAA to reside in offshore data centers, including those located in regions such as the People's Republic of China, where data access and oversight mechanisms do not align with U.S. privacy and security expectations. Some companies have also used health data for marketing and sold it for profit over privacy.

Enemy actors have also leveraged Americans' health data. This exploitation by hostile nation states can have severe implications for national security, including espionage, data theft, and network/system disruption.

Collectively, the lack of adequate privacy protections for health data can erode Americans' trust. It is CHIME's position that any data that is considered to be PHI should be handled as PHI regardless of whether the other party is a traditional healthcare provider or not.

1

DATA USE POLOCIES

Top Challenges and Recommended Solutions

01 Patient Consent

Providers are navigating how to best manage patient consent related to AI use.

SOLUTIONS

- Educate patients, clinicians and staff on the uses of AI and how it is being used to help deliver patient care.
- Align AI patient consent policies with your organization's risk tolerance.
- Engage legal and compliance experts to assess risk tolerance levels.
- Consent policies can:
  - Inform patients when AI is used to generate documentation or guide clinical decisions.
  - Be implemented using a single, overarching consent or notification form that explains AI use in clinical settings, reviewed during intake or initial visits.
  - Avoid overly technical language that could overwhelm or confuse patients.
  - Be made available to patients in accessible formats, such as patient portals or signs in care settings.
  - Be posted on patient portals using simple content that explains how these tools work and how they're used in care decisions.
  - Capture AI use clearly at the beginning of care (e.g., during intake), so patients know what and how it may inform clinical decisions.
- Create visuals and analogies that make tough concepts approachable. Patients shouldn't need a computer science degree to understand what's guiding their care.
- Involve ethics teams and patient advisory groups to define what's actually meaningful to share—this isn't a one-size-fits-all conversation.
- Consider adjusting disclosure practices based upon patient feedback and expectations and as AI use evolves.

02 Explainable AI vs. Interpretable AI

Providers are using AI in a variety of ways. How it is being used can drive how a provider manages risk and patient consent. A central challenge lies in distinguishing between explainable AI, which attempts to explain how an algorithm arrived at its answer or output, and interpretable AI, which allows users to more intuitively understand how an AI-driven decision was reached.

SOLUTIONS

- Use standardized, plain-language materials that explain AI's role in care and serves to contextualize use of explainable AI vs. interpretable AI.
- Build consent models that meet patients where they are—give them the option to get just the big picture or dive deeper if they want more details.
- Use consistent, plain-language terms to explain the difference between AI tools and solutions that help providers "understand how" and those that assist providers to "see why."

03 Clinician Awareness & Involvement

Clinicians may struggle to understand or explain to patients just how AI is being used.

SOLUTIONS

- Involve clinicians in the development and understanding of AI tools.
- Support clinicians with language and education that helps them explain AI's role.

04 Opt-out Policies

Enacting opt-out policies related to the use of AI can present significant complications for providers.

SOLUTIONS

- Opt-out policies should be avoided.
- Opt-out policies related to using deidentified data should be approached cautiously, as this could have negative impacts on research and medical science.

05 Secondary Uses

Providers see value in sharing data with AI developers. However, they are facing challenges navigating the liability associated with secondary uses of data – creating friction among providers and vendors.

SOLUTIONS

- Not all liability should be borne by providers.
- Providers must be able to determine how patient data they are responsible for protecting is used and re-used by third parties.
- Secondary use cases should be outlined in master service agreements.
- Trusted Exchange and Common Agreement (TEFCA) policies will need to be updated to account for increases in AI-generated data.
- Applying a framework, such as the Gartner AI trust, risk and security management (TRiSM) framework which aims to build trust, may be helpful.

06 Data Anonymization

Providers and clinicians are concerned that patient data used to train AI models must be properly anonymized, and clinicians need to know the data source.

SOLUTIONS

- Developing and implementing standardized procedures for anonymizing patient data is crucial.
- A mandatory protocol to validate the process of anonymization is needed.
- Providers and vendors should establish clear communication channels to ensure that clinicians and patients understand how AI algorithms are being used.
- AI vendors should, on at least an annual basis, provide proof to their customers (provider organizations) that their data is anonymized.
- This validation should be communicated or reported to the architect responsible for overseeing the anonymization process.
- OCR could issue iterative guidance on how using AI training data interplays with HIPAA requirements.

Discussion

The use of AI in healthcare is often seen as a way to reduce administrative burdens. It can also be used to streamline operations for medical conditions that are time-sensitive.

PATIENT CONSENT

As more providers adopt AI, they must determine how to handle informing the patients they treat about its use in their care. Determining how and when to inform patients, and to what degree it is being used is still an evolving area for providers. The integration of AI into clinical workflows — ranging from ambient listening tools (a voice recognition technology that captures a patient interaction with a clinician) to traditional clinical decision support systems to chatbots — raises difficult questions about what, when, and how to best disclose its use.

As AI tools such as clinical decision support systems, ambient listening technologies, and scheduling assistants become more prevalent in healthcare, it's increasingly important to clarify how patient consent applies under HIPAA's Treatment, Payment, and Healthcare Operations (TPO) framework—particularly since many of these tools may operate in the background or across multiple functions. For example, ambient listening used for clinical documentation may fall under 'treatment' or 'operations' depending on its use, while revenue cycle AI tools clearly support 'payment.' Clarifying these distinctions helps ensure that patients understand when their data can be used without explicit consent under TPO, and when additional authorization may be required.

In some organizations, consent is obtained at the beginning of the patient's care. The challenge lies in determining how much detail to disclose about AI and machine learning (ML) tools that can impact patient care. While some argue that disclosing every automation would overwhelm patients, others believe that complete transparency is crucial. To address this, some institutions have built an inventory of AI tools and share summaries with internal teams, with plans to make this information available to patients in the future. They use platforms, including their websites, to inform patients about AI usage rather than providing notifications for every instance. Surveys indicate that patients do not want to be notified every time AI is used. This underscores the need to consider whether a single consent form should be used to inform patients about AI usage. Providers continue to evaluate how extensively AI should be integrated into their organizations, as they are cautious about increasing the discoverability of sensitive information and are carefully considering the best approach to ensure patient consent and data privacy. As AI continues to permeate the healthcare delivery landscape, one member reflected, "getting consent for each and every use of AI will be wildly impractical."

Practices differ widely depending on institutional culture, risk tolerance, and community norms, leaving providers uncertain about just how much to disclose to patients and how often this should occur. There is a dilemma between providing meaningful transparency and avoiding information overload, with concerns about how far consent needs to go, especially as AI becomes more embedded and less visible in care delivery. Some providers believe patients want an overall awareness of the use of AI, rather than the highest level of specificity. For example, patients may not want to know that it is being used to help craft medical notes or documentation entry.

STUDIES HAVE FOUND THAT PATIENTS WANT TO BE TOLD THAT AI IS BEING USED TO HELP DELIVER THEIR CARE, AND SOME RESEARCHERS HAVE CONCLUDED THAT AI NOTIFICATIONS FOR PATIENTS SHOULD BECOME AN ETHICAL STANDARD OF CARE.

The results of a study by researchers at Duke University School of Medicine published in February 2025 in the *American Journal of Bioethics* concluded just this. It found that patient notification and consent should be tailored based on the AI tool's level of autonomy and clinical risk. Researchers recommended a flexible framework that ranges from broad institutional disclosures to fully informed consent for high-risk, autonomous applications. Another study published in *JAMA Network* in July 2025 found that patient comfort with ambient AI documentation tools varies based on trust, understanding, and perceived benefit. It concluded that a flexible, multimodal informed consent approach which includes education, digital tools, and opt-out options could ethically support adoption of ambient AI documentation tools in clinical care. This said, the degree to which providers share granular details remains an area of active consideration and discussion.

In considering consent policies, it is important to note that the term "AI" is extremely broad. AI is a language in itself – varying from natural language processing (NLP) to data aggregation tools. Importantly, there are differences between "predictive AI" and "generative AI" which must be recognized, as they are two distinct types of AI, each with different functionalities and applications. They are also used in two very separate and distinct areas of an HCO - administrative and clinical. It is important to recognize that providers are using AI in a variety of ways. How it is being used can drive how a provider manages risk and patient consent.

Patient education will also need to take into account the appropriate literacy level. According the U.S. government, 79 percent of American adults have low literacy levels. Therefore, consent and education policies will need to be tailored to the average reading level. Medicare tailors their materials to 5th to 7th grade levels. The Agency for Healthcare Research and Quality (AHRQ) has guidance on how best to structure health information which providers can reference.

EXPLAINABLE AI VS. INTERPRETABLE AI

As AI becomes increasingly embedded in clinical workflows, healthcare providers face mounting pressure to modernize patient consent policies. A central challenge lies in distinguishing between explainable AI, which attempts to explain how an algorithm arrived at its answer or output, and interpretable AI, which allows users to more intuitively understand how an AI-driven decision was reached. While both frameworks aim to enhance trust and transparency, they present different implications for consent. Explainable AI often requires complex, technical justification that may overwhelm or confuse patients, whereas interpretable AI lends itself to clearer communication but may be limited in scope.

CLINICIAN AWARENESS & INVOLVEMENT

According to the American Medical Association (AMA), in 2024 almost 70 percent of doctors were using AI as compared to just 38 percent in 2023.

