

Are we ready for Artificial Intelligence in Medicine?

Jeffrey David Iqbal, Rasita Vinay

Institute for Biomedical Ethics and History of Medicine, University of Zurich, Switzerland

The headlines keep coming in thick and fast on Artificial Intelligence (AI) in Medicine. There is rarely a day without a new research article being published with claims of superior accuracy or other performance measure in some medical function. Such research often promises impacts on screening, diagnosis, or monitoring, and increasingly, also on treatment and prediction. While the use of AI has been heralded to revolutionise medicine in numerous specialties, its use in broad clinical applications has been limited. Some countries, including Switzerland, still do not have widespread implementation of basic electronic health records systems upon which more advanced AI applications can be solidly built. Such intensive publicity exaggerating a technology's potential benefits must be referred to as a "hype".

Artificial Intelligence as a field developed in the 1950s, when computer technology advanced and its use expanded rapidly. Since then, periods of excitement about its potential and vast funding were often followed by periods of disillusionment and de-funding, which were termed "AI winters" in the early nineties [2]. The field has benefited from vast funding by both private and public sectors and the frenzy of research & development activity from Pharma/MedTech, start-ups and big tech players. With the onset of more widespread adoption [3], at least some of these developments will ultimately find their way into clinical prac-

tice. Technological developments and implementations often follow the hype cycle, and this will likely also extend to AI [4].

Whether we are ready for what is currently happening is another question rarely discussed. In this viewpoint, we would like to briefly introduce: 1) what artificial intelligence in medicine is, 2) what elements constitute readiness and 3) what main arguments can be brought forward by each side of the debate on whether we are ready for Medical AI.

What is artificial intelligence in medicine?

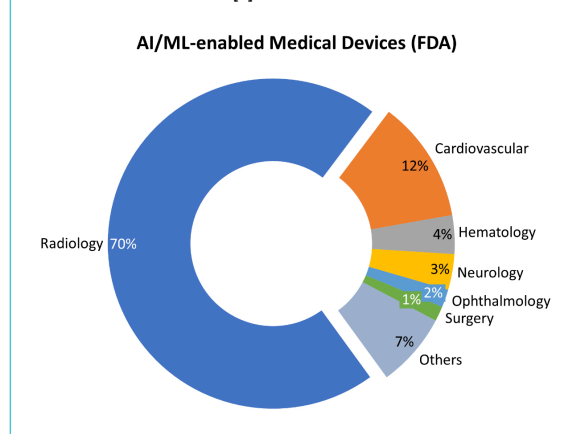
"Any sufficiently advanced technology is indistinguishable from magic." (Arthur C. Clarke, British futurist and author)

There is no single definition of what constitutes Artificial Intelligence. It is typically referred to as the field of computer systems able to perform non-physical tasks normally requiring human intelligence. In this definition, it refers to human intelligence that itself has no single definition, but most generally is the capacity to deal flexibly and effectively with practical and theoretical problems. AI subsumes technology fields, such as machine learning (the development and application of computer algorithms for transforming data into intelligent action) or deep learning (a type of machine learning technology using large artificial neural networks) and can apply to all medical functions and disciplines.

Here, it is important to distinguish that AI is not the same as digitisation, but rather a part of it. Digitisation in medicine refers to the wider field of using computer systems for the provision or support of healthcare delivery. Examples of technology within this sphere include Electronic Health Records (EHRs)—patient-specific health data in digital format—that belong to the Big Data category, and robotics—machines that are capable of physical actions normally performed by humans.

What is considered AI in Medicine has evolved over time. In 1954, a mechanical machine (a type of computer at the time) was proposed to assist clinicians in making diagnoses given the entry of individual symptoms [5]. In 1972, a Naïve Bayes-based System for acute abdominal pain developed at the University of Leeds took the approach to a digital format [6]. Such Computerized Decision Support

Figure 1: FDA approved/cleared/exempt Artificial Intelligence/Machine Learning-enabled devices from 1997 to Oct-2021 by specialty (FDA Lead Panel). Data Source: U.S. Food and Drug Administration: Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices [1].



