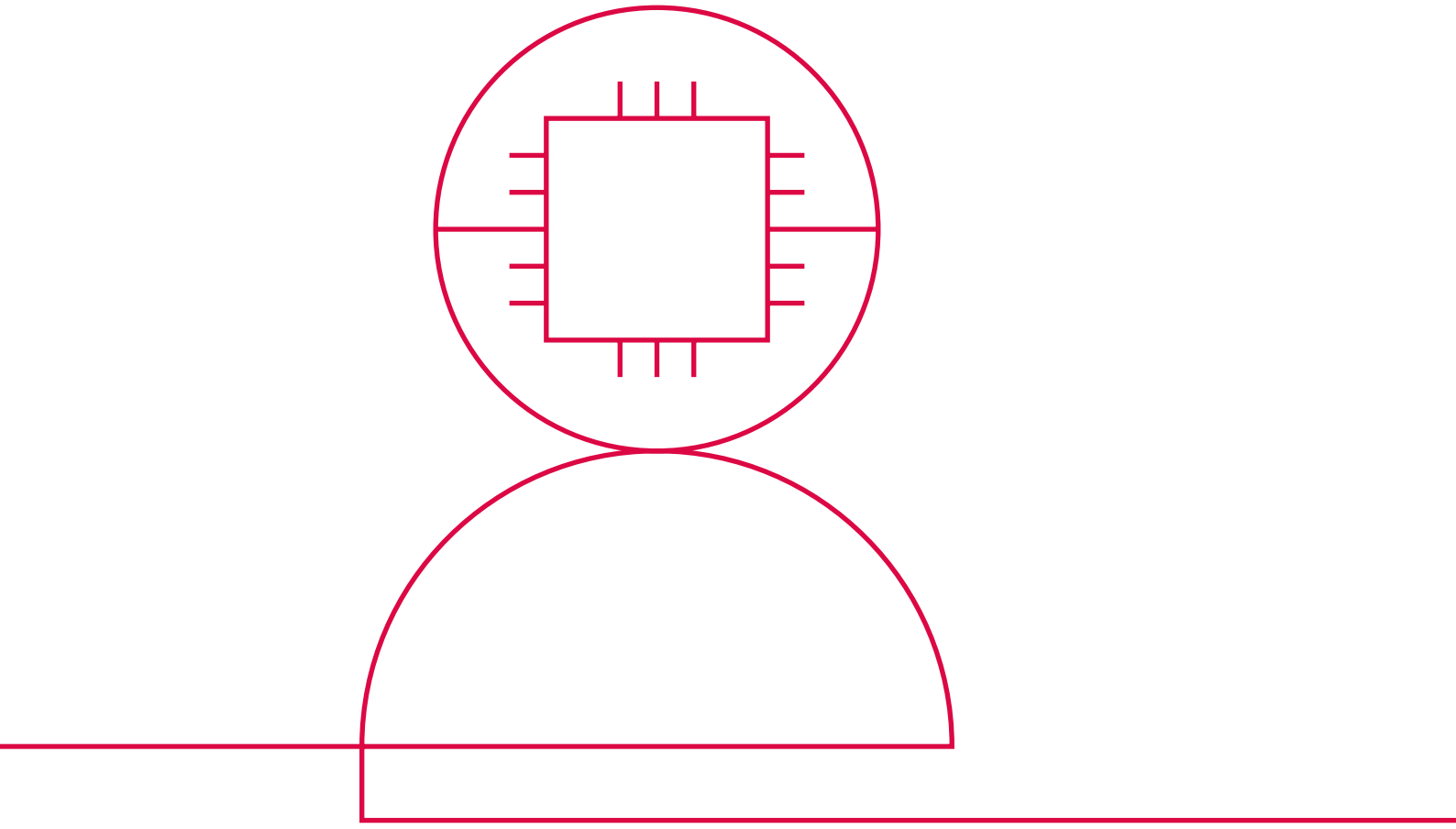
Academy of Medical Royal Colleges

January / 2019

Artificial Intelligence in Healthcare



Contents

Foreword

By any measure, Artificial Intelligence – the use of intelligent machines to work and react like humans – is already part of our daily lives. Facial recognition at passport control and voice recognition on virtual assistants such as Alexa and Siri are already with us. Driverless cars or 'companion' robots that 'care' for the elderly are undergoing trials and most commentators say will be commonplace soon.

As with automation after the industrial revolution, it is hard to think of any area of our lives that will not be affected by this nascent data driven technology. Artificial Intelligence is already with us in healthcare too. Google's DeepMind has taught machines to read retinal scans with at least as much accuracy as an experienced junior doctor. Babylon, the health app start-up, claims its chatbot has the capacity to pass GP exams although this is contested by the Royal College of General Practitioners.

And just as some say AI is going to provide instant relief to many of the pressures healthcare systems across the world are facing, others claim AI is little more than snake oil and can never replace human delivered care. It already has a role, but how far can that extend? It is difficult to imagine how the judgement around patient behaviours, reactions and responses and the subtleties of physical examination, particularly observation and palpation) can be anything other than human.

It will be for our politicians and ultimately the public to decide how far and in what ways AI impacts patient care across the UK.

This report is not meant to be an exhaustive analysis of all the potential AI holds or what all the implications for clinical care will be. It is instead a snapshot of 12 domains that will be most impacted by AI and looks at each from a clinical, ethical and practical perspective. The authors have, of necessity, limited the time horizon to the next few years. For this reason, we have left discussions about the impact of AI in surgery for the future. The report does however, consider how AI might affect the diagnostic disciplines, because that is already with us in some form.

Equally, it does not pretend to answer the myriad questions which will surely follow as this technology develops. More, this report is designed as a starting point for clinicians, ethicists, policy makers and politicians among others to consider in more depth.

Scientific progress is about many small steps and occasional big leaps. Medicine is no exception. Artificial Intelligence and its application in healthcare could be another great leap, like population-wide vaccination or IVF, but as this report sets out, it must be handled with care.

For me, the key theme that leaps from almost every page of this report is the tension between the tech mantra, 'move fast and break things' and principle enshrined in the Hippocratic Oath, 'First, do no harm.' This apparent dichotomy is one that must be addressed if we are all to truly benefit from AI. What, in other words, must we do to allow the science to flourish while at the same time keeping patients safe? Doctors can and must be central to that debate – the basis of which is set out here.

Professor Carrie MacEwen, Chair, AoMRC

About this report

The Academy of Medical Royal Colleges (the Academy) is grateful to NHS Digital for commissioning this work and to the many well-informed thinkers and practitioners from the worlds of AI, medicine, science, commerce and bio-ethics who so willingly gave up their time and knowledge to contribute to this work. They are listed at the end of this section and without them, this report would not have been possible.

The contents represent a series of one-to-one interviews conducted over the spring and summer of 2018 and two focus groups held in July 2018. Most quotes are attributed where practical while some other views have been aggregated to provide a more general view. Dr Farzana Rahman also interviewed many US commentators, academics and thinkers as she was based there at the time of writing. It is worth noting that there was overwhelming consensus among the participants on both sides of the Atlantic when discussing the domains the authors identified as areas for discussion.

These are:

- Patient safety
- The doctor and patient relationship
- Public acceptance and trust
- Accountability for decisions
- Bias, inequality and unfairness
- Data quality, consent and information governance
- Training and education
- Medical research
- The regulatory environment
- Intellectual property and the financial impact on the healthcare system
- Impact on doctors' working lives
- Impact on the wider healthcare system.

Each of the above was then considered from a clinical, ethical and practical perspective by the authors and contributors.

The scope of discussion of the possible implications of AI in future healthcare is almost limitless. This report focuses on the likely clinical impact of AI for doctors and patients in the near future, by which we mean certainly within the next five years, though more likely by the end of the decade. It does not consider in detail the potential effects of AI in non-clinical elements of healthcare: logistics, stock supply, patient flow and bed management, although in compiling this report it is clear there will be many. Neither does it address the specific impact on nurses, pharmacists and allied healthcare professionals, each of which would warrant their own report.

Many of the applications envisaged in the short term involve tools to support healthcare professionals, whereas looking further into the future, AI systems may exhibit increasing autonomy and independence. This report focuses more on AI as decision support tools rather than the decision making tools which, by common consensus, are much further away.

Dr Jack Ross, Dr Catherine Webb, Dr Farzana Rahman, AoMRC

AI will allow doctors to be more human

Dr Simon Eccles,
Chief Clinical Information Officer for Health and Care, NHS England, Department of Health and Social Care, NHS Improvement

Executive summary

Artificial Intelligence has already arrived in healthcare. Few doubt though, that we are only at the beginning of seeing how it will impact patient care. Not unsurprisingly, the pace of development in the commercial sector has outstripped progress by traditional healthcare providers – in large part because of the great financial rewards to be had.

Few doubt too that while AI in healthcare promises great benefits to patients, it equally presents risks to patient safety, health equity and data security.

The only reasonable way to ensure that the benefits are maximised and the risks are minimised is if doctors and those from across the wider health and care landscape take an active role in the development of this technology today. It is not too late.

That is not to say doctors should give up medicine and take up computational science, far from it – their medical and clinical knowledge are vital for their involvement in what is being developed, what standards need to be created and met and what limitations on AI should be imposed, if any.

