JAMA | Special Communication | AI IN MEDICINE

AI, Health, and Health Care Today and Tomorrow
The JAMA Summit Report on Artificial Intelligence

Derek C. Angus, MD, MPH; Rohan Khera, MD, MS; Tracy Lieu, MD, MPH; Vincent Liu, MD, MSc; Faraz S. Ahmad, MD, MS; Brian Anderson, MD; Sivasubramanium V. Bhavani, MD, MS; Andrew Bindman, MD; Troyen Brennan, MD, MPH; Leo Anthony Celi, MD, MPH, MSc; Frederick Chen, MD, MPH; I. Glenn Cohen, JD; Alastair Denniston, MA, PhD; Sanjay Desai, MD; Peter Embí, MD, MS; Aldo Faisal, PhD; Kadija Ferryman, PhD; Jackie Gerhart, MD; Marielle Gross, MD, MBE; Tina Hernandez-Boussard, PhD, MS, MPH; Michael Howell, MD, MPH; Kevin Johnson, MD, MS; Kristine Lee, MD; Xiaoxuan Liu, MBChB, PhD; Kimberly Lomis, MD; Alex John London, PhD; Christopher A. Longhurst, MD, MS; Ken Mandl, MD, MPH; Elizabeth McGlynn, PhD; Michelle M. Mello, MPhil, PhD, JD; Fatima Munoz, MD, MPH; Lucila Ohno-Machado, MD, PhD, MBA; David Ouyang, MD; Roy Perlis, MD, MSc; Adam Phillips, MD; David Rhew, MD; Joseph S. Ross, MD, MHS; Suchi Saria, PhD; Lee Schwamm, MD; Christopher W. Seymour, MD, PhD; Nigam H. Shah, MBBS, PhD; Rashmee Shah, MD, MS; Karandeep Singh, MD, MMSc; Matthew Solomon, MD, PhD; Kathryn Spates, JD, ACNP-BC; Kayte Spector-Bagdady, JD, MBE; Tommy Wang, PhD; Judy Wawira Gichoya, MD, MS; James Weinstein, MS, DO; Jenna Wiens, PhD; Kirsten Bibbins-Domingo, PhD, MD, MAS; for the JAMA Summit on AI

Editor's Note
Multimedia
Supplemental content
CME at jamacmelookup.com

IMPORTANCE Artificial intelligence (AI) is changing health and health care on an unprecedented scale. Though the potential benefits are massive, so are the risks. The JAMA Summit on AI discussed how health and health care AI should be developed, evaluated, regulated, disseminated, and monitored.

OBSERVATIONS Health and health care AI is wide-ranging, including clinical tools (eg, sepsis alerts or diabetic retinopathy screening software), technologies used by individuals with health concerns (eg, mobile health apps), tools used by health care systems to improve business operations (eg, revenue cycle management or scheduling), and hybrid tools supporting both business operations (eg, documentation and billing) and clinical activities (eg, suggesting diagnoses or treatment plans). Many AI tools are already widely adopted, especially for medical imaging, mobile health, health care business operations, and hybrid functions like scribing outpatient visits. All these tools can have important health effects (good or bad), but these effects are often not quantified because evaluations are extremely challenging or not required, in part because many are outside the US Food and Drug Administration's regulatory oversight. A major challenge in evaluation is that a tool's effects are highly dependent on the human-computer interface, user training, and setting in which the tool is used. Numerous efforts lay out standards for the responsible use of AI, but most focus on monitoring for safety (eg, detection of model hallucinations) or institutional compliance with various process measures, and do not address effectiveness (ie, demonstration of improved outcomes). Ensuring AI is deployed equitably and in a manner that improves health outcomes or, if improving efficiency of health care delivery, does so safely, requires progress in 4 areas. First, multistakeholder engagement throughout the total product life cycle is needed. This effort would include greater partnership of end users with developers in initial tool creation and greater partnership of developers, regulators, and health care systems in the evaluation of tools as they are deployed. Second, measurement tools for evaluation and monitoring should be developed and disseminated. Beyond proposed monitoring and certification initiatives, this will require new methods and expertise to allow health care systems to conduct or participate in rapid, efficient, and robust evaluations of effectiveness. The third priority is creation of a nationally representative data infrastructure and learning environment to support the generation of generalizable knowledge about health effects of AI tools across different settings. Fourth, an incentive structure should be promoted, using market forces and policy levers, to drive these changes.

CONCLUSIONS AND RELEVANCE AI will disrupt every part of health and health care delivery in the coming years. Given the many long-standing problems in health care, this disruption represents an incredible opportunity. However, the odds that this disruption will improve health for all will depend heavily on the creation of an ecosystem capable of rapid, efficient, robust, and generalizable knowledge about the consequences of these tools on health.

JAMA. doi:10.1001/jama.2025.18490
Published online October 13, 2025.

Author Affiliations: Author affiliations are listed at the end of this article.

Group Information: The JAMA Summit on AI participants appear in the Supplement.

Corresponding Author: Derek C. Angus, MD, MPH, Critical Care Medicine, UPMC, 3550 Terrace St, 614 Scaife Hall, Pittsburgh, PA 15261 (angusdc@upmc.edu).

The scope, scale, and speed with which artificial intelligence (AI) will transform health and health care are staggering.[1-4] AI is changing how and when individuals seek care and how clinicians interact with patients, establish diagnoses, and implement and monitor treatments. Indeed, there is considerable enthusiasm that AI, especially given recent advances (**Box**), could address long-standing challenges in the access, cost, and quality of health care delivery.[1,4] Yet, the optimal path for AI development and dissemination remains unclear. In contrast to drugs or more traditional medical devices, there is little consensus or structure to ensure robust, safe, transparent, and standardized evaluation, regulation, implementation, and monitoring of new AI tools and technologies.[3,5-7] Some challenges are long-standing for digital health information technology as a whole, albeit more prescient with the rise of AI, while others are specific to AI.

To ensure that innovation in AI is both encouraged and appropriately incorporated into health care delivery, alignment on how to address these challenges among AI developers, health care systems and professionals, payers, regulators, and patients is required.[4,5,7] A JAMA Summit convened multiple stakeholders to review the current state of AI in health and health care, focusing on how best to evaluate, regulate, and monitor AI tools and technologies and the implications for health care infrastructure and workforce. AI can influence health and health care in many ways (Box), but we limited discussion to tools and technologies used by clinicians, patients, and individuals with health or wellness concerns, and health care systems (**Table 1**). Although *AI* is a broad term that can include older technologies similar to those of traditional statistical and computer-based decision support applications, discussion was limited to more recent advances such as machine learning models using ensemble methods, deep learning, generative AI, and agentic AI (Box).[8] We used the term *AI tool* to represent any tool, technology, device, or application containing such AI.

Box. The Broad Nature and Use of Artificial Intelligence (AI)

AI is a broad term; for decades, AI consisted of rule-based representations of knowledge (software encoding logic statements like "if X, then Y") and prediction models (artificial neural networks), offering output similar to that of traditional computer software and statistical models. However, with advances in computing power and the availability of larger, more complex datasets over the last 2 decades, AI evolved rapidly. The following 3 advances help differentiate AI from prior digital technologies:

- **Deep learning:** development of deeper, more convoluted neural networks capable of interpreting large complex datasets to address specific yet complicated tasks (eg, computer vision).
- **Generative AI:** an extension of deep learning using so-called large language and foundation models capable of generating new content to address far broader task requests (eg, ChatGPT or Gemini).
- **Agentic AI:** an extension of deep learning and generative AI capable of autonomous decision-making (eg, the Tesla autopilot software for autonomous driving).

This article focuses on newer AI tools used by clinicians, health care systems, and individuals with health concerns. Newer AI technology used broadly across society can also influence health. Examples include:

- AI tools deployed in the life sciences (eg, tools to enhance drug discovery, improve the conduct of randomized trials, or interrogate health care data) could improve biomedical discovery.
- AI tools deployed to address social factors (eg, tools to improve housing affordability or food security) could alter downstream health.
- AI tools used broadly in social communication and interaction (eg, algorithms deployed in social media platforms) could affect mental health, spread information (or disinformation) about health and health care, and influence aspects of care-seeking behavior, such as attitudes about vaccination.
- Use of general efficiency hacks by biomedical scientists and health care professionals, such as using ChatGPT for biomedical report writing or Gemini for internet searches, could impact how medical information is summarized, distributed, and used.

AI Tools in Health and Health Care

Clinical Tools

Most AI tools in the medical literature fall under the category of clinical tools (ie, tools directly supporting the clinical activities of health care professionals). Examples include AI software for automated screening of diabetic retinopathy, AI software embedded in a portable echocardiography device to provide automated diagnosis, or a machine learning algorithm that scans the electronic health record (EHR) to provide sepsis alerts and treatment recommendations.[9-19] Many of these tools require US Food and Drug Administration (FDA) clearance as medical devices, and more than 1200 have been cleared, the majority of which are in medical imaging.[7,20] AI has transformed medical imaging, augmenting image interpretation and dramatically changing how radiologists and pathologists work, with adoption of some type of AI by 90% of US health care systems.[1,21-23] AI-based clinical decision support tools embedded in the EHR are also widely available, in part because some have been outside the FDA's oversight and because the cost to a health care system to have the tool turned on may be perceived as modest.[23] Despite good access to these native EHR AI tools, concern among clinicians and health care systems persists regarding their accuracy, value, and utility.[23,24]

Outside medical imaging and EHR-embedded applications, dissemination of clinical AI tools has been slower. The primary barrier is likely that the health care systems responsible for buying these tools do not believe they offer enough value. The costs of not just licensing a tool but also of ensuring adequate training, digital infrastructure, maintenance, and monitoring can be considerable.[23,25,26] There is no reimbursement for most AI tool use and, even in select cases when insurers provide reimbursement, it does not offset the full costs.[27] There can also be skepticism regarding whether the purported benefits will be realized in practice. Concerns about algorithmic bias, automation bias (favoring suggestions from automated systems while overlooking contradictory information), lack of generalizability, and insufficient trust and endorsement from clinicians and patients may further dampen health care systems' enthusiasm.[23,24,28-32]

Direct-to-Consumer (DTC) Tools

There is considerable enthusiasm among the public for DTC AI tools (ie, tools used by patients or individuals with health or wellness

Table 1. Categories of Artificial Intelligence (AI) Tools that Directly Influence Health and Health Care Delivery

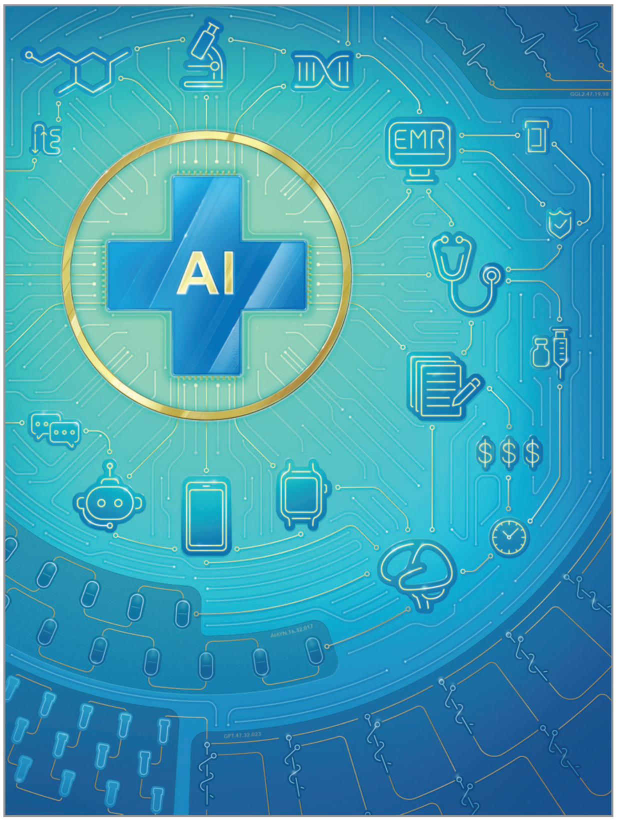
	Clinical	Direct to consumer	Business operations	Hybrid
Description	AI tools used by a clinician to support diagnostic or treatment decisions	AI tools used by a person with health or wellness concerns, without necessarily engaging with any health care system or professional	AI tools used by a health care system or professional to optimize aspects of health care delivery	AI tools that serve multiple purposes (eg, clinical and business operations)
Opportunity	Improve patient outcomes via: Better access to care, diagnostic accuracy, and compliance with best practice guidelines More personalized care	Improve patient and population health via: Improved disease surveillance and management Helping individuals adopt healthy lifestyles and prevention strategies More prompt, more personalized, and/or less costly access to care	Improve health care delivery via: Automating labor-intensive processes Reduced administrative burden Reduced waste Improved revenue generation	Improve health care delivery via: Reduced administrative burden on health care professionals Helping health care professionals execute care tasks Helping patients navigate health care delivery options
Examples	• AI software for autonomous screening of diabetic retinopathy • Embedded software in portable cardiac ultrasonography machine that provides automated diagnosis • EHR-based algorithm that generates sepsis alerts with treatment prompts	• Smartphone app that diagnoses and treats skin conditions • Chatbot that offers mental health support • Algorithm that uses biosensor data from smartwatch to detect falls or arrhythmias	• Algorithm that uses EHR data to optimize coding for billing • Software to optimize supply chain management • Software to optimize patient and staff scheduling	• Ambient AI that transcribes patient-clinician conversations to generate notes, bills, and treatment plans • Large language model that replies to patient secure messages • Web-based patient navigator that helps schedule appointments based on patient concerns
Current trends	• Many published evaluations • Considerable efforts to provide regulatory oversight • High adoption in medical imaging • High availability but variable adoption of embedded EHR tools • Low adoption outside imaging and EHR with concerns about trust, effectiveness, safety, and implementation costs	• Large number of apps and downloads • Scant data on health effects, with some concerns about safety and effectiveness • Very limited regulatory oversight, especially regarding health effects • Optimal business model unclear • Little integration across products or with health care delivery	• Rapidly growing market, with wide adoption by health care systems • Very limited evaluation of effects on health • Limited regulatory oversight • Concerns expressed by clinicians about untoward effects on patients	• Rapidly growing market, with wide interest and adoption by health care systems • Very limited evaluation of effects on health • Limited regulatory oversight • Concerns expressed by clinicians about untoward effects on patients

