Irish Journal of Medical Science (1971 -) (2022) 191:1991–1994
https://doi.org/10.1007/s11845-021-02853-3

BRIEF REPORT

Artificial Intelligence: the future of medicine, or an overhyped and dangerous idea?

Shubhangi Karmakar[1]

Received: 26 January 2021 / Accepted: 7 November 2021 / Published online: 16 November 2021

Abstract

Introduction Contemporary discourse on Artificial Intelligence (AI) in medicine is oft-sensationalised to the point of bearing no resemblance to its everyday impact and potential — either to proselytise it as a saviour or to condemn its perilous, amoral and sprawling reach.This report aims to unravel the paucity of understanding underpinning this hyperbolic duality, whilst addressing the potential clearly defining its ethical use poses to the semi-public healthcare models in Ireland and Europe.

Discussion The report contrasts the challenge of regulating the breakneck development of AI, with healthcare's necessity for stringent quality control in ethical technological development to ensure patients' well-being.Physical, practical and philosophical approaches to Artificial Intelligence in medicine are explored through Beauchamp and Childress' principles of delivering care with beneficence, non maleficence, justice and autonomy. AI is scrutinised under Kantian deontological, Benthamite utilitarian and Rawlsian perspectives on health justice. Actor Network theory is used to explain sociotechnical interactions governing human stakeholders developing ethical AI.These analyses operate firstly to define AI concisely, then ground it in its contemporary and future functions in healthcare. They highlight the importance of aligning medical AI with accepted ethical standards as a necessity of its integrated use across healthcare.

Conclusion This report concludes that balanced assessment of AI's role in healthcare requires improvement in three areas: improving clarity in definition of AI and its extant remit in medicine; aligning contemporary discourse on AI use with contemporary objective ethical, legal and system frameworks; and clearly identifying for dismissal a number of logical fallacies deliberately sensationalising AI's potential.

Keywords Information technology · Philosophical ethics · Public health ethics · Social control of science/technology · Technology/risk assessment

Introduction

"Artificial Intelligence" (AI) remains imprecisely defined, since Turing's first writings questioned whether a system "like a child brain — rather little mechanism, lots of blank sheets" could be designed to make choices, whose evolution and learning could be sped up, given engineering advances are "adequate for the requirements", by a human experimenter's "exercise of intelligence" [1].

AI's role in healthcare suffers from this paucity of understanding. Compared to the American model tying cost to quality of private healthcare [2], its challenges are unique in the Irish and European setting [3] where goals of semi-public healthcare and private healthcare entrepreneurship are separated under ethical, legislative and data system frameworks [4]. Here, the industry of AI appears to the ill-versed eye, to be developing at a breakneck speed with little concern for regulation, whilst medicine appears to rely on stringent quality control to ensure patients' well-being foremost.

This report strikes through the sensational and scrutinises the substantive arguments on the contemporaneous role of AI in healthcare, from the core principles of delivering care with beneficence, non-maleficence, justice and autonomy, under Kantian deontological, Benthamite utilitarian and Rawlsian ethical theories, and making note of the effect of Actor Network theory on human stakeholders in AI.

In doing so it demonstrates the feasibility of a future for AI that is neither "overhyped" nor "dangerous", but accessible, accountable and adaptable to the needs of the plurality of stakeholders in healthcare.

✉ Shubhangi Karmakar
karmakas@tcd.ie

[1] School of Medicine, Trinity College Dublin, Dublin 2, College Green, Ireland

Discussion

Defining Artificial Intelligence

The first attempts to workshop AI proposed it as the intersection of computing, mathematics, neural nets, language interpretation, chaos, learning and decision making in Dartmouth University [5]. To date, it largely incorporates the growth and change in the same fields, with the incorporation of two closely related concepts:

1. the "Internet of Things" (IoT), encompassing the connection of physical technologies and devices, to aggregate multi-source qualitative or quantitative data over time [6] and
2. "Big Data", analysing advances in gathering, storing and managing these data, to be processed by AI algorithms [7].

