

PERSPECTIVE

AI IN MEDICINE

Ethical Obligations to Inform Patients About Use of AI Tools

Michelle M. Mello, JD, PhD, MPhil; Danton Char, MD, MAS; Sonnet H. Xu

Permeation of artificial intelligence (AI) tools into health care tests traditional understandings of what patients should be told about their care. Despite the general importance of informed consent, decision support tools (eg, automatic electrocardiogram readers, rule-based risk classifiers, and UpToDate summaries) are not usually discussed with patients even though they affect treatment decisions. Should AI tools be treated similarly? The legal



Multimedia

doctrine of informed consent requires disclosing information that is material to a reasonable patient's decision to accept a health care service, and evidence suggests that many patients would think differently about care if they knew it was guided by AI. In recent surveys, 60% of US adults said they would be uncomfortable with their physician relying on AI,¹ 70% to 80% had low expectations AI would improve important aspects of their care,² only one-third trusted health care systems to use AI responsibly,³ and 63% said it was very true that they would want to be notified about use of AI in their care.⁴

To help health care leaders and practitioners navigate this thorny terrain, we provide a framework for deciding what patients should be told about AI tools. In contrast to prior work,⁵ we focus on tools that involve human oversight because use of autonomous AI in health care is rare.

Ask, Tell, or Neither?

Notwithstanding emphasis on a "right to notice and explanation" in prominent guidelines such as the Blueprint for an AI Bill of Rights,⁶ it is neither practicable nor desirable to tell patients about every AI tool used at a health care organization. At some hospitals, a patient's care may implicate dozens of AI tools, many operating in the background. This promises to compound the problem that disclosures and consents in health care are often perfunctory and legalistic, not meaningful and autonomy enhancing.^{7,8} Patients who want to be kept informed about their care in principle may struggle with the information overload that a hospital admission entails, making it important to assess what value additional disclosures would contribute. Furthermore, although consent should not substitute for other AI governance mechanisms, overemphasizing consent could undercut the impetus for oversight.⁹ Disclosures may also do harm, stoking distrust of clinicians' recommendations and communications even where evidence suggests that AI improves them.^{10,11} For instance, patients perceive AI-drafted emails to be more empathetic than those drafted by physicians, but if told that an email was drafted by AI, they are significantly less satisfied with it.¹² When AI-augmented care is superior, a consent regime can result in clinicians having to deliver suboptimal treatment, which may cause moral

distress.¹³ Offering opt outs also entails social costs: it may reduce the diversity of the patient data used to train AI tools.⁹ For these reasons, it should not be presumed that disclosure of AI tools always serves patients' interests.

Setting a patient disclosure policy should be part of every decision by a health care organization to deploy an AI tool. We propose that the policy—notify patients that the tool is being used, seek their consent to use it, or neither—should be driven by 2 determinations. First, how serious is the risk that using the tool could cause physical harm to patients?¹⁴ Second, to what extent do patients have a meaningful opportunity to exercise agency in response to a disclosure? When both risk and opportunity for agency are high, the disclosure obligation is at its apex. At a minimum, a robust notification is ethically required; where obtaining consent is feasible, that may be preferable. In contrast, when there is little risk of harm and no real prospect that patients can take action in response to information, it may be reasonable and even preferable not to provide notification. Illustrative examples are given in the **Table**.

Determining the likelihood that an AI tool may cause patient harm requires assessing (1) the risk that the tool poses, (2) the likelihood that errors will reach patients without being intercepted, and (3) the severity of the harm that could result.¹⁵ The greater the product of these factors, the stronger the disclosure obligation. Notably, the product may be high even if some factors are low. For example, an AI surgical visualization tool may have a low risk of error, but errors can instantaneously cause serious harm. A patient portal messaging tool may make potentially consequential mistakes infrequently but entail a low likelihood of clinicians correcting errors.¹⁶

What if the risk of error is higher for some patients than others due to poorer algorithmic performance in subgroups? When patients can be readily identified as members of a subgroup vulnerable to algorithmic bias, it may be desirable to give them tailored information even though the decision framework militates against disclosure to other patients. For example, when applying a predictive algorithm trained primarily on adult data to children, parents should be told that the physician's confidence in the output is lower than it would be for adults.