And while the Academy welcomes the use of Artificial Intelligence in healthcare and the significant opportunities and benefits it offers patients and clinicians, there are substantial implications for the way health and care systems across the UK operate and are organised. It is the Academy's view that while the UK's health and care systems were somewhat late to recognise the potential AI has when it comes to improving healthcare, the NHS in general and NHS Digital in particular are catching up fast. Both are taking a commendably 'real-world' approach in an environment which is traditionally slow to change.

The recent publication of the NHS Long Term Plan set out some admirable ambitions for the use of digital technology and while the Academy applauds these aspirations the day to day experience of many doctors in both primary and secondary care is often a world away from the picture painted in the plan. With many hospitals using multiple computer systems, which often don't communicate, the very idea of an AI enabled healthcare system seems far-fetched at best.

For AI to truly flourish, not only must IT be overhauled and made inter-operable, but the quality and extent of health data must be radically improved too. The workforce will need to be trained on its value and the need for accuracy and healthcare organisations will need to have robust plans in place to provide backup services if technology systems fail or are breached.

In view of this the Academy has identified seven key recommendations which politicians, policy makers and service providers would do well to follow.

Recommendations:

1. Politicians and policymakers should avoid thinking that AI is going to solve all the problems the health and care systems across the UK are facing. Artificial intelligence in everyday life is still in its infancy. In health and care it has hardly started – despite the claims of some high-profile players

2. As with traditional clinical activity, patient safety must remain paramount and AI must be developed in a regulated way in partnership between clinicians and computer scientists. However, regulation cannot be allowed to stifle innovation

3. Clinicians can and must be part of the change that will accompany the development and use of AI. This will require changes in behaviour and attitude including rethinking many aspects of doctors' education and careers. More doctors will be needed who are as well versed in data science as they are in medicine

4. For those who meet information handling and governance standards, data should be made more easily available across the private and public sectors. It should be certified for accuracy and quality. It is for Government to decide how widely that data is shared with non-domestic users

5. Joined up regulation is key to make sure that AI is introduced safely, as currently there is too much uncertainty about accountability, responsibility and the wider legal implications of the use of this technology

6. External critical appraisal and transparency of tech companies is necessary for clinicians to be confident that the tools they are providing are safe to use. In many respects, AI developers in healthcare are no different from pharmaceutical companies who have a similar arms-length relationship with care providers. This is a useful parallel and could serve as a template. As with the pharmaceutical industry, licensing and post-market surveillance are critical and methods should be developed to remove unsafe systems

7. Artificial intelligence should be used to reduce, not increase, health inequality – geographically, economically and socially.

It is said that artificial intelligence will deliver major improvements in quality and safety of patient care at reduced costs, with some observers even suggesting it represents an imminent revolution in clinical practice. Yet we are very early in the evidence cycle and it is unclear how true such predictions will prove to be.

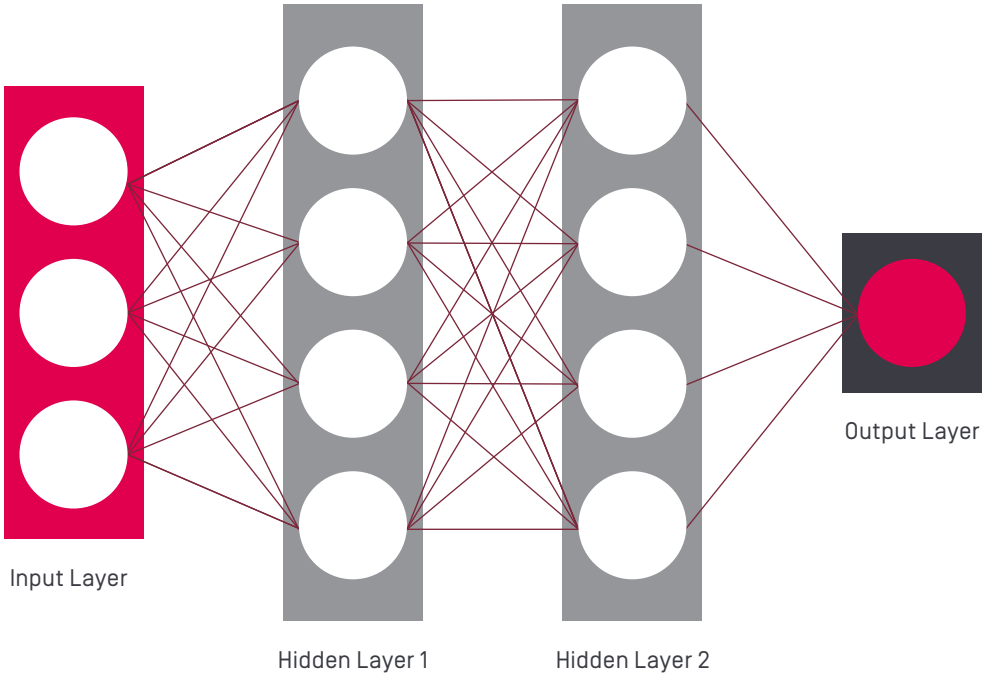
Clinicians, researchers, policy specialists and funding organisations are aware that something important may be emerging, but they have few tools for appraising the potential of AI to improve services.

Prof John Fox, Chairman,
OpenClinical CIC, Chief Scientific Officer, Deontics Ltd

What is AI? A primer for clinicians

Artificial intelligence describes a range of techniques that allow computers to perform tasks typically thought to require human reasoning and problem-solving skills. 'Good Old-Fashioned AI', which follows rules and logic specified by humans, has been used to develop healthcare software since the 1970s, though its impact has been limited. More recently there have been huge technological developments in the field of machine learning and especially with artificial neural networks, where computers learn from examples rather than explicit programming.

Figure 1: A deep neural network with hidden layers

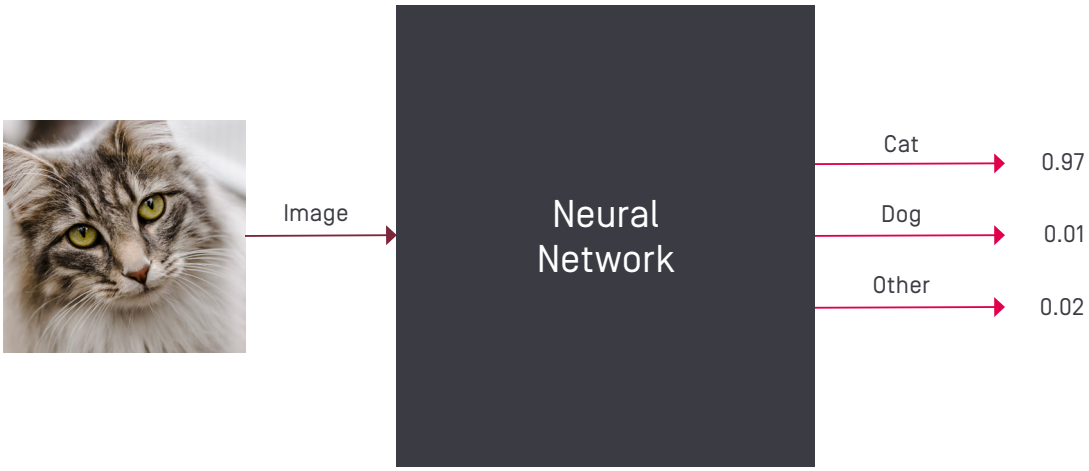


Neural networks function by having many interconnected 'neurons'. The connections between these neurons get stronger if they help the machine to arrive at the correct answer and weaken if they do not help to reach the correct answer. The system itself is made up of an input layer, some hidden layers and an output layer. There are a huge number of connections between each layer that can be refined. Over time, these billions of refinements can hone an algorithm that is very successful at the task.

For the purposes of this report we will use a broad definition of artificial intelligence, including machine learning, natural language processing, computer vision and chatbots. We will focus on 'narrow' AI which is designed for a specific application, rather than the more science fiction hopes of a generalised AI which can accomplish all and any tasks a human can.

'Artificial Neural Networks' are a common type of machine learning inspired by the way an animal brain works. They progressively improve their ability at a particular task by considering examples. Early image recognition software was taught to identify images that contain a face by analysing example images that have been manually labelled as 'face' or 'no face'. Over time, with a large enough data set and powerful enough computer, they will get better and better at this task. They are able to independently find connections in data.

Figure 2: Can a machine distinguish a cat from a dog?



Learn OpenCV (2017) Neural Networks: A 30,000 Feet View for Beginners.

There are three key limitations of these methods:

Explainability

Modern machine learning algorithms are often described as a 'black box'. Decisions are based on the huge number of connections between 'neurons' and so it is difficult for a human to understand how the conclusion was reached. This makes it difficult to assess reliability, bias or detect malicious attacks.

Data requirement

Neural networks need to be trained on a huge amount of accurate and reliable data. Inaccurate or misrepresentative data could lead to poorly performing systems. Health data is often heterogeneous, complex and poorly coded.

