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

AI will transform medicine and the health care system—
for better or for worse, depending on how it is built and applied.

February 5, 2024

Thank you for the opportunity to address the Committee. I am a professor and researcher at Berkeley, but my work on AI is grounded in another part of my identity, as a practicing emergency physician. Seeing patients, from academic hospitals in Boston to Tséhootsooí Medical Center in Fort Defiance, Arizona, has given me a window into the miracles of modern medicine—awe-inspiring innovations in the diagnosis and treatment of disease, that have extended and improved life for millions.

Getting those miraculous tests and treatments to the right patient at the right time, though, is difficult. It requires processing enormous amounts of information, much of it imperfect and uncertain; errors have life-and-death consequences. Senators, I cannot imagine what it is like to do your job, but my guess is that you face situations like that every day. Throughout my 10 years of practicing medicine, I have agonized over missed diagnoses, futile treatments, unnecessary tests, and more. The collective weight of these errors, in my view, is a major driver of the dual crisis in our health care system: suboptimal outcomes at very high cost.[1] AI holds tremendous promise as a solution to both problems.

By helping doctors and others in the health care system make better decisions, I believe AI can both improve health and reduce costs—a rare combination. Let me share a few concrete examples drawn from my own work, to convey why I am so optimistic.

[1] While economists and policy makers have traditionally focused on the role of misaligned financial incentives, a large and growing body of recent research indicates much of our health care system's inefficiency has its roots in human error. See Baicker, K., Mullainathan, S. and Schwartzstein, J., 2015. Behavioral hazard in health insurance. *The Quarterly Journal of Economics*, *130*(4), pp.1623-1667.

In the US alone, 300,000 people experience sudden cardiac death every year.[2] What makes these events so tragic is that many of them are preventable: had we known a patient was at high risk, we would have implanted a defibrillator in her heart, to terminate the potential arrythmias that cause sudden death, and save her life. Unfortunately, we are very bad at knowing who is at high risk. It's not just that we miss 300,000 opportunities every year to implant defibrillators and prevent those deaths. Even when we do implant defibrillators, we often do so in the wrong patients: up to one-third of patients end up never needing their defibrillator, meaning we increased their risk of complications and wasted resources without ever delivering a life-saving shock.[3]

Fifteen years ago, when I was in medical school, I was stunned by these numbers. Today, working with colleagues in the US and Sweden, we have trained an AI system to predict the risk of sudden cardiac death using just the waveform of a patients electrocardiogram (ECG). It performs far better than our current prediction technologies, based largely on human judgment. This means we have the potential to both save more lives *and* reduce waste, by ensuring that precious defibrillators are implanted in the right patients. It's rare to have an opportunity to both improve quality and reduce cost; normally we must choose. AI is a transformative new way for us to sidestep this dilemma entirely, and rebuild our health care system on a foundation of data-driven decision-making.

This principle—better human decisions through AI-driven predictions—extends far beyond sudden cardiac death. We've found similar opportunities to improve quality and reduce cost in settings ranging from invasive testing for heart attack[4] to mammograms for breast cancer prevention.[5] AI can also help diagnose social vulnerability: we have promising early results showing AI can find subtle signs of interpersonal violence in x-rays. This means physicians can help recognize victims of violence when they come to seek help in the ER—instead of missing those opportunities, as happens all too often today—and social services can connect them to the resources they need.[6] AI is also starting to drive innovation in the science of medicine, for example, by discovering entirely new classes of antibiotics in drug libraries that were passed over for decades by human researchers.[7]

Despite my great optimism, I worry that without concerted effort from researchers, the private sector, and government, AI may be on a path to do more harm than good in health care. So I'd also like to provide an example of how AI can go wrong. Working on this problem has taught me a lot about we can work together to ensure that AI systems are safe.

---

[2] Huikuri, H.V., Castellanos, A. and Myerburg, R.J., 2001. Sudden death due to cardiac arrhythmias. *New England Journal of Medicine*, 345(20), pp.1473-1482.
[3] Moss, A.J., Greenberg, H., Case, R.B., Zareba, W., Hall, W.J., Brown, M.W., Daubert, J.P., McNitt, S., Andrews, M.L. and Elkin, A.D., 2004. Long-term clinical course of patients after termination of ventricular tachyarrhythmia by an implanted defibrillator. *Circulation*, 110(25), pp.3760-3765.
[4] Mullainathan, S. and Obermeyer, Z., 2022. Diagnosing physician error: A machine learning approach to low-value health care. *The Quarterly Journal of Economics*, *137*(2), pp.679-727.
[5] Daysal, N.M., Mullainathan, S., Obermeyer, Z., Sarkar, S.K. and Trandafir, M., 2022. An Economic Approach to Machine Learning in Health Policy. *CEBI Working Paper*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4305806
[6] Williams, B., Oto, A., Ludwig, J., Graber, R., Obermeyer, Z. and Mullainathan, S., 2023. Making the invisible epidemic visible. https://www.brookings.edu/articles/making-the-invisible-epidemic-visible/
[7] Stokes J.M., et al. A deep learning approach to antibiotic discovery. *Cell*. 2020 Feb 20;180(4):688-702.

Five years ago, my colleagues and I uncovered evidence that a family of poorly-designed AI algorithms, built and used in both public and private sectors, contained large-scale racial bias.[8] These algorithms had a laudable goal: to identify patients at high risk of future health problems—exacerbations of chronic conditions like heart failure, diabetes, etc. The AI's predictions are used by health systems around the world to decide who gets access to extra help, in the form of 'care management' programs. In theory, this is a great use of AI, because these programs are a win-win: high-risk patients get the help they need to manage chronic conditions, reducing future flare-ups and complications; and the health care system saves the money it would have spent on the resulting ER visits and hospitalizations.

