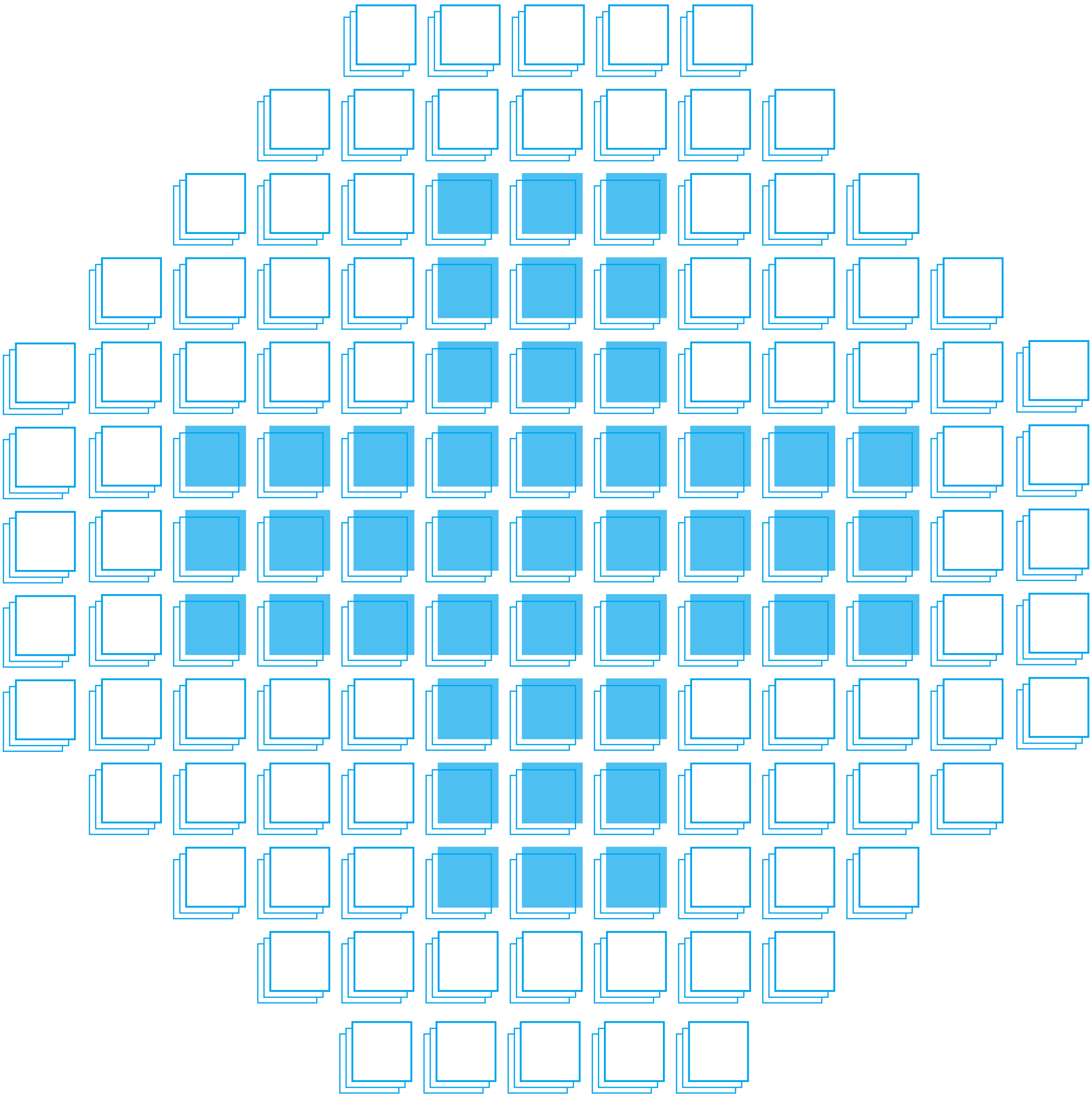
Advancing AI in the NHS



LEV TANKELEVITCH / ALICE AHN / RACHEL PATERSON
MATTHEW REID / EMILY HILBOURNE / BRYAN ADRIAANSE
GIORGIO QUER / DROR NIR / PARTH PATEL

POLYGEIA
STUDENTS SHAPING GLOBAL HEALTH POLICY

Acknowledgements

We and *Polygeia* as a whole are grateful to Dr Dror Nir, Director, *RadBee*, whose insights were valuable throughout the research, conceptualisation, and writing phases of this work; and to Dr Giorgio Quer, Senior Research Scientist, *Scripps Research Institute*; Dr Matt Willis, *Oxford Internet Institute, University of Oxford*; Professor Eric T. Meyer, *Oxford Internet Institute, University of Oxford*; Alexander Hitchcock, Senior Researcher, *Reform*; Windi Hari, Vice President Clinical, Quality & Regulatory, *HeartFlow*; Jon Holmes, co-founder and Chief Technology Officer, *Vivosight*; and Claudia Hartman, *School of Anthropology & Museum Ethnography, University of Oxford* for their advice and support.

Author affiliations

Lev Tankelevitch, *University of Oxford*
Alice Ahn, *University of Oxford*
Rachel Paterson, *University of Oxford*
Matthew Reid, *University of Oxford*
Emily Hilbourne, *University of Oxford*
Bryan Adriaanse, *University of Oxford*
Giorgio Quer, *Scripps Research Institute*
Dror Nir, *RadBee*
Parth Patel, *University of Cambridge*

All affiliations are at the time of writing.

Polygeia

Polygeia is an independent, non-party, and non-profit think-tank focusing on health and its intersection with technology, politics, and economics. Our aim is to produce high-quality research on global health issues and policies. With branches in Oxford, Cambridge, London, and New York, our work has led to policy reports, peer-reviewed publications, and presentations at the House of Commons and the European Parliament.

www.polygeia.com @Polygeia

Foreword

Almost every day, as MP for Cambridge, I am told of new innovations and developments that show that we are on the cusp of a technological revolution across the sectors. This technology is capable of revolutionising the way we work; incredible innovations which could increase our accuracy, productivity and efficiency and improve our capacity for creativity and innovation.

But huge change, particularly through adoption of new technology, can be difficult to communicate to the public, and if we do not make sure that we explain carefully the real benefits of such technologies we easily risk a backlash. Despite good intentions, the *care.data* programme failed to win public trust, with widespread worries that the appropriate safeguards weren't in place, and a failure to properly explain potential benefits to patients. It is vital that the checks and balances we put in place are robust enough to sooth public anxiety, and prevent problems which could lead to steps back, rather than forwards.

Previous attempts to introduce digital innovation into the NHS also teach us that cross-disciplinary and cross-sector collaboration is essential. Realising this technological revolution in healthcare will require industry, academia and the NHS to work together and share their expertise to ensure that technical innovations are developed and adopted in ways that prioritise patient health, rather than innovation for its own sake.

Alongside this, we must make sure that the NHS workforce whose practice will be altered by AI are on side. Consultation and education are key, and this report details well the skills that will be vital to NHS adoption of AI. Technology is only as good as those who use it, and for this, we must listen to the medical and healthcare professionals who will rightly know best the concerns both of patients and their colleagues.

The new Centre for Data Ethics and Innovation, the ICO and the National Data Guardian will be key in working alongside the NHS to create both a regulatory framework and the communications which win society's trust. With this, and with real leadership from the sector and from politicians, focused on the rights and concerns of individuals, AI can be advanced in the NHS to help keep us all healthy.

Daniel Zeichner
MP for Cambridge
Chair, All-Party Parliamentary Group on Data Analytics

Contents

Executive summary	**4**
Recommendations	**6**
Introduction	**8**
What is AI?	9
Applications in healthcare	9
Data governance	**16**
Data sharing	17
Consent and opt-outs	19
Technology	22
Digital infrastructure	**24**
High quality and quantity digital data	24
Broader IT infrastructure	26
Standards	**28**
Training data	29
Performance evaluation	30
Interpretability	31
Usability	33
Regulations	**34**
Regulatory guidelines	34
Regulatory frameworks	37
Supporting healthcare innovators	38
Brexit	39
The workforce	**40**
Estimating workforce impact	40
Re-thinking workforce impact	42
Developing capability for AI	43
Working with AI	44
Leadership	45
Legal liability	**47**
The duty of healthcare professionals	47
The duty of healthcare institutions and manufacturers	49
The public's role	49
References	**50**

Executive summary

Artificial intelligence (AI) has the potential to transform how the NHS delivers care. From enabling patients to self-care and manage long-term conditions, to advancing triage, diagnostics, treatment, research, and resource management, AI can improve patient outcomes and increase efficiency. Achieving this potential, however, requires addressing a number of ethical, social, legal, and technical challenges. This report describes these challenges within the context of healthcare and offers directions forward.

Data governance

AI-assisted healthcare will demand better collection and sharing of health data between NHS, industry and academic stakeholders. This requires a data governance system that ensures ethical management of health data and enables its use for the improvement of healthcare delivery. Data sharing must be supported by patients. The recently launched NHS data opt-out programme is an important starting point, and will require monitoring to ensure that it has the transparency and clarity to avoid exploiting the public's lack of awareness and understanding. Data sharing must also be streamlined and mutually beneficial. Current NHS data sharing practices are disjointed and difficult to negotiate from both industry and NHS perspectives. This issue is complicated by the increasing integration of 'traditional' health data with that from commercial apps and wearables. Finding approaches to valuate data, and considering how patients, the NHS and its partners can benefit from data sharing is key to developing a data sharing framework. Finally, data sharing should be underpinned by digital infrastructure that enables cybersecurity and accountability.

Digital infrastructure

Developing and deploying AI-assisted healthcare requires high quantity and quality digital data. This demands effective digitisation of the NHS, especially within secondary care, involving not only the transformation of paper-based records into digital data, but also improvement of quality assurance practices and increased data linkage. Beyond data digitisation, broader IT infrastructure also needs upgrading, including the use of innovations such as wearable technology and interoperability between NHS sectors and institutions. This would not only increase data availability for AI development, but also provide patients with seamless healthcare delivery, putting the NHS at the vanguard of healthcare innovation.

Standards

The recent advances in AI and the surrounding hype has meant that the development of AI-assisted healthcare remains haphazard across the industry, with quality being difficult to determine or varying widely. Without adequate product validation, including in real-world settings, there is a risk of unexpected or unintended performance, such as sociodemographic biases or errors arising from inappropriate human-AI interaction. There is a need to develop standardised ways to probe training data, to agree upon clinically-relevant performance benchmarks, and to design approaches to enable and evaluate algorithm interpretability for productive human-AI interaction. In all of these areas, *standardised* does not necessarily mean one-size-fits-all. These issues require addressing the specifics of AI within a healthcare context, with consideration of users' expertise, their environment, and products' intended use. This calls for a fundamentally interdisciplinary approach, including experts in AI, medicine, ethics, cognitive science, usability design, and ethnography.

Regulations

Despite the recognition of AI-assisted healthcare products as medical devices, current regulatory efforts by the UK Medicines and Healthcare Products Regulatory Agency and the European Commission have yet to be accompanied by detailed guidelines which address questions concerning AI product classification, validation, and monitoring. This is compounded by the uncertainty surrounding Brexit and the UK's future relationship with the European Medicines Agency. The absence of regulatory clarity risks compromising patient safety and stalling the development of AI-assisted healthcare. Close working partnerships involving regulators, industry members, healthcare institutions, and independent AI-related bodies (for example, as part of regulatory sandboxes) will be needed to enable innovation while ensuring patient safety.

The workforce

AI will be a tool for the healthcare workforce. Harnessing its utility to improve care requires an expanded workforce with the digital skills necessary for both developing AI capability and for working productively with the technology as it becomes commonplace. Developing capability for AI will involve finding ways to increase the number of clinician-informaticians who can lead the development, procurement and adoption of AI technology while ensuring that innovation remains tied to the human aspect of healthcare delivery. More broadly, healthcare professionals will need to complement their socio-emotional and cognitive skills with training to appropriately interpret information provided by AI products and communicate it effectively to co-workers and patients. Although much effort has gone into predicting how many jobs will be affected by AI-driven automation, understanding the impact on the healthcare workforce will require examining *how* jobs will change, not simply how many will change.

Legal liability

AI-assisted healthcare has implications for the legal liability framework: who should be held responsible in the case of a medical error involving AI? Addressing the question of liability will involve understanding how healthcare professionals' duty of care will be impacted by use of the technology. This is tied to the lack of training standards for healthcare professionals to safely and effectively work with AI, and to the challenges of algorithm interpretability, with "black-box" systems forcing healthcare professionals to blindly trust or distrust their output. More broadly, it will be important to examine the legal liability of healthcare professionals, NHS trusts and industry partners, raising questions about medical ethics, workforce training, product regulation, and public support.

Recommendations

1. **The NHS, the Centre for Data Ethics and Innovation, and industry and academic partners** should conduct a review to understand the obstacles that the NHS and external organisations face around data sharing. They should also develop health data valuation protocols which consider the perspectives of patients, the NHS, commercial organisations, and academia. This work should inform the development of a data sharing framework.
2. **The National Data Guardian and the Department of Health** should monitor the NHS data opt-out programme and its approach to transparency and communication, evaluating how the public understands commercial and non-commercial data use and the handling of data at different levels of anonymisation.
3. **The NHS, patient advocacy groups, and commercial organisations** should expand public engagement strategies around data governance, including discussions about the value of health data for improving healthcare; public and private sector interactions in the development of AI-assisted healthcare; and the NHS's strategies around data anonymisation, accountability, and commercial partnerships. Findings from this work should inform the development of a data sharing framework.
4. **The NHS Digital Security Operations Centre** should ensure that all NHS organisations comply with cybersecurity standards, including having up-to-date technology.
5. **NHS Digital, the Centre for Data Ethics and Innovation, and the Alan Turing Institute** should develop technological approaches to data privacy, auditing, and accountability that could be implemented in the NHS. This should include learning from Global Digital Exemplar trusts in the UK and from international examples such as Estonia.
6. **The NHS** should continue to increase the quantity, quality, and diversity of digital health data across trusts. It should consider targeted projects, in partnership with professional medical bodies, that quality-assure and curate datasets for more deployment-ready AI technology. It should also continue to develop its broader IT infrastructure, focusing on interoperability between sectors, institutions, and technologies, and including the end users as central stakeholders.
7. **The Alan Turing Institute, the Ada Lovelace Institute, and academic and industry partners in medicine and AI** should develop ethical frameworks and technological approaches for the validation of training data in the healthcare sector, including methods to minimise performance biases and validate continuously-learning algorithms.
8. **The Alan Turing Institute, the Ada Lovelace Institute, and academic and industry partners in medicine and AI** should develop standardised approaches for evaluating product performance in the healthcare sector, with consideration for existing human performance standards and products' intended use.
9. **The Alan Turing Institute, the Ada Lovelace Institute, and academic and industry partners in medicine and AI** should develop methods of enabling and evaluating algorithm interpretability in the healthcare sector. This work should involve experts in AI, medicine, ethics, usability design, cognitive science, and ethnography, among others.
10. **Developers of AI products and NHS Commissioners** should ensure that usability design remains a top priority in their respective development and procurement of AI-assisted healthcare products.

11. **The Medicines and Healthcare Products Regulatory Agency** should establish a digital health unit with expertise in AI and digital products that will work together with manufacturers, healthcare bodies, notified bodies, AI-related organisations, and international forums to advance clear regulatory approaches and guidelines around AI product classification, validation, and monitoring. This should address issues including training data and biases, performance evaluation, algorithm interpretability, and usability.

12. **The Medicines and Healthcare Products Regulatory Agency, the Centre for Data Ethics and Innovation, and industry partners** should evaluate regulatory approaches, such as regulatory sandboxing, that can foster innovation in AI-assisted healthcare, ensure patient safety, and inform on-going regulatory development.

13. **The NHS** should expand innovation acceleration programmes that bridge healthcare and industry partners, with a focus on increasing validation of AI products in real-world contexts and informing the development of a regulatory framework.

14. **The Medicines and Healthcare Products Regulatory Agency and other Government bodies** should arrange a post-Brexit agreement ensuring that UK regulations of medical devices, including AI-assisted healthcare, are aligned as closely as possible to the European framework and that the UK can continue to help shape Europe-wide regulations around this technology.

15. **The General Medical Council, the Medical Royal Colleges, Health Education England, and AI-related bodies** should partner with industry and academia on comprehensive examinations of the healthcare sector to assess which, when, and *how* jobs will be impacted by AI, including analyses of the current strengths, limitations, and workflows of healthcare professionals and broader NHS staff. They should also examine how AI-driven workforce changes will impact patient outcomes.

16. **The Federation of Informatics Professionals and the Faculty of Clinical Informatics** should continue to lead and expand standards for health informatics competencies, integrating the relevant aspects of AI into their training, accreditation, and professional development programmes for clinician-informaticians and related professions.

17. **Health Education England** should expand training programmes to advance digital and AI-related skills among healthcare professionals. Competency standards for working with AI should be identified for each role and established in accordance with professional registration bodies such as the General Medical Council. Training programmes should ensure that "un-automatable" socio-emotional and cognitive skills remain an important focus.

18. **The NHS Digital Academy** should expand recruitment and training efforts to increase the number of Chief Clinical Information Officers across the NHS, and ensure that the latest AI ethics, standards, and innovations are embedded in their training programme.

19. **Legal experts, ethicists, AI-related bodies, professional medical bodies, and industry** should review the implications of AI-assisted healthcare for legal liability. This includes understanding how healthcare professionals' duty of care will be affected, the role of workforce training and product validation standards, and the potential role of NHS Indemnity and no-fault compensation systems.

20. **AI-related bodies such as the Ada Lovelace Institute, patient advocacy groups and other healthcare stakeholders** should lead a public engagement and dialogue strategy to understand the public's views on liability for AI-assisted healthcare.

1 Introduction

The growing and ageing population, along with budgetary constraints, are placing immense pressure on the NHS.[1] The consequences of this pressure are evident, with primary care inaccessible for many patients and urgent care services overflowing.[2] In its update to the *Five Year Forward View*, the NHS has highlighted technology and innovation as key to achieving its aims of widening primary care access, alleviating the strain on urgent care services, improving patient safety, and increasing efficiency.[3]

Artificial intelligence (AI)—technology that enables the completion of tasks that would otherwise require some form of intelligence—has emerged as a tool with immense potential for transforming how the NHS delivers healthcare.[4] From enabling patients to self-care and manage long-term conditions, to advancing triage, diagnostics, treatment, research, and resource management, AI can improve care while reducing cost.[5]

The time is ripe for AI. There has been a nine-fold increase in the number of AI-related academic papers published each year since 1996.[6] At least £19 billion was invested in AI in 2016.[7] In the UK, there are over 200 AI startups, along with major tech companies, like Google, Microsoft, and IBM, that regularly collaborate with UK industries on AI projects.[8] Within this market, healthcare is consistently the top industry for investment.[9] The UK Government has recognised the value of AI, prioritising the establishment of the UK "as a world leader in new technologies such as artificial intelligence", setting out plans for a new Centre for Data Ethics and Innovation, and announcing a £300 million investment in new healthcare technology.[10]

Despite AI's potential to revolutionise healthcare, discussions around the technology have been accompanied by hype, both positive and negative.[11] The promises of AI have been touted at the expense of the ethical, social, legal, and technical challenges that it presents.[12] Likewise, fears around AI—including valid concerns about job automation and existential risks—are often highlighted without the appropriate context.[13] Rather than being one uniform *thing*, AI is best considered as a tool that can be applied for diverse purposes in myriad ways.

Fortunately, the field is progressing. Throughout this project's duration, several valuable reports were published on AI and its use in healthcare, including from *Reform*, the *House of Lords Select Committee on AI*, *Future Advocacy* and *Wellcome*, *Nesta*, the *Nuffield Council on Bioethics*, and the *House of Commons Science and Technology Committee*.[14] In this report, we hope to contribute to this timely discussion, focusing on AI's application to healthcare within the UK, the key policy issues that arise in this context, and possible directions forward.

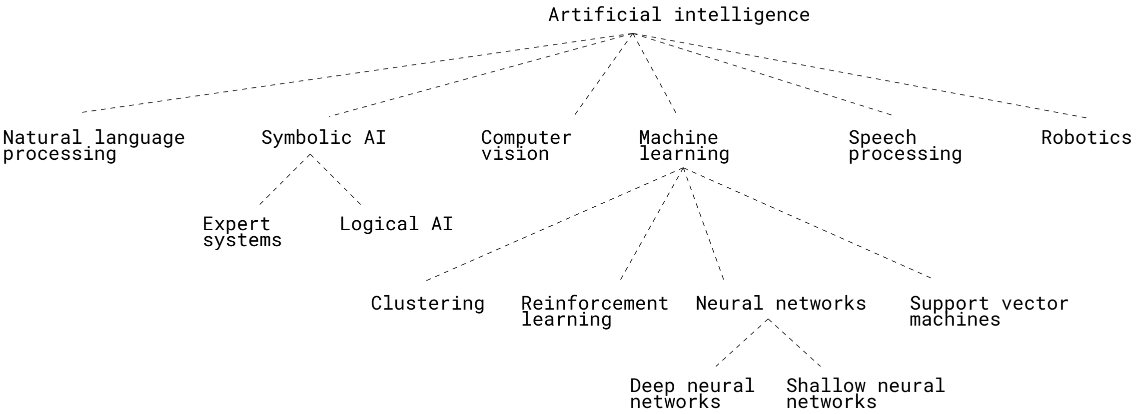
---

1 NHS England (2017). *Next steps on the NHS Five Year Forward View*.
2 Ibid.
3 Ibid.
4 Hall, W. and Pesenti, J. (2017). *Growing the artificial intelligence industry in the UK.*.
5 Stone, P. et al. (2016). *"Artificial Intelligence and Life in 2030." One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel*. Stanford University.
6 AI100 (2017). *AI Index 2017 Annual Report*.
7 Bughin, J. et al. (2017). *Artificial intelligence: the next digital frontier*? McKinsey Global Institute.
8 Kelnar, D. (2016). *Artificial Intelligence in the UK: Landscape and learnings from 226 startups*.
9 CB Insights Research (2017). *Up And Up: Healthcare AI Startups See Record Deals*.
10 HM Treasury (2017). *Autumn Budget 2017*.
11 Malik, O. (2016). "The Hype–and Hope–of Artificial Intelligence". *The New Yorker*.
12 Russell, S., Dewey, D., and Tegmark, M. (2015). "Research Priorities for Robust and Beneficial Artificial Intelligence". *AI Magazine* 36.4, pp. 105–114.
13 Etzioni, O. (2016). "Most experts say AI isn't as much of a threat as you might think". *MIT Technology Review*.
14 Harwich, E. and Laycock, K. (2018). *Thinking on its own: AI in the NHS.*. Reform; House of Lords Select Committee on Artificial Intelligence (2018). *AI in the UK: ready, willing and able?*; Fenech, M., Strukelj, N., and Buston, O. (2018). *Ethical, social, and political challenges of artificial intelligence in health.* Future Advocacy; Loder, J. and Nicholas, L. (2018). *Confronting Dr Robot: Creating a people-powered future for AI in health*. Nesta; Nuffield Council on Bioethics (2018). *Bioethics briefing note: Artificial intelligence (AI) in healthcare and research*. House of Commons Science and Technology Committee (2018). *Algorithms in decision-making*.

