REVIEW

Conceptualising Artificial Intelligence as a Digital Healthcare Innovation: An Introductory Review

Anmol Arora [ID]

School of Clinical Medicine, University of Cambridge, Cambridge, UK

Abstract: Artificial intelligence (AI) is widely recognised as a transformative innovation and is already proving capable of outperforming human clinicians in the diagnosis of specific medical conditions, especially in image analysis within dermatology and radiology. These abilities are enhanced by the capacity of AI systems to learn from patient records, genomic information and real-time patient data. Uses of AI range from integrating with robotics to creating training material for clinicians. Whilst AI research is mounting, less attention has been paid to the practical implications on healthcare services and potential barriers to implementation. AI is recognised as a "Software as a Medical Device (SaMD)" and is increasingly becoming a topic of interest for regulators. Unless the introduction of AI is carefully considered and gradual, there are risks of automation bias, overdependence and long-term staffing problems. This is in addition to already well-documented generic risks associated with AI, such as data privacy, algorithmic biases and corrigibility. AI is able to potentiate innovations which preceded it, using Internet of Things, digitisation of patient records and genetic data as data sources. These synergies are important in both realising the potential of AI and utilising the potential of the data. As machine learning systems begin to cross-examine an array of databases, we must ensure that clinicians retain autonomy over the diagnostic process and understand the algorithmic processes generating diagnoses. This review uses established management literature to explore artificial intelligence as a digital healthcare innovation and highlight potential risks and opportunities.

Keywords: machine learning, data, diagnostic algorithms, artificial intelligence, innovation

Introduction

Artificial intelligence (AI) is conventionally defined as the ability of a computer system to perform acts of problem solving, reasoning and learning.[1] Of these processes, independent learning is most important as it mitigates the need for human intervention to continually enhance the performance of the system. AI is recognised as an emergent enabling technology, with the potential to enable multidisciplinary innovation.[2] There is potential for AI systems to work more efficiently, with increased reliability, and for longer hours, than paid human workers. This is a significant area of innovation to many healthcare organisations which currently suffer from staff shortages, increasing demand pressures and strict financial constraints. Experts in the field have described the scope of AI to address longstanding deficiencies in healthcare and improve patient care.[3]

AI also has the potential to enhance medical devices, especially by working synergistically with robotic technologies in enhancing the rate and scope of their iterative improvements. Broadly speaking, medical devices work more effectively when they are dynamic and able to respond to individual patient needs, which is facilitated by AI integration. AI is beginning to be recognised as a "software as

Correspondence: Anmol Arora
School of Clinical Medicine, University of Cambridge, Trinity Hall, Trinity Lane, Cambridge, CB2 1TJ, UK
Tel +441223 332500
Email aa957@cam.ac.uk

a medical device" (SaMD) and one which requires particular attention to be paid to post-market performance.[4] Due to its ability to facilitate innovation at device, network, service and contents layers of care pathways, implementation of AI systems requires much consideration as to how they should sustainably be introduced. In the US, the Food and Drug Administration (FDA) is introducing a fresh set of regulation for AI-enhanced medical devices.[4] Whilst there has been much research into the uses of AI in healthcare and abstract discussion about its potential, there has been limited study into how AI may practically be incorporated with healthcare systems, regulated and monitored after market entry. As a digital innovation, AI presents several challenges to healthcare systems looking to realise its potential, many of which have been noted in the wider innovation literature.

This paper employs recognised management literature to conceptualise AI as a digital innovation, referring to established uses within healthcare. In doing so, it will offer an insight into the potential of AI to transform healthcare systems, by showing its uses and practical implications.

