Annals of Internal Medicine

POSITION PAPER

Artificial Intelligence in the Provision of Health Care: An American College of Physicians Policy Position Paper

Nadia Daneshvar, JD, MPH; Deepti Pandita, MD; Shari Erickson, MPH; Lois Snyder Sulmasy, JD; and Matthew DeCamp, MD, PhD; for the ACP Medical Informatics Committee and the Ethics, Professionalism and Human Rights Committee*

Internal medicine physicians are increasingly interacting with systems that implement artificial intelligence (AI) and machine learning (ML) technologies. Some physicians and health care systems are even developing their own AI models, both within and outside of electronic health record (EHR) systems. These technologies have various applications throughout the provision of health care, such as clinical documentation, diagnostic image processing, and clinical decision support. With the growing availability of vast amounts of patient data and unprecedented levels of clinician burnout, the proliferation of these technologies is cautiously welcomed by some physicians. Others think it presents challenges to the patient–physician relationship and the professional integrity of physicians. These dispositions are understandable, given the "black box" nature of some AI models, for which specifications and development methods can be closely guarded or proprietary, along with the relative lagging or absence of appropriate regulatory scrutiny and validation. This American College of Physicians (ACP) position paper describes the College's foundational positions and recommendations regarding the use of AI- and ML-enabled tools and systems in the provision of health care. Many of the College's positions and recommendations, such as those related to patient-centeredness, privacy, and transparency, are founded on principles in the ACP Ethics Manual. They are also derived from considerations for the clinical safety and effectiveness of the tools as well as their potential consequences regarding health disparities. The College calls for more research on the clinical and ethical implications of these technologies and their effects on patient health and well-being.

Ann Intern Med. doi:10.7326/M24-0146 **Annals.org**
For author, article, and disclosure information, see end of text.
This article was published at Annals.org on 4 June 2024.

The applications of artificial intelligence (AI) and machine learning (ML) (see the Glossary) in medicine have expanded steadily since the 1970s and continue to grow at a rapid rate. From January 2020 to October 2023, the U.S. Food and Drug Administration (FDA) reviewed, approved, authorized, or cleared more AI- and ML-enabled tools than it had in the preceding 25 years. Since November 2022, interest in AI has grown significantly alongside the rise of generative AI tools, such as OpenAI's ChatGPT, and the corresponding increase in mainstream media coverage.

The health care industry has been particularly excited about AI technology and what it may mean for the future of medicine and health care delivery. The amount of data that continues to be compiled about persons through various consumer- and patient-facing digital health applications is impossible for even the most astute of physicians to sift through and process, let alone apply to clinical decisions. With the worsening national shortage of clinicians and record levels of physician burnout, there is growing enthusiasm about the expansion of seemingly omniscient tools that guide health care practitioners in clinical decision making and assist with common sources of administrative burden. Furthermore, it is expected that AI technologies, which can process vast amounts of patient data from various sources to inform medical decisions, will enable more personalized, data-driven patient care. The expected benefits for patient-centered care and decision making are among the reasons that AI-enabled tools and systems may not only be expected but required in the future of medicine and health care.

Although data-driven care is a cornerstone of modern medicine, data-driven decision making can be complicated and fraught with error (1). Similarly, although AI tools can transform the practice of medicine in many beneficial ways, clinical decision support based on AI output without a basic understanding of AI technology can have serious, even fatal consequences for patients (2). Therefore, it is important to note that when being

* This paper, authored by Nadia Daneshvar, JD, MPH; Deepti Pandita, MD; Shari Erickson, MPH; Lois Snyder Sulmasy, JD; and Matthew DeCamp, MD, PhD, was developed for the ACP Medical Informatics Committee and the Ethics, Professionalism and Human Rights Committee. Individuals who served on the ACP Medical Informatics Committee from initiation of the project until its approval were Parag H. Mehta†, Jitendra Barmecha†, Elisa I. Choi†, Ellen J. Gelles†, Megan D. Hiles†, Ross W. Hilliard†, Isaure Hostetter†, Sharath Kharidi†, Lisa Rotenstein†, Matthew Sakumoto†, Lipika Samal†, and John J. Whyte†. Individuals who served on the ACP Ethics, Professionalism and Human Rights Committee from initiation of the project until its approval were Jan K. Carney†, Alejandro Moreno†, Omar T. Atiq†, Eileen D. Barrett†, Eduardo Brueta†, Angelique N. Collamer†, Charlene M. Dewey†, Kari L. Esbensen†, Amy K. Holbrook†, Diana Jung†, Kyle E. Karches†, Kenneth Prager†, Ikaasa Suri†, and Erik A. Wallace†. Approved by the ACP Board of Regents on 20 February 2024.
† Nonauthor contributor.

used for clinical decision making, the more appropriate term is "augmented" intelligence, meaning that it continues to incorporate human intelligence and is used as a tool to assist clinicians (3, 4). Extensive research is necessary to assess the short- and long-term risks and effects of the clinical use of AI on quality of care, health disparities, patient safety, health care costs, administrative burden, and physician well-being and burnout. It is critical to increase overall awareness of the clinical risks and ethical implications of using AI, including any measures that can be taken to mitigate the risks. Comprehensive educational resources are necessary to help clinicians, both in practice and in training, navigate this rapidly evolving area of technology, including improving their collective understanding of where the technology may be integrated in systems they already use and recognizing its implications.

Along with best practices, research, regulatory guidance, and oversight are needed to ensure the safe, effective, and ethical use of these technologies. This executive summary provides a synopsis of the American College of Physicians' (ACP) policy positions on the use of AI in the provision of health care. The full background, rationale, and policy recommendations can be found in Appendix 1 (available at Annals.org). The recommendations in this paper are intended to inform the College's advocacy regarding both predictive and generative AI policies. However, although we see great potential in generative AI benefiting both physicians and patients, because the landscape for this subset of AI technology is still evolving, it is too early to comment on its full scope and implications. The ACP will continue to consider this evolving technology as it matures. In addition, ACP recognizes that there may be challenges to implementing some of these recommendations due to the dynamic and evolving nature of AI technology.

Methods

The ACP Medical Informatics Committee, which addresses issues related to medical informatics and health information technology (IT), and the ACP Ethics, Professionalism and Human Rights Committee developed this position paper on the basis of a review of laws and regulations, ethical principles, and empirical studies and a broad review of literature about the uses of AI in medicine and health care. Databases searched included Google as well as Google Scholar, and search terms used included *artificial intelligence*, *internal medicine*, and *primary care*, among others. The authors also considered input from ACP's Board of Governors, Board of Regents, Council of Early Career Physicians, Council of Resident/Fellow Members, Council of Student Members, and Council of Subspecialty Societies in the development of its recommendations. This position paper and related recommendations were reviewed and approved by the ACP Medical Informatics Committee in January 2024, the ACP Ethics, Professionalism and Human Rights Committee in January 2024, and the Board of Regents in February 2024.

Recommendations/Position Statements

1. *ACP firmly believes that AI-enabled technologies should complement and not supplant the logic and decision making of physicians and other clinicians.*
2. *ACP believes that the development, testing, and use of AI in health care must be aligned with principles of medical ethics, serving to enhance patient care, clinical decision making, the patient–physician relationship, and health care equity and justice.*
3. *ACP reaffirms its call for transparency in the development, testing, and use of AI for patient care to promote trust in the patient–physician relationship. ACP recommends that patients, physicians, and other clinicians be made aware, when possible, that AI tools are likely being used in medical treatment and decision making.*
4. *ACP reaffirms that AI developers, implementers, and researchers should prioritize the privacy and confidentiality of patient and clinician data collected and used for AI model development and deployment.*
5. *ACP recommends that clinical safety and effectiveness, as well as health equity, must be a top priority for developers, implementers, researchers, and regulators of AI-enabled medical technology and that the use of AI in the provision of health care should be approached by using a continuous improvement process that includes a feedback mechanism. This necessarily includes end-user testing in diverse real-world clinical contexts, using real patient demographics, and peer-reviewed research. Special attention must be given to known and evolving risks that are associated with the use of AI in medicine.*
6. *ACP reaffirms that the use of AI and other emerging technologies in health care should reduce rather than exacerbate disparities in health and health care. To facilitate this effort:*
   a. *ACP calls for AI model development data to include data from diverse populations for which resulting models may be used.*
   b. *ACP calls on Congress, HHS, and other key entities to support and invest in research and analysis of data in AI systems to identify any disparate or discriminatory effects.*
   c. *ACP recommends that multisector collaborations occur between the federal government, industry, nonprofit organizations, academia, and others that prioritize research and development of ways to mitigate biases in any established or future algorithmic technology.*
7. *ACP recommends that developers of AI must be accountable for the performance of their models. There should be a coordinated federal AI strategy, built upon a unified governance framework. This strategy should involve governmental and nongovernmental regulatory entities to ensure:*

GLOSSARY

Algorithm: "A systematic procedure that produces—in a finite number of steps—the answer to a question or the solution of a problem" (5). "A set of rules that precisely define a sequence of operations" (6).