Abbreviation: EHR, electronic health record.

concerns).[33-35] Examples include an app that accesses a smartphone camera to help an individual self-diagnose a dermatologic condition, a chatbot offering mental health support, and an algorithm using biosensor data from a smartwatch to detect falls or arrhythmias.[36-39] There are currently more than 350 000 mobile health apps, with AI frequently embedded in the software.[35,40] Three of 10 adults worldwide have used a mobile health app, and the market is already more than $70 billion annually.[35] Although some products, such as those monitoring for arrhythmias, are regulated by the FDA, companies can often label DTC tools as low-risk general wellness products, and may avoid regulatory or reporting requirements.[33,41,42] Most of these tools do not connect with the health care delivery enterprise, limiting health care professionals' ability to access data and review or coordinate recommendations. Tools that do interact with health care professionals require considerable upfront investment to ensure data are integrated appropriately. Some health insurers have encouraged use of DTC wellness apps via subscription coverage or reward incentives, but most usage is not reimbursed.[43-45] Other barriers include the lack of high-quality evidence regarding health benefits; concerns about trust, usability, and privacy; integration (or not) with health care systems; and uncertainty regarding the right business model.[46,47] Nevertheless, as these barriers are overcome or circumvented, these tools, because they bypass much of the existing infrastructure of traditional health care, may represent truly disruptive innovation.[48]

Health Care Business Operations Tools

Health care systems are rapidly purchasing AI tools to boost system efficiencies and operating margins.[2,49,50] Examples include AI software to optimize bed capacity, revenue cycle management, patient and staff scheduling, supply chain, and reporting requirements.[2,49-57] One popular example is use of AI by health care systems to generate prior authorization requests and, perhaps not surprisingly, insurers are similarly adopting AI tools to evaluate those requests.[58-60] Use of AI to improve business operations is already ingrained in many industries.[61] However, the consequences for patients when health care delivery organizations adopt these tools are not well understood. For example, if a health care system implements



AI software to optimize operating room scheduling, there could be large effects (good or bad) on staffing costs and on access to time-sensitive surgical interventions. Although changes in access may affect hospital quality reports, attributing these changes to the tool may be missed because use of the tool, similar to other health care operational strategies, does not require any evaluation or regulatory review.

Hybrid Tools

Many AI tools support both business operations and some aspect of clinical care or patient experience. For example, so-called *AI scribes* that listen to patient-clinician conversations help operations by generating notes and bills (reducing documentation burden) while also providing clinical support, such as offering possible diagnoses or treatment recommendations for verification. Similarly, a web-based AI patient navigator tool may bring more patients into a health care system while providing better direction to the right health care professionals, generating system revenue and improving patient access. These tools are being adopted very rapidly by health care systems.[2,23] AI scribes in particular, although initially associated with mixed results, appear well received by patients and clinicians, especially when integrated with the EHR.[62-65] Every patient-clinician conversation in the US may soon be accompanied by a live interactive AI agent.

Challenges for the Evaluation of AI Tools

In theory, given the importance of AI tools' potential effects on patients' health outcomes (as well as on health care costs and workforce), methodologically rigorous evaluations should be undertaken to generate a solid evidence base to inform their dissemination (ie, production of generalizable knowledge about the conditions under which particular effects are realized).[66,67] In practice, despite wide acceptance that AI tools can have large effects on health, there is considerable debate regarding which tools require evaluation, how evaluations should be conducted, and who is responsible.

Which AI Tools Require Evaluation of Their Health Effects?

The evaluation of clinical AI tools in peer-reviewed publications seems broadly supported. Evaluation is required for many FDA clearances, such as the de novo software as a medical device pathway, and examples of untoward consequences when AI tools were disseminated without prior peer-reviewed evaluations, such as high missed case and false alarm rates by a sepsis alert system, sparked calls for mandatory evaluations before broad use.[68,69] There are multiple evaluations of DTC tools and, although some showed benefit, others raised safety concerns, such as advice that was harmful or contrary to guidelines and a lack of adequate support during mental health crises, prompting demands that these tools should also be routinely evaluated.[34,70] However, the DTC tools evaluated to date are a tiny fraction of the market, and negative findings do not appear to have hindered market growth.

There are strong opinions, but few peer-reviewed evaluations, regarding the health effects of business operations tools. For example, many US physicians believe that insurers' use of AI tools to deny prior authorization is having widespread negative consequences for patients, but no studies have examined whether and how using AI tools has affected denial rates.[60,71,72] Not surprisingly, insurers defend their use of such tools.[73] Although peer-reviewed evaluations of business operations tools are rare, customer testimonials and use cases, often with blended claims about both system efficiencies and health care quality, are routinely used to market these products.[74,75] The business operations tools for which published evaluations are available are those with hybrid function like AI scribes, though evaluation largely focuses on patient and clinician satisfaction and on clinician workflow rather than on the effects on health care quality and patient outcomes.[62-65]

Many stakeholders believe that all of these AI tools affect health, but the impetus for peer-reviewed evaluation of potential health consequences, and likelihood that their findings will affect adoption, appears strongest for those tools most proximate to clinicians. The lack of evaluation of the health effects of business operations tools is not unique to AI, but rather is consistent with that for all health care organizational strategies, despite calls for change.[76-78]

How Should Evaluations Be Conducted?

There are considerable methodologic challenges to the generation of transferable knowledge about the health consequences of health and health care AI tools, especially how to define the intervention and context, identify the mechanism of action, capture relevant outcomes, and infer causality.

Defining the Intervention and Context

To be generalizable, any evaluation must describe what the actual intervention is and in what context it is assessed. With an AI tool, the intervention is both the AI software and its delivery package: the human-computer interface and accompanying training. A poor human-computer interface or lack of training could considerably diminish the tool's effectiveness. The context includes both the task the user is addressing and the setting. Some tools are quite task specific (eg, a sepsis alert), but for generative or agentic AI tools capable of aiding in many tasks, defining exactly what tasks are being addressed in a particular evaluation can be challenging. Even when the task is narrow, a tool's performance can vary considerably by setting.[79] For example, a sepsis alert's effectiveness may depend on site of deployment (emergency department vs hospital ward; community vs teaching hospital), how it is incorporated into workflow, and many other supporting and competing activities and priorities in the workplace.[80,81] The intervention and context for DTC tools can also be very hard to define given the potentially wide range of relevant user characteristics (eg, digital literacy), customization options of the software, and, especially for generative and agentic AI, breadth of tasks. Defining the application and setting of use cases for business operations tools in a manner that is generalizable is similarly difficult.

Many of these challenges are well known in the evaluation of any complex health services or delivery intervention, and there are well accepted approaches in health services research and implementation science to deal with them. However, the delivery science of AI does present relatively unique issues.[82] For example, there are strategies to incorporate into an evaluation how a user's learning curve influences the effect of a complex intervention, but these strategies typically assume the tool is static. With some AI tools, there is the added complexity that the tool itself can also be learning (with improving or degrading predictive performance), in turn potentially changing user confidence.

Identifying the Mechanism of Action

Generalizability is aided by knowing not just if an intervention worked, but also how. For AI, assessing how an intervention works can be considered across 3 concentric layers. The inner layer is the tool's interpretability, a description of the mathematical model that drives the tool's decision-making processes. However, deep learning, generative, and agentic AI models are so complex that descriptions of their underlying mathematic structure may be hard to interpret and provide limited insight on their likely clinical performance. The middle layer is explainability, where the model's output decisions across a variety of settings (inputs) are used to provide a picture of how the tool behaves. Explainability aids transparency when the underlying mathematical structure is hard to interpret, but it can be difficult to know if the model's performance is adequately explained across all reasonable situations it might encounter.[83]

Explainability is typically explored post hoc in existing datasets, but the accuracy of an AI model when applied retrospectively is not necessarily indicative of how care decisions will be influenced by the tool in practice.[84] The outer layer, therefore, is a prospective evaluation of how the tool performs as an actual intervention in real-world settings.[85] This layer would ideally assess the extent to which the tool's performance depends on the other features of the intervention (the human-computer interface and user training) and the context in which the tool is deployed. For AI with wide applicability, such as generative and agentic AI tools, comprehensive assessment of their real-world performance is particularly daunting.

Capturing Outcomes

For tools used in health care systems (clinical, health care business operations, and hybrid AI tools), health consequences can be captured using EHR data and standard clinical research data collection approaches. User experience may also be important, and can be captured using standard mixed-methods approaches. The problem is not how to capture relevant outcomes, but rather that these approaches are time-consuming and expensive, potentially exceeding that to develop the tool in the first place. This problem poses a 2-fold threat: either investigators will be dissuaded from doing any evaluation or evaluations will be limited, potentially omitting key outcomes or restricting to more feasible settings (eg, an academic medical center), thus compromising generalizability. For DTC tools, capturing their effects on health outcomes has added complexity, especially when the intent is to mitigate rare events in otherwise well individuals. For example, to determine whether an AI tool using heart rate data from a smartwatch could identify atrial fibrillation, investigators had to enroll more than 400 000 individuals and provide a multilayered system involving a nationwide telehealth service, independent clinician adjudication, and referral mechanisms to individuals' own clinicians.[86]

Inferring Causality

Ultimately, the goal is to understand the effect of an AI tool, not simply whether its use is associated with, but not the cause of, a particular outcome. The standard design for causal inference in health care is the randomized clinical trial (RCT). However, few AI tools have been evaluated by RCTs.[80,87-91] RCTs are typically expensive, time-consuming, and focus on 1 or 2 interventions in 1 clinical setting. For AI tools, RCT designs used more commonly in health services research and implementation science (eg, cluster or stepped wedge designs with embedded qualitative components) may often be more applicable.[92] Regardless of the design choice, if an AI tool were to be evaluated across all reasonable use cases, each tool might require multiple RCTs. Given the rate at which AI tools are being developed, relying on RCTs, at least as currently conducted, as the default approach seems quite impractical. Newer RCT designs, such as platform trials and trials embedded in the EHR, may permit faster, cheaper RCTs.[93-100] Of note, many non–health care industries, including the insurance industry, have implemented systems for rapid randomized A/B testing to evaluate their business tools, especially digital technologies.[66,101-104] By contrast, health care systems rarely use randomized A/B testing to evaluate business operations, possibly because of perceived aversion to experimentation or beliefs that randomized data are unnecessary, unhelpful, or too hard to acquire.[66,78,105]

The alternative approach is to use observational data. For example, the smartwatch study described above used a single-group cohort study to well characterize the performance characteristics of the AI tool.[86] When comparing clinical outcomes for cohorts of individuals cared for with and without an AI tool, there are numerous quasiexperimental approaches to facilitate causal inference from observational data, albeit with the caveat that they often require additional data collection (eg, detailed information on variation in the setting in which the tool is used to allow identification of candidate instrumental variables) and significant statistical expertise.[92,106]

Who Is Responsible?