1.1 What is AI?

AI is a field encompassing a broad range of technologies that enable the completion of tasks that would otherwise require some form of intelligence (see Figure 1).[15] Within AI, *machine learning*—an approach in which algorithms learn from data rather being explicitly programmed—is currently the largest subfield, and the method behind most of today's applications of AI.[16] Within machine learning, *artificial neural networks*, such as *deep learning* algorithms, are one particularly powerful category of algorithms in use today.[17]

Figure 1: Branches of artificial intelligence



Source: JASON (2017). *Perspectives on Research in Artificial Intelligence and Artificial General Intelligence Relevant to DoD*, and other research. Note that categories and definitions of AI and its sub-fields are ambiguous and can overlap extensively. For example, computer vision can be considered an application of machine learning, rather than a sub-field. This list is not intended to be exhaustive.

Another important distinction is that between *general* and *narrow* AI.[18] General AI is the capability to perform many tasks flexibly, across a range of environments (and perhaps exhibiting sentience), more akin to a human.[19] However, this is currently considered a goal more than a reality.[20] In contrast, narrow AI is the capability to perform only very specific tasks (e.g., detecting stroke in medical images), and is the type of technology behind all of today's AI applications.[21] This distinction is relevant because general and narrow AI each present unique policy issues that require distinct approaches and timescales of action.[22] Because of its current deployment and immediate relevance to healthcare, this report focuses on narrow AI.

1.2 Applications in healthcare

AI's growth in healthcare is partly due to its synergy with other trends in health, such as preventative healthcare, self-care, and precision medicine.[23] AI also benefits from other trends in technology, including smartphones, Internet of Things devices, and

15 Russell, S. and Norvig, P. (2016). *Artificial Intelligence: A Modern Approach*.
16 The Royal Society (2017). *Machine learning: the power and promise of computers that learn by example*.
17 Jordan, M. I. and Mitchell, T. M. (2015). "Machine learning: Trends, perspectives, and prospects". *Science* 349.6245, pp. 255–260.
18 Russell and Norvig, *Artificial Intelligence*.
19 Ibid.
20 JASON (2017b). *Perspectives on Research in Artificial Intelligence and Artificial General Intelligence Relevant to DoD*. JSR-16-Task-003. The MITRE Corporation.
21 The Royal Society, *Machine learning: the power and promise of computers that learn by example*.
22 JASON, *Perspectives on Research in Artificial Intelligence and Artificial General Intelligence Relevant to DoD*.
23 Bhavnani, S. P. et al. (2017). "2017 Roadmap for Innovation—ACC Health Policy Statement on Healthcare Transformation in the Era of Digital Health, Big Data, and Precision Health: A report of the American College of Cardiology Task Force on Health Policy Statements and Systems of Care". *J. Am. Coll. Cardiol.* 70.21, pp. 2696–2718.

consumer-grade wearables, which have enabled industry and academia to easily collect large and diverse datasets to develop AI algorithms.[24] Likewise, the advancement of cloud computing has allowed companies to rapidly and cheaply scale up products, while enabling healthcare systems to benefit from the technology without as much investment in infrastructure.[25] These trends, along with significant advances in machine learning and computer hardware, have enabled companies to apply AI to almost every aspect of healthcare. We identified at least 43 companies applying AI to healthcare that are based or have offices in the UK (see Figure 2). In this section we provide an overview of the range of AI applications in healthcare, including examples from across the world.

Figure 2: AI healthcare companies in the UK

Personal health management

Triage

Diagnostics

Decision support

Patient monitoring

Treatment

Resource management

Medical research

Multiple

Source: online research. Note that this list is not intended to be exhaustive, company categorisations may be arbitrary, and the use of AI in each company was not verified.

Promotion of health, prevention of illness

With long-term conditions such as diabetes accounting for 70% of total health and social care spend in the UK, promoting health and preventing such conditions are core components of the NHS's plan to reduce demand for healthcare.[26] To help advance this goal, smartphone apps, wearables, and home sensors are being transformed through AI algorithms into personal health monitoring tools that can enable individuals to independently monitor and improve their health.[27]

For example, the app Lark uses an AI-powered virtual health coach, together with diet and

24 Bhavnani et al., "2017 Roadmap for Innovation—ACC Health Policy Statement on Healthcare Transformation in the Era of Digital Health, Big Data, and Precision Health: A report of the American College of Cardiology Task Force on Health Policy Statements and Systems of Care"; JASON, *Perspectives on Research in Artificial Intelligence and Artificial General Intelligence Relevant to DoD*.
25 Bhavnani et al., "2017 Roadmap for Innovation—ACC Health Policy Statement on Healthcare Transformation in the Era of Digital Health, Big Data, and Precision Health: A report of the American College of Cardiology Task Force on Health Policy Statements and Systems of Care".
26 NHS England, *Next steps on the NHS Five Year Forward View*; Iacobucci, G. (2017). "NHS in 2017: Keeping pace with society". *BMJ* 356, p. i6738.
27 JASON (2017a). *Artificial Intelligence for Health and Health Care*. JSR-17-Task-002. The MITRE Corporation.

exercise tracking, to help people lose weight and prevent the development of diabetes.[28] With over a million users, a recent study found the app's efficacy to be comparable to programs led by in-person healthcare professionals.[29] Similarly, the NHS is working together with Verily and Merck on algorithms that analyse vital signs to identify and engage individuals at risk of developing long-term conditions such as heart disease.[30]

Beyond promoting overall health, AI can help with early detection of specific conditions before complications develop. One example is atrial fibrillation, a heart condition that increases the risk of stroke.[31] Detecting it early is essential for preventing strokes, yet estimates indicate that over 400,000 people aged 64 years or older have gone undiagnosed in the UK.[32] Detection requires recording electrocardiogram (ECG) data as it occurs and having it analysed by a professional.[33] The Kardia Mobile ECG is a small smartphone-enabled, AI-driven ECG device which enables people to easily test for the condition at home, by capturing and automatically analysing ECG data.[34] Part of the NHS Innovation Accelerator, it is currently being used across 40 NHS organisations, with a potential savings of £968 per patient.[35] Still, it is possible to make detection of atrial fibrillation even simpler with an AI-powered app being developed by company Cardiogram to *passively* screen for the condition using data from smartwatches, eliminating the need for explicit testing and additional ECG devices.[36]

Patient intake and triage

More than one in ten people struggle to get a GP appointment, while 27% of appointments are potentially avoidable, with inappropriate patient referral a top reason.[37] AI has the potential to optimise the patient intake and triage process, reducing the burden on the NHS.

Babylon Health's *GP at hand* service and Sensely's *Ask NHS* both provide patients with an AI-powered triage service—a smartphone chat with a chatbot—to determine whether a GP appointment or other service is appropriate.[38] Other startups offering similar services, such as Your.MD and Ada Health, are catching up with over a million active users.[39] AI enables these apps to set up natural-sounding but automated conversations, in which patients are asked medical questions tailored to their symptoms, in line with learnt clinical pathways, and then referred to another service or provided with self-care information.[40]

Diagnosis

By automating, refining, and speeding up aspects of diagnosis, AI can be a powerful complementary tool for the healthcare professional, enabling them to focus on the doctor-patient relationship and the medical nuances of each patient.[41]

The NHS estimates that optimising the deployment of pathologists and diagnostic imaging services can improve healthcare delivery and save up to £130 million per year.[42] Medical imaging has been a popular target of research and development in AI, with its relatively standardised images and suitability for powerful deep learning algorithms.[43] CT scans, MRIs, radiographs, echocardiograms, and dermoscopy images are all being targeted for

---

28 Stein, N. and Brooks, K. (2017). "A Fully Automated Conversational Artificial Intelligence for Weight Loss: Longitudinal Observational Study Among Overweight and Obese Adults". *JMIR Diabetes* 2.2, e28.
29 Ibid.
30 Galea, A., Hough, E., and Khan, I. (2017). *Test Beds: The story so far*. NHS England.
31 NHS. *Atrial fibrillation*.
32 England, P. H. (2017b). *Atrial fibrillation prevalence estimates for local populations*.
33 National Institute for Health and Care Excellence (2016). *AliveCor Heart Monitor and AliveECG app (Kardia Mobile) for detecting atrial fibrillation: Guidance and guidelines*.
34 Ibid.
35 NHS Accelerator (2017). *AliveCor Kardia Mobile ECG*.
36 Sanches, J. M. et al. *Detecting Atrial Fibrillation using a Smart Watch—the mRhythm study*.
37 The Primary Care Foundation and the NHS Alliance. *Making Time in General Practice*. Tech. rep.
38 O'Hear, S. (2017b). "Babylon Health partners with UK's NHS to replace telephone helpline with AI-powered chatbot". *TechCrunch*; Armstrong, S. (2018). "The apps attempting to transfer NHS 111 online". *BMJ* 360, k156.
39 O'Hear, S. (2017c). "Your.MD raises $10M to grow AI-driven health information service and marketplace". *TechCrunch*; O'Hear, S. (2017a). *Ada is an AI-powered doctor app and telemedicine service*.
40 Armstrong, "The apps attempting to transfer NHS 111 online".
41 Quer, G. et al. (2017). "Augmenting diagnostic vision with AI". *The Lancet* 390, p. 221.
42 NHS England, *Next steps on the NHS Five Year Forward View*.
43 Giger, M. L. (2018). "Machine Learning in Medical Imaging". *Journal of the American College of Radiology* 15.3, pp. 512–520.

AI-powered diagnosis of cancers, fractures, and cardiovascular, respiratory, and eye diseases.[44]

By speeding up diagnosis, AI can not only save time, but also lives. In conditions like stroke, where each minute untreated increases the extent of brain damage, reducing the time to intervention remains a key challenge for the NHS.[45] Viz.ai's system analyses CT scans to automatically diagnose stroke and uses a smartphone app to alert specialists who can rapidly intervene.[46] With NHS trials reported to begin this year, the aim is to speed up and simplify patient transfers, cutting down on the so-called "door-to-needle" time.[47]

AI can also prevent misdiagnosis, reducing human error. A recent study found that over 50% of vertebral fractures were missed by radiologists at one NHS trust, with a clear majority of these errors made by non-specialist radiologists.[48] To avoid these errors, Zebra Medical Vision has recently developed an algorithm capable of automatically and routinely detecting such fractures.[49] With vertebral fractures costing the UK £1.5 billion, the potential health and economic impact of AI-assisted screening is substantial.[50]

"Cancer Research UK is currently exploring the potential of AI in the early detection of cancer, by taking a machine learning approach to examine patterns of symptoms and behaviours within accessible datasets that could indicate the presence of cancer."—Cancer Research UK[51]

Alongside medical imaging, AI has advanced precision medicine in complex diseases like cancer, enabling the analysis of genomic, molecular, and other data to personalise diagnosis and treatment.[52] With 476 genes, 3701 variants, 65 tumour types, and 97 drugs in one large oncology database, precision medicine would be virtually impossible to implement without AI.[53] For example, IBM's Watson is used to analyse genomic data together with databases of previous patients, clinical trials, and medical literature to determine the best cancer treatment option for a given patient.[54] Likewise, Sophia Genetics, currently deployed in the UK, has developed a system that draws on a growing database of over 180,000 patients across 400 hospitals worldwide to better diagnose cancers and other diseases.[55]

Treatment

In surgical environments, timing is key and resources, such as blood, are costly.[56] Gauss Surgical have developed FDA-approved technology that enables rapid and precise monitoring of blood loss in operating theatres and maternity wards.[57] By allowing earlier detection of haemorrhaging, it reduces the number of required blood transfusions—by at least 50% in one clinical study.[58]

Beyond the surgical room, monitoring patients in A&E and intensive care is a challenge, with conditions rapidly changing and staff juggling multiple patients—a challenge exacerbated by the NHS's staff shortage.[59] To streamline this work, Drayson Technologies, in partnership with the Oxford-based NHS trust, has developed an AI-assisted vital-sign

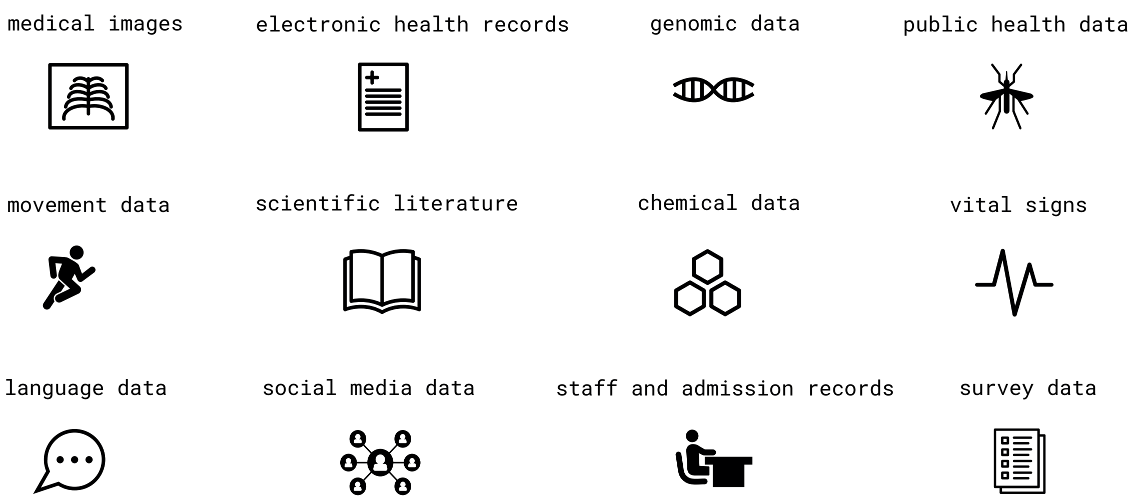
44 Giger, "Machine Learning in Medical Imaging".
45 Saver, J. L. (2006). "Time Is Brain—Quantified". *Stroke* 37.1, pp. 263–266.
46 Viz.ai (2018). *Viz.ai Granted De Novo FDA Clearance for First Artificial Intelligence Triage Software*.
47 Business Leader (2018). "AI start-ups set to change UK healthcare market".
48 Mitchell, R. M. et al. (2017). "Reporting of vertebral fragility fractures: can radiologists help reduce the number of hip fractures?" *Arch Osteoporos* 12.1, p. 71.
49 Business Wire (2017). "Zebra Medical Vision Announces New Algorithm to Detect Compression and Other Vertebral Fractures".
50 Ibid.
51 Cancer Research UK (2017). *Written evidence, Lords Select Committee on Artificial Intelligence*.
52 Bhavnani et al., "2017 Roadmap for Innovation—ACC Health Policy Statement on Healthcare Transformation in the Era of Digital Health, Big Data, and Precision Health: A report of the American College of Cardiology Task Force on Health Policy Statements and Systems of Care".
53 Chakravarty, D. et al. (2017). "OncoKB: A Precision Oncology Knowledge Base". *JCO Precis Oncol* 2017.
54 Rhrissorrakrai, K., Koyama, T., and Parida, L. (2016). "Watson for Genomics: Moving Personalized Medicine Forward". *Trends Cancer* 2.8, pp. 392–395.
55 Sophia Genetics. *https://www.sophiagenetics.com/company/about-us/about-us-details.html*.
56 Stokes Elizabeth A. et al. (2018). "Accurate costs of blood transfusion: a microcosting of administering blood products in the United Kingdom National Health Service". *Transfusion (Paris)* 0.0.
57 Gauss Surgical (2017). *Gauss Surgical Receives FDA Clearance for Second Generation Triton for Real-Time Monitoring of Surgical Blood Loss*.
58 Bernal, N. P. et al. (2017). "Accurate Measurement of Intraoperative Blood Loss during Wound Excision Leads to More Appropriate Transfusion and Reduced Blood Utilization". *Journal of Anesthesia & Clinical Research* 8.11, pp. 1–6.
59 NHS England, *Next steps on the NHS Five Year Forward View*.

monitoring system, which alerts hospital staff when patients deteriorate and allows for rapid observation of patient status.[60]

In mental health, AI is being applied to depression, whose treatment is a trial-and-error process, with up to 70% of patients unresponsive to the first round of antidepressants.[61] Spring Health has developed an algorithm that matches patients with the appropriate antidepressants based on a short questionnaire, in an effort to save the time and costs of ongoing treatment.[62]

AI is also being used as a treatment in itself, with AI-powered chatbots, such as Woebot, delivering smartphone-based Cognitive Behavioural Therapy in a conversational format to those unable to access traditional psychotherapy.[63] Such digital treatment can cheaply and easily scale, supporting the NHS's priority of expanding access to mental health treatment.[64]

Figure 3: Examples of data used in AI-assisted healthcare



Source: online research. Note that this list is not intended to be exhaustive.

Long-term condition management

Managing widespread and costly long-term conditions such as diabetes, dementia, and epilepsy is a key priority for the NHS, with self-care playing a central role.[65] AI is emerging as an enabler of the self-care approach, with tools that can monitor patients, provide guidance, and rapidly alert healthcare professionals as necessary.

Diabetes accounts for 10% of the NHS budget.[66] Alongside prevention, approaches to managing diabetes and avoiding complications are essential.[67] As a potential solution, Glooko has developed a digital diabetes management app which acts as a monitoring and decision support tool for patients and their healthcare team.[68] Early trials have

60 NIHR Oxford Biomedical Research Centre (2017). *Ground-breaking digital health deal agreed with Drayson Technologies.*
61 Al-Harbi, K. S. (2012). "Treatment-resistant depression: therapeutic trends, challenges, and future directions". *Patient Prefer Adherence* 6, pp. 369–388.
62 Chekroud, A. M. et al. (2016). "Cross-trial prediction of treatment outcome in depression: a machine learning approach". *The Lancet Psychiatry* 3.3, pp. 243–250.
63 Fitzpatrick, K. K., Darcy, A., and Vierhile, M. (2017). "Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial". *JMIR Mental Health* 4.2, e19.
64 NHS England, *Next steps on the NHS Five Year Forward View.*
65 Ibid.
66 Diabetes UK (2014). *The Cost of Diabetes.*
67 NHS England, *Next steps on the NHS Five Year Forward View.*
68 Freiherr, G. (2018). *How AI can help patients manage diabetes.*

demonstrated reductions in blood glucose levels, reflecting better management.[69]

For dementia, an ongoing NHS Test Bed is investigating a range of sensors that can automatically monitor patients at home, detect any unexpected events, and alert healthcare professionals, thereby preventing unplanned hospital admissions.[70]

For epilepsy, the NHS is piloting the myCareCentric Epilepsy app and wearable which monitors patients' daily health parameters to automatically detect seizures, alert clinicians, and help patients better manage their condition.[71] Estimates suggest that providing the app across the UK would save the NHS over £250 million by reducing the number of seizure-related deaths and hospital admissions.[72]

Medical research

AI is an indispensable tool for precision medicine initiatives like the UK Biobank (with half a million participants) and the US-based All of Us Research Program (with over 1 million participants), which are gathering and analysing massive amounts of diverse health-related data in an effort to accelerate research and improve health.[73]

AI is also being applied to the entire drug development pipeline in an effort to speed up and optimise the slow and often hit-and-miss process of finding effective therapies.[74] In many cases, existing drugs can be re-purposed for certain conditions, but identifying such matches can be difficult.[75] UK-based BenevolentAI is applying AI to systematically comb through clinical trial data and academic papers to find associations between disease states, drug targets, and drugs in search of promising candidates.[76] In other cases, potential drug targets are known but the right drug does not yet exist, so company Atomwise is applying AI to aid in drug design.[77] Tackling the next part of the pipeline, Antidote is applying AI to help match often hard-to-recruit patients with on-going clinical trials.[78]

"[AI] is also used for a variety of functions across research and development (R&D) including computer-assisted drug design, clinical trial data interpretation and clinical trial simulations such as pharmacological modelling."—The Academy of Medical Sciences[79]

Public health oversight

At a larger scale, AI can be used to manage public health by integrating a variety of data sources. AIME has developed algorithms, currently deployed in Malaysia and Brazil, that analyse public health, weather, and social media data to predict the timing and location of dengue fever outbreaks with 88% accuracy, up to three months in advance.[80]

AI can also be used to optimise public health interventions: analysing social networks on Facebook enables the identification and engagement of homeless youth that would be most influential in spreading awareness about HIV, with one pilot study finding a 25% increase in self-reported HIV testing.[81]

Resource management

---

69 Offringa, R. et al. (2017). "Digital Diabetes Management Application Improves Glycemic Outcomes in People With Type 1 and Type 2 Diabetes". *J Diabetes Sci Technol*.
70 Galea, Hough, and Khan, *Test Beds: The story so far*.
71 Microsoft Reporter (2018). *Tech that helps epilepsy patients 'could save NHS £250m'*.
72 Microsoft Reporter, *Tech that helps epilepsy patients 'could save NHS £250m'*; Dixon, P. A. et al. (2015). "National Audit of Seizure management in Hospitals (NASH): results of the national audit of adult epilepsy in the UK". *BMJ Open* 5.3, e007325.
73 Regalado, A. (2018). "UK Biobank Supercharges Medicine with Gene Data on 500,000 Brits". *MIT Technology Review*; "All of Us Research Program ushers in new era for technology-driven citizen science" (2018). *EurekAlert!*
74 Wein, W. (2016). "Drug development: successes, problems and pitfalls—the industry perspective". *ESMO Open* 1.1.
75 Nosengo, N. (2016). "Can you teach old drugs new tricks?" *Nature News* 534.7607, p. 314.
76 Medeiros, J. (2018). *This AI unicorn is disrupting the pharma industry in a big way*.
77 Shu, C. (2018). "Atomwise, which uses AI to improve drug discovery, raises $45M Series A". *TechCrunch*.
78 Molteni, M. (2018). "Meet the Company Trying to Democratize Clinical Trials With AI". *WIRED*.
79 Academy of Medical Sciences (2017). *Written evidence, Lords Select Committee on Artificial Intelligence*.
80 Startup Daily (2015). *Google Flu Trends may have failed, but MedTech startup AIME has uncovered the secret to predicting viral disease outbreaks*.
81 Yadav, A. et al. (2017). "Influence Maximization in the Field: The Arduous Journey from Emerging to Deployed Application". *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*. AAMAS '17. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems, pp. 150–158.