The second criterion in this framework asks what opportunity patients realistically have to exercise agency with respect to the AI tool in the context of the care setting. Agency could take 2 forms. For some tools, patients may be able to opt out. For example, patients may be able to choose surgery that does not involve an AI-guided robot or mammography without AI interpretation.¹⁷ In contrast, they probably cannot opt out of having their hospital use a predictive algorithm to determine whether to prestock blood in the operating room for their surgery.

Table. Illustrative Artificial Intelligence (AI) Use Cases and Recommendations for Patient Notification

Use case	Recommendation and rationale ^a
Consent	
AI-guided, nonautonomous surgical robot	Consent, because surgery typically involves considerable risk of harm and patients can choose nonrobotic surgery.
Tool that analyzes the genomic information of patients with cancer to predict responses to specific drugs and recommend treatment plans	Consent, because inaccuracy could lead to poor outcomes and patients can opt out of having the tool used.
Notification	
Predictive algorithm to identify patients who may have hypertrophic cardiomyopathy based on existing echocardiography data	Notification of those who screen positive, because although patients cannot avoid having the tool used, knowing that follow-up echocardiography is being recommended because of an AI algorithm may affect their decision about whether to accept the recommendation.
Generative AI tool that drafts replies to patient emails	Notification, because inaccuracy may cause harm and informed patients may be activated to question odd email replies.
Ambient listening tool that generates clinic visit summaries	Notification, because inaccuracy may cause harm and patients can check summaries for accuracy. (In some states, however, consent is legally required.)
Neither consent nor notification	
Predictive algorithm to inform decision whether to stock blood in operating room for potential transfusion	Neither, because patients cannot influence such operational decisions and ensuring blood availability does not replace patient consent for blood transfusion.
AI-assisted mammography interpretation	Neither, assuming AI-assisted interpretation outperforms human only, because exercising the choice to seek alternate care elsewhere would increase risk of harm.
Generative AI tool that drafts summaries of imaging findings dictated by radiologist	Neither, because risk of harm is low given that tool only summarizes radiologist's input.
Predictive algorithm for identifying inpatients for whom daily laboratory testing can be safely discontinued	Neither, because patients cannot opt out of use of the tool or demand that laboratory tests be continued; and because the determination to order laboratory tests comes from the clinician and the alert merely suggests laboratory tests might be safely discontinued.
Generative AI tool to help hospital file prior authorization requests with insurers	Neither, because risk of harm is low (given that denials can be appealed) and patients cannot influence such operational decisions or personally review prior authorization requests for accuracy.

^a Recommendations assume that the AI model has performed well in predeployment testing on a similar population.

For other tools, even though opting out is not possible, patients who understand that AI is involved may alter their behavior in ways that promote their interests. For instance, a patient who knows their physician's email message was drafted by AI may be more likely to question something that seems odd than a patient who thinks their physician was the only author. A daughter who knows that the hospital nurse caring for her mother did not write her own end-of-shift summary note may be more likely to log on to the electronic health record, check the note for errors and omissions, and alert the incoming nurse to the need for corrections. Therefore, health care organizations should assess whether disclosure enhances patients' ability to participate in their care even where no choice can be offered about having the tool used.

Ethical vs Legal Obligations

Our ethical framework is generally consonant with the legal doctrine of informed consent, which requires disclosure of material information. *Material* means information that would be important to a reasonable patient's decision about whether to accept a treatment—or in some states, a reasonable physician's decision about which treatment to offer. Information about trivial risks does not meet this standard, nor does information about things patients cannot avoid having used in their care. However, it is relevant if patients would be able to obtain different care elsewhere.

State laws may specify what patients must be told about certain uses of AI tools. Wiretapping laws may require 2-party consent for ambient visit recording. Some states have adopted laws requiring disclosure that certain patient communications were drafted by AI. These legal requirements reflect the core considerations in this ethical framework: 2-party consent laws allow patients to opt out of recording, and laws about patient emails contemplate that pa-

tients may ask questions about puzzling communications they know were written by bots.

