American Medical Association Principles for Augmented Intelligence Development, Deployment, and Use

Approved by AMA Board of Trustees on November 14, 2023

As the number of Augmented Intelligence (AI)-enabled health care tools and systems continue to grow, these technologies must be designed, developed, and deployed in a manner that is ethical, equitable, responsible, and transparent. With a lagging effort towards adoption of national governance policies or oversight of AI, it is critical that the physician community engage in development of policies to help inform physician and patient education, and guide engagement with these new technologies. It is also important that the physician community help guide development of these tools in a way that best meets both physician and patient needs, and help define their own organization's risk tolerance, particularly where AI impacts direct patient care. The AMA is committed to ensuring that AI can meet its full potential to advance clinical care and improve clinician well-being. This may only be accomplished by ensuring that physicians engage only with AI that satisfies rigorous standards to meet the goals of the quadruple aim,[1] advance health equity, prioritize patient safety, and limit risks to both physicians and patients.

These new principles build on earlier AMA policy development activities, including the 2018 foundational AMA AI policy, Augmented Intelligence in Medicine,[2] followed by 2019 policy for payment and coverage of AI.[3] However, as AI has rapidly developed beyond AI-enabled medical devices, new policy and guidance for adoption of both device and non-device uses of AI-enabled technologies is necessary to assist in deployment of these new advances to physicians and patients. These principles will serve as the foundation for AMA's evolving advocacy on AI.

The AMA is dedicated to providing continued guidance to physicians on how to best engage with new AI-enabled technologies with the understanding that policy development related to AI will likely continue to develop given the rapid pace of change in this space.

Oversight of Health Care Augmented Intelligence

There is currently no national policy or governance structure in place to guide the development and adoption of non-device AI. While the Food and Drug Administration (FDA) regulates AI-enabled medical devices, many types of AI-enabled technologies fall outside the scope of FDA oversight, including AI that may have clinical applications, such as some clinical decision support functions. While the Federal Trade Commission and the Health and Human Services Office for Civil Rights have oversight over some aspects of AI, their authorities are limited and not adequate to ensure appropriate development and deployment of AI generally, and specifically in the health care space. The AMA encourages a whole of government approach to implement governance policies that ensure overall and disparate risks to consumers and patients arising from AI are mitigated to the greatest extent possible.

In addition to government, health care institutions, practices, and professional societies share some responsibility for appropriate oversight and governance of AI-enabled systems and technologies. Beyond government oversight or regulation, purchasers and users of these technologies should have appropriate and sufficient policies in place to

---

[1] AI systems should enhance the patient experience of care and outcomes, improve population health, reduce overall costs for the health care system while increasing value, and support the professional satisfaction of physicians and the health care team.

[2] American Medical Association. (2018). "AI in Healthcare: A Report from the American Medical Association Board of Trustees." AMA, https://www.ama-assn.org/system/files/2019-08/ai-2018-board-report.pdf (Accessed September 14, 2023).

[3] American Medical Association. 2019). "AI in Healthcare: A Report from the American Medical Association Board of Trustees - 2019." https://www.ama-assn.org/system/files/2019-08/ai-2019-board-report.pdf (Accessed September 14, 2023).

ensure they are acting in accordance with the current standard of care. Similarly, clinical experts are best positioned to determine whether AI applications are high quality, appropriate, and whether the AI tools are valid from a clinical perspective. Clinical experts can best validate the clinical knowledge, clinical pathways, and standards of care used in the design of AI-enabled tools and can monitor the technology for clinical validity as it evolves over time.