Medical Uses of Artificial Intelligence

Potential uses of artificial intelligence are well described in literature, some of which are proven and some which are speculative. Thus far, studies have shown that AI can outperform human doctors in a number of specific tasks, affecting a range of fields such as dermatology, cardiology and radiology.[5–7] The deep learning capacity of AI would enable systems to detect characteristics and patterns unrecognisable to humans. This was evident with the Deep Patient initiative, in which a research group at Mount Sinai Hospital, New York trained a program using the electronic health records of 700,000 patients and then used the program to predict disease in another sample of 76,214 patients.[8] They noted that the results "significantly outperformed" those obtained from raw health record data and alternative feature learning strategies. However, since the program teaches itself patterns, it is difficult to understand how these types of programs yield the results they do.[9] There are numerous other similar projects which have attracted sensational media headlines about AI replacing doctors.

Media headlines and recent research have focussed on the diagnostic capabilities of machine learning systems. However, as illustrated in Figure 1, there are many uses of AI which will not be immediately visible to patients. These uses require "deductive" AI systems, which can analyse data to reveal patterns which would be infeasible for humans to recognise. It is important to note that "generative" AI systems are also emerging, which can create synthetic data after being trained on a pre-existing dataset, though their use in clinical practice is limited. They do, however, have the potential to improve patient outcomes indirectly, such as through creating training material for clinicians.[10] As AI becomes increasingly capable, research progressively aims to integrate different technologies to realise synergies, eg with robotics, Internet of Things and patient record management systems. AI has the potential to exhibit substantial generativity and as the technology evolves, novel uses will certainly emerge. Some have even speculated about the possibility that AI may be used to iteratively enhance itself and that its ability to do so may eventually outperform the ability of human programmers. This may initiate a recursive self-improvement cycle, whereby improvements also improve the improvement process.[11] Experts in the field refer to this possibility as an intelligence explosion, or "singularity".[12] At this point, it is important to note a criticism of the practical implications of this prediction. Thus far, applied AI has proved extremely capable at performing focussed tasks, but progress towards artificial general intelligence (AGI) has been less groundbreaking. It can be argued that the growing amount of speculation about "singularity" is premature and diverts our attention from more immediate social and ethical issues resulting from current day machine intelligence.[13]

Artificial Intelligence as a Transformational Innovation

The wide range of uses of AI within the healthcare sector highlight its aforementioned role as an enabling technology. When conceptualising the innovation, it is important to consider the pronounced generativity and scope for future innovation. This can be achieved by applying a layered architecture model, as proposed by Yoo et al (2010).[14] A version of this model, applied to AI in healthcare, is shown in Figure 2 which illustrates multiple layers of the technology. Separation of layers of the technology facilitates innovation at each layer, which could include:

- Innovation at the contents layer to integrate a wide range of forms of patient data.
- Innovation at the service layer to extend diagnostics to different diseases.
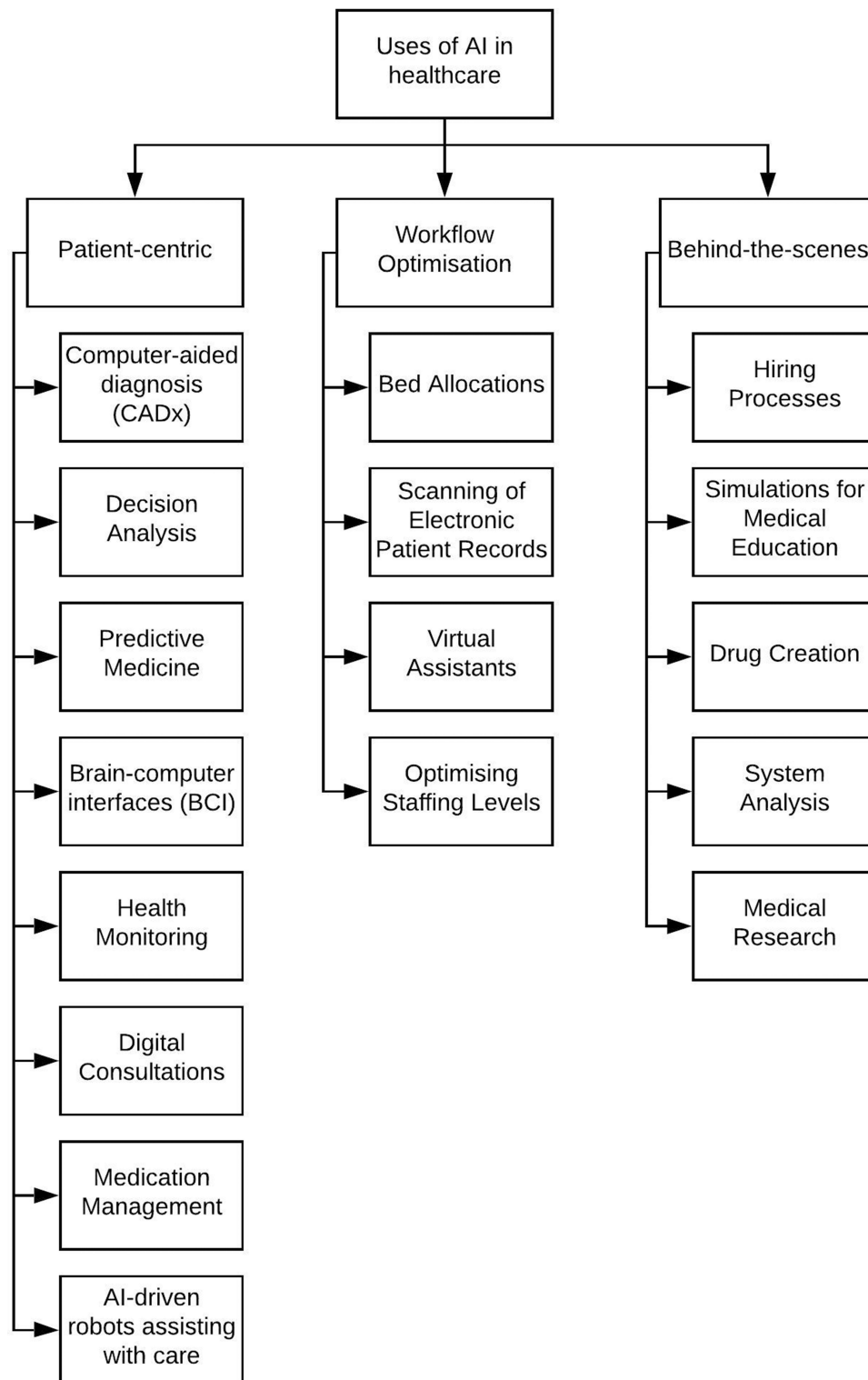
Figure 1 Overview of uses of AI in the healthcare industry.

- Innovation at the network layer to integrate information from different hospitals and expand datasets.

