

The Ethics of Medical AI and the Physician-Patient Relationship

SALLY DALTON-BROWN

Abstract: This article considers recent ethical topics relating to medical AI. After a general discussion of recent medical AI innovations, and a more analytic look at related ethical issues such as data privacy, physician dependency on poorly understood AI helpware, bias in data used to create algorithms post-GDPR, and changes to the patient-physician relationship, the article examines the issue of so-called robot doctors. Whereas the so-called democratization of healthcare due to health wearables and increased access to medical information might suggest a positive shift in the patient-physician relationship, the physician's 'need to care' might be irreplaceable, and robot healthcare workers ('robot carers') might be seen as contributing to dehumanized healthcare practices.

Keywords: Medical AI; GDPR; algorithm bias; care robots

Is artificial intelligence the 'stethoscope of the 21st century'?¹ AI detection systems claim impressive effectiveness ratings, such as 80 percent accuracy for a British AI program that detects heart disease, 95 percent for an American 'smart' microscope that detects blood infections, and 94 percent for a Japanese endoscopic system for diagnosing cancerous growths.² The US Food and Drug Administration approved the first software powered by artificial intelligence that replaces the need for a specialized doctor to interpret medical imagery in April 2018; IDX-DR software analyses images of the retina to detect whether a person with diabetes has a complication from diabetic retinopathy.³ Eye-screening technology, developed by the CSIRO (the Commonwealth Scientific and Industrial Research Organisation) in Australia, enables GPs to test diabetic patients for diabetic retinopathy—a condition affecting one in three diabetic people, and that can lead to blindness if untreated.⁴ Also in Australia, NEC is working on AI intelligence algorithms for eye tracking and facial recognition, as part of developing new diagnostic tools for detecting autism at a young age.⁵ The Queensland University of Technology and Children's Health Queensland are exploring the use of robots as healthy eating companions, designed to deliver nonjudgmental advice.²

But what are the ethical issues of concern in a highly disruptive technological environment where it is claimed that up to 40 percent of jobs will need to be restructured due to the rise of AI and automation in the next few decades? *The Economist* has argued that "AI will remain narrow, not general. Instead of wondering whether AI can replace a job, it is better to ponder whether it could replace humans at a specific task."⁶ Crunching big data⁷ and working diagnostically within specific parameters to a high degree of accuracy would indeed seem to be useful tools to assist clinicians. In 2016 IBM's Watson supercomputer was credited with "diagnosing in minutes the precise condition affecting a leukaemia patient in Japan that had been baffling doctors for months, after cross-referencing her information with 20m oncology records."⁸ A London hospital in 2016 teamed up with Google's DeepMind AI arm to develop a health app that allegedly could free up more than half a million hours spent annually on paperwork.⁹

The value approach to technological design suggests 12 human values that are implicated in technological design: human welfare, ownership and property, privacy, freedom from bias, universal usability, trust, autonomy, informed consent, accountability, identity, calmness, and environmental sustainability.¹⁰ The EU's Ethics in Science and New Technology Group's statement in March 2018 on AI suggests the following key concerns: safety, moral responsibility, manipulation ('data nudging'), value preservation, and the impact on current governance and regulation.¹¹

A more useful and succinct approach comes from a Stanford March 2018 study which identified the following four issues relating to medical AI as: issues of data privacy; physician dependency on AI helpware that is poorly understood; data used to create algorithms containing the bias of the companies or health care systems that developed them; and a change in the dynamics of the physician-patient relationship.¹²

Data privacy is a general concern, not one limited to medical AI.¹³ A machine that has no real memory of a patient may be more secure than a doctor who remembers certain facts about his patients, and the safety of patient files is an issue whether AI is used or not. Bitcoin technology offers potential solutions in this area.¹⁴ However, public trust, or rather the lack thereof, in data privacy was recently demonstrated in Australia with the introduction of an opt-out record-keeping system, 'My Health.' Concerns about security have been raised following the hacking of Singapore's health records database, as well as the potential issue of private companies getting access to such personal data. The government and Australia's opposition party continue to ponder how best to regulate AI, after calls from the country's Chief Scientist, Alan Finkel, for a voluntary scheme in terms of which companies would apply for a 'Turing stamp' for artificial intelligence providers, similar to the Fairtrade logo on coffee.¹⁵

