Groundhog Day for Medical Artificial Intelligence

BY ALEX JOHN LONDON

Following a boom in investment and overinflated expectations in the 1980s, artificial intelligence entered a period of retrenchment known as the "AI winter." With advances in the field of machine learning and the availability of large datasets for training various types of artificial neural networks, AI is in another cycle of halcyon days. Although medicine is particularly recalcitrant to change, applications of AI in health care have professionals in fields like radiology worried about the future of their careers and have the public tittering about the prospect of soulless machines making life-and-death decisions. Medicine thus appears to be at an inflection point—a kind of Groundhog Day on which either AI will bring a springtime of improved diagnostic and predictive practices or the shadow of public and professional fear will lead to six more metaphorical weeks of winter in medical AI.

AI applications in health care evoke strong reactions not only because of the specter of machines making life-and-death decisions but also because of the prospect of machines encroaching on realms of decision-making revered as the province of expert professionals. It is no accident that Aristotle uses medical decision-making as a paradigm for practical rationality (*phronesis*) and the expert physician as an analogue for the practically wise person (*phronimos*). Reverence for medical expertise as combining scientific knowledge and practical application is perhaps as old as medicine itself. The medical art is revered as requiring and embodying the perfection of aspects of ourselves that are in some sense the highest and most noble: practical rationality informed by knowledge, expressed in judgment, suffused with compassion.

There is, undoubtedly, profound truth in this portrait of medicine. Medicine blends science and art, and the skills that constitute the art must be learned in practice, by doing. But we must also beware of romanticizing human judgment. For decades, psychologists and decision scientists like Paul Meehl and Robyn Dawes have argued that simple algorithms for specific diagnostic and prediction tasks of-

Rational AI?

ten outperform clinicians. Despite robust evidence to this effect, professionals routinely overestimate their ability to perform such tasks and underestimate the value of actuarial methods for making health care decisions. Precisely because medical diagnostic and prediction decisions are intimately bound up with matters of life and death, perpetuating the neglect of highly accurate algorithmic decision tools is not a benign deference to professional prerogative. It is a potentially lethal hubris whose tithe is exacted in avoidable morbidity and mortality.

In many respects, contemporary AI systems are the more computationally robust and statistically sophisticated progeny of the relatively simple decision tools studied by Meehl and colleagues. Increasingly, such systems are leveraging previously unimaginable volumes of information to make diagnostic and predictive judgments. And in some cases, they do so more quickly, with greater specificity and sensitivity, than humans do. When human and machine intelligence seek to perform the same task, we must be prepared to use whichever approach offers the best prospect of more accurate and reliable performance on that task.

Having said that, it is crucial that the accuracy and reliability of AI systems be validated through rigorous, independent testing and that factors affecting their performance be transparently communicated and continuously monitored. Regulatory agencies, such as the U.S. Food and Drug Administration, have a critical role to play in enforcing strict standards for such testing and disclosure. And if health professionals are to intelligently incorporate AI systems into the art of medicine, medical schools and health systems will have to develop and foster a more robust culture of statistical literacy and fluency.

In the end, romanticism about either human judgment or machine intelligence is a recipe for failure. Ironically, the most difficult challenge posed by the springtime of AI in medicine remains a fundamentally human one: to know ourselves and our limitations, to control hubris in its many forms, to rely when necessary on the assistance of others (even if they are machines), and to strive always to do better.

Alex John London is the Clara L. West professor of ethics and philosophy and the director of the Center for Ethics and Policy at Carnegie Mellon University DOI: 10.1002/hast.842