



## 8. AI in Medicine

Angeliki Kerasidou and Xaroula (Charalampia) Kerasidou

### Abstract

AI promises major benefits for healthcare. But along with the benefits come risks. Not so much the risk of powerful super-intelligent machines taking over, but the risk of structural injustices, biases, and inequalities being perpetuated in a system that cannot be challenged because nobody actually knows how the algorithms work. Or, the risk that there might be no doctor or nurse present to hold your hand and reassure you when you are at your most vulnerable. There are many initiatives to come up with ethical or trustworthy AI and these efforts are important. Yet we should demand more than this. Technological solutionism and the urge to “move fast and break things” often dominate the tech industry but are inappropriate for the healthcare context and incompatible with basic healthcare values of empathy, solidarity, and trust. So how can such socio-political and ethical issues get resolved? It is at this juncture that we have the opportunity to imagine different futures. Using Grace's fictional story, this chapter argues that in order to shape the future of healthcare we need to decide whether, to what extent, and how—under what regulatory frameworks and safeguards—these technologies could and should play a part in this future. AI can indeed improve healthcare but, instead of casting ourselves loose and at the mercy of this seemingly inevitable technological drift, we should be actively paddling towards a future of our choice.

### Case study: Grace's Future

After several weeks of experiencing abdominal pain, Grace pays a visit to her doctor. There, a number of samples, including blood samples, are collected, and an MRI is performed. The examination data is then combined with her health records—collected both during medical visits and in real time through her smartphone. The doctor informs Grace that the results will be ready in 72 hours and the diagnosis will be promptly communicated to her, along with an action plan.

In the three days she awaits the results, Grace reflects on how things have changed since she was young. Although she still describes these medical visits as ‘going to the doctor’, she didn’t see a doctor in the flesh. From the virtual receptionist, who always greets her a little too cheerfully, to the polite but slightly bossy-sounding bot taking her medical history, to the MRI scanner barking instructions, the majority of tasks are now performed by intelligent machines. The only role left for humans is the delicate procedure of taking intravenous blood samples.

---

This is the manuscript of a chapter that has been accepted for publication by [Oxford University Press](#) in the book *Future Morality* edited by David Edmonds published in September 2021.

© the several contributors 2021

Monographs, or book chapters, which are outputs of Wellcome Trust funding have been made freely available as part of the [Wellcome Trust's open access policy](#)

Grace is thankful that phlebotomists have survived the introduction of AI in medicine because she has always been afraid of needles and is reassured by a human presence. She tries to start a conversation; “How’s your day been?”—“Busy! I have twenty more samples to take before the end of my shift”, replies the nurse without lifting his eyes.

The results duly appear on Grace’s smartphone within the promised period. It’s not good news. There’s a cancer diagnosis accompanied by an action plan. “Based on your individual results and the latest system upgrade”, reads the message, “the recommendation is for chemotherapy to commence immediately. Click *here* and *here* for more information about what this means for you. Click *Yes* if you consent.” The message explains that, following consent, appropriate healthcare facilities (hospital and pharmacy) will be informed and that details of the dates and times for treatment will follow.

Grace wonders how good the latest upgrade is. Even after all these years, data on some ethnic groups are incomplete and misdiagnoses common. She once heard the Secretary of Health proclaiming: “AI will benefit all of us! It will deliver a truly personalised and efficient healthcare system, saving tax-payer money.” Grace is not so sure. She could request to see a specialist for a face-to-face consultation, but the waiting list is long, and she cannot afford to pay out-of-pocket for such a ‘luxury’. She clicks *Yes*, and heads back to work.

## Augmented Patient-Centred Care

In recent decades, the doctor–patient relationship has undergone a transformation—from medical paternalism, where decisions are taken by the doctor with minimal, if any, involvement from the patient, to a model of care that places the patient at its centre.<sup>1</sup> Once, healthcare professionals were perceived as the unchallenged authority and patients as the ill or broken bodies that required fixing. Today, patients are seen as partners in the therapeutic encounter, contributing not only their own knowledge and expertise about their condition, but also their values that might bear upon the care plan. In the patient-centred model of care, effective communication and empathy, and recognition that different people might hold different values and priorities regarding their health, are the new professional and ethical norms.