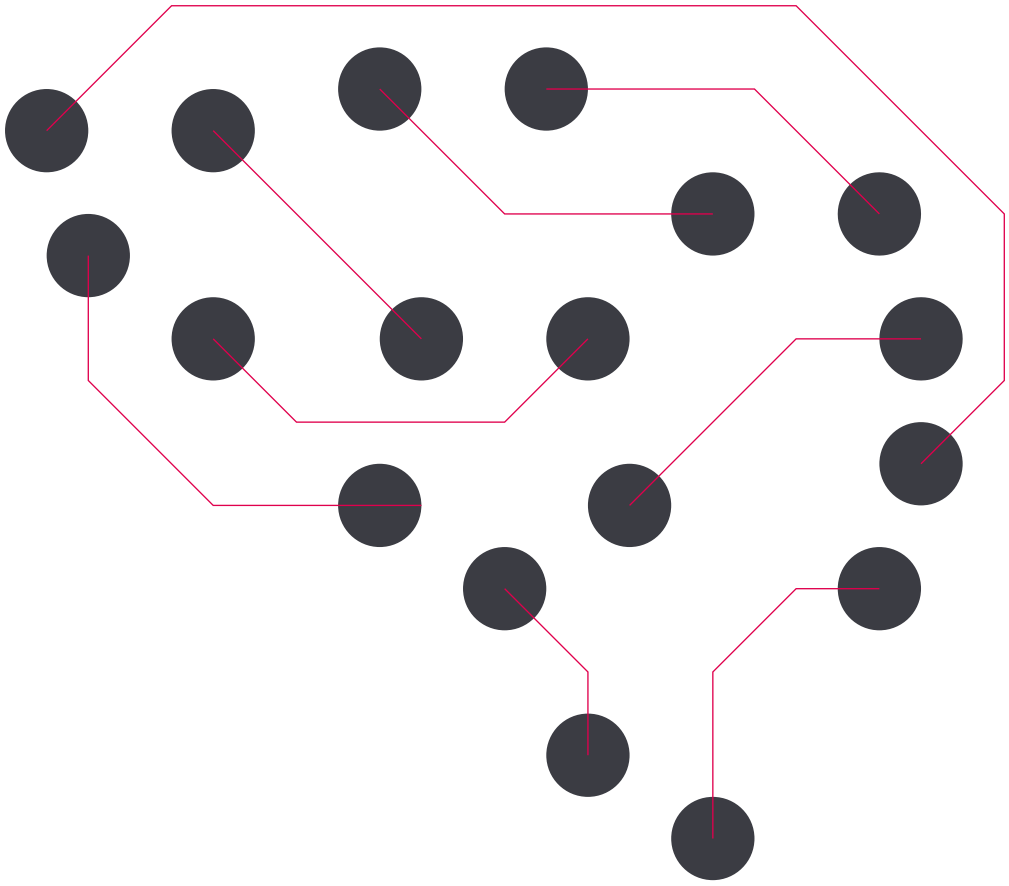
Transferability

Algorithms may be well optimised for the specific task they have been trained on but may be confidently incorrect on data it has not seen before.

The pitfalls of machine learning in healthcare:

- Training and testing on data that is not clinically meaningful
- Lack of independent blinded evaluation on real-world data
- Narrow applications that cannot generalise to clinical use
- Inconsistent means of measuring performance of algorithms
- Commercial developers' hype may be based on unpublished, untested and unverifiable results.

Domains



Patient safety

Central to the debate about the introduction of AI to healthcare is perhaps the most fundamental question: will patients be safe or safer? Proponents argue machines don't get tired, don't allow emotion to influence their judgement, make decisions faster and can be programmed to learn more readily than humans. Opponents say human judgement is a fundamental component of clinical activity and the ability to take a holistic approach to patient care is the essence of what it means to be a doctor.

Digitised clinical support tools offer a way to cut unwarranted variation in patient care. Algorithms could standardise tests, prescriptions and even procedures across the healthcare system, being kept up-to-date with the latest guidelines in the same way a phone's operating system updates itself from time to time. Advice on specialist areas of medicine normally only available through referral to secondary or tertiary services could be delivered locally and in real-time. Direct-to-patient services could provide digital consultations regardless of time of day, geography, or verbal communication needs including language.

However, algorithms could also provide unsafe advice. The tech mantra of 'move fast and break things' does not fit well when applied to patient care. As we shall see across the domains, evaluating whether an AI is safe will be challenging. It may be poorly programmed, poorly trained, used in inappropriate situations, have incomplete data and could be misled or hacked. And worse, dangerous AI could replicate harm at scale.

Clinical considerations:

— Algorithms could standardise assessment and treatment according to up-to-date guidelines, raising minimum standards and reducing unwarranted variation

— Artificial intelligence could improve access to healthcare, providing advice locally and in real-time to patients or clinicians and identifying red flags for medical emergencies like sepsis

— Decision support tools could be confidently wrong and misleading algorithms hard to identify. Unsafe AI could harm patients across the healthcare system.

Ethical issues:

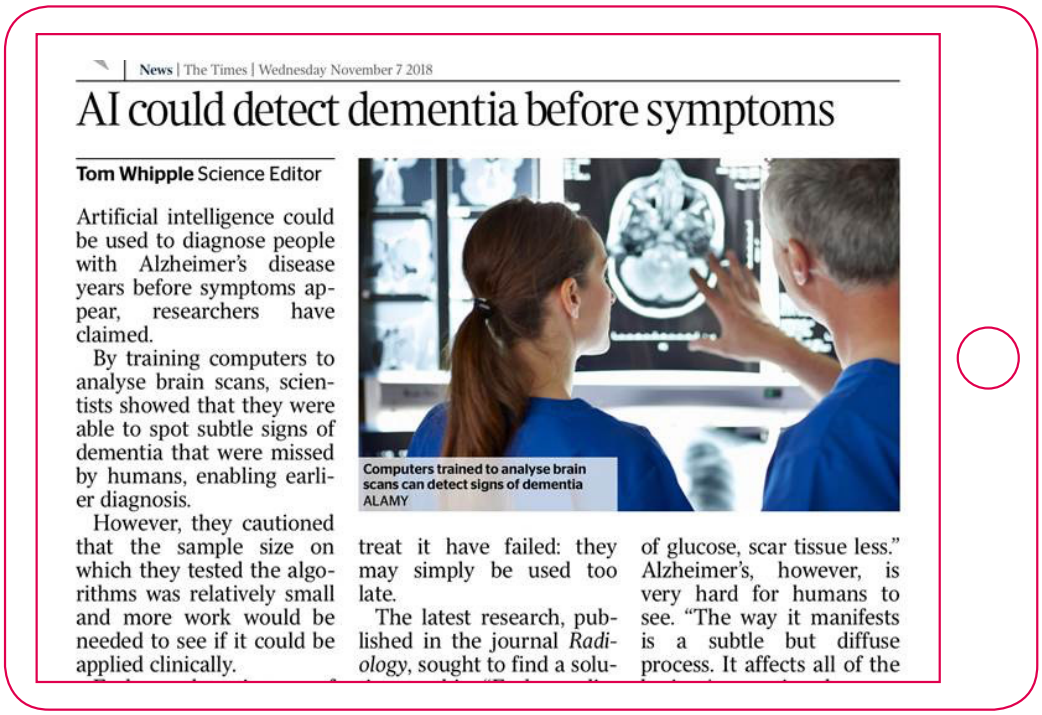
— The widespread introduction of new AI healthcare technology will help some patients but expose others to unforeseen risks. What is the threshold for safety on this scale – how many people must be helped for one that might be harmed? How does this compare to the standards to which a human clinician is held?

— Who will be responsible for harm caused by AI mistakes – the computer programmer, the tech company, the regulator or the clinician?

— Should a doctor have an automatic right to over-rule a machine's diagnosis or decision? Should the reverse apply equally?

Practical challenges:

- Human subtleties may be hard to digitise and machines may struggle to negotiate a pragmatic compromise between medical advice and patient wishes
- Few clinicians will be able to understand the 'black box' that neural networks use to make decisions and the code may be hidden as intellectual property. Should we expect them to trust its decision?
- A focus on measurable targets could lead to AI 'gaming' the system, optimising markers of health rather than helping the patient
- As clinicians become increasingly dependent on computer algorithms, these technologies become attractive targets for malicious attacks. How can we prevent them from being hacked?
- The importance of human factors and ergonomics risk being overlooked. Public, patients and practitioners should be engaged at the design phase and not left simply as end-users.



News | The Times | Wednesday November 7 2018

AI could detect dementia before symptoms

Tom Whipple Science Editor

Artificial intelligence could be used to diagnose people with Alzheimer's disease years before symptoms appear, researchers have claimed.

By training computers to analyse brain scans, scientists showed that they were able to spot subtle signs of dementia that were missed by humans, enabling earlier diagnosis.

However, they cautioned that the sample size on which they tested the algorithms was relatively small and more work would be needed to see if it could be applied clinically.

Computers trained to analyse brain scans can detect signs of dementia
ALAMY

treat it have failed: they may simply be used too late.

The latest research, published in the journal *Radiology*, sought to find a solu-

of glucose, scar tissue less." Alzheimer's, however, is very hard for humans to see. "The way it manifests is a subtle but diffuse process. It affects all of the

I do think the issues around patient safety are subject to some confusion. The fact is machines are better at numbers than humans, but you will always need a human, so I don't think it's an either/or choice. And in terms of safety, machines are much better at recognising things like rare diseases, simply because they are working from a bigger dataset, so you could argue that some patients will be significantly safer. That said, regulation is not really keeping up. We [in the commercial sector] need to work better with the regulators so they understand the technology and we can ensure patient safety is paramount.

At the moment AI isn't really viewed by regulators as anything much more than a novelty – a sort of glorified decision support tool. That's fine for now, but we are not far off going beyond that in some areas, so organisations like the Care Quality Commission will have to catch up.

Dr Hannah Allen, Associate Medical Director, Babylon Health.