- Innovation at the device layer to develop physical machinery as well as AI logical capability. AI technology itself could even be used to perpetuate innovation.
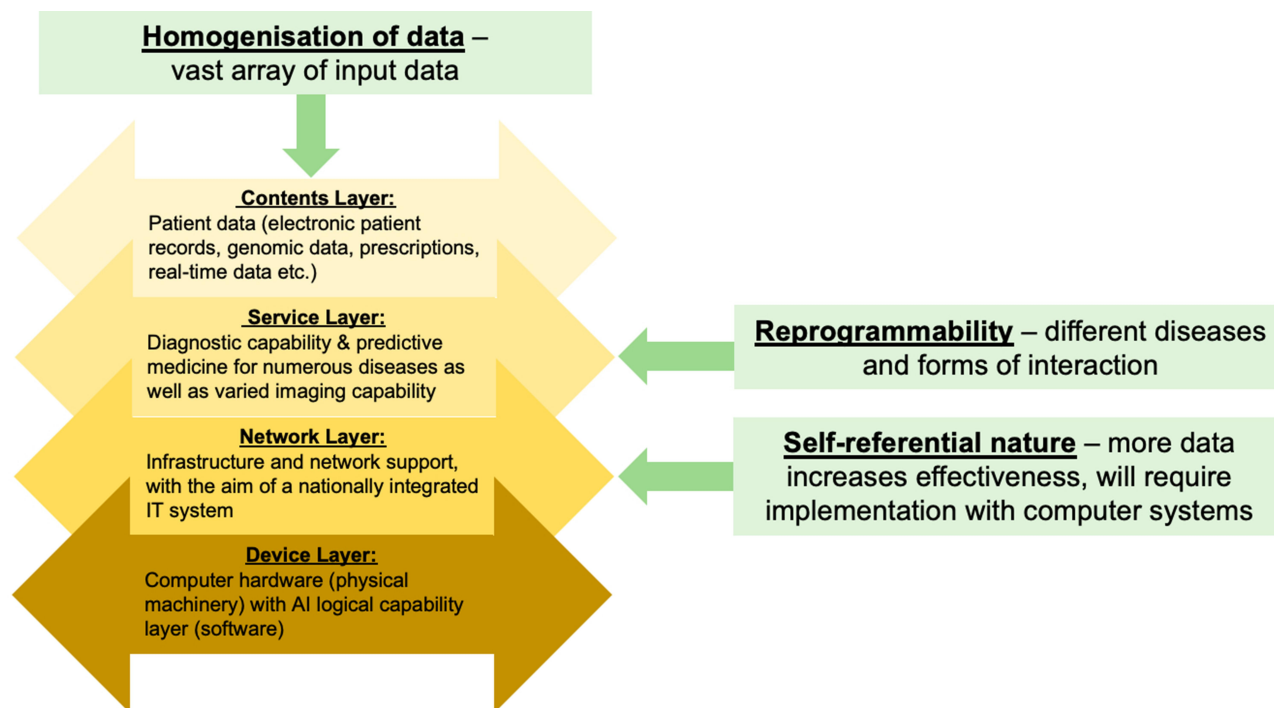
Figure 2 Application of the layered architecture model proposed by Yoo et al (2010),[14] illustrating the relevance of each proposed layer to AI as a digital healthcare innovation. The relevance of three key characteristics (homogenisation of data, reprogrammability and a self-referential nature) are also noted.

AI diagnostic systems rely on machine learning to identify patterns which would be infeasible to program. These systems require large and representative datasets to produce accurate outcomes. There is a credible threat of AI systems yielding biased decisions, either as a result of unrepresentative datasets or biased algorithms used to analyse those datasets. Before AI systems can be fully implemented, it must be ensured that training datasets are appropriate for the population that the algorithms will ultimately serve. For example, in dermatology, it has been noted that most algorithms are trained on Caucasian or Asian patients, but that these may yield inaccurate results if used for patients of other ethnicities unless the algorithms have been trained accordingly.[15] A wealth of information regarding individual patients is used by human doctors to assist decision making, including medical records, prescription history and real-time data. Much of the information, especially patients' symptoms, is not easily codifiable and therefore cannot easily be analysed. This issue of data homogenisation will be particularly difficult to overcome, especially considering that there is variability across healthcare systems in the form and style of medical records. Yoo et al (2010) argue that digital technologies possess three key characteristics (Figure 2): homogenisation of data, reprogrammability and a self-

referential nature. If homogenisation of data is achieved, this will permit reprogrammability as the logical capability of the technology can be altered to apply to a range of diseases. The self-referential nature of AI arises from its need to integrate with existing computer systems. As the technology diffuses and more data are acquired, the effectiveness of the technology will increase.[14] This adoption phase is a crucial stage of the innovation process, which is strongly affected by network effects. In this context, the term 'network effects' refers to the phenomenon by which if one healthcare facility begins to use a new system, the marginal benefit accrued by a second facility looking to incorporate the same system may increase as it can now benefit from a larger source of data being available. Vast change to existing IT infrastructure may be required before AI can be implemented. For example, in the United Kingdom, it has been noted that existing IT infrastructure is unfit for incorporating AI in the National Health Service (NHS) and fresh systems with relevant standards must be developed.[16] In this context, network effects may act against the adoption of new technologies, as organisations may be locked-in to old technologies.[17]