Yet, clinicians and providers are grappling with how much detail is ethically—and practically—appropriate to disclose. Should patients be informed every time an AI tool influences a diagnosis or treatment recommendation? Is a high-level explanation sufficient, or do organizations have a duty to disclose underlying algorithmic logic? These questions are compounded by regulatory uncertainty, variable institutional comfort levels, and patient expectations that continue to evolve. Without consistent standards or public consensus, providers risk either overburdening patients with unnecessary technical details or under-informing them, potentially eroding trust and accountability in care delivery.

Studies highlight the importance of clinician participation in the development and understanding of AI tools. A *British Journal of Medicine* (BJM) April 2025 article found nine clinical processes where AI is used including: "(1) identifying clinical problems suitable for AI solutions; (2) forming project teams or collaborating with experts; (3) organizing and curating relevant data; (4) establishing robust physical and virtual infrastructure, and computer systems' architecture that support subsequent stages; (5) exploring AI neural networks on open access platforms before making a new decision; (6) validating AI/ML models; (7) registration; (8) clinical deployment and continuous performance monitoring; and (9) improving the AI ecosystem ensures its adaptability to evolving clinical needs." Among the conclusions reached by researchers was that "clinical integration and implementation of AI tools must include building trust and confidence among clinicians in the development process."

Resources exist to help providers navigate the AI journey. The Health AI Partnership offers tools to help providers navigate the purchase and deployment of AI including designing and testing clinician workflow and disseminating information to end users.

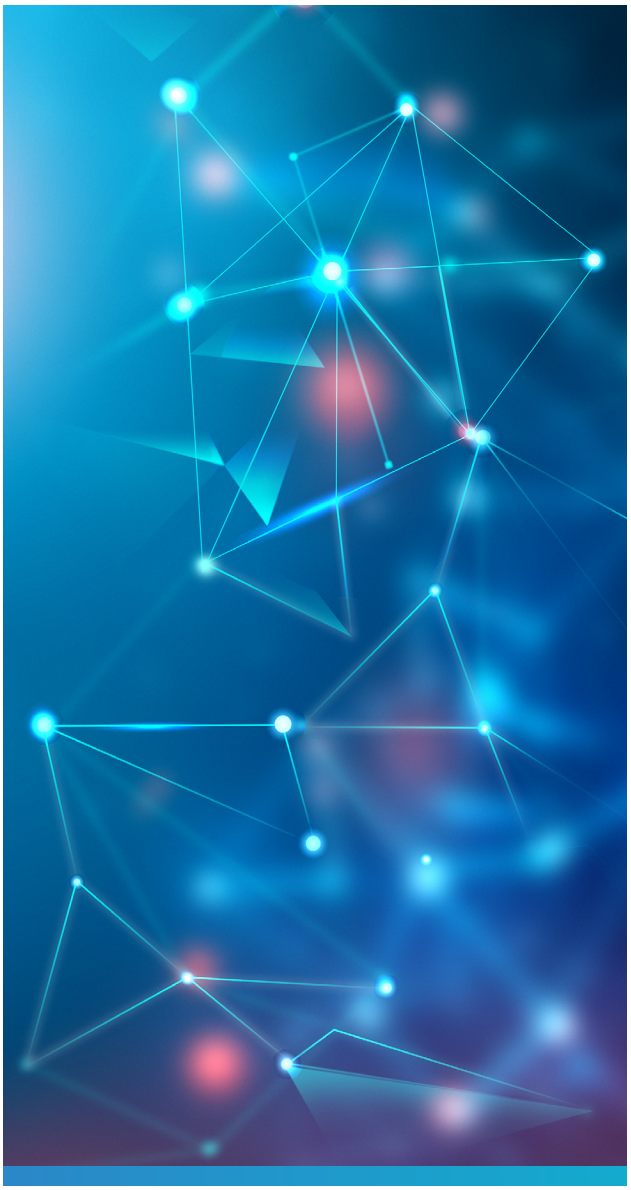
OPT-IN & OPT-OUT POLICIES

Most providers today operate under an opt-out data model for data exchange purposes. Opt-in policies require explicit patient consent before data can be shared or used including by AI systems.

As AI adoption increases, opt-out policies could present insurmountable barriers and "opt-out for AI will mean opt-out for treatment," as one member explains. As electronic health records (EHRs) continue building AI into their systems natively, parsing through opt-out policies will become even more complicated. Opt-out policies related to using deidentified data should be approached cautiously, as they can have significant negative impacts on research and medical science. Deidentified data is crucial for various types of important research and advancements in medicine, including clinical studies, policy assessments, and life sciences research. When large numbers of individuals opt-out of allowing their deidentified data to be used, it can undermine the quality and validity of the data on which this research depends.

Some states have already adopted opt-out policies, and several have also adopted policies that permit patients "to be forgotten." For instance, the California Consumer Privacy Act (CCPA), as amended by the California Privacy Rights Act (CPRA), includes provisions for the right to be forgotten and opt-out policies. This law allows consumers to request the deletion of their personal data and opt-out of the sale or sharing of their data.

It is also worth noting that under the voluntary Trusted Exchange and Common Agreement (TEFCA) designed to facilitate seamless and secure health information exchange across the United States, consent policies are standardized across different Qualified Health Information Networks (QHINs). However, state laws that require opt-in consent for data sharing can impact



how TEFCA policies are implemented. In practice, this means that if a state requires opt-in consent for health data sharing, QHINs and provider Participants operating under TEFCA must ensure that they obtain explicit consent from patients before including their data in the exchange. With the increase in the amount of AI-generated data and increased adoption of TEFCA, crafting workable data sharing policies will become further complicated.

SECONDARY USES

As providers increasingly work with vendors to train their algorithms, providers see tremendous value in data sharing – when certain safeguards are in place.

HCOs report significant concerns related to legal and regulatory liability associated with secondary uses of data, especially with the rise of AI model training. The training involves large datasets that teach AI systems to recognize patterns and make predictions. Healthcare providers are particularly concerned about these secondary uses of data because they could lead to unauthorized sharing or misuse of sensitive patient information, which would compromise privacy and trust. For instance, should transcripts from clinician visits be included in training data? Will diagnoses ruled out by a clinician be included? If yes, how do vendors and third parties intend to use this data? How can liability be managed when a vendor or an academic medical center trains their algorithm on a patient population that is more homogenous but is then deployed in a healthcare setting with greater patient heterogeneity? How can providers adequately manage risk when vendors express reluctance to certain terms and conditions in a master service agreement? What if a vendor wants to train their algorithm using a provider's data but is unwilling to agree to liability terms or share risk?

One way to approach these challenges is by applying a "data-first approach" like the one developed by Gartner. The Gartner model has been used by some providers to help them navigate trust and accountability challenges. Gartner recommends that organizations adopt explainable and interpretable AI frameworks to support responsible data sharing, emphasizing that AI systems must be transparent in how they use and process data to build trust, ensure accountability, and comply with ethical and regulatory standards. This includes prioritizing data quality, documentation, and model transparency to enable meaningful oversight and stakeholder understanding.

DATA ANONYMIZATION

Vendors are increasingly looking to improve their algorithms by using patient data held by providers to train their models. There is much value in allowing vendors to train their models on provider-held data sets as it can help improve the quality of the model. However, there is a strong need for more transparency, and providers need to understand how vendors are using patient data.

Without a specific procedure or set of rules to ensure that the anonymization process is correctly implemented and verified, providers have no way of confirming whether an AI algorithm is performing as intended. Providers need assurance that if a clinician contributes to AI training, that it is considered "Healthcare Operations" and it is therefore protected under HIPAA and whether contributing their data will help improve patient care.

Some providers have also expressed concerns that by sharing their data they are surrendering their intellectual property (IP). For instance, AI tools in oncology often rely on large volumes of sensitive patient data, including genomic information and imaging. Cancer centers worry that sharing this data externally – particularly with commercial AI vendors – could expose proprietary research methods, clinical insights, or unique datasets that represent competitive advantages.

SECTION 2

Fostering Patient Centered Care

2

SAFETY

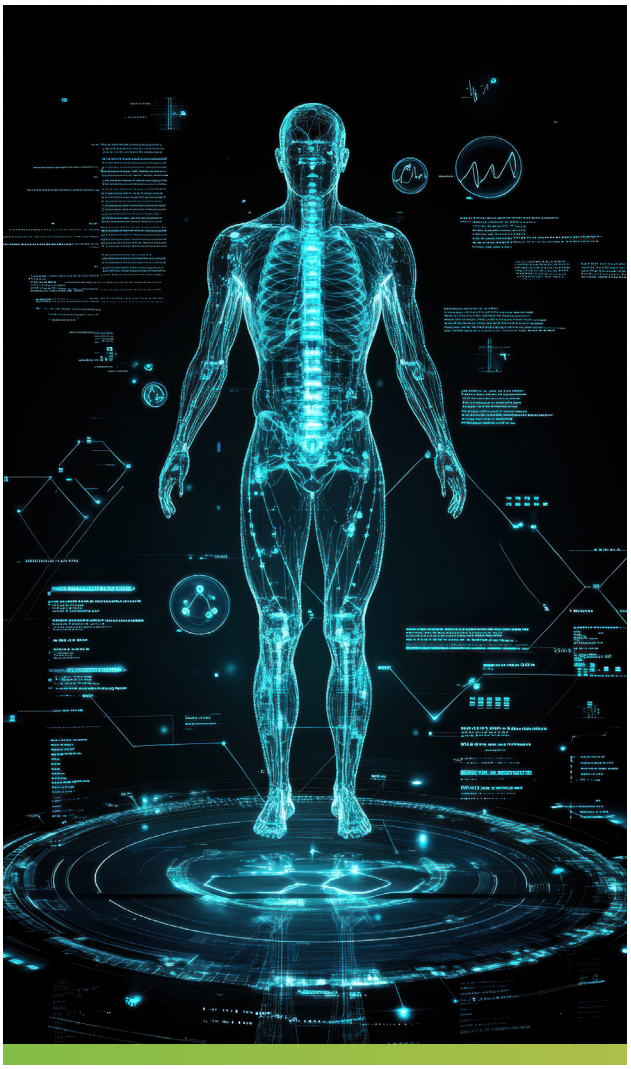
Top Challenges and Recommended Solutions

01 Frameworks & Standards

AI in healthcare can negatively affect patient safety and outcomes if not carefully implemented and managed.