The introduction of artificial intelligence (AI) into healthcare has the potential to disrupt this model by augmenting some aspects of it and challenging others. AI is often described as being more rational and reliable than humans. The judgement and reasoning of humans can be affected by an array of irrelevant factors, such as the amount of glucose in their bloodstream or the time of the day.<sup>2</sup> In healthcare, where decisions can be, literally, a matter of life or death, augmented rationality and impartiality could prove invaluable qualities. AI systems also have the potential to improve the personalised character of patient-centred care. Through the use of devices such as smartphones and wearables, AI tools could track our movements, our interactions with others, our eating habits and sleeping patterns. This could lead to highly-tailored health advice and treatment without the need for a consultation with a human doctor. Individual values and preferences could be directly entered into the system, and patients could be empowered to take active ownership and control of their health.

---

1 R. Kaba and P. Sooriakumaran, The evolution of the doctor–patient relationship, *International Journal of Surgery*, 5/1 (2007).

2 S. Danziger, J. Levav, and L. Avnaim-Pesso, Extraneous factors in judicial decisions, *Proceedings of the National Academy of Sciences*, 108/17 (2011).

That, at least, is the promise. But studies show that the effectiveness of AI to date has been exaggerated.<sup>3,4</sup> Hopes that machine learning algorithms could overcome the limitations of humans by being more rational, neutral, and objective have been undermined by evidence that such systems can perpetuate human prejudices, amplify bias, and make inaccurate predictions.<sup>5,6</sup> Consider, for example, the app which, when presented with exactly the same set of symptoms, was found to suggest a heart attack as a possible diagnosis if the user was a man, but merely a panic attack if was a woman. This discrepancy was explained by the “heart attack gender gap”: the unsettling finding that the overrepresentation of men in the medical and research data means that women are up to 50 percent more at risk of receiving a misdiagnosis for a heart attack than men.<sup>7</sup> Yet, faced with a technology that cannot be interrogated and yet is touted to exhibit superior decision-making abilities, there is a danger that doctors, nurses, and patients will be reluctant to challenge its recommendations.

AI developers are seeking technological solutions that could diagnose and mitigate bias; algorithms, it is said, can be improved by utilising more and ‘better’ data. However, the implication that social, political, and ethical problems are just mathematical riddles to be computed out, seems naive. Even though reliable data are crucial for safe healthcare, equally important is understanding the social context in which these health data are collected. Consider, for example, the US case of a medical algorithm designed to identify “high risk” patients in need of extra care based on healthcare expenditure data—this was found to dramatically underestimate the health needs of the sickest black patients. While race-blind, the algorithm, which guided the decision-making for millions of Americans, failed to take into account the long-standing health inequalities that black patients face because of barriers to accessing care, lack of insurance, mistrust of the medical system, etc.<sup>8</sup> For these reasons, black patients tend to spend less of their income on healthcare; it does not follow that their need for care is less acute. There are now increasing calls for a more inclusive and engaged agenda around health and AI which could see the patient-centred model of care expand in two ways: first, by establishing healthcare professionals as gatekeepers or mediators between patients and AI systems in order to ensure not only accessibility, but also safety, quality, and fairness of care;<sup>9</sup> second, by enlisting diverse voices and expertise from patient groups, community health workers, nurses, carers, and others in order to better understand the impact of AI technologies on the ground.<sup>10</sup>

3 M. Nagendran, Y. Chen, C. A. Lovejoy, A. C. Gordon, M. Komorowski, H. Harvey, E. J. Topol, J. P. A. Ioannidis, G. S. Collins, and M. Maruthappu, Artificial intelligence versus clinicians: Systematic review of design, reporting standards, and claims of deep learning studies, *BMJ* (2020), 368: m689. PubMed PMID: 32213531.