The scale of the innovation, in practical terms, can be assessed by considering the effect of the innovation on existing value chains. Innovations can be categorised as

innovations which change a product being delivered or innovations which change the process along a value chain. The use of AI in healthcare could be considered as both a product and process innovation. Indeed, the separation between the two types of innovation is inherently ambiguous, with Simonetti et al (1995) finding that 96.9% of innovations they studied fell into a "grey zone" between the two.[18] With these extensive effects on existing value chains, AI holds the potential to redefine capabilities of the healthcare industry in patient care and blur the boundaries of the industry as we see continued intrusion from large technological companies such as Google and Microsoft. In this way, there is potential for disruption of processes, markets and organisations within the healthcare industry, which Lucas et al (2013) emphasise as being characteristic of transformative innovations.[19]

Barriers to Adoption

As explored by Fichman et al (2014), digital process innovations offered by AI have the potential to change the administrative core, as well as the technical core, of large-scale healthcare organisations.[20] This may even lead to some degree of reformation of healthcare structures, or at least considerable changes to organisational practices. However, in order to fully appreciate the effects, we must consider the implementation period from the perspectives of the organisation, the state, the public and commercial companies. Contention between the priorities of these groups has the potential to delay the diffusion of innovation during the implementation period. Whilst all stakeholders may designate patient safety as a high priority, the balance of cost and ethical issues may heterogeneously affect decisions of the stakeholders.

The implementation period is well documented as being a key stage in the innovation process.[21] In this sense, it is not enough to simply invent a new product, we must also consider the wider surrounding organisational factors.[21] Technological innovation within the healthcare field is well represented by the punctuated model of equilibrium, as proposed by Loch and Huberman (1997).[22] This states that there are long periods of stability, punctuated by periods of rapid innovation. This is in part due to the long periods of testing that are required before innovations are recognised as safe and ethical as well as the slow rate of diffusion of technology. It is difficult to estimate the likely trajectory of AI; however, it can be predicted to surpass the advancements facilitated by technologies such as genetic sequencing,

electronic patient records and digital consultations which preceded it. AI will require cooperation with these previous technologies to realise its full potential. The ability of AI to cross-examine genomic information, patient histories and real-time patient data is critical to its ability to provide the most accurate diagnoses.

The rate of diffusion of electronic networks is considerably slower within healthcare systems than within society at large.[23] This is also complicated by the fact that AI implementation might require considerable reconfiguration of IT systems. It is difficult to allocate sufficient IT system investment to hospitals when there are competing demands for equipment and staff.[23] It is well recognised within the field that there is a moral imperative towards innovation in the healthcare industry, due to likely improvements in patient care.[24] However, even when IT systems indirectly improve patient outcomes, investment may be viewed less favourably than for items which directly improve patient care at the point of contact. For the politicians who make these decisions, cutting investment into staffing and equipment could be met with an unfavourable public response.

The economic effects of technologies such as artificial intelligence dictate that decisions regarding their implementation must be made at a national level.[25] In state-funded healthcare systems, the government's role is balancing the desire for the highest possible care standards with strict financial constraints. Relevant government policy and regulation is often influenced heavily by public opinion with input from stakeholders within the health service. In this way, it is not enough for companies to simply prove that their AI system is capable of reducing morbidity and mortality, they must gain the advocacy of key stakeholders and be acceptable to the public. The recent COVID-19 pandemic has accelerated implementation of tele- and digital health systems, which may assist in the drive towards digitalised healthcare delivery.