One of the things we try to do at CQC is to encourage providers to innovate, because we know this can lead to big improvements in the quality of care. But innovation needs to be done in the right way so that we see the greatest possible benefit, while we also make sure that people using services are kept safe. We know that AI is one of the fastest moving areas of innovation at the moment so it can be a big part of this story in the future, but it does present some new challenges and we're working with other regulators, government, providers and the public to make sure that we can respond in the right way.

Malte Gerhold, Executive Director of Strategy and Intelligence,
Care Quality Commission

The doctor and patient relationship

The nature of the relationship between clinicians and their patients has evolved as medicine as evolved. For centuries, the doctor held exclusive knowledge and issued 'orders'. Today, doctors are expected to take a holistic approach, providing care that is tailored to each patient's wishes and based on shared decision-making. The future use of AI technologies has the potential to cause a further seismic shift in the culture of interactions between clinicians and patients.

Much of this depends on the nature of the interface between the public and AI. Applications could range from a doctor-facing decision support tool, potentially unnoticed by the patient, to an autonomous AI system accessible from the patient's own devices, diagnosing and treating conditions without human clinical involvement.

As AI systems become more autonomous with a greater degree of direct-to-patient advice, a significant need arises to establish the role of clinicians in maintaining quality, safety, patient education and holistic support.

The psychological impact on both patients and doctors of the presence of AI must be anticipated, including an inherent reluctance to disagree with the recommendations of digital systems.

Clinical considerations:

- The holistic side of a consultation would be difficult to replicate with digital tools – doctors are better equipped to detect non-verbal signs, tone of voice and other subtle cues. Loss of this human contact could lead to reduced awareness of patients' loneliness, safeguarding, or social needs
- Will the doctor become a second opinion, a step in the quality assurance process, or an interpreter? In what contexts should clinical staff review AI-generated advice for quality-assurance and interpretation before it is accessible to a patient?
- There is a risk that lay people unfamiliar with medical data may under – or overestimate the severity of conditions and misunderstand the magnitude of risks.

AI will change the doctor-patient relationship. The doctor will need to behave differently – to learn how to interact with expert patients, who may have self-diagnosed with AI tools.

AI supporting clinicians is not that far away, but AI replacing clinicians is a long way off. AI is not ready to fully interpret a patient's nuanced response to a question, nor is it ready to replace examining patients – but it is very good at making differential diagnoses from results.

Dr Phil Koczan, GP. CCIO for digital integration and NHS England (London) and Clinical Advisor to the Professional Record Standards Body.

Ethical issues:

— Can a doctor be expected to act on the decisions made by a 'black box' AI algorithm? In deep neural networks, the reasons and processes underlying the decisions made by AI may be difficult to establish, even by skilled developers. Do doctors need to explain that to patients?

— Will clinicians bear the psychological stress if an AI decision causes patient harm? They could feel great responsibility for their role in the process without the power to modify or understand the contribution of the AI to the error

— Could the ready availability of a tool superficially appearing to 'replace' a doctor's advice diminish the value of clinicians in the eyes of the public and therefore reduce trust and degrade the quality of the doctor-patient relationship?

Practical challenges:

— If AI and doctor disagree, who will be perceived as 'right'? The degree of relative trust held in technology and in healthcare professionals may differ between individuals and generations

— Autonomous health advice and the interface with wearable devices may promote patients' health ownership and supported self-care but could result in increased health anxiety or health fatigue for some members of the public

— Reduced face-to-face contact could reduce opportunities for clinicians to offer health promotion interventions – this must be factored into systems.

Chatbots

Alder Hey Children's Hospital uses a chatbot, 'Olly', for children to discuss questions they have before surgery. Similar technologies are becoming available for a range of medical uses, such as online Cognitive Behavioural Therapy.

Some patients find chatbots to be an approachable and convenient way of gaining information and advice, including in areas where patients feel uncomfortable discussing issues that may embarrass them. Others may lament the loss of human contact.

Why are we not dictating to the IT providers what we want and need? Why are we beholden to them, not the other way around? We need technology that works for patients and makes our lives as healthcare professionals easier. So, the NHS needs to take control.

Professor Helen Stokes-Lampard, Chair,
Royal College of General Practitioners

Public acceptance and trust

By any measure, the concept of AI – how it works and what it can and cannot do, is complex. But, in the same way that few people need to know how a flight booking app works, so it is safe to assume that patients will not need to know the details of how AI works. They simply need to know that it does work and can be trusted to work reliably for them.

Gaining that trust will be one of the most essential steps to the development of AI in healthcare. For this reason, developers should continue to focus on the utility of AI to the individual rather than seek explicit approval from the outset. Health apps, chatbots that focus on young people and their mental health or home monitoring systems that learn our routines are good examples of this in that they are already proving their worth and their use can be easily monitored.

As AI embeds itself in our everyday lives through avenues other than health, acceptance and trust in the concept that a machine is making decisions that are in our best interests will increase. That said, the 'social licence' that AI enjoys so far is a precious commodity. Historic controversy over genetically modified food perhaps demonstrates the consequences when the trust between science and the wider public breaks down. It should also serve as a warning to AI developers that they should not take public acceptance and trust for granted.

Clinical considerations:

— There are no nationally agreed standards for quality. Should there be? And if so who should set them?

— Does the introduction of standards inevitably stifle opportunities for innovation?

— How can a patient or a clinician differentiate 'good' AI from 'bad' AI? A mental health app with a great user interface, may, for example, be based on very poor data.

Ethical issues:

— Should 'self-help' AI always be free for users or paid for? Does this risk creating a two-tier system when it comes to the quality of the AI itself?

— Should such apps or online resources be subject to marketing restrictions?

— How transparent should AI providers be about how the data is used?

Practical challenges:

— If there is greater acceptance of or reliance on AI among younger users, would that ultimately create a two-tier health system, with older patients more reliant on doctor delivered care because they don't trust the machines?

— If AI is 'over-sold' by developers and politicians and fails to deliver the promised benefits, there is a real risk that the public could reject the use of AI in healthcare altogether?

— Should patients be always given a choice about whether a doctor or an algorithm makes their diagnosis?

Trusting AI

Around 15% of Facebook users told US pollsters, Axios.com they would reduce their use of the social networking platform following the Cambridge Analytica scandal of 2018. Arguably, maintaining users' trust in a highly personal area such as their health is an even greater challenge.

It would take just a few news stories, fake or otherwise, about people being refused a mortgage because they had been using a mental health support app or see their insurance premiums rise because they had self-diagnosed a serious disease and the public's trust would evaporate overnight.

AI to support patient care is being developed in a variety of ways and has huge potential to support doctors and enable them to spend more time with patients. However, we musn't get carried away or think that the AI applications developed so far can replace a fully trained and qualified doctor. We need much more robust trials and evidence to work out how it can best be used.

So let's embrace it, evaluate it using the same rigorous standards we apply to any new medical innovation and educate ourselves on the opportunities AI offers to support great patient care.

Professor Andrew Goddard, President,
Royal College of Physicians of London

Accountability for decisions

Who should be held responsible when something goes wrong? It is a fundamental question at the heart of the conversation between clinicians, healthcare organisations, policy makers and AI developers. To what extent do we expect healthcare providers to understand the intricacies of AI technology and technology firms to understand the realities of clinical practice?

AI is rapidly developing and complex and there will be errors and unforeseen consequences. Technology companies are currently focusing on AI that will support clinicians, rather than replace clinical judgement – implying that accountability for mistakes remains with the clinician. But a line needs to be drawn between accountability for content and for operation. A clinician might be accountable for not using an algorithm or device correctly, but in the event of harm being caused by incorrect content rather than improper use, then the accountability must lie with those who designed and then quality assured it.

However, this line may not be so easy to define. Clinicians may find themselves incorrectly justifying decisions made by AI because of the well-documented concept known as automation bias. Here, humans can have a tendency to 'trust' a machine more than they might trust themselves. If the clinician is, in effect, 'rubber stamping' anything recommended by an algorithm, who is responsible if an error is made?

The risk of incomplete data

An AI algorithm designed to predict which patients with pneumonia could be safely discharged and treated as outpatients learnt, incorrectly, that patients who have a history of asthma have a lower risk of dying from pneumonia. This was because it was true from the training data, as patients with asthma usually went to ICU, received more aggressive care and so were less likely to die. The algorithm did not understand this and used the rule that if someone had asthma they were more likely to be treated as an outpatient. Given this, the research team decided to use a rule-based system that was more intelligible to humans, rather than neural networks, so they could identify and remove dangerous rules.

Machine learning algorithms can be hidden in the much vaunted 'black box', where the reasons behind the decision might not be explainable in a way that humans can understand. Combine this with the idea that the software itself may be unavailable to review for intellectual property reasons, the training data for privacy reasons and true accountability becomes even more impractical. Crucially, the patient and the clinician may be recommended a course of action or treatment without any real opportunity to check or challenge the approach taken by the machine.

Clinical considerations:

— The line between accountability for harms caused by faulty content and incorrect operation needs to be recognised

— The need to protect against automation bias and 'rubber stamping' of AI-generated recommendations must be considered

— Clinicians will need new skills to appraise new technology and enable them to agree or disagree confidently with AI-generated recommendations.

Ethical issues:

— Transparency of decisions may be key to empowering patients and gaining trust – but would an insistence on removing the 'black box' jeopardise the opportunity to realise the full potential of machine learning?

— The introduction of AI-generated recommendations alongside clinical judgement may change patients' views on who or indeed what to trust. Might this result in new ideas about what constitutes clinical negligence?