Even if there were agreement on the need and approach for an evaluation of the health effects of an AI tool, it is unclear who is responsible. For tools requiring FDA review, initial evaluation rests with the developer. However, the evaluation does not necessarily include assessment of real-world health effects.[107] For tools exempt from FDA clearance, developers would likely conduct evaluations commensurate with their claims. For example, a DTC tool developer may evaluate subscriber loyalty, while a developer of a revenue cycle management tool may wish to demonstrate that their tool improves revenue.[108,109] However, in neither instance will the developer necessarily assess the health consequences of their tool. Health care delivery organizations may be motivated to conduct their own evaluations, but many may not have the funds or the expertise to conduct thorough evaluations.[7,110-115] Governments may provide grants to fund some evaluations, but such funding is far from comprehensive. Patients and communities are stakeholders who are not responsible for evaluation, but whose perspectives are crucial. However, their perspectives are not routinely included.[116,117]

Challenges for the Regulation of AI Tools

Health and health care AI tools should be subject to a governance structure that protects individuals and ensures the tools achieve their potential benefits. For other health care interventions, regulatory oversight is an important part of that governance, assuring society and markets that an intervention is credible. However, the US has no comprehensive fit-for-purpose regulatory framework for health and health care AI. Reasons include the diverse and rapidly evolving nature of AI technology, the numerous agencies with jurisdiction over different types and aspects of AI, and a lack of regulatory frameworks specifically tailored for AI within these agencies.[41] Drug and traditional medical device development also benefits from international harmonization of regulatory standards.[7] Although there are efforts to harmonize standards for AI, there are important international differences. For example, both the Biden and Trump administrations have reduced regulatory burden, while the European Union has enacted a comprehensive framework for greater regulatory oversight.[41,118,119]

In the US, the FDA regulates any AI tool classified as a medical device (ie, a technology used to diagnose, treat, mitigate, cure, or prevent a disease or condition), regardless of whether it is used by a clinician or consumer.[7] The FDA applies a risk-based, function-specific approach to provide the least burdensome assurance of safety and effectiveness. This approach includes determining what types of devices to focus on and what level of evidence is required for marketing. Although the agency has reviewed AI-enabled devices for many years, its scope is limited by resources and congressional law.[41] For example, the 21st Century Cures Act excludes software (including AI software) from the definition of medical device if its function is to provide administrative support (eg, scheduling and billing), general wellness support, some types of clinical decision support, and a number of EHR and data management functions.[7,120] As such, many AI tools discussed here are exempt from FDA regulation.[120] Even for tools over which the FDA does have authority, as noted above, clearance does not necessarily require demonstration of improved clinical outcomes.[107] Notably, generative and agentic AI tools can be capable of so many tasks as to seriously challenge the traditional intended use framework for device regulation.[3,41] Rather, they more closely resemble health care professionals, raising the idea that states could one day license AI agents as digital physicians.

EHR-based AI tools that do not meet the definition of a medical device fall under the Assistant Secretary for Technology Policy/Office of the National Coordinator for Health Information Technology, which offers voluntary certification for health information technology systems, including EHR platforms, based on demonstration of a product's transparency, risk management, trustworthiness, and fairness.[121] Similarly, some laboratory-based clinical AI tools that do not meet the definition of a medical device can be used via the Centers for Medicare & Medicaid Services (CMS) clinical laboratory improvement amendments program.[122] Use of some business operations tools like prior authorization software are also subject to CMS guidelines, but only when used by entities subject to CMS regulation.[123] DTC general wellness tools fall under the Federal Trade Commission (FTC), whose focus is privacy rights, security of personal information, and protection from false advertising and unfair or deceptive practices.[42] Partly to fill gaps in federal oversight, there are numerous regulatory and certification efforts by state governments and professional medical societies.[41]

Challenges for the Responsible Use (Implementation and Monitoring) of AI Tools

The features of AI tools that are challenging for evaluation and regulation are also challenging for implementation and monitoring. For example, for tools used in health care systems, the high dependency of their effectiveness on user training and context means health care systems require considerable infrastructure and resources to ensure users are appropriately trained and the tool is used in the optimal setting.[26,110,112] Because the tool may be subject to little prior evaluation and regulation, guidelines on how to ensure optimal use may also be lacking.[7] Even if previously shown to be beneficial, whether the tool is actually beneficial in a particular setting or continues to be beneficial over time is uncertain.[7,41,107,110] However, without adequate sample size, infrastructure, expertise, and resources, a health care system will be unable to determine whether its use of the tool is beneficial and therefore when best to adopt or de-adopt it. For DTC tools that are outside FDA oversight but under the FTC's jurisdiction, the types of practices the FTC polices are much narrower than the full suite of problems that can arise from AI implementation and threaten patients' and clinicians' interests. Additional components of appropriate implementation are the need to show that a tool's use is both fair and patient-centered (eg, not financially toxic), but demonstrating that these criteria are met is not straightforward.[124-126] Former FDA Commissioner Robert Califf recently summarized the problem, stating "I have looked far and wide, I do not believe there's a single health system in the United States that's capable of validating an AI algorithm that's put into place in a clinical care system."[25]

Table 2. Strategies to Improve Development and Dissemination of Artificial Intelligence (AI) Tools in Health and Health Care

Strategy	Rationale	Examples of rollout
Engage all stakeholders in total product life cycle management	Traditional sequenced pathway from development to evaluation to regulatory approval (if required) to dissemination with monitoring (if required), with each step shepherded by different stakeholders, is not well suited for AI tools In particular, full evaluation of health consequences not possible until tool is disseminated Greater engagement of all relevant stakeholders in each phase of a tool's life cycle may augment development, dissemination, and impact on health	• Engage patients and clinicians in design and development • Partner developers with health care systems in deployment evaluations and safety mitigation • Engage regulators, health care systems, and developers in collaborative monitoring plan, including determination of need to capture health consequences
Develop and implement the right measurement tools for evaluation, regulation and monitoring	Without new methods, any effort to increase oversight and assurance will likely be cumbersome, expensive, and potentially ineffective In particular, no good tools exist to quickly and efficiently assess health consequences of AI tools across all relevant settings and use cases	• Integrate and deploy proposed safety and compliance approaches (eg, Joint Commission/CHAI collaboration) • Develop and promulgate novel methods to facilitate fast, efficient yet robust evaluations of health consequences • Adopt standards for evaluating health consequences (not just safety) during deployment
Build the right data infrastructure and learning environment	Without a better data infrastructure and learning environment, any effort to increase oversight and assurance will likely be cumbersome, expensive, and potentially ineffective In particular: Poor data collection, access, and interoperability hamper ability to quickly and efficiently construct and interrogate relevant nationally representative data on AI tool use and effects Institutions capable of robust deployment and assessment may be poorly representative of many important settings	• Create nationally representative retrospective health care data sandbox • Learn from the FDA Sentinel program and learning health system initiatives to support curation of regularly updated data on tool deployment within nationally representative cohort of health care systems • Eventually create federated platform capable of rapid prospective evaluations capable of robust causal inference • Provide training and resources for health care systems to conduct or participate in evaluations of health consequences of AI tools
Create the right incentive structure	Current incentives are not well aligned across different stakeholders, potentially impeding progress Market forces alone may not guarantee optimal development and dissemination of AI tools The strategic initiatives listed above may require specific incentives	• Federal funds (eg, akin to those provided by HITECH) to incentivize health care systems to adopt data interoperability standards • Federal research funding for novel methods development

Abbreviations: CHAI, Coalition for Health AI; FDA, US Food and Drug Administration; HITECH, the Health Information Technology for Economic and Clinical Health Act.

Potential Solutions

All stakeholders agree that health and health care AI should be fair, appropriate, valid, effective, and safe.[5,24,113,119,121,127-140] However, creating an environment that promotes innovation and dissemination in accordance with these principles will be extremely difficult. Three big issues dominate the landscape. First, the traditional linear pathway for health care interventions, composed of discrete steps from development and evaluation through regulatory authorization to monitored dissemination, each managed by specific stakeholders, does not easily fit AI because the tools are so broad and flexible, evolving rapidly, and hard to fully evaluate until embedded in practice. Many tools will enter practice with limited evaluation, possibly no regulatory review, and, unless new approaches are adopted, be monitored only for process compliance and safety, not for effectiveness. Second, with current methods and infrastructure, any attempt to impose comprehensive evaluation, implementation, and monitoring approaches would likely be prohibitively expensive. Third, there are many stakeholders, with no overarching incentive or accountability structure. Addressing these issues will require progress in 4 areas (**Table 2**).

Engage Stakeholders in Total Product Life Cycle Management

The first step, endorsed by multiple government agencies and nongovernmental organizations, is to recognize the need for holistic, continuous, multistakeholder, team-based management of AI tools across their entire life cycle, from development to deployment.[7,26,121,135-138,141-144] For example, greater engagement of patients and clinicians with developers in the design and development phases of an AI tool can enhance its transparency and trustworthiness.[26,116,142,144] Similarly, developers can help health care systems with deployment, monitoring, and fixing problems such as model hallucinations.[26,142] And the ability of regulators to reassure the public about the safety and effectiveness of AI tools requires collaborative engagement with both developers and health care systems.[7,120,143] These multistakeholder partnerships go beyond traditional seller-client or developer-regulator relationships, but will be key to successful total product life cycle management.

Develop and Implement Proper Evaluation and Monitoring Tools

There are many proposed standards and certification processes for AI tools.[145] For example, the quality of an evaluation could be judged using APPRAISE-AI, CONSORT-AI, DECIDE-AI, SPIRIT-AI, or TRIPOD+AI guidelines; a tool's design, intended use, and performance could be reported and labeled, similar to an FDA nutrition label, with a model card and its real-world use could be assessed locally by a health care system, which could itself be certified as *AI ready*.[110,112,128,135,136,146-153] Approaches adopted by individual health care systems could also serve as resources for other systems.[110,112] Together, these evaluation and monitoring approaches could provide so-called *algorithmovigilance*, akin to pharmacovigilance.[154] However, evidence is lacking about the extent to which these various initiatives are practical, sustainable, adequately accessible, able to provide the guarantees they purport to provide, or adequately comprehensive yet not unnecessarily redundant. Piloting, fine-tuning, and integrating them into a comprehensive package would be a significant advance. One step in this regard is The Joint Commission's partnership with the Coalition for Health AI (CHAI) to develop and roll out a certification process for health care systems' responsible use of AI.[155]

That said, the monitoring standards proposed thus far focus on safety and process compliance.[110,135,136,153,155,156] None address how effectiveness (ie, improved clinical outcomes) will be determined across different settings and over time. Effectively, there is a tacit assumption that proving effectiveness is either unnecessary or is sufficiently demonstrated by premarketing testing (and will be maintained assuming a tool's use remains safe). Here, setting standards and providing training and resources for health care systems to conduct or participate in evaluations of the causal effects of an AI tool during use (eg, routine use of randomized batched stepped wedge designs or nonrandomized interrupted time series) would be helpful. This approach could extend to business operations tools (eg, setting standards and providing the training and resources for greater use of A/B testing). Novel methods to aid logistic and analytic aspects of these causal inference evaluations will also be required, especially for DTC tools and tools based on generative or agentic AI. These evaluation methods themselves may rely heavily on AI.[106,157]

Build the Proper Data Infrastructure and Learning Environment

The generation of efficient, fast, robust, and generalizable information on the safety and effectiveness of AI tools requires a significant investment in data infrastructure and analytic capacity.[158,159] Some questions can be addressed within a single health care system, but even large, well-resourced health care systems with their own analytic and implementation specialist expertise may struggle to conduct routine swift yet robust analyses. Furthermore, many evaluations, especially if intended to provide reassurance that AI tools are safe and effective across diverse settings and populations, will require data from multiple health care systems or data that link DTC apps with health care. Organizations like CHAI propose the creation of retrospective datasets from partnering health care systems.[135] Such datasets could be curated and made accessible to developers as a sandbox for AI tool development and exploration of model performance across different populations and settings. However, evaluation of real-world use requires data be shared on a tool's use, which requires a mechanism to obtain regular data updates and include many data elements not just about patients and setting, but about tool use and clinical workflows. CHAI announced recently its intention to expand in this direction, although managing frequent data updates with new, potentially proprietary, information about specific tools will be very challenging.[160] Such a system could mimic pharmacovigilance efforts like the FDA-supported Sentinel initiative, though with richer information on tool use.[136,157,161] Though Sentinel has been successful, it required considerable effort to ensure data quality, interoperability, and sharing.[162]