Funding pressures and increasing demand on the NHS require optimising the use of infrastructure and staff resources to maintain quality of care. By analysing trends in patient intake, outcomes, and staff deployment, AI-powered tools can help the NHS "do more with less".[82]

Companies such as CareSkore have already developed products in this area targeting the US-based private healthcare system.[83] There is ample opportunity to apply similar technology to public healthcare systems. By analysing historical patient data together with holiday and flu patterns, one system trialled in several Paris hospitals could predict surges in admission rates up to 15 days in advance, enabling hospitals to allocate resources accordingly.[84] Similarly, one NHS Test Bed is developing "a demand and capacity dashboard to capture real-time data on patient flow and optimise bed and staff availability", in an effort to relieve pressure on mental health urgent care services.[85]

The way forward

NHS England Chief Executive, Simon Stevens, has stated, "we have a great opportunity to get smarter about the way we are using AI and machine learning with datasets to improve the quality of clinical care."[86] This opportunity, however, presents a set of ethical, social, legal, and technical challenges that should be understood and addressed. In the next sections, we describe these challenges and offer directions forward.

82 Nuffield Trust (2012). *Can NHS hospitals do more with less?*
83 Mannes, J. (2016). *CareSkore gets $4.3M to bring machine learning to preventive care.*
84 Marr, B. (2016). *Big Data In Healthcare: Paris Hospitals Predict Admission Rates Using Machine Learning.*
85 Galea, Hough, and Khan, *Test Beds: The story so far.*
86 Stevens, S. *Speech at Health and Care Innovation Expo 2017.*

2 Data governance

The development and deployment of AI-assisted healthcare will demand more health data to be collected and compiled from an increasing variety of sources, and shared with an increasing array of stakeholders.[87] This calls for a health data governance system that covers all aspects of "data management, data uses, and the technologies derived from it" (see Figure 4).[88] The central challenge for such a system is to be able to harness health data for the collective benefit of society while ensuring ethical data management.[89] Added complexity comes from the UK's public provisioning of healthcare, requiring data governance policy to navigate public-private partnerships with potentially competing interests and diverse stakeholders.[90] Thus, there is a need to develop a data-sharing framework, supported by technology infrastructure, that can advance AI-assisted healthcare ethically, efficiently, and with mutual benefit for all involved.[91]

Figure 4: Dimensions of data governance

who?

Who is accessing the data?
Are they trained and trusted?

where?

Where is data being accessed?
Is it safe and secure?

what?

What data is being accessed?
Is it personal and anonymised?

why?

Why is data being accessed?
What is the expected outcome?

Source: Based on The British Academy and The Royal Society (2017). *Data management and use: Governance in the 21st Century*

The right data governance system can help unlock the value of patient data for the NHS, the public, and industry.[92] IBM has spent billions of dollars acquiring healthcare companies that represent hundreds of millions of patients.[93] The NHS, with millions of cradle-to-grave records, may be no less valuable.[94] Access to health data is a necessary prerequisite for the development of AI-assisted healthcare. An appropriate data governance system can not only facilitate such access, but by instilling trust in patients and healthcare professionals alike, can stimulate further engagement with research and industry programmes that are important for clinical validation of AI products. Lastly, if such a system is implemented nationally, it can enable wider distribution of AI-assisted healthcare, beyond local hubs of digital health innovation.[95]

87 The Royal Society, *Machine learning: the power and promise of computers that learn by example*.
88 The British Academy and The Royal Society (2017). *Data management and use: Governance in the 21st century*.
89 Floridi, L. and Taddeo, M. (2016). "What is data ethics?" *Phil. Trans. R. Soc. A* 374.2083, p. 20160360.
90 Bell, S. J. (2017). *Life Sciences Industrial Strategy*.
91 Mittelstadt, B. and Floridi, L. (2015). *The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts*; Bell, *Life Sciences Industrial Strategy*.
92 Bell, *Life Sciences Industrial Strategy*.
93 Lohr, S. (2016). "IBM Buys Truven for $2.6 Billion, Adding to Trove of Patient Data". *The New York Times*.
94 Bell, *Life Sciences Industrial Strategy*.
95 Ibid.

2.1 Data sharing

For the NHS to realise its value, health data needs to be shared with those capable of transforming it into AI products and insights.[96] The NHS's limited resources entail that industry is central for this development.[97] The *Life Sciences: Industrial Strategy* report (LSIS) positions NHS-industry collaborations as one of the main drivers of a "significant transformation in the way healthcare is delivered in the UK."[98] A streamlined, mutually beneficial, and transparent data sharing framework must underpin such partnerships.[99]

Streamlined data sharing

Current NHS data sharing practices are disjointed, difficult to negotiate, and vary in their terms.[100] While citing existing ambitious data sharing projects such as the UK Biobank, the LSIS nevertheless concludes, "the most significant obstacles for them have been regulatory."[101] Such navigational obstacles may allow resource-rich companies to effectively monopolise the NHS data market, leaving the NHS unable to access a diverse range of innovators and to secure competitive deals.[102] As the UK Government's AI review underscored, there is currently "relatively little widely shared understanding of even the questions that organisations should consider when approaching data-sharing for AI."[103] A streamlined data sharing framework can therefore help level the playing field.

"We have tried to approach the NHS to see if there was a way to access some of this data but we have struggled to even find the right person to talk to."—Matteo Berlucchi, CEO, Your.MD[104]

NHS trusts are likewise unclear about their obligations, as evidenced by the controversial partnership between the Royal Free trust and DeepMind Health.[105] The Trust was found to be in breach of the Data Protection Act for a lack of transparency with patients, a failure to justify the need for the million shared records, and a failure to conduct adequate privacy assessments.[106] A streamlined data sharing framework should prevent such instances and instill confidence in trusts, patients, and industry stakeholders alike.[107]

Mutually beneficial data sharing

Another challenge will be establishing terms that will bring value not only to industry partners, but also to patients, via the NHS.[108] To manage commercial partnerships, the Government has proposed Data Trusts, "a set of relationships underpinned by a repeatable framework, compliant with parties' obligations to share data in a fair, safe and equitable way."[109] As many have emphasised, however, there is currently no consensus about how such a framework, commercial or otherwise, should look.[110]

Part of the challenge is the inherent difficulty of valuating data.[111] Data can be reused in many, often unpredictable, ways, and its value can increase through linkage with other datasets.[112] The British Academy and Royal Society highlight that, unlike with tangible goods, "value is typically derived from the combination and use of data rather than from

96 Bell, *Life Sciences Industrial Strategy*.
97 Huppert J. (2017). *Oral evidence, Lords Select Committee on Artificial Intelligence*.
98 Bell, *Life Sciences Industrial Strategy*.
99 Hall and Pesenti, *Growing the artificial intelligence industry in the UK*.
100 Perrin, N. (2017). *Oral evidence, Lords Select Committee on Artificial Intelligence*.
101 Bell, *Life Sciences Industrial Strategy*.
102 Ibid.
103 Hall and Pesenti, *Growing the artificial intelligence industry in the UK*.
104 Ibid.
105 Information Commissioner's Office (2017b). *Royal Free - Google DeepMind trial failed to comply with data protection law*.
106 Ibid.
107 The British Academy and The Royal Society, *Data management and use*.
108 Bell, *Life Sciences Industrial Strategy*; Hall and Pesenti, *Growing the artificial intelligence industry in the UK*; The British Academy and The Royal Society, *Data management and use*.
109 Hall and Pesenti, *Growing the artificial intelligence industry in the UK*.
110 Severs, M. (2017). *Oral evidence, Lords Select Committee on Artificial Intelligence*; Perrin, N., *Oral evidence, Lords Select Committee on Artificial Intelligence*; Bell, *Life Sciences Industrial Strategy*.
111 The British Academy and The Royal Society, *Data management and use*.
112 Ibid.

individual data points".[113] The value of data, including its outputs, also depends on the stakeholder. How the NHS and the public benefit from data sharing differs from commercial stakeholders. Developing protocols to valuate health data from different stakeholder perspectives will therefore be critical for a data sharing framework.[114]

"Essentially, the development of these algorithms is going to have to be a collaborative effort. In order to build these processes, it is important to develop the appropriate frameworks to support cross-sector data sharing and ensure that there are sufficient incentives on both sides, benefits for all and fairness in how those benefits are distributed."—Sobia Raza, Head of Science, PHG Foundation[115]

Data ownership or control is another source of uncertainty. Who owns or controls health data generated from a hospital visit?[116] Is it the patient, the NHS, or the technology company facilitating data collection? This becomes more complex when patient-generated health data (e.g., from wearables) become integrated into patients' health records. Increased digitisation and public-private data flows are creating ambiguity between the contexts of care provision and commercial development—a "context collapse", as termed by a Wellcome Trust report.[117] AI-assisted healthcare platforms provide care while simultaneously collecting patient data. As a result, patients are "unsure whether [they] are using a service or making a transaction."[118] Understanding how different stakeholders view and benefit from control of data will also be important for developing a mutually beneficial data sharing framework.

Examining and learning from existing agreements is an important next step.[119] For example, AI healthcare company Drayson Technologies has recently signed a profit-sharing agreement with the University of Oxford and Oxford University Hospitals NHS Foundation Trust—where much of the technology has been developed—ensuring that the NHS continues to receive royalties as the technology is licensed for use around the world.[120] Another approach, suggested by the LSIS, may be for the NHS to retain a "golden share" in all enterprises that arise from data-sharing agreements, ensuring that companies remain UK-based.[121] In any approach, it is critical that commercial interests, whether from NHS Trusts or from industry, do not interfere with other principles such as privacy, consent, security, or transparency.

Transparent data sharing

A data sharing framework should also be transparent with the public and other stakeholders about the terms and conditions of potential agreements.[122] The aim is to engender trust and confidence in a public that is currently under-informed and sceptical about the NHS's commercial partnerships.[123] The *care.data* programme should be a lesson learned, with its failure to secure the trust of both patients and healthcare professionals, and to dispel concerns about commercial data access, privacy, and security (see also *Public engagement* below).[124]

"...patients' medical records contain secrets, and we owe them our highest protection. Where we use them—and we have used them, as researchers, for decades without a leak—this must be done safely, accountably, and transparently."—Ben Goldacre, Senior Clinical Research Fellow, University of Oxford[125]

113 The British Academy and The Royal Society, *Data management and use*.
114 Perrin, N., *Oral evidence, Lords Select Committee on Artificial Intelligence*.
115 Raza, S. (2017). *Oral evidence, Lords Select Committee on Artificial Intelligence*.
116 The British Academy and The Royal Society, *Data management and use*; Hunter, P. (2016). "The big health data sale". *EMBO Rep.* 17.8, pp. 1103–1105.
117 Ipsos MORI and Wellcome Trust (2016). *The one-way mirror: Public attitudes to commercial access to health data*.
118 Hunter, "The big health data sale".
119 Hall and Pesenti, *Growing the artificial intelligence industry in the UK*.
120 NIHR Oxford Biomedical Research Centre, *Ground-breaking digital health deal agreed with Drayson Technologies*.
121 Bell, *Life Sciences Industrial Strategy*.
122 The British Academy and The Royal Society, *Data management and use*.
123 Caldicott, F., (2017). *Oral evidence, Lords Select Committee on Artificial Intelligence*; Staa, T.-P. v. et al. (2016). "Big health data: the need to earn public trust". *BMJ* 354, p. i3636.
124 Perrin, N., *Oral evidence, Lords Select Committee on Artificial Intelligence*.
125 Goldacre, B. (2014). *Care.data is in chaos. It breaks my heart*.

Recommendation 1

The NHS, the Centre for Data Ethics and Innovation, and industry and academic partners should conduct a review to understand the obstacles that the NHS and external organisations face around data sharing. They should also develop health data valuation protocols which consider the perspectives of patients, the NHS, commercial organisations, and academia. This work should inform the development of a data sharing framework.

2.2 Consent and opt-outs

Data sharing must be enabled by public support and underpinned by transparency and accountability.[126] The main challenge is finding an approach that maximises ethical data management and gains the public's confidence, while enabling efficient data sharing for the short- and long-term improvement of healthcare delivery.[127] The scenario of patients providing informed consent for each use of data at the relevant time remains impractical due to the lack of public awareness about data governance, the difficulty of predicting data usage, and the logistics of obtaining repeated consent.[128]

As an alternative, the National Data Guardian (NDG) proposed an opt-out approach, which has recently been launched as the NHS *national data opt-out programme*.[129] Patients decide whether they want to opt out of sharing health data for NHS service improvement and healthcare-related research.[130] Evidence suggests that an opt-out approach increases participation rates, thereby increasing the amount and diversity of available health data, both aspects which are essential to high-quality AI development (see also *Standards* section).[131] However, without the necessary patient engagement, transparency, and accountability, this approach risks exploiting patients' lack of awareness or understanding about data sharing.[132] It is essential that these elements are in place for the opt-out programme.

Commercial versus non-commercial data usage

One challenge is how to address the distinction between NHS and commercial data use in the opt-out programme.[133] Development of AI-assisted healthcare will partially depend on private companies, thereby tying this data use to commercialisation.[134] Currently, this distinction is not made.[135] The Wellcome Trust and the Health Foundation highlight the difficulty of distinguishing between commercial and non-commercial uses.[136] Private companies access data for many purposes, including direct care, and research groups often collaborate with commercial organisations (e.g., the publicly funded UK Biobank has an ongoing collaboration with Google).[137]

126 The British Academy and The Royal Society, *Data management and use*; National Data Guardian for Health and Care (2016). *Review of Data Security, Consent and Opt-Outs*.
127 National Data Guardian for Health and Care, *Review of Data Security, Consent and Opt-Outs*.
128 Mittelstadt and Floridi, *The Ethics of Big Data*; National Data Guardian for Health and Care, *Review of Data Security, Consent and Opt-Outs*; House of Commons Science and Technology Committee (2016b). *The Big Data Dilemma*.
129 NHS Digital (2018b). *National data opt-out programme*.
130 Ibid.
131 Chan, T. et al. (2016). "UK National Data Guardian for Health and Care's Review of Data Security: Trust, better security and opt-outs". *Journal of Innovation in Health Informatics* 23.3, pp. 627–632; The Royal Society, *Machine learning: the power and promise of computers that learn by example*; Wellcome Trust (2016). *Response to National Data Guardian for Health and Social Care's Review of Data Security, Consent and Opt-Outs*.
132 The British Academy and The Royal Society, *Data management and use*.
133 The Health Foundation (2016). *Response to National Data Guardian for Health and Social Care's Review of Data Security, Consent and Opt-Outs*.
134 Bell, *Life Sciences Industrial Strategy*; The Royal College of Radiologists (2017). *Written evidence, Lords Select Committee on Artificial Intelligence*.
135 NHS Digital, *National data opt-out programme*.
136 Wellcome Trust, *Response to National Data Guardian for Health and Social Care's Review of Data Security, Consent and Opt-Outs*; The Health Foundation, *Response to National Data Guardian for Health and Social Care's Review of Data Security, Consent and Opt-Outs*.
137 Wellcome Trust, *Response to National Data Guardian for Health and Social Care's Review of Data Security, Consent and Opt-Outs*.

"One of the things that worries members of the public is what use their data might be put to that involves making a profit for somebody other than the health service."—Dame Fiona Caldicott, National Data Guardian for Health and Care[138]

Nevertheless, the NHS is perceived as more trustworthy than for-profit organisations, and there is not enough public understanding of how and why commercial use of data occurs, with at least 17% of the public not wanting their health data used by commercial entities for any reason.[139] These findings suggest the need to consider this distinction more carefully and to increase public understanding of the purposes of data use (see also *Public engagement* below). Rather than distinguishing between commercial and non-commercial uses, the Wellcome Trust advocates for increased transparency about commercial data access to be embedded in the opt-out process.[140]

Levels of data anonymisation

The level of data anonymisation covered by the opt-out programme is another important issue. Currently, it covers only personally identifiable data, such as that including name, address, date of birth, postcode, or NHS number. De-identified or "de-personalised" data (with identifiers removed or encrypted), or fully anonymised data (presented as statistics or trends rather than at the individual level) can be shared regardless of the opt-out (see Figure 5).[141] This enables easier sharing of this data for use by the NHS, academics, and private companies for the purpose of improving healthcare delivery.[142]

Figure 5: Spectrum of identifiability

Source: Adapted from Understanding Patient Data (2017). *Identifiability demystified*.

However, there is ambiguity around how truly "non-personal" de-identified data is, with research projects requiring different levels of anonymisation depending on their aims.[143] Developing powerful AI technology often depends on having rich, well-linked data about each patient—precisely the level of granularity that can enable easy re-identification of patients.[144] Thus, there is a thin line between personal and de-identified data that is important to consider.

138 Caldicott, F., *Oral evidence, Lords Select Committee on Artificial Intelligence*.
139 Ipsos MORI and Wellcome Trust, *The one-way mirror: Public attitudes to commercial access to health data*.
140 Wellcome Trust, *Response to National Data Guardian for Health and Social Care's Review of Data Security, Consent and Opt-Outs*.
141 National Data Guardian for Health and Care, *Review of Data Security, Consent and Opt-Outs*; Understanding Patient Data (2017). *Identifiability demystified*.
142 National Data Guardian for Health and Care, *Review of Data Security, Consent and Opt-Outs*.
143 Wellcome Trust, *Response to National Data Guardian for Health and Social Care's Review of Data Security, Consent and Opt-Outs*; DeepMind (2017b). *Written evidence, Lords Select Committee on Artificial Intelligence*.
144 The Royal Society, *Machine learning: the power and promise of computers that learn by example*.

"We do not think that there is one level of data anonymisation that is 'good enough' for all research problems, as the required level of anonymisation can vary on a project-by-project basis."—Google DeepMind[145]

The Information Commissioner advocates that this issue can be addressed with accountability mechanisms to deter potential re-identification.[146] The General Data Protection Regulation (GDPR) has introduced new regulations and increased sanctions against illegal data re-identification.[147] Transparency about data sharing is another essential component, including "about what happens to de-identified or anonymous data, how it can be used and how it is safeguarded".[148] Monitoring data sharing and the opt-out programme will be important for evaluating whether the dual approach of accountability and transparency is effective at deterring illegal re-identification and gaining the public's confidence.

Recommendation 2

The National Data Guardian and the Department of Health should monitor the NHS data opt-out programme and its approach to transparency and communication, evaluating how the public understands commercial and non-commercial data use and the handling of data at different levels of anonymisation.

Public engagement

Finding the right approach to data sharing ultimately hinges both on understanding what the public finds important about data governance issues such as privacy, consent, and data sharing, and on gaining their confidence through robust transparency and accountability mechanisms.[149] The Royal Society and British Academy highlight that increased knowledge about the collection, sharing, and potential applications of data is linked to more positive attitudes and increased confidence among the public about data sharing.[150]

"Without effective public deliberation, conclusions cannot readily be drawn on public views, particularly about the uses of personal data and the desired benefits of such uses."—The Royal Statistical Society[151]

A public engagement strategy is essential given the currently low awareness and understanding of data sharing related issues.[152] The Wellcome Trust's *Understanding Patient Data* initiative is a valuable effort in this area, with its simplified explainers, workshops, and horizon scanning activities.[153]

145 DeepMind, *Written evidence, Lords Select Committee on Artificial Intelligence*.
146 Wellcome Trust, *Response to National Data Guardian for Health and Social Care's Review of Data Security, Consent and Opt-Outs*; The British Academy and The Royal Society, *Data management and use*; Denham E (2017). *Oral evidence, Lords Select Committee on Artificial Intelligence*.
147 The Royal Society and The British Academy (2017a). *Data Governance: Landscape Review*. Tech. rep.
148 Wellcome Trust, *Response to National Data Guardian for Health and Social Care's Review of Data Security, Consent and Opt-Outs*; The Royal Society and The British Academy, *Data Governance: Landscape Review*.
149 Caldicott, F., *Oral evidence, Lords Select Committee on Artificial Intelligence*.
150 The Royal Society and The British Academy (2017b). *Data governance: public engagement review*. The Royal Society and the British Academy.
151 Royal Statistical Society (2017). *Written evidence, Lords Select Committee on Artificial Intelligence*.
152 Ipsos MORI and Wellcome Trust, *The one-way mirror: Public attitudes to commercial access to health data*; The Royal Society and The British Academy, *Data governance: public engagement review*.
153 Wellcome Trust (2017). *Understanding Patient Data launches today*.

Recommendation 3

The NHS, patient advocacy groups, and commercial organisations should expand public engagement strategies around data governance, including discussions about the value of health data for improving healthcare; public and private sector interactions in the development of AI-assisted healthcare; and the NHS's strategies around data anonymisation, accountability, and commercial partnerships. Findings from this work should inform the development of a data sharing framework.