— Could a two-tier diagnostic service emerge, with only the wealthiest gaining access to human-led interpretation of test results or imaging? Or conversely, only the wealthiest gaining access to a perhaps superior machine-led interpretation of test results or imaging?

Practical challenges:

— Will technology companies be willing to take responsibility for the results of AI systems and will the NHS? Will there be a significantly increased workload for Accountable Officers and Chief Information Officers?

— Does the public sufficiently understand the concept of accountability? Would the public understand the (probably nuanced) question of machine accountability?

— Inadequate input would lead to inappropriate results – does the quality of data need to be standardised, or indeed kite-marked?

Bias, inequality and unfairness

Will AI provide more fair and objective decisions than humans, who are limited by our own personal experience and biases? Or will they collect and even amplify human prejudices, embedding discrimination within healthcare systems? If the training data isn't representative, or the goals inappropriately chosen, then the resulting AI tool could be deeply inequitable.

Machine learning algorithms being used outside of healthcare have been criticised for discriminating based on race, gender, age, poslcode and religion, while chat bots have been tricked into propagating hate speech. Artificial intelligence can 'learn' the wrong values and even become self-fulfilling – for example, an algorithm for helping with job hiring decisions might simply reward people who have the same background as those in the historical recruitment data, reinforcing its bias with every decision.

The 'black box' nature of neural networks makes it particularly hard to truly assess whether an AI is biased. Worse still, machine learning is very good at identifying proxies for characteristics, such as predicting race and socioeconomic group from names and postcodes. Tech companies such as IBM, Google, Microsoft and Facebook are all creating tools to help identify bias in algorithms.

Clinical considerations:

— Clinicians will need to be confident that decision support tools are valid for the patient in front of them, not just the specific group that made up the training data

— Algorithms can lead to wrong assumptions based on incomplete data, for example suggesting having asthma lowers a patient's risk of death from pneumonia (see 'the risk of incomplete data' on page 18)

— Doctors learn from errors through reflection and changing future practice. How can we stop algorithms from reinforcing their own behaviour when they make mistakes?

Ethical issues:

— Is it acceptable to stratify patients by factors such as age, race, postcode or socioeconomic group if this can improve outcomes, or would this negatively impact those patients? This is a big question for society and ethicists

— Do we have an ethical duty to encourage under-represented groups to provide more of their data to be used to train algorithms?

— Artificial intelligence has the potential to use the wide range of differences between us to provide truly individualised care – though this might be better for some people than others.

Practical challenges:

— If training data is only obtained by those who specifically volunteer and consent for their data to be used, algorithms will learn from unrepresentative datasets

— Algorithms could be 'loaded' with hidden preferences, such as favouring a particular drug manufacturer over another

— Artificial intelligence will need high quality labelled data from electronic health records. Is it clinicians' responsibility to make sure all data is recorded in a standardised machine-readable way?

'Precision medicine can be biased as most of the data is based on that of people with European ancestry. If the data is not representative for minority populations then it could be potentially harmful.'

Berk Ustun, Postdoctoral Fellow,
Center for Research in Computation and Society,
Harvard University

Racial bias in criminal justice algorithms

The Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) is an algorithmic risk score used to help judges in certain US states decide sentencing, by predicting a defendant's risk of reoffending. However, analysis of over 7,000 arrestees by the investigate journalists ProPublica argued there was a systematic bias against black defendants, who were inaccurately classified as future criminals at almost twice the rate of white defendants.

'There should be a notion amongst patients, society and the general population that there is a societal good in sharing their data to make sure that health related algorithms are as fair and beneficial as possible.'

Finale Doshi-Velez, Assistant Professor in Computer Science at the Harvard Paulson School of Engineering and Applied Science

Built in bias?

A paper in the Journal of the American Medical Association warned against the potential racial disparities that could come from relying on machine learning for skin cancer screenings. Algorithms using neural networks have been developed for the detection of melanoma, using publicly available images of melanoma, which is more prevalent in white skin. The technology is therefore more effective at detecting melanoma in white skin than black. However, even though melanoma is rarer in individuals with black skin, those with black skin have higher rates of mortality. This is not only due to the type of melanoma, but also because of poor detection rates and identification by physicians. Will this type of technology therefore promote bias and unfairness?

Data quality, consent and information governance

The UK Government and its health and social care systems have a legal duty to maintain the privacy and confidentiality of its citizens. Europe's 2018 General Data Protection Regulation (GDPR) offers additional privacy safeguards. However, the development of AI and machine learning algorithms relies on the use of large datasets. The accuracy and evolution of these algorithms depends on the availability of high volumes of good-quality data. Balancing these two areas raises a number of considerations.

Clinical considerations:

— The use of machine learning to guide decisions introduces a potential third-party into the doctor patient relationship. If data from this interaction is used for further algorithm development, assumptions of confidentiality and trust can rightly be challenged

— Should all clinicians be trained to critically assess data quality, computational robustness and information governance?

— Good quality AI depends on good quality data. With a few notable exceptions, the quality of patient level data is notoriously patchy in the NHS. Does the system have the resources, skills and appetite to improve it?

Ethical concerns:

— Whose data is it really? Does it belong to the patient (the source), the system (the collector and aggregator), or the developer (who adds value to the raw materials)?

— Patients do not, in the main, know that data about them and their disease is collected and used. When told, very few opt out. Does the NHS have a moral duty to tell everyone who uses the system?

— Does the value to wider society of data about a person's health trump an individual's right to withdraw consent for its use?

Practical issues:

— Many of these issues regarding privacy and information governance require careful and detailed explanation to patients. Is this practical or affordable?

— The General Data Protection Regulations (GDPR) say companies should be able to alter or delete personal data if requested. But what if the request is made after the data has been incorporated into an algorithm?

— The GDPR also say that companies should minimise the amount of data they collect and keep. This could stifle innovation and development, which would in turn negatively impact patients.

Issues around data quality and information governance lie at the heart of the debate around the development of AI in healthcare. Outside of data on cancer, rare diseases and congenital anomalies, patient data is generally of poor quality. Although this is improving, AI is generally being developed at pace outside of the system. While it would not be legally possible, let alone ethically acceptable to give developers access to patient identifiable data without abiding by the strict information governance protocols in place, there is an inherent paradox that the organisations that are most in need of rich, reliable and robust datasets to improve healthcare are unlikely to ever have access to them at any meaningful scale. This has led some to argue that AI in healthcare should be overseen by government.

The Government's code of conduct for data driven technologies

Innovators in the field of big data and artificial intelligence may come from sectors which are not always familiar with medical ethics and research regulation and who may utilise data sets and processing methods that can sit outside existing NHS safeguards. In order for people to know their data is being used for public good, fairly and equitably and their privacy and rights are safeguarded, we developed a set of principles that should be followed by anyone developing, testing or evaluating such technologies including commercial companies, NHS Trusts, Universities and Charities.

The 10 Principles in the Code of Conduct for data driven technologies enable the development and adoption of safe, ethical and effective data-driven* health and care technologies. The Code should be seen as complementary to the health research, medical device regulations and the CE mark process as well as other regulatory approvals. When used as part of an overarching strategy designed to create a trusted environment for data-driven technologies that is:

— The safest in the world

— Appropriately responsive to progress in innovation

— Ethical, legal, transparent and accountable

— Evidence-based

— Competitive and collaborative

— In alignment with the NHS Constitution

Dr Indra Joshi, NHS England,
Jess Morley, Technology Advisor, Dept. of Health and Social Care

*'Data-driven' means the outputs of the technology are based on data analysis and interpretation. In healthcare this currently includes technologies such as health apps, wearables, or software that automate interpretation to a greater (deep learning) or lesser extent (simple descriptive statistics)

We've got a real opportunity with AI based tech to gain time and efficiencies, but it has to be implemented in a safe and trusted way. We need to bring everyone with us on this journey of transformation.

Dr Indra Joshi, Digital Health and AI Clinical Lead, NHS England

Training and education

The adoption of AI in clinical practice will inevitably impact the training and education of clinicians, both through enhanced technological opportunities and through a shift in fundamental learning needs as professional working practices change.

Artificial intelligence could underpin sophisticated digital tools to support learning and development:

— AI could be incorporated into high-fidelity simulations generating clinical scenarios across a range of specialities to enhance training and revalidation

— With the pace of advancement of medical knowledge, the sheer volume of new information exceeds that with which an individual can keep pace in real time. Artificial intelligence has the potential to analyse large datasets across multiple sites to condense information for the clinician for practical use

— Combined with other digital technologies, AI could be used to personalise training by evaluating previous experiences, responses and outcomes to model the strengths and weaknesses of individual clinicians. Personalised medicine need not be for patients alone.

It is often suggested that AI will play a pivotal role in automating simple clinical tasks to free clinician time for more complex activities. Although attractive in terms of workforce utilisation and cost, there is the potential that losing skill in more basic tasks could undermine those needed for more complex work. It should be noted too that a review of the work needed to prepare the healthcare workforce for a digital future, by Dr Eric Topol, on behalf of Health Education England is to be published soon.

Clinical considerations:

— The automation of routine clinical tasks may skew the doctor's view of normality and impede their cognitive pattern recognition. For example, it is essential that clinicians understand the anatomical variants in a normal chest X-ray so that when faced with pathology they are able to confidently identify it as significant

— If certain human clinical skills are 'replaced' by AI, what happens if the technology fails?

— How would a machine's mistakes be detected?