Practical Implications of Artificial Intelligence Implementation

Risks of privacy, control and dependence on IT systems are well characterised as general risks associated with the use of AI.[26] In the field of healthcare, this corresponds to data privacy concerns, ambiguous accountability of decision making and automation bias. As the diagnostic capability of the technology improves, supervision becomes increasingly difficult. If the diagnostic capability of technology does begin to outperform that of human doctors,

this introduces a risk of clinicians becoming susceptible to automation bias. Automation bias is a well-studied phenomenon which prevails when the doctor has more trust in the diagnostic ability of the technology than their own judgement.[27] This would confound occasions when there is a disagreement between the human doctor and the system. The doctor would likely have the right to overrule the system but would have to justify this to the patient. On that basis, doctors should have at least a basic understanding of the algorithms, probabilities and datasets that are responsible for suggested diagnoses. This would help them to identify the causes of incorrect diagnoses and facilitate direct alterations to the system. We can assume that the doctor should have some degree of control over the system, but this is difficult with no experience in statistics or computer programming. The literature also notes a risk that AI algorithms may perpetuate and amplify healthcare inequalities through biased algorithms.[28] Within the wider machine learning literature, there are two recognised steps that must be taken to ensure that AI acts to reduce, or at least not exacerbate, inequalities:[29,30]

1. Datasets should be representative of the general population, with existing biases noted and accounted for. Effort should be spent on ensuring that there are data available for minority groups and those who are less served by existing systems.
2. Active measures should be taken to ensure that AI algorithms do not perpetuate existing biases.

Corrigibility, interruptibility and interpretability are widely recognised as important traits to incorporate with AI technology.[31,32] These traits do not insist that clinicians develop the systems themselves, but they must have sufficient understanding to be able to interpret, interrupt and correct computer systems if necessary. Thus, in order to prepare clinicians for the introduction of AI, we should pay attention to coding and data techniques that may become important for its correct use. In the long term, doctors may also need to develop some statistical expertise to communicate probabilities generated by AI software to patients, in order that they may continue to communicate effectively to patients in the wider context of their disease.

Where AI systems have been used so far, the "black box" problem has been raised as a potential obstacle.[33,34] The "black box" phenomenon is so-called as machine learning algorithms leading to diagnoses are suggested to resemble a "black box", with no way of looking inside to see exactly how an answer is reached. One effort to overcome the black box problem is the development of a system, by researchers at DeepMind and Moorfields Eye Hospital, to interpret optical coherence tomography (OCT) scans, triage patients and recommend treatments.[35] The system developed clearly illustrates the decision-making process of the machine learning system that led it to a decision and this can be scrutinised by the clinician interpreting the results.[35,36] This inspection and verification by clinicians is particularly important for artificial intelligence as a medical device because of the risk of systematic iatrogenic harm if the algorithms are flawed. It has been noted that machine learning algorithms must be extensively scrutinised, audited and debugged if they are to be used in clinical practice.[3] Indeed, it is recognised that whilst AI is unlikely to replace human doctors, it will facilitate the emergence of data-driven high-performance medicine with clinicians retaining some degree of control over the systems.[3]

Concluding Remarks

Conceptualisation of AI as a digital healthcare innovation, and the use of recognised frameworks to do so, reveals insight into the positive potential and the risks that AI offers. Whilst much of what has been discussed may appear to be hyperbolic, in places optimistic and in others pessimistic, it is important to acknowledge these as the two polarised viewpoints that shape 21st century AI debate. The optimism is sufficiently plausible that it warrants further research and the pessimism such that it warrants a cautious approach. AI has the potential to improve healthcare delivery through altering clinical practice as well as optimising workflows. As an innovation, AI exhibits three key characteristics: it is self-referential, reprogrammable and capable of pronounced generativity. There are a number of potential barriers which may delay adoption of emerging AI technologies and several risks associated with their use. It is important that this is considered before the technology is adopted, considering the path dependency typically exhibited in patterns of innovation diffusion. On this basis, we should proactively adapt organisational practices in order to cope with the transformational impacts of AI. Whilst the timescales for these impacts are uncertain, we should not assume that they are far away.