Crucially, any prospective evaluations of AI tools, especially involving randomization, would need an even more sophisticated and integrated collaboration across health care systems with real-time or near–real-time data access. Some individual health care systems have created the data and analytic capacity necessary to function as a learning health system.[163] Here, such a system would effectively be a multicenter learning health system collaborative. From a practical standpoint, a federated data approach may work best, where health care system data remain in place, reducing some logistic and contractual burdens. However, there are many financial, contractual, and operational challenges to the creation of such a collaborative.[158] That said, if successful and assuming the collaborative was representative of the breadth of patient populations and clinical settings seen nationally, it would not only inform the responsible use of AI tools across their entire life cycle, but could also support evaluation of other health care interventions and data interoperability solutions.[66,164-166] Such an initiative would require federal support, but aligns well with the Department of Health and Human Services' priorities.[167-170]

Create the Right Incentive Structure

The entire motivation to implement, adopt, and monitor health and health care AI tools requires adequate incentives for relevant stakeholders to participate. If left as an underregulated market, it is unclear whether the right tools will be developed, whether tools will be adopted in a manner that maximizes their beneficial effects on health while minimizing risk, and whether their effects will be measured properly. It is also possible that market forces will provide perverse incentives, such as adoption of tools that maximize developer profits or health system operating margins while inadvertently compromising health care quality or health outcomes.[171] Where there is little federal regulation, states may enact a patchwork of heterogeneous policies, making compliance overly burdensome for developers, and thus accidentally stifling investment.[41] It seems necessary, therefore, to develop legislation, regulation, and market designs that align incentives for appropriate multistakeholder engagement.

Some priorities may be achieved through the market. For example, the desire to engage patients and clinicians in the codesign of AI tools may be achieved spontaneously if developers anticipate individuals and health systems will be more willing to purchase such tools. Other priorities may be harder to achieve without specific policy levers. For example, health care systems face considerable costs standing up the digital infrastructure and technical expertise required to ensure they are meeting responsible use standards or to collaborate in federated learning initiatives. While some health care systems may make the necessary investments in response to market forces, such as the opportunity to foster commercial partnerships with developers or gain advantage over competitors, more uniform investment by all systems may require stronger government levers, such as financial incentives or regulatory requirements. A potential model is the Health Information Technology for Economic and Clinical Health Act: although numerous challenges with digital health information persist, it is notable that a relatively modest federal investment of $35 billion led to EHR adoption by more than 97% of health care systems within a decade.[167,172] Federal research funds may also be required to develop robust yet efficient evaluation methods for AI tools in real-world settings, fostering partnerships among developers, health services research/implementation science experts, and health care delivery systems. Health care market forces often fail to provide adequate care to vulnerable populations without government intervention, and the same will likely hold for the responsible use of AI in resource-poor settings. It is also likely that greater transparency with regard to the health effects of DTC or business operations tools will require greater regulation.

Implications for the Health Care Workforce

The impact of AI on the health care workforce will be wide-ranging. Clinicians may be excited by the potential benefits, worry about job displacement, and enjoy or resist requirements to improve their AI literacy. They may also have existential concerns like misalignment of human and AI ethos, goals, and principles. Although their comfort with AI is increasing, US physicians blame lack of regulatory oversight as the primary reason for their lack of trust and adoption of both clinical and health care business operations tools.[72] Health care worker unions are also raising concerns about unsafe and underregulated AI.[140] The workforce composition may also need to change, adding more experts in the development, implementation, and evaluation of AI tools.[173,174] And, of course, AI tools have the potential to augment how many clinicians work, albeit with key caveats.

First, AI tools can change which health care professional executes which task. For example, a portable echocardiography machine with AI-based interpretation upskills the ultrasound technician, potentially obviating the need for interpretation by a cardiologist or radiologist.[10] The benefit is improved access for patients by helping to close care gaps, especially in underserved settings. However, such tools could create friction between health care professional groups, eg, by challenging the scope of practice regulations. They may also change the required skills of different health care professionals, such as AI literacy. To anticipate and manage such changes, health care systems must think beyond educating a health care professional in how to use a particular AI tool and instead rethink entire organizational structures, workforce composition, skill distribution, and accountability across hierarchical levels. Of course, the most extreme example would be when a DTC tool obviates the need for an individual to seek professional care altogether. This potential will vary greatly depending on the health problem, and therefore affect specialties very differently, but such disruption seems inevitable.

Second, beyond focused learning to use any particular tool, there is a need to include a foundational understanding of AI for health care professionals in both training and continuing education.[175-178] Though many clinical tools may provide information that a health care professional can use without understanding the underlying technology (eg, computer tomography or whole genome sequencing), AI tools not only provide information but also contribute to judgment under conditions of uncertainty. Health care professionals should therefore better understand the strengths and weakness of their own decision-making and the susceptibilities and unintended consequences when sharing judgment tasks with an AI tool. These learning requirements also apply to health care administrators using business operations AI tools.

Third, new technology is often first available to, and adopted by, individuals and organizations with greater means and resources. If AI tools are to be developed and disseminated in a manner that is fair, equitable, and does not widen the digital divide, then any education and reorganization efforts must include those parts of health care delivery responsible for the most vulnerable groups. Of course, efforts to ensure fair access to AI must also be cognizant of the risks of deploying a tool with potential untoward consequences in settings poorly equipped to detect them.

Fourth, many AI tools are aimed at reducing the administrative burden on health care professionals (eg, medical record documentation or prior authorization appeals) on the premise that this burden contributes to burnout, low morale, and stress. However, this line of reasoning may have flaws. First, burnout, low morale, and stress are wicked problems; it is a tall ask to expect an AI tool to fix them.[179] Second, if freed from administrative tasks, clinicians may be asked to see more patients, which could also cause burnout. Third, focusing on tools to automate tasks such as prior authorization potentially misses the larger opportunity to rethink entirely the purpose and value of such tasks.

Fifth, because optimal AI tool development and dissemination requires much more integration with care delivery than that required of traditional technology development, health care professionals will need to understand that they are participants in the continuous learning, improvement, and evaluation cycle of these products. Although clinicians use traditional drugs and devices off label, the degree of uncertainty regarding the benefits and harms of an AI tool may be considerably greater. As long as clinicians know they have a voice in this effort, their contributions to improve the performance of AI tools could be a source of job satisfaction.

Additional Considerations

Though not discussed in detail here, important ethical and legal issues will affect adoption of health and health care AI. One issue is data rights, privacy, and ownership. Health and health care data are essential fuel for AI tools. Although one can ascribe where health and health care data originate, it is less clear who owns, or ought to own, the data, especially when aggregated and transformed, and what rights for privacy and use extend, or should extend, to whom.[180] The Health Insurance Portability and Accountability Act and US intellectual property law provide some guidance on privacy rights, security obligations, trade secrets, and ownership of tools developed from data, but are less clear on ownership of underlying data and do not protect against emergence of dark markets, reflect all ethical considerations, or fully align with state or international legislation, such as the EU's AI Act and General Data Protection Regulation.[180-187]

A second issue is how to provide ethical oversight of AI as it is deployed.[188] One argument is that decisions by health care systems to deploy AI tools are part of clinical care operations.[189,190] Any evaluation of the success or not of the deployment is quality improvement, not research, and exempt from the ethical oversight required of human subjects research.[191] The alternate argument is that AI tool deployment is rolled out under conditions of uncertainty, and the goal of any evaluation is to generate knowledge that would benefit future patients.[192] Thus, such evaluations are research, even if they are also quality improvement, and should fall under the purview of the local institutional review board and principles of the Common Rule.[66,193] Currently, both approaches occur, reflecting the broader debate regarding how best to provide ethical oversight of learning health systems.[66,194-196] An additional aspect of this problem is that, regardless of which body provides oversight (institutional review board or otherwise), the competency required to oversee ethical AI deployment may be lacking without adequate infrastructure, resources, and training.[110,112,197]

Finally, use of AI tools has thus far largely been voluntary. However, as their benefits become more established, failure to use an AI tool may be considered unethical or a breach of standard care. A health care system or professional may thus be liable in a malpractice suit for failing to use AI.[198,199] At the same time, if a plaintiff sues for an adverse outcome when care was provided in which an AI tool was involved, the question arises of whether liability rests with the health care professional, the health care system, or the developer of the tool. Though relevant case law is currently limited, developers, health care systems, and health care professionals will all have to adopt strategies to manage their liability risk.[139,200] These examples are just some of the many new issues that will need to be addressed as AI becomes more incorporated in health and health care.

Conclusions

AI will massively disrupt health and health care delivery in the coming years. The traditional approaches to evaluate, regulate, and monitor novel health care interventions are being pushed to their limits, especially with generative and agentic AI, and especially because the tools' effects cannot be fully understood until deployed in practice. Nonetheless, many tools are already being rapidly adopted, in part because they are addressing important pain points for end users. Given the many long-standing problems in health care, this disruption represents an incredible opportunity. However, the odds that this disruption will improve health for all will depend heavily on creation of an ecosystem capable of rapid, efficient, robust, and generalizable knowledge about the consequences of these tools on health.

ARTICLE INFORMATION

Accepted for Publication: September 5, 2025.

Published Online: October 13, 2025. doi:10.1001/jama.2025.18490

Author Affiliations: *JAMA*, Chicago, Illinois (Angus, Khera, Lieu, Perlis, Ross, Seymour); Yale University, New Haven, Connecticut (Angus, Khera); Kaiser Permanente, Pleasanton, California (Lieu, V. Liu, Bindman, Lee, Ouyang); Northwestern University, Chicago, Illinois (Ahmad); CHAI, Boston, Massachusetts (Anderson); Emory University, Atlanta, Georgia (Bhavani, Gichoya); Harvard T.H. Chan School of Public Health, Boston, Massachusetts (Brennan); MIT, Cambridge, Massachusetts (Celi); American Medical Association, Chicago, Illinois (Chen, Desai, Lomis); Harvard Law School, Cambridge, Massachusetts (Cohen); University of Birmingham, Birmingham, United Kingdom (Denniston, X. Liu); Vanderbilt University Medical Center, Nashville, Tennessee (Embí); Imperial College London, London, United Kingdom (Faisal); Universität Bayreuth, Germany (Faisal); Johns Hopkins University, Baltimore, Maryland (Ferryman, Gross, Saria); Epic, Verona, Wisconsin (Gerhart); Stanford University, Stanford, California (Hernandez-Boussard, Mello, N. H. Shah); Google, Mountain View, California (Howell); University of Pennsylvania, Philadelphia (Johnson); Microsoft AI, London, United Kingdom (X. Liu); Carnegie Mellon University, Pittsburgh, Pennsylvania (London); University of California, San Diego Health (Longhurst); Boston Children's Hospital, Boston, Massachusetts (Mandl); Harvard Medical School, Boston, Massachusetts (Mandl, Perlis); Kaiser Permanente Bernard J. Tyson School of Medicine, Pasadena, California (McGlynn); San Ysidro Health, San Diego, California (Munoz); Yale School of Medicine, New Haven, Connecticut (Ohno-Machado, Ross, Schwamm); Apple, Cupertino, California (Phillips); Microsoft, Redmond, Washington (Rhew, Weinstein); University of Pittsburgh, Pittsburgh, Pennsylvania (Seymour); Meta, Menlo Park, California (R. Shah); University of Utah, Salt Lake City, Utah (R. Shah); University of California, San Diego (Singh); Sutter Health, Sacramento, California (Solomon); The Joint Commission, Oakbrook Terrace, Illinois (Spates); University of Michigan Medical School, Ann Arbor (Spector-Bagdady); Health Care and Organizational Economist, Palo Alto, California (Wang); University of Michigan, Ann Arbor (Wiens); Editor in Chief, *JAMA* and the JAMA Network, Chicago, Illinois (Bibbins-Domingo).