- Health care AI must be designed, developed, and deployed in a manner which is ethical, equitable, responsible, and transparent.
- Use of AI in health care delivery requires clear national governance policies to regulate its adoption and utilization, ensuring patient safety, and mitigating inequities. Development of national governance policies should include interdepartmental and interagency collaboration.
- Compliance with national governance policies is necessary to develop AI in an ethical and responsible manner to ensure patient safety, quality, and continued access to care. Voluntary agreements or voluntary compliance is not sufficient.
- Health care AI requires a risk-based approach where the level of scrutiny, validation, and oversight should be proportionate to the potential overall or disparate harm and consequences the AI system might introduce. [See also Augmented Intelligence in Health Care H-480.939 at (1)]
- Clinical decisions influenced by AI must be made with specified human intervention points during the decision-making process. As the potential for patient harm increases, the point in time when a physician should utilize their clinical judgment to interpret or act on an AI recommendation should occur earlier in the care plan.
- Health care practices and institutions should not utilize AI systems or technologies that introduce overall or disparate risk that is beyond their capabilities to mitigate. Implementation and utilization of AI should avoid exacerbating clinician burden and should be designed and deployed in harmony with the clinical workflow.
- Medical specialty societies, clinical experts, and informaticists are best positioned and should identify the most appropriate uses of AI-enabled technologies relevant to their clinical expertise and set the standard of care for AI usage in their specific domain. [See Augmented Intelligence in Health Care H-480.940 at (2)]

When to Disclose: Transparency in Use of Augmented Intelligence-Enabled Systems and Technologies

As implementation of AI-enabled tools and systems continues to increase, it is essential that use of AI in health care be transparent to both physicians and patients. Transparency requirements should be tailored in a way that best suits the needs of the end users. Disclosure should contribute to physician and patient knowledge and not create unnecessary administrative burden. When AI is utilized in health care decision-making, that use should be disclosed and documented in order to limit risks to, and mitigate inequities for, both physicians and patients, and to allow each to understand how decisions impacting patient care or access to care are made. While transparency does not necessarily ensure AI-enabled tools are accurate, secure, or fair, it is difficult to establish trust if certain characteristics are hidden.

- When AI is used in a manner which directly impacts patient care, access to care, or medical decision making, that use of AI should be disclosed and documented to both physicians and/or patients in a culturally and linguistically appropriate manner. The opportunity for a patient or their caregiver to request additional review from a licensed clinician should be made available upon request.
- When AI is used in a manner which directly impacts patient care, access to care, medical decision making, or the medical record, that use of AI should be documented in the medical record.
- AI tools or systems cannot augment, create, or otherwise generate records, communications, or other content on behalf of a physician without that physician's consent and final review.
- When health care content is generated by generative AI, including by large language models, it should be clearly disclosed within the content that was generated by an AI-enabled technology.
- When AI or other algorithmic-based systems or programs are utilized in ways that impact patient access to care, such as by payors to make claims determinations or set coverage limitations, use of those systems or programs must be disclosed to impacted parties.

- The use of AI-enabled technologies by hospitals, health systems, physician practices, or other entities, where patients engage directly with AI should be clearly disclosed to patients at the beginning of the encounter or interaction with the AI-enabled technology.

What to Disclose: Required Disclosures by Health Care Augmented Intelligence-Enabled Systems and Technologies

Along with significant opportunity to improve patient care, all new technologies in health care will likely present certain risks and have limitations that physicians must carefully navigate during the early stages of clinical implementation of these new systems and tools. AI-enabled tools are no different and are perhaps more challenging than other advances as they present novel and complex questions and risks. To best mitigate these risks, it is critical that physicians understand AI-driven technologies and have access to certain information about the AI tool or system being considered, including how it was trained and validated, so that they can assess the quality, performance, equity, and utility of the tool to the best of their ability. This information may also establish a set of baseline metrics for comparing AI tools. Transparency and explainability regarding the design, development, and deployment processes should be mandated by law where possible, including potential sources of inequity in problem formulation, inputs, and implementation. Additionally, sufficient detail should be disclosed to allow physicians to determine whether a given AI-enabled tool would reasonably apply to the individual patient they are treating. Physicians should understand that, where they utilize AI-enabled tools and systems without transparency provided by the AI developer, their risks of liability for reliance on that AI will likely increase. The need for full transparency is greatest where AI-enabled systems have greater impacts on direct patient care, such as by AI-enabled medical devices, clinical decision support, and interaction with AI-driven chatbots. Transparency needs may be somewhat lower where AI is utilized for primarily administrative, practice-management functions.