2.3 Technology

A data governance system also requires the right technology infrastructure, including systems for cybersecurity, data privacy, and accountability (see also *Digital infrastructure* section).[154] This is not only necessary for ensuring privacy and security, but can also help reduce legal and procedural transaction costs embedded in NHS-industry data sharing partnerships, one of the main barriers to sensitive data sharing highlighted by the Government's AI review.[155]

Cybersecurity

A 2016 Care Quality Commission report highlighted the inadequacy of patient data handling and other aspects of cybersecurity across NHS England Trusts.[156] It concluded that daily practices do not reflect cybersecurity standards, with technology failing to meet users' needs and thereby leading to security-compromising workarounds.[157] There is also a lack of leadership with adequate cybersecurity training, and little external validation of cybersecurity systems.[158] The WannaCry hack provided further demonstration of the NHS's vulnerability to cyberattacks.[159] Cybersecurity standards are not new—it is a matter of having adequate infrastructure and ensuring compliance.[160]

Recommendation 4

The NHS Digital Security Operations Centre should ensure that all NHS organisations comply with cybersecurity standards, including having up-to-date technology.

Data privacy and anonymisation

In line with requirements imposed by the GDPR and the Anonymisation Code from the Information Commissioner's Office, data controllers should maintain technical infrastructure that minimises the risk of shared data re-identification.[161] A promising approach is that of *privacy by design*, in which privacy and security are baked directly into the data infrastructure.[162] For example, certain encryption methods can enable data processing by commercial partners to be done directly on encrypted data, without exposing the raw data outside of the NHS.[163] The aim is to find the approach that maximises anonymity and security while enabling streamlined data sharing and processing.

---

154 Bell, *Life Sciences Industrial Strategy*.
155 Hall and Pesenti, *Growing the artificial intelligence industry in the UK*.
156 Care Quality Commission (2016). *Safe data, safe care*.
157 Ibid.
158 Ibid.
159 National Audit Office (2017). *Investigation: WannaCry cyber attack and the NHS*..
160 Care Quality Commission, *Safe data, safe care*.
161 Information Commissioner's Office (2017a). *Big data, artificial intelligence, machine learning and data protection*.
162 Kum, H.-C. and Ahalt, S. (2013). "Privacy-by-Design: Understanding Data Access Models for Secondary Data". *AMIA Jt Summits Transl Sci Proc* 2013, pp. 126–130; Information Commissioner's Office, *Big data, artificial intelligence, machine learning and data protection*.
163 Check Hayden, E. (2015). "Extreme cryptography paves way to personalized medicine". *Nature News* 519.7544, p. 400.

Data auditing and accountability

Also required is technology that can authenticate users and monitor sensitive data sharing within the NHS and its partners, in line with what the Royal Society and British Academy call the "championing of accountability" in data governance.[164] One challenge is deciding which stakeholder to task with data auditing: the NHS, its industry partners, patients, or independent third parties?[165] To that end, distributed ledger technology (e.g., blockchain) has been touted as a solution which allows for multiple parties to independently audit data access without relying on any one mediator.[166] For example, DeepMind is developing the Verifiable Data Audit, a distributed ledger-like system which allows partnering hospitals to see how DeepMind is processing medical data in real time, with entries stating what and why particular data has been accessed.[167]

The GDPR recommends the use of a self-service system for individuals to access their own data.[168] This approach could pave the way for a patient-centred health data auditing system. For example, Estonia, with a fully functioning self-service system, is considering a "personal data market", in which each patient directly engages in data transactions with interested companies, and monitors data access.[169]

Recommendation 5

NHS Digital, the Centre for Data Ethics and Innovation, and the Alan Turing Institute should develop technological approaches to data privacy, auditing, and accountability that could be implemented in the NHS. This should include learning from Global Digital Exemplar trusts in the UK and from international examples such as Estonia.

164 Bell, *Life Sciences Industrial Strategy*; The Royal Society and The British Academy, *Data Governance: Landscape Review*.
165 The British Academy and The Royal Society, *Data management and use*.
166 Government Office for Science. *Distributed ledger technology: beyond block chain*. Tech. rep.
167 DeepMind (2017a). *Trust, confidence and Verifiable Data Audit*.
168 The British Academy and The Royal Society, *Data management and use*.
169 Priisalu, J. and Ottis, R. (2017). "Personal control of privacy and data: Estonian experience". *Health Technol.* 7.4, pp. 441–451.

3 Digital infrastructure

The development and deployment of AI depends on digital data.[170] However, the *Wachter Review* and the *Life Sciences: Industrial Strategy*, among others, have emphasised that there are issues with the NHS's existing data, including with its quality, organisation, and access.[171] Additionally, at a larger scale, there is a need to upgrade information technology (IT) infrastructure to enable integration of more advanced AI-assisted innovation.[172] A key focus should be *interoperability*: the ability of healthcare services and systems to seamlessly exchange and use electronic health information.[173] Urgent and effective digitisation of the NHS at multiple scales is essential for taking advantage of AI-assisted healthcare (see also *Data governance* section).[174]

"Currently, effective IT is not available and this will slow the implementation of AI."—Royal College of Radiologists[175]

3.1 High quality and quantity digital data

Development and validation of AI products often depends on having access to high quantities of digital data to ensure stable algorithm performance. For example, the rapid development of AI-based image recognition was partly enabled by the availability of large, labelled datasets such as ImageNet, which contains over 14 million images.[176]

However, digitisation of the NHS—including its data—varies dramatically across sectors and regions.[177] The primary care sector is almost 100% digitised, successfully managing a system of clinical records, prescriptions, referrals, and appointments.[178] Between GP practices, interoperability of Electronic Health Records is smooth, making transfer of care quick and relatively burden-less.[179]

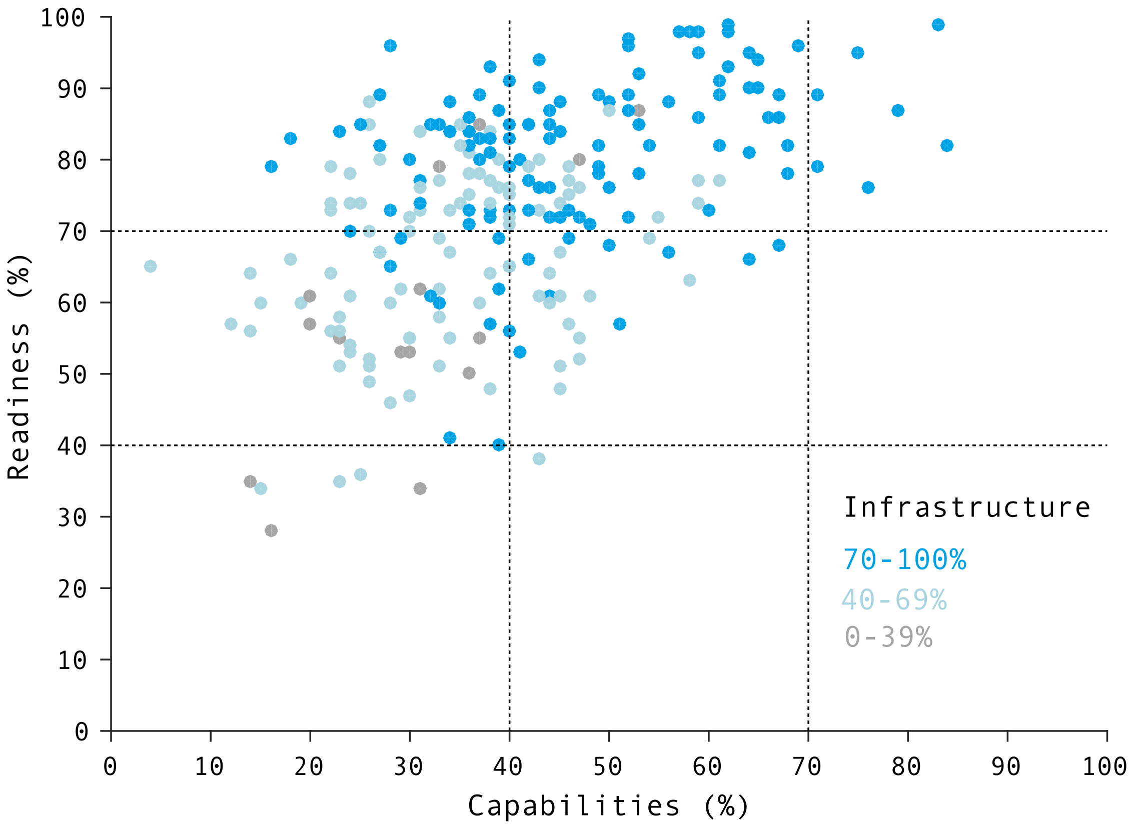
"...the NHS is a fantastic potential resource but is not yet equipped to capitalise on the data it collects."—Wellcome Trust and The Association of Medical Research Charities[180]

In contrast, secondary care lags substantially behind, with many trusts having not yet digitised their clinician notes.[181] The last major attempt to digitise secondary care, launched in 2002, was the costliest IT project ever undertaken in the NHS but failed dramatically to achieve its goals "largely because it was too centralised, failed to engage properly with trusts and their healthcare professionals, and tried to accomplish too much too quickly."[182] Since then, digitisation has been inconsistent. A 2016 *Digital Maturity Assessment*, conducted by NHS England, revealed that over half of the trusts in the assessment had an IT readiness score of below 40%, with only 3% of trusts achieving a readiness score of 70% or higher, the threshold deemed to represent a healthy level of

170 Bughin et al., *Artificial intelligence: the next digital frontier?*
171 Wachter, R. (2016). *Making IT work: harnessing the power of health information technology to improve care in England*. National Advisory Group on Health Information Technology in England; Bell, *Life Sciences Industrial Strategy*.
172 Wachter, *Making IT work: harnessing the power of health information technology to improve care in England*.
173 Ibid.
174 Bell, *Life Sciences Industrial Strategy*.
175 The Royal College of Radiologists, *Written evidence, Lords Select Committee on Artificial Intelligence*.
176 Markoff, J. (2012). "For Web Images, Creating New Technology to Seek and Find". *New York Times*.
177 Wachter, *Making IT work: harnessing the power of health information technology to improve care in England*.
178 Ibid.
179 Ibid.
180 Wellcome Trust and the Association of Medical Research Charities (AMRC) (2017). *Written evidence, Lords Select Committee on Artificial Intelligence*.
181 Wachter, *Making IT work: harnessing the power of health information technology to improve care in England*.
182 Ibid.

digital maturity (see Figure 6).[183] Recent estimates indicate that it will take until 2023 before all trusts are fully digitised.[184] As recently cited in the Commons Science and Technology Committee's report on algorithms in decision-making, "variability in NHS digitisation will mean that some trusts lag behind others in terms of improved healthcare access."[185] For AI development, variability in digitisation not only decreases the quantity of available data, but also limits the representativeness of data and thereby increases the risk of biases in AI algorithms.

Figure 6: 2016 Digital Maturity Assessment of NHS England trusts



Source: NHS England (2016). Each dot is one NHS England Trust. Dashed lines reflect MyNHS bandings. Ideally, all trusts should be blue (above 70% in infrastructure) and in the top right corner (above 70% in capabilities and readiness).

The quality of data also matters for AI performance, including its completeness, consistency, accuracy, representativeness, and linkage, among other aspects (see Figure 7 and the *Standards* section). Ensuring data quality is an ongoing challenge within the NHS.[186] For example, one way that NHS Digital measures data quality across trusts is by tracking how life-long diagnoses, such as autism, are persistently coded in patient episodes, under the assumption that such information should be provided in every record made for these patients.[187] However, data from 2017 indicates that up to 50% of episodes are missing such key information in some conditions.[188]

183 Wachter, *Making IT work: harnessing the power of health information technology to improve care in England*; *NHS England: Digital Maturity Assessment*.
184 Wachter, *Making IT work: harnessing the power of health information technology to improve care in England*.
185 House of Commons Science and Technology Committee, *Algorithms in decision-making*.
186 Wachter, *Making IT work: harnessing the power of health information technology to improve care in England*.
187 NHS Digital (2018a). *Data quality report on comorbidity diagnostic persistence*.
188 Ibid.

Figure 7: Aspects of data quality

Completeness	No values missing, ensuring that all data dimensions are available for use in AI development.
Consistency	Information represented in the same way across clinicians, providers, and systems, making it easy and reliable to combine data across the NHS.
Accuracy	No human or computer errors in information coding or in the medical assessments embedded in the data, minimising the impact that such errors have on AI algorithm performance.
Representativeness	Adequate representation of the target population to which an AI product will be applied, important for minimising biases in AI algorithms.
Linkage	Easily linked to other types of data about the same individual or population to enable richer datasets.

Source: Based on Keller et al. (2017). *The Evolution of Data Quality: Understanding the Transdisciplinary Origins of Data Quality Concepts and Approaches.*

Thus, there is a need to increase the digitisation, quality assurance, and organisation of data that can be used to develop AI algorithms.[189] To enable easy wins, these efforts can be partly guided by assessments of the AI innovation landscape, including the readiness of AI applications for deployment.[190] For example, AI-assisted medical imaging is closer to deployment than applications relying on Electronic Health Records, and medical images are already stored in a standard digital database, the Picture Archive and Communication System (PACS).[191] In line with this, the Royal College of Radiologists has suggested to focus quality assurance efforts on normal x-ray images in the PACS, which can be used to develop algorithms that refer radiologists only to abnormal images for further assessment.[192] In all cases, it is important that such targeted efforts ensure future interoperability of digitised data.

3.2 Broader IT infrastructure

To maximise the benefits of AI and enable future development, broader IT infrastructure also needs upgrading. For example, recent advances in wearable and environmental sensor technology can provide novel sources of data on patients, staff, and infrastructure, furthering the development of AI-assisted monitoring of health and NHS resources.[193] As highlighted by the *Wachter Review*, one of the key aims of digital infrastructure upgrades should be interoperability.[194] A patient's GP record should be easily linked to their medical images, consultants' notes, and any smartphone or wearable health data, no matter which hospital they visit across the country.[195]

"Without clear guidance at a national level for both interoperability and data access, enabling appropriate and controlled access for research to representative and joined-up datasets, the full potential for UK data to improve health and care will not be realised."—Professor Sir John Bell, Life Sciences Industrial Strategy 2017[196]

For developing AI products, interoperability enables the building of large and rich training datasets comprising many patients across hospitals or many sources of data about

189 Bell, *Life Sciences Industrial Strategy*.
190 PwC (2017). *What doctor? Why AI and robotics will define New Health*. PwC.
191 Bughin et al., *Artificial intelligence: the next digital frontier?*
192 The Royal College of Radiologists, *Written evidence, Lords Select Committee on Artificial Intelligence*.
193 Bughin et al., *Artificial intelligence: the next digital frontier?*
194 Wachter, *Making IT work: harnessing the power of health information technology to improve care in England*.
195 Bughin et al., *Artificial intelligence: the next digital frontier?*
196 Bell, *Life Sciences Industrial Strategy*.

individual patients.[197] For example, *Sophia AI*, a platform that provides AI-assisted diagnostics based on genomic data, operates through a cloud-based system which takes advantage of all data available to it across hospitals in the UK and around the world.[198] The more institutions use the platform and provide it with data, the more powerful its algorithms become.[199] During deployment of AI products, interoperability also enables patients to enjoy seamless healthcare services across hospitals, and between primary and secondary care.[200] Estonia's healthcare system, for example, has patient data integrated not only between GPs and hospitals, but also with emergency services, wherever they are.[201]

"The goal is not digitisation for digitisation's sake, but rather to improve the way care is delivered in the NHS, in part by using digital tools."—Professor Robert Wachter, University of California, San Francisco[202]

Efforts to upgrade infrastructure may benefit from identifying low-hanging fruit, such as cloud-based platforms that are easily accessible with only basic infrastructure. An additional approach is to learn from more advanced healthcare environments, as being done in the Global Digital Exemplars programme in which successfully digitised trusts serve as reference sites and partners for other trusts.[203] Further, input from healthcare professionals and patients, as domain experts and users, is crucial in the development of a digital NHS.[204] Digitisation of health services should not be presented to end users and stakeholders as an exercise purely being done "for its own sake", but instead should focus on the tangible benefits of increased digitisation, such as the use of AI to improve productivity.[205] Likewise, advanced digitisation should not come at the expense of basic needs in healthcare delivery.

"Our intention is that, in the future, hospitals won't merely choose an IT vendor, they will choose a hospital that they want to partner with and implement the same system, keeping the IT 80% the same and making only the 20% of changes that are absolutely necessary to meet local needs."—NHS England[206]

Recommendation 6

The NHS should continue to increase the quantity, quality, and diversity of digital health data across trusts. It should consider targeted projects, in partnership with professional medical bodies, that quality-assure and curate datasets for more deployment-ready AI technology. It should also continue to develop its broader IT infrastructure, focusing on interoperability between sectors, institutions, and technologies, and including the end users as central stakeholders.

197 Bell, *Life Sciences Industrial Strategy*.
198 Burgess, M. *The Science Museum's Robots exhibition gives some unsettling truths about humanity*.
199 Magistretti, B. *Swiss data analytics company Sophia Genetics could be Switzerland's next unicorn*.
200 Bughin et al., *Artificial intelligence: the next digital frontier?*
201 Priisalu and Ottis, "Personal control of privacy and data".
202 Wachter, *Making IT work: harnessing the power of health information technology to improve care in England*.
203 England, N. (2017a). *NHS England: Global Digital Exemplars*.
204 Wachter, *Making IT work: harnessing the power of health information technology to improve care in England*.
205 Ibid.
206 NHS England, *Next steps on the NHS Five Year Forward View*.

4 Standards

Amid the increase in popularity of AI-assisted healthcare, there is a need for validation of AI products to ensure their safety and efficacy in clinical settings.[207] Diverse stakeholders have raised concerns about the potential quality of AI products. The Institute of Electrical and Electronics Engineers has warned about the "gap between how AI/AS [artificial intelligence / autonomous systems] is marketed and their actual performance, or application".[208] Michael Osborne, professor of machine learning at the University of Oxford, says that "we are not where we would want to be in ensuring that the algorithms we deliver are completely verifiable and validated".[209] Similarly, the Academies of Medical Sciences and Medical Royal Colleges have called for increased validation of AI products.[210] The transformative potential of the technology should not by hampered by poor quality standards.

"There is a gap between the validation of algorithms and the validation of their implementation clinically."—Royal College of Radiologists[211]

These concerns are not without merit. For instance, one smartphone app claiming to detect cancer among skin lesions was taken to court in 2015 by the US Federal Trade Commission for failing to provide evidence for its claims.[212] A study has found that 3 out of 4 smartphone apps for skin cancer detection incorrectly classify almost a third of melanomas as unconcerning.[213] The accuracy of patient triage apps has also been challenged, with one study finding that such apps provided the appropriate recommendation just over half the time.[214] To be clear, there are important differences between companies in their approaches, with some taking time to publish studies and conduct real-world trials.[215] Overall, however, the trend appears to be that many claims are made without the support of necessary evidence.[216] As a *Nature* editorial points out, "Many reports of new AI diagnostic tools, for example, go no further than preprints or claims on websites."[217]

"Strong standards for auditing and understanding the use of AI systems 'in the wild' are urgently needed."—AI Now Institute[218]

The overarching aim of innovation is to improve existing healthcare delivery in terms of safety, efficacy, productivity, or human factors. To this end, the Medicines and Healthcare Products Regulatory Agency has called for the development of validation standards in addition to a clear regulatory framework to ensure the safety and efficacy of AI-assisted healthcare.[219] Achieving this requires addressing the specifics of AI technology within a healthcare context—issues including bias and training data, performance evaluation, and interpretability.

207 JASON, *Artificial Intelligence for Health and Health Care*.
208 The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems (2016). *Ethically Aligned Design: A Vision For Prioritizing Wellbeing With Artificial Intelligence And Autonomous Systems, Version 1*. IEEE.
209 House of Commons Science and Technology Committee (2016a). *Robotics and artificial intelligence*.
210 Royal Medical Colleges, A. of (2017). *Information and Digital Technologies Clinical: Requirements 2020*. Academy of Royal Medical Colleges; Academy of Medical Sciences, *Written evidence, Lords Select Committee on Artificial Intelligence*.
211 The Royal College of Radiologists, *Written evidence, Lords Select Committee on Artificial Intelligence*.
212 Federal Trade Commission (2015). ""Melanoma Detection" App Sellers Barred from Making Deceptive Health Claims". *Federal Trade Commission*.
213 Wolf, J. A. et al. (2013). "Diagnostic Inaccuracy of Smartphone Applications for Melanoma Detection". *JAMA Dermatol* 149.4, pp. 422–426.
214 Semigran, H. L. et al. (2015). "Evaluation of symptom checkers for self diagnosis and triage: audit study". *BMJ* 351, h3480.
215 JASON, *Artificial Intelligence for Health and Health Care*.
216 Nature (2018). "AI diagnostics need attention". *Nature* 555.7696, p. 285.
217 Ibid.
218 Campolo, A. et al. (2017). *AI Now 2017 Report*. AI Now Institute.
219 Medicines and Healthcare Products Regulatory Agency (2017b). *Written evidence, Lords Select Committee on Artificial Intelligence*.