Ethical issues:

— Doctors' time will be required to ensure data quality to train AI systems. Might this mean clinical experience during training will be reduced and could this have an adverse effect on patient care?

— Should public money be diverted from training healthcare professionals to training AI systems?

Practical considerations:

— How much do clinicians need to understand a tool in order to use it safely? Is it enough to simply be aware of its uses and limitations, or is more in-depth study required?

— Should there be a cross-cutting specialty of 'digital doctors', specifically trained within computer and data science and medicine? Or is collaboration with technology experts sufficient?

— Is there a risk that it might make the profession a less attractive career option?

The medical profession has long interacted with pharmaceutical companies. Medical students are educated to interpret and critique the output of clinical trials and strict marketing regulations are in place. As doctors seek the evidence behind pharmaceuticals, perhaps they should similarly be trained to appraise new healthcare technologies for safety and efficacy and understand their technical limitations and risks.

There are models of 'peaceful co-existence' – autopilots on planes for example, have improved airline safety without compromising the training of pilots. There is little reason why the same cannot be true for medicine.

AI in the critically ill and on intensive care

Artificial intelligence has the potential for good in the critically ill, whether on the ward or particularly on the ICU or HDU. It has great potential to help ensure clinicians are aware of or able to prioritise the sickest or the deteriorating patient and make sure they receive optimal and timely treatment. However, there are difficult problems to overcome from sensor error rejection or even calibration error. The system must be able to sense check a differential diagnosis. There are also issues between continuously recorded and intermittently recorded data.

There will be learning weaknesses. The data set or programming may make assumptions – for example the system might assume the most common cause of hypotension on ICU is septic shock because of the associations it has learnt, without realising or having the data necessary to detect cardiogenic shock or considering that a fall in cardiac output could be due to another cause such as a pneumothorax. This is sometimes an issue with the original dataset used for learning and the breadth of what the AI was asked to learn about.

Artificial intelligence in the critically ill then promises great potential in terms of optimising treatment and providing or stimulating timely interventions. It poses difficulties in a fast changing environment where decisions are needed quickly with rapid implications for care with potential issues with artefact rejection from monitoring, and continuous versus intermittent data sampling. AI may struggle with wider areas of care, may suffer from training datasets providing associations which fall down when the situation lies outside the main areas covered by its learning. Clinical staff may lose faith in AI (sometimes inappropriately) or struggle to cope with what to do when their assessment is not in line with the AI, especially if they are inexperienced. The use of AI in prognostication poses great ethical issues for the influence for AI on human made assessments. That said, none of these reasons are sufficient on their own not to welcome AI – but as with so much in medicine, we need to be cautious and understand the potential shortcomings from the start.

Professor Gary Mills, Consultant in Intensive Care Medicine and Anaesthesia, Sheffield Teaching Hospitals.

Medical research

Artificial intelligence is ideally suited to analysing the large and complex data sets used in medical research. Pharmaceutical companies are looking to AI to streamline the development of new drugs, researchers can use predictive analytics to identify suitable candidates for clinical trials and scientists can create more accurate models of biological processes.

But there are challenges as well – for example, what dataset do you test new hypotheses against? And, as data linkage is held by many as the key to unlocking our knowledge of disease, would an algorithm be capable of coming to common sense conclusions?

There are plenty of questions around how useful machine learning will be in practice. Does this approach lead to the ecological fallacy, where aggregate data provides false answers? Will it overwhelmingly generate multiple instances of correlation without knowledge of causation, wasting researchers' time and resources and misleading the public? In any case, clinical input will be needed for the foreseeable future, to ensure the validity and relevance of research.

Clinical considerations:

- The margins of clinical and research consent are becoming blurred as clinical management and outcomes become more and more dependent on big data and 'research' becomes immediately relevant for individual patient care
- Machine learning can sift through terabytes of data to find patterns and correlations that humans might miss, freeing researchers from some of the more mundane tasks and potentially enabling 'big finds' in cohort studies
- On the other hand, automated research risks generating multiple instances of correlation without knowledge of causation, wasting researchers' time and resources. Clinical input will be needed for the foreseeable future, to ensure the validity and relevance of research.

Cochrane and AI – Project Transform

With more and more research being published, it is increasingly difficult for clinicians to keep up to date. Systematic reviews aim to give a complete summary of the current best evidence, by bringing together data from multiple different studies. However, they are painstakingly labour intensive and can take years to research and write. Cochrane's Project Transform, in partnership with Microsoft, is using AI to speed up the process by which systematic reviews are conducted.

Machine learning can be used to automate the literature search by using 'text mining' to analyse trial reports. Artificial intelligence can be used to inspect thousands of randomised trials, identify and categorise them and select which are appropriate for the systematic review. This drastically speeds up the time taken to conduct the literature review – the Project Transform team estimate a 60-80% reduction in research effort.

Ethical issues:

— Could the ability of AI and machine learning to analyse large data sets quickly and inexpensively skew the research landscape away from traditional medical studies and divert funding and effort away from 'gold standard' research methods?

— Fully informed consent and anonymity may be challenging to achieve. Is a new model of consent needed?

— How do developers and researchers prevent an algorithm identifying an individual patient if it is only analysing small cohorts, such as when looking at rare diseases for example?

Practical challenges:

— AI research needs be thoroughly evaluated for its effectiveness, cost effectiveness and the risk of unintended consequences.

— Researchers from technological backgrounds will need to act in accordance with the key underpinning principles of ethical medical research, including professional standards on maintaining confidentiality, transparency and minimising adverse effects.

— Might there be a negative or positive impact on recruitment to studies?

Artificial intelligence and machine learning techniques can allow datasets to be analysed far more quickly, thoroughly and inexpensively. It may be though, that there is a risk that this may lead to a shift towards research solely focusing on analysing large data sets, skewing the research landscape away from traditional medical studies and diverting funding and effort away from 'gold standard' research methods.

Researchers from technological backgrounds will need to be made aware of the key underpinning principles of ethical medical research, including professional standards on maintaining confidentiality, transparency and minimising adverse effects.

Accidental identification

Analysis on big data can blur the line between quality improvement and research requiring specific ethical approval and explicit informed consent. Fully informed consent may prove difficult, particularly given the potential for models to generate unexpected findings and for future advances in technology which may be applied to the same dataset. Truly anonymised data may become harder and harder to achieve – machine learning provides the ability to de-anonymise and re-identify patients from fewer and fewer data points. Is a new model of consent therefore needed?

The regulatory environment

At the heart of the development of AI in healthcare are questions around the regulatory environment. As with all regulation, a balance must be struck between protecting the public, clinicians and the service and promoting growth and innovation. These are not mutually exclusive concepts and there are past examples of good practice – for example, with the development of the appropriate ethical and legal considerations which underpinned the development of In-Vitro Fertilisation. Indeed many point out that it was thanks to early focus on regulation that the science was allowed to flourish. Lessons can be drawn for the development of AI.

The challenges to regulators presented by AI are diverse – the impact it is likely to have on medical systems and devices, clinical practice, relationships between clinicians and patients (and between providers of health-related applications marketed direct to patients) mean that regulators will need to work in a complementary way to develop relevant and appropriate regulatory frameworks for AI. While many AI products will meet the definition of a medical device and would therefore fall under the regulatory jurisdiction of the MHRA, there are also implications for:

— General Medical Council – clinicians will need clear guidelines on the appropriate use of AI

— Medical defence organisations – the nature of negligence claims may change as patients adapt to the availability of AI-generated decisions and recommendations

— Care Quality Commission – will need to consider how AI systems are embedded and used in healthcare organisations and their impact on quality of care

— NHS Digital – will have a role in clinical risk management in the development of health IT systems.

The advent of AI is a potential game-changer for healthcare and regulatory processes will need to adapt. For example, the current approach to safety relies greatly on a structured approach to foreseeing hazards which can be avoided or mitigated. In the 'black box' of machine learning, it will not necessarily be possible to foresee potential hazards, so new ways of conducting clinical safety processes may be needed for AI. Similarly, the regulatory framework for medical devices will need to adapt to the world of AI.

Emerging technologies will need to be tested to make sure they are robust – but how? Should the regulation of products be based upon the process for development, such as minimum dataset standards and clinician involvement, or on the quality of the output ('real world testing')? The former would be less labour-intensive but could potentially miss those that have gone through the right process but generated the wrong result due to error or unknown component factors. The reality may be that safeguards need to be built into the whole chain from development through to production.

There is already a plethora of apps providing advice direct to patients. A balance needs to be struck between effective regulation and encouraging innovation. Should products that provide autonomous diagnosis and management require a 'licence to practice'? Could they prescribe? How would indemnity be managed? Would clinicians be left dealing with the aftermath of errors or bad advice from an AI system? It might be argued that the level of regulation should be varied according to the risks – for example psychiatric patients, the young and the elderly might be at particular risk from any 'bad advice' from digitised systems. If this is the case, should systems aimed at such groups be regulated more closely?

Clinical considerations:

— Regulatory oversight of the correct use of AI products by clinicians will be as important as regulation of the content

— Clinical input into quality assurance of data, evidence-based review and real-world testing will be needed

— Will doctors be required to 'pick up the pieces' from AI errors or bad advice?

Ethical issues:

— Vulnerable groups, such as patients with psychiatric illness, are at particular risk from any 'bad advice' from digitised systems. Should systems aimed at such groups be regulated more closely?

— Is there a risk that AI will drive unsustainable demand leading to rationing?

— Could regulation halt progress by stifling innovation and preventing the technological industry from working at its usual pace?