4 H. Salisbury, Prestidigitation, *BMJ* (2020), 368: m648. PubMed PMID: 32098837.

5 Z. Obermeyer, B. Powers, C. Vogeli, and S. Mullainathan, Dissecting racial bias in an algorithm used to manage the health of populations, *Science*, 366/6464 (October 2019), 447–53. PubMed PMID: 31649194.

6 T. C. Veinot, H. Mitchell, and J. S. Ancker, Good intentions are not enough: How informatics interventions can worsen inequality, *Journal of the American Medical Informatics Association*, 25/8 (August 2018), 1080–8, 10.1093/jamia/ocy052. PubMed PMID: 29788380.

7 S. Das, It’s hysteria, not a heart attack, GP app Babylon tells women, *Sunday Times*, 13 October 2019, available at: <https://www.thetimes.co.uk/article/its-hysteria-not-a-heart-attack-gp-app-tells-women-gm2vxbrqk>.

8 C. Y. Johnson, Racial bias in a medical algorithm favours white patients over sicker black patients, *The Washington Post*, 24 October 2019, available at: <https://www.washingtonpost.com/health/2019/10/24/racial-bias-medical-algorithm-favors-white-patients-over-sicker-black-patients/>.

9 Academy of Medical Royal Colleges, Artificial intelligence in healthcare, January 2019, available at: [https://www.aomrc.org.uk/wp-content/uploads/2019/01/Artificial\\_intelligence\\_in\\_healthcare\\_0119.pdf](https://www.aomrc.org.uk/wp-content/uploads/2019/01/Artificial_intelligence_in_healthcare_0119.pdf).

10 K. Crawford, R. Dobbe, T. Dryer, G. Fried, B. Green, E. Kaziunas, et al *AI Now 2019 Report* (AI Now Institute, 2019), available at: [https://ainowinstitute.org/AI\\_Now\\_2019\\_Report.html](https://ainowinstitute.org/AI_Now_2019_Report.html).

## Trust

Since Grace has consented to the recommended treatment plan, we might be tempted to say that she trusts her doctor to make the right decision for her. But does she? Is her attitude truly one of trust?

If we trust someone to do something for us, we believe they have the skill and knowledge to perform the entrusted action, and goodwill towards us, meaning that they will not try to wrong or hurt us. Trust takes time to establish and can easily be broken, which makes it important that beliefs in skill and goodwill are regularly reinforced through new, positive experiences. In a purely trust relationship, the trustor does not have guarantees, only her justified belief, that the trusted person will confirm her trust. Effectively, a trust relationship is a relationship of vulnerability, in which the trustor becomes vulnerable towards the trustee.

Grace seems to trust her doctor and this seems unsurprising. Medicine is one of the most trusted professions. We allow doctors to poke and prod us, to cut us open and stitch us up, we tell them our private thoughts, concerns, and even dark secrets. We trust doctors and nurses because we believe they know how to care for us; they have the specialist skill and knowledge, but also goodwill. And by trusting them to care for us, we make ourselves vulnerable to them. Patients invariably find themselves in a position of vulnerability towards their doctors, not least because of the knowledge and power imbalance between them.

Some claim that AI will make trust redundant. The increased accuracy, efficiency, and patient empowerment it promises could remove vulnerability from the doctor–patient relationship. And without vulnerability there is no reason for trust. People will be able to rely on their healthcare professionals and healthcare systems to always deliver what has been agreed. The institutional organisation of the medical and nursing professions with accreditation systems, professional governance structures, and treatment protocols has gone some way to remove vulnerability and establish more of a relationship of reliance between doctors and patients. AI could accelerate this process. But the downside of this operationalisation and subsequent loss of trust has been the increase in the litigation culture. Patients are treated more like customers expecting a certain service. When patients feel that the service promised has not been delivered, they use legal means to complain and demand compensation (as opposed to feeling gratitude when things go well, or betrayal when things go wrong, as in trust relationships). AI, with its promises for augmented rationality and reliability, is likely to accelerate this trend too. If patients are promised augmented care, they will expect augmented care, and will feel within their rights to complain when their expectations are not fully met.