For consideration under *contemporary* ethical, systems and legal frameworks, it is pertinent to conceptualise AI in its strictest practical sense in the *present day*. That is, as the digital and physical applications of computation and machine processing of Big Data, through decision and learning processes, which together create an "ecosystem of automation" [8] to produce outcomes which would be said to show "intelligence" by a human — noting carefully that since its inception in thought, the outcomes of AI "do not presuppose any feelings on the part of the machine", or allusions to "sentience" [1].

In brief, the function of AI can be understood as computing algorithms achieving outcomes not far removed from what Turing would have imagined his "Turing Test" to apply to, but accelerated in capacity coupled to physical and digital technological development following Moore's Law over the last 50 years, to an extent of "unreasonable effectiveness of data" [9], where even simple algorithms can reach highly accurate conclusions, due to "Machine Learning" adaptable programming on datasets, ideally with shared ontologies [10].

The role of medicine and medical ethics

Whilst any AI can be said to operate through objective millisecond binary decisions, with outcomes ascribed fixed values by stakeholders — for instance self-driving cars' crash prioritisation [11] — medicine works by implementing long-term decisions optimised by complex overlapping ontologies, such as guides of best practice, that call for subjective value decisions, utilising a plurality of objective data with discretionary nuance.

Medical practice aspires to a balance between Beauchamp & Childress' pillars: Beneficence, Non-maleficence, Autonomy and Justice [12]. The first two are largely seen as indisputable; hence, most ethical contradictions arise from differing values placed on individual and social distributive justice in healthcare.

Benthamite [13] and Millsian [14] utilitarianism prioritise social distributive justice in public healthcare, for "the most good for the most people", working by "rule" using a set guideline of best practice to maximise population wellbeing. An example would be cervical screening programmes, under the principle that screening the most people would reduce population incidence of cervical cancer over time [15].

Kantian Deontology [16] normatively centres autonomy above other priorities, in privatised healthcare approaches. This has particular weighting in decision at extremes such as "end of life" [12], and inspires individual patients to seek strong predictive values from screening such as in the Cervical Check Scandal [17] where they felt autonomy was undermined by systematic "deceit".

Finally, Rawlsian Justice [18] acts by the "liberty" and the "difference" principle, to maximise individual autonomy and help the least favoured first, so being born "under a veil of ignorance" of circumstance, everyone has the same opportunity in life. In the context of cervical screen, this contradicts Benthamites as it says whilst screening for diseases should be available widely, it should be promoted if there is likelihood of benefit and contradicts Kantians as it should be deniable if no benefit is foreseen, regardless of its informing individuals or population health.

Hence, conceptualised broadly under these three ethical frameworks, it is important to understand the role AI could play in healthcare by which agendas and ethical principles it serves to favour.

Reconceptualising AI in healthcare: past, present and future

A balanced understanding of the remits, restrictions and ramifications of AI in healthcare, encompassing individual, corporate and public stakes and governance in digital healthcare and personalised medicine, has suffered in public regard from a high degree of narrative bias, as the consequence of compound logical fallacies, and pertinent omission of theory in public understanding.

Hasty generalisations in definition by *mutatio elenchi* of the writers' deliberate misdirection to sensation, obfuscate AI's contemporaneous substantive significance in public discourse. Reductive assertions for "clickbait" journalism are coupled with confirmation bias *ad verecundiam* by approaching research "authorities in their field" who lack transparency about the scope

of existing AI utility when eager to overblow their own "next breakthrough" [19]. This disinforms public awareness of the extent of AI's past and current employ in research, development, delivery and monitoring of healthcare, and misinforms by scaremongering of a future "singularity" [20]. Finally, there exists a lack of public education about Actor Network Theory to appreciate sociotechnical and ethical interactions of developing AI solutions for medicine. A specific concern is that the lack of understanding of predilection of human designer-operators of technology to stray to ethical ambiguity as technology becomes more agnostic and powerful, leads to an attribution of anthropomorphised harms such as 'cultural bias' intrinsically to AI algorithms as tools, when in fact the social effects actually arise from the "context in which they are embedded and used" [21] by humans.