Practical challenges:

— Should regulation of products be based upon the process for development, such as minimum dataset standards and clinician involvement, or on the quality of the output?

— Should there be international standards?

— In evolving/learning systems, how could improvement and progress be monitored in an ongoing way?

— Should the level of regulation of products aimed at patients be proportionate to the risk?

Regulators need to focus on two broad issues in tandem – is the process correct and is the content correct? Both aspects will bring fresh challenges as AI, by its very nature, is dynamic. An algorithm which meets clinical standards on a Monday, may be a different algorithm on a Tuesday.

As things stand, the current regulatory environment is only capable of approving or not approving people, procedures, medicines, devices or institutions in a static context. It may be that a 'light touch' approach to regulation will move towards approving (or not) the provider of AI and not the AI itself.

Intellectual property and the financial impact on the heathcare system

Healthcare is big business. The development of AI tools requires significant resource and expertise, for which creators and investors of capital, time and specialist knowledge are likely to expect to reap rewards for successful products. The development of AI technologies requires access to meaningfully labelled data and clinical strategic design. There is potential for the NHS to profit from selling data, or at least recoup some costs. Indeed some commentators put the value of the data it holds at £15bn – potentially an attractive sum in the era of budget-constrained healthcare system.

The economic gains to be made from healthcare AI are significant and could be of marked financial benefit to the country of ownership. For UK PLC, should those gains be made by the NHS, the public, or corporations? Will it be fair and equal to those contributing data (the public), advice and skills?

Technological advancements in AI have the potential to dramatically change the landscape of the healthcare system. They could be used to promote integration of services and data, leading to more streamlined and efficient care pathways. Direct-to-patient AI technologies have the potential to replace the need for a medical consultation in some cases, providing reassurance, advice, or direct access to simple treatments.

However, there is also potential to drive a new demand through drastically increased ease of access, leading to a large increase in the number of other contacts with the health service – particularly where systems err on the side of caution for reasons of safety. This could improve early detection of serious conditions, but could also lead to over-investigation and a vast new source of financial demand.

Clinical considerations:

— Who owns the data – does it belong to each patient, the public as a whole, the NHS, or the government? Who should provide consent and who should reap the rewards from any monetisation?

— An organisation as vast and complex as the NHS, with finite resources, may struggle to keep pace with rapid advancements in technology

— Consideration must be given to ensuring sufficient human resources are in place and can provide back-up if the IT systems, on which AI is based, fail or are hacked

— AI systems could reduce the need for some face-to-face consultations, reducing the financial burdens of travel for patients as well as facility costs.

Ethical issues:

— Advances in healthcare AI have potential to improve care globally. Do high-income countries have a humanitarian duty to share data and technologies with resource-poor countries, where the potential benefits to provide a higher standard of care are very marked?

— If the technology is owned by a private company, their choice of business model may exacerbate health inequalities if payment is required for a higher standard of service

— Any public money or staff time invested in developing healthcare AI has the potential to benefit an enormous number of people if successful. It is a challenge to find an ethical balance between potential future population heath gains with an unknown financial impact and the use of resources to treat current patients via conventional methods

— New technologies need to prove their real-world cost-effectiveness, particularly if they generate new need rather than serving existing unmet need.

Practical challenges:

— Financial interests from collaboration with technology companies may generate conflicts of interest, in parallel to those with pharmaceutical companies

— Should the NHS fund AI research, or collaborate with private partners in exchange for data sharing?

— The duration of intellectual property rights over technology may be a source of controversy. Allowing open access for peer review could promote safety, faster development and rapid improvements. However, if the companies creating the products do not possess the intellectual property for a significant period, the development of the technology may not be commercially viable, stifling progress.

It remains to be seen on which elements of the system AI will have the greatest initial impact. Medical investigations could be automatically identified and ordered in advance of face-to-face consultations so that results are immediately available to achieve a more rapid diagnosis. Primary care-like systems could diagnose and triage directly to secondary care, avoiding the need for a GP consultation, while secondary care-like systems equipped with up-to-date treatment algorithms could support GPs to manage conditions traditionally requiring specialist input. We must remain cognisant that integration of new AI technologies into services will involve parties with a range of financial interests and manage this with due care to achieve equitable benefits for all.

Impact on doctors' working lives

At a time of widespread clinician burnout and a shortage of staff, AI offers the potential to automate some of the workload and reduce the burden of routine tasks. This could leave doctors free to engage in the more interesting and challenging work and could present opportunities to work more flexibly. Some have feared that certain experts may be 'replaced' by AI in the long-term, leading to unemployment, although the breadth of skills and attributes required of a doctor cannot be easily replicated.

Artificial intelligence tools supporting clinical decision making could empower clinicians to work confidently in a wider range of areas, providing 'as needed' access to support from a repository of up-to-date knowledge. Underlying this is an implicit trust that the technologies can be relied upon, which will generate tensions if disagreement or loss of faith occurs.

Artificial intelligence could change the type of person who would choose to become a doctor. If sophisticated AI in the future were to take on a dominant role in talking to patients, information-processing and decision-making, this reduction in direct patient interaction and shift in professional role and tasks could significantly alter the day-to-day nature of medicine as a career.

Clinical considerations:

— Successful AI could improve clinical efficiency, helping doctors by automating 'non-human' tasks thousands of times faster than humans possibly can

— Decision support tools could increase doctors' confidence in managing cases of clinical uncertainty, or less familiar types of condition

— What will be the medicolegal position for a clinician who disagrees with the AI?

Ethical issues:

— If the public begin to view some of the skills gained through medical school and clinical practice as 'replaceable', will this disempower the medical profession and its organisations?

— A reduction in the social element of the consultation and reduced need for 'problem-solving', could affect job satisfaction.

Practical challenges:

— Will AI lead to unemployment or the shortening of medical careers in certain areas, or will this be counteracted by ever-growing service demand?

— Clinical practice involves a host of varied skills in patient interaction, information synthesis and decision-making. If technology encroaches on some of these domains, will this fundamentally change what it is to be a clinician and the type of person who would choose to become one?

Artificial intelligence could fundamentally change the way doctors work, as well as their relationships with patients. Modern medicine is a necessarily cautious and risk-averse industry. Will doctors be steering the direction of medical AI, or be overtaken by the rapid pace of technological development?

Clinical engagement is required to achieve harmony between the professions and the burgeoning healthcare technology market and to shape the advancement and deployment of these technologies for the benefit of patients.

I have no problem with AI replacing radiologists if it removes the humdrum work. I don't see AI ever replacing radiologists in the more complex cases or in interventional radiology. We'll always need high-quality radiologists in large numbers because the service is so reliant on the work we do. I think we are a very long way off replacing humans with machines in diagnostics

Dr Nicola Strickland, President, Royal College of Radiologists

Nurses, along with other professions working in health and social care, want tools that support them in their work. Technology is not value free and embodies the assumptions of designers. We need conversations with citizens, nurses and designers about how work will change. If we understand the strengths of each we can meet the common challenges faced by our health system. The time has come to redesign work itself. Nurses will help shape that future and the tools they will be using, like AI. We see this report as playing a part in that crucial ongoing conversation

Ross Scrivener, Digital Resources Manager and e-Health Lead,
Royal College of Nursing

Impact on the wider healthcare system

However it cuts, there are two visions of an AI enabled healthcare system. We could see a utopian world, where health inequalities are reduced, where access to care is dramatically improved and quality and standards of care are continuously driven up as machines learn more about the conditions of the people they are treating. The dystopian, but also feasible outcome is that health inequalities increase, or the system becomes overwhelmed by 'the worried well' who have arrived at their GPs' surgery or the Emergency Department because they have erroneously been told to attend by their AI enabled Fitbit or smartphone. Equally worrying is a world where only the wealthy will be able to access the best AI delivered healthcare as those providers will be the only ones with pockets deep enough to access the best data and develop the best AI. The reality, as with most revolutionary developments, is that the future will be located somewhere between the two.

It is for policymakers, politicians, legislators, clinicians and ethicists to decide now how the wider healthcare system will be AI enabled and improved for future generations.

Clinical considerations:

— Artificial intelligence in healthcare will generate whole new industries and disciplines around data management, computational science and medical informatics. How should this be planned for?

— It could dramatically reduce the cost of care in some specialties such as diagnostics through earlier and more accurate diagnosis. Equally, it could give rise to a dramatic increase in demand as patients self-refer for care

— Public health could be equally revolutionised as AI offers the opportunity to micro-target people or groups at risk of disease.

Ethical concerns:

— If AI brings rapid progress in the treatment of some diseases, could those who are not signed up to a data sharing environment be excluded from those advances?

— Should the AI developed in advanced western societies be shared with less advanced economies? Will it be shared equally with disadvantaged UK citizens – those that suffer from homelessness, mental illness and poverty?

— Could an individual patient's health data influence the quality of treatment they receive?

Practical issues:

— Rapid advances in technology and science may result in change fatigue, leaving NHS staff demoralised and unable to keep pace. A key challenge is to keep clinicians engaged from the outset

— Issues of information governance, public acceptance, funding limitation and lack of clinical engagement may prevent the potential benefits from being realised

— An open AI healthcare landscape may encourage the proliferation of alternatives to well-tested and validated treatments. Does this jeopardise medical engagement and oversight?

What might the future have in store?