Physicians' dependency on helpware leads to questions of liability, of who exactly is to blame if a doctor accepts the wrong recommendation of an AI.¹⁶ The issue of how an AI might be 'credentialled' is a complex one; AI systems that claim to provide particular clinical services or benefits to patients such as unregulated online mental health services have given rise to lawsuits. Arguably however, AI issues of liability can be debated under the general rubric of evolving law. The more interesting question is that of the level of moral agency that can be ascribed to an algorithm. John P. Sullins proposes that intelligent machines (i.e., robots) are indeed moral agents when the machine has autonomous intentions and responsibilities. If the machine is seen as autonomous, then the machine can be considered a moral agent.¹⁷ Australian lawmakers are taking some steps to consider matters related to robot innovation, such as drone regulation and driverless vehicles, but these chiefly target issues of safety.

The next two issues identified in the Stanford study are the interlinked concerns of how data derived from diagnostic AI algorithms is both created and used, and the algorithm's impact on how clinicians make decisions and communicate with their patients.

Potential bias in algorithm development suggest that the diagnostic data process may contain flaws from the outset. The agenda of the data manager such as an insurance company, for example, might skew the process. Agendas aside, there is also the issue of 'loopthink,' i.e., "programming biases that would exclude outliers and reduce abstract or even ethical thinking on the part of an

AI program.”¹⁸ The algorithm ignores data that does not ‘fit.’ While this is not an issue limited to nonhuman diagnosticians when examining datasets, the problem is how AI might replicate that often intuitive diagnostic thinking that experienced practitioners allegedly possess. AI diagnoses are unsurprisingly more consistent than those of clinicians, as a study comparing AI diagnoses of oxygen deprivation in newborns versus clinicians’ diagnoses revealed—the latter group not only disagreed with each other, but on the data being re-presented to them later, disagreed with themselves.¹⁹ But this may not necessarily be a bad thing; it depends on whether one believes a consensual diagnosis evolved through debate is more accurate than one derived through strictly applied parameters.²⁰ In other words, the question is whether a little inconsistency through aggregated thinking, rather than algorithmic consistency, might lead to better patient outcomes. In analysing a cybersecurity AI program, researchers noted that:

...individuals who are making ‘outlier decisions’ should not be considered as ‘wrong’. In fact, in some cases they may represent people at the forefront of new knowledge creation. In these cases the individuals have an important role to play in challenging the ‘group think.’²¹

The concern about algorithm bias is part of a larger concern with the comprehensibility and accessibility of data to the patient. The EU’s new General Data Protection Regulation (GDPR) means that from May 25, 2018, according to section 4, article 22, users have a ‘right to explanation’ on all decisions made by automated or AI algorithm systems.²² Informed consent, not to mention patient dignity and autonomy, require that the basis for making the diagnosis should be comprehensible. The ‘black box’ issue, as it is termed with AI—is that the algorithm operates in a way incomprehensible to most non-specialists, who can only see input and output, without any understanding of the process itself.

Comprehensibility is only one aspect of communication, however. An AI diagnosis of a 50 percent likelihood of a fatal illness requiring strong intervention can be presented to a patient in various ways. If the AI diagnosis is one derived at through a range of factors resulting in a 40–60 percent likelihood of an outcome, then should the patient be apprised of *all* the variant factors that have gone into arriving at that possible outcome? And is this different from how a doctor might advise a patient when trying to predict a percentage chance of mortality?

