Stanford University
Human-Centered
Artificial Intelligence

Policy Brief
HAI Policy & Society
February 2024

# Understanding Liability Risk from Healthcare AI

Michelle M. Mello and Neel Guha

**ARTIFICIAL INTELLIGENCE (AI) HOLDS TREMENDOUS POTENTIAL TO TRANSFORM HEALTHCARE. But even amid vast opportunities to improve patient care and reduce costs, grave concerns about the wide-ranging risks of adopting AI tools persist. Attorneys worry about liability and litigation implications for healthcare organizations, which must comply with still evolving federal laws. Perhaps the most pressing legal question is: Who will be held responsible when AI tools contribute to patient injury?**

Perceptions about liability risk will influence physicians' and healthcare organizations' willingness to use AI tools. Outsized liability concerns can lead to conservative decision-making regarding AI innovation and adoption. In the past, older forms of clinical decision support—such as software to manage patient care and improve patient safety—have enabled healthcare organizations to prevent injuries and malpractice claims. In that sense, *not* adopting new technological tools could eventually be viewed as a harmful decision.

In our paper, "Understanding Liability Risk from Using Healthcare Artificial Intelligence Tools," we examined the challenges courts face in dealing with cases involving software errors. We further analyzed how AI tools can increase or mitigate legal risk before concluding

## Key Takeaways

Optimism about AI's tremendous potential to transform healthcare is tempered by concerns about legal liability: Who will be held responsible when the use of AI tools contributes to patient injury?

Case law on physical injury caused by AI or software systems is sparse. Our analysis of 51 such cases revealed that liability claims generally relate to harm caused by defects in software used to manage care or resources, physicians' use of software in making care decisions, or the malfunctioning of software embedded in medical devices.

The intangible and opaque nature of software and AI models poses significant challenges for holding software developers liable according to traditional rules governing product liability. Until tort doctrine evolves to address the impact of AI, plaintiffs may struggle to assert, let alone win, their legal claims.

We provide a risk assessment framework that will help healthcare organizations calibrate their approach to implementing and monitoring healthcare AI tools based on a careful assessment of the liability risk of each tool. Regulating healthcare AI should also take into account the different degrees of risk of harm.

Carefully negotiating licensing agreements with AI developers is an important avenue for healthcare organizations to mitigate liability risk.

with several risk-management recommendations for healthcare organizations, focusing on AI applications that have a "human in the loop." Our research will support healthcare organizations, physicians, patients, and policymakers as they weigh the potential benefits against the liability risks of AI use in medicine, while helping them navigate evolving liability issues to ensure the safe adoption of AI tools.

*...not adopting new technological tools could eventually be viewed as a harmful decision.*

# Introduction

There is sparse case law pertaining to AI-related liability in healthcare. Medical AI models are still relatively new and few personal injury claims have led to judicial opinions. In the software liability cases that have been decided to date, plaintiffs have grappled with a variety of challenges.

Typically, when a product injures a patient, courts look to well-established rules to determine how to allocate liability between the party using the product and the company that made it. The plaintiff must show that the defendant owed them a "duty of care," that the defendant's conduct fell below the "standard of care," and that this violation caused the injury. But making these determinations is much more complicated for AI and other software tools applied in healthcare settings.

Because software is not a tangible object, courts have been reluctant to apply product liability doctrines to AI-related injury claims. The doctrine of "preemption," meanwhile, prevents patients from making personal injury claims in state courts related to certain devices that have already been cleared by the Food and Drug Administration. Additionally, most states require that patients suing a product manufacturer demonstrate that there is a reasonable, safer, alternative design and that injury was foreseeable. Meeting these demands is technically difficult, given plaintiffs' limited ability to see into the "black box" of AI systems. Finally, plaintiffs suing physicians must show the decision to follow or depart from model predictions was "unreasonable." The tendency for models to perform well on some patient populations but not others—in addition to general problems of opacity—makes it difficult to prove that errors were reasonably foreseeable.