- When AI-enabled systems and technologies are utilized in health care, the following information should be disclosed by the AI developer to allow the purchaser and/or user (physician) to appropriately evaluate the system or technology prior to purchase or utilization:
    - Regulatory approval status
    - Applicable consensus standards and clinical guidelines utilized in design, development, deployment, and continued use of the technology
    - Clear description of problem formulation and intended use accompanied by clear and detailed instructions for use
    - Intended population and intended practice setting
    - Clear description of any limitations or risks for use, including possible disparate impact
    - Description of how impacted populations were engaged during the AI lifecycle
    - Detailed information regarding data used to train the model:
        - Data provenance
        - Data size and completeness
        - Data timeframes
        - Data diversity
        - Data labeling accuracy
    - Validation Data/Information and evidence of:
        - Clinical expert validation in intended population and practice setting and intended clinical outcomes
        - Constraint to evidence-based outcomes and mitigation of “hallucination” or other output error
        - Algorithmic validation
        - External validation processes for ongoing evaluation of the model performance, e.g., accounting for AI model drift and degradation
        - Comprehensiveness of data and steps taken to mitigate biased outcomes
        - Other relevant performance characteristics, including but not limited to performance characteristics at peer institutions/similar practice settings
        - Post-market surveillance activities aimed at ensuring continued safety, performance, and equity

    - Data Use Policy
        - Privacy
        - Security
        - Special considerations for protected populations or groups put at increased risk
    - Information regarding maintenance of the algorithm, including any use of active patient data for ongoing training
    - Disclosures regarding the composition of design and development team, including diversity and conflicts of interest, and points of physician involvement and review
- Physicians should carefully consider whether or not to engage with AI-enabled health care technologies if this information is not disclosed by the developer. As the risk of AI being incorrect increases risks to patients (such as with clinical applications of AI that impact medical decision making), disclosure of this information becomes increasingly important. [See also Augmented Intelligence in Health Care H-480.939]

Generative Augmented Intelligence

Generative AI is a type of AI that can recognize, summarize, translate, predict, and generate text and other content based on knowledge gained from large datasets. Generative AI tools are finding an increasing number of uses in health care, including assistance with administrative functions, such as generating office notes, responding to documentation requests, and generating patient messages. Additionally, there has been increasing discussion about clinical applications of generative AI, including use as clinical decision support to provide differential diagnoses, early detection, and intervention, and to assist in treatment planning. While generative AI tools show tremendous promise to make a significant contribution to health care, there are a number of potential risks and limitations to consider when using these tools in a clinical setting or direct patient care. To manage risk, health care organizations should develop and adopt appropriate polices that anticipate and minimize negative impacts. Physicians who are considering utilizing a generative AI-based tool in their practice should ensure that all practice staff are educated on the risks and limitations, including patient privacy concerns, and additionally, should have appropriate governance policies in place for its use prior to adoption.

- Generative AI should: (a) only be used where appropriate policies are in place within the practice or other health care organization to govern its use and help mitigate associated risks; and (b) follow applicable state and federal laws and regulations (e.g., HIPAA-compliant Business Associate Agreement).
- Appropriate governance policies should be developed by health care organizations and account for and mitigate risks of:
    - Incorrect or falsified responses; lack of ability to readily verify the accuracy of responses or the sources used to generate the response
    - Training data set limitations that could result in responses that are out of date or otherwise incomplete or inaccurate for all patients or specific populations
    - Lack of regulatory or clinical oversight to ensure performance of the tool
    - Bias, discrimination, promotion of stereotypes, and disparate impacts on access or outcomes
    - Data privacy
    - Cybersecurity
    - Physician liability associated with the use of generative AI tools
- Health care organizations should work with their AI and other health information technology (health IT) system developers to implement rigorous data validation and verification protocols to ensure that only accurate, comprehensive, and bias managed datasets inform generative AI models, thereby safeguarding equitable patient care and medical outcomes. [See Augmented Intelligence in Health Care H-480.940 at (3)(d)]
- Use of generative AI should incorporate physician and staff education about the appropriate use, risks, and benefits of engaging with generative AI. Additionally, physicians should engage with generative AI tools only when adequate information regarding the product is provided to physicians and other users by the developers of those tools.