Data accessibility—whether driven by AI or physician—can be seen within the larger context of the so-called ‘democratization’ of healthcare. The accessibility of knowledge through online information (no matter how misleading), and of health wearables that allow users to collect and track some of their own health data suggests a shift in the patient-physician paradigm. (A 2017 PWC report on medical AI noted the growing prevalence of consumer wellbeing apps, including wearables for early detection.) There are also ‘softbots,’ or online therapeutic avatars, that are allegedly producing good health care results.²³ Digital camera records of the wearer’s day have been shown to help dementia patients to recollect aspects of earlier experiences that have subsequently been forgotten, thereby acting as a retrospective memory aid.²⁴ Other AI programs assist those with ‘locked-in’ syndrome or with autism—the AI effectively operating as an extended brain function in much the way computers do, generally speaking. Previously, Nicholas Jewson’s object-oriented medical cosmology, as encapsulated under the rubric of ‘sick man

theory,' predicated medical knowledge as increasingly guarded by the gatekeeper physician.²⁵ According to some medical sociologists, the increasing distribution of medical knowledge amongst lay people has reversed Jewson's model.²⁶ The patient now faces both the added autonomy and the added responsibility deriving from access to a far greater range of complex data.

In terms of autonomy, an AI might be more 'ethical' than a human practitioner. Categorising an effective physician-patient relationship in terms of key processes such as autonomy, managing the power imbalance, and offering a competent service as a physician, an AI carer-physician might be more effective. A patient may trust an AI more, believing the AI's health advice is unbiased and not prone to human fallibility. An AI doctor will not be exhausted after a long shift, nor has it any concept of power.

Yet these aspects of the relationship relate to knowledge and decision-based virtues such as informed consent and autonomy, rather than to the empathic skills that reinforce patient dignity. For some patients, autonomy is arguably less important than empathy, and care less a question of democracy than of dignity.

The issue is that of how the AI 'values' the patient, of that human contact that can be a significant aspect of a trusting relationship. Are shared human values and experience (rather than programmed ethical parameters) required for ethical approaches that demonstrate respect? A European Parliament Committee report on the civil law of AI has stated that "human contact is one of the fundamental aspects of human care" and replacing humans with robots could "dehumanise caring practices."²⁷ The EU has called for a specialist commission to answer questions in relation to AI ethics in hospitals and health care institutions.²⁸ Japan devotes a third of its robotics funding to elderly²⁹ and end-of-life care robotics, thus the urgent need for ethical understanding of the issues relating to such a vulnerable population group.³⁰

Given that the basic tenets of care ethics are those of recognizing others' needs, taking on responsibility for addressing those needs competently, and the responsiveness of the care receiver to the carer, can an AI entirely fulfil the ethics of care? The first two acts, of recognizing and addressing patient needs, are not particularly problematic, but the third—the notion of responsiveness, of patient and physician acknowledging common humanity and responding appropriately and genuinely, might be. Yet, not all physicians are highly empathic, and empathy can be a detraction from care, given its focus on the agent rather than on the altruistic act. Whether a carer robot could be programmed to respond to emotional needs as well as to physical requires ethicists to ponder whether a *simulation* of empathy is sufficient. If empathy comes from shared human experiences, can those be recreated ethically and effectively?

Phrasing the argument in extreme terms—given that robots cannot die, how are they able to care for mortals, for example—exposes the irrelevance. One does not need to experience a concept in order to understand it or to demonstrate empathy. Palliative care workers for example are not resurrectors. Doctors with, for example, severe autism may struggle with empathy but not with care.

An ethics of care and an ethics of empathy are not identical. Care ethics is grounded in a mutual identification of 'humanness,' as the notion of responsiveness indicates. Carers are enmeshed within reciprocal, dependent interrelations. The recognition of the interests of both self and other, carer and cared for, within

such ethical practice is based on mutual recognition. Surely such recognition requires species recognition? However, an alien would be 'programmed' by an entirely different ontological-epistemological mindset, a care robot by those involved in a value-derived algorithmic process. (And probably also by an ethics committee.) It is a dehumanized product but not a dehumanized process. And one that brings us back to the ethics of the programmer. Given that the AI carer is the product of human design, although not an empathic agent, it can be an agent of care. Though this argument would mean that a toaster could be seen as ethical, the difference comes back to the degree and complexity of the simulation of human agency.