Conflict of Interest Disclosures: Dr Khera reported receiving grants from National Institutes of Health (NIH; 1R01HL167858, 1R01AG089981, and 5K23HL153775), Bristol Myers Squibb, BridgeBio, and Novo Nordisk and being an academic cofounder of Ensight-AI and Evidence2Health outside the submitted work. Dr V. Liu reported receiving grants from NIH during the conduct of the study. Dr Ahmad reported receiving nonfinancial support from Pfizer, grants from Atman Health, Tempus and Abiomed, and personal fees from AstraZeneca and Alnylam Pharmaceuticals outside the submitted work. Dr Bhavani reported receiving grants from NIH and Kaiser Permanente outside the submitted work. Dr Cohen reported being supported in part by a Novo Nordisk Foundation Grant for a scientifically independent International Collaborative Bioscience Innovation & Law Programme (Inter-CeBIL programme No. NNF23SA0087056); receiving grants from Novo Nordisk Foundation (NNF23SA0087056) during the conduct of the study serving as the chair of the ethics advisory board for Illumina, a member of the Bayer Bioethics Council a bioethics advisor for Bexorg, and an advisor for World Class Health and Manhattan Neuroscience; receiving compensation for speaking at events organized by Philips with the Washington Post as well as the Doctors Company; and being retained as an expert in health privacy, gender-affirming care, and reproductive technology lawsuits. Dr Embi reported having a patent for Vanderbilt Algorithmovigilance Monitoring and Operations System licensed to VigilAI, Inc. Dr Ferryman reported receiving personal fees from Merck KGaA and All of Us Research Program outside the submitted work. Dr Gross reported receiving grants from Yosemite to Johns Hopkins during the conduct of the study; being founder/CEO of de-bi, co. outside the submitted work; and having a patent for application No. 19/220,825 and PCT/US25/31199 pending through de-bi, co. that is broadly relevant to health systems transparency and data ownership. Dr Howell reported being employed by Google and owning equity in Alphabet outside the submitted work and having a patent related to understanding sound using deep learning licensed to Google (EP3762942B1). Dr Johnson reported receiving grants from NIH during the conduct of the study. Dr X. Liu reported being an employee of Microsoft AI and previously serving as a health scientist at Apple and receiving consulting fees from Hardian Health and Conceivable Life Sciences; and currently serving as a board member of the MHRA Airlock Steering Committee. Dr McGlynn reported receiving travel expense to summit from Gordon and Betty Moore Foundation and being employed by Kaiser Permanente during the conduct of the study. Dr Mello reported receiving personal fees from AVIA and Augmedix, and grants from Stanford Health Care, Patient-Centered Outcomes Research Institute, and Stanford Impact Labs outside the submitted work and spousal employment with Cisco Systems, Inc. Dr Perlis reported receiving personal fees from Genomind, Circular Genomics, and Alkermes outside the submitted work. Dr Ross reported receiving grants from the Food and Drug Administration, Johnson & Johnson, Medical Devices Innovation Consortium, AHRQ, NIH/National Heart, Lung, and Blood Institute, and Arnold Ventures outside the submitted work and serving as an expert witness at the request of

relator's attorneys, the Greene Law Firm, in a qui tam suit alleging violations of the False Claims Act and Anti-Kickback Statute against Biogen Inc. that was settled September 2022. Dr Saria reported serving as CEO and a board member of Bayesian Health, serving as a board member of the Coalition of Health AI, serving as an advisor for Century Health, serving as an advisory board member of Duality Tech, and receiving grants from Gordon and Betty Moore Foundation, FDA Center of Excellence, and NIH Center outside the submitted work. Dr Schwamm reported having a patent for US2024031761 pending for an AI enabled stroke classifier and serving as an unpaid content advisor on digital health to the Stroke editorial board and a voluntary member of a client hospital advisory committee for Abridge, an ambient AI clinical documentation company, on behalf of Yale New Haven Health System. Dr Seymour reported receiving grants from NIH during the conduct of the study and personal fees from Octapharma, Beckman Coulter, and Edwards LifeSciences outside the submitted work. Dr N. Shah reported being a cofounder of Prealize Health (a predictive analytics company) as well as Atropos Health (an on-demand evidence generation company) and serving on the board of the Coalition for Healthcare AI (CHAI), a consensus-building organization providing guidelines for the responsible use of artificial intelligence in health care and serving as an advisor to Opala, Curai Health, JnJ Innovative Medicines, and AbbVie Pharmaceuticals. Dr Singh reported receiving personal fees from Google through serving on Google's Consumer Health Advisory Panel and as a paid consultant during the conduct of the study. Dr Spector-Bagdady reported receiving grants from National Center for Advancing Translational Sciences and The Greenwall Foundation during the conduct of the study. Dr Wang reported receiving personal fees from Kaiser Permanente, Health AI Partnership, SCAN Foundation, Emerson Collective, and Gordon and Betty Moore Foundation outside the submitted work; and serving on the advisory council for Health AI Partnership. Dr Wawira Gichoya reported receiving grants from NIH, Clarity, and LUNIT during the conduct of the study and speaker fees from Cook Medical outside the submitted work. Dr Wiens reported receiving travel fees from Kaiser Permanente during the conduct of the study. No other disclosures were reported.

Group Information: The JAMA Summit on AI participants appear in the Supplement.

Disclaimer: Drs Angus, Khera, Lieu, Perlis, Ross, Seymour, and Bibbins-Domingo are editors at *JAMA* and the JAMA Network but were not involved in any of the decisions regarding review of the manuscript or its acceptance.

Disclaimer: The US Food and Drug Administration (FDA) participated in the JAMA AI Summit on October 15-16, 2024, but the contents of this article represent the views of the authors and do not necessarily represent the official views of, and are not an endorsement by, the FDA/Department of Health and Human Services or the US government. The views, findings, and interpretations contained in this article do not constitute FDA guidance, position on this matter, or legally enforceable requirements.

Additional Contributions: We dedicate this work to the memory of Atul Butte, a giant in the field whose advice in the planning of this JAMA Summit was invaluable.

REFERENCES

1. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. *Nat Med*. 2022;28(1):31-38. doi:10.1038/s41591-021-01614-0

2. Canada's Drug Agency. *2025 Watch List: Artificial Intelligence in Health Care Technologies*. Canadian Agency for Drugs and Technologies in Health; 2025.

3. Wachter RM, Brynjolfsson E. Will generative artificial intelligence deliver on its promise in health care? *JAMA*. 2024;331(1):65-69. doi:10.1001/jama.2023.25054

4. Alowais SA, Alghamdi SS, Alsuhebany N, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC Med Educ*. 2023;23(1):689. doi:10.1186/s12909-023-04698-z

5. Ratwani RM, Classen D, Longhurst C. The compelling need for shared responsibility of AI oversight: lessons from health IT certification. *JAMA*. 2024;332(10):787-788. doi:10.1001/jama.2024.12630

6. Wiens J, Saria S, Sendak M, et al. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med*. 2019;25(9):1337-1340. doi:10.1038/s41591-019-0548-6

7. Warraich HJ, Tazbaz T, Califf RM. FDA perspective on the regulation of artificial intelligence in health care and biomedicine. *JAMA*. 2025;333(3):241-247. doi:10.1001/jama.2024.21451

8. Howell MD, Corrado GS, DeSalvo KB. Three epochs of artificial intelligence in health care. *JAMA*. 2024;331(3):242-244. doi:10.1001/jama.2023.25057

9. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316(22):2402-2410. doi:10.1001/jama.2016.17216

10. Holste G, Oikonomou EK, Tokodi M, Kovács A, Wang Z, Khera R. Complete AI-enabled echocardiography interpretation with multitask deep learning. *JAMA*. 2025;334(4):306-318. doi:10.1001/jama.2025.8731

11. Kamran F, Tjandra D, Heiler A, et al. Evaluation of sepsis prediction models before onset of treatment. *NEJM*. 2024;1(3):AIoa2300032. doi:10.1056/AIoa2300032

12. Wong A, Otles E, Donnelly JP, et al. External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients. *JAMA Intern Med*. 2021;181(8):1065-1070. doi:10.1001/jamainternmed.2021.2626

13. Alyoubi WL, Shalash WM, Abulkhair MF. Diabetic retinopathy detection through deep learning techniques: a review. *Inform Med Unlocked*. 2020;20:100377. doi:10.1016/j.imu.2020.100377

14. Soh CH, Wright L, Baumann A, et al; AGILE-Echo investigators in Tasmania, Western New South Wales, Northern Victoria, Western Australia, Western Queensland and Northern Territory, Australia. Use of artificial intelligence-guided echocardiography to detect cardiac dysfunction and heart valve disease in rural and remote areas: rationale and design of the AGILE-echo trial. *Am Heart J*. 2024;277:11-19. doi:10.1016/j.ahj.2024.08.004

15. Boussina A, Shashikumar SP, Malhotra A, et al. Impact of a deep learning sepsis prediction model on quality of care and survival. *NPJ Digit Med*. 2024;7(1):14. doi:10.1038/s41746-023-00986-6

16. Arabi YM, Alsaawi A, Alzahrani M, et al; SCREEN Trial Group and the Saudi Critical Care Trials Group. Electronic sepsis screening among patients admitted to hospital wards: a stepped-wedge cluster randomized trial. *JAMA*. 2025;333(9):763-773. doi:10.1001/jama.2024.25982

17. Bhavani SV, Holder A, Miltz D, et al. The precision resuscitation with crystalloids in sepsis (PRECISE) trial: a trial protocol. *JAMA Netw Open*. 2024;7(9):e2434197-e2434197. doi:10.1001/jamanetworkopen.2024.34197

18. Henry KE, Hager DN, Pronovost PJ, Saria S. A targeted real-time early warning score (TREWScore) for septic shock. *Sci Transl Med*. 2015;7(299):299ra122. doi:10.1126/scitranslmed.aab3719

19. Adams R, Henry KE, Sridharan A, et al. Prospective, multi-site study of patient outcomes after implementation of the TREWS machine learning-based early warning system for sepsis. *Nat Med*. 2022;28(7):1455-1460. doi:10.1038/s41591-022-01894-0

20. US Food and Drug Administration. Artificial intelligence in software as a medical device. March 25, 2025. Accessed September 30, 2025. https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-software-medical-device

21. Shafi S, Parwani AV. Artificial intelligence in diagnostic pathology. *Diagn Pathol*. 2023;18(1):109. doi:10.1186/s13000-023-01375-z

22. Rajpurkar P, Lungren MP. The current and future state of AI interpretation of medical images. *N Engl J Med*. 2023;388(21):1981-1990. doi:10.1056/NEJMra2301725

23. Poon EG, Lemak CH, Rojas JC, Guptill J, Classen D. Adoption of artificial intelligence in healthcare: survey of health system priorities, successes, and challenges. *J Am Med Inform Assoc*. 2025;32(7):1093-1100. doi:10.1093/jamia/ocaf065

24. The American Medical Association. AMA augmented intelligence research: physician sentiments around the use of AI in health care: motivations, opportunities, risks, and use cases: shifts from 2023 to 2024. Published February 2025. Accessed September 23, 2025. https://www.ama-assn.org/system/files/physician-ai-sentiment-report.pdf

25. Tahir D. Health care AI, intended to save money, turns out to require a lot of expensive humans. *CBS News*. December 23, 2024. Accessed June 27, 2025. https://www.cbsnews.com/news/health-care-ai-cost-humans/

26. Esmaeilzadeh P. Challenges and strategies for wide-scale artificial intelligence (AI) deployment in healthcare practices: a perspective for healthcare organizations. *Artif Intell Med*. 2024;151:102861. doi:10.1016/j.artmed.2024.102861

27. Parikh RB, Helmchen LA. Paying for artificial intelligence in medicine. *NPJ Digit Med*. 2022;5(1):63. doi:10.1038/s41746-022-00609-6

28. Krishnamoorthy M, Sjoding MW, Wiens J. Off-label use of artificial intelligence models in healthcare. *Nat Med*. 2024;30(6):1525-1527. doi:10.1038/s41591-024-02870-6

29. Khanijahani A, Iezadi S, Dudley S, Goettler M, Kroetsch P, Wise J. Organizational, professional, and patient characteristics associated with artificial intelligence adoption in healthcare: a systematic review. *Health Policy Technol*. 2022;11(1):100602. doi:10.1016/j.hlpt.2022.100602

30. Jabbour S, Fouhey D, Shepard S, et al. Measuring the impact of AI in the diagnosis of hospitalized patients: a randomized clinical vignette survey study. *JAMA*. 2023;330(23):2275-2284. doi:10.1001/jama.2023.22295

31. Khera R, Simon MA, Ross JS. Automation bias and assistive AI: risk of harm from AI-driven clinical decision support. *JAMA*. 2023;330(23):2255-2257. doi:10.1001/jama.2023.22557

