Artificial Intelligence and Decision-Making in Healthcare: Prediction or Preferences?

M. Kate Bundorf
Duke University and NBER

Maria Polyakova
Stanford University and NBER

December 10, 2022

People make most of their decisions in the context of uncertainty. Because each of our decisions can have many possible outcomes, we can never know with certainty how the choices we make today will affect us tomorrow. In the canonical economic model, people make decisions under uncertainty by assessing both the likelihood of potential outcomes associated with alternative courses of action and their utility of each outcome and then choosing the alternative with the greatest expected utility (Morgenstern and Von Neumann, 1953; Savage, 1954). In this model, people need two types of information to make a choice among alternatives: the distribution of future outcomes associated with each alternative and how they value each of the possible future outcomes. Here we refer to the first as "prediction" and the second as "preferences." This distinction is consistent with the theoretical model developed by Agrawal et al. (2018a) which differentiates between prediction and "judgment" in understanding the impact of artificial intelligence on decision-making.

The distinction between prediction and preferences is central for understanding how artificial intelligence will impact health care treatment decisions. Uncertainty is at the heart of clinical decision making; both diagnosis and treatment are probabilistic. A well-working AI tool can be enormously beneficial for understanding the distribution of potential diagnoses and treatment outcomes. As **Obermeyer and Mullainathan** discuss in this volume, there are important challenges in generating predictions and deploying them in clinical settings including incomplete or biased data, challenges of communicating algorithmic predictions to human experts, and barriers to implementation. Overall, however, prediction is essential for clinical medicine and prediction technologies can reduce the cost or increase the quality of this important input into decision making (Agrawal et al., 2018b).

The ability of artificial intelligence to incorporate patient preference into decision making, however, is less straightforward. Clinical decision making requires interpreting information about probabilistic outcomes for patients. In many cases, this requires making trade-offs among uncertain, even if well-predicted outcomes. Patient preferences are critical for making these trade-offs. A challenge, which we explored empirically in the context of expert advice for the choice of health insurance plans (Bundorf et al., 2019), is that patients may not have full information about their own preferences. Analogously, in many clinical settings, patients may not have well-formed preferences over treatment outcomes. Indeed a key component of decision aids for health care treatment decisions is helping patients understand “either explicitly or implicitly, the value they place on potential benefits and harms” (Stacey et al., 2017). Physicians and other types of clinicians, due to their experience and expertise, naturally play a role in helping patients formulate their preferences. Can AI do this instead?

We propose that existing AI approaches to advising consumers based on their preferences are less well suited to health care decision-making than in other contexts. A common data-driven approach for capturing preferences is to link the decisions of millions of consumers with objective measures of choice satisfaction to predict whether someone with your characteristics is likely to be satisfied with a particular option. This “consumers-like-you” approach, however, has limitations for decision making in health care. It assumes consumers are making good decisions when we have lots of evidence that they don’t (Bernheim et al., 2019). Indeed, if consumers “like you” made great decisions, then they wouldn’t need anyone’s advice.

An alternative approach is to provide machine-based expert advice. This type of advice, however, imposes expert preferences on individual choices—in essence, assuming away preference heterogeneity across patients, the type of information we wanted to incorporate in the first place. Whether such computerized experts can replace today’s physicians depends on how much we think physicians are able to tailor their expert advice to each patient. Physicians are under increasing pressure to provide that type of personalized advice due to a movement toward “shared decision making,” in which patients are full partners with clinicians in health care decision-making as opposed to more passive recipients of expert advice (Resnicow et al., 2022).

The objective of shared decision making is to ensure that treatment decisions more fully reflect patient values. Patients may prefer tailored even if imperfect advice to a recommendation that is only right for an average patient.

As long as AI remains limited in its ability to “predict” patient preferences, it will remain a complement rather than a substitute for many physicians. As **Dranove and Garthwaite** in this volume recognize, medical specialties who focus less on patient relationships may be more substitutable by AI. Indeed, as AI tools become better at predicting the probabilities of different potential outcomes, the role of physicians is likely to shift more to the domain of helping patients formulate their preferences in response to AI-generated information.

To summarize, what are the implications of the distinction between prediction and preferences for the role of AI in healthcare decision making? AI’s current state as a data prediction exercise may limit its ability to inform decisions that are highly dependent on preferences or decisions for which patient preferences vary substantially. Clinicians who help patients translate prediction into decisions by incorporating patient preferences will have skills which are complementary to the strengths of AI. Paradoxically, this implies that in contrast to the fear of AI replacing the medical profession promulgated by some, increased integration of AI into a clinician’s daily routine may not replace physicians, but rather incentivize physicians to focus on what medical students often say motivated them to choose medicine - listening to the patient.

References

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb, “Prediction, judgment, and complexity: a theory of decision-making and artificial intelligence,” in “The economics of artificial intelligence: An agenda,” University of Chicago Press, 2018, pp. 89–110.

_ , _ , and _ , *Prediction machines: the simple economics of artificial intelligence*, Harvard Business Press, 2018.

Bernheim, B Douglas, Stefano DellaVigna, and David Laibson, *Handbook of Behavioral Economics-Foundations and Applications 2*, Elsevier, 2019.

Bundorf, M Kate, Maria Polyakova, and Ming Tai-Seale, “How do humans interact with algorithms? Experimental evidence from health insurance,” *NBER Working Paper*, 2019.

Morgenstern, Oskar and John Von Neumann, *Theory of games and economic behavior*, Princeton university press, 1953.

Resnicow, Ken, Delwyn Catley, Kathy Goggin, Sarah Hawley, and Geoffrey C Williams, "Shared decision making in health care: theoretical perspectives for why it works and for whom," *Medical Decision Making*, 2022, *42* (6), 755–764.

Savage, Leonard J, *The Foundations of Statistics*, John Wiley & Sons, 1954.

Stacey, Dawn, France Légaré, Krystina Lewis, Michael J Barry, Carol L Bennett, Karen B Eden, Margaret Holmes-Rovner, Hilary Llewellyn-Thomas, Anne Lyddiatt, Richard Thomson et al., "Decision aids for people facing health treatment or screening decisions," *Cochrane database of systematic reviews*, 2017, (4).