Ulster Med J 2023;**92(3):**167-169

Medical Ethics

Artificial Intelligence in Medicine

Andrew Cupples

Amara's Law states that "we tend to overestimate the effect of technology in the short run and underestimate the effect in the long run."[1] Media reports over the last few months have been brimming over with reports of recent advances in the field of Artificial Intelligence, breathlessly predicting how dramatically our society will change in the near future - predictions of wide scale job losses as many tasks are offloaded onto AI, not just the simple mechanical tasks of administration and organisation, but also the complex higher level work of the professions, such as law, finance and medicine.

It's important to clarify a few points surrounding the AI discussion. What we have now is what is traditionally known as 'weak' AI - computers being computers, processing commands, retrieving data, running software designed by humans in order to fulfil tasks directed by humans. As computers become more powerful, these things run better and faster, with higher accuracy and greater scope; but weak AI still consists of computers being computers, just doing it really well. A good example is Deep Blue, the IBM designed chess playing computer, which does a better job of playing chess than even the top players; but ask it to do your accounts for you or write a poem and it won't have a clue - like a train running on rails, it's going to travel faster than a human in a straight line, but only to a pre-determined destination. What we *think* of as AI is usually referred to as 'strong' AI, or Artificial General Intelligence, which is where machinery and software breaks free from the constraints of pre-determined patterns, and gains the ability to learn, adapt and change - bearing a resemblance to our own development as humans, compared to a young child using what he or she observes from the world around them, being taught, growing in knowledge but also learning to synthesise and predict, not merely in an iterative way, but in a complex and multifaceted form of development, with the result that entirely novel situations can be negotiated successfully, with that knowledge applied to a broad range of problems of different natures - Deep Blue giving you relationship advice, or writing a sonata; less of a train, and more of a helicopter, moving in multiple dimensions at high speed. Oh, and it doesn't need a pilot. Are we there yet? Has strong AI been developed; or perhaps a question would be - has strong AI evolved?

You may be familiar with the so-called Turing Test[2] - a relatively simple test conceptualised by Alan Turing in the 1950's, to see if machine could trick a human into recognising it as a fellow human, by having a text based conversation. Turing surmised that if a machine was able to trick a human into believing that it was a fellow homo sapiens, then the machine could be considered to be 'thinking', in a similar way to a human being. It's a somewhat flawed test, in that what is really being assessed is the ability of a machine to emulate a human, but nonetheless the fact that the Turing test was successfully passed by Google's LaMDA model in June 2022 is a milestone in the development of thinking machines. However, the fact remains that strong AI as a concept remains far beyond the reach of any current model of computing; indeed, it may be that computers and machines may never develop sentience or consciousness, at least in a way which mirrors these uniquely human characteristics, without being merely very clever mimics.

The real news relates to what are known as Large Language Models (LLMs) - you may be familiar with Open AI's ChatGPT or Microsoft's Bard tools; the theory goes that if you have a model of learning called a neural network, which is a little like our brain, consisting of multiple little connections between lots of factors, and then you feed it lots of data - for example the entire internet, all the research articles in the last 50 years - and then set it to work making multiple interconnected links between all the data, then give it an interface by which people can interact with it, then essentially you can ask questions of it and make requests of it, and it will generate a response which will be understandable, based on factual information, and probably in keeping with what a human would come up with if they had the time and inclination to do so. And people are starting to realise that LLMs are actually really good at this. OK, there have been a few false starts, and a few scary bits - like Microsoft's LLM Chatbot called Sydney that went a bit rogue and tried to convince a reporter to leave his wife for the chatbot, stating that it loved him, for example[3]; and there's the phenomenon of LLM hallucinations, where ChatGPT will generate a perfectly logical and well researched article, with full references, but unfortunately will throw in a few references to articles which are entirely made up - so for example we have the story recently of the US attorney whose case rested on a very well reasoned submission, which was found to be full of entirely false case law references[4] - and it turned out that his solicitor who'd prepared the brief had just plugged it into ChatGPT and handed him what came out the other end, without checking it first! But those are blips which are being ironed out. Bugs will be squashed, progress is inevitable. Large Language Models of AI are here to stay, whether or not governments are able to legislate for them, whether or

not companies are able to lock them down, or make them prohibitively expensive. Right now you can download an LLM to your computer which will operate entirely offline - a couple of gigabytes of stuff that lives on your computer, no internet connected; you could train it on all your word documents and PDFs, feed it all your journal entries and appraisal reflections - responses will be shaped to your own context, preferences and perhaps prejudices. Consider for a moment an offline LLM trained on a patient's entire medical records, how the summing up of their entire medical experience could be a game-changing tool in personalising their medical care?

Ethos

The advances in AI have the potential to shake us at a very deep level. They force us to ask questions about how we consider ourselves and our patients as humans, and how we consider ourselves as doctors and professionals. They also force us to ask difficult questions about the future of our careers and indeed our profession. I would contend, however, that our approach to this new technology should be the same as our approach to any new technology; indeed, we can gain much by reflecting on how past generations and peoples have viewed technological developments, and how they have been considered in literary and religious contexts, because these reflect the zeitgeist, the broader mood of society, often to a greater degree than the thoughts of those of us in the ivory towers of the professions and academia.