"Standards is one of the factors that can accelerate the development of artificial intelligence while helping build trust in the technology and promoting public acceptance."—British Standards Institution[220]

4.1 Training data

Most AI products today fundamentally consist of algorithms and *training data*, the information that algorithms use to learn predictive relationships.[221] Training data define how the algorithm and product will perform with new data when deployed in the real world.[222] Validating AI products includes understanding the quantity and quality of training data used, including how it was created and processed, how well it aligns with the product's definition of intended use, and how it relates to the intended real-world population.[223]

"If AI systems are to be regulated then the training/input data utilised is integral to the system as a whole. This is especially true in the heterogeneous, 'big data' medical research field."—Medicines and Healthcare Products Regulatory Agency[224]

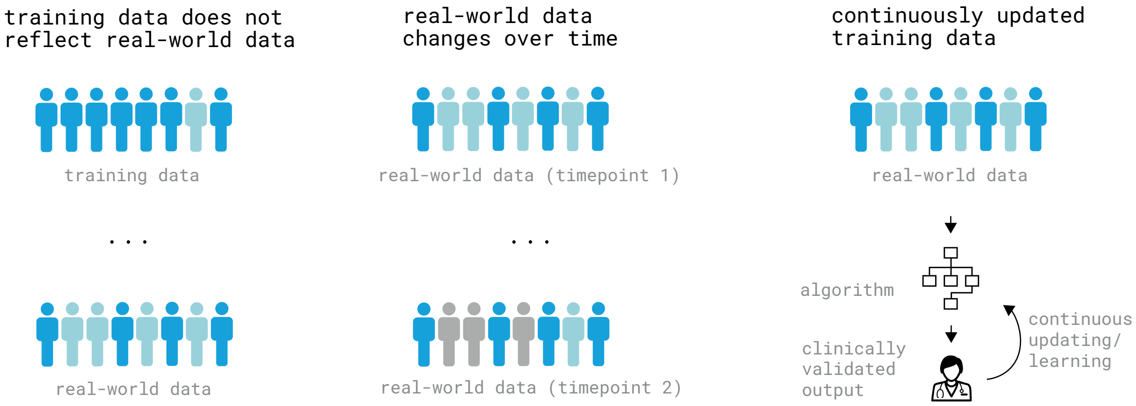
Without such validation, there is a risk of inadequate or skewed performance during deployment.[225] For example, biases present in the training data can lead to performance varying between populations, as was the case with policing algorithms that were found to discriminate against Black Americans.[226] Biases can be difficult to detect and can arise for many reasons, including an imbalance in the amount of data between populations, or in the way such data is labelled.[227]

"If someone is trying to sell you a black box system for medical decision support, and you don't know how it works or what data was used to train it, then I wouldn't trust it."—John Giannandrea, Head of AI, Google[228]

These concerns are pertinent to healthcare. For example, one study found that certain cardiovascular risk factors developed predominantly using White population data led to biased results in non-White individuals.[229] In other cases, issues can arise when the real-world data with which the product works changes over time (e.g., if the quality of medical image data changes over time).[230] AI products can also be designed to adaptively learn from real-world data during deployment, leading to continuous changes in the algorithms. This is a powerful approach, but one that requires careful management to avoid biases in the real-world data (e.g., hospitals that adopt AI innovation may disproportionately serve certain demographics).[231] All of these scenarios can lead to performance that—without appropriate validation—may not be expected during development (see also Figure 8). Considering the issues around training data in the specific context of healthcare is essential. For example, eliminating the influence of race and ethnicity may be desirable in the justice system, but can lead to sub-optimal performance in healthcare, where such information can be vital for appropriate decisions.[232]

220 British Standards Institution (2017). *Written evidence, Lords Select Committee on Artificial Intelligence.*
221 Jordan and Mitchell, "Machine learning".
222 Ibid.
223 Academy of Medical Sciences, *Written evidence, Lords Select Committee on Artificial Intelligence*; Clinical Decision Support Coalition (2017). *Voluntary Industry Guidelines for the Design of Medium Risk Clinical Decision Support Software*; Medicines and Healthcare Products Regulatory Agency, *Written evidence, Lords Select Committee on Artificial Intelligence.*
224 Medicines and Healthcare Products Regulatory Agency, *Written evidence, Lords Select Committee on Artificial Intelligence.*
225 JASON, *Artificial Intelligence for Health and Health Care.*
226 Kirchner, L. et al. (2016). "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And it's Biased Against Blacks." *ProPublica.*
227 Mittelstadt, B. D. et al. (2016). "The ethics of algorithms: Mapping the debate". *Big Data & Society* 3.2, p. 2053951716679679.
228 Hanly, K. (2017). "Google AI chief claims biased algorithms are a big danger". *Digital Journal.*
229 Gijsberts, C. M. et al. (2015). "Race/Ethnic Differences in the Associations of the Framingham Risk Factors with Carotid IMT and Cardiovascular Events". *PLoS One* 10.7, e0132321.
230 The Royal Society, *Machine learning: the power and promise of computers that learn by example.*
231 Ibid.
232 Char, D. S., Shah, N. H., and Magnus, D. (2018). "Implementing Machine Learning in Health Care — Addressing Ethical Challenges". *N. Engl. J. Med.* 378.11, pp. 981–983.

Figure 8: Potential scenarios involving training data



In the first scenario, the training data does not represent the intended real-world population. In the second, the training data does initially represent the real-world population, but the real-world data changes over time. In the third scenario, the algorithm is continuously updated (trained) based on real-world data. Note that this is not intended to be an exhaustive list of scenarios. Based on The Royal Society (2017). *Machine learning: the power and promise of computers that learn by example*.

Defining the boundaries in which an AI product should work as expected should be a primary requirement of quality assurance, as is common to many technologies and sectors.[233] Others have pointed out that rigorous testing is especially important with the use of "black box" algorithms, in which an intuitive understanding of their decision-making is obscured.[234] In the case of algorithms that learn continuously, regular post-market validation may be necessary to ensure their continued safety, though how this will be implemented is an open question.[235]

Recommendation 7

The Alan Turing Institute, the Ada Lovelace Institute, and academic and industry partners in medicine and AI should develop ethical frameworks and technological approaches for the validation of training data in the healthcare sector, including methods to minimise performance biases and validate continuously-learning algorithms.

4.2 Performance evaluation

How AI product performance is evaluated is another important issue that is closely related to the discussion around training data.[236] A variety of evaluation metrics are available, each of which reveals different aspects of performance, with no agreed upon standards for deciding between them.[237] For example, common to medical diagnosis are the measures of "sensitivity" (e.g., fraction of sick people correctly identified as sick) and "specificity" (e.g., fraction of healthy people correctly identified as healthy). Other measures can provide probability information alongside simple yes/no answers (e.g., "a malignant tumour with

233 International Medical Device Regulators Forum (2017). *Software as a Medical Device (SaMD): Clinical Evaluation.*
234 Mittelstadt et al., "The ethics of algorithms".
235 International Medical Device Regulators Forum, *Software as a Medical Device (SaMD): Clinical Evaluation.*
236 Ibid.
237 Amarasingham, R. et al. (2016). "Consensus Statement on Electronic Health Predictive Analytics: A Guiding Framework to Address Challenges". *EGEMS (Wash DC)* 4.1.

95% probability").[238]

"As a field, we should be aware of the dangers of convincing ourselves that we have solved a particular problem based on evidence provided by generic metrics that, while persuasive to a [machine learning] colleague, is insufficient for a domain expert."—Cynthia Rudin, Associate Professor, Duke University, and Kiri Wagstaff, Jet Propulsion Laboratory, California Institute of Technology[239]

Some research involves comparisons between algorithm and human performance, an approach which can questionable or complex.[240] For example, there are differences in the time and information available to algorithms and humans.[241] It is also unclear which standard of human performance is desired for comparison. For example, is the aim for AI products to be better than or as good as clinicians? Should comparisons be made to the best clinicians or the average clinician?

Evaluation of a given product also depends on the context, including a product's intended use and the amount of human oversight required.[242] For example, a probabilistic output that requires time to consider may not be a useful measure for a product intended to be used in fast-paced surgical settings—the aim of AI, as with all healthcare innovation, should always be to assist healthcare professionals in delivering care. Similarly, in some contexts, human interpretability of algorithms' decision-making may be as equally important as accuracy (see also below).[243]

Recommendation 8

The Alan Turing Institute, the Ada Lovelace Institute, and academic and industry partners in medicine and AI should develop standardised approaches for evaluating product performance in the healthcare sector, with consideration for existing human performance standards and products' intended use.

4.3 Interpretability

Algorithm interpretability is another important issue, defined here as *the ability of the user to understand an algorithm's decision-making process to appropriately evaluate its output for a healthcare decision*.[244] Sometimes referred to as "transparency" or "explainability", we distinguish interpretability from simply revealing a product's underlying code, which may be relevant from a software engineering or other regulatory standpoint, but which is not informative within a healthcare context.[245]

Algorithm interpretability is critical because it defines the user's dependence on the product, determining how much they can reasonably intervene in the healthcare decision. As such, some have argued that interpretability determines how much other product validation is needed to ensure safety and efficacy.[246] This has been the principle behind the voluntary industry guidelines developed by the US-based Clinical Decision Support Coalition (CDSC).[247] Representing industry and healthcare professionals, among others,

238 Eddy, D. M. et al. (2012). "Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–7". *Value Health* 15.6, pp. 843–850.
239 Rudin, C. and Wagstaff, K. L. (2014). "Machine learning for science and society". *Mach Learn* 95.1, pp. 1–9.
240 Bejnordi, B. E. et al. (2017). "Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer". *JAMA* 318.22, pp. 2199–2210.
241 Nature, "AI diagnostics need attention".
242 International Medical Device Regulators Forum (2016). *IMDRF Proposed Document: Software as a Medical Device (SaMD): Clinical Evaluation*.
243 Lipton, Z. C. (2016). "The Mythos of Model Interpretability". *arXiv:1606.03490 [cs, stat]*; Doshi-Velez, F. and Kim, B. (2017). "Towards A Rigorous Science of Interpretable Machine Learning". *arXiv:1702.08608 [cs, stat]*.
244 Clinical Decision Support Coalition, *Voluntary Industry Guidelines for the Design of Medium Risk Clinical Decision Support Software*; Doshi-Velez and Kim, "Towards A Rigorous Science of Interpretable Machine Learning".
245 Lipton, "The Mythos of Model Interpretability".
246 International Medical Device Regulators Forum, *IMDRF Proposed Document: Software as a Medical Device (SaMD): Clinical Evaluation*.
247 Clinical Decision Support Coalition, *Voluntary Industry Guidelines for the Design of Medium Risk Clinical Decision Support Software*.

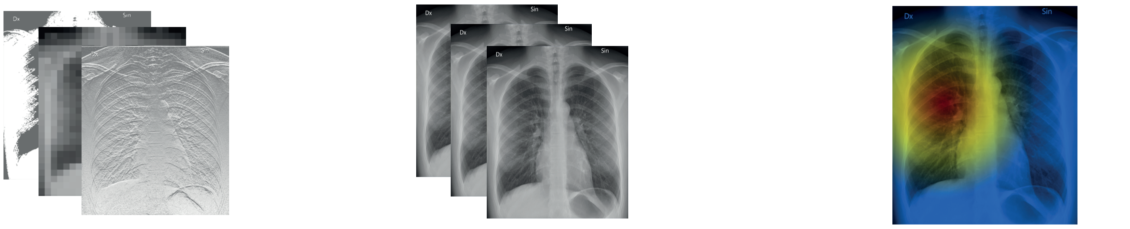
the CDSC states that "taking over, in any substantial way, the healthcare decision-making carries with it heightened responsibility for validation"—less interpretability may require more performance validation.[248] Algorithm interpretability, therefore, plays a key role in defining the interdependence between product and user, bearing not only on standards and regulations but also on legal liability (see *Legal liability* section).[249]

"The transparency of information on limitations with algorithms, clinical model, quality of data used to build the models, assumptions made, etc. can help users question the validity of output of the SaMD [software as a medical device] and avoid making incorrect or poor decisions."—International Medical Device Regulators Forum[250]

Interpretability has been a challenge in the AI field, with algorithms becoming increasingly more sophisticated and powerful, while approaches to informing users about how individual decisions are reached have lagged behind.[251] Current approaches include visualising algorithms' internal components, providing explanatory examples, or highlighting which aspects of the input data contributed to the decision (see Figure 9).[252] In parallel, there is a need to develop methods to *evaluate* interpretability, especially in real-world contexts.[253]

Figure 9: Approaches to interpretability

Visualisation | Explanation by example | Local explanation



"This is how the algorithm analysed the image." | "These other example images have led to the same decision." | "These areas of the image are most informative for the decision."

Cartoon illustration of three approaches being taken to enable interpretability, presented here in the context of a hypothetical algorithm designed to detect a health condition in an x-ray image. This is not an exhaustive description of approaches. Based on Lipton (2016). *The mythos of model interpretability.*

Part of the challenge of interpretability is its ambiguity, with its definition depending on the algorithm in question and the context in which it is applied, including the user's end goal and expertise.[254] The interpretability that a patient at home requires is different from that of a surgeon in an operating theatre. As such, developing it requires understanding the workflow and decision-making of healthcare professionals.[255] To this end, the American College of Radiology's Data Science Institute is bringing together multidisciplinary stakeholders, including AI researchers and radiographers, to develop AI-assisted healthcare that is directly informed by a clinical perspective.[256]

248 Clinical Decision Support Coalition, *Voluntary Industry Guidelines for the Design of Medium Risk Clinical Decision Support Software.*
249 The Royal College of Radiologists, *Written evidence, Lords Select Committee on Artificial Intelligence.*
250 International Medical Device Regulators Forum, *IMDRF Proposed Document: Software as a Medical Device (SaMD): Clinical Evaluation.*
251 Quer et al., "Augmenting diagnostic vision with AI".
252 Lipton, "The Mythos of Model Interpretability".
253 Doshi-Velez and Kim, "Towards A Rigorous Science of Interpretable Machine Learning".
254 Lipton, "The Mythos of Model Interpretability"; Doshi-Velez and Kim, "Towards A Rigorous Science of Interpretable Machine Learning".
255 Lipton, Z. C. (2017). "The Doctor Just Won't Accept That!" *arXiv:1711.08037 [stat].*
256 American College of Radiology Data Science Institute (2017). *The ACR Data Science Institute Structures Artificial Intelligence Development to Optimize Radiology Care.*

Recommendation 9

The Alan Turing Institute, the Ada Lovelace Institute, and academic and industry partners in medicine and AI should develop methods of enabling and evaluating algorithm interpretability in the healthcare sector. This work should involve experts in AI, medicine, ethics, usability design, cognitive science, and ethnography, among others.

4.4 Usability

Product usability—the ease and efficiency of a product's use and how it integrates into the clinical workflow—is another important determinant of AI product performance in the real world.[257] Though usability is not an aspect unique to AI, it is important that the novelty of AI products does not distract from poor design, as has happened in the last major digitisation attempt in the NHS.[258] Guidelines by the Medicines and Healthcare Products Regulatory Agency emphasise that poor usability design frustrates and impedes users, affects patient safety, and decreases the quality of any data collected.[259]

"Simply put, if usability is lacking, the completion of user tasks may be slower and more error-prone. Therefore, delivery of therapy will suffer and patient safety may be compromised."—Bob North, Human Centered Strategies[260]

Usability design is particularly important in AI-assisted healthcare because of what the International Medical Device Regulators Forum calls "the uniqueness of indirect contact between patients and SaMD [software as a medical device]".[261] An AI product's impact on patient health is often through the human interpretation of the product's output, such as a diagnostic recommendation, rather than through direct contact with the body. Usability design is therefore at the nexus between AI products and their users, and is closely related to algorithm interpretability.[262]

Usability design involves consideration of the product's interface, as well as the user's environment and expertise.[263] For example, designing for patients is different than for clinicians. Effective usability design often involves the users in the process and requires dedicated testing in real-world environments.[264]

Recommendation 10

Developers of AI products and NHS Commissioners should ensure that usability design remains a top priority in their respective development and procurement of AI-assisted healthcare products.

257 North, B. (2015). *The growing role of human factors and usability engineering for medical devices.* British Standards Institution.
258 Wachter, *Making IT work: harnessing the power of health information technology to improve care in England.*
259 Medicines and Healthcare Products Regulatory Agency (2017a). *Guidance on applying human factors to medical devices.*
260 North, *The growing role of human factors and usability engineering for medical devices.*
261 International Medical Device Regulators Forum, *IMDRF Proposed Document: Software as a Medical Device (SaMD): Clinical Evaluation.*
262 North, *The growing role of human factors and usability engineering for medical devices.*
263 Ibid.
264 North, *The growing role of human factors and usability engineering for medical devices*; Medicines and Healthcare Products Regulatory Agency, *Guidance on applying human factors to medical devices.*

5 Regulations

In addition to developing technical standards for AI product validation, it is important to find ways to implement such standards to ensure their uptake. This could involve publishing them as guidelines, implementing them as part of a regulatory framework, or a combination thereof. AI products have been rightly recognised as medical devices by the Medicines Healthcare Products Regulatory Agency (MHRA) and the European Commission.[265] However, there is now a need to complement these regulatory efforts with guidelines that can enable innovators to address questions concerning AI product classification, validation, and monitoring.[266] Moreover, given the relative novelty and the rapidly evolving research around AI-assisted healthcare, it is important to find the right approaches, regulatory or otherwise, which simultaneously protect patient safety, enable innovation, and promote industry uptake of standards.[267]

5.1 Regulatory guidelines

The MHRA currently considers AI products as medical devices if they perform functions such as prevention of disease, monitoring, diagnosis, or treatment.[268] Recently, the European Commission has enacted new legislation targeting digital health products, which comes into full effect in 2020.[269] This legislation introduces a new classification system for AI products which considers both the intended purpose and the overall risk assessment of such products (see Figure 10), along with increased clinical evaluation requirements and post-market surveillance, among other changes.[270] However, this legislation has not yet been accompanied by guidelines that can enable industry to confidently meet the requirements and address AI-specific concerns.

265 Medicines and Healthcare Products Regulatory Agency (2014). *Medical devices: software applications (apps)*; European Commission. *Revisions of Medical Device Directives*.
266 Academy of Medical Sciences, *Written evidence, Lords Select Committee on Artificial Intelligence*; The Royal College of Radiologists, *Written evidence, Lords Select Committee on Artificial Intelligence*.
267 Hall and Pesenti, *Growing the artificial intelligence industry in the UK*.
268 Medicines and Healthcare Products Regulatory Agency, *Medical devices*.
269 European Commission, *Revisions of Medical Device Directives*.
270 Ibid.

Figure 10: EU "Software as a Medical Device" Risk Classification

Class III	Software intended to provide information which is used to take decisions with diagnosis or therapeutic purposes if such decisions have an impact *that may cause death or an irreversible deterioration of a person's state of health.*
Class IIb	Software intended to provide information which is used to take decisions with diagnosis or therapeutic purposes if such decisions have an impact *that may cause a serious deterioration of a person's state of health or a surgical intervention.* Or, software intended to monitor physiological processes if it is intended *for monitoring of vital physiological parameters, where the nature of variations of those parameters is such that it could result in immediate danger to the patient.*
Class IIa	Software not falling into above categories but that is intended to provide information which is used to take decisions with diagnosis or therapeutic purposes, or software intended to monitor physiological processes.
Class I	All other software.

All classes except I will require certification by a notified body such as the British Standards Institution. Class I products will require self-declaration by the manufacturer.

Novel EU device classication which considers the intended use and risk associated with a device. Source: Adapted from the European Commission (2017).

The International Medical Device Regulators Forum (IMDRF), which includes representatives from EU and the MHRA, has also commented on clinical evaluation of AI products (see Figure 11).[271] They have gone the furthest at emphasising the role of training data, performance evaluation, and algorithm interpretability.[272] However, they remain ambiguous about whether and how these aspects should be evaluated for each class of products, and, confusingly, their risk classification scheme does not directly map onto the EU's own.[273]

For example, both the new EU regulations and the IMDRF guidelines mention the review of scientific literature as a source of clinical evidence.[274] However, this may prove to be inadequate for many AI products, given that a substantial amount of work on AI-assisted healthcare presents a "proof of principle", without demonstration of real-world performance.[275] This calls for consideration of alternative approaches that can ensure clinically meaningful performance of AI-assisted healthcare products.

271 International Medical Device Regulators Forum, *IMDRF Proposed Document: Software as a Medical Device (SaMD): Clinical Evaluation.*
272 Ibid.
273 Clinical Decision Support Coalition, *Voluntary Industry Guidelines for the Design of Medium Risk Clinical Decision Support Software*; International Medical Device Regulators Forum, *IMDRF Proposed Document: Software as a Medical Device (SaMD): Clinical Evaluation*; European Commission (2017). "Regulation (EU) 2017/745". *Official Journal of the European Union* 60.
274 European Commission, "Regulation (EU) 2017/745"; International Medical Device Regulators Forum, *Software as a Medical Device (SaMD): Clinical Evaluation.*
275 Nature, "AI diagnostics need attention".

Figure 11: IMDRF's clinical evaluation framework for AI products

Valid clinical association	Is there a valid clinical association between your AI product output and your AI product's targeted clinical condition?
Analytical validation	Does your AI product correctly process input data to generate accurate, reliable, and precise output data?
Clinical validation	Does use of your AI product's accurate, reliable, and precise output data achieve your intended purpose in your target population in the context of clinical care?

Source: Adapted from the International Medical Device Forum (2017).

Another example concerns the use of data acquired during real-world deployment, termed *post-market surveillance data*.[276] The EU and IMDRF emphasise the value of such data for validating AI product performance.[277] The IMDRF suggests that post-market data may even be used to change products' risk classification as it becomes necessary.[278] Additionally, it may be one way to ensure regular validation of continuously-learning algorithms.[279] Although a potentially powerful tool, it requires careful consideration of how to best integrate it into the UK's regulatory framework to manage patient safety.

"The biggest challenge will be in adapting regulation to address the individual features of fast changing AI algorithms. This is important because, while there are many potential healthcare benefits from AI, these technologies are not without considerable potential risks."—Medicines and Healthcare Products Regulatory Agency[280]

As such, there is a need to develop regulatory guidelines around how AI-assisted healthcare products should be classified, validated, and monitored.[281] To lead this development, the UK would benefit from the establishment of a unit within the MHRA that is dedicated to digital health. Such a unit should work together with manufacturers, notified bodies such as the British Standards Institution, healthcare bodies such the NHS and the National Institute for Health and Care Excellence, and AI-related bodies such as the Alan Turing Institute and the Centre for Data Ethics and Innovation. It should also closely work with the IMDRF and the European Commission, as the MHRA currently does, to build on existing working guidelines and ensure international harmonisation.