Artificial intelligence: "A collection of computer algorithms displaying aspects of human-like intelligence for solving specific tasks" (7).

Natural language processing: "A type of AI that refers to algorithms that employ computational linguistics to understand and organize human speech" (7).

Augmented intelligence: "A conceptualization of artificial intelligence that focuses on AI's assistive role, emphasizing that its design enhances human intelligence rather than replaces it" (4).

Machine learning: "A subset of AI that harnesses a family of statistical modeling approaches to automatically learn trends from the input data and improve the prediction of a target state" (7).

Neural network: "A computer program that operates in a manner inspired by the natural neural network in the brain" (8).

Deep learning: "A subset of ML consisting of multiple computational layers between the input and output that form a 'neural network' used for complex feature learning" (7).

Convolutional neural network: "A subset of DL techniques that is particularly efficient in AI-based pattern recognition. It is the foundation of many image processing AI algorithms, for instance in radiology" (7).

Chatbot: "A computer program that simulates human conversation with an end user" (9).

Clinical decision support software: "Software that is intended to provide decision support for the diagnosis, treatment, prevention, cure, or mitigation of diseases or other conditions" (10).

Computer vision: "Scientific field that deals with how computers process, evaluate, and interpret digital images or videos" (7).

Generative AI: "Deep-learning models that can generate high-quality text, images, and other content based on the data they were trained on" (11).

Health equity: "The attainment of the highest level of health for all people, where everyone has a fair and just opportunity to attain their optimal health regardless of race, ethnicity, disability, sexual orientation, gender identity, socioeconomic status, geography, preferred language, or other factors that affect access to care and health outcomes" (12).

Large language model: "Large language models use computational artificial intelligence algorithms to generate language that resembles that produced by humans" (13).

Risk: "The combination of the probability of occurrence of harm and the severity of the potential harm" (14).

Software: "Instructions that tell a computer what to do" (15).

   a. *the oversight of the development, deployment, and use of AI-enabled medical tools;*
   b. *the enforcement of existing and future AI-related policies and guidance; and*
   c. *mechanisms to enable and ensure the reporting of adverse events resulting from the use of AI.*

8. *ACP recommends that in all stages of development and use, AI tools should be designed to reduce physician and other clinician burden in support of patient care.*
9. *ACP recommends that training be provided at all levels of medical education to ensure that physicians have the knowledge and understanding necessary to practice in AI-enabled health care systems.*
10. *ACP recommends that the environmental impacts of AI and their mitigation should be studied and considered throughout the AI cycle.*

CONCLUSION

The expansion of AI and ML technologies in health care systems means that physicians are encountering new tools that they were not previously aware of or do not yet fully understand. To ensure maximum benefit and minimum harm to patients from these new technologies, and to ensure that they are used in alignment with the ethical responsibilities of physicians and the medical profession, more guidance, regulatory oversight, research, and education are needed for physicians, other clinicians, and health care systems. The ACP hopes that the recommendations in this position paper will inform the development of policies in this critical area of health IT.

From American College of Physicians, Washington, DC (N.D., S.E.); University of California Irvine Health, Laguna Niguel, California (D.P.); American College of Physicians, Philadelphia, Pennsylvania (L.S.S.); and University of Colorado Anschutz Medical Campus, Aurora, Colorado (M.D.).

Financial Support: Financial support for the development of this position paper came exclusively from the ACP operating budget.

Disclosures: Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M24-0146.

Corresponding Author: Nadia Daneshvar, JD, MPH, American College of Physicians, 25 Massachusetts Avenue NW, Suite 700, Washington, DC 20001; e-mail, ndaneshvar@acponline.org.

Author contributions are available at Annals.org.

References

1. **Obermeyer Z, Powers B, Vogeli C, et al.** Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366:447-453. [PMID: 31649194] doi:10.1126/science.aax2342
2. **Lyell D, Wang Y, Coiera E, et al.** More than algorithms: an analysis of safety events involving ML-enabled medical devices reported to the FDA. J Am Med Inform Assoc. 2023;30:1227-1236. [PMID: 37071804] doi:10.1093/jamia/ocad065
3. **McNemar E.** How does artificial intelligence compare to augmented intelligence? HealthITAnalytics. 24 September 2021. Accessed at https://healthitanalytics.com/news/how-does-artificial-intelligence-compare-to-augmented-intelligence on 10 January 2024.
4. **AMA.** Augmented intelligence in medicine. Accessed at www.ama-assn.org/practice-management/digital/augmented-intelligence-medicine on 14 February 2024.
5. **The Editors of Encyclopaedia Britannica.** Algorithm. Encyclopaedia Britannica. Accessed at www.britannica.com/science/algorithm on 11 January 2024.
6. **Stone HS.** Introduction to Computer Organization and Data Structures. McGraw-Hill; 1971.
7. **Adler-Milstein J, Aggarwal N, Ahmed M, et al.** Meeting the moment: addressing barriers and facilitating clinical adoption of artificial intelligence in medical diagnosis. NAM Perspect. 2022;2022. [PMID: 36713769] doi:10.31478/202209c
8. **Zwass V.** Neural network. Encyclopaedia Britannica. Accessed at www.britannica.com/technology/neural-network on 11 January 2024.
9. **IBM.** What is a chatbot? Accessed at www.ibm.com/topics/chatbots on 11 January 2024.
10. **U.S. Food and Drug Administration.** Clinical decision support software: guidance for industry and Food and Drug Administration staff. Accessed at www.fda.gov/media/109618/download on 20 July 2023.
11. **Martineau K.** What is generative AI? IBM. 20 April 2023. Accessed at https://research.ibm.com/blog/what-is-generative-AI on 20 January 2024.
12. **Centers for Medicare & Medicaid Services.** Health equity. Accessed at www.cms.gov/pillar/health-equity on 12 February 2024.
13. **Clusmann J, Kolbinger FR, Muti HS, et al.** The future landscape of large language models in medicine. Commun Med (Lond). 2023;3:141. [PMID: 37816837] doi:10.1038/s43856-023-00370-1
14. **U.S. Food and Drug Administration.** Applying human factors and usability engineering to medical devices: guidance for industry and Food and Drug Administration staff. Accessed at www.fda.gov/regulatory-information/search-fda-guidance-documents/applying-human-factors-and-usability-engineering-medical-devices on 10 August 2023.
15. **The Editors of Encyclopaedia Britannica.** Software. Encyclopaedia Britannica. Accessed at www.britannica.com/technology/software on 11 January 2024.

Author Contributions: Conception and design: N. Daneshvar, S. Erickson, D. Pandita, L. Snyder Sulmasy.
Analysis and interpretation of the data: N. Daneshvar, S. Erickson, D. Pandita, L. Snyder Sulmasy.
Drafting of the article: N. Daneshvar, M. DeCamp, S. Erickson, D. Pandita, L. Snyder Sulmasy.
Critical revision of the article for important intellectual content: N. Daneshvar, M. DeCamp, S. Erickson, D. Pandita, L. Snyder Sulmasy.
Final approval of the article: N. Daneshvar, M. DeCamp, S. Erickson, D. Pandita, L. Snyder Sulmasy.
Provision of study materials or patients: D. Pandita.
Administrative, technical, or logistic support: N. Daneshvar.
Collection and assembly of data: N. Daneshvar, D. Pandita.

Appendix 1: Background and Rationale

History

Artificial intelligence first emerged as a discipline in the 1950s when mathematicians and computer scientists collectively began exploring ways to create intelligent machines capable of emulating human reasoning and decision making (16). By the 1960s, a widely cited initial definition of AI as "the science of making machines do things that would require intelligence if done by men" emerged (17). Sets of programmed rules that computers processed and based outputs on (rule-based expert systems) were in development for use in biomedical contexts. One of these programs included the first chatbot or conversational software, ELIZA, which was trained to play the role of a psychotherapist. A concept ahead of its time considering today's large language models (LLMs) (see the Glossary) (for example, ChatGPT), the program was named after George Bernard Shaw's fictional character Eliza Doolittle, "to emphasize that it may be incrementally improved by its users, since its language abilities may be continually improved by a 'teacher'" (18–20).