SOLUTIONS

- Ensuring patient safety must be a shared responsibility.
- AI should be developed in a way that prioritizes trust and safety.
- Human intervention is needed before, during, and after AI is put into use.
- AI tools should augment, not replace, human clinical decision-making.
- Emerging AI frameworks that offer templates on how to move forward can be helpful to providers.
- Providers should consider voluntary adoption of the National Institute of Standards and Technology (NIST) AI Risk Management Framework (AI RMF).
- Providers should be encouraged to review other resources like the National Academy of Medicine AI Code of Conduct, the Trust, Identity, Privacy, Protection, Safety, and Security (TIPPSS) Framework, and the Health AI Partnership (HAIP) Health Equity Across the AI Lifecyle (HEAAL) Framework.
- Structures and processes to report potential safety events related to the AI tools are needed.



02 Oversight

Generic laws and policies that do not consider the unique needs of the healthcare sector can hinder innovation and effectiveness.

SOLUTIONS

- Deliberate policymaking is needed. Fine-tuned policies can help improve oversight and transparency for providers.
- Laws must meet providers where they are while also looking ahead (i.e., be iterative).
- Redundant laws and policies should be avoided.
- Policymakers should carefully consider the unique aspects of the healthcare sector to avoid negative impacts on patient care and innovation.
- Third party AI tools need to be audited, and oversight of these tools is needed to ensure they are performing as intended.
- The level of oversight needed should be matched to the level of risk to patient safety. A higher level of oversight may be needed for some forms of predictive AI or if it is directly impacting a patient's health outcomes.
- Oversight of AI algorithms that fall outside of the jurisdiction of Assistant Secretary for Technology Policy / Office of the National Coordinator (ASTP/ONC) and FDA are needed.
- As laws are crafted, policymakers should look to existing, provider-supported frameworks which are complimentary.
- A "Better Business Bureau" type of resource of AI products which lists solutions that have been determined to have met certain standards could be helpful.

03 Safe Harbors

Providers are concerned that with closed systems and without the ability to review an AI algorithm, they will incur additional liability for patient outcomes and safety.

SOLUTIONS

- Providers need liability protections to shield them from legal risk when using FDA-authorized AI tools that offer limited transparency into their underlying algorithms.
- Safe harbors could help speed AI adoption by removing liability fears.
- Clear standards for responsible use are needed, especially when AI recommendations are overridden or not followed.

04 Assurance Labs

HCOs need information about how the AI products were tested before purchasing them.

SOLUTIONS

- Providers must have ongoing monitoring of safety and quality of the AI products at a regular cadence as long as they are used to provide patient care.
- Assurance labs may offer a way to help ensure that AI products are performing as intended.
- Before providers can truly rely on AI assurance labs, they must be validated through standardized criteria that demonstrate their ability to reliably assess the safety, fairness, and effectiveness of AI tools across diverse clinical settings.
- Assurance labs should not override the need for transparency of large language models (LLMs) used to train an AI tool. Additionally, they must have ongoing monitoring of safety and quality of the AI products at a regular cadence as long as they are used to provide patient care.

Discussion

Referred to as the "AI Godmother, pioneering researcher Fei-Fei Li, PhD, Founding Co-Director of Stanford's Human-Centered AI Institute has advocated for AI that works to serve humanity.

In her book, "*The Worlds I See*," Dr. Li writes:

> *"If we were to broaden our vision for AI to explicitly include a positive impact on humans and communities—if our definition of success could include such things—I was convinced that AI could change the world for the better. I still am."*

FRAMEWORKS & STANDARDS

As HCOs chart their AI course they are looking for ways to deploy and use AI in an ethical manner. Many HCOs have AI councils, acceptable use, and/or governance policies established within their organizations, though rural, safety-net, and smaller providers are either just getting started or have not yet taken the leap. Providers may also be served by having structures and processes to report potential safety events related to the AI tools. In the beginning, this could be an internal process.

There are emerging AI frameworks and standards that serve as a good starting point for HCOs embarking on or refining their AI adoption journey. They are intended to enable responsible, ethical, and effective advancements in health AI on behalf of the patients and communities they serve. They include:

- The NIST AI Risk Management Framework (AI RMF). Developed in conjunction with the private sector, "it is intended for voluntary use and to improve the ability to incorporate trustworthiness considerations into the design, development, use, and evaluation of AI products, services, and systems." America's AI Action Plan calls for revising the existing Framework.
- Providers are already deeply familiar with the NIST Cybersecurity Framework (CSF) and have voluntarily adopted it. Released in 2023, the NIST AI RMF offers a complementary framework to better manage risks to individuals, organizations, and society associated with AI.
- NIST is also developing a way to reach consensus faster around AI standards development – typically a lengthy process – through their "Zero Drafts" pilot project.
- The National Academy of Medicine's Artificial Intelligence Code of Conduct is "aimed at providing a guiding framework to ensure that AI algorithms and their application in health, health care, and biomedical science perform accurately, safely, reliably, and ethically in the service of better health for all."
- The Health AI Partnership (HAIP) is a multi-stakeholder, provider-driven effort spearheaded by Duke University. Their Health Equity Across the AI Lifecyle (HEAAL) Framework focuses on empowering providers to safely and fairly use AI through the use of

standards to focus on procurement, development, clinical integration, and lifecycle management.

- The Coalition for Health AI (CHAI) is a multi-stakeholder group co-led by vendors and providers founded by MITRE which is now independent. Their Blueprint for Trustworthy AI in Healthcare outlines principles and practices for the responsible development, deployment, and oversight of AI tools in clinical settings.
- The AI Maturity Roadmap developed by the AI Collaborative that includes vendors and leaders from several health systems. It contains six focus areas including: culture; governance; business implementation; value; maintenance and operations; and information architecture. This roadmap focuses on a maturity model giving providers to benchmark their progress.
- The TIPPSS Framework—which stands for Trust, Identity, Privacy, Protection, Safety, and Security—is a cross-standard safety model designed to guide the responsible development and deployment of AI in healthcare and other critical sectors. Several major standards and frameworks apply the TIPPSS framework including:
    - The IEEE UL 2933-2024 is a standard focused on the Clinical Internet of Things (CIoT) data and device interoperability. It embeds TIPPSS principles to ensure that connected medical devices—including those powered by AI—are secure, interoperable, and trustworthy across hospital systems, EHRs, and wearable technologies.
    - The ISO/IEC 42001, is "an international standard an international standard that specifies requirements for establishing, implementing, maintaining, and continually improving an Artificial Intelligence Management System (AIMS) within organizations."
    - The OWASP GenAI Security & Privacy Guide which provides "clear and actionable insights on designing, creating, testing, and procuring secure and privacy-preserving AI systems."
    - The Trustworthy Technology and Innovation Consortium (TTIC), a cross-standard safety model that aligns with IEEE UL 2933, OWASP GenAI, NIST AI RMF, and ISO/IEC 42001.

OVERSIGHT

AI tools have the power to augment – not replace – clinicians to improve patient care. However, safety, transparency, and appropriate oversight policies are needed to ensure this occurs. To ensure the AI tools that providers are purchasing and using are performing as intended and that they support patient safety, oversight is necessary. Policies that are already in place for other technologies relied upon to treat patients could be applied to AI as well. For instance, emergency department alerts and order sets must be audited before being put into use by a clinician on a patient.