Disclosure

The author reports no conflicts of interest in this work.

References

1. Paschen U, Pitt C, Kietzmann J. Artificial intelligence: Building blocks and an innovation typology. *Bus Horiz*. 2019. doi:10.1016/j.bushor.2019.10.004

2. European Commission [Internet]. Re-Finding Industry: Report from the High-Level Strategy Group on Industrial Technologies. Publications Office of the EU; 2018. Available from: https://ec.europa.eu/research/industrial_technologies/pdf/re_finding_industry_022018.pdf. Accessed 12 April 2020.

3. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019;25(1):44–56. doi:10.1038/s41591-018-0300-7

4. US Food and Drug Administration [Internet]; January 2020. Available from: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device. Accessed April 13, 2020.

5. Haenssle HA, Fink C, Schneiderbauer R, et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Ann Oncol*. 2018;29(8):1836–1842. doi:10.1093/annonc/mdy166

6. Steele AJ, Denaxas SC, Shah AD, Hemingway H, Luscombe NM. Machine learning models in electronic health records can outperform conventional survival models for predicting patient mortality in coronary artery disease. *PLoS One*. 2018;13(8):e0202344. doi:10.1371/journal.pone.0202344

7. Rajpurkar P, Irvin J, Zhu K, et al. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. *arXiv*. 2017;2:1–58.

8. Miotto R, Li L, Kidd BA, Dudley JT. Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records. *Sci Rep*. 2016;6(1):26094. doi:10.1038/srep26094

9. Krisberg K Artificial Intelligence Transforms the Future of Medicine [Internet]. Available from: https://news.aamc.org/research/article/artificial-intelligence-transforms-future-medicine/. Accessed 3 March 2019.

10. Arora A. Disrupting clinical education: Using artificial intelligence to create training material. *Clin Teach*. 2020;17(4):357–359. doi:10.1111/tct.13177

11. Bostrom N. Superintelligence: Paths, Dangers, Strategies. OUP Oxford; 2014.

12. Tegmark M. Life 3.0: Being Human in the Age of Artificial Intelligence. 01. Allen Lane; 2017.

13. Aicardi C, Fothergill BT, Rainey S, Stahl BC, Harris E. Accompanying technology development in the Human Brain Project: From foresight to ethics management. *Futures*. 2018;102:114–124. doi:10.1016/j.futures.2018.01.005

14. Yoo Y, Henfridsson O, Lyytinen K. Research Commentary—The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research. *Info Sys Res*. 2010;21(4):724–735. doi:10.1287/isre.1100.0322

15. Chan S, Reddy V, Myers B, Thibodeaux Q, Brownstone N, Liao W. Machine Learning in Dermatology: Current Applications, Opportunities, and Limitations. *Dermatol Ther*. 2020;10(3):365–386. doi:10.1007/s13555-020-00372-0

16. AHSN Network AI Report [Internet]. https://wessexahsn.org.uk/img/news/AHSN%20Network%20AI%20Report-1536078823.pdf. Accessed July 21., 2020.

17. Katz M, Shapiro C. Systems Competition and Network Effects. *J Econ Perspect*. 1994;8(2):93–115. doi:10.1257/jep.8.2.93

18. Simonetti R, Archibugi D, Evangelista R. Product and process innovations: How are they defined? How are they quantified? *Scientometrics*. 1995;32(1):77–89. doi:10.1007/BF02020190

19. Lucas HC, Agarwal R, Clemons EK, Sawy OA, Weber B. Impactful Research on Transformational Information Technology: An Opportunity to Inform New Audiences. *MIS Quarterly*. 2013;37(2):371–382. doi:10.25300/MISQ/2013/37.2.03