One argument would suggest that there is no reason why *acts* of care cannot be accepted as synonymous with *thoughts* of care. The patient could accept an action as one backed by an intent to care, whether that intent is produced by a human desire to help those in need, or a program, and assess not the agent, but the beneficence of the act. For the patient, the AI physician's intent is recognized as humane, if not human.

Can a robot carer simulate 'recognition,' empathy, and human experiences, sufficiently to fulfil the requirements of an ethics of care? Or will it always be required to work alongside a human practitioner to ensure that the patient literally receives that valuable human touch? Some patients do not in fact distinguish between AI and human physicians, the danger thus that of becoming attached to AI diagnosticians, even forming attachments to them, vulnerable patients thus becoming unable to determine if they are interacting with a machine or not.³¹ The ethical issue may be not that of physical touch that suggests care and recognition of human suffering, but the exact opposite, namely the danger of anthropomorphizing AIs.

For patients, a vulnerable population group, the issue is not necessarily even ethical, but philosophical; one of how we define ourselves as human. The problem of the cyborgism of a robot carer is one extreme of the debate begun with the use of the first prostheses and other artificial aids to human health. Being cared for by AIs can be controlled to ensure that the ethics of care are not violated. But the fact that there is real concern about whether such carers might increase human isolation and alienation suggests that AI might offer an efficient solution to physical needs but has yet to provide a sufficient answer to that universal need to recognize something human, fallible, and mortal.

Notes

1. Norman A. Your future doctor may not be human. This is the rise of AI in medicine. 31 Jan 2018; available at <https://futurism.com/ai-medicine-doctor/> (last accessed 9 Sept 2018).
2. Artificial intelligence: Is this the future of early cancer detection? 29 Oct 2017; available at https://www.eurekalert.org/pub_releases/2017-10/sh-aii102317.php (last accessed 2 Nov 2018).
3. The FDA just opened the door to let AI make medical decisions on its own. 13 Apr 2018; available at <http://www.bioethics.com/archives/43050> (last accessed 2 Nov 2018).
4. Johnston R. Australian scientists are using artificial intelligence to prevent blindness in people with diabetes. 20 Sept 2017; available at <https://www.gizmodo.com.au/2017/09/australian-scientists-are-using-artificial-intelligence-to-prevent-blindness-in-people-with-diabetes/> (last accessed 9 Sept 2018).
5. See https://au.nec.com/en_AU/press/201710/researchers-explore-nec-artificial-intelligence-medical-diagnostic-tools.html.
6. AI, Radiology and the future of work. *The Economist*, 9 June 2018.

