

The background features a dark field with vertical lines in shades of blue and purple. Overlaid on this are several large, semi-transparent circles in vibrant colors: red, yellow, purple, and orange. The text is white and positioned in the upper left quadrant.

Confronting Dr Robot

Creating a people- powered future for AI in health

Summary

John Loder and
Lydia Nicholas
May 2018



nesta
Health Lab

About Nesta

Nesta is a global innovation foundation. We back new ideas to tackle the big challenges of our time.

We use our knowledge, networks, funding and skills - working in partnership with others, including governments, businesses and charities. We are a UK charity but work all over the world, supported by a financial endowment.

To find out more visit www.nesta.org.uk

About Nesta Health Lab

Nesta Health Lab is committed to being a centre of expertise on people-powered and data-driven health: we work with partners from the health, care, voluntary, community and social enterprise sectors to test and scale new ways for people to remain healthy.

Our work focuses on new sources of support, which make it possible for people to be more involved in their health, such as peer support; new sources of data, which improve people's knowledge about their health, such as citizen-generated smartphone data; and new sources of innovation, that generate new solutions, including our 100 day People Powered Results method for transforming systems.

Over the past few years, we have backed over 100 local health systems and individual organisations with more than £20 million of funding.

www.nesta.org.uk/health-lab



Confronting Dr Robot

Creating a people-powered future for AI in health

Executive summary

Artificial Intelligence could become part of the front door to healthcare. It could make the health system simpler, more accessible, more responsive, more sustainable, and put patients more in control. But there's a risk that the public could experience it more as a barrier than an open door, blocking access to care, offering opaque advice and dehumanising healthcare in every sense. We're now at a crucial moment when decisions are being made which will determine whether the technology develops into People Powered AI.

Artificial intelligence (AI) looks like it could be one of the transformative technologies of our era. Healthcare is rich in the data that AI thrives on, and in the kinds of questions that it can tackle. While the use of AI in healthcare is at an earlier stage than the hyperbole surrounding the technology might suggest, it is developing at pace, and this raises both significant opportunities and risks.

AI has delivered some striking results. There have been research trials that successfully use machine learning on images from, for example, radiology, dermatology and ophthalmology, to diagnose to a level of accuracy that matches clinicians' own abilities. This, and other AI developments, have led to the suggestion that machines are poised take the place of doctors.

However, today's AI is narrow and not capable of the holistic thinking and complex judgement required for many clinical tasks. While there are significant areas of medicine where more narrow applications of decision-making rules and expert pattern matching predominate, the path towards AI replacing humans is not solely determined by technical capability. Technology implementation will need to address trust, accountability and similar factors. And, at the same time, humans remain especially good at certain tasks, such as learning to identify rare situations from small amounts of data.

This jump to focusing on whether or not AI could replace doctors also potentially distracts from some far more immediate and likely applications of AI in health. It is far easier for AI to be adopted where there are no or few good alternatives on offer, than in areas where humans are effective and trusted. Areas where AI could be effective, and where there are few good alternatives include:

Advice and triage before seeing a doctor

When should someone first seek help from the health service?

Proactive care

When is the right time to intervene in the face of worsening symptoms?

Automated second opinion

How does the diagnosis and treatment I am getting compare to alternative options?

There are already AI products in the market in these areas, although the evidence base is not always sufficient. If AI was adopted in these areas it would be in a hugely influential position over our health and care. This could bring great benefits, but also comes with significant risks that need to be proactively managed and mitigated.

Applications of AI

1. Advice and triage before seeing a doctor

Most people find it hard to know exactly when to seek appropriate medical help. Twenty per cent of GP¹ appointments and 19 per cent of A&E attendances² are for minor medical problems that could be treated at home. This unnecessary or preventable demand creates significant pressure on health resources.

AI is already being used to offer healthcare advice and diagnoses directly to consumers. People can get access to a diagnostic app or access to a chatbot to share symptoms with, and use this advice to decide whether or not to seek further medical help. In this way AI is beginning to provide a form of triage into the healthcare system and, if developed correctly, could help solve a major issue: how to support the appropriate use of limited healthcare resources including reducing the unnecessary use of health services.