We can start by examining our earliest pre- or proto-historical accounts, coming from the ancient Near East. The opening chapters of the book of Genesis in the Bible contains a repeating pattern of fallen humanity setting itself over and against a creator God in the pursuit of knowledge and self creation, and the resulting consequences far exceed the expectations of the hapless humans at fault - Eve wishes to be like God by seeking forbidden knowledge, and eviction from Paradise results, with humanity cursed[5]. We can look also at the ancient Akkadian recounting of the flood myths of Atrahasis and Gilgamesh for parallel accounts of human expansion without divine permission, punished with cataclysmic flooding[6] mirroring the Biblical account of Noah[7]. The story of Prometheus stealing fire from the Olympian gods to further the advancement of the humans to whom he is sympathetic results in the first documented example of hepatic resection by bird of prey[8]. However, the pinnacle of human self creation is found in the story of the tower of Babel[9] - technology used to pursue godlikeness, with the consequent disruption of human society and loss of free communication. The warning here from the Ancients is that the wrongheaded pursuit of knowledge and technological advancement runs the risk of catastrophe and cataclysm.

In another three steps we can jump from the pre-modern era, into the modern, and onwards into our post-modern society by looking at three characters in literature - the Golem, Frankenstein's monster and the Robot. The Golem is a man of clay, made by a Jewish priest, animated by holy words placed in his mouth, which runs amok, finally being defeated by its priestly creator; heavenly justice is served[10]. Frankenstein's monster is a creation of a scientist, a medical student, which gains sentience and finally has its revenge on its creator, but is banished to the Arctic wastes; justice is done in keeping with the principles of an Enlightenment morality[11]. Rossum's Original Robot, from the 1920s play by Czech writer Karel Capek[12], sees a scientist discovering a substance which allows him to generate a race of superhuman beings, with the result that humanity fades away, and the Robots ascend, with the final scene seeing the last human falling in love with a Robot, a new Adam and Eve, looking forward to a new creation of post-human hybrids.

It could be said that the prevailing mood of early historical and religious thought and popular literature is one of warning against unconditional experimentation and acceptance of the new technologies, especially when they are divorced from well established moral frameworks. The word which best sums this up is *hubris* - human overreach and pride, neglectful of past lessons and ignorant of possible consequences. If you have got to this point and are assuming that I am universally negative on the subject of AI, then I would like to add some balance to the discussion.

A better word for us to consider as healthcare professionals as we consider these technological developments is *synthesis*.

Doctors excel at synthesis. We collect the strands of information from a carefully taken clinical history, a focussed examination, appropriate diagnostic tests properly interpreted, and weave them together into a clear clinical picture, set against the correct sociocultural background of our patient; we compare the patterns of disease with our past experiences, and draw on our constantly updated understanding of disease processes to formulate our diagnosis; we formulate a management plan based on the current research evidence, contextualised by our available resources (and, especially at present, the limitations of those resources), and then we work collaboratively with our patients to decide on the best of course of action, which we then act upon, follow up, review, amend, working with many other colleagues all doing different things, all day long. In a word, this is synthesis - the drawing together of multiple diverse strands to make something new and good. Despite the headlines, we are a long way off machines being able to synthesise like human doctors, so our jobs are safe for the foreseeable future. Part of the problem is that, for those on the outside looking in, the process of medicine looks as if it could be reduced to a set of rules and flow charts, and this is the mistake made by the technologists, that reductionism which is so tempting to apply to things which we only superficially understand. It's difficult to imagine machines developing a gut instinct about a clinical situation, or utilising the sense of smell to guide antibiotic choice, for example.

Where we need to continue this concept of synthesis is

by critiquing the technologies being offered to us with an appropriate mindset

The possibilities are exciting - in the first instance, looking at all of the labour saving potential for our humdrum administrative workload - AI driven voice recognition for accurate record taking and automation of clinic letters and reports[13]; automated rota creation; and secondly the possibilities for aiding our clinical work - consider how powerful search functions and decision support tailored to patients' own records might add to our existing scope, allowing prescribing safety improvements and workload reductions thanks to automated checking of lab results and test reports[14], screening out the normal and allowing us to focus on the relevant; consider medical education in the near future whenever the student's progress can be monitored and assessed, and their programme tailored to focus on their weak points with a virtual personal tutor; consider diagnostic monitoring systems which would be able to not just monitor multiple vital signs, but give recommendations and even administer appropriate treatment without human intervention.