Tort law, which apportions liability for injury or loss, has a history of evolving to adapt to technological changes—and it will here too. To better understand the current state of play and the continuing evolution of liability risk for healthcare AI, we reviewed 803 court cases and studied the salient issues addressed in 51 judicial decisions related to physical injuries from AI and other software (in both health- and non-health-related contexts).

# Research Outcomes

Known cases related to medical software or AI have clustered around three main scenarios, each of which is instructive for understanding the different types of potential AI healthcare liability.

*Tort law (...) has a history of evolving to adapt to technological changes—and it will here too.*

First are cases where defects in software used to manage care or resources cause harm to patients, who in turn sue the developer of the software and/or the hospital for negligently maintaining it. In Lowe v. Cerner, for example, the court upheld the plaintiffs' claim that a drug-management software product had a defective user interface, which led physicians to mistakenly believe they had scheduled medication they hadn't. Ambrose v. St. Joseph's Hospital of Atlanta is a case involving a hospital alleged to have harmed patients by failing to update software on a surgical microscope. Because the software was managing administrative functions, the plaintiffs were able to sue for negligence instead of filing a medical malpractice lawsuit.

In the second type of cases, patients sue when harm occurs after physicians consult software to make care decisions, such as a technician screening patients for conditions or a doctor generating medication regimens. Plaintiffs sue the developer for erroneous software design and/or the physician for relying on erroneous software recommendations. In the 2023 case Sampson v. HeartWise Health Systems Corporation, physicians followed the output of a software program for cardiac health screening that wrongly classified a young adult with a family history of congenital heart defects as "normal." The patient died weeks later.

Malpractice claims against physicians of this kind are generally determined based on what other specialists would have done. The Sampson case was allowed to proceed to trial based on the allegation that the physicians should have scrutinized the erroneous software recommendation or independently reached the right decision. But courts have varied in their approaches to claims against software developers. In one non-medical case, design defect claims were dismissed on the basis that algorithms aren't products and applying tort doctrine could implicate First Amendment free speech rights. In other cases, courts allowed medical and ordinary negligence claims against the software developer for violating the standard of care. And in Sampson, the court dismissed the ordinary negligence claim because the developer's licensing agreement gave physicians final decision-making responsibility and the developer wasn't a "healthcare provider" under state law. Collectively, these cases point to a future where developer liability varies depending on private contracts and jurisdictional variation.

*These cases point to a future where developer liability varies depending on private contracts and jurisdictional variation.*

The third group of cases we examined concerns situations where software embedded within devices malfunctions. Patients can sue physicians and hospitals and allege that they negligently used, installed, or maintained devices, including implantables, surgical robots, and monitoring tools. In Seargeant v. Orthopedic Associates Medical Clinic, physicians, a technician, and a clinic were sued after a human error during a reprogramming of an infusion pump led to a lethal morphine dosing. Other cases target developers for defects in manufacturing, design, and warnings, though plaintiffs often fail to meet the court's demand that they identify specific design flaws rather than point to the software failure itself as a defect.

In sum, our case review identified three emerging trends:

1. **Plaintiffs struggle to sustain claims when they cannot identify specific design defects in software.** It's not enough simply to show that the software produced an error; plaintiffs must show how and why the error occurred. The difficulty of understanding how AI models produce their outputs makes this a burdensome task.
2. **AI algorithms perform differently for different groups of patients.** This makes it challenging to prove a physician should have known the output wasn't reliable for a particular patient.
3. **Courts appear not to distinguish AI from traditional software.** This raises the risk that cases about one kind of software will have impacts on other kinds, even if they are distinct and should be treated as such.

*One of the most important steps is to resist lumping all AI applications together, as courts have done.*

## Policy Discussion

The future of tort doctrine on healthcare AI is highly uncertain. There is a limited body of case law around harm stemming from software use in medical settings, with even fewer cases related specifically to AI. What's more, courts do not clearly distinguish cases involving healthcare software and healthcare AI—when nuanced questions pertaining to liability matter for patients, physicians, policymakers, and the deployment of AI in healthcare. While awaiting clarification of how tort doctrine will evolve, healthcare organizations and clinicians should take steps to manage possible liability.