By the 1970s, AI in medicine became a subject of formal scientific inquiry. Some of the first exploratory uses of computational methods in medicine included the development of consultation programs for assisting physicians with the diagnosis and treatment of bacterial infections (21, 22) and glaucoma (23). Early milestones included the development of INTERNIST-1, an experimental system to aid in internal medicine diagnoses (24), and DXplain (The General Hospital Corporation), a clinical decision support system initially released in the mid-1980s that is still available today (25). By 1995, the FDA had approved the first AI- and ML-enabled medical tool: an automated reader of cervical cytology slides (26).

Terminology

Many aspects of AI have been controversial since its inception, and even its nomenclature has been no exception. The moniker of AI has generated plenty of discontentment, with some prominent public figures even expressing displeasure with the term, arguing it is an inaccurate, perhaps even dangerous misnomer that can cause misconceptions and risks, such as through raising expectations, assumptions, and trust of the technology (27, 28).

Definitions

The term *artificial intelligence* can be better defined. Each regulatory agency uses its own terms and definitions for tools that can use algorithms and AI technology, such as clinical decision support software and predictive decision support interventions (DSIs) (definitions for these terms can be found in the Glossary and elsewhere in this discussion).

For the purposes of this paper, we use the term artificial intelligence to mean "a collection of computer algorithms displaying aspects of human-like intelligence for solving specific tasks" (7). In addition, ACP believes it is important to note that when being used for clinical decision making, the more appropriate term is *augmented* intelligence, as this emphasizes that its design should enhance human intelligence rather than replace it (4).

Current State

Most current applications of AI in health care implement predictive AI, which uses algorithms to analyze data and make predictions. Generative AI, on the other hand, uses algorithms to identify statistical patterns in large data sets and generates new content (for example, text, images, music, videos) on the basis of those patterns (29). Both predictive and generative AI use ML (see the Glossary), but generative AI can often use deep learning (see the Glossary) (11). The contrast between these 2 types of AI is expected to diminish with advances in AI technology, however (29). The statistical methods behind AI are not entirely new, nor is the idea of using model-based predictions (consider, for example, predictive scores like CHA_2DS_2-VASc). However, the rapid pace of AI development, its potential scope, and the fact that some AI models are so-called black boxes whose inner workings and predictive variables are unknown or unknowable suggest AI needs ongoing review.

Applications of natural language processing (NLP) (see the Glossary) have also expanded. This type of AI has been implemented in ambient clinical documentation technology (that is, "digital scribes") (30), for example, as well as in certain early prediction tools for sepsis (31). A subset of NLP, LLMs, use deep learning methods (particularly transformers, which are a subset of neural networks) to conduct NLP tasks (32).

Deep learning tools are also being applied in areas of medicine that rely heavily on diagnostic imaging and pattern recognition (for example, radiology, pathology, dermatology, and ophthalmology) (33). Convolutional neural networks (see the Glossary), which implement deep learning, enable the field of computer vision, which "deals with how computers process, evaluate, and interpret digital images or videos" (7). The image classification performance of these tools has been similar to that of physicians (34, 35).

Generative AI and LLMs

Generative AI uses LLMs in producing various forms of new content. The first free, widely accessible generative AI tool, ChatGPT ("GPT" is an acronym for "generative pretrained transformer"), was released to the public for testing in November 2022. Shortly thereafter, it surpassed Instagram as the fastest-growing consumer application

ever (36). ChatGPT-3, an updated version of the original ChatGPT LLM, was trained on more than 45 terabytes of data, including 175 billion parameters, from books, academic papers, and all of Wikipedia (36). Some technology leaders believe that the most powerful version of ChatGPT yet—GPT-4—is poised to revolutionize medicine (37). Although ChatGPT's developers have not disclosed the volume of parameters on which GPT-4 was trained, it is estimated that more than 1 trillion parameters were used (38).

In November 2023, the American Medical Association released updated principles for the development, deployment, and use of health care AI. The American Medical Association's principles focus on the need for oversight, when transparency and disclosure of AI use is recommended and what information should be included in disclosures, generative AI, physician liability for using AI, cybersecurity and data privacy, and the use of AI and automated decision-making systems by payors (39).

Regulatory Oversight of Health Care Industry AI Development and Use

U.S. Department of Health and Human Services. The federal agency charged with regulatory oversight of medical device products (for example, AI- and ML-enabled devices) is the FDA, which is a division of the U.S. Department of Health and Human Services (HHS). In addition to overseeing commercial AI products through the FDA, HHS has implemented AI in various contexts. After releasing an AI strategy in early 2021 (40), HHS released its Trustworthy AI Playbook in September 2021 to guide the use of AI within HHS (41). Between 2022 and 2023, HHS more than tripled its publicly reported AI use cases, which are in various stages of development ("development and acquisition," "initiation," "implementation," and "operation and maintenance"), to 163 use cases (42). Most of the use cases have been adopted by the National Institutes of Health ($n = 47$) and the FDA ($n = 44$), although many other agencies are using the technology for various purposes (42). See Appendix 2 for HHS AI use case examples.

FDA. In the 28 years from 1995 when the FDA approved its first medical AI tool until October 2023, the FDA reviewed, approved, authorized, or cleared 694 AI- and ML-enabled medical devices (43). Most of these (478 [69%]) were approved, authorized, or cleared after January 2020—an indication of the rapid proliferation of AI in medicine (43). Of these 694 devices, the vast majority (531 [77%]) have been radiology tools; most others have been for cardiovascular care (71 [10%]); and others have been for neurology, hematology, ophthalmology, and gastroenterology/urology, as well as other areas of medicine (43).

Office of the National Coordinator for Health IT. In November 2023, the Office of the National Coordinator for Health IT released the "Health Data, Technology, and Interoperability: Certification Program Updates, Algorithm Transparency, and Information Sharing (HTI-1) Final Rule," with provisions aimed at increasing the transparency of predictive DSIs, which the rule defines as "technology that supports decision-making based on algorithms or models that derive relationships from training data and then produce an output that results in prediction, classification, recommendation, evaluation, or analysis." The rule establishes new requirements for evidence-based DSIs and predictive DSIs supplied by health IT developers through their health IT modules.

Recommendations and Rationale

1. *ACP firmly believes that AI-enabled technologies should complement and not supplant the logic and decision making of physicians and other clinicians.*

Artificial intelligence– and ML-enabled tools hold great promise for informing, guiding, and improving patient care. In some instances, AI- and ML-enabled tools may even outperform clinicians. However, ACP believes that a physician's training and observations must remain the central tenet of patient care. Technology, especially AI, can be misguided through inappropriate, incomplete, or flawed data inputs, leading to flawed outputs, which can mislead physicians into ill-informed decision making with dangerous consequences for patients. Therefore, physicians must be careful not to place too much faith or decision-making power in AI- and ML-enabled tools.

Historically, however, humans have been prone to overreliance on automated systems (44), including those enabled by AI and ML. "Automation complacency" (45) and "automation bias," or the tendency to "trust algorithms without sufficient skepticism," are not uncommon among users of such systems (44). Furthermore, experts have been just as likely as beginners to overly rely on autonomous systems, and this overreliance is mostly unaffected by training (46). Even though skepticism is warranted, being overly skeptical of AI can have ramifications, too: It could mean that patients and physicians will not experience a benefit that AI could offer. A balanced approach to AI technologies is in order.

The language that physicians, technology developers, and society use to describe the qualities and features of AI tools has substantial implications for how society thinks about and uses AI. Some have cautioned against using anthropomorphic language in discussing AI and its applications, suggesting that the psychological effects of doing so can create human-centric expectations and assumptions and increase trust and reliance on systems (28, 47–49). As one way to mitigate the risks of overreliance and automation bias, ACP believes that everyone, especially persons in the fields of AI academia and health care, has a responsibility to avoid anthropomorphizing AI models and their functionality, including using appropriate AI terminology that does not ascribe human characteristics to AI models.