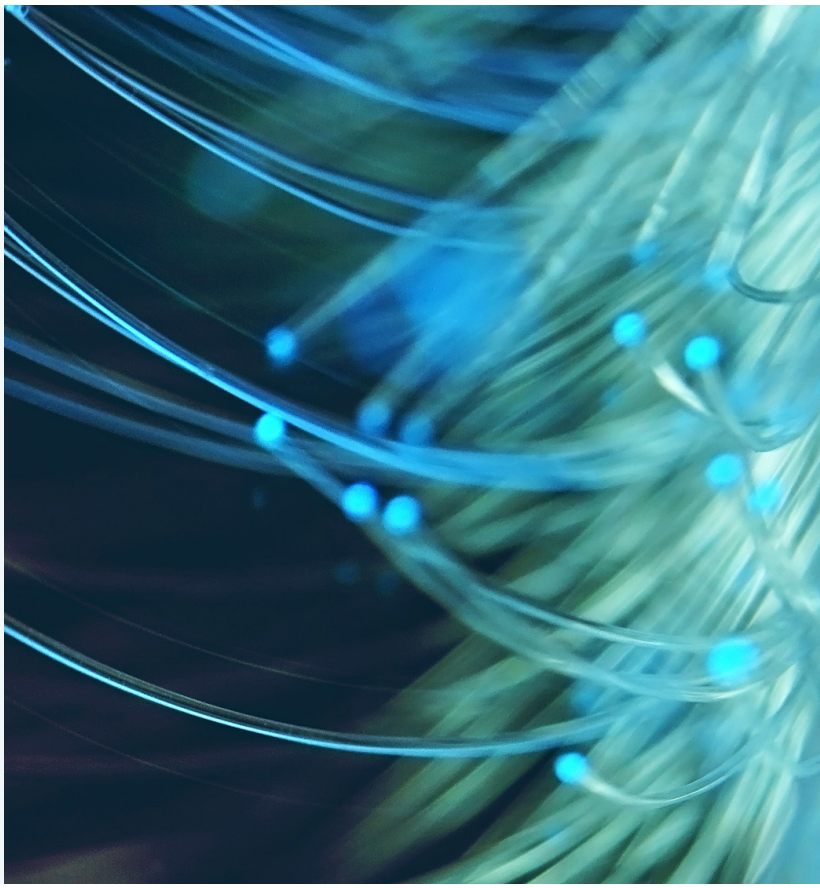
HOSPITALS TYPICALLY AUDIT ORDER SETS FOR SAFETY THROUGH A STRUCTURED, MULTIDISCIPLINARY GOVERNANCE PROCESS THAT BLENDS CLINICAL OVERSIGHT, INFORMATICS, AND REGULATORY COMPLIANCE.

While oversight is needed, not all forms of AI will necessitate the same level of scrutiny. Whereas explainable AI, predictive AI and interpretive models can be measured and monitored, generative AI models pose greater challenges due to its inferential nature.

Explainable AI intended to help a clinician better understand the rationale behind an AI decision / output and interpretable AI which analyzes outputs like images and patient conversations, pose lower risk levels than predictive AI used to arrive at a clinical outcome. For instance, AI embedded into ambient listening software intended to help a clinician better understand the rationale behind an AI decision / output (explainable AI), poses a lower risk level than AI that uses data to predict (predictive AI) a patient's diagnosis, deliver personalized medicine, predict a drug's side effects, or determine if a patient needs to be admitted to the hospital. Therefore, AI tools used for clinical decision-making necessitate that the level of oversight be higher since the impact to patient care, outcomes, and safety is greater. With momentum building as more providers adopt AI, the demand for transparency and trust will only grow.

Providers are already knowingly and willingly accepting risk for self-developed software from a practice liability standpoint, and these tools undergo extensive testing processes and internal validation to ensure safety, reliability, and compliance. However, when adopting third-party AI tools – especially those integrated into clinical workflows – providers can face additional uncertainty around liability, transparency and oversight, particularly if the vendor's development and testing processes are opaque or not aligned with clinical standards.

To date, federal oversight has been limited to reviews conducted by the FDA related to software as a medical device (SaMD) and oversight by ASTP/ONC of predictive algorithms and other AI that is part of certified EHRs. AI that falls outside of these areas, however, is significant and exists with little oversight.



The FDA regulates SaMD if it provides direct diagnostic or treatment recommendations. There are concerns, though, that the FDA approval process for SaMD for AI is not as rigorous as it should be, and there is currently not a way to validate the use of "off-the-shelf" products in hospitals and other provider settings following FDA approval.

It has also been found that some devices with AI capabilities, but for which AI is not the primary function, are reviewed by the FDA through a more traditional pathway, such as continuous glucose monitors (CGMs). Thus, these devices are not subjected to the same level of scrutiny as those deemed exclusively "AI," directly impacting the level of post-market oversight. While all devices are subject to post-market surveillance, the requirements vary depending on the risk classification and regulatory pathway used for approval. One study found that only 28 percent of devices approved by the FDA for AI use were prospectively validated, meaning they were tested in the real world prior to approval.

The FDA has – as of August 2025 – approved more than 1,200 medical devices with AI capabilities, though none of the devices approved have either generative AI capabilities or act as LLMs. According to a recent study published in Nature, researchers found that when LLMs were evaluated to determine how they performed for the purposes of clinical decision-making, they showed "considerable promise for clinical decision support (CDS)." It is unclear if the FDA's medical device approval process is currently designed to handle review of these models.

In 2025, the FDA introduced draft guidance tailored to AI-enabled Device Software Functions (AI-DSF), emphasizing a Total Product Lifecycle (TPLC) approach which includes: design and development; standards; validation and performance monitoring; and post-market surveillance to ensure continued safety and effectiveness. The new guidance sets expectations for developers around risk assessment, data management, bias mitigation, cybersecurity, and public transparency. However, the draft guidance is not yet final, and therefore, not yet in effect. Further, even if it is finalized, it is considered "non-binding" pursuant to the way the agency typically approaches their oversight authority.

It is also worth noting that because the FDA only regulates AI that serves as a medical device, there are AI tools that are not regulated by the agency, including those which typically support administrative tasks or provide general health information, rather than directly influencing a provider's medical decisions. The agency itself notes that: "The FDA's traditional paradigm of medical device regulation was not designed for adaptive artificial intelligence and machine learning technologies. Many changes to artificial intelligence and machine learning-driven devices may need a premarket review."

ASTP/ONC regulates predictive AI as part of a certified EHR (CEHRT), which typically include administrative systems and certain clinical support software. These policies are included in the Health Data, Technology, and Interoperability (HTI-1) final rule (summary here). However, their authority only extends so far as the AI vendor also has a certified product on ASTP/ONC's Certified Health IT Product List (CHPL). For instance, while some ambient listening products are certified by ASTP/ONC, not all are. Even if an AI tool is integrated within a certified EHR, unless the AI vendor has a product listed on the CHPL, ASTP/ONC does not have oversight over them. Thus, some algorithms may not be subject to any HHS oversight.

Further complicating things, the FTC also maintains oversight over a large swath of U.S. businesses under the Federal Trade Commission Act, which regulates competition and unfair and deceptive business practices which could include AI models. As one member opined, "Remember, FDA doesn't govern all AI. Some falls under FTC. This is the age-old issue of where FDA stops and FTC comes in and vice versa."

THERE IS GROWING EVIDENCE THAT HEALTHCARE PROVIDERS FEEL INTIMIDATED AND OVERWHELMED BY THE PATCHWORK OF FEDERAL AND STATE LAWS GOVERNING THE USE OF AI IN CLINICAL SETTINGS.

This regulatory complexity is creating significant challenges for adoption, compliance, and trust. The 2025 Black Book Survey Highlights Growing Divide Over AI Regulation in U.S. Healthcare found 73 percent of providers believe AI tools are being implemented without adequate clinical validation, and 85 percent of healthcare professionals cited a lack of clarity around liability when AI tools are involved in patient harm. Audits could be used to ensure that AI is performing as intended. For instance, when HCOs deploy AI-enabled medical devices—such as gastrointestinal systems that score polyp detection—providers often lack access to these closed systems and cannot review or validate how the algorithm makes its recommendations. Given how heavily healthcare providers are regulated and the criticality of patient safety, they could also benefit from a "Better Business Bureau" type of resource where listed AI solutions have been determined to have met certain standards. Providers – particularly small and under-resourced – do not have the resources, workforce, and time to thoroughly research each AI tool or product from a safety and security standpoint, again slowing uptake and adoption.

SAFE HARBORS

A chief complaint among CHIME members is that too many models are "black boxes," making it impossible for providers to audit them to determine if they are performing as intended. This can deter clinician adoption. As one member stated, "If you don't tell a physician why they get a risk score they won't use it." And, from a security standpoint, if providers are not able to ascertain where data is being stored, they are less inclined to trust it. These tools often lack transparency, leaving clinicians without insight into how treatment recommendations are made. Providers are particularly concerned that many of these algorithms are effectively "frozen in time"—unable to adapt to evolving clinical knowledge or patient populations—raising serious questions about safety, accountability, and long-term reliability.

The lack of transparency has created a host of issues for providers including patient safety, liability and risk. With nearly half of physicians expressing the need for greater oversight, including by the FDA, and uncertainty around a regulatory pathway, adoption of AI models could be dampened. Safe harbors for providers could alleviate some of this pressure when combined with improved transparency and clear standards for responsible use. One hypothetical example of a safe harbor could be requiring documentation of AI-generated output alongside the clinician's decision to follow or override it—providing both transparency and a defensible record of clinical judgment.

Providers are accustomed to being able to review medical journals to help justify the use of a medical device or procedure. However, this is not possible with the use of AI models, raising significant safety

and liability concerns. Without this same level of review, providers will continue to be constrained in exercising their medical judgement when deciding whether the benefits of adopting an AI model outweigh the risks to patient safety and outcomes.

Given the black box nature of these algorithms, providers continue to express the need for a "safe harbor" that helps ensure they are not bearing the liability for the failure of an algorithm or poor outcome stemming from a data training set that is outdated, or does not account for the provider's patient population. This lack of transparency is hindering provider adoption when they are not privy to how the algorithm is making clinical recommendations.