Correspondence:
Jeffrey David Iqbal, PhD
Institute for Biomedical
Ethics and History of Medi-
cine (IBME)
Winterthurerstrasse 30
CH-8006 Zürich
[jeffreydavid.iqbal\[at\]ujzh.ch](mailto:jeffreydavid.iqbal[at]ujzh.ch)

Systems (CDSS) are now in increasing use and are hardly considered magic though that is how they were viewed at the time.

This evolving yet blurred definition of what constitutes AI seems to be somewhat unsatisfactory, yet a conclusive list of technologies or performance criteria to fall within the definition of AI remains elusive and is yet to be recognised in literature.

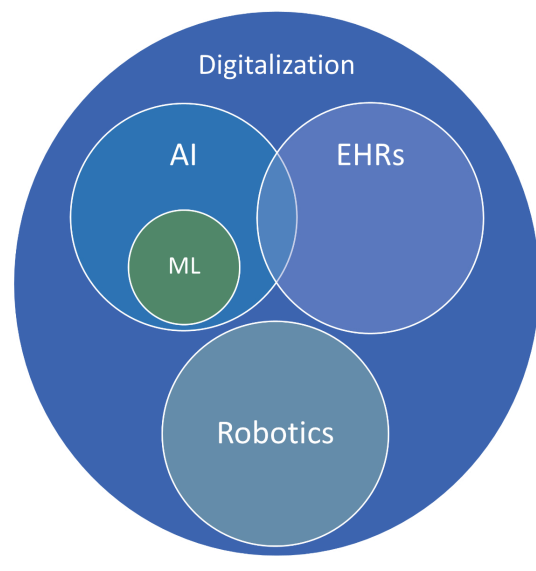
When can we speak of readiness?

There are several concepts subsumed under the term readiness. According to common dictionaries, it can either refer to a state of preparation for use or prompt willingness to use. While some stakeholders, e.g. Big Tech companies or health insurers, may be willing to use AI technologies in medicine, the question of preparation must be considered to maximise benefit and minimise harm for patients.

Readiness thus constitutes the preparedness by all relevant stakeholders in healthcare delivery. In the execution of medical functions, readiness ensures the safety and effectiveness of AI-based procedures. The fact that we are already using some AI-based tools is therefore not sufficient basis to claim that we are ready for their use as we could be using them without being prepared for their safe and effective use.

The preparedness of stakeholders for the use of a technology is evaluated according to three interdependent dimensions. The first is the technology/product that is to be employed and its maturity/readiness-for-use. Second, the users of the technology need to have sufficient knowledge and ability in its use. Third, if such use occurs in wider organisational or societal contexts, these need to be considered.

Figure 2: A rough conceptual attempt at differentiating between the spaces of digitisation, AI and robotics.



What speaks for readiness and what against?

Technology

AI-based technologies in medicine need to be evaluated on their purpose, safety and efficacy to determine their readiness. Such evaluations need to be conducted on the individual product-level rather than broader technology classes.

Consider the example of online symptom checkers. The idea of a related diagnosis-making machine originates from the 1950s and has been implemented in a limited way today in many examples of web-based applications. There, patients can enter their symptoms to see whether they should see a physician or not and are sometimes provided with an initial diagnosis and treatment recommendations for a limited set of diseases and conditions. A well-known example is the symptom checker provided by WebMD [7]. The risks to the patient-consumer associated with the use of such technologies can be well-controlled if set to err on the side of caution, i.e. recommending a physician visit if the symptoms do not reach the threshold of certainty for a non-critical diagnosis that an experienced and well-trained human physician could provide.

On the other hand, there are also technologies that are clearly not yet ready for deployment and are often used as examples in research literature. Or there are technologies that are deployed despite evidence they either lack adequate evaluation or they cause harm by making wrongful predictions [8–10].