Critiques of the Framework

Unlike some prior work, this framework does not require disclosure of all AI tools that are high risk in the sense that they perform an important clinical function or are deployed on seriously ill patients.^{14,18} This may seem counterintuitive. However, most tools in use today entail a human reviewing and acting on model output. The key question therefore is whether the residual risk to patients is low once one considers the likelihood that the human in the loop will successfully detect errors. Capacity constraints, automation bias, and users' unfamiliarity with a tool's weaknesses may undercut efficacy.¹⁹

This framework also excludes several factors that would seem salient to disclosure obligations. First, AI tools arguably are not meaningfully different from other decision support tools or risk-producing aspects of care about which patients are not informed. This line of reasoning assumes that current disclosure practices are optimal, however. Furthermore, AI is a cognitive intervention likely to function with less human oversight than prior generations of decision aids. Second, some commentators consider risk of discrimination to be an independent reason to disclose use of algorithmic tools.⁷ But although use of attributes such as race in algorithms creates obligations to monitor and mitigate bias, the risk of discrimination can be considered within the general rubric of risk of harm described above.

Third, perhaps disclosure policies should consider whether using the AI tool requires sending identifiable patient information outside the health system, because many patients feel uneasy about third-party use of their data. A difficulty, however, is that the line

Box. Sample Patient Consent for an Ambient Scribe Tool for Clinic Visits^a**Modality**

Information is distributed via paper information sheet when patient checks in for clinic visit; nurse elicits oral consent when patients are brought to examination room.

Information Sheet

Apex Hospital uses an artificial intelligence (AI) tool called Intelligent Ear to help your physician take notes on your visit today and summarize them in your medical record. Under California law, you have the right not to have this tool used.

Intelligent Ear makes an audio recording of your physician visit and uses AI to write out the conversation in full. It then uses a different kind of AI to generate a summary of the important parts of the conversation. Your physician will have the opportunity to review and edit this summary before making it part of your medical record.

The audio recording is securely sent to the company that developed Intelligent Ear, but the company does not keep any information about you once the visit summary is created.

Apex Hospital is using this tool because of evidence that it can improve care. Having this type of aid enables physicians to keep

their focus on you during your visit and save time writing notes at the end of the day. Testing showed Intelligent Ear generates good-quality summaries and that physicians find it helpful.

The tool performs somewhat less well for people with accented English, people with speech impediments, and conversations with more than 2 speakers. To help avoid inaccuracies, your physician has been trained to review summaries with special care in those circumstances. Apex Hospital is using this tool after studying the results of a pilot test at Apex, and continues to monitor the tool's performance to make sure it is working well for all types of patients.

We are excited to be able to offer this tool in our continuing efforts to deliver the best care possible. If you do not wish to have the tool used in your visit today, please let your nurse know.

Nurse Script

Did you have any questions about the scribe tool that the physician uses? Is it OK with you if the physician uses it today?

^a For use in a state where 2-party consent for recording conversations is legally required. In other states, notification without nurse confirmation of consent would be ethically sufficient.

between deidentified and potentially identifiable data is dissolving, making this an unstable basis for distinguishing cases where disclosure is and is not necessary. A related issue is that inferences about people can be drawn from the data of others like them, sapping the protective power of consent.⁹

Fourth, some may argue that disclosure policies should vary with the complexity of the AI tool because consent is not meaningful if a patient cannot truly understand what they are accepting. In our view, it is less important for patients to comprehend how an algorithm works than to grasp the risks and benefits associated with using it. Other medical products also work through mechanisms that most laypersons cannot understand, yet this complexity does not alter the extent to which ordinary consent practices apply.

What Should Be Disclosed?

Disclosures need not be extensive to serve their ethical purpose. Using plain and concrete language, patient notifications should cover (1) the fact that an AI tool is being used; (2) what functions it performs; (3) a basic description of how it works, including the clinician's role; (4) why the organization believes it improves care; (5) a basic description of how the organization monitors the tool's performance, including differential performance across patient subgroups; and (6) where applicable, what choices the patient has about having the tool used in their care.^{20,21} Some evidence suggests that patients may also like to receive information about data privacy and whether the tool is cleared by the US Food and Drug Administration.²² Where patient consent, as opposed to notification, is deemed appropriate, additional information about known risks, benefits, and alternatives should be involved, including information about the risks associated with not using the AI tool.²³ But even here, detailed technical information, such as the specific machine learning method the algorithm uses, will not be useful to most patients.