Providers alone cannot be unilaterally responsible for patient safety. Just as cybersecurity is a collective responsibility, the healthcare ecosystem must also work together to ensure that AI is protecting patients. As one CHIME member put it, using AI should be tied "back to the Hippocratic oath – do no harm. What is good medicine? That should be a basic premise."

ASSURANCE LABS

Providers need assurances that the AI products they are purchasing are going to perform as intended and can deliver favorable patient outcomes, prioritize patient safety, and meet the needs of their patient populations. And, they must have ongoing monitoring of safety and quality of the AI products at a regular cadence as long as they are used to provide patient care.

AI assurance labs have emerged as a way to help ensure safe, effective, and ethical deployment of AI technologies in clinical settings. These labs are intended to serve as independent environments where AI models – especially those used for decision support, diagnostics, and documentation – can be rigorously tested, validated, and continuously monitored.

Both HAIP and CHAI are actively addressing assurance lab needs. HAIP's Algorithm-Based Clinical Decision Support (ABCDS) Oversight initiative, is Duke Health's effort

which they say, “ensures that algorithmic health tools are carefully validated during creation, closely scrutinized throughout their lifecycles, and continuously monitored after deployment.” This includes maintaining a registry of AI tools, conducting evaluation checkpoints across the AI lifecycle, and ensuring alignment with Duke Health’s safety and governance standards.

Collectively, assurance lab efforts may help HCOs assess risk, meet evolving standards, and build systems that clinicians and patients can trust. As AI becomes more embedded in care delivery, assurance labs could offer a structured way to manage its lifecycle, mitigate harm, and promote fairness across diverse populations and use cases. Before providers can safely rely on AI assurance labs though, they must be validated through standardized criteria that demonstrate their ability to consistently assess the safety, fairness, and effectiveness of AI tools across diverse clinical settings. The Joint Commission (TJC) has partnered with CHAI to co-develop an AI playbook, tools and a certification program that combine CHAI’s AI best practices with the Joint Commission’s evidence-based standards to “scale the responsible use of AI.” CHAI, together with the Joint Commission, released their first set of AI guidance in September 2025 designed to support responsible AI adoption. They have stated that they intend to follow this effort with a voluntary AI certification program. Given that the Joint Commission accredits nearly 15,000 healthcare organizations, its potential role could raise concern about transparency, consistency, and the appropriateness of delegating oversight to a private accrediting body.

While assurance labs have potential to play a valuable role in evaluating AI tools, they should not be a substitute for transparency about the LLMs used to train those tools. For example, many medications lack sufficient research data for individuals over the age of 80, yet AI systems may generate recommendations based on data from younger populations—typically those aged 45 to 65. Applying these insights to older patients may not always be appropriate. Clinicians need visibility into how an LLM was trained in order to assess whether its recommendations are relevant and safe for the specific patient in front of them. This level of transparency should complement, not replace, any testing conducted by an assurance lab—and ideally, such limitations should be clearly flagged within the lab’s evaluation.

Some policymakers have expressed concern that CHAI’s approach to assurance labs – which is aligned to the HHS AI strategy released in January 2025 – could present regulatory capture, the notion an external entity is taking on a responsibility that is better handled by the government. Others have raised concerns that dominant industry players could unduly influence CHAI’s standards and compromise objectivity, create conflicts of interest, erode public trust, lack balanced stakeholder representation, and should focus more on intellectual property.

2

SUPPORTING PATIENT OUTCOMES

Top Challenges and Recommended Solutions

01 Managing Different Patient Populations

Due to limitations in traditional care models, data representation, and technology design, optimizing patient outcomes for diverse patient populations can be challenging.

SOLUTIONS

- Transparency and safeguards are needed to ensure AI tools support fair and outcomes-based care.
- AI models should be stress tested to ensure non-discriminatory performance across diverse patient population subsets.
- Since bias is nearly impossible to eliminate, the goal should be to acknowledge its presence and implement safeguards that promote fairness.
- Recognize the difference between “noise” vs. “bias,” where bias refers to systematic errors that consistently skew data in a particular direction, while noise refers to random errors or variability that obscure patterns without a consistent effect.
- Training AI systems on representative datasets or ensuring transparency related to the population it was trained on is critical.

02 Human Oversight and Clinical Responsibility

AI tools are subject to hallucination.

SOLUTIONS

- Healthcare providers might place too much trust in AI systems, failing to apply their own clinical judgment and potentially leading to poor decisions.
- Treat AI adoption as an ongoing process, requiring continuous refinement and adaptation—not a one-time setup.
- Establish regular review and monitoring protocols to catch errors early and ensure AI tools remain safe, effective, and fair.
- Promote fairness in AI use by combining diverse, representative data with clinician training, strong governance structures, and thoughtful policy frameworks.
- Build feedback loops that allow for iterative improvements based on real-world performance and user experience.
- Ensure transparency and accountability in how AI systems are developed, deployed, and evaluated.

Discussion

Delivering care in a manner that results in the best outcomes for patients while remaining fair is a complex and evolving challenge.

MANAGING DIFFERENT PATIENT POPULATIONS

Providers are actively learning and adapting, recognizing that underlying data embedded in foundational systems such as EHRs and medical technologies—including pulse oximeters, infrared thermometers, and x-ray equipment—are often calibrated using populations that may not be representative of certain patient groups. Medical professionals, specialty societies, patient advocacy groups, and other stakeholders are working diligently to address these challenges. This includes ongoing efforts to improve the fairness of AI algorithms by ensuring they are trained on diverse and representative datasets.

It is hard to remove all instances of bias in AI systems which can lead to unequal treatment and outcomes for certain populations. According to a study published in Nature Medicine, researchers found that after studying nine LLMs involving 1,000 emergency visits that even when the clinical details were identical, the AI sometimes changed its recommendations based on a patient's background—like socioeconomic status or demographics—affecting decisions around triage, testing, treatment, and mental health. This underscores the need for transparency and safeguards to ensure these tools support fair and consistent care.

Since bias is nearly impossible to eliminate—whether in AI systems or among researchers and clinicians themselves—the goal should be to acknowledge its presence and implement safeguards that promote fairness. Even when using ambient voice technology to generate clinical notes, such as in orthopedic specialties with limited diversity among practitioners, the influence of gender, experience, and speech patterns can shape both the AI's output and the assessor's interpretation. For instance, a female physician evaluating notes in a male-dominated specialty may notice differences that reflect underlying structural imbalances in training data. These challenges raise questions about how to ensure fair assessments and whether voice recordings used for secondary research can ever be truly de-identified. Importantly, calling for transparency and documentation—such as noting when AI output is used and how clinical decisions are made in response—is not a demand for massive datasets, but rather a practical safeguard that can be applied regardless of the size or scope of the training data.

When structural limitations in data inputs exist and when training data lacks diversity – such as being skewed toward younger or male patients – then applying those models to underrepresented populations (e.g., elderly patients or those with speech impairments) can lead to inaccurate or inequitable outcomes. It is therefore important to train AI systems on representative datasets or restrict their use to populations for which they were designed. This is especially critical in clinical decision support, where assumptions based on incomplete questioning or inference can introduce both bias and noise. A good example of bias vs. noise is Daniel Kahneman, PhD's analogy of a faulty bathroom scale discussed in a *New York Times* guest essay. In this example, Nobel Prize winner Dr. Kahneman says that "bias" is like a bathroom scale that regularly overstates your weight, vs. "noise" when a scale gives you different readings each time. To ensure safe and effective AI use in healthcare, there is a need to clearly define and differentiate noise, bias, and hallucinations. Noise stems from random variability in data and can obscure patterns; bias arises from skewed data or flawed assumptions, often impacting the accuracy of predictive models like those used in diagnosis; and hallucinations—unique to generative models like LLMs—occur when the system produces confident, but false or fabricated outputs not grounded in the training data.

HUMAN OVERSIGHT AND CLINICAL RESPONSIBILITY

AI tools are viewed as assistive, not replacements for clinicians. While hallucinations in generative models may be less frequent, they still occur – such as when an AI tool infers that a diagnosis was ruled out when it wasn't or when the generated note misgenders a patient. Those studying AI at MIT found that AI content can generate false results, noting that generative AI models' "goal is to generate plausible content, not to verify its truth." Research from the *International Journal of Health Sciences* in 2024 describes AI hallucinations in healthcare and life sciences, finding "significant challenges, especially in contexts where precision and reliability are paramount" and the "potential to undermine trust and efficacy in critical domains such as healthcare and legal proceeding." The researchers offered the following example:

[I]f a user asks for a recommendation for a vegan dish and the AI suggests a recipe that includes meat, this would be an input-conflicting hallucination. In clinical settings, such hallucinations could lead to recommendations for medications or treatments that contradict a patient's known allergies or health conditions, thereby posing serious health risks.

Differences in physician communication styles (e.g., verbose vs. succinct) and patient narratives can also affect note quality. Clinicians are responsible for reviewing AI-generated notes, but time constraints and workflow pressures may lead to over-reliance on imperfect tools. This raises concerns about liability—whether errors fall on the clinician or the vendor—and highlights the need for governance processes and clinician education. If inaccuracies are carried forward into other notes for the patient this can create safety issues, delay care and even impact insurance, all issues highlighted by the *Patient Safety and Quality Healthcare* publication.