Anthropomorphic language can include terms such as *think*, *learn*, and *hallucinate*. Some have argued that describing certain AI outputs as "hallucinations," in addition to being inaccurate and anthropomorphic, is not ideal because of stigmatization (50). *Confabulation* may be a better term, although it may raise similar concerns (50, 51).

2. *ACP believes that the development, testing, and use of AI in health care must be aligned with principles of medical ethics, serving to enhance*

patient care, clinical decision making, the patient–physician relationship, and health care equity and justice.

The use of AI in health care must align with principles of medical ethics (52). Physicians as individuals and medicine as a profession have a duty to put patient interests first. Artificial intelligence technology, when appropriately used, should serve to enhance patient care; decision making based in patient values, interests, and preferences; and the patient–physician relationship. Physician duties articulated by the College in the ACP Ethics Manual are based on the ethical principles of *beneficence*—"the duty to promote good and act in the best interest of the patient"; *nonmaleficence*—"the duty to do no harm to the patient"; *respect for patient autonomy*—"the duty to protect and foster a patient's free, uncoerced choices" and the related duty of truth-telling; and *justice*—the equitable distribution of "the life-enhancing opportunities afforded by health care" (52).

Maintaining the patient–physician relationship requires care. Technology, including EHRs, can be used in ways that support or detract from this relationship (53). Artificial intelligence should be implemented in ways that do not harm or interfere with this relationship but instead enhance and promote the therapeutic alliance between patient and physician. Ambient clinical documentation (that is, audio technologies that transform natural patient–physician communication into a draft clinical note format), for example, could promote better communication and connection during a visit, freeing a physician from the distraction of staring at a computer screen. Other technologies, such as conversational agents or chatbots, have the potential to support chronic care management and patient access; however, if their use by patients becomes required or burdensome, they could interfere with patient–physician relationships (54).

Physicians should advocate for AI tools that support putting the patient–physician relationship first and, as a matter of respect for patient uniqueness and dignity, support informed decision making according to the patient's needs, values, and preferences. The promise of AI is that it can summarize data and information more quickly and comprehensively to inform decision making; the peril of AI is that it could exert too much influence on clinical decision making and be subject to hidden or unconscious biases or agendas, such as those of developers and or deployers (for example, if AI outputs exclude a clinically indicated option because of cost). At the individual patient level, this means ensuring AI tools do not replace or interfere with patient–physician decisions. At a system level, this means ensuring patients, physicians, and other clinicians are involved in the design, development, implementation, and testing of AI tools and systems throughout the entire product cycle (**Appendix Figure 1**), to advocate for solutions to the problems they face and to ensure that the tools do not exacerbate health disparities.

3. *ACP reaffirms its call for transparency in the development, testing, and use of AI for patient care to promote trust in the patient–physician relationship. ACP recommends that patients, physicians, and other clinicians be made aware, when possible, that AI tools are likely being used in medical treatment and decision making.*

Artificial intelligence transparency is important for patients as well as physicians and other clinicians. In a July 2021 position paper, "Health Information Privacy, Protection, and Use in the Expanding Digital Health Ecosystem," the College emphasized the need for transparency in AI use (55). Specifically, ACP called for increased transparency around the collection and use of personal health information (for example, "methods of deidentification, timelines, allowable disclosures, and when consent is needed") in AI development (55). The ACP has also advocated that clinical entities should develop clear policies that relate to the aggregation of data and their use and release for purposes other than direct care of the patient (for example, performance aggregation and reporting and research). We reaffirm these views, especially for the use of patient data to train AI models.

Even if patients are not, at present, explicitly informed of all the ways technology is involved in their care—for example, they may or may not be told about computer-assisted electrocardiogram or mammography interpretation—the newness of AI and its potential for clinically significant effects on care suggests that honesty and transparency about its use are paramount. At least 1 study has indicated that patients would prefer to have options when it comes to the use of AI, including the choice of refusing certain uses of AI in their care (56). Focus groups conducted in November 2019 that inquired about patient choice and autonomy found that "the preservation of choice" contributed to participants' comfort with the use of AI in their care, and that participants believed that "patients should have the right to choose to have an AI tool used in their care and be able to opt-out of AI involvement if they felt strongly" (56).

Other studies have found that patients have mixed views about the use of AI in their care. A December 2019 survey of patient perspectives of AI found that an overwhelming majority of respondents considered it important for them to be told when an AI program has played a big role in their diagnosis or treatment (887 [95.7%]; $n = 926$) (57). A smaller but still substantial majority of respondents also considered it important for them to be told when an AI program has played a small role in their diagnosis or treatment (801 [86.5%]) (57). The authors also found that "most respondents were very concerned or somewhat concerned about AI's unintended consequences, including misdiagnosis (91.5%), privacy breaches (70.8%), less time with clinicians (69.6%), and higher health care costs (68.4%). A higher proportion of respondents who self-identified as being members of racial and ethnic minority groups indicated being very concerned about these issues, compared with White respondents" (57).

Although we respect patient autonomy and encourage transparency with patients when AI is being used in their care, most physicians do not have a clear understanding of where AI and ML are already in use or integrated in their own health care systems or practices. Knowing where these technologies are being implemented in existing workflows and EHR systems can help physicians determine when it may be wise to question

software output and better gauge the risks of using AI-enabled systems on an individual patient. Recognizing these needs, a group of practicing physicians collaborated in creating "The Physicians' Charter for Responsible AI" to guide the development, testing, and use of AI tools in clinical practice (58). The Charter "stems from a growing concern [. . .] about the rapid pace of AI and how it will be implemented in healthcare." The College agrees with the authors that the views of practicing physicians are necessary in guiding the incorporation of AI into health care.

To promote patient safety and transparency, ACP calls for transparency, clarity, and education for physicians and other clinicians about where AI-generated information is entering workflows and ensuring that physician knowledge and discretion can supersede AI-generated "defaults," when appropriate. In addition, there must be readily available mechanisms for clinicians to call attention to, and for AI developers to correct, errors in AI-generated output or decision making. Physicians must seek to be aware of AI uses in their practice and exercise clinical and professional judgment in making appropriate disclosures, just as they routinely do in other aspects of care.

We recognize that the feasibility of transparency with patients regarding AI use and availability of opt-out mechanisms can be variable and dependent on the nature of the AI tools themselves; the systems in which they may be integrated; and perhaps most significantly, the level of disclosure or transparency, if any, with physicians. For instance, transparency with patients about the integration of AI into certain devices, such as a glucometer or insulin pump, and the possibility of opting out of using such a device may be reasonably feasible. However, transparency with patients about AI integration into EHR systems and other common sources of information for physicians (for example, search engines) may not be as feasible, especially given that physicians are often not made aware of the integration. Questions also remain about whether disclosure or the ability to opt-out is necessary or feasible when AI tools are used to assist with or complete tasks commonly associated with administrative burden. Recognizing that some AI tools may operate outside the patient–physician relationship, health systems may need to notify patients directly about those uses (for example, as part of consent processes). New approaches will likely be necessary to address whether and how to disclose AI use and obtain consent in specific circumstances.

In addition, to appropriately empower physicians with the tools they need to deliver health care safely and effectively, the College believes that the FDA and Office of the National Coordinator for Health Information Technology should engage patients, physicians, and others in health care in developing standard, easy-to-understand model labels and standardized reporting checklists that include indicators for model quality and performance, use criteria (for example, appropriate patient populations), and other important safety and effectiveness information. These labels must be updated as models are changed and updated. Training and testing data sources and attributes should be transparent and accessible to physicians, regulators, and auditors. Algorithmic impact assessments, audits, and post-market monitoring are also recommended.