- Clinicians should be aware of the risks of patients engaging with generative AI products that produce inaccurate or harmful medical information (e.g., patients asking chatbots about symptoms) and should be prepared to counsel patients on the limitations of AI-driven medical advice.
- Governance policies should prohibit the use of confidential, regulated, or proprietary information as prompts for generative AI to generate content.
- Data and prompts contributed by users should primarily be used by developers to improve the user experience and AI tool quality and not simply increase the AI tool's market value or revenue generating potential.

Physician Liability for Use of Augmented Intelligence-Enabled Technologies

The question of physician liability for use of AI-enabled technologies presents novel and complex legal questions and potentially poses risks to the successful clinical integration of AI-enabled technologies. As legal theories of liability and accountability for AI continue to evolve, the AMA will continue to advocate to ensure that physician liability for the use of AI-enabled technologies is limited and adheres to current legal approaches to medical malpractice.

- Current AMA policy states that liability and incentives should be aligned so that the individual(s) or entity(ies) best positioned to know the AI system risks and best positioned to avert or mitigate harm do so through design, development, validation, and implementation. [See Augmented Intelligence in Health Care H-480.939]
    - Where a mandated use of AI systems prevents mitigation of risk and harm, the individual or entity issuing the mandate must be assigned all applicable liability.
    - Developers of autonomous AI systems with clinical applications (screening, diagnosis, treatment) are in the best position to manage issues of liability arising directly from system failure or misdiagnosis and must accept this liability with measures such as maintaining appropriate medical liability insurance and in their agreements with users.
    - Health care AI systems that are subject to non-disclosure agreements concerning flaws, malfunctions, or patient harm (referred to as gag clauses) must not be covered or paid and the party initiating or enforcing the gag clause assumes liability for any harm.
- When physicians do not know or have reason to know that there are concerns about the quality and safety of an AI-enabled technology, they should not be held liable for the performance of the technology in question.

Data Privacy and Augmented Intelligence

Data privacy is highly relevant to AI development, implementation, and use. The AMA is deeply invested in ensuring individual patient rights and protections from discrimination remain intact, that these assurances are guaranteed, and that the responsibility falls with the data holders. AI development, training, and use requires assembling large collections of health data. AI machine learning is data hungry; it requires massive amounts of data to function properly. Increasingly, more electronic health records are interoperable across the health care system and, therefore, are accessible by AI trained or deployed in medical settings. AI developers may create legal arrangements, e.g., business associate agreements, that bring them under the Health Insurance Portability and Accountability Act (HIPAA) Privacy and Security Rules. Yet even HIPAA cannot protect patients from the "black box" nature of AI which makes the use of data opaque. AI system outputs may also include inferences that reveal personal data or previously confidential details about individuals. This can result in a lack of accountability and trust and exacerbate data privacy concerns. Often, AI developers and implementers are themselves unaware of exactly how their products use information to make recommendations.

It is unlikely that physicians or patients will have any clear insight into a generative AI tool's conformance to state or federal data privacy laws. Large language models (LLM) are trained on data scraped from the web and other digital

sources (including HIPAA-covered environments),[4] yet few, if any, controls are available to help users protect the data they voluntarily enter in a chatbot query. For instance, there are often no mechanisms in place for users to request data deletion or ensure that their inputs are not stored or used for future model training. While tools designed for medical use should align with HIPAA, many "HIPAA-compliant" generative tools rely on antiquated notions of deidentification, i.e., stripping data of personal information. With today's advances in computing power, data can easily be reidentified. Rather than aiming to make LLMs compliant with HIPAA, all health care AI-powered generative tools should be designed from the ground up with data privacy in mind.

The AMA's Privacy Principles were designed to provide individuals with rights and protections and shift the responsibility for privacy to third-party data holders. While the Principles are broadly applicable to all AI developers, e.g., entities should only collect the minimum amount of information needed for a particular purpose, the unique nature of LLMs and generative AI warrant special emphasis on entity responsibility and user education.