Acknowledging this degree of sensationalisation, AI must transcend the polarised false dichotomy of "paradise or perdition", as it is philosophically unfair to characterise as "dangerous", "overhyped" or a "future", what is a staunch, secure mainstay in contemporary healthcare, and the harms of bias by stakeholder agenda are due to human design.

An example of beneficence close to home, Irish companies are a growing global developer in AI and healthcare for the last decade, not through sporadic breakthroughs or controversy but creating "everything from [barcodes] to RFID [providing] a hardware and infrastructure-driven AI solution set to get the most reliable quality control and traceability" for physical healthcare such as drugs and devices, from source to end-user [22].

Both barcodes and radio-frequency identification [RFID] are AI IoT utilised globally since 1940s–70 s, but contrary to the disinformative scaremongering of a dangerous future of public data harvesting for purely private entities [23], they are primarily used to fulfil ethical practice, through accountability and tracking of private supply of drugs and devices to public bodies such as the HSE.

Looking at AI's future potential, its utility has been researched in cervical cancer screening, and reported as having "outperformed human experts" in eye-catching press release titles. This appeals immediately to utilitarianism, deontology and Rawlsian ethics as it implies that "AI", or rather automation processing exponentially more patient samples at a greater accuracy rate, forms a more cost- and resource-effective intervention with better individualised information, even for low-resourced areas [24] — but also raises misinforming concerns of AI causing human redundancy.

However, once sensationalism is removed, the substantive data show that whilst the algorithm (requiring human expertise to design) often outperformed colposcopist ratings and cytology, it remained ineffective at "training for the subtleties of visual appearance, which is a difficult skill", and was only useful at highlighting where imaging qualities were low but not interpreting them [25].

This indicates that the nuanced discretion and experience of a human will remain essential to the operation and implementation of any AI automation, to counter the particular dangers of lacking beneficence and non-maleficence particularly for individual justice and patients affected by the lowest resource availability and therefore poorer quality imaging and interventions.

Conclusion

A balanced assessment of AI's role in healthcare remains an elusive pursuit, suffering from the following:

1. A lack of clarity in definition of AI and the systems it encompasses — including but not limited to the roles of IoT, Big Data and Machine Learning processes in *current* and *future* utility;
2. A lack of information matching *contemporary* use of AI to *contemporary ethical, legal and system frameworks* in healthcare across different global public and private systems; and
3. A number of logical fallacies deliberately sensationalising AI's future risks and capacities, as well as underplayed understanding of Actor-Network Theory, contributing to anthropomorphisation and attribution of hyperbolic intrinsic ill-intent, biases and threat of human redundancy, to basic algorithm functions.

If ethical pillars, and principles driving core medical practice such as utilitarianism, deontology or Rawlsian justice are clearly laid out at every step of analysis, with the example of the need for and optimisation of cervical screening from different perspectives illustrated throughout this report, it is possible to emerge from scandals and scaremongering to a discussion of the substantive utility, agendas, risks and benefits of AI and healthcare, where it is used for two-way private and public sector accountability, allowing access to better black-and-white diagnostic performance whilst remaining adaptable to "grey" scenarios with continued human expertise and input.

References

1. Turing AI (1950) Computing machinery and intelligence. Mind LIX(236):433–460
2. Burke L, Ryan A (2014) The complex relationship between cost and quality in US health care. Virtual Mentor 16(2):124–130
3. Kritikos M (2019) Artificial Intelligence ante portas: legal & ethical reflections. [Internet]. Europarl europa eu [cited 22 Jan 2021]. Available from: https://www.europarl.europa.eu/at-your-service/files/be-heard/religious-and-non-confessional-dialogue/events/en-20190319-artificial-intelligence-ante-portas.pdf