4. *ACP reaffirms that AI developers, implementers, and researchers should prioritize the privacy and confidentiality of patient and clinician data collected and used for AI model development and deployment.*

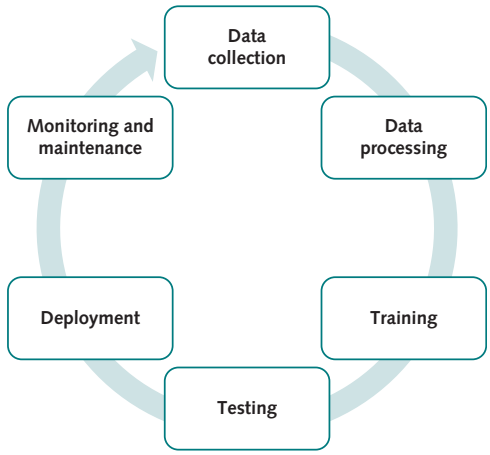
The ACP Ethics Manual (seventh edition) defines privacy as "freedom from unauthorized intrusion" and confidentiality as "a matter of respecting the privacy of patients, encouraging them to seek medical care and discuss their problems candidly, and preventing discrimination on the basis of their medical conditions" (52). In its July 2021 position paper, "Health Information Privacy, Protection, and Use in the Expanding Digital Health Ecosystem," the College emphasized the need for transparency in AI. The ACP also recommended federated learning as a potential solution to privacy issues involved with AI (55). Federated learning enables AI tools to be exposed to very large, geographically diverse data sets without the need for sharing sensitive clinical data (55). We reaffirm the critical need for privacy, confidentiality, and transparency in the use of patient and clinician data (for example, physician practice data, quality and safety metrics, and so forth) in all phases of AI cycles. If patient, physician, or other clinician data must be used for the development of AI models, the data should first be deidentified and aggregated. We note, however, that deidentification of data, particularly if the data is unstructured, can be a substantial challenge. We also renew our call for comprehensive federal privacy legislation, with special provisions regarding privacy protections for AI data sets included in such legislation.

5. *ACP recommends that clinical safety and effectiveness, as well as health equity, must be a top priority for developers, implementers, researchers, and regulators of AI-enabled medical technology and that the use of AI in the provision of health care should be approached by using a continuous improvement process that includes a feedback mechanism. This necessarily includes end-user testing in diverse real-world clinical contexts, using real patient demographics, and peer-reviewed research. Special attention must be given to known and evolving risks that are associated with the use of AI in medicine.*

Successful integration of AI in medicine will require the involvement of end-user physicians and clinicians, and when appropriate, patients, in the development and testing of AI tools and systems to ensure clinically appropriate AI uses and applications and to maximize the feasibility of AI adoption. Vendors, model developers, and health care systems should ensure that models are trained, maintained, and updated using the latest clinical practice guidelines. In addition, the safety, utility, and applicability of AI models are dependent on the quality and attributes of the data used for their development (59) (**Appendix Figure 2**). Flawed data can contribute to false-positive and false-negative results with clinically significant health and safety implications for patients (59).

Since the 1980s, ACP has advocated that efforts to assess new and emerging technologies should ensure that they are safe and effective before they become a part of common

Appendix Figure 1. The AI cycle.



A depiction of the cyclical and iterative nature of AI model development and implementation. AI = artificial intelligence.

medical practice. The ACP has further recommended that health IT systems be tailored to emphasize patient safety improvement (62). More recently, ACP noted the following:

> The movement of automated, AI-based systems into [diagnosis, therapy selection, and population health management] is a cause for concern by many physicians and others–specifically when considering care decisions regarding diagnosis and therapy selection [citation omitted]. There is justifiable concern that what may be initially presented as an assistant could easily become a risk to physician autonomy and a risk to patient safety. [. . .] These concerns must be addressed satisfactorily before these technologies are permitted to enter the clinical workflows, and more research on the potential effects of the use of AI, as well as any emerging technology, in clinical workflows is needed (63).

Furthermore, context- and system-specific data, model design, use, and testing are necessary to ensure clinical safety and effectiveness (64). Pragmatic randomized controlled trials are recommended to test AI models in real-world settings. In addition, ACP recommends that the use of AI in the provision of health care should be approached by using a continuous improvement process that includes a feedback mechanism or other feedback process, such as a Plan-Do-Study-Act cycle (65). Continuous improvement processes should be implemented with the understanding that they can contribute to model deterioration when model-informed decisions are introduced into the model (66). Caution is necessary to ensure that these improvement processes do not give rise to more risk than they are intended to mitigate.

Given the rapid evolution of AI-enabled tools, there will be unknown risks that will need to be managed as the field changes. Beyond intrinsic AI problems that may arise from the development stages of the AI cycle, there are also a range of postdeployment risks upon implementation due to human factors, including biases (67).

An analysis of adverse events involving FDA-approved ML devices that were submitted to the FDA's Manufacturer and User Facility Device Experience identified 266 events involving 25 devices (2). Consequences of events were categorized as "harm," "near miss events with potential to harm if not for intervention to prevent it," "hazards with potential to cause harm," "consequences for health care delivery without specific patient harm," "no consequences for health care delivery," and "complaints which generally describe the users [sic] experience but do not indicate harm, hazard, or systemic problems qualifying for other categories" (2). Although 14 of these events were voluntarily reported, almost all 266 events qualified as 1 of 3 kinds of events subject to mandatory adverse event reporting: malfunctions ($n = 238$), injuries ($n = 25$), or death ($n = 1$) (2) (**Appendix Figure 3**).

6. *ACP reaffirms that the use of AI and other emerging technologies in health care should reduce rather than exacerbate disparities in health and health care. To facilitate this effort:*
   a. *ACP calls for AI model development data to include data from diverse populations for which resulting models may be used.*
   b. *ACP calls on Congress, HHS, and other key entities to support and invest in research and analysis of data in AI systems to identify any disparate or discriminatory effects.*
   c. *ACP recommends that multisector collaborations occur between the federal government, industry, nonprofit organizations, academia, and others that prioritize research and development of ways to mitigate biases in any established or future algorithmic technology.*

As a matter of both equity and safety, data used to train AI models should be carefully selected and assessed for suitability for the intended populations, locations, and uses of the resulting AI model (**Appendix Figure 2**). Diverse populations should include those that are underrepresented, socially marginalized, and disadvantaged. The ACP previously advocated for efforts to ensure that use of new technologies like AI does not increase health care disparities, noting that "AI, ML, and other algorithmic technology, if not implemented with caution and appropriate regulations, can embed implicit biases into health care decision-making systems, which can in turn threaten patient health and quality of care" (68). The College's concerns have been warranted given the findings of studies confirming that biases have been perpetuated by clinical algorithms to the detriment of Black patients and other underserved or underrepresented populations (for example, female patients, Hispanic patients, and patients with Medicaid insurance) (1, 69). For example, studies have indicated that certain dermatology-related algorithms perform worse on darker skin tones than lighter ones (70, 71). As Thadaney-Israni and Verghese (72) have pointed out, "Flawed or incomplete data sets that are not inclusive can automate inequality." Other studies such as a December 2023 comparative effectiveness review from the Agency for Healthcare Research and Quality found that of 17 studies evaluating the effect of 18 algorithms on

Appendix Figure 2. The example of IBM's Watson.

A well-known example of AI failure is IBM's application of its Watson ML technology to health care. Watson was designed to mine vast data sets seeking patterns that could lead to better treatment (60). For example, Watson's Oncology Expert Advisor was programmed to process health records, medical publications, and physicians' notes to provide treatment recommendations in real time (60). However, the program was unable to "read" physicians' notes, nor could it apply the medical literature it mined to the specific circumstances of individual patients, often dispensing guidance that was inconsistent with local conditions or the nuances of a particular patient's case (60). Another product, Watson for Oncology, which was trained on a minimal number of synthetic patient cases with synthetic data, resulted in recommendations of "unsafe and incorrect" cancer treatments (61).

AI = artificial intelligence; ML = machine learning.

racial and ethnic disparities in health and health care, 11 studies identified 13 algorithms that may perpetuate or exacerbate racial and ethnic disparities, 5 studies identified 4 algorithms that may lessen disparities, and 1 study found that the algorithm studied likely had no effect on disparities (73). Therefore, research evaluating the effect of AI technology on the practice of medicine, patient access to care, and the quality and effectiveness of patient care, including assessments of whether AI use in medicine contributes to or drives biased or discriminatory health practices or inequitable health outcomes, is necessary.

Bias in AI is not only a data problem, it is also an ethical problem. How analyses are done—that is, which variables are chosen—is important. In a notable example, health care costs were used as a proxy for health care needs (1); such an assumption can result in bias when less money is spent on certain patient populations, despite their needs. In addition, biases can arise regarding when and how algorithms are deployed (that is, if they are used for certain patient groups and not others) (74, 75).