20. Fichman RG, Dos Santos BL, Zheng Z. (Eric). Digital Innovation as a Fundamental and Powerful Concept in the Information Systems Curriculum. *MIS Quarterly*. 2014;38(2):329–A15. doi:10.25300/MISQ/2014/38.2.01

21. Garud R, Tuertscher P, Ven AHV. Perspectives on Innovation Processes. *Acad Manag Ann*. 2013;7(1):775–819. doi:10.1080/19416520.2013.791066

22. Loch CH, Huberman BA. A Punctuated-Equilibrium Model of Technology Diffusion. *Manage Sci*. 1997;45(2):160–177. doi:10.1287/mnsc.45.2.160

23. Dawson S, Sausman C, eds. Future Health Organizations and Systems. Palgrave Macmillan; 2005.

24. Lawton Robert. The Business of Healthcare Innovation Edited by Lawton Robert Burns. Cambridge University Press; 2005.

25. Dodgson M, Gann DM, Salter A. The Management of Technological Innovation: Strategy and Practice (Revised Ed.). Oxford: Oxford University Press; 2008.

26. Newell S, Marabelli M. Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of 'datification. *J Strategic Inf Sys*. 2015;24(1):3–14. doi:10.1016/j.jsis.2015.02.001

27. Academy of Medical Royal Colleges. Artificial Intelligence in healthcare [Internet]; 2019. Available from: http://www.aomrc.org.uk/reports-guidance/artificial-intelligence-in-healthcare/. Accessed 2 March 2019.

28. Panch T, Mattie H, Atun R. Artificial intelligence and algorithmic bias: implications for health systems. *J Glob Health*. 2019;9(2). doi:10.7189/jogh.09.020318

29. Nordling L. A fairer way forward for AI in health care. *Nature*. 2019;573(7775):S103–S105. doi:10.1038/d41586-019-02872-2

30. Geneviève LD, Martani A, Shaw D, Elger BS, Wangmo T. Structural racism in precision medicine: leaving no one behind. *BMC Med Ethics*. 2020;21(1):17. doi:10.1186/s12910-020-0457-8

31. Dafoe A AI Governance: A Research Agenda. Governance of AI Program, Future of Humanity Institute University of Oxford.2018. Available from: https://www.fhi.ox.ac.uk/wp-content/uploads/GovAIAgenda.pdf. Accessed 9 March 2019.

32. Hernandez-Orallo J, Martınez-Plumed F, Avin S, O hEigeartaigh S. Surveying Safety-relevant AI Characteristics, Centre for the Study of Existential Risk [Internet]. January 2019. Available from: http://ceur-ws.org/Vol-2301/paper_22.pdf. Accessed 9 March 2019.

33. London AJ, Intelligence A. Black-Box Medical Decisions: Accuracy versus Explainability. *Hastings Center Rep*. 2019;49(1):15–21. doi:10.1002/hast.973

34. Baum S Healthcare must overcome AI's "Black Box" problem. MedCity News [Internet] 2019. Accessed July 24. https://medcitynews.com/2019/01/healthcare-must-overcome-ais-black-box-problem/.

35. Fauw JD, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med*. 2018;24(9):1342–1350. doi:10.1038/s41591-018-0107-6

36. Deepmind. A major milestone for the treatment of eye disease [Internet]. Accessed July 24., 2020. https://deepmind.com/blog/article/moorfields-major-milestone.

Medical Devices: Evidence and Research

Dovepress

Publish your work in this journal

Medical Devices: Evidence and Research is an international, peer-reviewed, open access journal that focuses on the evidence, technology, research, and expert opinion supporting the use and application of medical devices in the diagnosis, monitoring, treatment and management of clinical conditions and physiological processes. The identification of novel devices and optimal use of existing devices which will lead to improved clinical outcomes and more effective patient management and safety is a key feature of the journal. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit http://www.dovepress.com/testimonials.php to read real quotes from published authors.

Submit your manuscript here: https://www.dovepress.com/medical-devices-evidence-and-research-journal