It normally takes an average of 17 years for evidence-based practices (EBPs) to find their way into clinical practice [11]. The creation of such EBPs, technological development, takes between 2–5 years for medical devices and 10–15 years for pharmaceuticals [12]. This time ensures some level of maturity and provides an opportunity to recognize and eliminate safety issues, proving efficacy beyond the controlled conditions of clinical trials in the real world. It is a path yet to be pursued for AI technologies of today.

Users of AI technology

Preparedness for the use of a technology depends on the users, their knowledge, skills, and desire for use.

The example of a symptom checker is clearly directed at patient-consumers, rather than Healthcare Professionals (HCPs). The tool recommends over-the-counter or otherwise accessible treatment options in cases where illnesses/conditions requiring further treatment options are very unlikely. Otherwise, it directs the patient-consumer to seek the attention of HCPs. While there are some patient-consumers with low health literacy or language understanding leading to missed or wrong diagnoses and potential consequences, the risk is low. On the other hand, it can lead to users seeking HCP attention that otherwise would not have done so.

Other technologies are directed at HCP use and require significant knowledge and skills resulting from appropriate training. Given the current interest and advancements in AI technologies, medical training systems and curricula have not yet incorporated AI sufficiently but should be adopted in combination with professional clinical experience.

Strengthened medical education in AI would allow for medical students to familiarise themselves with the concepts and foundations of AI and to be able to apply them with a critical mindset in a medical context [13].

Context of use

Beyond the technology and its purpose, as well as its users (HCPs, patient-consumers), the context of use must be considered for preparedness. Safety and efficacy, in the classical sense, may be too narrow considerations within pharmaceutical and medical device established practices. AI applications in medicine need to fulfil further standards of fairness, uphold privacy, ensure autonomy and be transparent, among other principles [14]. These societal expectations provide context for use and thus need to be considered when evaluating preparedness. To date, many advanced AI applications have limitations in transparency that can be described as “black boxes”.

On the other hand, the discourse is polarised by putting up higher requirements on one side and technological promises on the other. Underlying healthcare delivery is an economic context that in many countries globally is increasingly strained. The healthcare spend relative to gross domestic product (GDP) in percentage has increased in almost all of OECD (Organization for Economic Cooperation and Development) member states between 2000 and 2019 [15]. Notable examples include the United States with an increase from 12.5 to 17.0% and Switzerland increasing from 9.8 to 12.5%, making these two healthcare systems the most expensive in the world. Mega-trends around aging populations, increasing prevalence of chronic disease, as well as a global labour supply shortage in the medical professions, have deteriorated the outlook further. From an economic perspective, we are overdue for any technology that can control the expansion of costs for healthcare delivery. However, it remains unproven whether Medical AI will deliver net cost savings on a systemic level.

Conclusion

Artificial Intelligence in Medicine is a definition currently relying on and limited by the emulation of human intelligence and has been shifting since its advent in the 1950s. Readiness can best be understood as preparedness for safe and effective use and can be evaluated along three dimensions: technology, users of the technology, and context of use.

As the definition of what constitutes AI shifts towards ever-advancing technologies for which there may often be a time lag in training of users, as well as in knowledge and skills development, we must conclude that we may not be ready for what is considered Artificial Intelligence today. It could be argued that we are ready for the AI technology that was developed a few years back and is now considered standard practice today.

At a debate evening focusing on the readiness for Medical AI, an audience of more than 200 students, researchers and members of the general public were asked to vote on their opinion both pre- and post-debate. The debate took place between an academic and politician, who argued against readiness, and an academic and business executive argu-

ing for readiness. Prior to the debate, the audience voted in favour of readiness (54% for, 21% against, 25% were undecided). The majority was even more clear in the post-debate vote (63% for, 26% against, 11% remained undecided).

Bill Gates and Paul Allen put out the vision of “a computer on every desktop and in every home” at the onset of Microsoft. What seemed unlikely at the time has long been the reality in many parts of the world. Much of our medical routine and processes of today will be superseded in 20–30 years’ time. We may not be ready, but we better catch up with technology and shape that future.

Potential competing interests

Both authors have completed and submitted the International Committee of Medical Journal Editors form for disclosure of potential conflicts of interest. No potential conflict of interest was disclosed.