7. *ACP recommends that developers of AI must be accountable for the performance of their models. There should be a coordinated federal AI strategy, built upon a unified governance framework. This strategy should involve governmental and nongovernmental regulatory entities to ensure:*
   a. *the oversight of the development, deployment, and use of AI-enabled medical tools;*
   b. *the enforcement of existing and future AI-related policies and guidance; and*
   c. *mechanisms to enable and ensure the reporting of adverse events resulting from the use of AI.*

Independent, nongovernmental regulatory bodies, such as the Joint Commission, are necessary for oversight of AI models procured, developed, or deployed by governmental bodies. However, governmental bodies, such as the HHS Office of Inspector General, should also be involved in the oversight of AI-related technology used by governmental and nongovernmental entities. The College also supports recent calls for outcome-focused regulations (76) and for a public-private partnership to establish a "nationwide network of health AI assurance laboratories" (77). Regulatory oversight and performance monitoring are necessary throughout the entire tenure of the AI-enabled technology. If the use of AI-enabled technology is suspected to have contributed to an adverse event, the event and relevant details should be reported to the appropriate regulatory bodies, such as the FDA, and logged in public databases (for example, the FDA Manufacturer and User Facility Device Experience database). The College also supports proposals for HHS Office for Civil Rights enforcement against violations of section 1557 of the Affordable Care Act (which prohibits discrimination by covered health programs and activities based on race, sex, color, national origin, age, or disability) for the use of clinical algorithms in discriminatory ways (78, 79).

8. *ACP recommends that in all stages of development and use, AI tools should be designed to reduce physician and other clinician burden in support of patient care.*

Reducing unnecessary time, administrative, cognitive, and other burdens should be priorities in the design and development of AI-enabled devices to allow physicians to better care for patients. Artificial intelligence tools can be used to reduce administrative burden by performing patient intake, scheduling, and prior authorization functions, for example, and can decrease cognitive burden, for instance, by helping physicians get to the right diagnoses and treatments faster. Artificial intelligence should support more time for direct patient care by physicians and other clinicians. Artificial intelligence products should be sufficiently tested (that is, used by actual end users under real-world circumstances and all intended use contexts) before deployment to ensure usability and to identify and address problems and technologic burdens that may arise for clinicians and other members of the care team. Any mechanisms for clini-

Appendix Figure 3. Example of death associated with ML device.

The event that resulted in death was described by Lyell and colleagues (2) as follows:

"Insulin dosing software – An anonymous voluntary report expressed concerns over the aggressiveness with which an insulin dosing system treated hyperglycemia, with rates of change in blood sugar levels double that of other hospitals. Such rapid changes were described as causing patients to develop metabolic and EKG changes, leading to patients requiring intubation and 'several unfortunate outcomes including one patient death.' The report did not detail specific events and these events could not be confirmed by the manufacturer."

EKG = electrocardiogram; ML = machine learning.

cians to provide feedback on the performance of or any issues with the AI tool should not be burdensome to the clinician. The effects of AI-enabled burden reduction tools on burnout should be assessed.

New payment initiatives, especially those for value-based care, must support the use of AI technology as a mechanism to reduce burden and ideally improve quality. The ACP has previously advocated that "payment policies should create incentives for physicians and other health professionals to use health information technologies that have the functions and capabilities needed to improve clinical decision-making at the point of care, including functions designed to support care consistent with evidence-based guidelines, care coordination, and preventive and patient-centered care" (80). The ACP has also called for all involved parties to "support the development, adoption and use of innovative technologies that seamlessly enable enhanced and coordinated patient-centered care" (63). Artificial intelligence–enabled tools have the potential to improve the quality of patient care and reduce health care costs, thereby promoting value-based care—and it is critical that they take on this role, rather than contribute to clinician burden.

9. ACP recommends that training be provided at all levels of medical education to ensure that physicians have the knowledge and understanding necessary to practice in AI-enabled health care systems.

To enable safe deployment, comprehensive educational training programs and resources are needed at the undergraduate medical education, graduate medical education, and attending physician levels to address the knowledge gaps of current health care professionals. Education and training regarding the foundational concepts of AI; the ethics; clinically effective and appropriate uses of AI in medicine; and the risks and unintended consequences of AI use, including its effect on health disparities, should be incorporated into educational materials.

Physicians are far less likely to use AI tools if they do not understand, or trust, the output of AI systems. Therefore, to increase and improve AI use and usefulness, the creation and dissemination of clear and comprehensive educational materials to clinicians and end users of AI is crucial. Training should ensure that physicians remain able to make appropriate clinical decisions independently, in the absence of AI decision support should such technology become unavailable and more importantly, for vigilance against errors in AI generated or guided decisions to protect patient safety.

10. ACP recommends that the environmental impacts of AI and their mitigation should be studied and considered throughout the AI cycle.

Environmental health is defined as the health effects associated with environmental factors, such as air pollution, water contamination, and climate change. The College has called for the health sector to adopt environmentally sustainable and energy-efficient practices to aggressively reduce its greenhouse gas emissions (81), particularly given that it accounts for 8.5% of U.S. greenhouse gas emissions overall (82). Although estimates of health care technology's toll on greenhouse gas emissions are limited, we know that substantial computational resources involving immense amounts of data and energy consumption are required for the development of AI models (83). Well before the widespread proliferation of generative AI, it was reported that training just 1 AI model for NLP can emit 284 019.13 kilograms of carbon dioxide equivalent—nearly 5 times the lifetime emissions of the average American car (84). The carbon footprint of generative AI tools like LLMs is likely much higher (85, 86).

However, the utility of AI for climate change mitigation and the extent of the technology's inherent contribution to climate change are still being explored and may be up for debate (87). Given the divergent commentary and findings of existing studies about the relationship between AI and greenhouse gas emissions (87, 88), we believe that efforts to quantify the effect of AI on climate are necessary but also that the dearth of standardized measures impedes our ability to address the potentially negative climate effects of AI (83, 89, 90).

Conclusion

Advances in AI and ML tools have presented physicians and health care systems with various opportunities and challenges. Despite the historical use of advanced statistical concepts in the provision of health care, a new generation of technologic tools that incorporate modern AI and ML technologies warrant caution as well as optimism. The widespread clinical health and safety implications of these advanced AI and ML tools are not yet fully understood by physicians, and the totality of risks for patients have yet to be identified. Along with best practices, research, regulatory guidance, and oversight are needed to ensure the safe, effective, and ethical use of these technologies. The ACP hopes that the recommendations in this position paper will inform the development of guidance and policies related to the use of AI and ML tools in the provision of health care.

Appendix 2: HHS AI Use Cases

Use cases at the National Institutes of Health include "detecting clinicians' attitudes [emotions, biases, and burnout] through clinical notes" (National Library of Medicine) and assisting adjudicators in "identifying evidence on function" from lengthy case records in the Social Security Administration disability determination process (Clinical Center) (42). Use cases at the FDA, specifically at the National Center for Toxicological Research, involve detecting sex disparities in opioid drug safety signals using the FDA Adverse Events Reporting System Public Dashboard and the platform X (formerly known as Twitter) to determine whether X data can serve as an early warning system for opioid issues specific to women; guiding the selection of drugs for treating COVID-19 by mining publicly available adverse drug event data ("including the agency's database, public databases, and social media data"); and "assessing and mitigating bias in drug labeling documents" (42). Other use cases include detecting, validating, and reporting biologic-related adverse events from EHRs (FDA Center for Biologics Evaluation and Research, Office of Biostatistics

and Pharmacovigilance); detecting stimulant and opioid misuse and illicit use from clinical notes (Centers for Disease Control and Prevention National Center for Health Statistics); and program integrity efforts (for example, fraud and abuse detection) (Centers for Medicare & Medicaid Services) (42). Use cases mostly involve AI systems that were contracted for development ($n = 77$), although others are developed by HHS ("in-house") ($n = 53$), commercial off-the-shelf ($n = 2$), or involved unspecified developers ($n = 31$) (42).