32. Biro JM, Handley JL, Malcolm McCurry J, et al. Opportunities and risks of artificial intelligence in patient portal messaging in primary care. *NPJ Digit Med*. 2025;8(1):222. doi:10.1038/s41746-025-01586-2

33. Mandl KD. How AI could reshape health care-rise in direct-to-consumer models. *JAMA*. 2025;333(19):1667-1669. doi:10.1001/jama.2025.0946

34. De Freitas J, Cohen IG. The health risks of generative AI-based wellness apps. *Nat Med*. 2024;30(5):1269-1275. doi:10.1038/s41591-024-02943-6

35. Grand View Research. mHealth market size, share & trends analysis report (2024-2030). Accessed June 27, 2025. https://www.grandviewresearch.com/industry-analysis/mhealth-market

36. Wongvibulsin S, Yan MJ, Pahalyants V, Murphy W, Daneshjou R, Rotemberg V. Current state of dermatology mobile applications with artificial intelligence features. *JAMA Dermatol*. 2024;160(6):646-650. doi:10.1001/jamadermatol.2024.0468

37. Haque MDR, Rubya S. An overview of chatbot-based mobile mental health apps: insights from app description and user reviews. *JMIR Mhealth Uhealth*. 2023;11(1):e44838. doi:10.2196/44838

38. Triantafyllidis A, Kondylakis H, Katehakis D, et al. Smartwatch interventions in healthcare: a systematic review of the literature. *Int J Med Inform*. 2024;190:105560. doi:10.1016/j.ijmedinf.2024.105560

39. Reeder B, David A. Health at hand: a systematic review of smart watch uses for health and wellness. *J Biomed Inform*. 2016;63:269-276. doi:10.1016/j.jbi.2016.09.001

40. Wells C, Spry C. *An Overview of Smartphone Apps: CADTH Horizon Scan*. Canadian Agency for Drugs and Technologies in Health; 2022.

41. Mello MM, Cohen IG. Regulation of health and health care artificial intelligence. *JAMA*. 2025;333(20):1769-1770. doi:10.1001/jama.2025.3308

42. Simon DA, Shachar C, Cohen IG. Skating the line between general wellness products and regulated devices: strategies and implications. *J Law Biosci*. 2022;9(2):lsac015. doi:10.1093/jlb/lsac015

43. Abu Dabrh AM, Reddy K, Beech BM, Moore M. Health & wellness coaching services: making the case for reimbursement. *Am J Lifestyle Med*. 2024;19(6):800-813. doi:10.1177/15598276241266784

44. Gordon WJ, Landman A, Zhang H, Bates DW. Beyond validation: getting health apps into clinical practice. *NPJ Digit Med*. 2020;3(1):14. doi:10.1038/s41746-019-0212-z

45. Powell AC, Bowman MB, Harbin HT. Reimbursement of apps for mental health: findings from interviews. *JMIR Ment Health*. 2019;6(8):e14724. doi:10.2196/14724

46. He X, Zheng X, Ding H. Existing barriers faced by and future design recommendations for direct-to-consumer health care artificial intelligence apps: scoping review. *J Med Internet Res*. 2023;25:e50342. doi:10.2196/50342

47. Giebel GD, Speckemeier C, Abels C, et al. Problems and barriers related to the use of digital health applications: scoping review. *J Med Internet Res*. 2023;25:e43808. doi:10.2196/43808

48. Bower JL, Christensen CM. Disruptive technologies: Catching the wave. *Harv Bus Rev*. 1995;73(1):43-53.

49. Adner R, Weinstein J. GenAI could transform how health care works. *Harvard Business Review*. November 27, 2023. Accessed September 23, 2025. https://hbr.org/2023/11/genai-could-transform-how-health-care-works

50. Chen M, Decary M. *Artificial Intelligence in Healthcare: An Essential Guide for Health Leaders*. Sage Publications; 2020:10-18.

51. Boussina A, Krishnamoorthy R, Quintero K, et al. Large language models for more efficient reporting of hospital quality measures. *NEJM AI*. 2024;1(11):AIcs2400420. doi:10.1056/aics2400420

52. Mahmoudian Y, Nemati A, Safaei AS. A forecasting approach for hospital bed capacity planning using machine learning and deep learning with application to public hospitals. *Healthc Anal (N Y)*. 2023;4:100245. doi:10.1016/j.health.2023.100245

53. Jalil MS, Mehedy MTJ, Saeed M, Snigdha EZ, Khan N. Optimizing revenue cycle management in healthcare: AI and IT solutions for business process automation. *The American Journal of Engineering and Technology*. 2025;7(03):141-162. doi:10.37547/tajet/Volume07Issue03-14

54. Knight DR, Aakre CA, Anstine CV, et al. Artificial intelligence for patient scheduling in the real-world health care setting: a metanarrative review. *Health Policy Technol*. 2023;12(4):100824. doi:10.1016/j.hlpt.2023.100824

55. O'Callahan K, Sitters S, Petersen M. 'You make the call': improving radiology staff scheduling with AI-generated self-rostering in a medical imaging department. *Radiography (Lond)*. 2024;30(3):862-868. doi:10.1016/j.radi.2024.03.014

56. Deveci M. Effective use of artificial intelligence in healthcare supply chain resilience using fuzzy decision-making model. *Soft Comput*. Published online July 17, 2023. doi:10.1007/s00500-023-08906-2

57. Khan FS, Masum AA, Adam J, Karim MR, Afrin S. AI in healthcare supply chain management: enhancing efficiency and reducing costs with predictive analytics. *J Comput Sci Technol Stud*. Published online November 17, 2024. doi:10.32996/jcsts.2024.6.5.8

58. Busis NA, Khokhar B, Callaghan BC. Streamlining prior authorization to improve care. *JAMA Neurol*. 2024;81(1):5-6. doi:10.1001/jamaneurol.2023.4324

59. Anand S. Automating prior authorization decisions using machine learning and health claim data. *Int J Artif Intel, Data Sci, Machine Learning*. 2022;3(3):35-44. doi:10.63282/3050-9262.IJAIDSML-V3I3P104

60. Shachar C, Killelea A, Gerke S. *AI and Health Insurance Prior Authorization: Regulators Need to Step Up Oversight*. Health Affairs Forefront; 2024.

61. Bharadiya JP, Thomas RK, Ahmed F. Rise of artificial intelligence in business and industry. *J Eng Res Rep*. 2023;25(3):85-103. doi:10.9734/jerr/2023/v25i3893

62. Liu T-L, Hetherington TC, Dharod A, et al. Does AI-powered clinical documentation enhance clinician efficiency? a longitudinal study. *NEJM AI*. 2024;1(12):AIoa2400659. doi:10.1056/AIoa2400659

63. Tierney AA, Gayre G, Hoberman B, et al. Ambient artificial intelligence scribes to alleviate the burden of clinical documentation. *NEJM Catalyst*. 2024;5(3)23.0404. doi:10.1093/jamia/ocae295

64. Cao DY, Silkey JR, Decker MC, Wanat KA. Artificial intelligence-driven digital scribes in clinical documentation: pilot study assessing the impact on dermatologist workflow and patient encounters. *JAAD Int*. 2024;15:149-151. doi:10.1016/j.jdin.2024.02.009

65. Tierney AA, Gayre G, Hoberman B, et al. Ambient Artificial Intelligence Scribes: Learnings after 1 Year and over 2.5 Million Uses. *NEJM Catalyst*. 2025;6(5):CAT.25.0040. doi:10.1056/CAT.25.0040

66. Angus DC, Huang AJ, Lewis RJ, et al; JAMA Summit on Clinical Trials Participants. The integration of clinical trials with the practice of medicine: repairing a house divided. *JAMA*. 2024;332(2):153-162. doi:10.1001/jama.2024.4088

67. Jackson GP, Shortliffe EH. Understanding the evidence for artificial intelligence in healthcare. *BMJ Qual Saf*. 2025;34(7):421-424. doi:10.1136/bmjqs-2025-018559

68. Habib AR, Lin AL, Grant RW. The epic sepsis model falls short-the importance of external validation. *JAMA Intern Med*. 2021;181(8):1040-1041. doi:10.1001/jamainternmed.2021.3333

69. Kwong JCC, Nickel GC, Wang SCY, Kvedar JC. Integrating artificial intelligence into healthcare systems: more than just the algorithm. *NPJ Digit Med*. 2024;7(1):52. doi:10.1038/s41746-024-01066-z

70. Grundy Q. A review of the quality and impact of mobile health apps. *Ann Rev Public Health*. 2022;43:117-134. doi:10.1146/annurev-publhealth-052020-103738

71. Lubell J. How AI is leading to more prior authorization denials. *AMA News Wire*. March 10, 2025. Accessed September 26, 2025. https://www.ama-assn.org/practice-management/prior-authorization/how-ai-leading-more-prior-authorization-denials

72. American Medical Association. 2024 AMA prior authorization physician survey. Accessed June 27, 2025. https://www.ama-assn.org/system/files/prior-authorization-survey.pdf

73. Pearce K. Blue Cross Blue Shield of Massachusetts uses artificial intelligence to speed review time, automate authorizations, and eliminate administrative costs. PRNewswire. October 12, 2022. Accessed June 27, 2025. https://newsroom.bluecrossma.com/2022-10-12-BLUE-CROSS-BLUE-SHIELD-OF-MASSACHUSETTS-USES-ARTIFICIAL-INTELLIGENCE-TO-SPEED-REVIEW-TIME,-AUTOMATE-AUTHORIZATIONS-ELIMINATE-ADMINISTRATIVE-COSTS

74. Palantir. Palantir for hospitals. Accessed June 27, 2025. https://www.palantir.com/offerings/palantir-for-hospitals/

75. DataBricks Data Solutions. Case studies: explore our case studies to learn how we empower healthcare and life sciences organizations with Data-Driven AI Solutions. Accessed June 27, 2025. https://dnamic.ai/data-case-studies/

76. Olsen LM, Nabel EG, McGinnis JM, McClellan MB. Evidence-based medicine and the changing nature of health care: 2007 IOM Annual Meeting Summary. National Academies Press; 2008.

77. McGinnis JM, Olsen L, Goolsby WA, Grossmann C. *Engineering a Learning Healthcare System: A Look at the Future: Workshop Summary*. National Academies Press; 2011.

78. Finkelstein A, Taubman S. Health care policy. Randomize evaluations to improve health care delivery. *Science*. 2015;347(6223):720-722. doi:10.1126/science.aaa2362

79. Cao J, Zhang X, Shahinian V, et al. Generalizability of an acute kidney injury prediction model across health systems. *Nat Mach Intell*. 2022;4(12):1121-1129. doi:10.1038/s42256-022-00563-8

80. Angus DC. Do sepsis alerts help? *JAMA*. 2025;333(9):759-760. doi:10.1001/jama.2024.25818

81. Wong A, Cao J, Lyons PG, et al. Quantification of sepsis model alerts in 24 us hospitals before and during the COVID-19 pandemic. *JAMA Netw Open*. 2021;4(11):e2135286-e2135286. doi:10.1001/jamanetworkopen.2021.35286

82. Li RC, Asch SM, Shah NH. Developing a delivery science for artificial intelligence in healthcare. *NPJ Digit Med*. 2020;3(1):107. doi:10.1038/s41746-020-00318-y

83. Shah NH, Milstein A, Bagley SC. Making machine learning models clinically useful. *JAMA*. 2019;322(14):1351-1352. doi:10.1001/jama.2019.10306

84. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. *Lancet Digit Health*. 2021;3(11):e745-e750. doi:10.1016/S2589-7500(21)00208-9

85. He B, Kwan AC, Cho JH, et al. Blinded, randomized trial of sonographer versus AI cardiac function assessment. *Nature*. 2023;616(7957):520-524. doi:10.1038/s41586-023-05947-3

86. Perez MV, Mahaffey KW, Hedlin H, et al; Apple Heart Study Investigators. Large-scale assessment of a smartwatch to identify atrial fibrillation. *N Engl J Med*. 2019;381(20):1909-1917. doi:10.1056/NEJMoa1901183

87. Han R, Acosta JN, Shakeri Z, Ioannidis JPA, Topol EJ, Rajpurkar P. Randomised controlled trials evaluating artificial intelligence in clinical practice: a scoping review. *Lancet Digit Health*. 2024;6(5):e367-e373. doi:10.1016/S2589-7500(24)00047-5