Entity Responsibility:

- Entities should make information available about the intended use of generative AI in health care and identify the purpose of its use. Individuals should know how their data will be used or reused, and the potential risks and benefits.
- Individuals should have the right to opt-out, update, or forget use of their data in generative AI tools. These rights should encompass AI training data and disclosure to other users of the tool.
- Generative AI tools should not reverse engineer, reconstruct, or reidentify an individual's originally identifiable data or use identifiable data for nonpermitted uses, e.g., when data are permitted to conduct quality and safety evaluations. Preventive measures should include both legal frameworks and data model protections, e.g., secure enclaves, federated learning, and differential privacy.

User Education:

- Users should be provided with training specifically on generative AI. Education should address:
    - legal, ethical, and equity considerations,
    - risks such as data breaches and re-identification,
    - potential pitfalls of inputting sensitive and personal data, and
    - the importance of transparency with patients regarding the use of generative AI and their data.

[See Augmented Intelligence in Health Care H-480.940 at (4),(5)]

Augmented Intelligence Cybersecurity

Data privacy relies on strong data security measures. There is growing concern that cyber criminals will use AI to attack health care organizations. AI poses new threats to health IT operations. AI-operated ransomware and AI-operated malware can be targeted to infiltrate health IT systems and automatically exploit vulnerabilities. Attackers using ChatGPT can craft convincing or authentic emails and use phishing techniques that entice people to click on links—giving them access to the entire electronic health record system.

AI is particularly sensitive to the quality of data. Data poisoning is the introduction of "bad" data into an AI training set, affecting the model's output. AI requires large sets of data to build logic and patterns used in clinical decision-making. Protecting this source data is critical. Threat actors could also introduce input data that compromises the

[4] Feathers, T., et al. "Facebook is receiving sensitive medical information from hospital websites. The Markup. June 16, 2022." https://themarkup.org/pixel-hunt/2022/06/16/facebook-is-receiving-sensitive-medical-information-from-hospital-websites.

overall function of the AI tool. Failure to secure and validate these inputs, and corresponding data, can contaminate AI models—resulting in patient harm.

Because stringent privacy protections and higher data quality standards might slow model development, there could be a tendency to forgo essential data privacy and security precautions. However, strengthening AI systems against cybersecurity threats is crucial to their reliability, resiliency, and safety.

AI cybersecurity considerations:

- AI systems must have strong protections against input manipulation and malicious attacks.
- Entities developing or deploying health care AI should regularly monitor for anomalies or performance deviations, comparing AI outputs against known and normal behavior.
- Independent of an entity's legal responsibility to notify a health care provider or organization of a data breach, that entity should also act diligently in identifying and notifying the individuals themselves of breaches that impact their personal information.
- Users should be provided education on AI cybersecurity fundamentals, including specific cybersecurity risks that AI systems can face, evolving tactics of AI cyber attackers, and the user's role in mitigating threats and reporting suspicious AI behavior or outputs.

Payor Use of Augmented Intelligence and Automated Decision-Making Systems

Payors and health plans are increasingly using AI and algorithm-based decision-making in an automated fashion to determine coverage limits, make claim determinations, and engage in benefit design. Payors should leverage automated decision-making systems that improve or enhance efficiencies in coverage and payment automation, facilitate administrative simplification, and reduce workflow burdens. While the use of these systems can create efficiencies such as speeding up prior authorization and cutting down on paperwork, there is concern these systems are not be designed or supervised effectively—creating access barriers for patients and limiting essential benefits.

Increasingly, evidence shows that payors are using automated decision-making systems to deny care more rapidly, often with little or no human review. This manifests in the form of increased denials, stricter coverage limitations, and constrained benefit offerings. For example, a payor allowed an automated system to cut off insurance payments for Medicare Advantage patients struggling to recover from severe diseases, forcing them to forgo care or pay out of pocket. In some instances, payors instantly reject claims on medical grounds without opening or reviewing the patient's medical record. There is also a lack of transparency in the development of automated decision-making systems. Rather than payors making determinations based on individualized patient care needs, reports show that decisions are based on algorithms developed using average or "similar patients" pulled from a database. Models that rely on generalized, historical data can also perpetuate biases leading to discriminatory practices or less inclusive coverage.[5,6,7,8]

We must ensure that automated decision-making systems do not reduce needed care, nor systematically withhold care from specific groups. Steps should be taken to ensure that these systems are not overriding clinical judgement. Patients and physicians should be informed and empowered to question a payor's automated decision-making.

