How Much Can Potential Jurors Tell Us About Liability for Medical Artificial Intelligence?

W. Nicholson Price II[1,2], Sara Gerke[2,3], and I. Glenn Cohen[2,3]

[1]University of Michigan Law School, Ann Arbor, Michigan; [2]Project on Precision Medicine, Artificial Intelligence, and the Law, Petrie-Flom Center for Health Law Policy, Biotechnology, and Bioethics at Harvard Law School, Harvard University, Cambridge, Massachusetts; and [3]Harvard Law School, Harvard University, Cambridge, Massachusetts

See the associated article on page 17.

Artificial intelligence (AI) is rapidly entering medical practice, whether for risk prediction, diagnosis, or treatment recommendation. But a persistent question keeps arising: What happens when things go wrong? When patients are injured, and AI was involved, who will be liable and how? Liability is likely to influence the behavior of physicians who decide whether to follow AI advice, hospitals that implement AI tools for physician use, and developers who create those tools in the first place. If physicians are shielded from liability (typically medical malpractice liability) when they use AI tools, even if patient injury results, they are more likely to rely on these tools, even if the AI recommendations are counterintuitive. On the other hand, if physicians face liability from deviating from standard practice, whether an AI recommends something different or not, the adoption of AI is likely to be slower, and counterintuitive rejections—even correct ones—are likely to be rejected. In this issue of *The Journal of Nuclear Medicine,* Tobia et al. (*1*) offer an important empiric look at this question, which has significant implications as to whether and when AI will come into clinical use.

In 2019, we offered a set of possibilities to explore physician liability for the use of AI to make treatment decisions, concluding that under existing law the safest path for physicians was to use AI as a confirmatory tool but ultimately to stay within the existing standard of care when AI recommendations stray from that standard (Fig. 1, columns 1–6) (*2*). The study of Tobia et al. takes the scenarios we suggested and asks a simple but important question: How would potential jurors (here, 2,000 individuals in an online vignette study) view our scenarios in terms of whether the physician's actions were reasonable when a patient was harmed? Their results, partially shown in abbreviated form in Figure 1 (column 7), offer an important additional dimension to our doctrinal analysis. Essentially, potential jurors indicated that following the standard of care could result in no liability for physicians—but independently, so could following the advice of the AI system. If

a physician followed an AI recommendation to deviate from the normal standard of care, and the AI was actually incorrect, with patient injury resulting (Fig. 1, scenario 4), potential jurors were fairly likely to find that decision reasonable. In fact, potential patients seemed to find following the AI's recommendations more important, from a reasonableness perspective, than following the preexisting standard of care.

IMPLICATIONS FOR LIABILITY IN MEDICAL AI

These results suggest that, at least with respect to potential jurors and lay understanding, the use of AI might be closer to the standard of care than we might think. Following the advice of AI already reduces the risk of liability for injury that results from deviations from the standard of care. The contrary is not yet true, though: deviating from nonstandard AI recommendations is not yet viewed as subjecting a physician to liability on its own. This pattern, if it holds, should help physicians feel more at ease using AI to help them make decisions—and not just relying on AI to confirm what they already think. Developers and hospitals might similarly be more willing to create and implement AI tools in practice.

However, the results also suggest that we remain at the moment in a liminal zone; standards of care protect physicians, but so does AI. Results are uncertain and are likely to remain so (at least until and unless the use of AI itself becomes a new standard of care).

These are important results of a well-designed study. Of course, there are complications in translating a study of this sort into real life and investigating the implications. With respect to how legal cases will turn out, there are 2 axes of additional complexity: how the law circumscribes the role of a jury, and how the jury functions in practice.

Although the jury is central to liability, very few civil cases actually make it to juries (*3*). Most cases settle—though settlement negotiations are certainly informed by what juries are likely to do. Perhaps more importantly, judges have multiple opportunities to resolve cases against patients without trial. If in our hypothetical scenarios, among other things, a judge determines no reasonable juror could find for the plaintiff because it is beyond dispute that the physician followed the standard of care, the judge would not let the case go to the jury and instead would find for the physician. To illustrate: in scenario 3 in Figure 1, a physician followed the standard of care and rejected the AI's nonstandard advice. Because the physician unambiguously followed the standard of care, even though prospective jurors found this to be the second-least-reasonable scenario, a real case with these facts would likely be resolved in favor of the physician before it got to the jury. To be sure, battling medical experts often dispute what the standard of care

Scenario	AI recommendation	AI accuracy	Physician action	Patient outcome	Legal outcome (probable)	Empirical Study Findings (Tobia et al.)
1	Standard of care	Correct	Rejects	Injury	Liability	☑ ★
2		Incorrect (standard of care is incorrect)	Follows	Injury	No liability	☑ ★★★★
3	Nonstandard care	Correct (standard of care is incorrect)	Rejects	Injury	No liability	◇ ★★
4		Incorrect	Follows	Injury	Liability	☒ ★★★

FIGURE 1. Comparison of potential legal outcomes under current law according to analysis of Price et al. (2) and empiric study findings of Tobia et al. (1). ★ = agreement that physician decision was reasonable (highest is ★★★★; lowest is ★). Greater agreement indicates lower likelihood of liability; ☑ = study results confirming our analysis of current tort law; ◇ study results suggesting that jury outcome may also be liability; ☒ study results suggesting that jury might decide no liability.

actually requires, but a clear standard of care—whether established by documents or experts—will mean no physician liability, even if a jury would prefer the physician had followed the AI instead.