88. Topol EJ. Welcoming new guidelines for AI clinical research. *Nat Med*. 2020;26(9):1318-1320. doi:10.1038/s41591-020-1042-x

89. Martindale APL, Llewellyn CD, de Visser RO, et al. Concordance of randomised controlled trials for artificial intelligence interventions with the CONSORT-AI reporting guidelines. *Nat Commun*. 2024;15(1):1619. doi:10.1038/s41467-024-45355-3

90. Angus DC. Randomized clinical trials of artificial intelligence. *JAMA*. 2020;323(11):1043-1045. doi:10.1001/jama.2020.1039

91. Wu E, Wu K, Daneshjou R, Ouyang D, Ho DE, Zou J. How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals. *Nat Med*. 2021;27(4):582-584. doi:10.1038/s41591-021-01312-x

92. Curran GM, Landes SJ, McBain SA, et al. Reflections on 10 years of effectiveness-implementation hybrid studies. *Front Health Serv*. 2022;2:1053496. doi:10.3389/frhs.2022.1053496

93. Angus DC, Alexander BM, Berry S, et al; Adaptive Platform Trials Coalition. Adaptive platform trials: definition, design, conduct and reporting considerations. *Nat Rev Drug Discov*. 2019;18(10):797-807. doi:10.1038/s41573-019-0034-3

94. Berry SM, Connor JT, Lewis RJ. The platform trial: an efficient strategy for evaluating multiple treatments. *JAMA*. 2015;313(16):1619-1620. doi:10.1001/jama.2015.2316

95. Park JW, Liu MC, Yee D, et al; I-SPY 2 Investigators. Adaptive randomization of neratinib in early breast cancer. *N Engl J Med*. 2016;375(1):11-22. doi:10.1056/NEJMoa1513750

96. Huang DT, McVerry BJ, Horvat C, et al; UPMC REMAP-COVID Group, on behalf of the REMAP-CAP Investigators. Implementation of the randomized embedded multifactorial adaptive platform for COVID-19 (REMAP-COVID) trial in a US health system: lessons learned and recommendations. *Trials*. 2021;22(1):100. doi:10.1186/s13063-020-04997-6

97. McCreary EK, Bariola JR, Minnier TE, et al. The comparative effectiveness of COVID-19 monoclonal antibodies: a learning health system randomized clinical trial. *Contemp Clin Trials*. 2022;119:106822. doi:10.1016/j.cct.2022.106822

98. Huang DT, McCreary EK, Bariola JR, et al. The UPMC OPTIMISE-C19 (Optimizing Treatment and Impact of Monoclonal Antibodies Through Evaluation for COVID-19) trial: a structured summary of a study protocol for an open-label, pragmatic, comparative effectiveness platform trial with response-adaptive randomization. *Trials*. 2021;22(1):363. doi:10.1186/s13063-021-05316-3

99. Goldhaber NH, Jacobs MB, Laurent LC, et al. Integrating clinical research into electronic health record workflows to support a learning health system. *JAMIA Open*. 2024;7(2):ooae023. doi:10.1093/jamiaopen/ooae023

100. Horwitz LI, Kuznetsova M, Jones SA. Creating a learning health system through rapid-cycle, randomized testing. *N Engl J Med*. 2019;381(12):1175-1179. doi:10.1056/NEJMsb1900856

101. Gallo A. A refresher on A/B testing. *Harvard Business Review*. June 28, 2017. Accessed January 21, 2024. https://hbr.org/2017/06/a-refresher-on-ab-testing

102. Sawant N, Namballa CB, Sadagopan N, Nassif H. Contextual multi-armed bandits for causal marketing. *arXiv*. Preprint posted October 2, 2018. doi:10.48550/arXiv.1810.01859

103. Kohavi R, Thomke S. The surprising power of online experiments. *Harv Bus Rev*. 2017;95(5):74-82.

104. Solomon J. A/B testing in the insurance industry: what top firms are doing. *Kameleoon*. Accessed June 29, 2025. https://www.kameleoon.com/blog/ab-testing-insurance-industry#:~:text=A/B

105. Meyer MN, Heck PR, Holtzman GS, et al. Objecting to experiments that compare two unobjectionable policies or treatments. *Proc Natl Acad Sci U S A*. 2019;116(22):10723-10728. doi:10.1073/pnas.1820701116

106. Dahabreh IJ, Bibbins-Domingo K. Causal inference about the effects of interventions from observational studies in medical journals. *JAMA*. 2024;331(21):1845-1853. doi:10.1001/jama.2024.7741

107. Muralidharan V, Adewale BA, Huang CJ, et al. A scoping review of reporting gaps in FDA-approved AI medical devices. *NPJ Digit Med*. 2024;7(1):273. doi:10.1038/s41746-024-01270-x

108. Azad-Khaneghah P, Neubauer N, Miguel Cruz A, Liu L. Mobile health app usability and quality rating scales: a systematic review. *Disabil Rehabil Assist Technol*. 2021;16(7):712-721. doi:10.1080/17483107.2019.1701103

109. Pennington R. *Artificial intelligence (AI) and its Opportunity in Healthcare Organizations Revenue Cycle Management*. RCM; 2023.

110. Callahan A, McElfresh D, Banda JM, et al. Standing on FURM ground: a framework for evaluating fair, useful, and reliable AI models in health care systems. *NEJM Catalyst*. 2024;5(10):CAT.24.0131. doi:10.1056/CAT.24.0131

111. Longhurst CA, Singh K, Chopra A, Atreja A, Brownstein JS. A call for artificial intelligence implementation science centers to evaluate clinical effectiveness. *NEJM*. 2024;1(8):AIp2400223. doi:10.1056/AIp2400223

112. Bedoya AD, Economou-Zavlanos NJ, Goldstein BA, et al. A framework for the oversight and local deployment of safe and high-quality prediction models. *J Am Med Inform Assoc*. 2022;29(9):1631-1636. doi:10.1093/jamia/ocac078

113. Reddy S, Rogers W, Makinen VP, et al. Evaluation framework to guide implementation of AI systems into healthcare settings. *BMJ Health Care Inform*. 2021;28(1):e100444. doi:10.1136/bmjhci-2021-100444

114. Goldstein J, Weitzman D, Lemerond M, Jones A. Determinants for scalable adoption of autonomous AI in the detection of diabetic eye disease in diverse practice types: key best practices learned through collection of real-world data. *Front Digit Health*. 2023;5:1004130. doi:10.3389/fdgth.2023.1004130

115. Castro M, Bishop D, Weitzman D, Ramirez R. Enhancing diabetic eye disease detection through autonomous artificial intelligence implementation in a federally qualified health center. *Diabetes*. 2024;73(suppl 1). doi:10.2337/db24-57-OR

116. Winter PD, Carusi A. (De)troubling transparency: artificial intelligence (AI) for clinical applications. *Med Humanit*. 2023;49(1):17-26. doi:10.1136/medhum-2021-012318

117. Adus S, Macklin J, Pinto A. Exploring patient perspectives on how they can and should be engaged in the development of artificial intelligence (AI) applications in health care. *BMC Health Serv Res*. 2023;23(1):1163. doi:10.1186/s12913-023-10098-2

118. Edwards L. The EU AI Act: a summary of its significance and scope. *Artif Intell*. 2021;1:25.

119. European Parliament. EU AI Act: first regulation on artificial intelligence. Accessed June 29, 2025. https://www-europarl-europa-eu.pitt.

idm.oclc.org/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

120. US Food and Drug Administration. Artificial intelligence and machine learning in software as a medical device. Accessed June 27, 2025. https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device

121. US Dept of Health and Human Services. Health data, technology, and interoperability: certification program updates, algorithm transparency, and information sharing (HTI-1) final rule. Accessed June 27, 2025. https://www.federalregister.gov/documents/2024/01/09/2023-28857/health-data-technology-and-interoperability-certification-program-updates-algorithm-transparency-and

122. Zhang DY, Venkat A, Khasawneh H, Sali R, Zhang V, Pei Z. Implementation of digital pathology and artificial intelligence in routine pathology practice. *Lab Invest*. 2024;104(9):102111. doi:10.1016/j.labinv.2024.102111

123. Moundas C. *CMS Finalizes New Electronic Prior Authorization Requirements For Payers And Providers. Mondaq Business Briefing*. NA-NA; 2024.

124. Khera R, Butte AJ, Berkwits M, et al. AI in medicine: JAMA's focus on clinical outcomes, patient-centered care, quality, and equity. *JAMA*. 2023;330(9):818-820. doi:10.1001/jama.2023.15481

125. Eaneff S, Obermeyer Z, Butte AJ. The case for algorithmic stewardship for artificial intelligence and machine learning technologies. *JAMA*. 2020;324(14):1397-1398. doi:10.1001/jama.2020.9371

126. Jain SS, Mello MM, Shah NH. Avoiding financial toxicity for patients from clinicians' use of AI. *N Engl J Med*. 2024;391(13):1171-1173. doi:10.1056/NEJMp2406135

127. Youssef A, Pencina M, Thakur A, Zhu T, Clifton D, Shah NH. External validation of AI models in health should be replaced with recurring local validation. *Nat Med*. 2023;29(11):2686-2687. doi:10.1038/s41591-023-02540-z

128. Hernandez-Boussard T, Bozkurt S, Ioannidis JPA, Shah NH. MINIMAR (minimum information for medical ai reporting): developing reporting standards for artificial intelligence in health care. *J Am Med Inform Assoc*. 2020;27(12):2011-2015. doi:10.1093/jamia/ocaa088

129. You JG, Hernandez-Boussard T, Pfeffer MA, Landman A, Mishuris RG. Clinical trials informed framework for real world clinical implementation and deployment of artificial intelligence applications. *NPJ Digit Med*. 2025;8(1):107. doi:10.1038/s41746-025-01506-4

130. World Health Organization. *Ethics and Governance of Artificial Intelligence for Health*. World Health Organization; 2021.

131. Labkoff S, Oladimeji B, Kannry J, et al. Toward a responsible future: recommendations for AI-enabled clinical decision support. *J Am Med Inform Assoc*. 2024;31(11):2730-2739. doi:10.1093/jamia/ocae209

132. Li F, Ruijs N, Lu Y. Ethics & AI: a systematic review on ethical concerns and related strategies for designing with AI in healthcare. *AI*. 2022;4(1):28-53. doi:10.3390/ai4010003

133. Garba-Sani Z, Farinacci-Roberts C, Essien A, Yracheta JM. *ACCESS AI: A New Framework for Advancing Health Equity in Health Care AI*. Health Affairs Forefront; 2024.

134. Dankwa-Mullan I, Scheufele EL, Matheny ME, et al. A proposed framework on integrating health equity and racial justice into the artificial intelligence development lifecycle. *J Health Care Poor Underserved*. 2021;32(2):300-317. doi:10.1353/hpu.2021.0065

135. Shah NH, Halamka JD, Saria S, et al. A nationwide network of health AI assurance laboratories. *JAMA*. 2024;331(3):245-249. doi:10.1001/jama.2023.26930

136. Embí PJ, Rhew DC, Peterson ED, Pencina MJ. Launching the trustworthy and responsible AI network (TRAIN): a consortium to facilitate safe and effective AI adoption. *JAMA*. 2025;333(17):1481-1482. doi:10.1001/jama.2025.1331

137. Matheny M, Israni ST, Ahmed M, Whicher D. *Artificial Intelligence in Health Care: The Hope, the Hype, the Promise, the Peril*. National Academy of Medicine; 2019:10. doi:10.17226/27111

138. Health AI. Health AI partnership is a multi-stakeholder collaborative. Accessed June 29, 2025. https://healthaipartnership.org/about-ai-health-partnership

139. Matheny ME, Goldsack JC, Saria S, et al. Artificial intelligence in health and health care: priorities for action. *Health Aff (Millwood)*. 2025;44(2):163-170. doi:10.1377/hlthaff.2024.01003

140. National Nurses United. National Nurses United survey finds A.I. technology degrades and undermines patient safety. Accessed June 29, 2025. https://www.nationalnursesunited.org/press/national-nurses-united-survey-finds-ai-technology-undermines-patient-safety

141. Ng MY, Kapur S, Blizinsky KD, Hernandez-Boussard T. The AI life cycle: a holistic approach to creating ethical AI for health decisions. *Nat Med*. 2022;28(11):2247-2249. doi:10.1038/s41591-022-01993-y

142. Sidebottom R, Lyburn I, Brady M, Vinnicombe S. Fair shares: building and benefiting from healthcare AI with mutually beneficial structures and development partnerships. *Br J Cancer*. 2021;125(9):1181-1184. doi:10.1038/s41416-021-01454-2