One of the most important steps is to resist lumping all AI applications together, as courts have done. The potential AI use cases in healthcare vary widely—from image analysis to drug regimen development—and some tools are riskier than others. Our proposed framework for assessing healthcare AI liability risk, which draws on past literature, conceptualizes risk as a function of four major factors and recommends calibrating adoption decisions and post-deployment safety monitoring based on these risk indicators:

1. **The likelihood and nature of errors** (based on the AI model, its training data, its task design, and how it is integrated into clinical workflow).
2. **The likelihood that humans or another system will detect errors before they harm patients** (which depends in part on how much time with and visibility into the AI tool humans have).
3. **The potential harm if errors are not caught** (especially for tools that perform critical clinical functions or are used in caring for patients with serious health conditions).
4. **The likelihood that injuries would garner compensation in the tort system** (which turns on, among other things, the severity of the injury, the ease of proving negligence, and the causal relationship between the AI tool and the injury).

*Policymakers could help prevent injuries by adopting policies to help hospitals and physicians obtain the information they need to use AI tools safely.*

It is also important to recognize that healthcare organizations are in a buyer's market where many AI developers are jockeying to secure health system contracts and to access patient data. This buyer's market gives healthcare purchasers the opportunity to bargain for terms that minimize liability risk. For instance, licensing agreements should require developers to provide information that allows healthcare organizations to effectively assess and monitor risk, including information on assumptions regarding the model's ingested data, validation processes, and recommendations for auditing model performance.

Indemnification clauses are another important tool for informing liability. Purchasers can use indemnification clauses to establish who pays up, for example requiring that developers pay for errors in the model's output while hospitals pay for errors arising from poor deployment or misuse of the AI technology. Contracts between healthcare organizations and AI developers could also specify minimum insurance requirements and the healthcare organization's responsibilities to monitor systems post-deployment.

Policymakers could help prevent injuries by adopting policies to help hospitals and physicians obtain the information they need to use AI tools safely—for example, disclosure requirements AI developers must follow about how AI models are trained and tasked. Policymakers and physicians may also want to consider establishing guidelines for informing patients when AI is used in diagnostic or treatment decisions to provide a basis for informed consent. Lawyers, for their part, will need to become more literate in healthcare AI to effectively litigate cases.

AI applications have tremendous potential to improve the quality of care and boost patient outcomes. But the potential for injuries means that healthcare organizations, physicians, patients, and policymakers must carefully weigh the risks against the benefits of adoption. Responsibly managing liability will be crucial to harnessing the vast potential of healthcare AI innovations.

Reference: The original article is accessible at Michelle M. Mello and Neel Guha, "**Understanding Liability Risk from Using Healthcare Artificial Intelligence Tools**," *The New England Journal of Medicine*, 390 (January 2024): 271-278, https://pubmed.ncbi.nlm.nih.gov/38231630/. Copies of the article can be obtained from the authors (mmello@law.stanford.edu).

**Michelle M. Mello** is a professor of law at Stanford Law School and professor of health policy at Stanford University School of Medicine.

**Neel Guha** is a fourth-year JD/PhD student in computer science at Stanford University.

Stanford University's Institute for Human-Centered Artificial Intelligence (HAI) applies rigorous analysis and research to pressing policy questions on artificial intelligence. A pillar of HAI is to inform policymakers, industry leaders, and civil society by disseminating scholarship to a wide audience. HAI is a nonpartisan research institute, representing a range of voices. The views expressed in this policy brief reflect the views of the authors. For further information, please contact **HAI-Policy@stanford.edu.**

HAI **Stanford University** Human-Centered Artificial Intelligence

**Stanford HAI:** 353 Jane Stanford Way, Stanford CA 94305-5008

**T** 650.725.4537 **F** 650.123.4567 **E** HAI-Policy@stanford.edu **hai.stanford.edu**