UNIVERSITY OF CALIFORNIA, BERKELEY

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO

SANTA BARBARA • SANTA CRUZ

ZIAD OBERMEYER
ASSOCIATE PROFESSOR
SCHOOL OF PUBLIC HEALTH
UC BERKELEY

2121 BERKELEY WAY WEST
BERKELEY, CALIFORNIA 94720-7360
ZOBERMEYER@BERKELEY.EDU
TEL (510) 664-5390; FAX (510) 643-5056

Testimony Before the Committee on Oversight & Government Reform
Lowering the Cost of Healthcare: Technology's Role in Driving Affordability

December 6, 2025

Summary

AI can lower healthcare costs *and* improve quality—a rare combination. Unlike past innovations, AI is not just a new technology: it provides a new way of *making decisions about* technology. This means it can help doctors and patients allocate the amazing products of biomedical research to those who need it. The federal government can help accelerate progress towards these goals in several ways.

1. US health records are a vast resource for American leadership in health AI, but these data are dramatically under-used because of red tape and frictions in access. HHS should update its guidance to encourage streamlined deidentification of AI-ready data (physiological streams, notes, images) under existing Safe Harbor standards. Vendors should not be allowed to choke off access. And the federal agencies that hold health data should set aggressive targets for rapid delivery to those who wish to use it to improve patient care, while using modern data management methods and stringent legal enforcement to protect privacy.

2. AI progress is driven by transparency: public benchmarks track performance, driving competition and innovation. By contrast, when regulators evaluate health AI, they rely on lengthy and opaque processes that waste time and money. Targeted amendments to the CURES Act are needed, but meanwhile the FDA should take advantage of the fact that AI systems can be evaluated in a clear, quantitative way. Refocusing evaluation on measuring predictive performance in representative, previously unseen data would allow the FDA to streamline its processes. This will aid smaller, innovative new entrants and help foster a more dynamic market.

3. CMS could be one of the world's largest consumers of AI, and reap the largest benefits. It should create new mechanisms (codes or care models) that pay for AI to achieve specific cost reduction and quality improvement goals, setting the price based on the value these tools create. If CMS signals it is willing to pay for high-value AI, it will unleash the full power of the market and ensure it is pointed in directions that create value for patients and society.

Increasing data availability, focusing regulatory processes on what matters, and creating a market for high-value health AI will allow countless new researchers and companies to flourish. This will in turn build and sustain American leadership in this critical area.

Thank you for this opportunity to address the Committee. I am a physician and researcher at UC Berkeley, where I build and evaluate artificial intelligence (AI) tools to improve health and healthcare. I am also the co-founder of a company, Dandelion Health, and a non-profit, Nightingale Open Science, both dedicated to creating the data assets needed to catalyze research and commercial development in health AI. I am affiliated with a number of other public and private entities, but the views I present are entirely my own, based on my experiences and research.

Rather than speaking in general terms about the promise of AI, let me tell you about a specific example of how AI can transform the health system. I'll then highlight what I've learned along the way about how the federal government can accelerate progress.

Case study: AI can lower healthcare costs and improve quality—a rare combination

My research tackles the problem of sudden cardiac death. Every year, over 300,000 Americans drop dead with no warning, the result of fatal cardiac arrhythmias.[1] This problem showcases both the highs and lows of our health system. On the one hand, medical research has delivered a miraculous cure: a defibrillator implanted in the heart can detect and terminate otherwise fatal arrhythmias. On the other hand, doctors struggle to get defibrillators into the right patients. Exhibit A is the hundreds of thousands who die every year without defibrillators, because doctors never suspected their risk. Exhibit B is another astonishing fact: two-thirds of defibrillators never fire.[2] Doctors implant them with the goal of stopping the arrhythmias that cause death, but these patients never go on to develop the arrhythmia. So we perform invasive, risky surgery and spend up to $50,000 per patient, but get absolutely no benefit.

The traditional debate here is: are we implanting too many defibrillators, or too few? But the existence of both over- and under-use suggests this old saw misses the point. The better question is: how do we get defibrillators into those who need them, and spare the hearts and pocketbooks of those who don't?[3]

This is where AI can help. My colleagues and I have trained an AI system to predict the risk of sudden cardiac death, using patients' electrocardiogram (ECG). By looking only at a waveform—a tracing of the electrical field generated by the beating heart—the AI system makes risk predictions far more accurate than anything doctors use today. Working with Providence, one of the nation's largest health systems, and the International Heart Institute of Montana, we are already starting to test how this AI system can eliminate wasteful procedures and get precious defibrillators to patients who truly need them.