143. US Food and Drug Administration. Total product lifecycle considerations for generative AI enabled devices. Accessed June 29, 2025. https://www.fda.gov/media/182871/download

144. Hernandez-Boussard T, Lee AY, Stoyanovich J, Biven L. Promoting transparency in AI for biomedical and behavioral research. *Nat Med*. 2025;31(6):1733-1734. doi:10.1038/s41591-025-03680-0

145. Lu JH, Callahan A, Patel BS, et al. Assessment of adherence to reporting guidelines by commonly used clinical prediction models from a single vendor: a systematic review. *JAMA Netw Open*. 2022;5(8):e2227779. doi:10.1001/jamanetworkopen.2022.27779

146. Kwong JCC, Khondker A, Lajkosz K, et al. APPRAISE-AI tool for quantitative evaluation of AI studies for clinical decision support. *JAMA Netw Open*. 2023;6(9):e2335377. doi:10.1001/jamanetworkopen.2023.35377

147. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK; SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Lancet Digit Health*. 2020;2(10):e537-e548. doi:10.1016/S2589-7500(20)30218-1

148. Vasey B, Nagendran M, Campbell B, et al. Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *BMJ*. 2022;377(5):924-933. doi:10.1038/s41591-022-01772-9

149. Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ; SPIRIT-AI and CONSORT-AI Working Group. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Lancet Digit Health*. 2020;2(10):e549-e560. doi:10.1016/S2589-7500(20)30219-3

150. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. 2024;385:e078378. doi:10.1136/bmj-2023-078378

151. Mitchell M, Wu S, Zaldivar A, et al. Model cards for model reporting. presented at: Proceedings of the Conference on Fairness, Accountability, and Transparency; 2019; Atlanta, GA. doi:10.1145/3287560.3287596

152. Stoyanovich J, Howe B. Nutritional labels for data and models. *Q Bull Comput Soc IEEE Tech Comm Data Eng*. 2019;42(3).

153. Fox A. Epic leads new effort to democratize health AI validation. May 28, 2024. Accessed June 29, 2025. https://www.healthcareitnews.com/news/epic-leads-new-effort-democratize-health-ai-validation

154. Embi PJ. Algorithmovigilance: advancing methods to analyze and monitor artificial intelligence-driven health care for effectiveness and equity. *JAMA Netw Open*. 2021;4(4):e214622-e214622. doi:10.1001/jamanetworkopen.2021.4622

155. The Joint Commission. The Joint Commission and Coalition for Health AI join forces to scale the responsible use of AI in delivering better healthcare. June 11, 2025. Accessed June 29, 2025. https://www.jointcommission.org/resources/news-and-multimedia/news/2025/06/the-joint-commission-and-coalition-for-health-ai-join-forces/

156. Fleisher LA, Economou-Zavlanos NJ. Artificial intelligence can be regulated using current patient safety procedures and infrastructure in hospitals. *JAMA Health Forum*. 2024;5(6):e241369-e241369. doi:10.1001/jamahealthforum.2024.1369

157. Desai RJ, Matheny ME, Johnson K, et al. Broadening the reach of the FDA Sentinel system: a roadmap for integrating electronic health record data in a causal analysis framework. *NPJ Digit Med*. 2021;4(1):170. doi:10.1038/s41746-021-00542-0

158. Mandl KD, Gottlieb D, Mandel JC. Integration of AI in healthcare requires an interoperable digital data ecosystem. *Nat Med*. 2024;30(3):631-634. doi:10.1038/s41591-023-02783-w

159. Shah NH, Entwistle D, Pfeffer MA. Creation and adoption of large language models in medicine. *JAMA*. 2023;330(9):866-869. doi:10.1001/jama.2023.14217

160. Beavins E. CHAI embarks on post-deployment monitoring for AI as FDA oversight lags. Accessed July 15, 2025. https://www.fiercehealthcare.com/ai-and-machine-learning/chai-embarks-post-deployment-monitoring-ai

161. Platt R, Brown JS, Robb M, et al. The FDA Sentinel Initiative—an evolving national resource.

N Engl J Med. 2018;379(22):2091-2093. doi:10.1056/NEJMp1809643

162. Brown JS, Maro JC, Nguyen M, Ball R. Using and improving distributed data networks to generate actionable evidence: the case of real-world outcomes in the Food and Drug Administration's Sentinel system. *J Am Med Inform Assoc*. 2020;27(5):793-797. doi:10.1093/jamia/ocaa028

163. El-Kareh R, Brenner DA, Longhurst CA. Developing a highly-reliable learning health system. *Learn Health Syst*. 2022;7(3):e10351. doi:10.1002/lrh2.10351

164. McDonald PL, Foley TJ, Verheij R, et al. Data to knowledge to improvement: creating the learning health system. *BMJ*. 2024;384:e076175. doi:10.1136/bmj-2023-076175

165. Franklin JB, Marra C, Abebe KZ, et al; JAMA Summit on Clinical Trials Participants. Modernizing the data infrastructure for clinical research to meet evolving demands for evidence. *JAMA*. 2024;332(16):1378-1385. doi:10.1001/jama.2024.0268

166. Weeks WB, Spelhaug J, Weinstein JN, Ferres JML. Bridging the rural-urban divide: an implementation plan for leveraging technology and artificial intelligence to improve health and economic outcomes in rural America. *J Rural Health*. 2024;40(4):762-765. doi:10.1111/jrh.12836

167. Abbasi AB, Layden J, Gordon W, et al. A unified approach to health data exchange: a report from the US DHHS. *JAMA*. 2025;333(12):1074-1079. doi:10.1001/jama.2025.0068

168. Collins FS, Schwetz TA, Tabak LA, Lander ES. ARPA-H: accelerating biomedical breakthroughs. *Science*. 2021;373(6551):165-167. doi:10.1126/science.abj8547

169. ARPA. ARPA-H announces site selections by launching nationwide health innovation network. September, 2023. Accessed June 29, 2025. https://arpa-h.gov/news-and-events/arpa-h-launches-nationwide-health-innovation-network

170. ARPA. PRECISE-AI: performance and reliability evaluation for continuous modifications and useability of artificial intelligence. Accessed June 29, 2025. https://arpa-h.gov/explore-funding/programs/precise-ai

171. Mandl KD. Unseen commercial forces could undermine artificial intelligence decision support. *NEJM*. 2025;2(3):AIp2400922. doi:10.1056/AIp2400922

172. Halamka JD, Tripathi M. The HITECH era in retrospect. *N Engl J Med*. 2017;377(10):907-909. doi:10.1056/NEJMp1709851

173. Car J, Topol EJ. Advocating for a master of digital health degree. *JAMA*. 2025;333(9):753-754. doi:10.1001/jama.2024.27365

174. Petersson L, Larsson I, Nygren JM, et al. Challenges to implementing artificial intelligence in healthcare: a qualitative interview study with healthcare leaders in Sweden. *BMC Health Serv Res*. 2022;22(1):850. doi:10.1186/s12913-022-08215-8

175. Ötleş E, James CA, Lomis KD, Woolliscroft JO. Teaching artificial intelligence as a fundamental toolset of medicine. *Cell Rep Med*. 2022;3(12):100824. doi:10.1016/j.xcrm.2022.100824

176. Russell RG, Lovett Novak L, Patel M, et al. Competencies for the use of artificial intelligence–based tools by health care professionals. *Acad Med*. 2023;98(3):348-356. doi:10.1097/ACM.0000000000004963

177. Triola MM, Rodman A. Integrating generative artificial intelligence into medical education: curriculum, policy, and governance strategies. *Acad Med*. 2024;10:1097.

178. Car J, Ong QC, Erlikh Fox T, et al; Digital Health Systems Collaborative. The digital health competencies in medical education framework: an international consensus statement based on a Delphi study. *JAMA Netw Open*. 2025;8(1):e2453131-e2453131. doi:10.1001/jamanetworkopen.2024.53131

179. Sinskey JL, Margolis RD, Vinson AE. The wicked problem of physician well-being. *Anesthesiol Clin*. 2022;40(2):213-223. doi:10.1016/j.anclin.2022.01.001

180. Shachar C, Cadario R, Cohen IG, Morewedge CK. HIPAA is a misunderstood and inadequate tool for protecting medical data. *Nat Med*. 2023;29(8):1900-1902. doi:10.1038/s41591-023-02355-y

181. Liddell K, Simon DA, Lucassen A. Patient data ownership: who owns your health? *J Law Biosci*. 2021;8(2):lsab023. doi:10.1093/jlb/lsab023

182. Mikk KA, Sleeper HA, Topol EJ. The pathway to patient data ownership and better health. *JAMA*. 2017;318(15):1433-1434. doi:10.1001/jama.2017.12145

183. Assistant Secretary for Technology Policy. HIPAA versus state laws. Accessed June 29, 2025. https://www.healthit.gov/topic/hipaa-versus-state-laws

184. Vest JR. Geography of community health information organization activity in the United States: implications for the effectiveness of health information exchange. *Health Care Manage Rev*. 2017;42(2):132-141. doi:10.1097/HMR.0000000000000103

185. Mandl KD, Perakslis ED. HIPAA and the leak of "deidentified" EHR data. *N Engl J Med*. 2021;384(23):2171-2173. doi:10.1056/NEJMp2102616

186. Cassel C, Bindman A. Risk, benefit, and fairness in a big data world. *JAMA*. 2019;322(2):105-106. doi:10.1001/jama.2019.9523

187. European Parliament. What is GDPR, the EU's new data protection law? Accessed June 29, 2025. https://gdpr.eu/what-is-gdpr/

188. Mello MM, Shah NH, Char DS. President Biden's executive order on artificial intelligence: implications for health care organizations. *JAMA*. 2024;331(1):17-18. doi:10.1001/jama.2023.25051

189. Smith M, Sattler A, Hong G, Lin S. From code to bedside: implementing artificial intelligence using quality improvement methods. *J Gen Intern Med*. 2021;36(4):1061-1066. doi:10.1007/s11606-020-06394-w

190. Feng J, Phillips RV, Malenica I, et al. Clinical artificial intelligence quality improvement: towards continual monitoring and updating of AI algorithms in healthcare. *NPJ Digit Med*. 2022;5(1):66. doi:10.1038/s41746-022-00611-y

191. Whicher D, Kass N, Saghai Y, Faden R, Tunis S, Pronovost P. The views of quality improvement professionals and comparative effectiveness researchers on ethics, IRBs, and oversight. *J Empir Res Hum Res Ethics*. 2015;10(2):132-144. doi:10.1177/1556264615571558

192. Faden RR, Kass NE, Goodman SN, Pronovost P, Tunis S, Beauchamp TL. An ethics framework for a learning health care system: a departure from traditional research ethics and clinical ethics. *Hastings Cent Rep*. 2013;(Spec No):S16-S27. doi:10.1002/hast.134

193. Kass NE, Faden RR, Angus DC, Morain SR. Making the ethical oversight of all clinical trials fit for purpose. *JAMA*. 2025;333(1):75-80. doi:10.1001/jama.2024.0269

194. Kass NE, Faden RR. Ethics and learning health care: the essential roles of engagement, transparency, and accountability. *Learn Health Syst*. 2018;2(4):e10066. doi:10.1002/lrh2.10066

195. Spector-Bagdady K, Jagsi R. Big data, ethics, and regulations: Implications for consent in the learning health system. *Med Phys*. 2018;45(10):e845-e847. doi:10.1002/mp.12707

196. Lantos JD. The regulation of clinical research: what's love got to do with it? *Learn Health Syst*. 2017;2(1):e10048. doi:10.1002/lrh2.10048

197. Johnson KB, Horn IB, Horvitz E. Pursuing equity with artificial intelligence in health care. *JAMA Health Forum*. 2025;6(1):e245031. doi:10.1001/jamahealthforum.2024.5031

198. Price WN II, Gerke S, Cohen IG. Potential liability for physicians using artificial intelligence. *JAMA*. 2019;322(18):1765-1766. doi:10.1001/jama.2019.15064

199. Mello MM, Guha N. Understanding liability risk from using health care artificial intelligence tools. *N Engl J Med*. 2024;390(3):271-278. doi:10.1056/NEJMhle2308901

200. Mello MM, Guha N. Understanding liability risk from using health care artificial intelligence tools. *N Engl J Med*. 2024;390(3):271-278. doi:10.1056/NEJMhle2308901