The US Food and Drug Administration (FDA) has already taken a similar approach with its new Digital Health program, stocked with experts on AI, and those with "hands-on development experience with a [digital health] product's full life cycle".[282] It aims to focus regulatory work on digital health into one core unit, avoiding the fragmentation of expertise across healthcare areas.[283]

276 International Medical Device Regulators Forum, *Software as a Medical Device (SaMD): Clinical Evaluation.*
277 European Commission, "Regulation (EU) 2017/745"; International Medical Device Regulators Forum, *Software as a Medical Device (SaMD): Clinical Evaluation.*
278 International Medical Device Regulators Forum, *Software as a Medical Device (SaMD): Clinical Evaluation.*
279 Medicines and Healthcare Products Regulatory Agency, *Written evidence, Lords Select Committee on Artificial Intelligence.*
280 Ibid.
281 Ibid.
282 FDA Center for Devices and Radiological Health (2017a). *Digital Health Innovation Action Plan.*
283 Ibid.

Recommendation 11

The Medicines and Healthcare Products Regulatory Agency should establish a digital health unit with expertise in AI and digital products that will work together with manufacturers, healthcare bodies, notified bodies, AI-related organisations, and international forums to advance clear regulatory approaches and guidelines around AI product classification, validation, and monitoring. This should address issues including training data and biases, performance evaluation, algorithm interpretability, and usability.

5.2 Regulatory frameworks

The new EU legislation is a significant step forward, but simultaneously increases the burden on both manufacturers and regulators. If implemented inappropriately, it risks stifling AI-assisted innovation of healthcare.[284] Although the interdependency between regulation and innovation makes it difficult to tease apart cause and effect, evidence suggests that regulations of emerging areas should be flexible, incentive-based, and underpinned by efficient implementation.[285] For example, the commercial development of genetically engineered microorganisms in the 1990s was advanced partly by intellectual property rights, which provided an incentive, combined with a streamlined regulatory framework, which provided clarity and certainty to manufacturers.[286]

"It is important to establish further proportionate regulatory processes around AI that maintain appropriate safeguards whilst also fostering a facilitative environment for innovation in this field."—Academy of Medical Sciences[287]

Enforcing the EU's requirements for medical devices therefore calls for a clear and efficient regulatory framework within which manufacturers can develop AI products and bring them to market. Such a framework should be grounded in an evidence-based assessment of risk—"principally high quality science, informed by a rigorous understanding of benefits and costs", advocates the European Risk Forum.[288] It should be directly tied to the technical standards around AI-assisted healthcare (see *Standards* section). Ultimately, this will provide patients with access to healthcare that is at once timely, innovative, safe, and effective.

One approach is *regulatory sandboxing*, in which regulators work together with manufacturers and healthcare bodies, among others, to develop standards and policies that are iteratively adapted as the impact of AI products and policies becomes clearer.[289] This experimental approach is being successfully trialled since 2016 by the UK's Financial Conduct Authority for the regulation of financial technology, and has recently been advocated by DeepMind Health's independent review panel for the healthcare sector.[290] Given the high-consequence nature of healthcare, this approach may be most appropriate for low-risk AI products, though any insights gained could inform regulatory policies for higher risk products.

284 Clinical Decision Support Coalition, *Voluntary Industry Guidelines for the Design of Medium Risk Clinical Decision Support Software*; Hall and Pesenti, *Growing the artificial intelligence industry in the UK*.
285 Knut Blind (2012). *The Impact of Regulation on Innovation*. Nesta.
286 Wrubel, n., Krimsky, n., and Anderson, n. (1997). "Regulatory Oversight of Genetically Engineered Microorganisms: Has Regulation Inhibited Innovation?" *Environ. Manage.* 21.4, pp. 571–586.
287 Academy of Medical Sciences, *Written evidence, Lords Select Committee on Artificial Intelligence*.
288 European Risk Forum (2016). *Risk Regulation and Innovation*.
289 Fenwick, M., Kaal, W. A., and Vermeulen, E. P. M. (2016). "Regulation Tomorrow: What Happens When Technology is Faster than the Law?" *Lex Research Topics in Corporate Law & Economics Working Paper No. 2016-8*.
290 Financial Conduct Authority (2017). *Regulatory sandbox lessons learned report*. Bracken, M. et al. (2017). *DeepMind Health Independent Review Panel Annual Report*.

"The idea behind sandboxes and test beds is risk management—to test out new ideas in safe environments that minimise negative risks but also make the most of positive risks."—Geoff Mulgan, Chief Executive Officer, Nesta[291]

The FDA's Digital Health program is also experimenting with regulation with its new *company pre-certification* pilot program.[292] This approach shifts the regulatory focus from being product-based to company-based. Pre-certified companies demonstrating high standards of design, validation, and quality management, are able to submit less (or no) information for regulatory approval of certain products.[293] In return, the FDA gains in-depth access to companies' software development and quality management strategies—valuable information that will feed back into the FDA's regulatory guidance and policies.[294] It will be important to track how this program develops, and to consider how similar regulatory innovation can work in the UK, with its more complex regulatory arrangements including the MHRA, notified bodies, and the European Commission.[295]

Recommendation 12

The Medicines and Healthcare Products Regulatory Agency, the Centre for Data Ethics and Innovation, and industry partners should evaluate regulatory approaches, such as regulatory sandboxing, that can foster innovation in AI-assisted healthcare, ensure patient safety, and inform on-going regulatory development.

5.3 Supporting healthcare innovators

Providing companies with resources and opportunities that enable better product validation is a complementary strategy to ensure synergy between regulation and innovation. For example, the Accelerated Access Review (AAR) recommends creating a digital health catalyst, funded through public and private investment, that would provide support for late-stage testing of high quality digital health products in a real-world environment, such as the NHS.[296]

Similarly, the AAR recommends the creation of a strategic commercial unit within the NHS that can establish partnerships with innovative companies.[297] Such win-win agreements would provide companies with clinical data for their products and access to the NHS market, while the NHS would benefit from flexible pricing and early access to promising healthcare innovation.[298] A key example of this is the ongoing NHS Test Beds programme, a collaboration between the NHS and industry to develop and pilot the use of AI algorithms and other innovations in the context of real-world patient care in seven NHS sites.[299] If proven successful, such programmes should be expanded to match the growing AI market.

Recommendation 13

The NHS should expand innovation acceleration programmes that bridge healthcare and industry partners, with a focus on increasing validation of AI products in real-world contexts and informing the development of a regulatory framework.

291 Mulgan, G. (2017). *Anticipatory Regulation: 10 ways governments can better keep up with fast-changing industries*.
292 FDA Center for Devices and Radiological Health (2017b). *Digital Health Software Precertification (PreCert) Program*. WebContent.
293 Ibid.
294 Ibid.
295 Van Norman, G. A. (2016). "Drugs and Devices: Comparison of European and U.S. Approval Processes". *JACC: Basic to Translational Science* 1.5, pp. 399–412.
296 UK Government (2016). *Accelerated Access Review: Final Report*. Independently chaired report, supported by the Wellcome Trust. UK Government.
297 Ibid.
298 Ibid.
299 Galea, Hough, and Khan, *Test Beds: The story so far*.

5.4 Brexit

The future of AI-assisted healthcare regulation in the UK should be considered in light of Brexit and the UK's future relationship with the European Medicines Agency (EMA).[300] Assessment of medical devices by the MHRA independently of the EMA would require a significant increase in funding and in the workforce.[301] The potential long-term divergence between UK and EU requirements for product safety could also pose challenges for UK manufacturers.[302] AI products are particularly relevant in this scenario, given the relative ease of bringing digital health products to an international market.

Aligning the UK with EU regulations could be achieved by the MHRA seeking continued mutual recognition with the EMA, which would almost certainly require a commitment by the UK to existing and future EU rules.[303] Alternatively, the UK may seek unilateral recognition of products approved by the EMA, although this would diminish the position of the UK in the global market and its ability to shape legislation.[304]

Recommendation 14

The Medicines and Healthcare Products Regulatory Agency and other Government bodies should arrange a post-Brexit agreement ensuring that UK regulations of medical devices, including AI-assisted healthcare, are aligned as closely as possible to the European framework and that the UK can continue to help shape Europe-wide regulations around this technology.

300 Bell, *Life Sciences Industrial Strategy*.
301 Brexit Health Alliance (2018). *Brexit and the impact on patient access to medicines and medical technologies*.
302 Ibid.
303 Bell, *Life Sciences Industrial Strategy*.
304 Ibid.

6 The workforce

AI's purpose is to serve as a tool in the hands of the healthcare workforce, with the overarching aim of improving healthcare delivery. It is then inevitable that AI will impact the healthcare workforce—its structure, function, organisation, and volume—and including its clinical, administrative, and management staff. Conversely, the healthcare workforce will help determine the uptake of AI-assisted healthcare. The nature of these interactions may not be trivial or straightforward, requiring careful consideration and foresight.

"It is vital that we have a realistic, constructive and balanced discussion on the opportunities and challenges AI will bring to the UK workforce."—techUK[305]

6.1 Estimating workforce impact

In an effort to understand AI's impact on the workforce, many studies have attempted to quantify the number and proportion of jobs likely to be cut through AI-related job automation. In a widely-cited set of studies, Frey and Osborne analysed the relationship between job skills and the likelihood of AI-related automation of jobs across a range of sectors.[306] In the UK, they found that roughly a third of jobs are at a high risk of being automated—an estimate that mirrors an analysis from PwC, and that extrapolates to 15 million potentially automatable jobs over the coming decades.[307] A study by Nesta and the Oxford Martin School combined the strengths of expert opinion, analysis of broad socio-economic trends, and quantitative analysis of job data.[308] They suggest of a more optimistic but also more uncertain future: 20% of UK jobs will face a decline through automation, 8% will increase in demand, and 70% of jobs have no reliable prediction at all.[309]

Arntz et al. (2016) took a different approach by analysing component tasks within each job, rather than jobs as a whole.[310] They found that, as of 2012, only 10% of UK jobs had a high risk of having over 70% of their component tasks automated.[311] Similarly, the McKinsey Global Institute found that, globally, only 5% of jobs can be entirely automated.[312]

Although useful, such studies of job automation should be taken with a grain of salt.[313] Most only consider the technological, rather than the economical or structural, feasibility of automation.[314] Adoption of technology does not necessarily parallel its advancement, and it is difficult to meaningfully factor in trends such as globalisation, demographic change, or political uncertainty.[315] Many of these job estimates do not consider potential increases in productivity or economic growth indirectly created by AI, nor do they consider that workers with the same job can perform different tasks.[316]

Even if taken at face value, the above studies all indicate that the healthcare sector—in

305 techUK (2017). *Written evidence, Lords Select Committee on Artificial Intelligence*.
306 Frey, C. B. and Osborne, M. A. (2013). "The future of employment: how susceptible are jobs to computerisation?" *Technol. Forecasting Social Change* 114, pp. 254–280; Frey, C. et al. (2015). *Technology at Work: The future of innovation and employment*. Frey, C. B. et al. (2016). *Technology at Work v2. 0: The future is not what it used to be*.
307 Frey et al., *Technology at Work v2. 0*; Berriman, R and Hawksworth, J (2017). *Will robots steal our jobs? The potential impact of automation on the UK and other major economies*. PwC.
308 Bakhshi, H et al. (2017). *The Future of Skills: Employment in 2030*. Nesta.
309 Ibid.
310 Arntz, M., Gregory, T., and Zierahn, U. (2016). "The Risk of Automation for Jobs in OECD Countries".
311 Ibid.
312 Manyika, J. et al. (2017). *A future that works: Automation, employment and productivity*. McKinsey Global Institute.
313 Bakhshi et al., *The Future of Skills*.
314 Ibid.
315 Bakhshi et al., *The Future of Skills*; Harriss, L and Ennis, J (2016). *Automation and the Workforce (POSTnote 534)*. Parliamentary Office of Science and Technology.
316 Harriss and Ennis, *Automation and the Workforce (POSTnote 534)*; Arntz, Gregory, and Zierahn, "The Risk of Automation for Jobs in OECD Countries"; Brandes, P. and Wattenhofer, R. (2016). "Opening the Frey/Osborne Black Box: Which Tasks of a Job are Susceptible to Computerization?" *arXiv:1604.08823 [cs]*.

terms of available jobs—has a relatively brighter future ahead.[317] Frey and Osborne found that almost all the 25 "least automatable" jobs reside in the healthcare sector.[318] Likewise, a report by the Royal Society of Arts has found that only 1 out of 37 business leaders surveyed in the UK believed that automation will cut a high level of healthcare jobs.[319] This is likely due to healthcare's unique skill requirements, its predicted growth in demand, and AI-related job creation.

Healthcare's unique skills requirements

A key factor is the unique combination of skills required in healthcare (see also Figure 12).[320] The McKinsey Global Institute found that healthcare was protected from full automation due to its requirements to manage people, conduct expert decision-making, and perform physical activity in unpredictable environments.[321] The importance of interpersonal and higher-order cognitive skills in the healthcare sector has also been emphasised in reports by Nesta, the World Bank, and the UK Commission for Employment and Skills[322]

"The complexity of the interaction between a physician and patient. It routinely requires empathy and nuance, as well as expertise, complex decision making, context shifting, and unpredictable physical activity—often all at the same time. That's human terrain."—Jack Stockert, Managing Director, Health2047[323]

Figure 12: Top knowledge and skills important for future UK job demand

✓Judgement + decision-making	✓Critical thinking	✓Social perceptiveness
✓Fluency of ideas	✓Instructing	✓Operations analysis
✓Active learning	✓Education + training	✓Psychology
✓Learning strategies	✓Management of Personnel Resources	✓Time management
✓Originality	✓Coordination	✓Oral comprehension
✓Systems evaluation	✓Inductive reasoning	✓Memorisation
✓Deductive reasoning	✓Problem sensitivity	✓Speaking
✓Complex problem-solving	✓Information ordering	✓Oral expression
✓Systems analysis	✓Active listening	✓Category flexibility
✓Monitoring	✓Administration + management	✓Sociology + anthropology

Adapted from Bakhshi et al. (2017) *The Future of Skills: Employment in 2030*

Healthcare's predicted growth in demand

The continuing growth in healthcare demand serves as another protective factor against AI-related job losses.[324] UK-based studies situate healthcare as one of the sectors with the largest number of new job opportunities in the coming decades.[325] There is already a shortage of GPs and nurses in the UK.[326] The average GP's workload has steadily

317 Frey et al., *Technology at Work v2. 0*; Bakhshi et al., *The Future of Skills*; Manyika et al., *A future that works: Automation, employment and productivity*.
318 Frey and Osborne, "The future of employment".
319 Dellot, B. and Stephens, F.-W. (2017). *The age of automation: Artificial Intelligence, robotics and the future of low-skilled work.* The Royal Society of Arts.
320 Bakhshi et al., *The Future of Skills.*
321 Manyika et al., *A future that works: Automation, employment and productivity.*
322 Bakhshi et al., *The Future of Skills*; Cunningham Paula, W. V. V. (2016). "Employer Voices, Employer Demands, and Implications for Public Skills Development Policy Connecting the Labor and Education Sectors". *World Bank Research Observer* 31.1, pp. 102–134; Howat, C, Lawrie, M, and Sutton, R (2015). *Sector insights: skills and performance challenges in the health and social care sector.* UK Commission for Employment and Skills.
323 "How AI can help doctors — to a point" (2017). *Axios.*
324 Frey et al., *Technology at Work v2. 0.*
325 Bakhshi et al., *The Future of Skills.*
326 Baird, B. et al. (2016). *Understanding pressures in general practice.* The King's Fund.

increased in the last decade, and figures from 2016 suggest that 20% more nurses left the profession than joined, citing unsustainable workforce pressures.[327] All of this is likely to be compounded by the impact of Brexit.[328] Rather than gutting jobs, AI may thus help to ameliorate the workforce pressure that the healthcare sector faces by both optimising workflows and reducing service demand through innovations in patient triaging and patient self-care.[329]

"The NHS is now struggling to cope all year round. It is a pressure cooker and with bed occupancy at such constantly high levels and community services stretched, there is nowhere for the pressure to escape to. It would now take very little for hospitals to be fully overwhelmed."—Lara Carmona, Director of Policy, International and Parliamentary affairs, Royal College of Nursing[330]

AI-related job creation

Some impact studies have considered the potential increase in jobs to result from AI and automation. The Pew Research Center found that half of surveyed technology experts predicted that robotics and automation would create jobs at a similar rate than it displaced them.[331] An analysis of job data between 2001-2015, found that, although technology is likely to have displaced over 800,000 jobs in the UK, it also created nearly 3.5 million higher paying jobs over the same period.[332] This trend is also very much present in healthcare. Numerous reports have highlighted the demand for workers with both clinical and data analytics or informatics experience.[333] "We worry most about the relative absence of a well-trained, professional informatics workforce", concludes the Wachter Review on NHS digitisation.[334] Implementing AI-assisted healthcare requires a large, highly skilled workforce, opening up new job opportunities along the way.[335]

6.2 Re-thinking workforce impact

Because most research has focused on the number of jobs either lost or gained, the complexities of employment and AI may be missed.[336] More than inducing a shift in employment, AI may provide an opportunity for a re-structuring of the labour force and the nature of work.[337] For example, AI-assisted automation may provide clinicians and other healthcare professionals with additional time to spend on important and rewarding tasks, such as patient engagement.[338] It may also empower workers, such as nurses and nursing assistants, to undertake more independent decision-making and management.[339]

"We believe that jobs are more likely to evolve than to be eliminated in the wake of AI's development. The question then becomes one of technology's impact on job quality rather than job quantity."—The Royal Society of Arts[340]

It is thus equally important to understand *how* jobs will change, rather than simply *how many* will change.[341] The Future of Healthcare project is an example of this, with its focus on the task content of jobs in primary care, and how its various workflows can be impacted by automation.[342] The American College of Radiology's new Data Science Institute is

327 Baird et al., *Understanding pressures in general practice*; The Nursing and Midwifery Council (2017). *The NMC Register*. The Nursing and Midwifery Council.
328 Tobin, J. (2016). *NHS and Social Care Workforce: Implications of Leaving the European Union*. House of Lords Library.
329 PwC, *What doctor?*; Gretton, C and Honeyman, M (2016). *The digital revolution*. The King's Fund.
330 Campbell, D. (2016). "Demand for NHS care is dangerously high, says thinktank". *The Guardian*.
331 Smith, A. and Anderson, J. (2014). *AI, Robotics, and the Future of Jobs*. Pew Research Center.
332 Stewart, I., Debapratim, D., and Cole, A. (2015). *Technology and people: The great job-creating machine*. Deloitte.
333 Monitor Deloitte (2015). *Digital Health in the UK.*. The UK Office for Life Sciences; Wachter, *Making IT work: harnessing the power of health information technology to improve care in England*.
334 Wachter, *Making IT work: harnessing the power of health information technology to improve care in England*.
335 Monitor Deloitte, *Digital Health in the UK*.
336 IEEE (2017). *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, Version 2*. IEEE.
337 Ibid.
338 Dellot and Stephens, *The age of automation*.
339 Health Education England and The Royal College of Nursing (2017). *Improving Digital Literacy*. Health Education England and The Royal College of Nursing.
340 Dellot and Stephens, *The age of automation*.
341 Bakhshi et al., *The Future of Skills*; IEEE, *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, Version 2*.
342 Willis, M (2017). *Health Automation: Computerisation and automation in general practice services*. Nottingham, UK.

studying the implementation of AI tools in the radiologist's workflow.[343] Likewise, each relevant professional body in the UK, such as the Medical Royal Colleges, should lead their own impact assessment of AI.

"It is important to focus the analysis on how employment structures will be changed by automation and AI rather than on solely dwelling on the number of jobs that might be impacted. The analysis should focus on how current task content of jobs are changed based on a clear assessment of the automatibility of the occupational description of such jobs."—Institute of Electrical and Electronics Engineers[344]

Recommendation 15

The General Medical Council, the Medical Royal Colleges, Health Education England, and AI-related bodies should partner with industry and academia on comprehensive examinations of the healthcare sector to assess which, when, and *how* jobs will be impacted by AI, including analyses of the current strengths, limitations, and workflows of healthcare professionals and broader NHS staff. They should also examine how AI-driven workforce changes will impact patient outcomes.

Moreover, the Science and Technology Committee makes it clear that the UK has a "digital skills crisis", with almost a quarter of the UK lacking basic IT skills, let alone an understanding of AI.[345] This includes the healthcare sector, with a review by Health Education England concluding that "the need for leadership and a strategic approach to digital literacy acquisition is clear."[346]

Thus, the key AI-related employment issue in healthcare is meeting workforce demands: having a large enough workforce with the digital skills, as well as the "un-automatable" interpersonal and cognitive skills necessary for both developing AI capability, and for working productively with the technology as it becomes commonplace.[347]

"Professional bodies representing relevant clinicians and health professionals are very important stakeholders to involve when considering the use of AI, particularly with regard to professional education, training and workforce planning."—Cancer Research UK[348]

6.3 Developing capability for AI

To support the development, deployment, and maintenance of AI technology, the field of health informatics must expand within the NHS, as highlighted by the Wachter Review on NHS digitisation.[349] A broad field lying at the intersection of healthcare, data science, and computer science, health informatics uses data to drive the planning, management, and delivery of healthcare, including the use of AI and software development.[350] However, it has traditionally been far-removed from the "coal-face" of clinical practice.[351] To bridge the gap between these two areas, the Wachter Review recommends that each trust have a multi-disciplinary cohort of clinician-informaticians.[352] Clinician-informaticians should support evidence-based procurement of AI technology: identifying opportunities where it can improve healthcare delivery, evaluating the market for potential solutions, and

343 American College of Radiology (2017). "ACR Data Science Institute™ to Guide Artificial Intelligence Use in Medical Imaging". *PR Newswire*.
344 IEEE, *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, Version 2*.
345 House of Commons Science and Technology Committee, *The Big Data Dilemma*; Ecorys UK (2016). *Digital skills for the UK economy*. UK Department for Business, Innovation & Skills; Department for Culture, Media & Sport.
346 Health Education England (2016). *Literature review: Examining the extent to which digital literacy is seen as a challenge for trainers*. Health Education England.
347 Monitor Deloitte, *Digital Health in the UK*; Frey et al., *Technology at Work v2. 0*; Bakhshi et al., *The Future of Skills*.
348 Cancer Research UK, *Written evidence, Lords Select Committee on Artificial Intelligence*.
349 Wachter, *Making IT work: harnessing the power of health information technology to improve care in England*.
350 Shortliffe, E. H., Detmer, D. E., and Munger, B. S. (2016). "Clinical Informatics: Emergence of a New Profession". *Clinical Informatics Study Guide*. Springer, Cham, pp. 3–21.
351 Wachter, *Making IT work: harnessing the power of health information technology to improve care in England*.
352 Ibid.

facilitating adoption of the technology. This cohort, with dual expertise in clinical care and health informatics, should ensure that AI innovation remains tied to the human aspect of healthcare delivery, by taking into account both clinician and patient perspectives.[353]

Recommendation 16

The Federation of Informatics Professionals and the Faculty of Clinical Informatics should continue to lead and expand standards for health informatics competencies, integrating the relevant aspects of AI into their training, accreditation, and professional development programmes for clinician-informaticians and related professions.