As you can see from this example, AI is fundamentally different from other new technologies. Historically, new tools tend to increase costs: the tools themselves are expensive, but more importantly, deciding who should get which of the myriad new tools becomes ever more complex. As with defibrillators, this generates both over-use and under-use at large scale in the health system, lowering quality and increasing costs. AI, by contrast, is not just a new technology: it is a way to produce critical information that can help us make better decisions about technology. This lets us sidestep the usual dilemma—improve quality *or* reduce cost?—and instead do both. Defibrillators are far from the only example of where AI can improve quality and reduce cost: there is a growing list of examples from my own work and that of colleagues, ranging from heart attack[4] to cancer screening[5] and beyond.

I hope the potential to save lives *and* cut wasteful spending is as exciting to you as it is to me. If so, let me highlight a few lessons I've learned along the way about where the federal government can help make progress towards these goals by accelerating innovative health AI.

1. Multiple federal agencies can help bring high-value health data online for AI

The US has one of the largest strategic reserves of high-quality health data in the world. Unfortunately, much of these data—especially the data most useful for AI: complex physiological data like images and waveforms; clinical notes and reports—are inaccessible to those who want to build health AI products. It took a full 10 years for me to build up the data assets needed for my research on sudden cardiac death. And ultimately, I did the work with collaborators in Sweden, where—despite European data regulations—it was quicker and easier than in the United States. These facts speak to a major threat to American leadership in health AI. There are several things the government can do to fix this problem.

First, vendors of electronic health records commonly restrict access to AI-ready physiological data, or impose integration fees. This hampers innovation and is inconsistent with the goals of the 21st Century CURES Act. The Office of the National Coordinator for Health Information Technology (ONC) should vigorously enforce its interoperability and information-blocking rules to ensure that vendors must make high-value data (physiologic streams and images) available at reasonable, cost-based fees.

Second, there is great uncertainty around the legal standard for deidentification of images and notes under HIPAA Safe Harbor. This means that companies must use Expert Determination, engaging consultants to manually create bespoke data strategies—despite the fact that high-quality turnkey technical solutions exist. HHS should identify rapid, reliable methods for deidentification of AI-ready data like images and notes, and issue guidance that these methods meet the Safe Harbor standard. This would ensure minimal risks to privacy while accelerating the development of clinical AI tools.

Third, federal agencies that hold health data should publicly report how long it takes them to deliver it to those requesting it to advance patient care, and set aggressive timeline targets for delivery—days, not months. These data are a public good: they are paid for by all of us, via taxes and insurance, and they have huge value to all of us, via scientific discovery and clinical AI tools. By the end of this session, hundreds of Americans will have experienced sudden cardiac deaths preventable with AI; breast cancers are metastasizing right now that AI could have caught at the patient's last mammogram. Delays in data access mean lives lost. And yet, the myriad layers of permissions and approvals mean that accessing any federal data takes months or years. This discourages and delays innovation. It also opens the door for ideological bias: the government should not decide which questions get asked, but rather ensure that any question can be asked, to foster innovation while minimizing risks.

None of this contradicts the need to preserve patient privacy. In fact, modern data management means that some of the most sensitive data in the world are now stored in secure computing infrastructures, where they can *both* be easily accessed by those who need them, and highly protected from misuse: with multiple layers of operational and legal protections, and draconian law enforcement for those who break

the law. This would both allow rapid access, and enhance privacy and security: instead of sending CD-ROMs in the mail, or adding permissions and approvals, we must ensure access is safe by design.

More broadly, the extreme risk aversion around accessing data is a major threat to the United States' goal of winning the AI race. In some parts of the world, draconian regulations prevent anyone from doing nearly anything with any data. In other parts of the world, a privileged few have unfettered access to any data, with no regard to ethics or privacy. The United States must chart a middle ground between these two extremes if it wishes to win: access to health data must be easy and fast for those who want to push forward medicine in ethical ways, and privacy must be maintained using today's highly secure, cutting-edge technologies. This would contribute to a key priority outlined in the President's AI Action Plan: Build World-Class Scientific Datasets.