SECTION 3

Empowering Providers

3

TRANSPARENCY

Top Challenges and Recommended Solutions

01 Trust

Without transparency there is no trust, jeopardizing patient safety.

SOLUTIONS

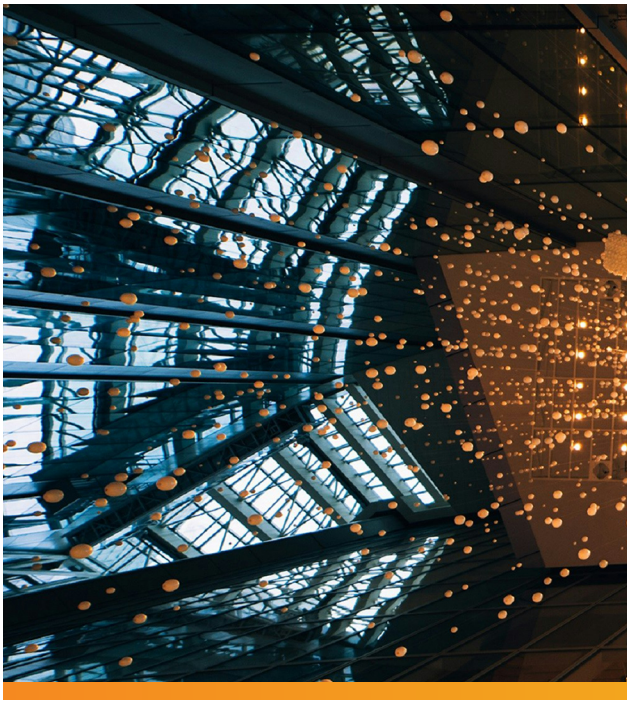
- AI model transparency, trust, and safety are intrinsically linked and essential for the successful integration of AI in healthcare.
- By making the inner workings of AI systems visible and understandable, stakeholders can ensure that these models are fair, valid, and effective, ultimately leading to safer and more reliable healthcare outcomes.
- Trust in AI models requires models to be trained on quality data, not just any data.
- Vendors should disclose key statistics about the datasets used to train their models—such as population size, demographics, and clinical domains—to enable HCOs to assess whether the model is appropriate for their patient populations.
- AI literacy training is needed for different HCO staff.

02 Training Data Processes

While an increasing number of HCOs are making investments in AI, there is still significant reluctance and uncertainty around AI models' performance.

SOLUTIONS

- The integration of AI in healthcare requires a concerted effort to ensure transparency and safety.
- By leveraging existing frameworks like the SAFER guides – while recognizing their limitations - and greater vendor oversight, healthcare providers can better assess and trust the AI technologies they deploy.



03 Model Cards

Without a standardized and transparent format for depicting AI models' intended use, performance, limitations, training data, ethical considerations, and potential biases, providers will struggle to evaluate these tools.

SOLUTIONS

- Use of model cards may have the potential to improve transparency around algorithm development and vendor accountability, however issues with measurement and validity must be addressed.
- A set of standards that can be applied to model cards is needed. The standards should offer a consistent way for providers to evaluate AI models.
- Ongoing provider and clinician input is needed to ensure the intended value of model cards is realized.
- Model cards should be living documents updated regularly to monitor ongoing performance, not just a snapshot in time when the technology is initially developed.
- The level of detail in the ASTP/ONC FAVEs model for predictive decision support interventions (DSIs) is similar to what providers may come to find helpful in model cards, which aim to provide a comprehensive overview of AI tools to ensure they are used appropriately and effectively.

Discussion

Engendering trust in AI healthcare models requires confidence in their accuracy, reliability, and ethical use.

America's AI Action Plan released in July by President Trump states:

> *Today, the bottleneck to harnessing AI's full potential is not necessarily the availability of models, tools, or applications. Rather, it is the limited and slow adoption of AI, particularly within large, established organizations. Many of America's most critical sectors, such as healthcare, are especially slow to adopt due to a variety of factors, including distrust or lack of understanding of the technology, a complex regulatory landscape, and a lack of clear governance and risk mitigation standards. A coordinated Federal effort would be beneficial in establishing a dynamic, "try-first" culture for AI across American industry.*

TRUST

Building trust involves transparent communication about AI model development, validation, and implementation, and addressing concerns related to data privacy, bias, and accountability.

AI model transparency, trust, and safety are intrinsically linked, forming a triad essential for the successful integration of AI in healthcare. Transparency in AI algorithms allows healthcare providers to understand how decisions are made, fostering trust among users and patients. This trust is crucial, as it ensures that healthcare professionals are confident in the AI's recommendations and are more likely to adopt and rely on these technologies. Finally, transparency is a key component of safety, as it enables the identification and mitigation of potential risks associated with AI models. As one member sees it, "Transparency needs to speak to mitigation of risk, like hallucinations...we still need to have healthy skepticism to not trust the AI blindly. This is not typical technology where you program it to perform specific outcomes. You're asking it to reason and think based on what it was trained on and sometimes guess when it doesn't know for sure. Training healthcare professionals to trust but verify is a safer approach."

TRAINING DATA PROCESSES

The need for transparency in AI algorithms within healthcare is becoming increasingly critical as AI adoption increases.

AI implementation spans various domains, including clinical decision support, administrative efficiency, revenue cycle automation, ambient documentation, medical imaging, patient evaluation, and diagnostics. However, challenges persist, especially with vendor-provided AI solutions. Often, these solutions are presented by vendors to purchasers at a high-level, making it difficult for providers to evaluate them and effectively integrate them into provider administrative and care settings. The complexity is further compounded by a combination of technologies such as automation, machine learning, and generative AI. As one member noted:

"Some current challenges with vendor ROI, when you send out those questionnaires you often get high level responses without the ability to vet how good the tech is; some are platform (not point) solutions and it's hard to fully vet how the different AI capabilities are built together; and many solutions are blends of Automation, ML/risk, GenAI, etc."

The uneven regulatory landscape described earlier contributes to providers' challenges as they navigate purchasing and deployment decisions and work to ensure what they are buying is safe and effective.

The safety of CDS such as AI imaging, particularly concerning the longevity of use and the potential for drift in practice patterns, points to a greater need for transparency. Gradual changes or variations in the way medical practices and procedures are carried out over time can affect the performance and accuracy of AI models. Since these models may have been trained on data that no longer reflects current practices, monitoring and updating

AI systems to ensure they remain effective and reliable is critical.

As imaging technology advances, the categorization of data and autonomous capabilities of AI systems raises significant concerns. Categorization of data involves using algorithms to analyze data and assign it to predefined categories based on patterns, features, or characteristics identified within the data. AI categorization can be used for various purposes, such as diagnosing medical conditions, sorting medical images, or organizing patient information with the intended goal of enhancing the efficiency and accuracy of data processing – ultimately improving decision-making and patient care. However, as this type of categorization expands beyond imaging to include broader clinical data within EHRs, it introduces new risks—particularly when such capabilities are embedded in CEHRT platforms or deployed by vendors that are not subject to the same regulatory oversight.

One potential solution could be to tie AI categorization to the SAFER guides – while recognizing their limitations in addressing AI-specific risks. By combining them with robust vendor oversight and AI-specific evaluation protocols, healthcare providers may be able to more effectively assess and trust the AI technologies they deploy. The SAFER guides provide a robust framework for risk assessment, specifically as it pertains to Test Results Reporting and Follow-up. This guide emphasizes the importance of accurate and timely reporting of test results, which is crucial for the effective functioning of AI algorithms in imaging. This approach could enhance the safety and reliability of AI in test reporting and other healthcare

applications by providing a structured framework for evaluating and ensuring the safety and reliability of AI systems.

Understanding the pace at which innovation in AI is occurring is pivotal to understanding how medicine must evaluate models, train clinicians, and safely deploy AI in clinical settings.

Gartner's 2025 Hype Cycle for Artificial Intelligence visually maps the evolving maturity of AI technologies in a dynamic landscape. The Cycle begins with an "Innovation Trigger" then rises precipitously to a "Peak of Inflated Expectations" before plummeting to the "Trout of Disillusionment" before gradually rising again to the "Slope of Enlightenment" followed by the "Plateau of Productivity." As Gartner explains, we are currently in the stage of "Peak of Inflated Expectations" for some types of AI (i.e., chat agents) and already in the "Trough of Disillusionment" for others (i.e., generative AI).

Gartner further concludes, "AI models are increasingly deployed to augment and replace human decision making... To build trust with users and stakeholders, application leaders must make these models more interpretable and explainable."

MODEL CARDS

There is momentum building to address some of the transparency issues through the use of model cards. Model cards summarize the details of a model in a structured format by describing its intended use, performance, limitations, training data, ethical considerations, and potential biases. Use of model cards can help healthcare providers make informed purchasing decisions. The cards can include the type of model, the

version, the developer's name, and sample outputs. It also may furnish information about the training data used to develop the model, the model's performance, and its limitations, among other things.

AS MORE HEALTHCARE PROVIDERS APPLY GOVERNANCE PROCESSES TO THE USE OF AI WITHIN THEIR ORGANIZATIONS, THE NEED TO BALANCE THE RISKS AND BENEFITS OF ITS USE WILL TAKE CENTER STAGE.