However, this use of consumer-facing AI could also generate a flood of unnecessary demand (from false positives or generally risk-averse advice); a new source of error in the system (from false negatives or other mistakes); and could widen health inequalities depending on the underlying business model. This area of development for AI also creates a new and urgent regulatory challenge.

These challenges need to be tackled, however, because this looks set to be a growth market and we are increasingly likely to see the use of AI as a highly influential entry point to the healthcare system.

A likely future is one where AI is a common first point of contact for health, and a front door to the health system - a highly influential position.

1. <http://selfcarejournal.com/wp-content/uploads/2015/09/IMS-1.3.105-16.pdf>

2. https://www.pagb.co.uk/content/uploads/2016/06/PAGB_AE_Executive_Summary_June-2015.pdf

2. Proactive care

AI is also becoming capable of extracting signals from real-time data and giving an early warning that a health problem is getting worse. For example, by listening to the breathing sounds of those with congestive heart failure to spot signs of deterioration. This could enable help to be directed to the people who need it in a more timely way, leading to a healthcare system that is more dynamic and responsive in the way it cares for people. Or it could generate a great deal of unnecessary concern, replace individualised conversations with standardised analytics, and generate an oppressive degree of monitoring.

AI would have an influence over who gets treated and when they get treated. While this could make the system much more dynamic, it could also make it more impenetrable, more unequal and less individualised.

3. Automated second opinion

AI could sit alongside doctors and offer an opinion on the same patient, for both diagnosis and treatment. This could be used to give patients a digital second opinion, which would enable them to challenge and advocate for their care more easily. It would also allow mass comparisons of physicians' decisions centrally, perhaps with a view to understanding variation in care better. This use of AI could have significant potential influence on power dynamics within health, as well as creating additional cost pressure on the system. The evidence that AI can do this reliably is not there, but products like IBM Watson claim to come close to this sort of functionality.

Both patients and managers will find it tempting to compare diagnosis and treatment to AI-generated opinions. This could be helpful if applied to the narrower questions where AI is competent, but high risk if misapplied to judgements beyond the capability of the technology or quality of underlying datasets.

If successful in these three categories, AI would be embedded into the workflows of the healthcare system, offering advice on who gets treated, when, and with what. This could make a significant positive impact on the health system. However, it also carries with it a great deal of responsibility, not only in terms of safety, accuracy and efficacy, but also in its impact on power and control for both patients and professionals. So, realising the significant potential benefits rests upon the appropriate development and use of the technology.

Consequences for power and autonomy

Like many digital technologies, AI can either democratise or centralise, empower or disempower, depending on the way it is implemented. It is easy to see how poorly designed and executed AI would be problematic:

- Poorly designed triage and prioritising systems could make the system even harder to access, making healthcare even **less simple** for patients.
- Increasing use of data-hungry analytics could **squeeze out dialogue**, which could make it harder to surface key details about the individual not captured in datasets.
- Opaque AI could **reduce transparency and accountability**.
- An over monitored patient is not helped to understand what the AI is saying and why, making it harder to control their care and leading to **reduced autonomy**.

Further, it is likely that those worst affected by these changes will be those with hard to diagnose conditions, complex social and health needs, and who already face disadvantage.

However, it is equally possible to imagine a future where the patient is significantly more empowered:

- AI makes it **simpler** to know when to seek help and get to the right person.
- Ensure patient and professionals are more prepared for and have more time for their conversation, so **more dialogue**.
- Patients find it easier, via home diagnostics and chatbot advice to understand their condition and to ask for help when they need it, so have **increased autonomy**.
- The ability to get a digital second opinion **increases transparency and accountability**.

Also the clinician could be freed from a lot of low priority work, fed useful insights, and more able to intervene at the right time.

Which of these futures turns out to be the real one is not purely determined by the technology itself, but by the choices that are made in its implementation. We need to be as careful in thinking about how new technology integrates with key relationships and pathways as with how it integrates key technical systems.

Therefore, in addition to core questions of whether it is **safe** and **effective**, we should also be applying the following principles to deliver what we call **People Powered AI**.