2. FDA evaluation of AI should be simple and transparent: Does AI predict what it says it does?

A key driver of AI progress is transparency. Public benchmarks like ImageNet or MMLU provide an objective yardstick for anyone to track the predictive performance of new algorithms. This has kept AI developers honest, produced intense competition and innovation, and catalyzed the tremendous recent advances we've all seen in image and language models.[6,7]

By contrast, when the Food and Drug Administration (FDA) evaluates health AI, it relies on complex, opaque processes that require large amounts of time and money, for both the agency and the applicant. It's important to recognize that the FDA was dealt a tough hand here. It is required to regulate health AI under the Federal Food, Drug, and Cosmetic (FD&C) Act (passed in 1938 and amended in 1976): health AI is, broadly speaking, a technology used for medical purposes that is not a drug. But the only available regulatory approach to such technology is to treat it like a physical device. The FDA valiantly attempted to square this circle by introducing the idea of "software as a medical device." But no amount of conceptual maneuvering can get around the fact that software is obviously not a medical device, and should not be regulated as such. Congress tried to help in the CURES Act of 2016, by carving out some software from this nonsensical regulation. The intent of CURES was correct, but with the benefit of hindsight, the carve-outs it introduced were arbitrary and confusing: for example, an AI that uses ECG waveforms must be regulated as a device, but an AI that uses laboratory results *for the exact same purpose* is exempt; the concept of physician "independent review" of recommendations is vague.

The result of this regulatory morass is an approval process that requires lengthy back and forth with FDA staff, costs nearly $1m for an army of consultants and regulatory experts to complete, and takes months or even years. This friction has perverse consequences, most obviously narrowing the pool of AI that can ever touch patients. It also benefits entities with deep expertise in government relations rather than those with the best AI products. It encourages locking up insights in the private sector: if my lab wanted to get our sudden cardiac death AI approved, we would need to form a company simply to raise money for FDA clearance, working against the President's AI Action Plan: Encouraging Open-Source and Open-Weight AI. Finally, it torques the market to serve the needs of big businesses—large hospitals and insurers—rather than patients. At the same time, recent reviews of FDA-authorized AI devices have found that many rely on relatively small, single-site datasets and provide limited information on real-world performance, calibration, or bias.[8,9] Thus despite its cost, the current FDA process does not consistently produce the transparent, benchmark-style evidence that would best support decisions.

The best solution would be to amend CURES to clarify that the exemption from burdensome device regulation applies to all AI predictive algorithms—even those that use physiological signals, or image data—provided they meet risk-based performance and safety standards.

Even without new legislation, the FDA should recognize that the basis for regulating AI—whether we call it a device or something else—is rigorous evaluation of the information it produces. AI systems make predictions on some measurable quantity. As a result, AI is quite easy to regulate: the quality of AI predictions for a given task can be straightforwardly evaluated using data.[1] Such evaluation data must be representative of the whole population—not just the handful who live near academic medical centers in Boston, Rochester, and San Francisco, where most current tools are trained.[10] My own prior work has shown that proper evaluation can catch and prevent a range of racial and other biases, ensuring that AI works for everyone, not just a privileged few. Critically, evaluation data must also be new: the AI cannot have been exposed to these data in its training process. But most currently FDA-approved tools provide few such guarantees, meaning 'overfitting' could be a major problem.

The FDA could build capacity to perform such evaluations itself. Or it could work with professional societies, who bring additional depth of expertise. For example, the company I co-founded, Dandelion, is working with the American Heart Association to create a robust evaluation ecosystem for cardiovascular AI.[11] Because Dandelion has created one of the largest health datasets in the world, linking complex physiological data to longitudinal patient outcomes, this joint effort can transparently measure quality. This work has taught us critical lessons, on how to assemble data and use it for evaluation, that will one day help us replicate this model at scale for a broader evaluation pipeline, consistent with the President's AI Action Plan priority of Building an AI Evaluations Ecosystem.

3. CMS should pay for algorithms that help reduce costs and improve quality

At your next doctor's visit, you may see your doctor using AI scribe software to make her more efficient (e.g., write her note, automate administrative tasks). But it's unlikely that she will be using any clinical AI tools to improve your health. The problem is not slow adoption of clinical AI tools: it's the lack of high-quality tools entering the development pipeline to begin with. The private sector is under-investing in health AI, in large part because of paralyzing uncertainty over the prices that insurers will pay.

CMS is, by a wide margin, the largest purchaser of health technology in the world. This gives it enormous power to shape the future of health AI in directions that serve patients, and sustain American leadership. CMS should be willing to pay for AI that creates value. For example, an AI system that makes an early diagnosis of heart attack or breast cancer, in addition to saving lives, can also save money for hospitalizations and expensive treatments for late-stage disease. AI that saves lives and reduces costs is worth paying for, and the price can be set based on the value it creates.

[1] This applies only to the kinds of evaluation currently performed by the FDA around *predictive accuracy*. Naturally, to understand the *effect* of an AI tool on patient health outcomes in practice, a randomized trial is needed.