---

[5] Obermeyer, Ziad, et al. "Dissecting racial bias in an algorithm used to manage the health of populations." Science 366.6464 (2019): 447-453. https://www.science.org/doi/10.1126/science.aax2342.
[6] Ross, C., Herman, B. (2023) "Medicare Advantage Plans' Use of Artificial Intelligence Leads to More Denials." https://www.statnews.com/2023/03/13/medicare-advantage-plans-denial-artificial-intelligence/ (Accessed September 14, 2023).
[7] Rucker, P., Miller, M., Armstrong, D.(2023). "Cigna and Its Algorithm Deny Some Claims for Genetic Testing, ProPublica Finds." https://www.propublica.org/article/cigna-pxdx-medical-health-insurance-rejection-claims (Accessed September 14, 2023).
[8] Ross, C., Herman, B. (2023). "Medicare Advantage Algorithms Lead to Coverage Denials, With Big Implications for Patients." https://www.statnews.com/2023/07/11/medicare-advantage-algorithm-navihealth-unitedhealth-insurance-coverage/ (Accessed September 14, 2023).

There should be stronger regulatory oversight, transparency, and audits when payors use these systems for coverage, claim determinations, and benefit design. [See Use of Augmented Intelligence for Prior Authorization D-480.956; Prior Authorization and Utilization Management Reform H-320.939]

- Use of automated decision-making systems that determine coverage limits, make claim determinations, and engage in benefit design should be publicly reported, based on easily accessible evidence-based clinical guidelines (as opposed to proprietary payor criteria), and disclosed to both patients and their physician in a way that is easy to understand.
- Payors should only use automated decision-making systems to improve or enhance efficiencies in coverage and payment automation, facilitate administrative simplification, and reduce workflow burdens. Automated decision-making systems should never create or exacerbate overall or disparate access barriers to needed benefits by increasing denials, coverage limitations, or limiting benefit offerings. Use of automated decision-making systems should not replace the individualized assessment of a patient's specific medical and social circumstances and payors' use of such systems should allow for flexibility to override automated decisions. Payors should always make determinations based on particular patient care needs and not base decisions on algorithms developed on "similar" or "like" patients.
- Payors using automated decision-making systems should disclose information about any algorithm training and reference data, including where data were sourced and attributes about individuals contained within the training data set (e.g., age, race, gender). Payors should provide clear evidence that their systems do not discriminate, increase inequities, and that protections are in place to mitigate bias.
- Payors using automated decision-making systems should identify and cite peer-reviewed studies assessing the system's accuracy measured against the outcomes of patients and the validity of the system's predictions.
- Any automated decision-making system recommendation that indicates limitations or denials of care, at both the initial review and appeal levels, should be automatically referred for review to a physician (a) possessing a current and valid non-restricted license to practice medicine in the state in which the proposed services would be provided if authorized and (b) be of the same specialty as the physician who typically manages the medical condition or disease or provides the health care service involved in the request prior to issuance of any final determination. Prior to issuing an adverse determination, the treating physician must have the opportunity to discuss the medical necessity of the care directly with the physician who will be responsible for determining if the care is authorized.
- Individuals impacted by a payor's automated decision-making system, including patients and their physicians, must have access to all relevant information (including the coverage criteria, results that led to the coverage determination, and clinical guidelines used).
- Payors using automated decision-making systems should be required to engage in regular system audits to ensure use of the system is not increasing overall or disparate claims denials or coverage limitations, or otherwise decreasing access to care. Payors using automated decision-making systems should make statistics regarding systems' approval, denial, and appeal rates available on their website (or another publicly available website) in a readily accessible format with patient population demographics to report and contextualize equity implications of automated decisions. Insurance regulators should consider requiring reporting of payor use of automated decision-making systems so that they can be monitored for negative and disparate impacts on access to care. Payor use of automated decision-making systems must conform to all relevant state and federal laws.