Principles for People Powered AI	Test
<p>Control. AI should give citizens a clearer and more timely understanding of their health and what should be done, in ways that support greater citizen confidence and control.</p>	<p>Patients should report higher levels of understanding of their condition, control of their health and confidence to manage it.</p>
<p>Simplicity. Well implemented AI should make it quicker and easier for patients to get a resolution to their problem. This requires clarity about the types of problem AI can deal with, and well defined boundaries beyond which human input is required, to avoid AI becoming an additional barrier.</p>	<p>Patients report that it takes less time, fewer steps, and less frustration to get to a resolution of their problem.</p>
<p>Dialogue. The conversation between doctor and patient should remain central. AI should support conversations - ensuring that they are with the right people, that it happens at the right time, and providing the information that supports it. AI should not degrade conversations by over-standardising or taking up unnecessary time.</p>	<p>Patients and professionals should report having higher quality conversations: more time to talk, clearer communication, better mutual understanding and more confidence in the decisions made.</p>
<p>Equity. AI should not be used in ways that exacerbate health inequalities. AI should help all citizens, and most particularly those who face the most challenges and disadvantage in relation to their health and wellbeing.</p>	<p>All previously mentioned metrics, analysed for equity.</p>
<p>Accountability. It must be possible for AI to be understood, questioned and held to account, otherwise AI could fundamentally disempower users - both citizens and health professionals. Without accountability (and the transparency underpinning it), the rest of the People Powered AI principles are hard to achieve - control, simplicity, dialogue and equity all require AI that can be understood and held to account by its users.</p>	<p>Pending European legislation (GDPR) allows for a right to an explanation of a decision from an algorithm. This should be maintained to provide the ability to scrutinise decisions and improve performance.</p>

These are principles that apply to any form of healthcare that aims to be humane and person-centred, but are not presently being applied to the design, development and implementation of AI.

AI development is being driven by private companies, who are not directly incentivised to think from the system point of view. If we sleepwalk into a situation where a small number of tech companies have already monopolised access to the data to build AI, and are selling into a health service which does not fully understand the technology they are buying, the capacity for public influence of AI will be much reduced.

There is currently a window of opportunity to shape the future of AI in health. Policymakers should set rules for AI and ownership of public data that ensure the public gets not only value for any data it decides to share, and privacy elsewhere, but also AI products that deliver maximum public benefit. This requires that both the providers and users of AI understand the technology, have the tools to shape the market, can understand the needs of citizens, and are able to work through the complexities of implementation. This can be achieved through the following four recommendations:

Recommendations

1 Public and clinical scrutiny:

Involve citizens and clinical professionals in the upstream design, development and implementation of the technology. This should include the requirement of mechanisms, such as public panels made up of citizens, that ensure technology development and implementation takes account of the demands and perspectives of citizens and healthcare professionals and ensures that People Powered AI principles are applied.

2 Controlled tests in real-world conditions:

Enable real-world experimentation of AI in designated test sites, with non-AI comparators, to understand how AI works in complex systems before wider take-up 'in the wild'.

3 Proactive market design:

System leaders actively engage in market design to maximise public benefit and ensure a plural market with genuine choice. This should include regulation that is upstream and proactive ('anticipatory regulation'), clarity over who owns both algorithms and data, and requiring adherence to key design principles, such as People Powered AI principles. Market design should also foster a diversity of new entrants to the market including procurement processes that work for smaller companies and market structures that support a diverse range of R&D activities.

4 Decision-makers equipped to be informed users:

Create a new cadre of public leaders and decision-makers with the technical skills, authority and institutional levers to scrutinise, manage and deploy AI in a responsible way. This should include incorporating artificial intelligence into medical education and health management training to enable the frontline workforce to be informed users of the technology.



58 Victoria Embankment
London EC4Y 0DS

+44 (0)20 7438 2500

information@nesta.org.uk

 [@nesta_uk](https://twitter.com/nesta_uk)

 www.facebook.com/nesta.uk

www.nesta.org.uk

May 2018

Nesta is a registered charity in England and Wales with company number 7706036 and charity number 1144091. Registered as a charity in Scotland number SCO42833. Registered office: 58 Victoria Embankment, London, EC4Y 0DS.