The optimal modality for conveying information will vary depending on the tool. Where full verbal disclosures are infeasible, online and printed information sheets are useful alternatives.⁵

For AI-generated clinical summaries and patient emails, a simple disclosure that the text was generated by an AI tool and reviewed by a clinician can direct interested patients to additional information via a hyperlink. An example of a simple disclosure is provided in the Box.

Conclusions

Health care practitioners are right to ask questions about what patients should be told about AI use, but they should not be left to grapple with them on their own. When a health care organization decides to deploy an AI tool, it should decide whether notification or consent is ethically required.²⁴ If it is, the organization should identify steps it can take to ensure that this policy is applied consistently.

To decide that disclosure is not required at the point of care is not to conclude that organizations have no transparency obligations concerning their use of AI. Two complementary steps can help discharge these obligations. First, health care organizations' websites and written materials distributed to patients should explain how the organization uses AI to deliver better, safer, more efficient care and protect staff well-being.⁵ They should describe the types of AI tools currently in use and how the organization ensures their safe and responsible use. This kind of transparency honors patients' desire to know how their care may be affected by AI and may help some make decisions about where to seek care. It also strengthens organizations' incentives to create strong governance processes for AI so that they have a compelling story to tell about their own accountability.

Second, health care organizations should explore ways to involve patient representatives in AI governance processes.²⁵ For example, structured consultations with patients who have had significant health care experiences and received training on AI can help ensure that organizational leaders understand patient perspectives on important decisions. Both of these steps recognize that more important than merely communicating risks of AI tools is taking action to minimize them.

ARTICLE INFORMATION

Author Affiliations: Stanford Law School, Stanford, California (Mello); Department of Health Policy, Stanford University School of Medicine, Stanford, California (Mello, Xu); Department of Anesthesia, Stanford University School of Medicine, Stanford, California (Char).

Corresponding Author: Michelle M. Mello, JD, PhD, MPhil, Stanford Law School, 559 Nathan Abbott Way, Stanford, CA 94305 (mmello@law.stanford.edu).

Published Online: July 21, 2025.
doi:10.1001/jama.2025.11417

Conflict of Interest Disclosures: Dr Mello reported grants from Stanford Health Care, the Patient-Centered Outcomes Research Institute, and Stanford Impact Labs during the conduct of the study and personal fees from AVIA for serving as a member of AVIA's Generative AI Collaborative and Augmedix for a presentation; Dr Mello's spouse is an executive with Cisco Systems Inc, which develops artificial intelligence tools and products that power such tools. Dr Char reported grants from the Patient-Centered Outcomes Research Institute and Stanford Impact Labs during the conduct of the study. Ms Xu reported grants from Stanford University's Humanities Research Intensive Fellowship program during the conduct of the study.