References

1. U.S. FDA. Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices [Internet]; 2021 [cited 2021 Nov 23]. Available from: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices>
2. Crevier D. AI - the tumultuous history of the search for artificial intelligence. New York: BasicBooks; 1993.
3. IQVIA Institute. Digital Health Trends 2021 [Internet]; 2021 [cited 2021 Aug 27]. Available from: https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/digital-health-trends-2021/iqvia-institute-digital-health-trends-2021.pdf?_=1628089218603
4. Gartner Inc. The 4 Trends That Prevail on the Gartner Hype Cycle for AI, 2021 [Internet]; 2021 [cited 2021 Nov 20]. Available from: <https://www.gartner.com/en/articles/the-4-trends-that-prevail-on-the-gartner-hype-cycle-for-ai-2021>
5. Nash FA. Differential diagnosis, an apparatus to assist the logical faculties. *Lancet*. 1954 Apr;266(6817):874–5. [http://dx.doi.org/10.1016/S0140-6736\(54\)91437-3](http://dx.doi.org/10.1016/S0140-6736(54)91437-3). PubMed. 0140-6736
6. de Dombal FT, Leaper DJ, Staniland JR, McCann AP, Horrocks JC. Computer-aided diagnosis of acute abdominal pain. *BMJ*. 1972 Apr;2(5804):9–13. <http://dx.doi.org/10.1136/bmj.2.5804.9>. PubMed. 0007-1447
7. Web MD. LLC. WebMD Symptom Checker with Body Map [Internet]; 2020. Available from: <https://symptoms.webmd.com/>
8. Wong A, Otles E, Donnelly JP, Krumm A, McCullough J, DeTroyer-Cooley O, et al. External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalized Patients. *JAMA Intern Med*. 2021 Aug;181(8):1065–70. <http://dx.doi.org/10.1001/jamainternmed.2021.2626>. PubMed. 2168-6114
9. Thompson HM, Sharma B, Bhalla S, Boley R, McCluskey C, Dligach D, et al. Bias and fairness assessment of a natural language processing opioid misuse classifier: detection and mitigation of electronic health record data disadvantages across racial subgroups. *J Am Med Inform Assoc*. 2021 Oct;28(11):2393–403. <http://dx.doi.org/10.1093/jamia/ocab148>. PubMed. 1527-974X
10. Szalavitz M. A Drug Addiction Risk Algorithm and Its Grim Toll on Chronic Pain Sufferers [Internet]. *WIRED*; 2021 [cited 2021 Aug 26]. Available from: <https://www.wired.com/story/opioid-drug-addiction-algorithm-chronic-pain>
11. Bauer MS, Damschroder L, Hagedorn H, Smith J, Kilbourne AM. An introduction to implementation science for the non-specialist. *BMC Psychol*. 2015 Sep;3(1):32. <http://dx.doi.org/10.1186/s40359-015-0089-9>. PubMed. 2050-7283
12. Van Norman GA. Drugs, Devices, and the FDA: Part 2: An Overview of Approval Processes: FDA Approval of Medical Devices. *JACC Basic Transl Sci*. 2016 Jun;1(4):277–87. <http://dx.doi.org/10.1016/j.jaccbts.2016.03.009>. PubMed. 2452-302X
13. Karaca O, Çalıřkan SA, Demir K. Medical artificial intelligence readiness scale for medical students (MAIRS-MS) - development, validity and reliability study. *BMC Med Educ*. 2021 Feb;21(1):112. <http://dx.doi.org/10.1186/s12909-021-02546-6>. PubMed. 1472-6920
14. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. *Nat Mach Intell*. 2019;1(9):389–99. <http://dx.doi.org/10.1038/s42256-019-0088-2>. 2522-5839

15. OECD. OECD Health Expenditure Indicators [Internet]. OECD; 2017 [cited 2021 Nov 20]. Available from: <https://doi.org/> <http://dx.doi.org/10.1787/8643de7e-en> .