Unfortunately, a subtle-seeming choice in the AI's design caused untold harm: a gap between what the algorithms were *supposed* to predict (health care needs) and what they *actually* predicted (health care costs). The AI's goal was to identify patients with high future health needs. But AI is extremely literal—it predicts a specific variable, in a specific dataset—and there is no variable available called 'future health *needs*.' So instead, the AI developers chose to predict a proxy variable that *is* present in health datasets: future healthcare *costs*. Spending on health care seems like a reasonable proxy for health needs. After all, sick people generate health costs. But because of discrimination and barriers to access, underserved patients who need health care often don't get it. This means Black patients—and also poorer patients, rural patients, less-educated patients, and all those who face barriers to accessing health care when they need it—get less spent on their health care than their better-served counterparts, even though they have the same underlying health conditions.[9] Low costs do not necessarily mean low needs.

Tragically, the AI ignored these simple facts. It predicted—accurately—that Black patients would generate lower costs, and thus deprioritized them for access to help with their health. The result was racial bias that affected important decisions for hundreds of million patients every year. Senator Wyden, I was heartened by the letters that you and Senator Booker sent to executives at major insurance companies in the wake of that study—I believe that had a great impact. Unfortunately, many of the biased algorithms we studied remain in use today. And similar dynamics were highlighted by a recent investigation of AI products used to deny claims[10]: in all these cases, AI learns from historical data, with all its biases and inequities, and encodes those past practices in policy. So those underserved patients whose claims have been denied by humans in our past datasets—often for unjust reasons—will have their claims denied by AI at scale, forever, unless we can re-align AI with our society's goals.

Fortunately, there are a number of specific things that programs under this Committee's jurisdiction can do to ensure that AI produces the social value we all want. I believe that Medicare, Medicaid, CHIP, and child welfare programs stand to realize enormous benefits from AI: well-designed products can both improve the quality of services and reduce their cost. As a result, these programs

[8] Obermeyer, Z., Powers, B., Vogeli, C. and Mullainathan, S., 2019. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), pp.447-453.
[9] While our work focused on demonstrating one specific bias—Black vs. White patients—this is not an issue of race alone: any populations with a wedge between the care they need and the care they get will be similarly affected.
[10] Ross, C. and Herman, B., 2023. Denied by AI. *STAT News.* March, 13, 2023. https://www.statnews.com/2023/03/13/medicare-advantage-plans-denial-artificial-intelligence/

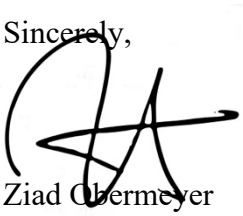
should be willing to pay for AI—but they should not simply accept the flawed products that the market often produces. Rather, they should take advantage of their market power to articulate clear criteria for what they will pay for, and how much. I believe this will harness the tremendous innovative power of the market, and ensure it is pointed in the right direction.

Based on my research, as well as my work with federal regulators and state Attorneys General, I believe that programs in the Committee's jurisdiction should explicitly evaluate AI algorithms for reimbursement on a small set of targeted criteria. AI developers must be transparent about the output of their algorithms. If an algorithm predicts health care costs, the developer should not be able to claim that it predicts "health risks" or "health needs"—unfortunately, many cost-predictors currently do exactly this. Algorithms' outputs should be evaluated for accuracy in a completely independent dataset, both overall and in protected groups, in keeping with good machine learning practice.[11] I emphasize that this approach focuses on the *output* of the algorithm—the accuracy of its predictions on a transparently-stated target—and thus does not require 'opening the black box': algorithms can be evaluated simply based on the predictions they produce. This avoids compromising trade secrets, and means purchasers (and similarly, regulators) do not need to evaluate the *inputs* of algorithms, or understand the many reasons why they might be biased—technical problems with a complex model, non-representative training data, use of an explicit race correction, etc. Instead, we can focus on one simple question: is the algorithm predicting what it's supposed to predict, accurately and equitably?[12] Finally, AI products should be valued and reimbursed according to established principles from health economics and outcomes research. If an AI results in an earlier diagnosis of heart attack or breast cancer, for example, that generates value to patients in the form of life-years, and to the health care system in the form of downstream costs avoided. The sooner public programs lay out what they are looking for, the sooner the market can deliver safe and effective AI products to solve the urgent problems they face.

I should note that my applied work has resulted in collaborations with a number of public and private entities, but the views I present are entirely my own, based on my experiences and research.

Many thanks again for this opportunity. I look forward to answering your questions.

Sincerely,



Ziad Obermeyer

[11] Accessing data for such evaluations is a non-trivial problem, but there are emerging solutions. For example, a company I co-founded, Dandelion Health, offers a free public service for the evaluation of algorithm performance and equity. This service, which is philanthropically supported by the Gordon and Betty Moore Foundation and the SCAN Foundation, allows any AI developer to securely upload the algorithm to Dandelion's computing environment. Dandelion will run the algorithm on its diverse national dataset and deliver back a report on the algorithm's performance, both overall and across key geographic, racial, ethnic, age, gender, and socioeconomic groups. More details are at https://dandelionhealth.ai/validation.

[12] This is analogous to the process by which the FDA regulates information about drugs: pharmaceutical companies must be transparent about the primary outcome a drug is intended to improve, and the drug's impact on that outcome is assessed in a rigorous randomized trial.