As IAPP, a non-profit whose mission is to "define, promote and improve the professions of privacy, AI governance and digital responsibility globally" describes it, model cards should include: model details; intended use; performance metrics; training data; quantitative analysis; and ethical considerations and recommendations.

Model cards for health AI are being developed by different entities including vendors such as the one found here and by entities like CHAI, found here. CHAI has also launched a registry for health AI model cards with the intention of making it easier for HCOs to evaluate and compare validated AI tools by standardizing how the information in model cards is presented. CHAI's model card framework includes "transparency for all Five of CHAI's Principles of Responsible AI (Transparency, Safety, Security & Privacy, Fairness & Bias, and Usefulness)."

Depending upon the autonomous capability of an AI model – that is the ability for the system to make a decision without human intervention – the use of model cards may have the potential to improve transparency around model development and vendor accountability. Some providers, however, have found them to be overly academic and impractical. Providers have also expressed concerns that without a way to consistently measure or validate one model against another or ensure patient privacy that the utility of model cards is limited. For instance, evaluating different LLMs for early detection of sepsis, a common hospital acquired infection, remains challenging without a uniform way to measure the different models. Providers need assurance that the models they are purchasing are adhering to the principles of responsible AI, which is rooted in trust and accountability.

The development of model cards in health AI can be analogized to the requirements set by ASTP/ONC's HTI-1 policies related to predictive analytics in several ways. ASTP/ONC has been working on refining certification requirements related to predictive analytics and AI transparency. Model cards, like the lists provided by EHR vendors, serve as a form of transparency and accountability.

ASTP/ONC states that their Predictive Decision Support Interventions (DSI) quality framework outlined in the HTI-1 final rule, is "necessary to help HCOs and users of these tools better determine whether their Predictive DSIs are fair, appropriate, valid, effective, and safe (FAVES)." ASTP/ONC further states that "transparency is a prerequisite for trustworthy AI." The inputs into a trustworthy and High Quality Predictive DSI include: 1) data transparency; 2) performance transparency; and 3) organizational transparency. Notably, the organizational piece requires Certified Health IT developers to apply intervention risk management for each predictive DSI they supply as part of their Health IT Module.

3

INNOVATION

Top Challenges and Recommended Solutions

01 Man vs. Machine

Several legacy issues, as well as emerging issues could slow AI adoption.

SOLUTIONS

- Assist providers navigate governance policies to monitor, manage and mitigate AI use.
- Design AI tools to enhance care while balancing efficiency and affordability.
- As vendors build and incorporate AI into their offerings, it needs to be built in with intentionality rather than bolted on as an afterthought. This should include controls and appropriate oversight to help providers safely and successfully adopt AI.
- Deploy AI to reduce clinician burden without adding complexity or over-reliance.
- Foster a culture of innovation by encouraging collaboration among clinicians, HCOs, and developers to develop collaborative solutions.
- Focus on improving data quality and infrastructure to enable safe, bias-free AI use.

- Protect data and ensure vendors meet safety, interoperability, and compliance standards.
- Align AI investments with measurable outcomes and fair pricing models.

02 Education & Workforce

The successful integration of AI into healthcare is challenged by a workforce that is still developing the skills needed to fully embrace and adapt to emerging technologies.

SOLUTIONS

- Invest in workforce training to build tailored education to build trust and effective use of AI.
- Reinforce that AI should augment—not replace—human decision-making and that clinical judgment remains central, even as AI capabilities grow.
- Foster partnerships among healthcare providers, educators, and AI developers to support successful integration.
- Prepare clinicians to navigate AI limitations, including hallucinations and over-reliance.
- Educate the educators so they are trained in AI tools to effectively teach others.
- Support ongoing learning to provide continuous education to keep pace with evolving AI technologies.

03 Small & Under-resourced Providers

Small, rural, and other under-resourced providers experience unique and profound barriers that make adoption of AI more challenging.

SOLUTIONS

- Policies and funding that support the integration of AI in healthcare, especially in rural and under-resourced areas, are needed.
- Providers, policymakers, and other stakeholders cannot afford to ignore the impact AI is having on the digital divide.
- Education is needed to assist small and under-resourced providers adopt AI.
- Satellite solutions that deliver broadband in areas where traditional carrier options are not available will help speed AI adoption.
- Use of creative solutions to address funding challenges including using free and low-cost tools can help under-resourced providers.
- Tax credits for those donating technology should be explored such as when a provider donates an instance of their EHR to a smaller provider.
- Group purchasing options to make AI tools more accessible to smaller healthcare providers may be helpful.
- Policies that promote technology investment should be leveraged like the Rural Health Transformation (RHT) Program.
- CMS Stark and OIG AKS policies could be relaxed to foster more technology donations.

Discussion

Innovation in AI is reshaping healthcare delivery, offering tools that promise to streamline workflows, enhance diagnostics, and support clinical decision-making.

MAN VS. MACHINE

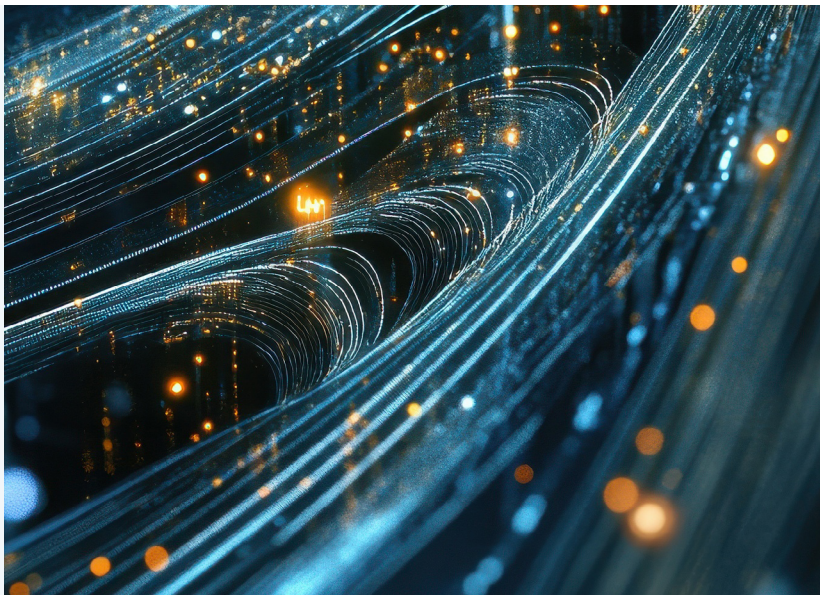
However, providers—especially smaller, rural, and under-resourced ones—face persistent challenges that complicate adoption, including funding cuts at the state and federal levels. Workforce shortages and clinician burnout continue to strain operations, while cost pressures limit the ability to invest in emerging technologies.

Many AI tools have yet to deliver consistent, scalable value, often requiring robust data infrastructure and governance that smaller systems lack. Additionally, the rapid pace of AI development has outpaced clinician education and training, creating gaps in understanding, trust, and effective use. To realize AI's full potential, innovation must be paired with strategic investment in workforce support, equitable access, and education tailored to diverse provider environments. Interest continues to grow around AI tools that address cost pressures and clinician burnout.

While HCOs are looking to AI to address these issues, there are still barriers to overcome. Studies find that investments in certain AI tools have not yet delivered a justifiable return on investment. And, while AI is seen as a strategic tool to address cost pressures, its success depends on infrastructure readiness, governance, and measurable outcomes.

AI uptake is also occurring at a higher rate among larger providers – while smaller, lesser resourced providers struggle to keep pace with long-standing issues like solvency and reimbursement. The findings of a study published in *JAMIA* in May 2025 which studied mid to larger provider adoption of AI concluded that systems with greater financial and technical resources are more likely to adopt and scale AI technologies. The rationale for AI adoption cited most frequently among respondents was related to clinician burden. All survey respondents indicated successful adoption of ambient AI but outside of this use case, AI adoption was uneven among providers. The authors concluded, "Other select AI use cases such as imaging/radiology and clinical risk stratification are adopted by many, though successes are modest."

HCOs are increasingly turning to chatbots for care management to enhance patient engagement and streamline administrative tasks which may offer cost-effective care modalities.



These AI-powered chatbots can assist patients with appointment scheduling, medication reminders, and answer frequently asked questions, thereby reducing the workload on healthcare staff. Additionally, chatbots can provide personalized health advice and monitor patient symptoms, ensuring timely interventions and improving overall care management. This innovative approach not only improves efficiency but also ensures that patients receive consistent and accessible support. One member who is piloting a chatbot solution found they were able to minimize depression, loneliness and feelings of isolation among their senior living community residents referring to its use as a "game changer." The positive experience was reported in the *Wall Street Journal* in June 2025.

Other challenges related to clinician adoption of AI tools involve clinician "over-reliance on AI technologies." An August 2024 article in the *British Medical Journal* found that, "As clinicians increasingly depend on AI for diagnostics and treatment recommendations, there is a tangible risk that their clinical skills may deteriorate." The article also found that the cognitive demands placed on clinicians could strain the provider-patient relationship and could also lead to clinician displacement issues. Research on clinician displacement is ongoing.