REFERENCES

1. Tyson A, Pasquini G, Spencer A, Funk C. 60% of Americans would be uncomfortable with provider relying on AI in their own health care. Accessed May 22, 2025. <https://www.pewresearch.org/science/2023/02/22/60-of-americans-would-be-uncomfortable-with-provider-relying-on-ai-in-their-own-health-care/>
2. Nong P, Ji M. Expectations of healthcare AI and the role of trust: understanding patient views on how AI will impact cost, access, and patient-provider relationships. *J Am Med Inform Assoc*. 2025;32(5):795-799. doi:10.1093/jamia/ocaf031
3. Nong P, Platt J. Patients' trust in health systems to use artificial intelligence. *JAMA Netw Open*. 2025;8(2):e2460628. doi:10.1001/jamanetworkopen.2024.60628
4. Platt J, Nong P, Carmona G, Kardia S. Public attitudes toward notification of use of artificial intelligence in health care. *JAMA Netw Open*. 2024;7(12):e2450102. doi:10.1001/jamanetworkopen.2024.50102
5. Elmore M, Economou-Zavlanos N, Pencina M. A heuristic for notifying patients about AI: from institutional declarations to informed consent. *Am J Bioeth*. 2025;25(3):121-123. doi:10.1080/15265161.2025.2458422
6. Office of Science and Technology Policy, The White House. Blueprint for an AI Bill of Rights. Accessed May 22, 2025. <https://bidenwhitehouse.archives.gov/ostp/ai-bill-of-rights/>
7. Bridges K. Race in the machine: racial disparities in health and medical AI. *Va Law Rev*. 2024;110(2):243-339.
8. Spector-Bagdady K, London AJ. Disclosure as abolition in medicine: disentangling autonomy from beneficence and justice in artificial intelligence. *Am J Bioeth*. 2025;25(3):1-3. doi:10.1080/15265161.2025.2458424
9. Evans, BJ, Bihorac A. Co-creating consent for data use—AI-powered ethics for biomedical AI. *NEJM AI*. 2024;1(7). doi:10.1056/Alpc2400237
10. Rosenbluth T. That message from your doctor? it may have been drafted by AI. *New York Times*. September 24, 2024. Accessed May 22, 2025. <https://www.nytimes.com/2024/09/24/health/ai-patient-messages-mychart.html>
11. Sharko M, Cole CL. Integrating artificial intelligence support in patient care while respecting ethical principles. *JAMA Netw Open*. 2025;8(3):e250462. doi:10.1001/jamanetworkopen.2025.0462
12. Cavalier JS, Goldstein BA, Ravitsky V, et al. Ethics in patient preferences for artificial intelligence-drafted responses to electronic messages. *JAMA Netw Open*. 2025;8(3):e250449. doi:10.1001/jamanetworkopen.2025.0449
13. Binkley CE, Pilkington BC. Informed consent for clinician-AI collaboration and patient data sharing: substantive, illusory, or both. *Am J Bioeth*. 2023;23(10):83-85. doi:10.1080/15265161.2023.2250289
14. Rose SL, Shapiro D. An ethically supported framework for determining patient notification and informed consent practices when using artificial intelligence in health care. *Chest*. 2024;166(3):572-578. doi:10.1016/j.chest.2024.04.014
15. Mello MM, Guha N. Understanding liability risk from using healthcare artificial intelligence tools. *N Engl J Med*. 2024;390(3):271-278. doi:10.1056/NEJMhle2308901
16. 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
17. Andrews M. Mammography AI can cost patients extra: is it worth it? KFF Health News. January 10, 2024. Accessed May 22, 2025. <https://kffhealthnews.org/news/article/artificial-intelligence-mammography-extra-cost/>
18. Pflanzner M. Balancing transparency and trust: reevaluating AI disclosure in healthcare. *Am J Bioeth*. 2025;25(3):153-156. doi:10.1080/15265161.2025.2457700
19. Patil SV, Myers CG, Lu-Myers Y. Calibrating AI reliance: a physician's superhuman dilemma. *JAMA Health Forum*. 2025;6(3):e250106. doi:10.1001/jamahealthforum.2025.0106
20. Hurley ME, Lang BH, Kostick-Quenet KM, Smith JN, Blumenthal-Barby J. Patient consent and the right to notice and explanation of AI systems used in health care. *Am J Bioeth*. 2025;25(3):102-114. doi:10.1080/15265161.2024.2399828
21. Yang YT. Beyond disclosure: rethinking patient consent and AI accountability in healthcare. *Am J Bioeth*. 2025;25(3):151-153. doi:10.1080/15265161.2025.2457688
22. Stroud AM, Miller JE, Barry BA. Cultivating patient-centered healthcare artificial intelligence transparency: considerations for AI documentation. *Am J Bioeth*. 2025;25(3):129-131. doi:10.1080/15265161.2025.2457775
23. Ploug T, Holm S. Consenting to and contesting AI use—different functions but both are necessary. *Am J Bioeth*. 2025;25(3):124-126. doi:10.1080/15265161.2025.2457721
24. Millen M, Tai-Seale M, Longhurst CA. A call for disclosure when using AI for patient communications. Published online May 9, 2025. *NEJM AI*. doi:10.1056/Alp2401167
25. Callahan A, McElfresh D, Banda JM, et al. Standing on FURM ground: a framework for deploying fair, useful, and reliable AI models in health care systems. *NEJM Catal*. 2024;5(10). doi:10.1056/CAT.24.0131