Today, obese patients are routinely advised to lose weight before surgery. Should patients who are not obese, but who, according to their fitbit lead sedentary lives and who, according to their supermarket loyalty cards, buy more than three bottles of wine a week and have a high cholesterol diet, equally be denied access to surgery? Or charged some sort of premium?

It may sound far-fetched and certainly the Academy of Medical Royal Colleges as well as the Royal College of Surgeons of England would object in the strongest possible terms, but it is feasible now and society is already partially used to it. Most people who smoke accept the need to be charged more for private health insurance. With more and more data about people's lifestyles collected from sources outside the healthcare system, such as their mobile phones and fitness or diet apps, could or should these be accessed by healthcare providers to assess the likely outcome of a particular intervention? Is there an ethical imperative to begin to do this to reduce the overall cost of healthcare? Or conversely, is there an ethical imperative to beef-up legislation to ensure it cannot happen?

It is for politicians and the public to decide where the balance should lie. This report does not have a particular view other than to argue that the principles on which the NHS was founded – that good healthcare should be available to all regardless of wealth, should apply to the introduction and use of this game changing technology, just as they have done to any other significant clinical development over the last 70 years.

Potential problems await every time there is human interaction with software or hardware, so the potential for blame shifting seems limitless. The way through this is to ensure appropriate lay, professional and industrial governance and make this clear.

Professor Jo Martin, President,
Royal College of Pathologists

Glossary

Algorithm

A step by step mathematical method of solving a problem. It is commonly used for data processing, calculation and other related computer and mathematical operations.

App

An abbreviation of application. Computer software, or a program – most commonly a small, specific one used for mobile devices.

Artificial Intelligence (AI)

The simulation of human intelligence processes by machines, especially computer systems. These processes include learning (the acquisition of information and rules for using the information), reasoning (using rules to reach approximate or definite conclusions) and self-correction.

Automation Bias

The propensity for humans to favour suggestions from automated decision-making systems and to ignore contradictory information made without automation, even if it is correct.

Babylon

Babylon is a subscription health service provider that enables users to have virtual consultations with doctors and health care professionals via text and video messaging through its mobile application.

Black Box

In science, computing and engineering, a black box is a device, system or object which can be viewed in terms of its inputs and outputs (or transfer characteristics), without any knowledge of its internal workings. Its implementation is 'opaque' and is therefore referred to as 'black.'

Chatbot

An artificial intelligence (AI) program that simulates interactive human conversation by using key pre-calculated user phrases and auditory or text-based signals.

Cognitive Behavioural Therapy

A type of psychotherapy in which negative patterns of thought about the self and the world are challenged in order to alter unwanted behaviour patterns or treat mood disorders such as depression.

DeepMind

DeepMind Technologies Ltd. is a firm based in the United Kingdom that works on artificial intelligence problems. It is part of the Google Alphabet group.

Machine Learning

An application of AI that provides systems with the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.

General Data Protection Regulations (GDPR)

The legal framework which sets guidelines for the collection and processing of personal information of individuals within the European Union (EU). It came into effect across the EU on 25 May 2018.

Information Governance (IG)

The management of information at an organisation.

Neural Network

A series of algorithms that endeavours to recognise underlying relationships in a set of data through a process that mimics the way the human brain operates.

Deep Neural Network

A neural network with a certain level of complexity, a neural network with more than two layers. Deep neural networks use sophisticated mathematical modelling to process data in complex ways.

NHS Digital

The national information and technology partner to the health and social care system. Its roles include:

- Supplying information and data to the health service
- Providing technological infrastructure
- Acting as the guardian of patient data
- Advising health and care on cyber and data security.

Terabyte

A unit of information equal to one million million bytes. One terabyte could store 130,000 digital images.

Further Reading

Much has been and doubtless will be written about the use of AI in healthcare and in society more widely. The reports and guidance the authors of this report found most useful were:

Reform	Thinking on its own: AI in the NHS
Nesta	Confronting Dr Robot
The AHSN Network	Accelerating AI in health and care
Nuffield Council on Bioethics	AI in healthcare and research
PWC	Why AI and robotics will define New Health
Future Advocacy	Ethical, social and political challenges of AI in health
House of Commons	Algorithms in decision making
Information Commissioner	Big data, AI, machine learning and data protection
House of Lords	AI in the UK (Chapter 7 Healthcare and AI)
CMO's annual report 2018	Machine learning for individualised medicine
CMO's annual report 2018	Emerging technologies in healthcare
RCS England	The Future of Surgery
Dept of Health and Social Care	Code of Conduct for data driven technologies

Thanks

The Academy of Medical Royal Colleges is grateful to the following people, who gave time to speak to us, attend focus groups, give advice, read and comment on the various drafts.

Paul Alexander, Policy and Academic Research Manager, Royal College of Radiologists

Dr Hannah Allen, Associate Medical Director, Babylon Health.

Professor Richard Ashcroft, Professor of Bioethics, Queen Mary University of London

Tim Atkins, Head of Strategy, Care Quality Commission

Professor Maureen Baker, Chair, Professional Record Standards Body

Professor Ted Baker, Chief Inspector of Hospitals, Care Quality Commission

Jayne Black, Joint Head of Policy and Campaigns (London) at the Royal College of Physicians

Chris Brooks, General Medical Council (GMP guidance)

Hannah Burd, Principal Advisor, Behavioural Insights Team

Dr John Chisholm, Chair of the British Medical Association's Ethics Committee

Shirley Cramer CBE, Chief Executive of the Royal Society for Public Health

Professor Finale Doshi-Velez, Assistant Professor in Computer Science at the Harvard Paulson School of Engineering and Applied Science

Professor Don Detmer Professor Emeritus and Professor of Medical Education at University of Virginia

Dr Simon Eccles, Chief Clinical Information Officer for Health and Care, NHSE, NHSI, DHSC

Professor Bobbie Farsides, Professor of Clinical and Biomedical Ethics, Brighton & Sussex Medical School

Dr Matthew Fenech, Affiliate Consultant, Artificial Intelligence in Health at Future Advocacy

Dr Tom Foley, Senior Clinical Lead, NHS Digital

Professor John Fox, Chairman, OpenClinical CIC, Chief Scientific Officer, Deontics Ltd

Malte Gerhold, Executive Director of Strategy and Intelligence, Care Quality Commission

Professor Andrew Goddard, President, Royal College of Physicians of London

Rose Gray, Policy Manager, Cancer Research UK

Professor Nina Hallowell, Associate Professor, Nuffield Department of Population Health, Oxford

Dr Steven Hamblin, Chief Technology Officer at Drayson Technologies

Eleonora Harwich, Head of Digital and Technological Innovation, Reform

Dr Matt Hoghton, Medical Director, Clinical Innovation and Research, Royal College of General Practitioners

Dr Ian Hudson, Chief Executive of the Medicines and Healthcare Products Regulatory Agency

Dr Cían Hughes, Applied Artificial Intelligence Researcher at DeepMind

Dr Indra Joshi, Digital Health and AI Clinical Lead, NHS England

Dr Catherine Kelly, Clinical Advisor, Digital Health and Care, Scottish Government

Mr Richard Kerr, Chair of the Royal College of Surgeons' Commission on the Future of Surgery

Professor Ronald Kessler, McNeil Family Professor of Health Care Policy at Harvard Medical School

Benedict Knox, Head of Communications, Healthwatch England

Dr Phil Koczan GP, CCIO for digital integration and NHS England (London) and Clinical Advisor to the Professional Record Standards Body.

Professor Sir Robert Lechler, President, Academy of Medical Sciences

Professor Martin Marshall, Vice-Chair, Royal College of General Practitioners (RCGP)

Phil Martin, Assistant Director, Education Policy, General Medical

Professor Gary Mills, Consultant in Intensive Care Medicine and Anaesthesia, Sheffield Teaching Hospitals

Jess Morley, Technology Advisor, Department of Health and Social Care

Dr Imran Rafi, Chair of the RCGP's Clinical Innovation and Research Centre (CIRC)

Dr Navin Ramachandran, Consultant Radiologist UCLH & Co-founder of OpenCancer and PEACH

Dr Jem Rashbass, National Director for Disease Registration, Public Health England

Peter Rees, Chair, Patient Lay Committee, Academy of Medical Royal Colleges

Nripsuta Saxena, Computer Scientist specializing in Algorithm Bias, University of Southern California

Ross Scrivener, e-Health Lead, Royal College of Nursing

Professor Martin Severs, Medical Director, NHS Digital

Vibha Sharma, Regulation Policy Manager, General Medical Council

Professor Helen Stokes-Lampard, Chair, Royal College of General Practitioners

Dr Nicola Strickland, President, Royal College of Radiologists

Dr Kenji Takeda, Director of the Microsoft Azure for Research Program

Dr Berk Ustun, Postdoctoral Fellow, Center for Research in Computation and Society, Harvard University

Professor Stephen Wilkinson, Professor of Bioethics in the Department of Politics, Philosophy and Religion, Lancaster University

Dr Naho Yamazaki, Head of Policy, Academy of Medical Sciences

Thanks must also go to Rosie Carlow for her diligent proof reading and progress chasing and James Taylor for the great design and layout work as well as meeting some tight deadlines.

Max Prangnell, Communications Director, AoMRC.

Academy of Medical Royal Colleges
10 Dallington Street
London
EC1V 0DB
United Kingdom

Telephone: +44 (0)20 7490 6810
Facsimile: +44 (0)20 7470 6811
Email: academy@aomrc.org.uk
Website: aomrc.org.uk

Registered Charity Number:
1056565