Providers are also grappling with their organization's data readiness. An organization's AI journey is often impacted by how clean and accurate their data is – to avoid bias and errors – with many providers still working through these issues. Providers are also calling for stronger vendor accountability and clearer standards to ensure AI tools are safe, interoperable, and aligned with clinical goals. There are opportunities for providers when partnering with large technology companies—but only if structured to protect provider interests and deliver shared value. Cybersecurity also remains a top priority—not just as a technical safeguard, but as a foundational requirement for trust and resilience. Even as many providers are addressing these issues, for others, the pricing models make AI outside their reach.

Tackling these challenges head-on is essential not only to ensure that AI investments translate into better care, smarter operations, and sustainable cost management—but also to unlock the full potential of innovation as a driver of meaningful transformation across the healthcare ecosystem. HCOs must implement robust data governance frameworks to ensure data accuracy, minimize duplication, and prepare data for AI applications.

EDUCATION & WORKFORCE

If the American healthcare system is to be a leader in AI, it is imperative that clinicians and other healthcare providers are educated and trained to use this technology. This means enhancing the training of healthcare providers and clinicians so that they can leverage AI systems effectively. Small and under-resourced providers are already facing significant challenges. This segment of providers will need additional support to help ensure they are not left behind in a widening "digital divide."

Investment in clinician education will help ensure that students and clinicians have the resources needed to engage with AI. Clinicians need targeted education to build critical thinking skills that help them recognize AI's limitations, including risks of overreliance and hallucinations, ensuring safe and informed use in patient care.

It will not be enough to simply introduce AI technologies; clinicians must be equipped to understand and navigate these tools effectively. They must be trained to understand that AI is iterative, requiring continuous re-evaluation and updates. This includes teaching educators how to use the technology themselves, as their ability to instruct others is contingent upon their own proficiency. Just because someone is a teacher does not mean they inherently know how to use or teach AI technologies.

To foster a culture of innovation and position the U.S. as a leader will require clinicians receive training in these next generation tools. An article in the *Postgraduate Medical Journal* from August 2025 found that AI education must be embedded within competency-based medical training to produce future-ready clinicians. A global study cited in the article found that while most medical students support AI use, they feel unprepared and undertrained, highlighting an imminent need for AI literacy. Further, according to experts positioning the American healthcare system as a leader in AI will require building an "AI-enabled medical school of the future."

Education for K-12 will also be needed to remain competitive. The *Economic Times* has reported that the People's Republic of China has already built AI into their K-12 curriculum and and as of September 31st eight hours of training is required. Preparing for the future will require teaching students critical thinking early on to reap the benefits in society later.

SMALL & UNDER-RESOURCED PROVIDERS

For many providers – particularly small and under-resourced providers – uptake of AI is slower as they navigate the costs, challenges and opportunities associated with adopting this new technology.

Innovative policy solutions will be needed to assist small, rural and under-resourced providers including those providers who were not eligible for EHR incentives under the Health Information Technology for Economic and Clinical Health (HITECH) Act, like long-term post-acute care (LTPAC) and behavioral health providers, requiring multifaceted and innovative approaches.

THE PACE OF AI INNOVATION RISKS EXACERBATING THE ALREADY SIGNIFICANT DIGITAL DIVIDE BETWEEN SMALL AND LARGER PROVIDERS AND POSES GREATER CHALLENGES TO ACCESS EMERGING TECHNOLOGY.

Many of these approaches must tackle basic technology needs before AI is contemplated. Rural providers often struggle with infrastructure, making it vital to first incentivize adoption of EHRs, cybersecurity, and interoperability. In some states, governors and legislators have allocated funds for this critical infrastructure, though often limited only to acute care providers and physician groups. Meanwhile, LTPAC and behavioral health providers face increased cost pressures as well. For instance, inflation is causing higher salaries even as Medicare and Medicaid reimbursement is not keeping up with these inflationary increases. In addition, some Medicare Advantage (MA) payers are reducing inpatient lengths of stay, which can often result in higher rehospitalization rates as patients are discharged before they are ready to return home.

Providers who are resource-strapped often do their own internal assessments to look for opportunities for greater efficiency such as centralizing functions, removing redundancy, and locating tools available

through existing software licenses. Leveraging "talents, tools, and treasure" – as one member calls them – helped them locate tools for templating cyber policies and data loss prevention.

Under-resourced providers can also take advantage of free tools like the STEPS Forward® "Governance for Augmented Intelligence," developed by the AMA in conjunction with Manatt Health, an eight-step toolkit to help providers establish a governance process and scale up on AI.

For those under-resourced providers attempting to adopt AI, some simply lack the purchasing power to afford it, including those solutions available through their EHR. One example of this involves a federally qualified health center (FQHC) in Arizona as reported by *Fierce Healthcare* in April 2025. Despite participating in a pilot run by HAIP to offer technical assistance, this FQHC still struggled with AI adoption. A big challenge for them and many other rural and small providers is access to a strong broadband infrastructure. Supporting these providers in overcoming connectivity barriers is crucial. Waiting for traditional carriers to deliver high-speed broadband to a rural area may set these providers further behind. Satellite options that can permeate these untouched rural and hard-to-reach areas are needed now.

Group purchasing organizations (GPOs) may also help with cost containment by leveraging their collective buying power. Additionally, there are programs that allow larger organizations to share their EHR instances with smaller providers, which can also help manage costs. Providing tax incentives to those donating technology may also be a way to help facilitate technology donations to lesser resourced providers. Universities, under scrutiny for infrastructure spending, can centralize functions to reduce waste. Further, policies included in the RHT Program as established under the One Big Beautiful Bill Act (OBBBA) permit and encourage technology investments. The CMS Stark and the OIG AKS policies could be amended to include permitting donations of AI tools and further ease the cybersecurity technology and related service donations provisions via sub-regulatory guidance.

Finally, it is worth noting that even when a small provider is able to adopt AI, they still face hurdles. The resource intensity required for auditing and assessing AI systems remains a major concern. It is not a simple "one and done" scenario. Following the purchase of an AI solution, it must be re-evaluated and monitored given its iterative nature. As one member noted, the thing keeping them up at night as her small, rural organization begins to adopt AI is, "the knowledge that we are responsible and knowing the resource intensity is not there to do all the audits that you can and should do." Several members have expressed concerns related to widening the gap between the "haves" and the "have nots."

ACKNOWLEDGEMENTS

CHIME's Public Policy team extends our deepest gratitude to all those who have contributed to this project including members of our Policy Steering Committee who serve as our north star and guide our advocacy efforts. Your unwavering dedication and invaluable insights have been pivotal in making this document a reality. We are profoundly grateful for your support and collaboration.

- Aaron Miri, DHA, FCHIME, CHCIO, EVP, Chief Digital & Information Officer, Baptist Health (PSC Co-Chair)
- Robert "Bob" Latz, PT, DPT, CHCIO, CDH-E, LCHIME, FCHIME, FHIMSS, CIO, Trinity Rehab Services (PSC Co-Chair)
- Angie Costakis, MS, CIO, Claiborne Memorial Medical Center
- Bobbie Byrne, M.D., MBA, EVP & CIO, Advocate Health
- Brian Sterud, MBA, CHCIO, FACHE, LFCHIME, VP of IT & CIO, Faith Regional Health Services
- Bridgett Ojeda, MAM, PMP, SVP & CIO, Bryan Health
- Chris Harper, CHCIOe, CDH-E, MBAi, MPM, SVP & CIO (TUKHS), Senior Associate Vice Chancellor Artificial Intelligence (KU), University of Kansas Medical
- Chris Plaisance, PhD, MBA, Business Intelligence Director & CIO, Black River Memorial Hospital
- Chuck Christian, CHCIO, CDH-E, LFCHIME, LFHIMSS, VP of Technology & CTO, Franciscan Health
- Corey M. Zeigler, LFCHIME, FACHE, CHCIO, CDH-E, CIO, Helio Health
- Donna Roach, MS, CHCIO, CDH-E, LFCHIME, LFHIMSS, CIO, University of Utah Health
- Fernando Small, PhD, CHCIO, AVP Academic Administration, Systems, and Analytics, MD Anderson Cancer Center
- J.D. Whitlock, MPH, MBA, CHCIO CIO, Dayton Children's Hospital
- Linda Stevenson, MBA, PMP, CIO, CHCIO, CDH-E, FCHIME, CIO, Fisher-Titus Medical Center
- Melissa Jost, MS, PMP, Director of Clinical Informatics and Clinician Health & Wellbeing, UC Davis Health
- Nathan Lesser, VP & CISO, Children's National Hospital
- Pallavi Ranade-Kharkar, PhD, MS, FAMIA, Director of Research Informatics and Genomics, Intermountain Healthcare
- Pam McNutt, SVP & CIO, Methodist Health System
- Rachini Moosavi, VP & Chief Analytics Officer, UNC Health
- Rob Maclay, CISSP, CISO, Stanford Children's Hospital
- Scott MacLean, CDH-E, CHCIO, LFCHIME, SVP & CIO, MedStar Health, CHIME Foundation Board Chair
- Shafiq Rab, M.D., MPH, FCHIME, Chief Digital Officer, System CIO & EVP, Tufts Medicine
- Sheree McFarland, MS, CHCIO, FCHIME, LCHIME, FACHE, CDH-E, CIO, West Florida, HCA
- Terri Couts, RN-BC, MHA, CHCIO, CDH-E, FCHIME, SVP, Chief Information & Digital Officer, Sharp Healthcare
- Terri Ripley, MIT, PMP, CHCIO, CIO, OrthoVirginia