The standard of care is especially protective of physicians because it is not unitary; if physicians can show that they have followed the standard as practiced by a "respectable minority" of the medical field, that too will be enough to prevent liability (4). So even when the use of AI becomes the standard of care (5), older rubrics will protect physicians who follow them for some time. None of this is captured in the study design Tobia et al. have adopted, which asks how a juror would view the case; in the U.S. system, only a tiny fraction of such cases ever make it to a juror. That said, settlement does typically take place in the shadow of law, such that anticipation of what a juror might do will feed back into settlement.

A second complication arises within the role of the jury. Unlike in vignette studies, jurors do not make decisions in a vacuum; they are instructed in the law by the judge and then engage in deliberative decision making. From observing only individual juror decisions, there is no reliable way to predict how collective jury verdicts will arise, though median jurors do tend to predict damage outcomes (6).

To some extent, these are standard critiques of mock juror research designs, and Tobia et al. have been careful in their design and not overclaimed.

FUTURE WORK AND BROADER IMPLICATIONS

The study of Tobia et al. should serve as a useful beachhead for further work to inform the potential for adoption of AI into medical practice. We note 3 avenues: hospital adoption, negligence more broadly, and AI deference outside medicine.

One key question, and one that Tobia et al. raise, is how liability factors directly into hospital decisions to adopt AI systems. With respect to liability, how do prospective jurors weigh the decisions described here, in which a physician interacts with AI, against situations in which there is no AI in the picture at all? If, as Tobia et al. suggest, following AI provides a different and independent liability-lessening effect, then hospitals may well be more eager to adopt AI systems. But that question should be independently evaluated. The relationship between hospital willingness to adopt an AI and the ramifications for physicians who do not follow it is complicated, mediated by (among other things) malpractice insurance, reimbursement rules, and whether the use of AI becomes part of the standard of care that hospitals themselves must follow. Stepping back,

as Tobia et al. note, the question of adoption will depend on more than liability; the impact of AI on quality and cost of care may matter more. Moreover, even if the use of AI increased the likelihood of liability for any given injury (a result Tobia et al. suggest is unlikely), if AI decreases the rate of error, as all hope it will, liability might still cut in favor of adoption.

A more policy-oriented question asks whether negligence is an effective framework for governing quality and redressing injuries when AI is involved. Selbst argues that the inscrutability of AI makes negligence fundamentally problematic because it interposes an obfuscatory layer between human actors and the consequences of their actions (7). Among other things, AI can replicate biases that exist in the medical system (8), but in such a way that the tort system cannot identify that bias (7,9). Questions of causation are also likely to be endemic, raising questions about the right framework for AI going forward. Rather than negligence, should a no-fault liability system be imposed for AI in medical or other contexts, or should some other system be created?

Finally, the study raises questions about AI liability more broadly: does the protective effect of following AI recommendations also apply to other domains, such as automated vehicles? Here we suggest caution. In medicine, some AI may be an inscrutable decisionmaker granted some level of trust and deference—but so, often, is a physician. For domains where the impacted individuals have their own understanding of the underlying system, deference to algorithmic recommendations might not develop as easily as deference can be transferred between different inscrutable actors in medicine.

DISCLOSURE

REFERENCES

1. Tobia K, Nielsen A, Stemitzer A. When does physician use of AI increase liability? *J Nucl Med.* 2021;62:17–21.
2. Price WN, Gerke S, Cohen IG. Potential liability for physicians using artificial intelligence. *JAMA.* 2019;322:1765–1766.
3. Galanter M. The Seegers Lecture: the decline of trials in a legalizing society. *Valparaiso University Law Rev.* 2017;51:559–577.
4. Mello MM. Of swords and shields: the role of clinical practice guidelines in medical malpractice. *University Pa Law Rev.* 2001;149:645–710.
5. Froomkin AM, Kerr I, Pineau J. When AIs outperform doctors: confronting the challenges of a tort-induced over-reliance on machine learning. *Ariz Law Rev.* 2019;61:33–99.
6. Campbell J, Chao B, Robertson C, Yokum DV. Countering the plaintiff's anchor: jury simulations to evaluate damages arguments. *Iowa Law Rev.* 2016;101:543–571.
7. Selbst AD. Negligence and AI's human users. *Boston University Law Rev.* 2020;100:1315–1376.
8. Benjamin R. Assessing risk, automating racism. *Science.* 2019;366:421–422.
9. Chamallas M. Architecture of bias: deep structures in tort law. *University Pa Law Rev.* 1998;146:463–531.