6.4 Working with AI

Digital literacy—"capabilities which fit someone for living, learning, working, participating and thriving in a digital society"—is a prerequisite for the traditional healthcare workforce to be AI-ready.[354] All healthcare professionals should be able to work comfortably in a digital environment, with an understanding of data governance principles, including data privacy, security, and consent.[355] They should be able to grasp the value of data and the importance of data integrity, not only for immediate patient outcomes, but also for the future use of such data in the development of AI and other digital health innovation. Digital literacy is already a focus for the National Information Board and Health Education England, but the need for it is becoming more urgent with the increasing pace of technological development.[356]

Beyond basic digital literacy, the healthcare workforce should be able to interact effectively with AI technology.[357] For the near future, most AI will serve as a tool in human hands, rather operate entirely autonomously. This requires that any information provided by AI products, such as decision support tools, is interpreted appropriately by healthcare professionals, and integrated with their own medical training.[358] This could entail an understanding of the kind of data AI products rely on, how data is processed, and the degree of uncertainty around any decision support.[359] Moreover, workers should be able to effectively communicate any associated risks to patients or co-workers. Although part of the burden lies with manufacturers and regulators to ensure that available technology meets standards of interpretability and usability (e.g., see *Standards* and *Regulations* sections), part of the challenge involves appropriate workforce training.[360]

"Increased use of AI will lead to changes in the skillset required for professionals, and training programmes should reflect this to allow staff to maximise on the opportunities afforded by AI. As such, there is a need to identify and address any gaps in capability to ensure the necessary training for the integration, manipulation and analysis of the data within appropriate ethical and regulatory frameworks."—Academy of Medical Sciences[361]

Radiologists and pathologists are examples of professionals that may need to prepare for such changes, with some suggesting that they be re-defined as "information specialists".[362] In their new role, instead of scanning dozens of images by eye, radiologists and pathologists

353 Wachter, *Making IT work: harnessing the power of health information technology to improve care in England*; Sood, H., NcNeil, K., and Keogh, B. (2017). "Chief clinical information officers: clinical leadership for a digital age". *BMJ* 358, j3295.
354 Health Education England and The Royal College of Nursing, *Improving Digital Literacy*.
355 Health Education England and The Royal College of Nursing, *Improving Digital Literacy*; House of Commons Science and Technology Committee, *Robotics and artificial intelligence*.
356 Health Education England and The Royal College of Nursing, *Improving Digital Literacy*.
357 Cancer Research UK, *Written evidence, Lords Select Committee on Artificial Intelligence*; Academy of Medical Sciences, *Written evidence, Lords Select Committee on Artificial Intelligence*.
358 The Royal Society, *Machine learning: the power and promise of computers that learn by example*; International Medical Device Regulators Forum, *IMDRF Proposed Document: Software as a Medical Device (SaMD): Clinical Evaluation*.
359 International Medical Device Regulators Forum, *IMDRF Proposed Document: Software as a Medical Device (SaMD): Clinical Evaluation*; Clinical Decision Support Coalition, *Voluntary Industry Guidelines for the Design of Medium Risk Clinical Decision Support Software*.
360 IEEE, *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, Version 2*.
361 Academy of Medical Sciences, *Written evidence, Lords Select Committee on Artificial Intelligence*.
362 Jha, S. and Topol, E. J. (2016). "Adapting to Artificial Intelligence: Radiologists and Pathologists as Information Specialists". *JAMA* 316.22, pp. 2353–2354.

would be guided by automatic AI-driven analysis of such images, and then "interpret the important data, advise on the added value of another diagnostic test, such as the need for additional imaging, anatomical pathology, or a laboratory test, and integrate information to guide clinicians".[363] The Royal College of Radiologists has embraced such a role for its professionals, arguing that AI can alleviate their workload, while allowing them to focus on more complex decision-making.[364] To master this role, such professionals would need to complement their medical training with a basic understanding of AI and data science.[365]

"The medical syllabus needs to start incorporating not just medical statistics but some basics of data science."—Hugh Harvey, Consultant Radiologist, UK Radiology Informatics Committee member[366]

A similar approach has been advanced by the Royal College of Nursing, which advocates for a re-thinking of the profession from one in which nurses perform a routine collection of tasks (a "nursing is doing" model) to one in which nurses are equipped with a deeper understanding of medical systems, and enabled to exercise clinical judgment (a "nursing is knowing" model).[367] By providing richer information and decision support, AI technology will further empower nurses in the provision of care.[368]

"We need nurses and midwives that are properly informed, trained and equipped. We need a workforce that is involved in the design, development and deployment of technology in healthcare."—Royal College of Nursing[369]

Crucially, any training strategy cannot focus solely on the technical aspects of AI tools. Equally important is the development of the "un-automatable" skills.[370] As recent reports have highlighted, the most in-demand skills will continue to be socio-emotional and higher-order cognitive skills: empathy, creativity, and team-work, as well as evaluation, decision-making, and active learning.[371] Not only will such skills protect healthcare professionals from job displacement, but they will be essential for working with AI.

Recommendation 17

Health Education England should expand training programmes to advance digital and AI-related skills among healthcare professionals. Competency standards for working with AI should be identified for each role and established in accordance with professional registration bodies such as the General Medical Council. Training programmes should ensure that "un-automatable" socio-emotional and cognitive skills remain an important focus.

6.5 Leadership

AI will not impact the healthcare system until its value is realised by those in a position of responsibility to promote change. The Wachter Review has emphasised the profound lack of leaders in most trusts with adequate training in both clinical care and informatics.[372] Leaders with expertise in AI are even rarer. An on-going study from King's College London has surveyed 20 governmental departments who have cited leadership as one of the key

363 Jha and Topol, "Adapting to Artificial Intelligence".
364 The Royal College of Radiologists, *Written evidence, Lords Select Committee on Artificial Intelligence*.
365 Jha and Topol, "Adapting to Artificial Intelligence".
366 Harvey, H (2017). *Oral evidence, Lords Select Committee on Artificial Intelligence*.
367 Health Education England and The Royal College of Nursing, *Improving Digital Literacy*.
368 PwC, *What doctor?*
369 Health Education England and The Royal College of Nursing, *Improving Digital Literacy*.
370 Bakhshi et al., *The Future of Skills*.
371 Bakhshi et al., *The Future of Skills*; Cunningham, "Employer Voices, Employer Demands, and Implications for Public Skills Development Policy Connecting the Labor and Education Sectors"; Howat, Lawrie, and Sutton, *Sector insights*.
372 Wachter, *Making IT work: harnessing the power of health information technology to improve care in England*.

challenges facing public bodies in using algorithms for decision-making.[373]

To address this, the Wachter Review has recommended that there be at least one Chief Clinical Information Officer (CCIO) per NHS trust that has the autonomy and authority to drive change.[374] CCIOs are clinical practitioners with expertise in digital health systems that are responsible for overseeing the strategic aims of each trust, including the design, implementation, and use of health informatics.[375] CCIOs should also have expertise in AI-assisted healthcare innovation, with a grasp of its unique opportunities, risks, and pitfalls. They should foster a digital- and AI-ready workforce within their trust by supporting skills training for their frontline workers and by recruiting clinician-informaticians and others that can enable AI capability.[376] To address this, the NHS Digital Academy has recently launched with the aim of training CCIOs. It will be important that its training programme adapts to the rapidly evolving fields of AI-assisted healthcare and AI ethics.[377]

"The dearth of professional, well-supported CCIOs with appropriate authority and resources is an enormous obstacle to successful deployment and benefits realisation of health IT at the trust level."—Professor Robert Wachter, University of California, San Francisco[378]

Recommendation 18

The NHS Digital Academy should expand recruitment and training efforts to increase the number of Chief Clinical Information Officers across the NHS, and ensure that the latest AI ethics, standards, and innovations are embedded in their training programme.

373 Griffiths, A, Rothstein, H, and Demeritt, D (2017). *Written evidence submitted to the Science and Technology Select Committee on Algorithms in Decision-Making.*
374 Wachter, *Making IT work: harnessing the power of health information technology to improve care in England*; Sood, NcNeil, and Keogh, "Chief clinical information officers".
375 Wachter, *Making IT work: harnessing the power of health information technology to improve care in England.*
376 Ibid.
377 Sood, NcNeil, and Keogh, "Chief clinical information officers".
378 Wachter, *Making IT work: harnessing the power of health information technology to improve care in England.*

7 Legal liability

The adoption of AI in the healthcare sector raises legal questions related to liability.[379] An obvious example is who should be held responsible for a misdiagnosis when AI is involved in the clinical process: the doctor, hospital, or manufacturer? Members of the European Parliament have recently pushed for an updated liability framework around robotics and AI.[380] In the UK, stakeholders such as the Wellcome Trust and the Association of Medical Research Charities, as well as legal and AI experts, have all called for clarification on how the UK's system of liability in healthcare will deal with emerging AI technology.[381]

"We need to ensure a clear chain of human accountability, responsibility and liability for decisions that an algorithm makes that impacts on human lives whether it be job selection, medical procedures or car insurance."—Noel Sharkey, Professor of AI and Robotics at Sheffield University, Co-Director at the Foundation for Responsible Robotics[382]

As with the standards and regulations of AI-assisted healthcare products, the challenge to the legal framework that relates to professional liability depends on whether AI replaces or augments healthcare professionals.[383] Existing liability law in healthcare assumes the role of a human in the medical process.[384] Thus, situations in which the human aspect is entirely removed will likely require the legal system to re-think the liability framework, similar to current questions around autonomous vehicles.[385] The replacement of human professionals by AI, however, is unlikely in the foreseeable future.[386] Instead, the case of human-AI interaction is a more pressing question for existing liability laws, presenting its own unique challenges. Arguably more than any other issue, the question of liability illustrates the interdisciplinary nature of the challenges that AI brings to healthcare, including questions about medical ethics, workforce training, product regulation, and public support.

7.1 The duty of healthcare professionals

In the healthcare sector, the most likely source of legal liability will be the duty of care imposed by the law of negligence.[387] Healthcare professionals have the duty to use reasonable care and skill in diagnosing and treating patients, "providing a service of no less a quality than that to be expected, based on the skills, responsibilities, and range of activities within their profession".[388] In a clinical negligence claim, the patient must prove three things: (1) that they are owed a duty of care; (2) that there was a breach of that duty; and (3) that the patient suffered harm because of the breach.[389] The second point—the breach of duty—is the element that is becoming more complicated by the introduction of AI in healthcare.[390]

AI, like other technology, is a tool that currently requires human interpretation for safe and

379 Wellcome Trust and the Association of Medical Research Charities (AMRC), *Written evidence, Lords Select Committee on Artificial Intelligence*.
380 European Parliament (2017). *Robots and artificial intelligence: MEPs call for EU-wide liability rules*.
381 Wellcome Trust and the Association of Medical Research Charities (AMRC), *Written evidence, Lords Select Committee on Artificial Intelligence*; Reed, C., Kennedy, E., and Silva, S. N. (2016). "Responsibility, Autonomy and Accountability: legal liability for machine learning"; DeepMind, *Written evidence, Lords Select Committee on Artificial Intelligence*.
382 All Party Parliamentary Group on Artificial Intelligence (2017). *Ethics and Legal in AI: Decision Making and Moral Issues*. 2. All Party Parliamentary Group on Artificial Intelligence.
383 Reed, Kennedy, and Silva, "Responsibility, Autonomy and Accountability: legal liability for machine learning".
384 Ibid.
385 Ibid.
386 Bughin et al., *Artificial intelligence: the next digital frontier?*
387 Reed, Kennedy, and Silva, "Responsibility, Autonomy and Accountability: legal liability for machine learning".
388 Kline, R. and Khan, S. (2013). *The Duty of Care of healthcare professionals*. Public World.
389 Reed, Kennedy, and Silva, "Responsibility, Autonomy and Accountability: legal liability for machine learning".
390 Ibid.

effective use in healthcare.[391] Unlike other technology, however, a lot of existing AI remains a "black box" for typical human users, with a lack of standards for the level of interpretability required for safe and effective use (see *Standards* section). According to Nicola Perrin, head of *Understanding Patient Data*, "the transparency and the explanation of how a decision has been made are going to be crucial" to liability.[392] The level of interpretability in an AI product constrains the interaction between the user and the technology, with "black-box" systems forcing healthcare professionals to blindly trust or distrust their output.[393]

Put another way, leaving the ultimate clinical decision to a human professional—placing the burden of liability onto them—requires enough interpretability in the technology for that user to safely and effectively integrate its output with their own knowledge.[394] At the other extreme, requiring AI products to be validated such that their safe and effective use is *independent* of the human user, will at least partly shift the burden of liability to the manufacturer.[395]

"...if the [clinical decision support system] does not enable the intended user to sufficiently understand the recommendation made by the software and equally importantly, the basis for the recommendation, such [clinical decision support system] runs the risk of being used as a substitute for the user's expertise and judgment. In such cases, the software may need to be validated to a higher degree..."—Clinical Decision Support Coalition[396]

Additional ambiguity arises because there is also a lack of standards for the training that healthcare professionals should have to work safely and effectively with AI (see *Workforce* section).[397] Likewise, the Royal College of Radiologists says that "it is not clear at what point, failure to use an AI system would become negligent".[398] The standard of care changes gradually, depending on "how widely the technology has been adopted by others working in the same field".[399] This transition between innovation and standard practice, likely to occur in the next decade, is the period of highest uncertainty regarding liability.

These uncertainties make it difficult to currently assign liability when things go wrong, because "the application of the law often depends on what a human knew, or ought to have known, at the time the liability arose", says Chris Reed, Professor of Law at Queen Mary University of London.[400] Ultimately, the challenge is determining whether an AI product was inadequately validated, or whether the human user misread it from a lack of expertise.[401] Developing an appropriate liability framework, therefore, relies both on defining validation standards for AI products and training standards for healthcare professionals.[402]

"If something goes wrong, it depends whether the system was designed badly or whether the clinician misread it."—Julian Huppert, MP, Chair of the Independent Review Panel for DeepMind Health[403]

391 International Medical Device Regulators Forum, *IMDRF Proposed Document: Software as a Medical Device (SaMD): Clinical Evaluation*.
392 Perrin, N., *Oral evidence, Lords Select Committee on Artificial Intelligence*.
393 Price, N. W. "Medical Malpractice and Black-Box Medicine". *Big Data, Health Law, and Bioethics*. Cambridge University Press.
394 Ibid.
395 Clinical Decision Support Coalition, *Voluntary Industry Guidelines for the Design of Medium Risk Clinical Decision Support Software*.
396 Ibid.
397 Cancer Research UK, *Written evidence, Lords Select Committee on Artificial Intelligence*; Academy of Medical Sciences, *Written evidence, Lords Select Committee on Artificial Intelligence*.
398 The Royal College of Radiologists, *Written evidence, Lords Select Committee on Artificial Intelligence*.
399 Reed, Kennedy, and Silva, "Responsibility, Autonomy and Accountability: legal liability for machine learning"; Price, "Medical Malpractice and Black-Box Medicine"; Berner, E. S. and La Lande, T. J. (2007). "Overview of Clinical Decision Support Systems". *Clinical Decision Support Systems*. Ed. by Hannah, K. J., Ball, M. J., and Berner, E. S. New York, NY: Springer New York, pp. 3–22.
400 Reed, Kennedy, and Silva, "Responsibility, Autonomy and Accountability: legal liability for machine learning".
401 Huppert J., *Oral evidence, Lords Select Committee on Artificial Intelligence*.
402 Reed, Kennedy, and Silva, "Responsibility, Autonomy and Accountability: legal liability for machine learning"; Price, "Medical Malpractice and Black-Box Medicine".
403 Huppert J., *Oral evidence, Lords Select Committee on Artificial Intelligence*.

Recommendation 19

Legal experts, ethicists, AI-related bodies, professional medical bodies, and industry should review the implications of AI-assisted healthcare for legal liability. This includes understanding how healthcare professionals' duty of care will be affected, the role of workforce training and product validation standards, and the potential role of NHS Indemnity and no-fault compensation systems.

7.2 The duty of healthcare institutions and manufacturers

In practice, the existing arrangement of NHS Indemnity means that trusts are held liable for their employees' acts of negligence.[404] As suggested by the PHG Foundation, NHS indemnity can simply be extended to cover AI products, a solution that seems to have some initial public support.[405] For example, a Royal Society survey found that the most common answer (32%) provided by the public regarding who should be held liable for AI-related errors was "the organisation the operator and machine work for".[406]

In addition to covering the performance of healthcare professionals, institutions have the duty to procure safe and effective medical equipment and provide sufficient training for its use.[407] The challenge will be to weigh the responsibility of the healthcare institutions with that of AI manufacturers, including determining whether the product was negligently placed on the market (e.g., without adequate validation in real-world environments).[408] One approach is to introduce a system of no-fault compensation, which can be shared by healthcare institutions and manufacturers.[409]

7.3 The public's role

Another important question is whether the issue of liability should be resolved through common law or legislation.[410] The Law Society notes that "one of the disadvantages of leaving it to the courts [...] is that the common law only develops by applying legal principles after the event when something untoward has already happened."[411] One risk is that a messy, high-profile case can halt innovation and adoption, particularly in healthcare, where the slow pace of innovation is well known.[412] As with other policy areas, an important guiding strategy will be understanding how the public views and values liability systems.[413] For example, the Royal Society found that the public preferred humans to be ultimately responsible for decisions in personally sensitive areas.[414]

Recommendation 20

AI-related bodies such as the Ada Lovelace Institute, patient advocacy groups and other healthcare stakeholders should lead a public engagement and dialogue strategy to understand the public's views on liability for AI-assisted healthcare.

404 British Medical Association (2012). *NHS medical indemnity*.
405 PHG Foundation (2017). *Written evidence, Lords Select Committee on Artificial Intelligence*.
406 The Royal Society, *Machine learning: the power and promise of computers that learn by example*.
407 Price, "Medical Malpractice and Black-Box Medicine".
408 Reed, C. (2017). *Written evidence, Lords Select Committee on Artificial Intelligence*.
409 Reed, *Written evidence, Lords Select Committee on Artificial Intelligence*; Price, "Medical Malpractice and Black-Box Medicine".
410 House of Commons Science and Technology Committee, *Robotics and artificial intelligence*.
411 The Law Society (2016). *Written evidence submitted to the House of Commons Select Committee on Science and Technology*.
412 Dixon-Woods, M. et al. (2011). "Problems and promises of innovation: why healthcare needs to rethink its love/hate relationship with the new". *BMJ Quality & Safety* 20.Suppl 1, pp. i47–i51.
413 The Royal Society, *Machine learning: the power and promise of computers that learn by example*.
414 Ipsos MORI (2017). *Public views of Machine Learning*. The Royal Society.

References

Academy of Medical Sciences (2017). *Written evidence, Lords Select Committee on Artificial Intelligence*.

AI100 (2017). *AI Index 2017 Annual Report*.

Al-Harbi, K. S. (2012). "Treatment-resistant depression: therapeutic trends, challenges, and future directions". *Patient Prefer Adherence* 6, pp. 369–388.

"All of Us Research Program ushers in new era for technology-driven citizen science" (2018). *EurekAlert!*

All Party Parliamentary Group on Artificial Intelligence (2017). *Ethics and Legal in AI: Decision Making and Moral Issues*. 2. All Party Parliamentary Group on Artificial Intelligence.

Amarasingham, R. et al. (2016). "Consensus Statement on Electronic Health Predictive Analytics: A Guiding Framework to Address Challenges". *EGEMS (Wash DC)* 4.1.

American College of Radiology (2017). "ACR Data Science Institute™ to Guide Artificial Intelligence Use in Medical Imaging". *PR Newswire*.

American College of Radiology Data Science Institute (2017). *The ACR Data Science Institute Structures Artificial Intelligence Development to Optimize Radiology Care*.

Armstrong, S. (2018). "The apps attempting to transfer NHS 111 online". *BMJ* 360, k156.

Arntz, M., Gregory, T., and Zierahn, U. (2016). "The Risk of Automation for Jobs in OECD Countries".

Baird, B. et al. (2016). *Understanding pressures in general practice*. The King's Fund.

Bakhshi, H et al. (2017). *The Future of Skills: Employment in 2030*. Nesta.

Bejnordi, B. E. et al. (2017). "Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer". *JAMA* 318.22, pp. 2199–2210.

Bell, S. J. (2017). *Life Sciences Industrial Strategy*.

Bernal, N. P. et al. (2017). "Accurate Measurement of Intraoperative Blood Loss during Wound Excision Leads to More Appropriate Transfusion and Reduced Blood Utilization". *Journal of Anesthesia & Clinical Research* 8.11, pp. 1–6.

Berner, E. S. and La Lande, T. J. (2007). "Overview of Clinical Decision Support Systems". *Clinical Decision Support Systems*. Ed. by Hannah, K. J., Ball, M. J., and Berner, E. S. New York, NY: Springer New York, pp. 3–22.

Berriman, R and Hawksworth, J (2017). *Will robots steal our jobs? The potential impact of automation on the UK and other major economies*. PwC.

Bhavnani, S. P. et al. (2017). "2017 Roadmap for Innovation—ACC Health Policy Statement on Healthcare Transformation in the Era of Digital Health, Big Data, and Precision Health: A report of the American College of Cardiology Task Force on Health Policy Statements and Systems of Care". *J. Am. Coll. Cardiol.* 70.21, pp. 2696–2718.