Web References

16. **Turkle S.** The new philosophers of artificial intelligence. In: Turkle S, ed. The Second Self: Computers and the Human Spirit. MIT Pr; 2005:219-244.
17. **Minsky M.** Semantic Information Processing, MIT Pr; 1968.
18. **Weizenbaum J.** ELIZA—a computer program for the study of natural language communication between man and machine. Communications of the ACM. 1966;9:36-45.
19. **Turkle S.** The Second Self: Computers and the Human Spirit. MIT Pr; 2005.
20. **Natale S.** The ELIZA effect: Joseph Weizenbaum and the emergence of chatbots. In: Natale S, ed. Deceitful Media: Artificial Intelligence and Social Life after the Turing Test. Oxford Univ Pr; 2021.
21. **Shortliffe EH, Axline SG, Buchanan BG, et al.** An artificial intelligence program to advise physicians regarding antimicrobial therapy. Comput Biomed Res. 1973;6:544-560. [PMID: 4589706] doi:10.1016/0010-4809(73)90029-3
22. **Shortliffe EH, Davis R, Axline SG, et al.** Computer-based consultations in clinical therapeutics: explanation and rule acquisition capabilities of the MYCIN system. Comput Biomed Res. 1975;8:303-320. [PMID: 1157471] doi:10.1016/0010-4809(75)90009-9
23. **Weiss S, Kulikowski CA, Safir A.** Glaucoma consultation by computer. Comput Biol Med. 1978;8:25-40. [PMID: 620517] doi:10.1016/0010-4825(78)90011-2
24. **Miller RA, Pople HE, Myers JD.** INTERNIST-1, an experimental computer-based diagnostic consultant for general internal medicine. N Engl J Med. 1982;307:468-476. [PMID: 7048091] doi:10.1056/NEJM198208193070803
25. **Barnett GO, Cimino JJ, Hupp JA, et al.** DXplain. An evolving diagnostic decision-support system. JAMA. 1987;258:67-74. [PMID: 3295316] doi:10.1001/jama.258.1.67
26. **U.S. Food and Drug Administration.** Premarket approval - PAPNET (R) TESTING SYSTEM. Accessed at www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P940029 on 11 July 2023.
27. **Mims C.** Why artificial intelligence isn't intelligent. The Wall Street Journal. 31 July 2021. Accessed at www.wsj.com/articles/why-artificial-intelligence-isnt-intelligent-11627704050 on 16 July 2023.
28. **Murgia M.** Sci-fi writer Ted Chiang: 'The machines we have now are not conscious.' Financial Times. 2 June 2023. Accessed at www.ft.com/content/c1f6d948-3dde-405f-924c-09cc0dcf8c84 on 12 June 2023.
29. **Abdullahi A.** Generative AI vs. predictive AI: what's the difference? eWeek. 11 January 2024. Accessed at www.eweek.com/artificial-intelligence/generative-ai-vs-predictive-ai/ on 20 January 2024.
30. **Tran BD, Latif K, Reynolds TL, et al.** "Mm-hm," "Uh-uh": are non-lexical conversational sounds deal breakers for the ambient clinical documentation technology? J Am Med Inform Assoc. 2023;30:703-711. [PMID: 36688526] doi:10.1093/jamia/ocad001
31. **Goh KH, Wang L, Yeow AYK, et al.** Artificial intelligence in sepsis early prediction and diagnosis using unstructured data in healthcare. Nat Commun. 2021;12:711. doi:10.1038/s41467-021-20910-4 [33514699]
32. **Abdullahi A.** What is a large language model (LLM)? A complete guide. eWeek. 15 February 2024. Accessed at www.eweek.com/artificial-intelligence/large-language-model/ on 4 January 2024.
33. **Esteva A, Chou K, Yeung S, et al.** Deep learning-enabled medical computer vision. NPJ Digit Med. 2021;4:5. [PMID: 33420381] doi:10.1038/s41746-020-00376-2
34. **Esteva A, Kuprel B, Novoa RA, et al.** Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542:115-118. [PMID: 28117445] doi:10.1038/nature21056
35. **McKinney SM, Sieniek M, Godbole V, et al.** International evaluation of an AI system for breast cancer screening. Nature. 2020;577:89-94. [PMID: 31894144] doi:10.1038/s41586-019-1799-6
36. **Dennean K, Gantori S, Kurnia Limas D, et al.** Let's chat about ChatGPT. UBS. 22 February 2023. Accessed at https://secure.ubs.com/public/api/v2/investment-content/documents/XILxY9V9P5RazGpDA1Cr_Q?apikey=Y8VdAx8vhk1P9YXDlEOo2Eoco1fqKwDk&Accept-Language=de-CH on 11 January 2024.
37. **Lee P, Goldberg C, Kohane I.** The AI Revolution in Medicine: GPT-4 and Beyond. Pearson; 2023.
38. **Schreiner M.** GPT-4 architecture, datasets, costs and more leaked. The Decoder. 11 July 2023. Accessed at https://the-decoder.com/gpt-4-architecture-datasets-costs-and-more-leaked/ on 14 November 2023.
39. **American Medical Association.** Principles for augmented intelligence development, deployment, and use. Accessed at www.ama-assn.org/system/files/ama-ai-principles.pdf on 29 November 2023.
40. **U.S. Department of Health and Human Services.** Artificial intelligence strategy. Accessed at www.hhs.gov/sites/default/files/final-hhs-ai-strategy.pdf on 17 August 2023.
41. **U.S. Department of Health and Human Services.** Trustworthy AI (TAI) playbook. Accessed at www.hhs.gov/sites/default/files/hhs-trustworthy-ai-playbook.pdf on 19 July 2023.
42. **U.S. Department of Health and Human Services.** HHS Artificial Intelligence Use Cases Inventory FY23 (CSV). Accessed at www.hhs.gov/sites/default/files/hhs-ai-use-cases-2023-public-inventory.csv on 26 September 2023.
43. **U.S. Food and Drug Administration.** Artificial intelligence and machine learning (AI/ML)-enabled medical devices. Accessed at www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices on 11 January 2024.
44. **Solove DJ, Matsumi H.** AI, algorithms, and awful humans. Accessed at https://fordhamlawreview.org/wp-content/uploads/2024/03/Vol.-92_Solove-Matsumi-1923-1940.pdf on 8 April 2024.
45. **Lin CT.** Automation complacency, the stepladder of AI in EHR's, "Writing a note is how I think". WHAT NOW? Accessed at https://ctlin.blog/2023/05/16/automation-complacency-the-stepladder-of-ai-in-ehrs-writing-a-note-is-how-i-think-what-now/ on 14 November 2023.
46. **Zerilli J, Knott A, Maclaurin J, et al.** Algorithmic decision-making and the control problem. Minds and Machines. 2019;29:555-578. doi:https://doi.org/10.1007/s11023-019-09513-7
47. Talking about AI in human terms is natural—but wrong. The Economist. 22 June 2023. Accessed at www.economist.com/culture/2023/06/22/talking-about-ai-in-human-terms-is-natural-but-wrong on 11 July 2023.
48. **Grimm CM.** The danger of anthropomorphic language in robotic AI systems. Brookings Institution. 18 June 2021. Accessed at www.brookings.edu/articles/the-danger-of-anthropomorphic-language-in-robotic-ai-systems/ on 11 July 2023.
49. **Alabed A, Javornik A, Gregory-Smith D.** AI anthropomorphism and its effect on users' self-congruence and self–AI integration: a theoretical framework and research agenda. Technol Forecast Soc Change. 2022; 182:121786. doi:10.1016/j.techfore.2022.121786
50. **Hatem R, Simmons B, Thornton JE.** Chatbot confabulations are not hallucinations. JAMA Intern Med. 2023;183:1177. [PMID: 37578766] doi:10.1001/jamainternmed.2023.4231
51. **Brender TD.** Chatbot confabulations are not hallucinations-reply. JAMA Intern Med. 2023;183:1177-1178. [PMID: 37578769] doi:10.1001/jamainternmed.2023.3875
52. **Snyder Sulmasy L, Bledsoe TA; ACP Ethics, Professionalism and Human Rights Committee.** American College of Physicians