4. Nuffield Bioethics (2019) Artificial intelligence (AI) in healthcare and research [Internet]. Nuffieldbioethics org [cited 22 Jan 2021]. Available from: https://nuffieldbioethics.org/wp-content/uploads/Artificial-Intelligence-AI-in-healthcare-and-research.pdf
5. McCarthy J, Minsky M, Rochester N, Shannon C (1955) A proposal for the dartmouth summer research project on artificial intelligence [Internet]. Www-formal Stanford Edu [cited 22 Jan 2021]. Available from: http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html
6. Wortmann F, Flüchter K (2015) Internet of Things. Bus Inf Syst Eng 57(3):221–224
7. O'Leary D (2013) Artificial Intelligence and Big Data. IEEE Intell Syst 28(2):96–99
8. Towers Clark C (2019) Big Data, IoT and AI, Part One: Three sides of the same coin. Forbes [Internet]. [cited 22 Jan 2021]. Available from: https://www.forbes.com/sites/charlestowersclark/2019/02/15/big-data-iot-and-ai-part-one-three-sides-of-the-same-coin/#5f9cb40a69da
9. Halevy A, Norvig P, Pereira F (2009) The unreasonable effectiveness of data. IEEE Intell Syst 24(2):8–12
10. Konys A (2016) Ontology-based approaches to big data analytics. Hard and Soft Computing for Artificial Intelligence. Multi Sec 355–365
11. Li M (2019) Another self-driving car accident, another AI development lesson. Towards Data Science [Internet]. [cited 22 Jan 2021] Available from: https://towardsdatascience.com/another-self-driving-car-accident-another-ai-development-lesson-b2ce3dbb4444
12. Page K (2012) The four principles: can they be measured and do they predict ethical decision making?. BMC Med Ethics 13(1)
13. Bentham J (2012) Introduction to the principles of morals and legislation. Dover Public
14. Mill J (1875) A System of logic, ratiocinative and inductive. Longmans, Green, Reader and Dyer, London
15. Williams J, Carter S (2016) An empirical study of the 'under-screened' in organised cervical screening: experts focus on increasing opportunity as a way of reducing differences in screening rates. BMC Medic Ethic 17(1)
16. Alexander L, Moore M (2020) Deontological ethics; The Stanford Encyclopedia of Philosophy. Winter: Metaphys Res Lab Stanford Univer
17. Carswell S (2018) CervicalCheck scandal: what is it all about?. The Irish Times [Internet]. [cited 22 Jan 2021] Available from: https://www.irishtimes.com/news/health/cervicalcheck-scandal-what-is-it-all-about-1.3480699
18. Rawls J (2005) A theory of justice. Cambridge (Mass.): Belknap Press Harvard University Press
19. Ray T (2019) Why is AI reporting so bad?. ZDNet [Internet]. [cited 22 Jan 2021] Available from: https://www.zdnet.com/article/why-is-a-i-reporting-so-bad
20. Purcell B (2018) Is the singularity coming?. Arc Digital Media [Internet] [cited 22 Jan 2021] Available from: https://arcdigital.media/is-the-singularity-coming-ef8580d4ce97
21. Cresswell K, Worth A, Sheikh A (2010) Actor-Network Theory and its role in understanding the implementation of information technology developments in healthcare. BMC Med Info Decis Making 10(1)
22. Karmakar S (2019) Building global footprint by keeping it real. Business Post [Internet] [cited 22 Jan 2021] Available from: https://www.businesspost.ie/commercial-reports/building-global-footprint-by-keeping-it-real-86cb0969
23. Woods K, Whyte B (2019) Cancer charity meets with health office over GMI data controversy. Busin Post [Internet] [cited 22 Jan 2021] Available from: https://www.businesspost.ie/health/cancer-charity-meets-with-health-office-over-gmi-data-controversy-9864231e
24. National Cancer Institute (2019) AI approach outperformed human experts in identifying cervical precancer. [Internet] Available from: https://www.cancer.gov/news-events/press-releases/2019/deep-learning-cervical-cancer-screening
25. Hu L, Bell D, Antani S et al (2019) An observational study of deep learning and automated evaluation of cervical images for cancer screening. Obstet Gynecol Surv 74(6):343–344