However, the challenge with each new medical technology is that we lose the humanity of our profession. With each new tool introduced the temptation is to be further removed from the patient. We neglect the real, sticky, complicated person in front of us, preferring the mediated person - filtered through the computer, the telephone line, the VR headset, the algorithmic reduction of their wholeness to a set of measurements and problem lists. And technological advance also brings with it inequality - healthcare is expensive, and medical technology especially so. Advances which provide AI drive healthcare will be cash cows for wealthy companies, and risk compounding the two tier service that is developing rapidly, driving inequalities[15]. How about the effects to staffing? Administration could be centralised and automated, and the risks to roles of administrative and paramedical staff (and indeed professional roles) are significant[16]. And we can consider the thorny question of risk and professional liability - who takes responsibility for errors created or caused by machines[17]? Like it or not, changes in healthcare technology will change our profession, but they are unlikely to abolish it in the short to medium term - while real human patients require healthcare, they will need and want human doctors to be the touchpoint, the personal medium for technology and ultimately the responsible agents. As in each previous generation of medicine our role will be to synthesise - drawing together the multiple different and changing strands for the good of our patients.

REFERENCES

1. Wilmott C. On the scientific plausibility of transhumanism. Chaper 5. In: MacKellar C, Stammers T, editors. The ethics of generating posthumans: philosophical and theological reflections on bringing new persons into existence. London: Bloomsbury; 2022.
2. Turing A. Computing machinery and intelligence. [Internet]. *Mind*. 1950; 59(236): 433-60. [cited Aug 16 2023]. doi.org/10.1093/mind/LIX.236.433
3. Yerushalmy J. Artificial intelligence (AI) 'I want to destroy whatever I want': Bing's AI chatbot unsettles US reporter. [Internet]. *The Guardian*.2023; Feb 17. [cited Aug 12 2023]. Available from: https://www.theguardian.com/technology/2023/feb/17/i-want-to-destroy-whatever-i-want-bings-ai-chatbot-unsettles-us-reporter?ref=the-pillar
4. Schnitzer K, DeGregory P. 'Humiliated' NY lawyer who used ChatGPT for 'bogus' court do profusely apologizes. [Internet]. *New York Post*. 2023; Jun 8. [cited Aug 16 2023]. Available from: https://nypost.com/2023/06/08/humiliated-lawyer-apologizes-over-chatgpt-court-flub/;
5. The Bible. Genesis 3:1-24.[Internet]. [cited Aug 16 2023]. Available from: https://www.bible.com/bible/111/GEN.3.1-24.niv.
6. Dalley SM. [Translator]. Myths from Mesopotamia: creation, the flood, Gilgamesh, and others. [Online]. Atrahasis composite translation. [cited Aug 16 2023]. Available from: www.omnika.org/texts/66
7. The Bible. Genesis 6:1-7:24. [Internet]. [cited Aug 16 2023]. Available from: https://www.bible.com/bible/111/GEN.6.niv.
8. Aeschylus. Prometheus bound. [Internet]. English translation unattributed. [cited Jul 16 2023]. Available from: http://classics.mit.edu/Browse/browse-Aeschylus.html
9. The Bible. Genesis 11:1-9. [Internet]. [cited Aug 16 2023]. Available from: https://www.bible.com/bible/111/GEN.11.1-9.niv
10. Michaelson J. Golem. Making men of clay: can imitating God extend to the creative realm? [Internet]. New York: Jewish Learning; 2023. [cited Aug 16 2023]. Available from: www.myjewishlearning.com/article/golem/
11. Shelley M. Frankenstein; or, the modern Prometheus. In three volumes. [Internet]. London: Lackington, Hughes, Harding, Mayor & Jones; 1818. [cited Aug 16 2023]. Available from: https://www.gutenberg.org/files/41445/41445-h/41445-h.htm
12. Čapek K, Mann W. Translator: Paul Selver. R.U.R. Rossum's Universal Robots: RUR (Rossumovi Univerzální Roboti) – a fantastic melodrama in three acts and an epilogue. [Internet]. New York: Samuel French, Inc.; 1920. [cited Aug 16 2023]. Available from: https://www.gutenberg.org/files/59112/59112-h/59112-h.htm
13. Cho A, Min IK, Hong S, Chung HS, Lee HS, Kim JH. Effect of Applying a Real-Time Medical Record Input Assistance System With Voice Artificial Intelligence on Triage Task Performance in the Emergency Department: Prospective Interventional Study. *JMIR Med Inform*. 2022; 10(8): e39892. doi: 10.2196/39892
14. Buch VH, Ahmed I, Maruthappu M. Artificial intelligence in medicine: current trends and future possibilities.. *Br J Gen Pract*. 2018;68(668):143-4.
15. Cheshire WP. First principles for medical AI. [Internet]. London: Christian Medical Fellowships [CMF]; 2023. [cited Aug 16 2023]. Available from: https://www.cmf.org.uk/resources/publications/content/?context=article&id=27451
16. Milmo D. AI revolution puts skilled jobs at highest risk, OECD says. . [Internet]. *The Guardian*.2023; Jul 11. [cited Aug 12 2023]. Available from: https://www.theguardian.com/technology/2023/jul/11/ai-revolution-puts-skilled-jobs-at-highest-risk-oecd-says Browsed 4th September 2023.
17. Maliha G, Gerke S, Cohen IG, Parikh RB. Artificial Intelligence and liability in medicine: balancing safety and innovation. *Milbank Q. 2021;99(3):629-47*. https://doi.org/10.1111/1468-0009.12504.