Ethics Manual: Seventh Edition. Ann Intern Med. 2019;170:S1-S32. [PMID: 30641552] doi:10.7326/M18-2160
53. **Sulmasy LS, López AM, Horwitch CA; American College of Physicians Ethics, Professionalism and Human Rights Committee.** Ethical implications of the electronic health record: in the service of the patient. J Gen Intern Med. 2017;32:935-939. [PMID: 28321550] doi:10.1007/s11606-017-4030-1
54. **McGreevey JD, Hanson CW, Koppel R.** Clinical, legal, and ethical aspects of artificial intelligence-assisted conversational agents in health care. JAMA. 2020;324:552-553. [PMID: 32706386] doi:10.1001/jama.2020.2724
55. **Rockwern B, Johnson D, Snyder Sulmasy L; Medical Informatics Committee and Ethics, Professionalism and Human Rights Committee of the American College of Physicians.** Health information privacy, protection, and use in the expanding digital health ecosystem: a position paper of the American College of Physicians. Ann Intern Med. 2021;174:994-998. [PMID: 33900797] doi:10.7326/M20-7639
56. **Richardson JP, Smith C, Curtis S, et al.** Patient apprehensions about the use of artificial intelligence in healthcare. NPJ Digit Med. 2021;4:140. [PMID: 34548621] doi:10.1038/s41746-021-00509-1
57. **Khullar D, Casalino LP, Qian Y, et al.** Perspectives of patients about artificial intelligence in health care. JAMA Netw Open. 2022;5:e2210309. [PMID: 35507346] doi:10.1001/jamanetworkopen.2022.10309
58. **Cardillo A, Collins W, Cotliar D, et al**. The Physicians' Charter for responsible AI: a practical guide to developing, testing, and using AI tools in clinical practice. Accessed at https://physicianscharter.ai/ on 2 October 2023.
59. **Cahan EM, Hernandez-Boussard T, Thadaney-Israni S, et al.** Putting the data before the algorithm in big data addressing personalized healthcare. NPJ Digit Med. 2019;2:78. [PMID: 31453373] doi:10.1038/s41746-019-0157-2
60. **Gainty C.** From a 'deranged provocateur to IBM's failed AI superproject: the controversial story of how data has transformed healthcare. The Conversation. 16 January 2023. Accessed at https://theconversation.com/from-a-deranged-provocateur-to-ibms-failed-ai-superproject-the-controversial-story-of-how-data-has-transformed-healthcare-189362 on 11 August 2023.
61. **Ross C, Swetlitz I.** IBM's Watson supercomputer recommended 'unsafe and incorrect' cancer treatments, internal documents show. STAT. 25 July 2018. Accessed at www.statnews.com/2018/07/25/ibm-watson-recommended-unsafe-incorrect-treatments/ on 11 August 2023.
62. **American College of Physicians.** Patient safety in the office-based practice setting. Accessed at www.acponline.org/acp_policy/policies/patient_safety_in_the_office_based_practice_setting_2017.pdf on 10 August 2023.
63. **Erickson SM, Outland B, Joy S, et al; Medical Practice and Quality Committee of the American College of Physicians.** Envisioning a better U.S. health care system for all: health care delivery and payment system reforms. Ann Intern Med. 2020;172:S33-S49. [PMID: 31958802] doi:10.7326/M19-2407
64. **Leveson NG.** The Therac-25: 30 years later. Computer. 2017;50:8-11.
65. **Institute for Healthcare Improvement.** How to improve: model for improvement. Accessed at www.ihi.org/resources/how-to-improve on 4 January 2024.
66. **Vaid A, Sawant A, Suarez-Farinas M, et al.** Implications of the use of artificial intelligence predictive models in health care settings: a simulation study. Ann Intern Med. 2023;176:1358-1369. [PMID: 37812781] doi:10.7326/M23-0949
67. **Kostick-Quenet KM, Gerke S.** AI in the hands of imperfect users. NPJ Digit Med. 2022;5:197. [PMID: 36577851] doi:10.1038/s41746-022-00737-z
68. **Fox W; American College of Physicians.** Comments re: Center for Medicare and Medicaid Innovation White Paper: Driving Health System Transformation - A Strategy for the CMS Innovation Center's Second Decade. Accessed at https://assets.acponline.org/acp_policy/letters/acp_comments_on_cmmi_driving_health_system_transformation_white_paper_2022.pdf on 10 August 2023.
69. **Seyyed-Kalantari L, Zhang H, McDermott MBA, et al.** Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. Nat Med. 2021;27:2176-2182. [PMID: 34893776] doi:10.1038/s41591-021-01595-0
70. **Adamson AS, Smith A.** Machine learning and health care disparities in dermatology. JAMA Dermatol. 2018;154:1247-1248. [PMID: 30073260] doi:10.1001/jamadermatol.2018.2348
71. **Daneshjou R, Vodrahalli K, Novoa RA, et al.** Disparities in dermatology AI performance on a diverse, curated clinical image set. Sci Adv. 2022;8:eabq6147. [PMID: 35960806] doi:10.1126/sciadv.abq6147
72. **Thadaney-Israni S, Verghese A.** Humanizing artificial intelligence. JAMA. 2019;321:29-30. [PMID: 30535297] doi:10.1001/jama.2018.19398
73. **ECRI-Penn Medicine Evidence-based Practice Center.** Impact of healthcare algorithms on racial and ethnic disparities in health and healthcare. Comparative Effectiveness Review No. 268. AHRQ Publication No. 24-EHC004. Agency for Healthcare Research and Quality; 2023. doi:10.23970/AHRQEPCCER268
74. **Rajkomar A, Hardt M, Howell MD, et al.** Ensuring fairness in machine learning to advance health equity. Ann Intern Med. 2018;169:866-872. [PMID: 30508424] doi:10.7326/M18-1990
75. **DeCamp M, Lindvall C.** Mitigating bias in AI at the point of care. Science. 2023;381:150-152. [PMID: 37440631] doi:10.1126/science.adh2713
76. **Ayers JW, Desai N, Smith DM.** Regulate artificial intelligence in health care by prioritizing patient outcomes. JAMA. 2024;331:639-640. [PMID: 38285467] doi:10.1001/jama.2024.0549
77. **Shah NH, Halamka JD, Saria S, et al.** A nationwide network of health AI assurance laboratories. JAMA. 2023;331:245-249. [PMID: 38117493] doi:10.1001/jama.2023.26930
78. **U.S. Department of Health and Human Services.** Section 1557 of the Patient Protection and Affordable Care Act. Accessed at www.hhs.gov/civil-rights/for-individuals/section-1557/index.html on 11 January 2024.
79. **Shachar C, Gerke S.** Prevention of bias and discrimination in clinical practice algorithms. JAMA. 2023;329:283-284. [PMID: 36602791] doi:10.1001/jama.2022.23867
80. **American College of Physicians.** Controlling health care costs while promoting the best possible health outcomes. American College of Physicians; 2009.
81. **Crowley R, Mathew S, Hilden D; Health and Public Policy Committee of the American College of Physicians.** Environmental health: a position paper from the American College of Physicians. Ann Intern Med. 2022;175:1591-1593. [PMID: 36279541] doi:10.7326/M22-1864
82. **Eckelman MJ, Huang K, Lagasse R, et al.** Health care pollution and public health damage in the United States: an update. Health Aff (Millwood). 2020;39:2071-2079. [PMID: 33284703] doi:10.1377/hlthaff.2020.01247
83. **Dhar P.** The carbon impact of artificial intelligence. Nat Mach Intell. 2020;2:423-425. doi:10.1038/s42256-020-0219-9
84. **Strubell E, Ganesh A, McCallum A.** Energy and policy considerations for deep learning in NLP. Presented at the 57th Annual Meeting of the Association for Computational Linguistics (ACL). Florence, Italy, June 2019. arXiv. Preprint posted online 5 June 2019. doi:10.48550/arXiv.1906.02243
85. **Heikkilä M**, We're getting a better idea of AI's true carbon footprint. MIT Technology Review. 14 November 2022. Accessed at www.technologyreview.com/2022/11/14/1063192/were-getting-a-better-idea-of-ais-true-carbon-footprint/ on 15 November 2023.
86. **Luccioni AS, Viguier S, Ligozat A.** Estimating the carbon footprint of BLOOM, a 176B parameter language model. J Mach Learn Res. 2023;24:11990-12004.
87. **Chasan A**, Some experts see AI as a tool against climate change. Others say its own carbon footprint could be a problem. CBS News.

26 August 2023. Accessed at www.cbsnews.com/news/artificial-intelligence-carbon-footprint-climate-change/ on 8 September 2023.
88. **Cowls J, Tsamados A, Taddeo M, et al.** The AI gambit: leveraging artificial intelligence to combat climate change-opportunities, challenges, and recommendations. AI Soc. 2023;38:283-307. [PMID: 34690449] doi:10.1007/s00146-021-01294-x
89. **Bloomfield PS, Clutton-Brock P, Pencheon E, et al.** Artificial intelligence in the NHS: climate and emissions. J Clim Chang Health. 2021;4:100056. doi:10.1016/j.joclim.2021.100056
90. **Heikkilä M**, Why we need to do a better job of measuring AI's carbon footprint. MIT Technology Review. 15 November 2022. Accessed at www.technologyreview.com/2022/11/15/1063202/why-we-need-to-do-a-better-job-of-measuring-ais-carbon-footprint/ on 7 September 2023.