Bracken, M. et al. (2017). *DeepMind Health Independent Review Panel Annual Report*.

Brandes, P. and Wattenhofer, R. (2016). "Opening the Frey/Osborne Black Box: Which Tasks of a Job are Susceptible to Computerization?" *arXiv:1604.08823 [cs]*.

Brexit Health Alliance (2018). *Brexit and the impact on patient access to medicines and medical technologies*.

British Medical Association (2012). *NHS medical indemnity*.

British Standards Institution (2017). *Written evidence, Lords Select Committee on Artificial Intelligence*.

Bughin, J. et al. (2017). *Artificial intelligence: the next digital frontier*? McKinsey Global Institute.

Burgess, M. *The Science Museum's Robots exhibition gives some unsettling truths about humanity*.

Business Leader (2018). "AI start-ups set to change UK healthcare market".

Business Wire (2017). "Zebra Medical Vision Announces New Algorithm to Detect Compression and Other Vertebral Fractures".

Caldicott, F., (2017). *Oral evidence, Lords Select Committee on Artificial Intelligence*.

Campbell, D. (2016). "Demand for NHS care is dangerously high, says thinktank". *The Guardian*.

Campolo, A. et al. (2017). *AI Now 2017 Report*. AI Now Institute.

Cancer Research UK (2017). *Written evidence, Lords Select Committee on Artificial Intelligence*.

Care Quality Commission (2016). *Safe data, safe care*.

CB Insights Research (2017). *Up And Up: Healthcare AI Startups See Record Deals*.

Chakravarty, D. et al. (2017). "OncoKB: A Precision Oncology Knowledge Base". *JCO Precis Oncol* 2017.

Chan, T. et al. (2016). "UK National Data Guardian for Health and Care's Review of Data Security: Trust, better security and opt-outs". *Journal of Innovation in Health Informatics* 23.3, pp. 627–632.

Char, D. S., Shah, N. H., and Magnus, D. (2018). "Implementing Machine Learning in Health Care — Addressing Ethical Challenges". *N. Engl. J. Med.* 378.11, pp. 981–983.

Check Hayden, E. (2015). "Extreme cryptography paves way to personalized medicine". *Nature News* 519.7544, p. 400.

Chekroud, A. M. et al. (2016). "Cross-trial prediction of treatment outcome in depression: a machine learning approach". *The Lancet Psychiatry* 3.3, pp. 243–250.

Clinical Decision Support Coalition (2017). *Voluntary Industry Guidelines for the Design of Medium Risk Clinical Decision Support Software*.

Cunningham Paula, W. V. V. (2016). "Employer Voices, Employer Demands, and Implications for Public Skills Development Policy Connecting the Labor and Education Sectors". *World Bank Research Observer* 31.1, pp. 102–134.

DeepMind (2017a). *Trust, confidence and Verifiable Data Audit*.

DeepMind (2017b). *Written evidence, Lords Select Committee on Artificial Intelligence*.

Dellot, B. and Stephens, F.-W. (2017). *The age of automation: Artificial Intelligence, robotics and the future of low-skilled work*. The Royal Society of Arts.

Denham E (2017). *Oral evidence, Lords Select Committee on Artificial Intelligence*.

Diabetes UK (2014). *The Cost of Diabetes*.

Dixon, P. A. et al. (2015). "National Audit of Seizure management in Hospitals (NASH): results of the national audit of adult epilepsy in the UK". *BMJ Open* 5.3, e007325.

Dixon-Woods, M. et al. (2011). "Problems and promises of innovation: why healthcare needs to rethink its love/hate relationship with the new". *BMJ Quality & Safety* 20.Suppl 1, pp. i47–i51.

Doshi-Velez, F. and Kim, B. (2017). "Towards A Rigorous Science of Interpretable Machine Learning". *arXiv:1702.08608 [cs, stat]*.

Ecorys UK (2016). *Digital skills for the UK economy*. UK Department for Business, Innovation & Skills; Department for Culture, Media & Sport.

Eddy, D. M. et al. (2012). "Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–7". *Value Health* 15.6, pp. 843–850.

England, N. (2017a). *NHS England: Global Digital Exemplars*.

England, P. H. (2017b). *Atrial fibrillation prevalence estimates for local populations*.

Etzioni, O. (2016). "Most experts say AI isn't as much of a threat as you might think". *MIT Technology Review*.

European Commission. *Revisions of Medical Device Directives*.

European Commission (2017). "Regulation (EU) 2017/745". *Official Journal of the European Union* 60.

European Parliament (2017). *Robots and artificial intelligence: MEPs call for EU-wide liability rules*.

European Risk Forum (2016). *Risk Regulation and Innovation*.

FDA Center for Devices and Radiological Health (2017a). *Digital Health Innovation Action Plan*.

FDA Center for Devices and Radiological Health (2017b). *Digital Health Software Precertification (PreCert) Program*. WebContent.

Federal Trade Commission (2015). ""Melanoma Detection" App Sellers Barred from Making Deceptive Health Claims". *Federal Trade Commission*.

Fenech, M., Strukelj, N., and Buston, O. (2018). *Ethical, social, and political challenges of artificial intelligence in health*. Future Advocacy.

Fenwick, M., Kaal, W. A., and Vermeulen, E. P. M. (2016). "Regulation Tomorrow: What Happens When Technology is Faster than the Law?" *Lex Research Topics in Corporate Law & Economics Working Paper No. 2016-8*.

Financial Conduct Authority (2017). *Regulatory sandbox lessons learned report*.

Fitzpatrick, K. K., Darcy, A., and Vierhile, M. (2017). "Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial". *JMIR Mental Health* 4.2, e19.

Floridi, L. and Taddeo, M. (2016). "What is data ethics?" *Phil. Trans. R. Soc. A* 374.2083, p. 20160360.

Freiherr, G. (2018). *How AI can help patients manage diabetes*.

Frey, C. B. and Osborne, M. A. (2013). "The future of employment: how susceptible are jobs to computerisation?" *Technol. Forecasting Social Change* 114, pp. 254–280.

Frey, C. B. et al. (2016). *Technology at Work v2. 0: The future is not what it used to be*.

Frey, C. et al. (2015). *Technology at Work: The future of innovation and employment*.

Galea, A., Hough, E., and Khan, I. (2017). *Test Beds: The story so far*. NHS England.

Gauss Surgical (2017). *Gauss Surgical Receives FDA Clearance for Second Generation Triton for Real-Time Monitoring of Surgical Blood Loss*.

Giger, M. L. (2018). "Machine Learning in Medical Imaging". *Journal of the American College of Radiology* 15.3, pp. 512–520.

Gijsberts, C. M. et al. (2015). "Race/Ethnic Differences in the Associations of the Framingham Risk Factors with Carotid IMT and Cardiovascular Events". *PLoS One* 10.7, e0132321.

Goldacre, B. (2014). *Care.data is in chaos. It breaks my heart*.

Government Office for Science. *Distributed ledger technology: beyond block chain*. Tech. rep.

Gretton, C and Honeyman, M (2016). *The digital revolution*. The King's Fund.

Griffiths, A, Rothstein, H, and Demeritt, D (2017). *Written evidence submitted to the Science and Technology Select Committee on Algorithms in Decision-Making*.

Hall, W. and Pesenti, J. (2017). *Growing the artificial intelligence industry in the UK*.

Hanly, K. (2017). "Google AI chief claims biased algorithms are a big danger". *Digital Journal*.

Harriss, L and Ennis, J (2016). *Automation and the Workforce (POSTnote 534)*. Parliamentary Office of Science and Technology.

Harvey, H (2017). *Oral evidence, Lords Select Committee on Artificial Intelligence*.

Harwich, E. and Laycock, K. (2018). *Thinking on its own: AI in the NHS*. Reform.

Health Education England (2016). *Literature review: Examining the extent to which digital literacy is seen as a challenge for trainers*. Health Education England.

Health Education England and The Royal College of Nursing (2017). *Improving Digital Literacy*. Health Education England and The Royal College of Nursing.

HM Treasury (2017). *Autumn Budget 2017*.

House of Commons Science and Technology Committee (2016a). *Robotics and artificial intelligence.*

House of Commons Science and Technology Committee (2016b). *The Big Data Dilemma*.

House of Commons Science and Technology Committee (2018). *Algorithms in decision-making*.

House of Lords Select Committee on Artificial Intelligence (2018). *AI in the UK: ready, willing and able?*

"How AI can help doctors — to a point" (2017). *Axios*.

Howat, C, Lawrie, M, and Sutton, R (2015). *Sector insights: skills and performance challenges in the health and social care sector*. UK Commission for Employment and Skills.

Hunter, P. (2016). "The big health data sale". *EMBO Rep.* 17.8, pp. 1103–1105.

Huppert J. (2017). *Oral evidence, Lords Select Committee on Artificial Intelligence*.

Iacobucci, G. (2017). "NHS in 2017: Keeping pace with society". *BMJ* 356, p. i6738.

IEEE (2017). *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, Version 2*. IEEE.

Information Commissioner's Office (2017a). *Big data, artificial intelligence, machine learning and data protection*.

Information Commissioner's Office (2017b). *Royal Free - Google DeepMind trial failed to comply with data protection law*.

International Medical Device Regulators Forum (2016). *IMDRF Proposed Document: Software as a Medical Device (SaMD): Clinical Evaluation*.

International Medical Device Regulators Forum (2017). *Software as a Medical Device (SaMD): Clinical Evaluation.*

Ipsos MORI (2017). *Public views of Machine Learning*. The Royal Society.

Ipsos MORI and Wellcome Trust (2016). *The one-way mirror: Public attitudes to commercial access to health data*.

JASON (2017a). *Artificial Intelligence for Health and Health Care*. JSR-17-Task-002. The MITRE Corporation.

JASON (2017b). *Perspectives on Research in Artificial Intelligence and Artificial General Intelligence Relevant to DoD*. JSR-16-Task-003. The MITRE Corporation.

Jha, S. and Topol, E. J. (2016). "Adapting to Artificial Intelligence: Radiologists and Pathologists as Information Specialists". *JAMA* 316.22, pp. 2353–2354.

Jordan, M. I. and Mitchell, T. M. (2015). "Machine learning: Trends, perspectives, and prospects". *Science* 349.6245, pp. 255–260.

Kelnar, D. (2016). *Artificial Intelligence in the UK: Landscape and learnings from 226 startups*.

Kirchner, L. et al. (2016). "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And it's Biased Against Blacks." *ProPublica*.

Kline, R. and Khan, S. (2013). *The Duty of Care of healthcare professionals*. Public World.

Knut Blind (2012). *The Impact of Regulation on Innovation*. Nesta.

Kum, H.-C. and Ahalt, S. (2013). "Privacy-by-Design: Understanding Data Access Models for Secondary Data". *AMIA Jt Summits Transl Sci Proc* 2013, pp. 126–130.

Lipton, Z. C. (2016). "The Mythos of Model Interpretability". *arXiv:1606.03490 [cs, stat]*.

Lipton, Z. C. (2017). "The Doctor Just Won't Accept That!" *arXiv:1711.08037 [stat]*.

Loder, J. and Nicholas, L. (2018). *Confronting Dr Robot: Creating a people-powered future for AI in health*. Nesta.

Lohr, S. (2016). "IBM Buys Truven for $2.6 Billion, Adding to Trove of Patient Data". *The New York Times*.

Magistretti, B. *Swiss data analytics company Sophia Genetics could be Switzerland's next unicorn*.

Malik, O. (2016). "The Hype–and Hope–of Artificial Intelligence". *The New Yorker*.

Mannes, J. (2016). *CareSkore gets $4.3M to bring machine learning to preventive care*.

Manyika, J. et al. (2017). *A future that works: Automation, employment and productivity*. McKinsey Global Institute.

Markoff, J. (2012). "For Web Images, Creating New Technology to Seek and Find". *New York Times*.

Marr, B. (2016). *Big Data In Healthcare: Paris Hospitals Predict Admission Rates Using Machine Learning*.

Medeiros, J. (2018). *This AI unicorn is disrupting the pharma industry in a big way*.

Medicines and Healthcare Products Regulatory Agency (2014). *Medical devices: software applications (apps)*.

Medicines and Healthcare Products Regulatory Agency (2017a). *Guidance on applying human factors to medical devices*.

Medicines and Healthcare Products Regulatory Agency (2017b). *Written evidence, Lords Select Committee on Artificial Intelligence*.

Microsoft Reporter (2018). *Tech that helps epilepsy patients 'could save NHS £250m'*.

Mitchell, R. M. et al. (2017). "Reporting of vertebral fragility fractures: can radiologists help reduce the number of hip fractures?" *Arch Osteoporos* 12.1, p. 71.

Mittelstadt, B. and Floridi, L. (2015). *The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts*.

Mittelstadt, B. D. et al. (2016). "The ethics of algorithms: Mapping the debate". *Big Data & Society* 3.2, p. 2053951716679679.

Molteni, M. (2018). "Meet the Company Trying to Democratize Clinical Trials With AI". *WIRED*.

Monitor Deloitte (2015). *Digital Health in the UK*. The UK Office for Life Sciences.

Mulgan, G. (2017). *Anticipatory Regulation: 10 ways governments can better keep up with fast-changing industries*.

National Audit Office (2017). *Investigation: WannaCry cyber attack and the NHS*.

National Data Guardian for Health and Care (2016). *Review of Data Security, Consent and Opt-Outs*.

National Institute for Health and Care Excellence (2016). *AliveCor Heart Monitor and AliveECG app (Kardia Mobile) for detecting atrial fibrillation: Guidance and guidelines*.

Nature (2018). "AI diagnostics need attention". *Nature* 555.7696, p. 285.

NHS. *Atrial fibrillation*.

NHS Accelerator (2017). *AliveCor Kardia Mobile ECG*.

NHS Digital (2018a). *Data quality report on comorbidity diagnostic persistence*.

NHS Digital (2018b). *National data opt-out programme*.

NHS England (2017). *Next steps on the NHS Five Year Forward View*.

NHS England: Digital Maturity Assessment.

NIHR Oxford Biomedical Research Centre (2017). *Ground-breaking digital health deal agreed with Drayson Technologies*.

North, B. (2015). *The growing role of human factors and usability engineering for medical devices*. British Standards Institution.

Nosengo, N. (2016). "Can you teach old drugs new tricks?" *Nature News* 534.7607, p. 314.

Nuffield Council on Bioethics (2018). *Bioethics briefing note: Artificial intelligence (AI) in healthcare and research*.

Nuffield Trust (2012). *Can NHS hospitals do more with less?*

Offringa, R. et al. (2017). "Digital Diabetes Management Application Improves Glycemic Outcomes in People With Type 1 and Type 2 Diabetes". *J Diabetes Sci Technol*.

O'Hear, S. (2017a). *Ada is an AI-powered doctor app and telemedicine service*.

O'Hear, S. (2017b). "Babylon Health partners with UK's NHS to replace telephone helpline with AI-powered chatbot". *TechCrunch*.

O'Hear, S. (2017c). "Your.MD raises $10M to grow AI-driven health information service and marketplace". *TechCrunch*.

Perrin, N. (2017). *Oral evidence, Lords Select Committee on Artificial Intelligence*.

PHG Foundation (2017). *Written evidence, Lords Select Committee on Artificial Intelligence*.

Price, N. W. "Medical Malpractice and Black-Box Medicine". *Big Data, Health Law, and Bioethics*. Cambridge University Press.

Priisalu, J. and Ottis, R. (2017). "Personal control of privacy and data: Estonian experience". *Health Technol.* 7.4, pp. 441–451.

PwC (2017). *What doctor? Why AI and robotics will define New Health*. PwC.

Quer, G. et al. (2017). "Augmenting diagnostic vision with AI". *The Lancet* 390, p. 221.

Raza, S. (2017). *Oral evidence, Lords Select Committee on Artificial Intelligence*.

Reed, C. (2017). *Written evidence, Lords Select Committee on Artificial Intelligence*.

Reed, C., Kennedy, E., and Silva, S. N. (2016). "Responsibility, Autonomy and Accountability: legal liability for machine learning".

Regalado, A. (2018). "UK Biobank Supercharges Medicine with Gene Data on 500,000 Brits". *MIT Technology Review*.

Rhrissorrakrai, K., Koyama, T., and Parida, L. (2016). "Watson for Genomics: Moving Personalized Medicine Forward". *Trends Cancer* 2.8, pp. 392–395.

Royal Medical Colleges, A. of (2017). *Information and Digital Technologies Clinical: Requirements 2020*. Academy of Royal Medical Colleges.

Royal Statistical Society (2017). *Written evidence, Lords Select Committee on Artificial Intelligence*.

Rudin, C. and Wagstaff, K. L. (2014). "Machine learning for science and society". *Mach Learn* 95.1, pp. 1–9.

Russell, S., Dewey, D., and Tegmark, M. (2015). "Research Priorities for Robust and Beneficial Artificial Intelligence". *AI Magazine* 36.4, pp. 105–114.

Russell, S. and Norvig, P. (2016). *Artificial Intelligence: A Modern Approach*.

Sanches, J. M. et al. *Detecting Atrial Fibrillation using a Smart Watch—the mRhythm study*.

Saver, J. L. (2006). "Time Is Brain—Quantified". *Stroke* 37.1, pp. 263–266.

Semigran, H. L. et al. (2015). "Evaluation of symptom checkers for self diagnosis and triage: audit study". *BMJ* 351, h3480.

Severs, M. (2017). *Oral evidence, Lords Select Committee on Artificial Intelligence*.

Shortliffe, E. H., Detmer, D. E., and Munger, B. S. (2016). "Clinical Informatics: Emergence of a New Profession". *Clinical Informatics Study Guide*. Springer, Cham, pp. 3–21.

Shu, C. (2018). "Atomwise, which uses AI to improve drug discovery, raises $45M Series A". *TechCrunch*.

Smith, A. and Anderson, J. (2014). *AI, Robotics, and the Future of Jobs*. Pew Research Center.

Sood, H., NcNeil, K., and Keogh, B. (2017). "Chief clinical information officers: clinical leadership for a digital age". *BMJ* 358, j3295.

Sophia Genetics. *https://www.sophiagenetics.com/company/about-us/about-us-details.html*.

Staa, T.-P. v. et al. (2016). "Big health data: the need to earn public trust". *BMJ* 354, p. i3636.

Startup Daily (2015). *Google Flu Trends may have failed, but MedTech startup AIME has uncovered the secret to predicting viral disease outbreaks*.

Stein, N. and Brooks, K. (2017). "A Fully Automated Conversational Artificial Intelligence for Weight Loss: Longitudinal Observational Study Among Overweight and Obese Adults". *JMIR Diabetes* 2.2, e28.

Stevens, S. *Speech at Health and Care Innovation Expo 2017*.

Stewart, I., Debapratim, D., and Cole, A. (2015). *Technology and people: The great job-creating machine*. Deloitte.

Stokes Elizabeth A. et al. (2018). "Accurate costs of blood transfusion: a microcosting of administering blood products in the United Kingdom National Health Service". *Transfusion (Paris)* 0.0.

Stone, P. et al. (2016). *"Artificial Intelligence and Life in 2030." One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel*. Stanford University.

techUK (2017). *Written evidence, Lords Select Committee on Artificial Intelligence*.

The British Academy and The Royal Society (2017). *Data management and use: Governance in the 21st century*.

The Health Foundation (2016). *Response to National Data Guardian for Health and Social Care's Review of Data Security, Consent and Opt-Outs*.

The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems (2016). *Ethically Aligned Design: A Vision For Prioritizing Wellbeing With Artificial Intelligence And Autonomous Systems, Version 1*. IEEE.

The Law Society (2016). *Written evidence submitted to the House of Commons Select Committee on Science and Technology*.

The Nursing and Midwifery Council (2017). *The NMC Register*. The Nursing and Midwifery Council.

The Primary Care Foundation and the NHS Alliance. *Making Time in General Practice*. Tech. rep.

The Royal College of Radiologists (2017). *Written evidence, Lords Select Committee on Artificial Intelligence*.

The Royal Society (2017). *Machine learning: the power and promise of computers that learn by example*.

The Royal Society and The British Academy (2017a). *Data Governance: Landscape Review*. Tech. rep.

The Royal Society and The British Academy (2017b). *Data governance: public engagement review*. The Royal Society and the British Academy.

Tobin, J. (2016). *NHS and Social Care Workforce: Implications of Leaving the European Union*. House of Lords Library.

UK Government (2016). *Accelerated Access Review: Final Report*. Independently chaired report, supported by the Wellcome Trust. UK Government.

Understanding Patient Data (2017). *Identifiability demystified*.

Van Norman, G. A. (2016). "Drugs and Devices: Comparison of European and U.S. Approval Processes". *JACC: Basic to Translational Science* 1.5, pp. 399–412.

Viz.ai (2018). *Viz.ai Granted De Novo FDA Clearance for First Artificial Intelligence Triage Software*.

Wachter, R. (2016). *Making IT work: harnessing the power of health information technology to improve care in England*. National Advisory Group on Health Information Technology in England.

Wein, W. (2016). "Drug development: successes, problems and pitfalls—the industry perspective". *ESMO Open* 1.1.

Wellcome Trust (2016). *Response to National Data Guardian for Health and Social Care's Review of Data Security, Consent and Opt-Outs*.

Wellcome Trust (2017). *Understanding Patient Data launches today*.

Wellcome Trust and the Association of Medical Research Charities (AMRC) (2017). *Written evidence, Lords Select Committee on Artificial Intelligence*.

Willis, M (2017). *Health Automation: Computerisation and automation in general practice services*. Nottingham, UK.

Wolf, J. A. et al. (2013). "Diagnostic Inaccuracy of Smartphone Applications for Melanoma Detection". *JAMA Dermatol* 149.4, pp. 422–426.

Wrubel, n., Krimsky, n., and Anderson, n. (1997). "Regulatory Oversight of Genetically Engineered Microorganisms: Has Regulation Inhibited Innovation?" *Environ. Manage.* 21.4, pp. 571–586.

Yadav, A. et al. (2017). "Influence Maximization in the Field: The Arduous Journey from Emerging to Deployed Application". *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*. AAMAS '17. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems, pp. 150–158.

POLYGEIA 2018