There is another way that AI in medicine could reduce trust. Technological advances in machine learning and cloud computing have resulted in the *datafication* of our lives and bodies, rendering our every action and activity—how many steps we take, or what we post on social media—as “actionable” health data, which can be collected and deployed in addressing health challenges such as diabetes or mental health problems. This has opened up the clinical space to new actors and created a fertile ground for new clinical–corporate alliances—the former eager to find solutions in an environment in which money is short, the latter eager to get hold of valuable personal healthcare data to train their AI models. Importantly, these new actors are not bound by the same moral commitments and fiduciary duties as traditional healthcare providers, nor by the same legal and regulatory frameworks.<sup>11</sup> Although people seem to understand the need for public–private partnerships, they remain wary about what they regard as unaccountable private companies profiteering from the use of their health data; data that were entrusted to their doctors to be used in their own care and for the common good.<sup>12,13</sup> A strong legal and regulatory framework could redress these issues and help establish a successful

11 B. Mittelstadt, Principles alone cannot guarantee ethical AI, *Nature Machine Intelligence* (2019), 1–7.

12 Ipsos Mori, The one-way mirror: public attitudes to commercial access to health data (Wellcome Trust, March 2016), available at: <https://wellcome.ac.uk/sites/default/files/public-attitudes-to-commercial-access-to-health-data-wellcome-mar16.pdf>.

relationship with these new actors. What is needed, beyond that, is the development of a new moral attitude and culture that would allow AI developers to be ethically and socially reflexive about their work and its real world impact.

## Empathy

Grace read the diagnosis and treatment plan on her smartphone. There was no standard consultation with a doctor. She could have asked for one, if she had the time to wait, or the money to pay for it. She chose not to. What more would the doctor say to her, anyway? All she needed to know was in that text.

Efficiency, in the use of time and other resources, is important in healthcare and, with our ageing population, is likely to become more so. By using AI tools to triage, diagnose, and treat patients, healthcare systems could use resources more prudently without, theoretically, compromising accuracy and effectiveness.

However, the quest for greater efficiency might have a negative effect on other fundamental healthcare values, such as empathy.<sup>14</sup> When every minute counts, and resources cannot stretch to cover all needs, holding someone's hand, and listening to their worries and concerns, are the things health services are tempted to leave out. Although the patient-centred model of care put empathy back on the professional map, economic imperatives have curtailed the ability of healthcare professionals to engage with their patients in an empathetic way. Indeed, it is an interesting observation that, on the one hand, we try to develop more personalised, even empathetic machines, attentive to our individual needs, whilst, on the other, healthcare professionals are left to operate in time- and resource-austere environments, forcing them to become more detached and mechanical in their dealings with patients.

Some proponents of AI argue, however, that employing AI systems could actually help healthcare professionals exercise empathy. Such systems will free up time, they maintain, for doctors and nurses to sit by the patient, engage with them, and care for them in a more humane way than has been possible so far.<sup>15</sup> This optimistic vision should not be discounted. Such a future is indeed possible. But its actualisation depends less on AI and more on those making healthcare decisions valuing empathy enough to adequately fund it.

Whatever side one takes, the constructed and oft-repeated dichotomy between a caring human and a cold impersonal machine is, well, artificial. Humans have always been shaped by and with technology and this is as true in the healthcare sector as any other. Empathy and caring manifest themselves with, and through, technology. From warming the stethoscope so as not to chill a naked body, to pain relief, to a well-designed prosthesis, to an efficient AI logistics system that won't leave an A&E patient without a much needed bed, caring is a combined achievement of humans and nonhumans. It is a choreography of practices between doctors, patients, nurses, cleaners, machines, algorithms, drugs, needles, institutions, regulations—to mention only a few—requiring constant rebalancing of complex needs and shifting tensions.<sup>16</sup> Finding the right balance is not a technological problem—it is an ethical, social, and political one, that societies will have to negotiate for themselves.