CMS and the FDA are already taking exciting steps in this direction. The new Advancing Chronic Care with Effective, Scalable Solutions (ACCESS) Model will implicitly reward health systems that use AI tools to optimize chronic disease care. The FDA will create a new pathway for AI-enabled devices in this model that relaxes up-front approval requirements. These are very positive steps forward.

But CMS can and should go further. By defining new service codes for AI, that are not specific to any one vendor's product, CMS could send a powerful signal about the kinds of products it wants to see, and how much those products are worth. CMS has clear authority to define such codes with no new legislation or sponsorship from professional societies. CMS along with FDA can jointly articulate specific evidentiary standards—e.g., a certain level of predictive performance against a given outcome—needed for approval and payment under this new code. This will unlock private investment by resolving uncertainty: with a clear goalpost to aim for, the market will be pointed in the right direction.

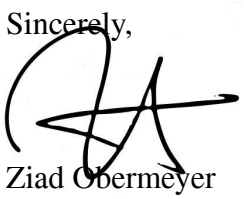
Conclusion

I am very optimistic about the potential of health AI to transform our health system, reduce its enormous costs, and improve its quality. Accelerating progress towards these goals is critical for both patients and the financial solvency of the United States. Health is also a critical arena for American leadership in AI more broadly: in addition to the intrinsic importance of health, the vastness of health data and the ability to learn from real outcomes at scale makes it a fertile development ground for new technical approaches.

The stakes are not abstract. Every delay in getting high-quality data into secure environments, or in approving and paying for effective AI tools, shows up as real patients who die of preventable heart attacks, strokes, and cancers. At the same time, other countries are moving quickly to align their data, evaluation, and payment systems around AI.

The United States has deep advantages: large and diverse health-care datasets; world-leading universities and companies; and a Medicare program whose payment policies shape the world market for medical technology. If we make it easier to use data safely, if we focus regulation on transparent evidence of how AI performs in the real world, and if we are willing to pay for tools that improve outcomes and lower costs, we can both save lives and secure American leadership in health AI. If we do not, those lives and that leadership will go elsewhere.

Sincerely,



Ziad Obermeyer
Associate Professor and Blue Cross of California Distinguished Professor, UC Berkeley
Nonresident Senior Scholar, USC Schaeffer Institute
Research Associate, National Bureau of Economic Research

References

1. Al-Khatib, S. M. *et al.* 2017 AHA/ACC/HRS guideline for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: Executive summary: A report of the American college of cardiology/American heart association task force on clinical practice guidelines and the heart rhythm society: A report of the American college of cardiology/American heart association task force on clinical practice guidelines and the heart rhythm society. *Circulation* **138**, e210–e271 (2018).

2. Merchant, F. M., Quest, T., Leon, A. R. & El-Chami, M. F. Implantable cardioverter-defibrillators at end of battery life: Opportunities for risk (re)-stratification in ICD recipients. *J. Am. Coll. Cardiol.* **67**, 435–444 (2016).

3. Baicker, K. & Obermeyer, Z. Overuse and underuse of health care: New insights from economics and machine learning. *JAMA Health Forum* **3**, e220428 (2022).

4. Mullainathan, S. & Obermeyer, Z. Diagnosing physician error: A machine learning approach to low-value health care. *Q. J. Econ.* **137**, 679–727 (2022).

5. Yala, A., Lehman, C., Schuster, T., Portnoi, T. & Barzilay, R. A deep learning mammography-based model for improved breast cancer risk prediction. *Radiology* **292**, 60–66 (2019).

6. Donoho, D. Data science at the singularity. *Harvard Data Science Review* **6**, (2024).

7. Hardt, M. The emerging science of machine learning benchmarks. *https://mlbenchmarks. org* (2025).

8. Chouffani El Fassi, S. *et al.* Not all AI health tools with regulatory authorization are clinically validated. *Nat. Med.* **30**, 2718–2720 (2024).

9. Muralidharan, V. *et al.* A scoping review of reporting gaps in FDA-approved AI medical devices. *NPJ Digit. Med.* **7**, 273 (2024).

10. Kaushal, A., Altman, R. & Langlotz, C. Geographic distribution of US cohorts used to train deep learning algorithms. *JAMA* **324**, 1212–1213 (2020).

11. American Heart Association. The American Heart Association AI Assessment Lab, Powered by Dandelion Health. *www.heart.org* https://www.heart.org/en/professional/quality-improvement/aha-ai-assessment-lab.