7. Ludwin SK, Murray TJ. Dilemmas in medical ethics in the age of big data. *Multiple Sclerosis Journal* 2017;23(10):1306–8.
8. However, the same system has just consumed five years and \$62m (£51m) in an unsuccessful attempt to transform care at the University of Texas MD Anderson Cancer Centre. Vize R. Technology could redefine the doctor-patient relationship, *The Guardian* 11 Mar 2017; available at <https://www.theguardian.com/healthcare-network/2017/mar/11/artificial-intelligence-nhs-doctor-patient-relationship> (last accessed 9 Sept 2018).
9. See Lynsey Barber’s analysis of 22 Nov 2016; available at <http://www.cityam.com/254167/google-deepmind-wants-make-doctors-paperwork-obsolete-new> (last accessed 21 Oct 2018).
10. Kahn PH Jr, Ishiguro H, Friedman B, Kanda T, Freier NG, Severson RL, Miller J. What is a human? Toward psychological benchmarks in the field of human–robot interaction. *Interaction Studies* 2007;8(3):363–90. See also Sharkey NE. The ethical frontiers of robotics. *Science* 2008; 322(2):1800–1.
11. See http://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf (last accessed 30 Oct 2018).
12. See <https://med.stanford.edu/news/all-news/2018/03/researchers-say-use-of-ai-in-medicine-raises-ethical-questions.html> (last accessed 9 Sept 2018).
13. Zapusek T. Artificial intelligence in medicine and confidentiality of data. *Asia Pacific Journal of Health Law and Ethics* 2017:105–26.
14. Amoako-Agyei E. Blockchain and AI merge for healthcare in new partnership deal. 23 Feb 2018, at <https://www.coinstaker.com/blockchain-ai-merge-healthcare-new-partnership/> (last accessed 29 Oct 2018).
15. See <https://www.innovationaus.com/2018/07/AI-ethics-hits-political-nerve> (last accessed 30 Oct 2018).
16. Oakden-Rayner L. Artificial intelligence won’t replace doctors soon but it can help with diagnosis. *The Conversation* 19 Sept 2017; available at <http://www.abc.net.au/news/2017-09-19/ai-wont-replace-doctors-soon-but-it-can-help-diagnosis/8960530> (last accessed 9 Sept 2018).
17. Sullins JP. Open questions in robethics. *Philosophy and Technology* 2011;24(3):233–8.
18. Cheshire WP. Loopthink: A limitation of medical artificial Intelligence. *Ethics and Medicine* 2017;33(1):7–12.
19. Garibaldi I. Application of simulated annealing fuzzy model tuning to umbilical cord acid-base interpretation. *IEEE Transactions on Fuzzy Systems* (1999);7:72–84.
20. My thanks to Professor Uwe Aickelin from the University of Melbourne for a very enlightening conversation on 8 May 2018 at Jimmy Watson’s, on algorithms, inconsistency and decision making, and for ongoing advice on reference material.
21. Miller S, Appleby S, Garibaldi JM, Aickelin U. Towards a more systematic approach to secure systems design and analysis. *International Journal of Secure Software Engineering* 2013;4 (1):11–30.
22. See <https://medium.com/syncedreview/will-new-eu-regulations-starve-data-hungry-deep-learning-models-25403795d26c>; and <https://www.eugdpr.org/key-changes.html> (last accessed 9 Sept 2018).
23. Hamet P, Tremblay J. Artificial intelligence in medicine. *Metabolism - Clinical and Experimental* (2017);69:36–40.
24. Hoey J, Zutis K, Leuty V, Mihailidis A. A tool to promote prolonged engagement in art therapy: design and development from arts therapist requirements. *Conference on Computers and Accessibility* 2010:211–8.
25. Reller T. AI’s revolutionary role in healthcare: Experts debate the pros and cons of artificial intelligence, which is changing many aspects of patient care. 2 June 2017; available at <https://www.elsevier.com/connect/ais-revolutionary-role-in-healthcare> (last accessed 3 Sept 2018).
26. Nicolson MA. Nicholas Jewson and the disappearance of the sick man from medical cosmology, 1770–1870. *International Journal of Epidemiology* 2009;38(3):639–42.
27. The EU Legal Affairs Committee *Report with recommendations to the Commission on Civil Law Rules on Robotics*; available at http://europa.eu/rapid/press-release_IP-18-3362_en.htm (last accessed 16 Sept 2018).
28. Clarke H, Burger J, Paxton V. Australia: Do robotics and artificial intelligence have a future in the healthcare space? 21 Feb 2018; available at <http://www.mondaq.com/australia/x/675750/Healthcare/> (last accessed 6 Sept 2018).
29. In *Robot and Frank*, a 2012 film about an aging ex-burglar cared for by a therapy robot, the human wins out over the artificial, as the robot is quickly turned to crime by his patient.

The Ethics of Medical AI and the Physician-Patient Relationship

30. What doctor? Why AI and robotics will define New Health; available at <https://www.pwc.com/gx/en/industries/healthcare/publications/ai-robotics-new-health.html> (last accessed 2 Aug 2019).
31. Luxton DD. Recommendations for the ethical use and design of artificial intelligent care providers. *Artificial Intelligence in Medicine* 2014;62:1–10.