13 H. Van Mil, Foundations of fairness: Views on uses of NHS patient data and NHS operational data (February 2020), *Understanding Patient Data*, available at: <https://understandingpatientdata.org.uk/news/accountability-transparency-and-public-participation-must-be-established-third-party-use-nhs>.

14 A. Kerasidou, Empathy and efficiency in healthcare at times of austerity, *Health Care Analysis*, 27/3 (2019).

15 E. Topol, The Topol Review: Preparing the healthcare workforce to deliver the digital future (National Health Service, 2019).

16 A. Mol, I. Moser, and J. Pols, Care: putting practice into theory. In A. Mol, I. Moser, and J. Pols (eds), *Care in Practice: On Tinkering in Clinics, Homes and Farms* (transcript Verlag, 2010), pp. 7–27.

## Conclusion

Grace's story is neither the worst nor the most exciting future one can imagine for healthcare. It sounds pretty mundane, believable, and almost inevitable. In most economically advanced countries, there will still be a functioning healthcare system, people will still be able to access services one way or another, get treatment and hopefully get better as a result. Maybe there won't be enough 'human touch' or empathy in the system, nurses and other professionals will often be overworked, specialist services will be difficult to access unless one can afford to pay, and AI will occasionally get things wrong—perhaps more often for some groups than others.

But even though Grace's story might sound inevitable, the future is never determined. So, is this the best possible outlook we can imagine?

AI promises major benefits for healthcare. But along with the benefits come risks. Not so much the risk of powerful super-intelligent machines taking over, but the risk of structural injustices, biases, and inequalities being perpetuated in a system that cannot be challenged because nobody actually knows how the algorithms work. Or the risk that there might be no doctor or nurse present to hold your hand and reassure you when you are at your most vulnerable. There are many initiatives to come up with ethical or trustworthy AI and these efforts are important. Yet we should demand more than this. Technological solutionism and the urge to “move fast and break things” often dominate the tech industry but are inappropriate for the healthcare context and incompatible with basic healthcare values of empathy, solidarity, and trust.

So how can such socio-political and ethical issues get resolved?

It is at this juncture that we have the opportunity to imagine different futures. We, the patients, doctors, nurses, carers—not merely figured as a group of users, consumers, or clients but as a collective of active citizens who are essential in the shaping of these technological futures—need to claim our place at the forefront of this process. Technology workers, researchers, medical professionals and activists, civil society and patient groups, need to form coalitions that push for accountability, for ongoing monitoring of AI systems, for the termination of AI projects that prove not to be good enough for those with the least power to protest, and even for the opportunity to say no to some of these technologies at the design stage and before their implementation.

To shape the future of healthcare we need to decide whether, to what extent, and how—under what regulatory frameworks and safeguards—these technologies could and should play a part in this future. AI can indeed improve healthcare but, instead of casting ourselves loose and at the mercy of this seemingly inevitable technological drift, we should be actively paddling towards a future of our choice.

In many respects “Grace's future” mirrors our present reality. She lives in a world with increased technology but the fundamental sociopolitical and ethical issues that characterise our present remain unresolved. Surely Grace deserves a better future—as do we.

## Further Reading

D'Ignazio, C. and Klein, L. F., *Data Feminism* (MIT Press, 2020).

Eubanks, V., *Automating Inequality: How High-tech Tools Profile, Police, and Punish the Poor* (St Martin's Press, 2018).

Noble, S. U., *Algorithms of Oppression: How Search Engines Reinforce Racism* (NYU Press, 2018).

Topol, E., *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again* (Basic Books, 2019).